Norwegian School of Economics
Bergen, Fall 2021

Impact of hull cleaning and crew performance on bunker consumption

Classification and optimization of underwater hull cleaning intervals under data uncertainty

Magnus Halvorsen Wik & Didrik Sverdrup Ånensen

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. Most importantly, we would like to extend our sincere gratitude for the pivotal counselling, feedback and discussions provided by our supervisor, Roar Os Ådland. His extensive insight and knowledge from the field of maritime economics has been essential, giving valuable advice and constructive criticism throughout the process. Furthermore, introducing us to his extensive network within the industry and connecting us with the ship owner for data. We want to thank the ship owner and the contact person for providing the data, and for rapid and thorough responses to any questions or issues we met. Their inputs have been of high importance for understanding the data provided and for evaluation of the results. Finally, we express our appreciation for the generous grants provided by the Norwegian Shipowners' association.

Norwegian School of Economics

Bergen, December 2021

Magnus Halvorsen Wik

Didrik Sverdrup Ånensen

Abstract

This thesis investigates the impact of underwater hull cleaning and crew performance on bunker consumption using noon reports. Biofouling imposes increased resistance on oceangoing vessels over time, and hull cleanings are subsequently performed to remove marine growth and reduce resistance. With uncertain data, a classification model is proposed to identify hull cleaning dates. The hull cleaning dates classified by the proposed model outperform the company-reported dates in terms of fitting expectations.

A model to economically optimize hull cleaning intervals is further defined, achieving savings of 0.3 – 1.4 % over a three-year period, by applying optimal intervals to vessels with two hull cleanings. Adding an additional hull cleaning resulted in fuel savings of 2.1 – 3.2 %. Although results are found under strict assumptions, they are similar to savings made by advanced continuous monitoring systems. Individual crew members are analyzed to find whether certain crew over- or underperforms in terms of fuel expenditure. Findings suggest that several masters and chief engineers have significant deviations in mean consumption even after controlling for all known covariates, although the causality of deviations remains unexplained.

Using fixed effects regression models, the impact of hull cleaning on fuel consumption is estimated to be approximately 1 % for Panamax and Medium Range vessels, and 9 % for Suezmax vessels. Crew members are estimated to explain between 3 – 4 % of variation in fuel consumption. Several machine learning models were tested to measure effects on prediction accuracy. Linear models achieved prediction accuracies of 63.5 – 67.5 %, increasing by 3 – 8 %, while advanced non-linear models achieved prediction accuracies of 77.1 – 78.2 %, increasing by 2.5 – 3 %. The thesis' findings contributes to existing literature by quantifying the impact of underwater hull cleaning and crew performance on bunker consumption under data uncertainty, by providing a model to identify hull cleanings and to observe potential savings of optimizing intervals.

Keywords – Underwater Hull Cleaning, Crew, Classification, Optimization, Machine Learning, Shipping

Contents

1 Introduction **1**

2 Literature review **4**

3 Data **7**
- 3.1 Presentation of data 8
- 3.2 Discussion of variables 8
- 3.3 Data-cleaning process 11
  - 3.3.1 Comments on cubic law 13
  - 3.3.2 Descriptive statistics of variables 15

4 Classifying hull cleanings **17**
- 4.1 OLS regression models 18
- 4.2 Comparison of prediction accuracy 20
- 4.3 Expectations 21
- 4.4 Classification algorithm 22
- 4.5 Fuel consumption profiles 24
- 4.6 Application of classification model 25
- 4.7 Comparison of fuel consumption profiles 26

5 Optimizing hull cleaning intervals **29**
- 5.1 Assumptions 29
- 5.2 Theoretical optimization model 31
- 5.3 Interval optimization 31
  - 5.3.1 Changing the shock size 33
  - 5.3.2 Changing the number of hull cleanings 34
- 5.4 Case study 36

6 Impact of individual crew performance **39**
- 6.1 Analysis of masters 39
  - 6.1.1 ANOVA testing 42
  - 6.1.2 T-testing 42
- 6.2 Analysis of chief engineers 43

7 Results and discussion **45**
- 7.1 Regression output 45
- 7.2 Variance decomposition 49
- 7.3 Prediction accuracy 51

8 Limitations and further reasearch **54**

9 Conclusion **57**

References **59**

Appendix **63**
- A1 Geographic locations 63

A2 Outliers in different vessels segments 64
A3 Fuel consumption profiles 65
A3.1 Comparison of fuel consumption profiles 65
A3.2 Classified hull cleaning dates for vessels with no basis of comparison 67
A3.3 Fuel consumption profiles for vessles with no basis of comparison . 68
A4 Sensitivity of parameters for optimization 70
A4.1 Interval sensitivity to shock size 70
A4.2 Interval sensitivity with increasing slope numbers 71
A4.3 Interval sensitivity changing shock size and growth factor 72
A4.4 Corresponding graphs with three performed hull cleanings 73
A5 Significant crew members with 5 % significance level 75
A6 Variance decomposition for ballast observations 77

List of Figures

3.1 Division of bearing and wind direction 10
3.2 Outliers for Medium Range vessels 13
3.3 Cubic law patterns of vessel segments 14
3.4 Speed-consumption relationship of the log transformed variables 15
4.1 Comparison of two port calls for vessel P1 23
4.2 Detected and company-reported hull cleaning dates 26
4.3 Comparison of selected fuel consumption profiles 27
5.1 Hull cleaning shock parameter δ 30
5.2 Optimization with $\delta = 0.5$ 33
5.3 Changing shock size δ 34
5.4 Increase in consumption when a HC deviates from the optimal timing . . 38
6.1 Boxplot of Panamax masters' mean consumption controlled for external variables 40
6.2 Boxplot of Suezmax masters' mean consumption controlled for external variables 41
6.3 Boxplot of Medium Range masters' mean consumption controlled for external variables 41
A1.1 Trading routes for all Panamax vessels 63
A1.2 Trading routes for all Suezmax vessels 63
A1.3 Trading routes for all Medium Range vessels 63
A2.1 Cut-off lines for outliers for all segments 64
A3.1 Comparison of fuel consumption profiles 65
A3.1 Comparison of fuel consumption profiles 66
A3.1 Comparison of fuel consumption profiles 67
A3.2 Classified hull cleaning dates for vessels with no basis of comparison . . . 67
A3.3 Classified fuel consumption profiles 68
A3.3 Classified fuel consumption profiles 69
A4.1 Optimal intervals for various shock sizes 70
A4.2 Optimal intervals for various shock sizes illustrated with $\delta = 0.9$ 70
A4.3 Changing growth factor g 71
A4.4 Optimal intervals for various growth factors 72
A4.5 Comparison of growth factors with different shock sizes 72
A4.6 Fuel consumption slopes for different values of δ with three hull cleanings 73
A4.7 Optimal intervals for various shock sizes with three hull cleanings 73
A4.8 Different growth factors with three hull cleanings 74

List of Tables

3.1 Variables selected for analysis . . . 11
3.2 Descriptive statistics for Panamax vessels . . . 16
3.3 Descriptive statistics for Suezmax vessels . . . 16
3.4 Descriptive statistics for Medium Range vessels . . . 16
4.1 Regression for all vessel segments split by interaction term on wind . . . 19
4.2 Performance metrics for machine learning methods . . . 21
5.1 Effects on consumption with different shock sizes . . . 35
5.2 Effects on consumption with different consumption increases . . . 36
5.3 Effects of optimal timings with two hull cleanings . . . 37
5.4 Effects of optimal timings with three hull cleanings . . . 38
6.1 Masters with significantly different consumption . . . 43
6.2 Chief engineers with significant differences in consumption . . . 44
7.1 Regression for different sets of regressors with fixed effects (FE) for Panamax vessels . . . 46
7.2 Regression for different sets of regressors with fixed effects (FE) for Suezmax vessels . . . 47
7.3 Regression for different sets of regressors with fixed effects (FE) for Medium Range vessels . . . 48
7.4 ANOVA for Panamax vessels . . . 49
7.5 ANOVA for Suezmax vessels . . . 50
7.6 ANOVA for Medium Range vessels . . . 50
7.7 Performance metrics for Panamax vessels . . . 52
7.8 Performance metrics for Suezmax vessels . . . 52
7.9 Performance metrics for Medium Range vessels . . . 53
A5.1 Significant chief engineers with 5 % significance level . . . 75
A5.2 Significant chief engineers with 5 % significance level . . . 76
A6.1 Variance decomposition for Panamax vessels in ballast . . . 77
A6.2 Variance decomposition for Suezmax vessels in ballast . . . 77
A6.3 Variance decomposition for Medium Range vessels in ballast . . . 77

1 Introduction

While most people live their lives without offering much thought to the endless supply chains and logistics that bring them their everyday products, shipping remains the most important conveyor of trade volumes. It has been described by the UN (2016) as the backbone of global trade and economy, and in 2018 shipping accounted for 80 % of global trade by volume and 70 % of global trade by value (UNCTAD, 2018). Furthermore, shipping is by a substantial margin the most environmental friendly means of cargo transport, measured by $\frac{CO^2}{km \cdot kg}$(IMO, 2009). However, the shipping industry alone is responsible for 2.89 % of global anthropogenic emissions, increased from 2.76 % in 2012, despite carbon intensity reductions of 1-2 % per year since 2015 (IMO, 2020). This means that absolute emissions in shipping grow faster than the improvements in fuel efficiency. The International Maritime Organization (IMO) has during the last decade imposed reductions on CO^2 and sulfur emissions (IMO, 2019). Reaching emission goals solicits effective solutions to reduce fuel consumption within everything from ship design (Hochkirch and Bertram, 2010) to route optimization (Kobayashi et al., 2014).

Holding trading patterns and ship-specific attributes constant, two factors that influence fuel consumption heavily is the condition of the hull and crew performance. Deteriorating hull and propeller performance is assumed to account for approximately 10 % of fuel costs and emissions for oceangoing vessels (Copernicus, 2021). For instance, the development of biofouling can drastically decrease fuel efficiency (Lindholdt et al., 2015). Underwater hull cleanings are amongst the most widespread methods of combating biofouling, although several other abatement methods exist (IMO, 2020). Adland et al. (2018) showed that there are significant decreases in fuel consumption following underwater hull cleaning procedures, also referred to as simply *hull cleanings*.

While biofouling can certainly have an impact on the performance of oceangoing vessels, the effects of crew members on fuel consumption is a topic which has not yet been subject to analysis. Managing director Peter Knudsen of *Blueflow Energy Management* stated that the most important aspect of saving fuel is the crew, referring to the importance of having the right competence on board. Without the capability to correctly operate and understand signals from different sources of data, it would be futile to introduce new technology and increase energy efficiency (VPO, 2018). Masters and chief engineers may impact bunker consumption through their ability to maintain stable speeds, efficiently accelerate/decelerate, and handle the rudder in a proficient manner.

The contribution of this thesis is fourfold. Although hull cleaning has been shown to have significant effects on fuel consumption, these procedures are not necessarily well documented. First, section 4 proposes a classification algorithm utilizing OLS regressions to detect hull cleaning dates under data uncertainty. The model builds upon the findings of Adland et al. (2017) to identify effects of hull cleanings on fuel consumption. Next, fuel consumption profiles using detected hull cleaning dates are visualized and discussed in regard to *a priori* expectations. This is valuable for ship owners who want to analyze fuel consumption with uncertain data or deficient hull cleaning dates.

Second, section 5 performs economical optimization of hull cleaning intervals. Further, the section investigates how assumptions about fuel consumption behavior affect the optima. Optimization of hull cleaning intervals has not been analyzed in existing literature, probably because of the complex relationship between the development of biofouling over time and fuel consumption. The section introduces a theoretical optimization problem to solve the general economical optima of hull cleaning intervals, subject to changes in various parameters. Optimization of hull cleaning intervals leads to lower fuel consumption as a result of minimizing external resistance over a dry dock interval. This is of interest for shipping companies who want to reduce fuel costs by improving their basis of decision-making regarding the timing of hull cleaning procedures.

Third, building on the information gathered on the effects of hull cleanings, section 6 isolates the effects of crew members, and ascertains whether and to what extent individual crew can influence fuel consumption. Modelling the effects of masters and chief engineers is however dependent on the interval in which they sail, due to the effects of biofouling and hull cleanings. Hull cleaning intervals should therefore be accounted for before looking at their impact. This section analyzes the effects on fuel consumption in terms of individual differences between crew members. This is interesting for ship owners who want to better understand to what degree individual crew members can affect fuel consumption, and the importance of having well trained, competent crew.

Fourth, section 7 quantifies the impact of hull cleaning and crew using fixed effects regression models. Next, the section examines the extent to which various drivers explain variation in fuel consumption, by fitting variance decompositions. Lastly, several linear and non-linear prediction models are implemented and compared, to discern how the thesis' findings affect prediction accuracy. This is interesting for shipping companies who are curious about the drivers of fuel consumption, and the total effects of hull cleaning and crew. The findings are useful for improving modelling and predictions of bunker consumption under data uncertainty. The discovered effects can consequently be used as a basis for further analysis and decision making to save fuel costs and emissions.

2 Literature review

Meng et al. (2016) describes the fuel efficiency P_E of a ship as a function of vessel speed V and resistance R_T such that $P_E = V \cdot R_T$. Resistance is further decomposed to three types of resistance:

$$R_T = R_F + R_R + R_A \tag{2.1}$$

where R_F represents the frictional resistance, including total deterioration of the hull and propeller. R_R represents the residual resistance (i.e. primarily waves), and R_A represents the air resistance mainly caused by wind. Lindholdt et al. (2015) suggests that the frictional resistance causes 70-90 % of the ships total resistance.

Biofouling, defined by Hellio and Yebra (2009) as "*the undesirable accumulation of microorganisms, algae and animals on structures submerged in seawater*", is among the top contributors to decreasing fuel efficiency (Hakim et al., 2017). Schultz (2007) found that light slime or deteriorated coating could impose penalties of 11 % on total resistance, heavy slime layers could penalize resistance by 20 %, while heavy calcareous fouling inflicted resistance penalties of 80 %. To combat biofouling, various antifouling techniques has been subject to thousands of years of development in line with the industry's importance to society (Dafforn et al., 2011). It is estimated that antifouling coatings save the shipping industry \$60 billion and 384 million tons of CO^2-emissions on an annual basis (Bressy and Lejars, 2014).

Although several antifouling paints have been shown to improve fuel efficiency (Kojima et al., 2016; Yang et al., 2014; Tripathi, 2016), studies suggest an average decline in vessel performance of 15-20 % over a typical 4 to 5-year sailing interval (IMO, 2011). Farkas et al. (2021) suggested that biofilm layers could increase fuel consumption by between 671 to 4153 tons per year for post-panamax tankers, leading to a potential cost increase of more than \$1.75 million. Consequently, performing underwater hull cleanings is a widespread approach to keep fuel efficiency near initial levels (i.e. dry dock levels). In dry dock, a complete overhaul of the hull is performed with new antifouling coating applied. In contrast, hull cleanings are generally performed by either divers or magnetic robots, removing marine growths from the hull with rotating brushes (Lindholdt et al., 2015). The effect is therefore greater in periods after a dry dock compared to hull cleanings.

Adland et al. (2018) measured the marginal effect of hull cleanings using a *difference-in-difference* estimator. Given that lnC_{vt} is the log-consumption of vessel v at time t, and w is the observable time window before and after the hull cleaning, the model is defined as:

$$lnC_{vt} = \delta_w \cdot I_{vt}^{AFTER_w} + X_{vt}\beta + \vartheta_v + \varepsilon_{vt} \tag{2.2}$$

where δ_w is the difference-in-difference estimator measuring the effect of hull cleanings, the dummy variable $I_{vt}^{AFTER_w}$ is equal to 1 after a hull cleaning has been performed, and 0 before. β is the vector of estimated coefficients for the external covariates X_{vt}, the time-invariant vessel-specific (i.e. fixed) effects are given by ϑ_v, and ε_{vt} represents random perturbation with $E[\varepsilon_{vt}] = 0$ and $Var(\varepsilon_{vt}) = \sigma^2$. The paper observed an average reduction in fuel consumption of 9 % following underwater hull cleanings.

The difference-in-difference model presented above is based on the previous work by Adland et al. (2017). To estimate and visualize fuel consumption profiles for sailing intervals with several hull cleanings, Adland et al. (2017) presented bunker consumption as a logarithmic function of a given set of variables and a time trend, equal to:

$$lnC_t = \sum_k \delta_k \cdot I_{tk} + X_t \cdot \theta + f(t) + \varepsilon_t \tag{2.3}$$

where the subscript t represents a given date, and k is signalling a hull cleaning with $k = \{1, 2, ..., K\}$. δ is denoted as the fuel efficiency effect for the period after a given hull cleaning k performed at time t. Further, $X_t \cdot \theta$ displays the effects of vessel characteristics such as weather conditions, speed, or draft, to isolate fuel consumption by excluding external factors. $f(t)$ is given as a cubic time trend such that $f(t) = \tau_1 \cdot t + \tau_2 \cdot t^2 + \tau_3 \cdot t^3$ and ε_t is residual perturbation with $E[\varepsilon_t] = 0$. A linear time trend after each hull cleaning procedure is added, which considers differences in consumption growth for each interval. It is implemented into the equation as $\delta_k = \alpha_k + \beta_k \cdot t$ expanding equation 2.3 to:

$$lnC_t = \sum_k (\alpha_k + \beta_k \cdot t) \cdot I_k + X_t \cdot \theta + f(t) + \varepsilon_t \tag{2.4}$$

Although the positive effects of underwater hull cleanings seem clear-cut, such procedures have certain drawbacks. If the fouling has reached a level where soft brushes are no longer effective, the procedure runs the risk of mechanically damaging the coating and further inducing corrosion on the hull (Lindholdt et al., 2015). Such harm will likely lead to a substantially swifter return of biofouling and thus lead to a larger increase in fuel consumption over time. While hull cleaning negates the global oceanic threat of spreading invasive marine species (Adland et al., 2018), certain antifouling paints enact adverse effects onto the environment when scrubbed off (Lindholdt et al., 2015). When considering whether to perform underwater hull cleanings, the trade-off between harmful effects on the ship and environment versus the cost savings achieved must be carefully weighted. The potentially harmful environmental effects is why the topic is still being discussed at the IMO, even though it is recognized as an effective abatement method (IMO, 2020).

Reducing fuel consumption and carbon emissions is key for most ship owners and charterers. On-board sensory data and software technology for continuous monitoring of the hull condition and ship performance has been developed by several players in recent years (Lande, 2017). Decision support services and software tools combined with innovations in data analytics have created new opportunities for owners and operators pursuing increased efficiency and fleet performance (StormGeo, 2021). Using decision support tools, crew can closely observe and take action once performance is dropping below a certain threshold. This has opened the possibility of optimizing the timing of hull cleanings to maintain efficient performance by gauging the vessel resistance, and continuously compare it to vessels of similar build (Lande, 2017).

These types of continuous decision support tools are still not very widespread. However, an unnamed user of the CASPER[1] software reported annual fuel savings of 1-3% across their fleet, amounting to approximately \$4 million, by performing favorable hull cleanings. With annual subscription costs of \$700,000, such decision support systems can lead to extensive fuel cost savings (GCaptain, 2012).

[1]CASPER is a software developed by *Propulsion Dynamics* enabling ships to undergo hydrodynamic mapping, which is needed to acquire the appropriate dimensions and baseline data for comparison with in-service performance data.

3 Data

Withing shipping, noon reports are a common practice to provide vessel updates to the shipping company. Noon reports are prepared and sent with a sample frequency of 24 hours, at noon local time, by the chief engineer (Wankhede, 2021). The reports include information of the ship's consumption and underlying conditions at specific points in time (Smith et al., 2013). They consist of standardized data and are among other things used to observe how speed and external environmental forces impact the ship's performance. In most cases, noon reports are used to evaluate the fuel consumption for a specific vessel in comparison to vessels of similar type.

However, one of the main issues with noon reports is the data granularity. Since the data is based on a combination of 24-hour averages and snapshots, the interpretations can sometimes be misleading and/or confusing. This specifically impacts the data on weather conditions and ship speed where we cannot capture events of accelerations/decelerations, manoeuvrings or sudden changes in weather that heavily impacts fuel consumption within those hours (Smith et al., 2013). In addition, most weather data is recorded as snapshots[2], meaning that changes in conditions during the previous 24 hours are not considered.

Furthermore, noon reports will not always be accurate due to numerous sources of errors. For instance, human error might be prevalent as all information has to be manually input by the chief engineer. Misinterpretations of inputs, and usage of different units or roundings between reporters, could all lead to errors in the assessment of noon reports. For this reason, continuous monitoring systems have been developed in order to improve data accuracy, the speed of acquisition, and allowing for higher resolution of data (Smith et al., 2013). Few vessels currently have such systems installed, and for the time being, noon reports are far more widespread than continuous monitoring. To address the presence of uncertainty in noon reports, a thorough cleaning process is required before using the data for analytic purposes.

[2]Other ship owners may have a policy that weather reports are estimated as 24-hour averages.

3.1 Presentation of data

The data in this thesis is comprised from 31,620 unprocessed noon reports spanning from the 1st of January 2018 to the 31st of December 2020 from an anonymous international shipping company. The fleet consists of three different classes (Suezmax, Panamax and Medium Range), each of which have six to eleven sister vessels of similar build. Data registered by the various vessels' Automatic Identification System (AIS) during the same period will also be utilized. To achieve an overview of typical trading routes and frequent locations, positional data from each vessel class are presented in appendix A1. Although similar, noticeable takeaways are that Panamax ships have no trans-pacific voyages, predominantly operating in the Atlantic Ocean. Medium Range vessels are evenly spread worldwide with no clear area of operations, while Suezmax vessels generally sail between the Americas and the Middle-East/Asia.

3.2 Discussion of variables

Due to environmental restrictions, usage of fuel types with different sulphur concentrations can vary depending on the current area and vessel. Total consumption is calculated only from the main propulsion engines, with no differentiation between fuel types. Masters and chief engineers are represented by unique numbers, which anonymously represent the actual masters and chief engineers in charge at the time of the noon report.

From April 2019, the data owner transitioned to a new reporting system containing additional weather and vessel operational data, such as wave and swell height, trim, and slip. As the variables are not available for the entire period, there is a trade-off between the total number of observations and additional weather and ship data. Alternatively, external weather data from ERA5 with hourly measurements could be utilized, but due to differences in resolution this poses a problem. ERA5 data is measured only on certain coordinates, leading to incongruence between local conditions and vessel conditions, as well as rounded coordinates resulting in missing values[3]. Thus, although some weather variables have been shown to be significant, it was decided that the inclusion of weather variables from external sources was not worth the modest increase in explanatory power

[3] When the nearest rounded coordinate is ashore, there are no measurements of water-specific variables.

exhibited in previous research (Adland et al., 2017; Nilsson and Nilsson, 2021).

The AIS data includes metrics such as longitude and latitude, draught, and heading of the vessel. Since draught is only measured in noon reports from the new system, 43 % of these observations are missing. AIS data is therefore used to measure draught. The course of the ship is manually registered, but suffers from missing values. Furthermore, the AIS heading is recorded as a snapshot, which could differ from the 24-hour course. To achieve a similar basis for all noon reports, the course of the ship is calculated as the bearing between the previous and current coordinates. The calculated bearings do not consider the actual direction of the ship, which factors in currents. This thesis therefore assumes that ships sail in straight line unaffected by currents. An issue with this assumption is that ships travelling along or around coastlines, is regularly forced to adjust the course of the ship. However, it is deemed that the advantages of this assumption outweighs the minimal difference between the calculated bearing and the reported course.

Port calls refer to intermediate stops for ships on scheduled voyages. When stopping, the ship will either be moored or at anchor, which indicates that the total consumption for a day in port should be close to zero. However, port calls are not easily discernible in the data. This is partly because they are not reported explicitly, and partly because speed reported by AIS and reported consumption have apparent discrepancies. In this thesis, a port stay is defined as two consecutive days with close to zero consumption[4]. The day before and after a port stay have reduced consumption compared to sailing days[5] due to manoeuvring at slow speeds close to shore, in addition to idle time. Such observations will be removed in the cleaning process because consumption analysis is only meaningful on sailing days (Adland et al., 2017). The initiation of a port call is therefore flagged prior to the day heading into port, to avoid removing flagged port calls when cleaning the data. The requirement of having two consecutive days with near zero consumption is to account for misreportings in data, thus reducing the risk of incorrectly flagging port calls amid trips. As the flagging of port calls is not perfect, it is important to emphasize that the flagging of port calls is not directly related to the data cleaning process, but for usage in later analysis.

[4]While port stays usually have zero consumption, manoeuvring within ports occasionally leads to low, non-zero consumption.

[5]Sailing days are here defined as days where the ship is believed to be sailing continuously.

Relative wind direction is estimated using the actual wind direction and bearing of the vessel. Since the data of wind direction is reported on eight different levels (N, NE, E etc.), all bearings are transformed into equal levels to match them. Bearings above 337.5 and below 22.5 degrees is set as N, while bearings between 22.5 and 67.5 degrees are set as NE and so on. This division is illustrated by figure 3.1. Since wind direction is defined as the direction from which it originates, the vessel is presumed to sail in headwind when the vessel course and wind direction are equal. If the course and wind direction are opposite, the vessel is sailing with wind from astern and so on.

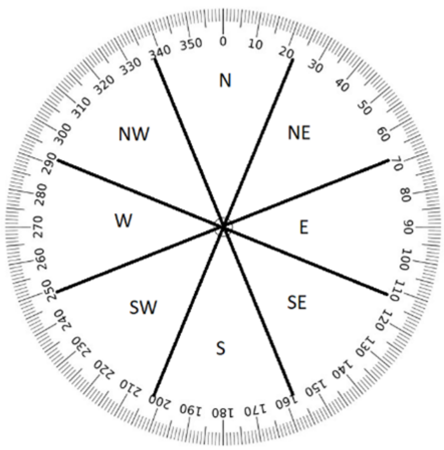


Figure 3.1: Division of bearing and wind direction

Extensive vessel information was also provided and implemented into the dataset. Stopford (2008) states that older ships generally have higher operating costs due to physical degradation over time, while Rakke (2016) observed that the age of the engine could affect fuel efficiency as much as 10 %. Therefore, vessel age could be an important factor to examine trends in fuel consumption. To isolate the effects of only underwater hull cleanings, dry dock dates are also important to document. The dry dock dates are determined by internal data, and the time since last dry dock is calculated and implemented into the data set. However, two linear time trends would describe the same variance, hence resulting in multicollinearity in the models. Since this thesis only examines observations between two dry docks for each vessel, the time since last drydock is chosen to represent the linear time trend.

Further variables could include effects that are not quantifiable using only noon reports and averaged weather data. For instance, close monitoring of the hull and detailed engine parameters would be advantageous, and Uyanık et al. (2020) achieved substantially higher adjusted R^2 when including high resolution engine data. The inclusion of RPM as a variable was considered, but Yu et al. (2021) expectedly observed a strong collinear relationship with speed. Since the presence of multicollinearity in the model would lead to inaccurate coefficient standard errors, the variable was excluded.

Table 3.1 shows the final selection of variables which will be utilized for further analysis, grouped by data source.

Table 3.1: Variables selected for analysis

Data Source	Parameter	Units
Noon reports	Date	dd/mm/yyyy
	Speed through water	kts
	Vessel name	ID
	Vessel segment	Type
	Wind speed	Beaufort[a]
	Relative wind direction	Directions[b]
	Wind speed : direction	Beaufort : Direction
	Cargo status	Laden/Ballast
	Master	ID
	Chief engineer	ID
AIS data	Draft	m
	Longitude	deg
	Latitude	deg
	Bearing	deg
Internal data	Time since dry dock	weeks
	Vessel age	weeks

[a]The Beaufort scale measures wind speed on a scale of 0 (calm) to 12 (hurricane force) (RMS, 2018).
[b]Directions consist of cardinal directions (N, E, S, W) and ordinal directions (NE, SE, SW, NW).

3.3 Data-cleaning process

The raw data consists of 31,620 noon reports across three years from 29 oceangoing vessels. All observations with consumption less than 2.5 tons/day are removed, dropping 13,409 observations. This removes observations where vessels have been stationary along with erroneous reports. Observations from the 1$^{\text{st}}$ of January 2018 are removed because variables such as time difference and bearing are dependent on the previous report, removing an extra 19 observations. Following Adland et al. (2018), noon reports where average speed is below seven knots or above 15 knots are removed from the dataset, dropping 805 observations. Speeds below seven knots indicate days with manoeuvring and/or pilotage, and are closely related to port calls. As the vessels have set design speeds, speeds over this threshold occurs either due to abnormalities or misreportings.

Reported average speed displays speed through water and is only measured when the vessel is sailing, thus excluding the impact of idle time. Calculating average speed as the total sailing distance since last report divided by the time since last report, gives the actual average speed over ground for the last 24 hours. Large variations in the difference between reported average speed and calculated average speed indicate periods of idle time or lower speeds, which is undesirable as the analysis should only consider sailing days. These variables are then used to further filter out port calls and days heading in and out of port. Since currents make up some of the difference between the two metrics, they are differentiated even when the vessel is sailing continuously. By setting the maximum difference between calculated and reported average speed to 2.5 knots, differences due to currents are allowed, while irregularities due to inconsistent sailing speeds are filtered out. This removes 2,741 observations.

Furthermore, all observations that have the exact same coordinates as the previous report are removed from the dataset, resulting in the exclusion of 337 observations. This also filters out duplicate reports made by the same vessel on the same day. Observations with a time difference since last noon report of more than 25 hours or less than 23 hours are rejected[6]. This allows for differences in time zones when a vessel sails from one time zone to another, and removes another 212 observations. Observations with missing values for either wind direction or wind speed are further removed, excluding 117 observations.

Removal of outliers is important for Ordinary Least Squares (OLS) regression, as they tend to have high leverage. Observations with wind speed higher than 10 on the Beaufort scale are removed to account for extreme weather, dropping eight observations. Draught values outside the 0.5 and 99.5 percentile for each vessel segment are removed from the dataset, excluding 255 observations. In figure 3.2, remaining outliers are removed for the Medium Range class based on visual identification. The cut-offs are established by the solid lines, which are based on either absolute fuel consumption or the relationship

[6] Although some reports may have higher time differences due to leaving port in the morning, and as such do not submit noon reports for the morning hours alone, these fall under the category of days in and out of port and are therefore undesirable.

between consumption and speed. Each vessel segment has distinct cut-offs, and plots of the remaining data for each segment are shown in appendix A2. In total, six outliers are removed from the dataset.

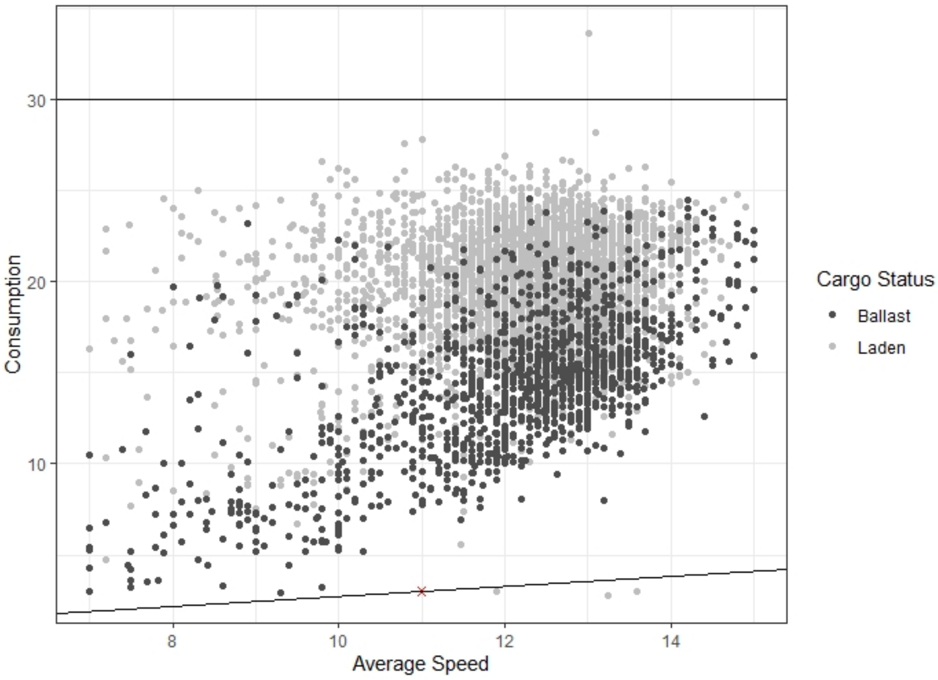
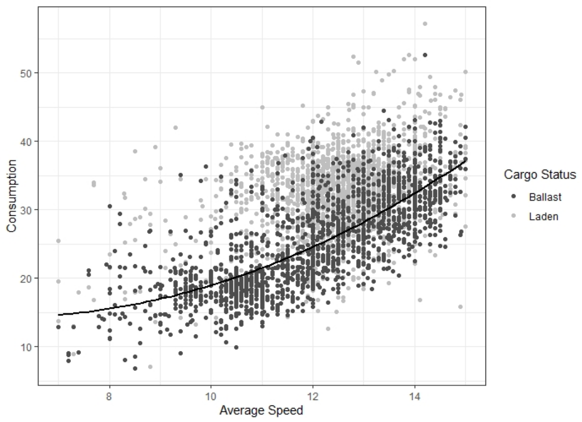


Figure 3.2: Outliers for Medium Range vessels

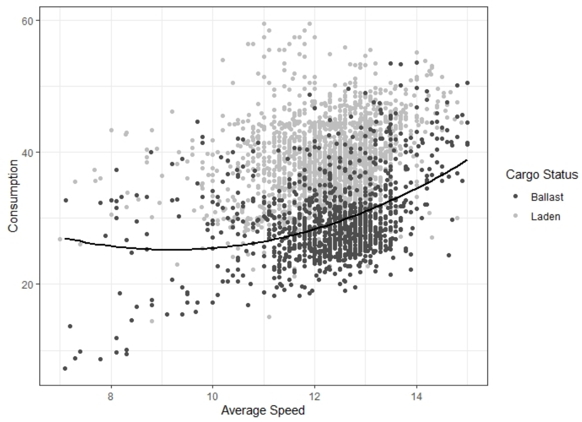
3.3.1 Comments on cubic law

To determine whether the remaining data set conforms to the expected fuel consumption behavior of cargo ships, the data can be examined in relation to the cubic law. A widely used assumption within shipping is that fuel consumption follows a cubic law, which states that the bunker consumption of ships can be approximated well as the current vessel speed to the power of three (Meng et al., 2016). Wang and Meng (2012) suggested that this is indeed a good approximation in the absence of historical data, where it is shown that the exponent is between 2.7 and 3.3 for speeds below 20 knots. For speeds above 20 knots, Kontovas and Psaraftis (2011) deemed that an exponent of 4 was appropriate. Adland et al. (2020) proposed that while the cubic rule seems to hold true near vessel design speeds, the exponent is instead between 1.7 and 2.3 at the ships' typical operating speeds.

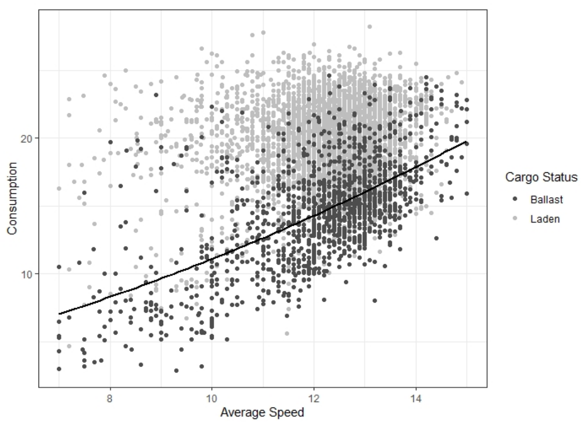
From the cleaned data, the correlation between speed and fuel consumption is found by applying a polynomial OLS regression line to fit observations for vessels in ballast condition. Only ballast observations are considered because masters have more flexibility regarding sailing speed when the ship is in ballast, compared to the stricter instructed speeds of laden vessels, usually between 12-14 knots. Thus, the variance of speed for laden vessels is likely too small to identify similar patterns. The observations are plotted prior to controlling for differing operating conditions. Consequently, observations at lower speeds could be affected by harsh weather, although the most extreme conditions are excluded from the data set. In figures 3.3a and 3.3b, the tendencies of a cubic law seem clear. Assuming a power-relationship between consumption and speed, there is indeed a resemblance of a non-linear relationship. In figure 3.3c however, the relationship looks rather linear. Still, the positive correlation indicates that speed is an important variable for predicting fuel consumption. It is therefore assumed that the remaining data set represents typical shipping data in theory, and that the results in this thesis can likely be extended to be representative for other vessels.



(a) Panamax



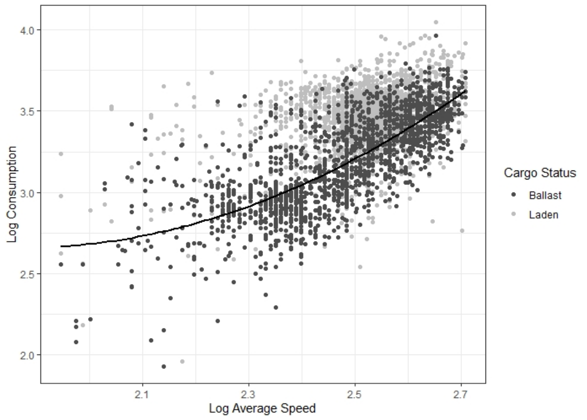
(b) Suezmax



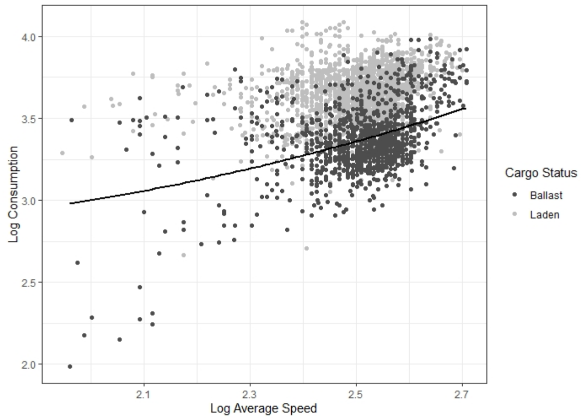
(c) Medium Range

Figure 3.3: Cubic law patterns of vessel segments

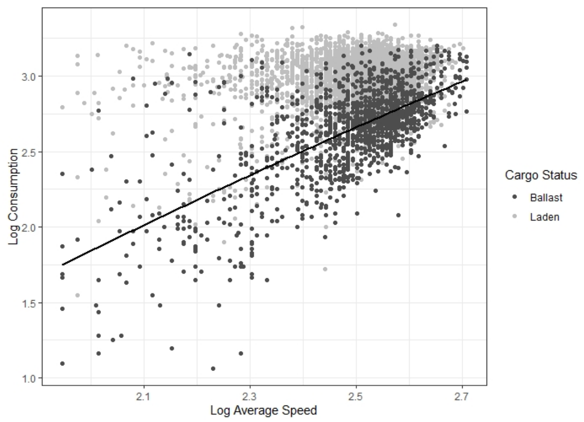
Considering the regression lines in figure 3.3, performing a log-transformation of the speed and consumption variables is likely beneficial, as there is a non-linear relationship between speed and consumption. By taking the logarithm of both variables, the relationship is linearized to a larger extent, enabling classic OLS-regressions and linear models to better fit the data (Dahly, 2017). The effective relationship between variables therefore becomes non-linear, while preserving a linear model (Benoit, 2011). This becomes clear in figure 3.4b, where the correlation between speed and consumption in log-log space is markedly more linear compared to figure 3.3b.



(a) Panamax



(b) Suezmax



(c) Medium Range

Figure 3.4: Speed-consumption relationship of the log transformed variables

3.3.2 Descriptive statistics of variables

The cleaned data set consists of 13,711 observations across all vessel segments. Summary statistics for the final selection of variables for each vessel class is presented in tables 3.2 to 3.4. In general, the external variables are fairly similar for each class. Medium Range vessels have substantially lower consumption compared to the other two classes, because they are significantly smaller. Furthermore, Medium Range vessels are considerably younger than the other classes, which could affect the fuel efficiency relative to Suezmax and Panamax vessels. The mean draught ranges from 9.3m (Panamax) to 13.1m (Suezmax).

Table 3.2: Descriptive statistics for Panamax vessels

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Wind Speed	4,135	4.67	1.39	0	4	6	9
Draught	4,135	9.27	1.88	6.80	7.40	11.00	12.70
Consumption	4,135	29.48	7.19	6.90	24.50	34.60	57.20
Speed	4,135	12.34	1.37	7.00	11.60	13.30	15.00
Vessel age	4,135	594.55	116.30	343	490	691	785

Table 3.3: Descriptive statistics for Suezmax vessels

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Wind Speed	3,312	4.62	1.43	0	4	6	9
Draught	3,312	13.07	3.15	9	9.2	15.8	17
Consumption	3,312	35.28	7.26	7.30	29.87	40.20	59.50
Speed	3,312	12.29	1.04	7.00	11.80	12.90	15.00
Vessel age	3,312	367.07	54.59	256	322	409	482

Table 3.4: Descriptive statistics for Medium Range vessels

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Wind Speed	6,264	4.45	1.33	0	3	5	9
Draught	6,264	10.37	1.67	6.80	8.50	11.80	12.60
Consumption	6,264	18.70	3.75	2.90	16.70	21.40	28.20
Speed	6,264	12.30	1.12	7.00	11.90	13.00	15.00
Vessel age	6,264	152.31	66.44	0	106	202	309

4 Classifying hull cleanings

Due to the documentation of performed hull cleanings sometimes being inadequate, as they are not necessarily registered in noon reports, hull cleaning dates are often uncertain. Using known dates, Adland et al. (2018) found significant drops in fuel consumption after hull cleanings, net all other predictors. The aim of this classification model is to reverse engineer these findings and use the mathematical model presented in section 2 as a foundation, where Adland et al. (2018) used a difference-in-difference estimator to measure the causal effect of hull cleanings on fuel consumption. However, the described approach necessitates the knowledge of when hull cleanings have been performed, and equally important, when they have not been performed. Without this information, a classification based on the same approach would contradict the assumption that the intervention[7] occurs independent of the outcome (CPH, 2013). In contrast, a classification model with no training data entails the identification of hull cleaning dates by identifying the expected outcome.

To classify hull cleaning dates, it is necessary to utilize a prediction model that captures the effects of the included predictors well, but where the variance of unobserved predictors remains. In more advanced prediction models there is a higher risk of overfitting. This implies that the existing predictors capture variance to a larger extent, including the variation around hull cleanings and biofouling that is yet to be explained. In terms of a bias-variance trade-off, advanced models increase the variance by decreasing the bias. Bias is used to quantify how much the average accuracy of the method changes as input from data changes, while variance describes how sensitive the method is to the chosen input data (James et al., 2013). OLS regressions thus have higher bias and lower variance compared to more complex non-linear techniques (Fortmann-Roe, 2012). This section seeks to observe the effect of sudden drops and changes in the fuel consumption slope over time. It is therefore desirable to use models with low variance to perceive such effects in the error term. With low variance, high interpretability, and low computational time, standard OLS regression will be used for further analysis.

[7]The intervention would in this case be the performance of a hull cleaning.

4.1 OLS regression models

The observed results from the OLS regression for each class, with fixed effects for specific vessels, are shown in table 4.1. The OLS estimator is given by:

$$lnC_{vt} = \alpha + \sum_i \beta_i X_{vit} + \vartheta_v + \varepsilon_{vt} \tag{4.1}$$

where lnC_{vt} is the log consumption of vessel v at time t, α is the constant term, and β_i is the vector of coefficients for the vessel characteristics i, given by X_{vit}. Fixed vessel effects are given by ϑ_v, and the residual term is ε_{vt} with $E[\varepsilon_{vt}] = 0$ and $Var(\varepsilon_{vt}) = \sigma^2$. This includes models with and without an interaction term between wind speed and wind direction, with crosswind set as the reference category. The interaction term is included to quantify the effects of additional wind strength given a set wind direction. For instance, the benefit of having wind from astern is dependent on the wind speed[8]. To make sensible estimates of the time trend variable, observations occurring before a vessel's first dry dock and after a vessel's second dry dock are removed. Vessels with observable dry dock periods of less than two years are also excluded, dropping two Panamax vessels and five Medium Range vessels from further analysis. As a consequence, a further 3,226 observations are removed.

When excluding the interaction term, all variables are significant on a 1 % level explaining between 55 % to 70 % of the variance in log fuel consumption. The coefficients are as expected, with log speed, draught, and headwind increasing resistance, and wind from astern reducing resistance. Considering columns (1), (3), and (6), increasing speed with 1 % leads to an increase in fuel consumption of 1.43 % for the Panamax class, while only increasing consumption by 1.08 % for Suezmax vessels. Since draught indicates how deep the ship extends below the waterline, it makes sense that consumption increases as the load of the vessel is increased. Deeper draught also increases frictional resistance as a larger portion of the hull is beneath the waterline. If draught increases with one unit, fuel consumption increases with 4.7 %[9] for Panamax, 3.8 % for Suezmax vessels, and 8.8 % for

[8] It could be argued that when wind speed increases above a certain level, the benefit of having wind from astern is reduced as wave height rises. In that case, speed is also likely to decrease.

[9] Since the logarithm of consumption is the dependent variable, the marginal effect of non logarithmic variables are given as the exponential coefficient subtracted by one: $\%\Delta lnC_t = 100 \cdot (exp(\beta_i) - 1)$.

Table 4.1: Regression for all vessel segments split by interaction term on wind

	Dependent variable:					
	Log Fuel Consumption					
	Panamax	Panamax	Suezmax	Suezmax	Medium	Medium
	(1)	(2)	(3)	(4)	(5)	(6)
Log Speed	1.430***	1.414***	1.081***	1.044***	1.381***	1.354***
	(0.022)	(0.023)	(0.031)	(0.031)	(0.028)	(0.028)
Draught	0.046***	0.047***	0.037***	0.036***	0.077***	0.077***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Wind Beaufort (WB)	0.045***	0.041***	0.038***	0.040***	0.069***	0.056***
	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)
Wind Front (F)	−0.014	0.058***	−0.064**	0.043***	0.033	0.028***
	(0.030)	(0.009)	(0.027)	(0.008)	(0.030)	(0.009)
Wind SideFront (SF)	−0.016	0.028***	−0.023	0.033***	0.004	0.027***
	(0.026)	(0.008)	(0.026)	(0.008)	(0.026)	(0.008)
Wind SideBack (SB)	0.065**	−0.049***	0.039	−0.043***	0.123***	−0.046***
	(0.029)	(0.008)	(0.026)	(0.008)	(0.027)	(0.008)
Wind Back (B)	0.121***	−0.075***	−0.023	−0.076***	0.118***	−0.079***
	(0.033)	(0.009)	(0.031)	(0.009)	(0.031)	(0.009)
Weeks since dry dock	0.0003***	0.0003***	0.001***	0.001***	0.0004***	0.0004***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
F : WB	0.013**		0.022***		−0.003	
	(0.006)		(0.005)		(0.006)	
SF : WB	0.009		0.012**		0.004	
	(0.005)		(0.005)		(0.006)	
SB : WB	−0.026***		−0.019***		−0.040***	
	(0.006)		(0.006)		(0.006)	
B : WB	−0.044***		−0.012*		−0.046***	
	(0.007)		(0.007)		(0.007)	
Constant	−0.864***	−0.811***	0.133	0.220***	−1.748***	−1.630***
	(0.062)	(0.061)	(0.085)	(0.083)	(0.077)	(0.076)
Observations	3,432	3,432	3,281	3,281	3,772	3,772
R^2	0.704	0.695	0.559	0.550	0.622	0.611
Adjusted R^2	0.703	0.694	0.556	0.548	0.620	0.609
Note:						*p<0.1; **p<0.05; ***p<0.01

Medium Range vessels. Additionally, vessels increase consumption by between 0.03 - 0.1 % per week, likely due to biofouling, aging, and other sources of deterioration of the hull.

The models with interaction terms between wind speed and direction are considered for further analysis due to the previously explained logic, as well as exhibiting increased adjusted R^2. To examine the performance of OLS regressions compared to other methods, several machine learning techniques are compared in terms of prediction accuracy in the next section.

4.2 Comparison of prediction accuracy

The number of tools available for prediction purposes has grown immensely. While classic methods such as regular OLS regression are still relevant, new prediction models are constantly in the works. During the last couple of decades, the computational power of computers has doubled every two years, following Moore's law (Moore, 1998). This has led to an increase in the commercial availability of advanced linear and non-linear machine learning algorithms. In this section, several of these methods will be compared against OLS regression in terms of prediction accuracy and explanatory power.

Both Nilsson and Nilsson (2021) and Uyanık et al. (2020) explored several machine learning techniques to find which models achieved the highest degree of accuracy given a set of variables. The same dependent variable and regressors, including fixed vessel effects, are applied for all techniques with the same random number generator. Hence, the metrics can be used to determine which models perform the best for predicting fuel consumption. Each algorithm uses a validation set approach, allocating 80 % of observations into a training set for model fitting and parameter tuning, and 20 % into a hold-out test set for each vessel class. Since all metrics are based on out-of-sample accuracy, the prediction errors are unbiased. While K-fold cross validation generally results in less bias, it is substantially more demanding in terms of computational power (James et al., 2013). Table 4.2 show the prediction errors and explanatory power for every technique on all segments.

The number of observations across vessel classes are comparable, ranging from roughly 3,300 to 3,800. In general, advanced regression trees perform the best for all vessel segments, with similar R^2 and RMSE. Extreme Gradient Boosting (xGBoost) and Random Forest perform the best. In terms of linear models, there is little to no difference between standard OLS and shrinkage techniques like Ridge, Lasso or Partial Least Squares. More advanced algorithms, such as xGBoost and Random Forest, are highly dependent on the fine-tuning of hyperparameters to achieve optimal results, while regular OLS is easier to use and interpret. For Panamax vessels, the differences between OLS and the best performing models are moderate, while regression trees decisively outperform OLS for Suezmax and Medium Range vessels. It is interesting to see whether the inclusion of hull cleaning and crew performance can shrink the differences in performance. This is further discussed in section 7.

Table 4.2: Performance metrics for machine learning methods

	Panamax			*Suezmax*			*Medium Range*		
	RMSE	R2	MAE	RMSE	R2	MAE	RMSE	R2	MAE
Linear	0.1571	68.93%	0.1152	0.1415	53.28%	0.1049	0.1541	59.64%	0.1178
Lasso	0.1571	68.96%	0.1151	0.1414	53.31%	0.1049	0.1542	59.56%	0.1178
Ridge	0.1579	68.97%	0.1154	0.1414	53.27%	0.1054	0.1542	59.59%	0.1179
PCR	0.1571	68.95%	0.1152	0.1415	53.28%	0.1049	0.1541	59.64%	0.1178
PLS	0.1610	67.40%	0.1192	0.1417	53.15%	0.1051	0.1545	59.38%	0.1185
Bagging	0.1760	60.90%	0.1233	0.1501	47.39%	0.1164	0.1698	51.17%	0.1252
Extra Trees	0.1368	76.71%	0.0889	0.0999	76.83%	0.0713	0.1232	74.67%	0.0893
Boosting	0.1383	76.43%	0.0908	0.1029	75.59%	0.0738	0.1255	73.46%	0.0896
Random Forest	0.1341	77.89%	0.0863	0.0978	78.49%	0.0702	0.1228	75.46%	0.0895
xGBoost	0.1346	77.34%	0.0873	0.0951	78.90%	0.0682	0.1251	73.62%	0.0854
BART	0.1450	73.63%	0.0966	0.1114	71.03%	0.0807	0.1280	72.08%	0.0960
	n = 3432			*n = 3281*			*n = 3772*		

4.3 Expectations

To gauge the credibility of empirical fuel consumption profiles, it is useful to clarify the thesis' *a priori* expectations. As discussed in section 2, the development of biofouling and worsening conditions of the hull, increases resistance and therefore fuel consumption over time. Hull cleanings are subsequently performed to remove the cause of the increased resistance. After performing a hull cleaning, most of the negative effects connected to biofouling are expected to disappear instantly, leading to a negative and instant shock in fuel consumption levels. While hull cleanings have an impact on short-term fuel consumption, it could also affect the growth rate of consumption over time due to the abrasion of antifouling coating. It is therefore expected that following a hull cleaning, the biofouling rate and hence the increase in consumption over time is either equal or larger than before.

4.4 Classification algorithm

The proposed classification algorithm looks at a vessel's consumption preceding and following port calls, to determine at which port calls the vessel has undergone an underwater hull cleaning. To control for external conditions in consumption, the residuals of the OLS regression in equation 4.1 is considered. The fitted values $\widehat{lnC_t}$ from the regression are given by:

$$\widehat{lnC_t} = \alpha + \sum_i \beta_i X_{it} \tag{4.2}$$

where $\widehat{lnC_t}$ is estimated based on equation 4.1. Controlling observations for all variables using equation 4.2 gives:

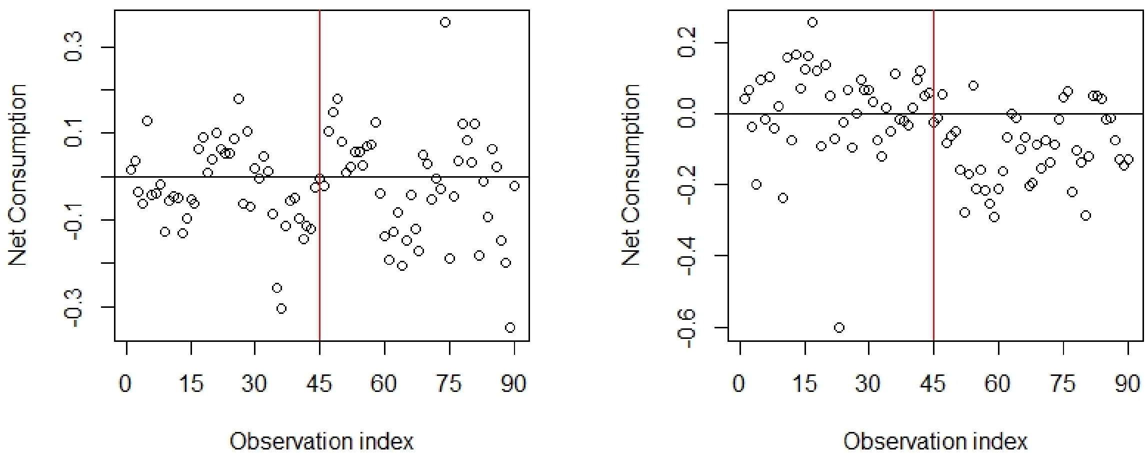
$$lnC_t^* = lnC_t - \widehat{lnC_t} = \varepsilon_t \tag{4.3}$$

Thus, the observed log fuel consumption controlled for external variables is equal to the remaining variance[10], such that lnC_t^* is given by ε_t. The log consumption controlled for all external variables is therefore referred to as lnC_t^*. The classification model will compare the lnC_t^* of the observations closest to a port call to look for significant drops in fuel consumption. For all port calls p of a vessel, the time window to be analyzed is defined as $p \pm w$, where the number of observations considered before and after port calls are specified by w. The algorithm initially uses an OLS-estimate to quantify the effect δ on lnC_t^* before and after the port call p. Similar to the model by Adland et al. (2018) described in section 2, this is given by:

[10]To get the actual controlled fuel consumption, the constant term α would have to be added. There is however no difference in results, but the interpretation of the lnC_t^* values differs slightly.

$$lnC_t^* = \delta \cdot I_t^{AFTER_w} + \epsilon_t \tag{4.4}$$

where $I_t^{AFTER_w}$ is a dummy variable indicating 1 if an observation is after the port call and 0 otherwise, and ϵ_t is the new residual term with $E[\epsilon_t] = 0$ and $Var(\epsilon_t) = \sigma^2$. The hull cleaning effect δ is then further analyzed to see whether the level of consumption before and after the port call has been reduced significantly. Specifically, using a significance level of s, if δ is significant and negative, the port call p is stored as a possible hull cleaning, along with the p-value and coefficient of δ. In figure 4.1a, observations of lnC_t^* before and after vessel P1's first port call are displayed. The first port call does not have a significant drop in log fuel consumption controlled for external variables following the port call (represented by the red vertical line). In figure 4.1b however, a significant drop is detected in the observations following the eighth port call of the same vessel. Port call 8 is therefore classified as a possible hull cleaning.



(a) Port call 1, no significant drop

(b) Port call 8, significant drop

Figure 4.1: Comparison of two port calls for vessel P1

The minimum time passed between two hull cleanings is measured by m, to ensure that the effect of a single hull cleaning is only considered once. If two port calls happen within the same week and a hull cleaning is performed on one of them, both port calls will perceive similar effects on consumption. m is intended to limit the realistic minimum interval between sequential hull cleanings. Hull cleanings are then determined in the order of the most significant drops in fuel consumption, subject to the constraint that another hull cleaning cannot happen within m months before or after. Possible hull cleanings are defined as port calls where δ is negative and significant. The possible hull cleaning with the lowest p-value for δ is automatically identified as a hull cleaning. The second most significant possible hull cleaning is then identified as a hull cleaning, but only if it is more than m days away from the first identified hull cleaning, and so on. The process is repeated for all vessels within the segment, and for all segments.

4.5 Fuel consumption profiles

From the classification algorithm, a vessel's hull cleanings are defined by $k = \{1, 2, 3, \ldots, K\}$, where K is the total number of hull cleanings. Depending on the amount of classified hull cleanings for a unique vessel, observations are arranged into K+1 intervals. In contrast to the model in equation 2.3 proposed by Adland et al. (2017), the effect of each hull cleaning is not added together, but has a unique impact on fuel consumption. Hence, the logarithmic value of fuel consumption lnC_t for an observation in week t, is given as:

$$lnC_t = \alpha + \tau_{kt} \cdot X_{kt} + \theta \cdot Y_t + \varepsilon_t \tag{4.5}$$

where X_{kt} is a dummy variable indicating 1 if week t is between the k^{th} and $k^{th} + 1$ hull cleaning and 0 otherwise, and τ_{kt} is the effect of hull cleaning k at time t. Y_t is a set of vessel characteristics based on the chosen variables with the associated vector of coefficients θ, and ε_t is the residual term with $E[\varepsilon_t] = 0$ and $Var(\varepsilon_t) = \sigma^2$. Within Y_t, one of the vessel characteristics is a linear time trend in the form of weeks since last dry dock. For the purpose of identifying hull cleaning effects, a non-linear time trend was deemed superfluous.

As mentioned in section 4.3, it is expected that the increase in fuel consumption grows steeper for each subsequent hull cleaning. Therefore, an interaction term between the hull cleaning and the linear time trend is added to the equation:

$$\tau_{kt} = \alpha_k + \beta_k \cdot t \tag{4.6}$$

where α_k is the shock in consumption after a given hull clean k, and β_k determines the slope of the hull clean interval dependent on time t. This extends equation 4.5 to:

$$lnC_t = (\alpha_k + \beta_k \cdot t) \cdot X_{kt} + \theta \cdot Y_t + \varepsilon_t \tag{4.7}$$

which is used for visual representation of fuel consumption over time.

4.6 Application of classification model

Before applying the algorithm and model on actual observations, all statistical parameters must be determined. Following the findings of Adland et al. (2018), the number of days w that define the observable time windows is set to 45, as this maximized the marginal effect of hull cleanings. Time windows of less than $2w$ are not considered, to have sufficient and comparable number of observations. The level of significance s is set to 1 %, as this was determined to be a reasonably strict level of certainty. Further, the minimum time interval between hull cleanings m is set to 150 days (about five months), as it is assumed that having more than one hull cleaning within five months does not make economical sense. This is because biofouling has likely not imposed sufficient negative effects on consumption within this period to justify the cost of hull cleaning procedures.

To ensure that remaining observations provide meaningful comparisons, ships with no company-reported hull cleanings are removed, as well as ships with no hull cleaning intervals longer than m. This leads to all remaining Medium Range vessels being excluded from further analysis, as well as one Panamax and three Suezmax vessels. Observations in hull cleaning intervals shorter than m are also removed. In figure 4.2, the green boxes represent dry docks, the orange boxes represent company-reported hull cleaning dates, while blue boxes represent the hull cleaning dates detected by the classification model. The red box indicates that the classification model agrees with the company-reported date. Further, the white intervals represent time periods outside our interest, either due to dry docks or because of an inadequate number of observations following a company-reported hull cleaning. Grey intervals represent periods of interest with observations, in other words the data used in the analysis.

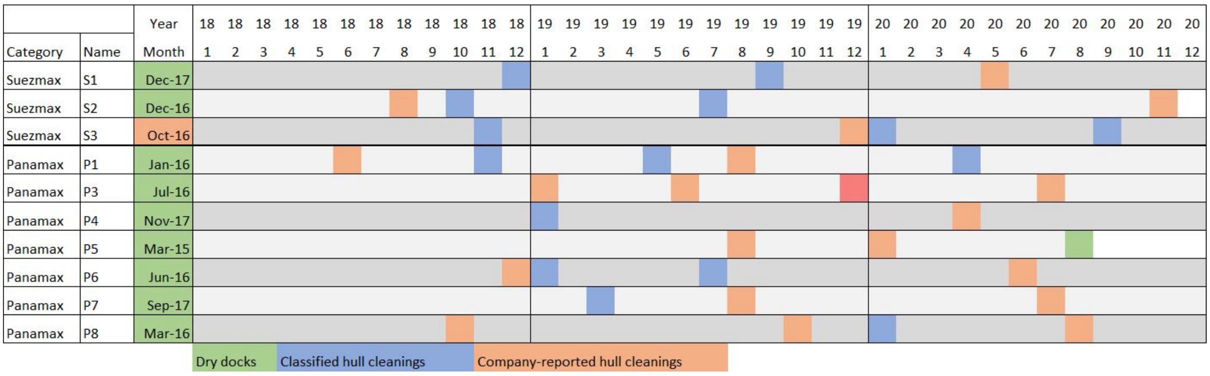


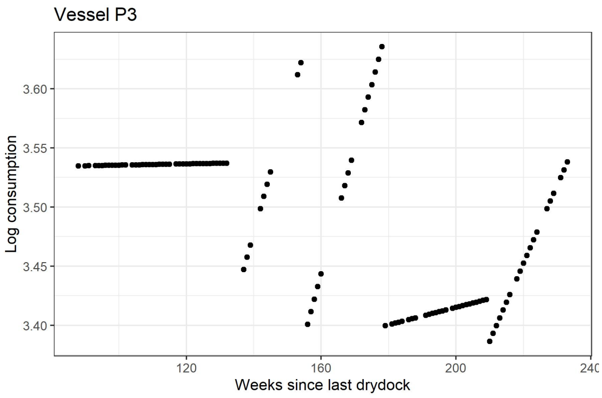
Figure 4.2: Detected and company-reported hull cleaning dates

The classification model generally disagrees with the company-reported dates, as observed in figure 4.2. Only on one occasion does the model agree with reported dates, although several others are rather close. For vessel P5, the model detects no hull cleanings, and the vessel is consequently excluded from further analysis. To measure the extent to which the classification model is successful, fuel consumption profiles of observations using both reported and classified hull cleaning dates are compared.

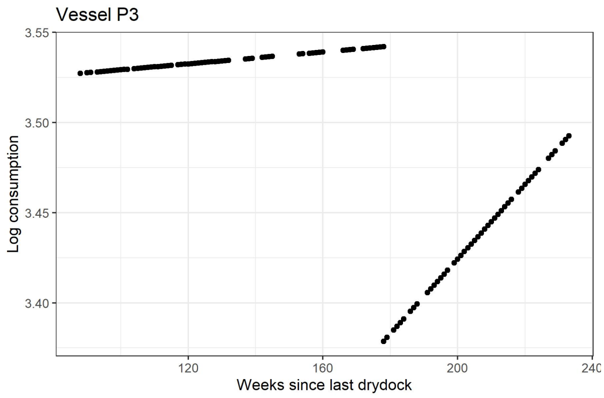
4.7 Comparison of fuel consumption profiles

To identify the effects of hull cleanings for each vessel, the predicted values found by an OLS regression based on equation 4.7 are used. The graphs are drawn holding all external variables constant, thus only varying hull cleanings and time trends. Some fuel consumption profiles are shown in figure 4.3, where jumps in consumption signal hull cleanings and gaps in the slopes represent weeks with no observations. Applying the described model to the dates provided by the company leads to largely nonsensical consumption profiles. Shown on the left-hand side of figure 4.3, consumption is often increasing immediately after hull cleanings and with wildly inconsistent, nonsensical slopes.

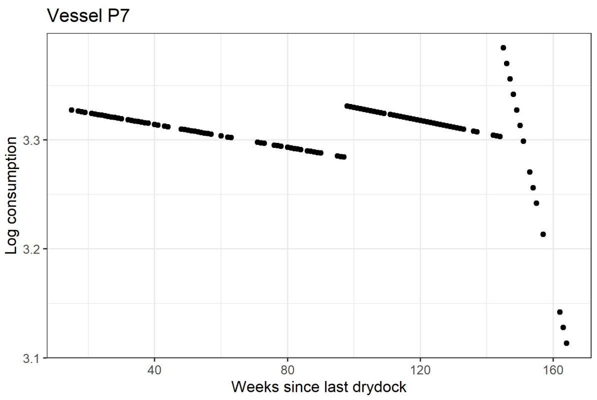
However, when applying the same model to the hull cleaning dates identified by the classification model, the results are markedly better in terms of fitting expectations. Fuel consumption profiles based on the classification model are shown on the right-hand side of figure 4.3. Since the classification model requires a hull cleaning to have a negative effect on fuel consumption, upward shocks are quite rare. It is further apparent that fuel consumption slopes seem to rise for the vast majority of intervals. Negative trends do occur occasionally, which means that there is still some variation that neither the prediction nor classification model is able to account for. Interestingly, there is no recurring evidence of steeper interval slopes following hull cleanings, although figure 4.3b and especially figure 4.3f conforms to expectations. All fuel consumption profiles, with comparisons whenever applicable, are displayed in appendix A3. This indicates that the abrasion of antifouling coating and ensuing hull corrosion may not have as severe effects as previously suggested.



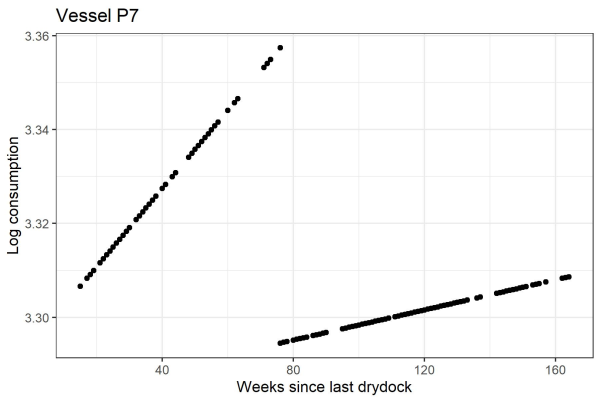
(a) Reported HC vessel P3



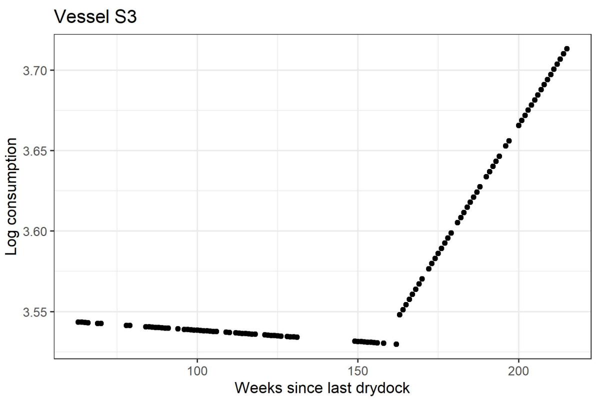
(b) Detected HC vessel P3



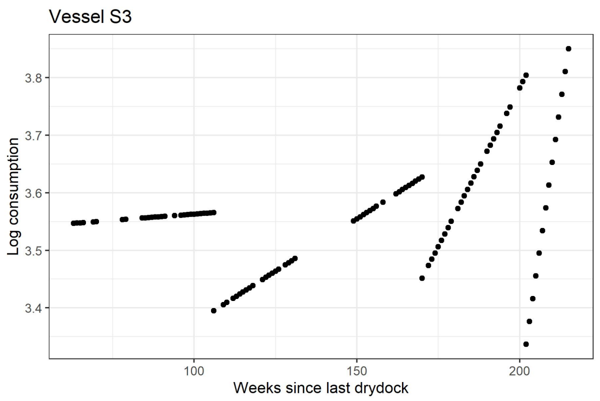
(c) Reported HC vessel P7



(d) Detected HC vessel P7



(e) Reported HC vessel S3



(f) Detected HC vessel S3

Figure 4.3: Comparison of selected fuel consumption profiles

Since the classification model allows for hull cleaning detection on vessels with no reported occurrences, it can be further tested on all vessels with an observable dry dock interval of more than two years. As dry docks are sometimes performed during the three-year period for Medium Range vessels, only observations before the second dry dock are considered. In appendix A3.2 and A3.3, the hull cleanings identified for vessels with no basis of comparison are presented, along with their respective fuel consumption profiles. The algorithm seems to classify hull cleanings that generally yield logical fuel consumption profiles in terms of fitting expectations. The extent to which the dates found by the classification model is able to explain variance in the data is discussed in section 7, where various external effects are measured in terms of prediction power.

5 Optimizing hull cleaning intervals

The optimal timing of hull cleanings remains an important unresolved issue within shipping. While the sheer complexity of the numerous factors affecting vessel performance makes this a hard problem to solve, the possible cost savings could be substantial. This section aims to find economical optimizations of hull cleaning intervals, based on the findings in section 4.

5.1 Assumptions

Section 4.5 presented fuel consumption profiles based on dates found by the classification model. However, random variations and unknown effects lead to certain traits in the profiles that are likely not present. Fuel consumption slopes, as seen in figure 4.3 and appendix A3, do not seem to follow any obvious pattern in terms of steepness following a hull cleaning. In practical terms, an economical optimization for a vessel that follows a pattern which is illogical and likely a result of arbitrary or unknown variation, is not of any value. To illustrate this further: if a master is considering performing a hull cleaning, but the optimal timing according to a given decision rule is dependent on whether the unknown post-cleaning slope is increasing or decreasing in steepness, the decision rule is worthless. Additionally, it is not intuitive for a vessel to consume less fuel after a regular hull cleaning compared to a dry dock [11]. In other words, to create a decision rule based on economical optimization, certain assumptions about the behavior of fuel consumption slopes have to be in place. This means that the optimization model to some degree strays away from the empirical findings in section 4.5, and instead explores theoretical models to improve applicability. This section initially presupposes fuel consumption to be weekly aggregated for the duration of a three-year interval, which is further discussed later.

Although Adland et al. (2018) found a 9 % drop in fuel consumption following a hull cleaning, the findings assume that the hull cleanings are spread out to a certain extent. If two hull cleanings were performed within two months of each other, one would expect the second hull cleaning to reduce consumption by substantially less than 9 %. This is because biofouling would not have had time to impose significant negative effects on performance.

[11] Some self-polishing coatings need time to smooth out to reach the point of minimum resistance, resulting in resistance reductions in the weeks following a dry dock.

Assuming that hull cleanings are efficient at restoring the hull condition to near dry dock levels, it is expected that the absolute reduction in fuel consumption is a function of time, but never dropping below dry-dock levels. To avoid the problem of subsequent hull cleanings, two assumptions are implemented. First, hull cleaning shocks are measured as a percentage drop from the difference between the current consumption level and the initial dry dock level. For instance, if the hull cleaning shock parameter[12] is $\delta = 0.5$, the initial consumption level is $C_0 = 50$, and the consumption of the previous week w is $C_{w-1} = 56$, the post-shock consumption level would be $C_w = 56 - (56 - 50) \cdot 0.5 = 53$. This example is further illustrated by figure 5.1, where the delta (δ) leads to a drop equal to half of the periodic increase. Second, to further prevent unnatural cleaning patterns, the hull cleaning intervals are not allowed to occur more than once every m weeks.

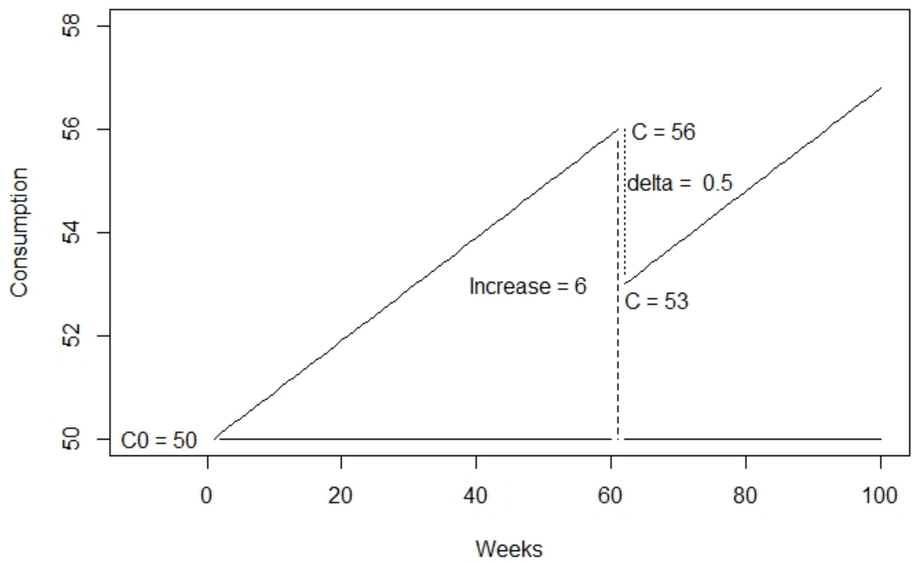


Figure 5.1: Hull cleaning shock parameter δ

The consumption level used is the back-transformed value of log consumption controlled for external variables, explained in section 4.4. For presentation purposes, logarithmic values are sometimes difficult to interpret, which is why the optimization model utilizes back-transformed values. This includes weekly consumption and growth rates based on data from the Panamax class, even though they are still hypothetical estimates. Further, the slope numbers of fuel consumption profiles are henceforth assumed to be constant,

[12] δ is in this section redefined compared to section 4, although they are related in terms of purpose.

due to the lack of evidence of steeper slopes in the empirical analysis in section 4.5.

5.2 Theoretical optimization model

This thesis will now build further upon the model created by Adland et al. (2017), presented in section 2. Contrary to their model, weeks are used as the measurement of time. This is because it is not necessary to determine the exact day of the optimal hull cleaning, and because weekly aggregated data makes the model less computationally demanding. Weeks and hull cleanings are denoted as w and k respectively, where weeks are given as $w \in W = \{1, 2, 3, \ldots, n\}$ and hull cleanings are given as $k \in K = \{1, 2, \ldots, h\}$. n is the number of weeks and h is the number of hull cleanings. The objective will thus be to minimize the sum of the vessel's consumption across all weeks, given by:

$$MIN : \sum_{w \in W} C_w \tag{5.1}$$

The consumption level C_w for week w is given subject to:

$$C_w = C_{w-1} + l - (C_{w-1} - C_0) \cdot \delta \cdot X_{k,w} \qquad \forall\, w \in W, k \in K \tag{5.2}$$

where l is the absolute increase in consumption per week, and $X_{k,w}$ is dummy variable indicating 1 if hull cleaning k is performed in week w and 0 otherwise. l represents the increased vessel resistance due to biofouling and other negative externalities affecting the hull and propeller.

5.3 Interval optimization

As previously discussed, the optimization is primarily designed based on hypothetical assumptions. The initial consumption C_0 and the consumption increase l are however empirically estimated to give a reasonably realistic interpretation of consumption over time. In the model, consumption is only observed once each week, whereas it is realistically reported anywhere from zero to seven days a week. Using the average of the ten first fitted values[13] for each Panamax vessel, the initial daily consumption level is approximated to be

[13] Estimated using back-transformed values of equation 4.7 from section 4.5

30 tons a day. To better compare initial weekly consumption C_0 with reality, the empirical daily consumption is multiplied by four, since the average number of observations per week for a given vessel is $3.65 \approx 4$. Thus, C_0 is set to $30 \cdot 4 = 120$ tons/week. The monthly consumption increase is estimated to be between 0.15 % - 3 % of C_0, based on the findings of Gundermann and Dirksen (2016) on added monthly resistance caused by biofouling.

The hull cleaning shock parameter δ is examined using values between 0.3 – 0.9. Since hull cleaning shocks are only described empirically using regular percentage drops, this estimate range is wide and based solely on achieving the desired behavior. The number of weeks n is tested for values corresponding to 2, 3 and 4 years, in other words $n = 156 \pm 52$, while the number of hull cleanings h is set to either 2 or 3. Further, the minimum number of weeks allowed between each hull cleaning is set to $m = 24$, which is approximately six months. The parameters n, h and m are heavily restricted because increasing either of these is exponentially more computationally demanding due to the number of possible combinations. This is why the interval constraint m is larger than the five months assumed in section 4.6.

Running the optimization with a constant fuel consumption increase of[14] $l = 0.3$ and a shock percentage of $\delta = 0.5$, the graph in figure 5.2 represents the optimal hull cleaning intervals. Note that the x-axis represents the percentiles of the optimal hull cleanings over a given period. Initially, a three-year period (156 weeks) is assumed. If for instance the first hull cleaning happens in week 56, this corresponds to the 36$^{\text{th}}$ interval percentile as $56/156 \approx 36$ %. Similarly, if the second hull cleaning happens in week 99, it corresponds to the 63$^{\text{rd}}$ percentile ($99/156 \approx 63$ %). This is done to provide a more general interpretation of the optimization which remains the same for all values of n.

In figure 5.2, there is an obvious pattern present. Hull cleaning intervals with constant slopes are decided by a certain threshold after which it will always drop, and will consequently always drop to the same consumption level. The pattern from figure 5.2 is in fact always present when holding slope numbers constant, unless other constraints are binding. Further, as long as the slope is constant the graph will always be centred around the middle, thus resembling an ambigram. It is therefore worth noting that due to the theoretical nature of the optimization, the optimal intervals are equal regardless of

[14]Based on monthly increase of 1 %, $l = C_0 \cdot \frac{0.01}{4}$.

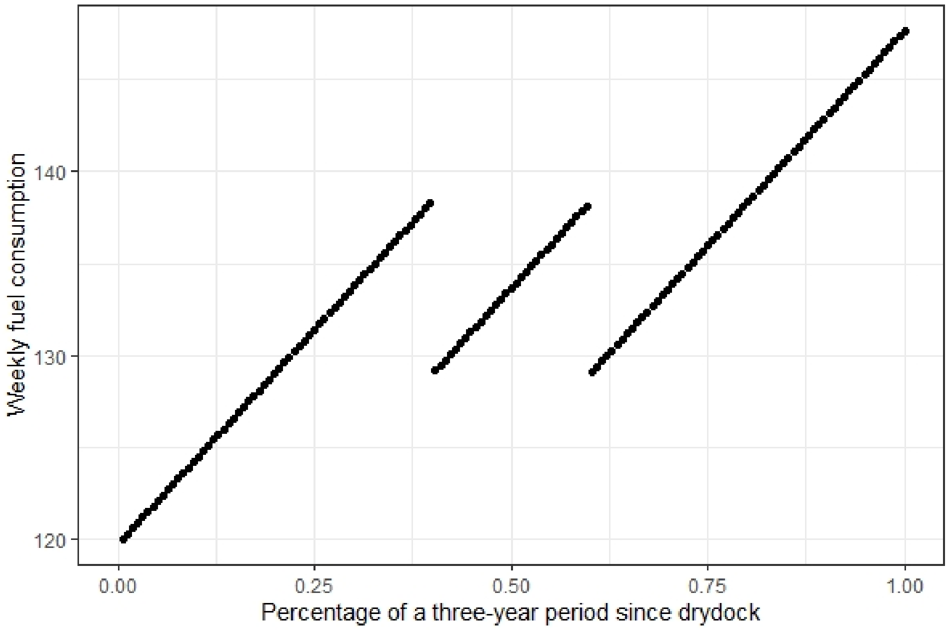


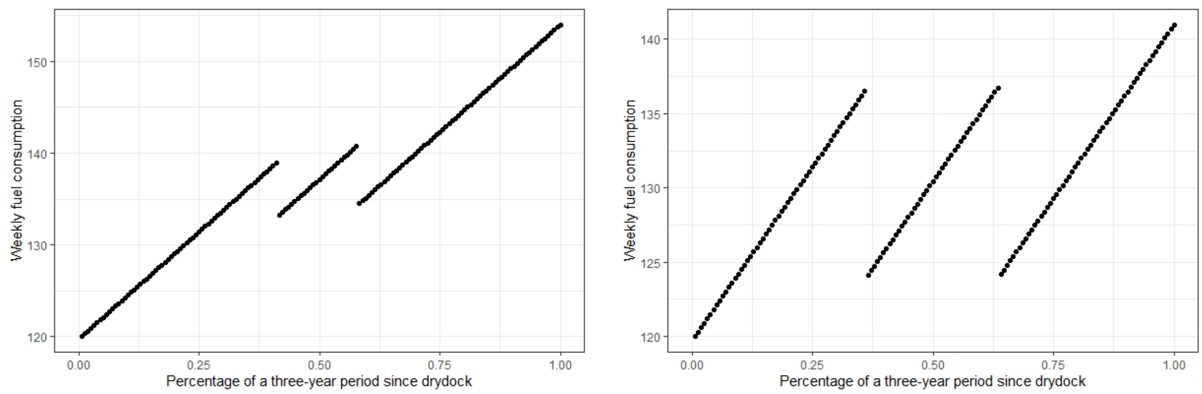
Figure 5.2: Optimization with $\delta = 0.5$

the initial consumption estimates C_0 and l. Figure 5.2 would consequently be identical for any value of these, with the only change being a different scale on the y-axis. For purposes other than estimation of cost benefits, it is therefore not necessary to estimate these parameters.

5.3.1 Changing the shock size

When looking at different shock sizes, it is apparent that the magnitude has an effect on the optimal interval. In figure 5.3a, the drop is only 25 %, and differs from the described pattern as the minimum interval constraint m is binding. From figures 5.3a and 5.3b, it is also noticeable that both hull cleaning intervals move toward the middle of the dry-dock interval as the shock decreases. In general, the sensitivity of hull cleaning intervals for different shocks is shown in Appendix A4.1.

In section 4.3, preliminary expectations of consumption growth are discussed. However, if the slope numbers are not constant, the behavior of optimal intervals differ. In appendix A4.2, it is briefly discussed what happens to the optimal intervals when faced with increased consumption growth after each subsequent hull cleaning. Keeping shocks constant, the optimal hull cleaning intervals are also dependent of the length of the period n, although the effect is rather small. For periods between 2-4 years, the deviations in optimal intervals



(a) $\delta = 0.25$ **(b)** $\delta = 0.75$

Figure 5.3: Changing shock size δ

vary by less than 1 percentage point for the first shock, and remains exactly equal for the second shock. For the scenario in which three hull cleanings are performed, the corresponding graphs are available in appendix A4.4, with the same interpretations as presented in this section.

5.3.2 Changing the number of hull cleanings

So far, the analysis has looked at the optimal timings given different shock sizes for performed hull cleanings. It is further interesting to observe the difference in total fuel consumption when varying the number of hull cleanings. All calculations are based on the optimal intervals found in the last sections.

Assuming a weekly growth of 0.25 % ($l = 0.3$), a single Panamax vessel with no hull cleanings would amass 22,347 tons of fuel consumption over a three-year period. When changing the shock parameter, the decrease in overall consumption performing three instead of two hull cleanings, ranges between 1.31 % and 1.80 %. The largest effects are seen with $\delta = 0.55$ and $\delta = 0.6$, with fuel savings of 1.80 %. This leads to savings of around 370 tons of fuel over three years. Whether the fuel savings are worth it, depends on the trade-off between hull cleaning and fuel costs. All consumption differences when varying the shock parameter δ are displayed in table 5.1.

Table 5.1: Effects on consumption with different shock sizes

	Two hull cleanings			*Three hull cleanings*				*Effect*	
Shock	Consumption	HC1	HC2	Consumption	HC1	HC2	HC3	Difference	Percentage
0.25	21507.26	40%	57%	21226.13	32%	49%	65%	281.13	1.31%
0.3	21367.51	41%	58%	21057.21	33%	49%	66%	310.3	1.45%
0.35	21233.99	41%	58%	20900.57	33%	49%	66%	333.42	1.57%
0.4	21106.7	41%	58%	20755.62	33%	49%	66%	351.08	1.66%
0.45	20984.95	40%	59%	20621.76	33%	49%	66%	363.19	1.73%
0.5	20868.08	40%	60%	20498.4	33%	49%	66%	369.68	1.77%
0.55	20755.8	39%	61%	20383.67	32%	50%	68%	372.13	1.79%
0.6	20647.8	38%	62%	20276.11	31%	50%	69%	371.69	1.80%
0.65	20543.89	38%	62%	20175.06	30%	50%	70%	368.83	1.80%
0.7	20443.84	37%	63%	20079.97	29%	50%	71%	363.87	1.78%
0.75	20347.41	36%	63%	19990.35	28%	50%	71%	357.06	1.75%
0.8	20254.43	35%	64%	19905.64	28%	50%	72%	348.79	1.72%
0.85	20164.7	35%	65%	19825.5	27%	50%	73%	339.2	1.68%
0.9	20078.05	35%	65%	19749.6	26%	50%	74%	328.45	1.64%

Furthermore, it could be valuable to observe the sensitivity of performing two or three hull cleanings when changing l. In table 5.2, it is apparent that fuel savings grow bigger when fuel consumption increases. Assuming an increase of $l = 0.05$ tons per week, the difference between performing two and three hull cleanings is only 0.31 %. Meanwhile, if $l = 0.9$, increasing to three hull cleanings results in a reduction of 4.93 % in total fuel consumption. A reasonable consumption increase of $l = 0.3$ would lead to a 1.75 % decrease in total consumption. Compared to performing no hull cleanings at all, the decrease would be 8.9 % and 10.5 % for two and three hull cleanings respectively. In other words, there is a noticeable difference in consumption depending on the number of hull cleanings performed.

Table 5.2: Effects on consumption with different consumption increases

	Two hull cleanings			*Three hull cleanings*				*Effect*	
Growth	Consumption	HC1	HC2	Consumtpion	HC1	HC2	HC3	Change	Percentage
0.05	18991.23	36%	63%	18931.72	28%	50%	71%	59.51	0.31%
0.1	19262.47	36%	63%	19143.45	28%	50%	71%	119.02	0.62%
0.15	19533.7	36%	63%	19355.17	28%	50%	71%	178.53	0.91%
0.2	19804.94	36%	63%	19566.9	28%	50%	71%	238.04	1.20%
0.25	20076.17	36%	63%	19778.62	28%	50%	71%	297.55	1.48%
0.3	20347.41	36%	63%	19990.35	28%	50%	71%	357.06	1.75%
0.35	20618.64	36%	63%	20202.07	28%	50%	71%	416.57	2.02%
0.4	20889.88	36%	63%	20413.79	28%	50%	71%	476.09	2.28%
0.45	21161.11	36%	63%	20625.52	28%	50%	71%	535.59	2.53%
0.5	21432.34	36%	63%	20837.24	28%	50%	71%	595.1	2.78%
0.55	21703.58	36%	63%	21048.97	28%	50%	71%	654.61	3.02%
0.6	21974.81	36%	63%	21260.69	28%	50%	71%	714.12	3.25%
0.65	22246.05	36%	63%	21472.41	28%	50%	71%	773.64	3.48%
0.7	22517.28	36%	63%	21684.14	28%	50%	71%	833.14	3.70%
0.75	22788.52	36%	63%	21895.86	28%	50%	71%	892.66	3.92%
0.8	23059.75	36%	63%	22107.59	28%	50%	71%	952.16	4.13%
0.85	23330.98	36%	63%	22319.31	28%	50%	71%	1011.67	4.34%
0.9	23602.22	36%	63%	22531.04	28%	50%	71%	1071.18	4.54%
0.95	23873.45	36%	63%	22742.76	28%	50%	71%	1130.69	4.74%
1	24144.69	36%	63%	22954.48	28%	50%	71%	1190.21	4.93%

5.4 Case study

Whilst this optimization is highly theoretical with several strict assumptions, it would be interesting to see the effect of optimal hull cleaning intervals on actual vessels. The classification model detected two performed hull cleanings on two separate Panamax vessels, P2 and P6. These vessels are further used in an empirical study based on the theoretical optimization of hull cleanings. The case study examines how moving the two hull cleanings to their respective optima affects total fuel consumption, as well as exploring the impact of adding a third hull cleaning.

For vessel P2, the classification model detected hull cleanings at week 164 and 248. The observations started when P2 had 118 weeks since last dry dock. For this vessel to be compared to the theoretical optima, hull cleanings are instead assumed to happen in week 46 (29 %) and 130 (83 %)[15]. From section 5.3, it is given that the optimal hull cleaning intervals with $\delta = 0.75$ and $l = 0.3$, occurs in week 56 (36 %) and 99 (63 %). In this section, only the consumption increase l is tested in terms of sensitivity, keeping δ constant. Table 5.3 displays the percentage savings in consumption for both vessel P2 and P6.

Table 5.3: Effects of optimal timings with two hull cleanings

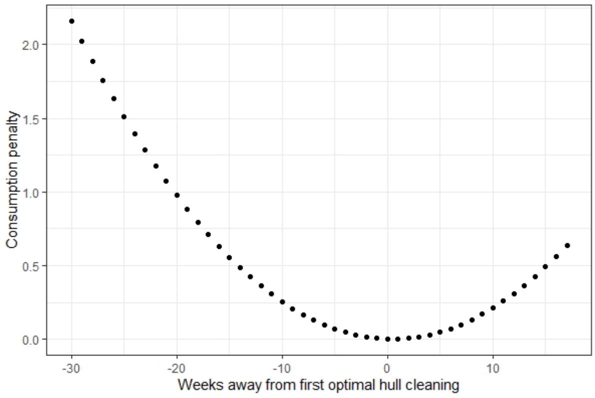
	Actual timing			*Optimal timing*			*Effect*	
	Consumption	HC1	HC2	Consumption	HC1	HC2	l	Difference
Vessel P2	20641.09	29%	83%	20347.41	36%	63%	0.3	1.42%
	21921.81	29%	83%	21432.34	36%	63%	0.5	2.23%
	23202.54	29%	83%	22517.28	36%	63%	0.7	2.95%
Vessel P6	20415.21	35%	52%	20347.41	36%	63%	0.3	0.33%
	21545.34	35%	52%	21432.34	36%	63%	0.5	0.52%
	22675.48	35%	52%	22517.28	36%	63%	0.7	0.70%

Expectedly, the closer the classified hull cleanings are to the optimal interval, the lower the savings in consumption are. Vessel P6 performs its first hull cleaning almost at the optimal time, while the second is performed 11 percentage points (18 weeks) too early. This results in savings of 0.33 % in total consumption over three years. On the other hand, vessel P2 misses the first hull cleaning by seven percentage points (12 weeks), and performs the second hull cleaning 20 percentage points (31 weeks) too late. The difference in total consumption is therefore 1.42 %, which indicates that the vessel could have saved over 263 tons of fuel over three years. The same principles are applied for consumption increases of $l = 0.5$ and $l = 0.7$, for which the potential fuel savings are even bigger for both vessels.

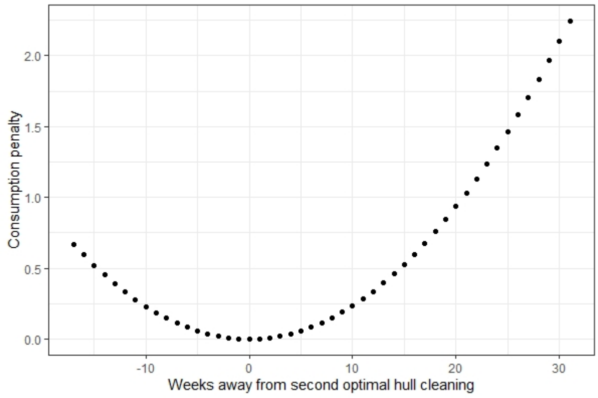
In figure 5.4, the sensitivity of fuel consumption loss regarding hull cleaning timings are displayed. For the first hull cleaning in figure 5.4a, the extra expenditure exponentially decreases from over 2 % when performed 30 weeks early, to 0 % at the optimal timing. When performing hull cleanings later than the optimal timing, fuel consumption is once

[15] All optimizations are performed on a 156-week interval starting in week 1. Since most vessels have performed their first dry dock several weeks before the observed period, direct comparisons cannot be made. The observed weeks are consequently scaled so hull cleanings happen in the same relative interval.

again increasing exponentially. Similar conclusions can be drawn when looking at the timing of the second hull cleaning in figure 5.4b, however, it is mirrored compared to figure 5.4a.



(a) First hull cleaning



(b) Second hull cleaning

Figure 5.4: Increase in consumption when a HC deviates from the optimal timing

Finally, increasing from two to three hull cleanings within the three-year period has considerable advantages in terms of fuel savings. Table 5.4 displays the differences in consumption for both vessel P2 and P6 using three different values of l. The difference is predictably substantially larger than simply performing the two hull cleaning intervals optimally. Adding an extra hull cleaning increases fuel savings with between 2.08 % to 4.37 % depending on the consumption increase l for vessel P6. For vessel P2, the corresponding savings are 3.15 % to 6.54 %. As previously discussed, the growth rate has no impact on the optimal intervals, but affects the total savings over the three-year period.

Table 5.4: Effects of optimal timings with three hull cleanings

	Actual timing			*Optimal timing*				*Effect*	
	Consumption	HC1	HC2	Consumption	HC1	HC2	HC3	l	Difference
Vessel P2	20641.09	29%	83%	19990.35	28%	50%	71%	0.3	3.15%
	21921.81	29%	83%	20837.24	28%	50%	71%	0.5	4.95%
	23202.54	29%	83%	21684.14	28%	50%	71%	0.7	6.54%
Vessel P6	20415.21	35%	52%	19990.35	28%	50%	71%	0.3	2.08%
	21545.34	35%	52%	20837.24	28%	50%	71%	0.5	3.29%
	22675.48	35%	52%	21684.14	28%	50%	71%	0.7	4.37%

6 Impact of individual crew performance

In this section, individual differences between key crew members are analyzed. To find the effects of key crew members on bunker consumption, it is again desirable to control for as many external factors as possible. Hull cleanings must be included to make fair comparisons between masters, since a master sailing right after a hull cleaning is expected to have lower fuel consumption relative to a master sailing right before a hull cleaning. Consequently, once hull cleanings have been identified, the impact of individual crew members on fuel consumption can be found.

In this section, the differences in fuel consumption between masters and chief engineers are statistically tested to find outliers. For the purpose of identifying individual masters, pairing masters and chief engineers would not isolate the effects of single crew members. Further, using a standard OLS regression approach with dummy variables for each master would not measure a master's deviation from the mean, but rather measure the deviation from the mean of the arbitrary reference master. Thus, a new testing scheme is proposed, using ANOVA and t-testing. When finding the effects of masters, it could be argued that the omittance of chief engineers could lead to biased estimators following the problem of *omitted variable bias* (Hanck et al., 2019). The assumption is therefore made that masters are, by and large, responsible for a vessel's consumption. This suggests that if a master is surrounded by only inefficient (or excellent) chief engineers, the deviation in terms of consumption is still credited to the master. Chief engineers are then tested in terms of deviation from their respective master's mean.

6.1 Analysis of masters

The concept of consumption controlled for all external variables (C_t^*) will be further used to separate masters in terms of fuel consumption. Using a similar regression to equations 4.3 and 4.5, C_t^* is estimated by:

$$C_t = \tau_{kt} \cdot X_{kt} + \theta \cdot Y_t + \varepsilon_t \tag{6.1}$$

$$C_t^* = \varepsilon_t \tag{6.2}$$

The regression model uses the non-transformed variables consumption and speed. While log transforming these variables likely represents the actual relationship between speed and consumption better, it is somewhat penalized in terms of inference validity. This is further discussed in section 8. The residuals of the OLS regressions are again considered as the consumption controlled for external variables for each vessel, with a mean of zero. C_t^* is in this case further scaled so that each vessel has $\sigma^2 = 1$. This is to ensure comparable C_t^* values between vessels within the same segment.

The regression model is estimated on each specific vessel. Only masters that have more than 50 observations are included. This implies that external effects such as weather likely differs slightly from ship to ship. This is however necessary when including the interval specific time trends, as the various vessels have different hull cleaning dates which would lead to nonsensical results if time trends were pooled for an entire segment. When looking at all observations from a specific vessel, we nevertheless assume that there are enough observations to avoid overfitting weather data. In addition, vessel fixed effects are automatically taken care of with this approach. In figures 6.1 to 6.3, boxplots are shown to compare the means and variance of C_t^* for different masters.

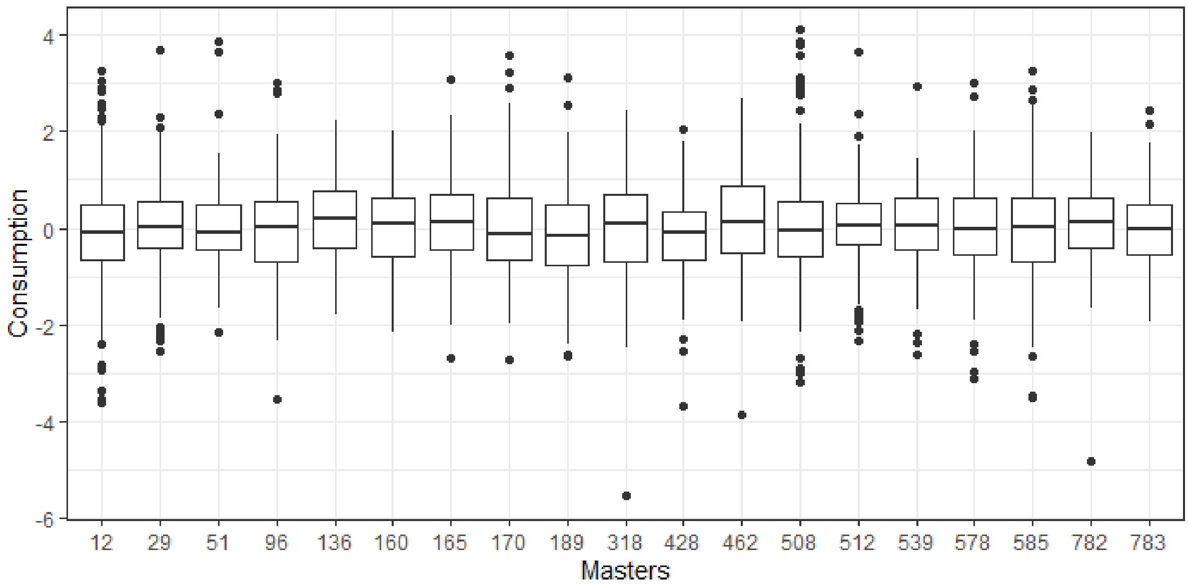


Figure 6.1: Boxplot of Panamax masters' mean consumption controlled for external variables

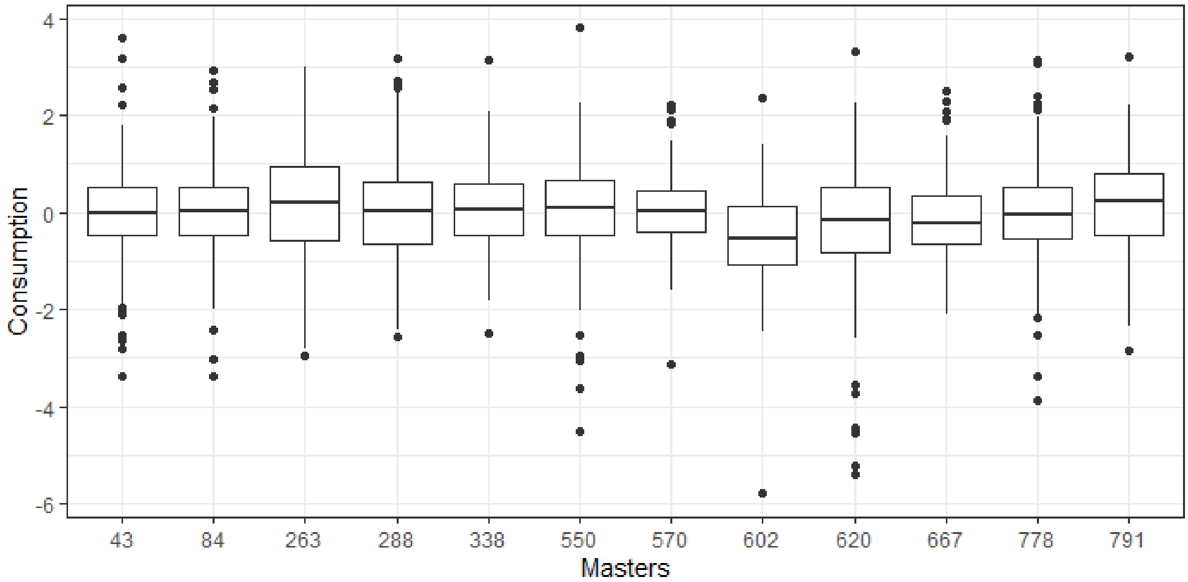


Figure 6.2: Boxplot of Suezmax masters' mean consumption controlled for external variables

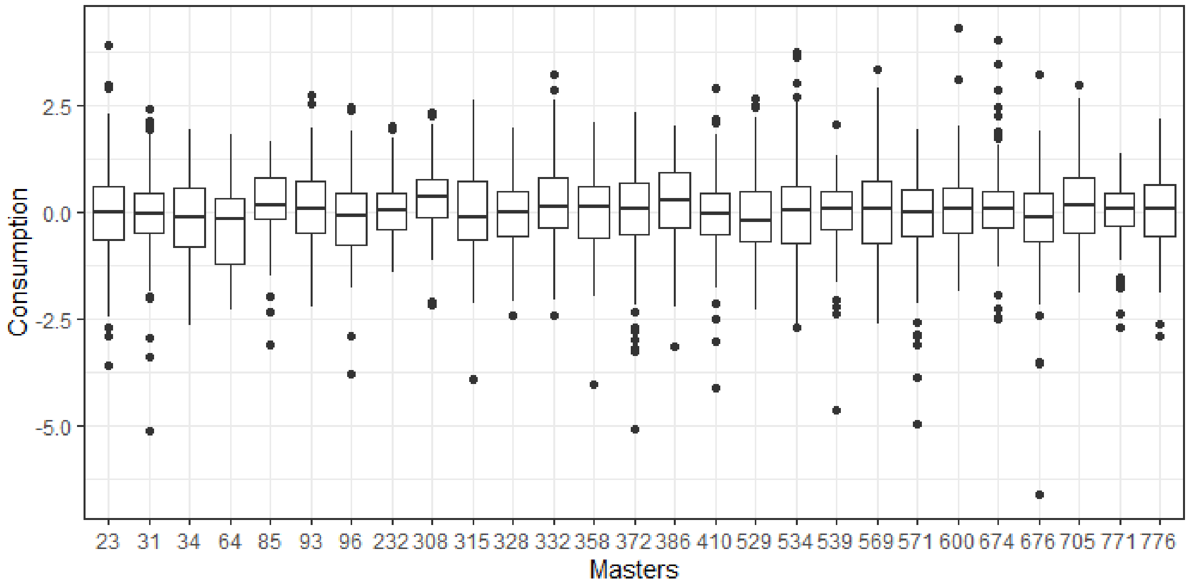


Figure 6.3: Boxplot of Medium Range masters' mean consumption controlled for external variables

For Medium Range vessels, master 308 appears to have the highest consumption, while for Suezmax vessels master 602 seems to have the lowest. For Panamax vessels and other masters, the differences seem to be somewhat less obvious and implores the question of whether or not there are actually significant differences between masters within the segment. A sensible first step is consequently to consider the C_t^* of masters in an ANOVA.

6.1.1 ANOVA testing

Prior to proceeding with statistical testing, it should be clear that the data follows the normality assumptions of ANOVA and t-testing to ensure valid inference. It follows from the central limit theorem that for sample sizes above $n = 30$, the sampling distribution approximates the standard normal distribution (Kwak and Kim, 2017). Autocorrelation is still present in C_t^* values, which will not prevent unbiased estimators, although inference validity may suffer slightly. The residuals are homoscedastic, achieved when setting $\sigma^2 = 1$. By plotting histograms of C_t^* values, it is obvious that observations are indeed resembling a normal distribution. With a sample size of $n \geq 50$ for all masters, the observations for each specific master should be normally distributed as well.

A one-sided ANOVA is fit to determine whether any of the masters have significantly different means (McDonald, 2014). The null hypothesis is that all masters have equal means, while the alternative hypothesis is that at least one master deviates in terms of mean C_t^*. For Panamax vessels, the ANOVA is not rejected, with a p-value of 0.83. Thus, we cannot conclude that there are definitely differences between Panamax masters. For Suezmax and Medium Range masters, the ANOVA is rejected with p-values of $2.5 \cdot 10^{-4}$ and $3.5 \cdot 10^{-5}$ respectively. Considering that the ANOVA only tests to see whether there are differences in means relative to each other, it is not certain that any one master significantly differs in terms of C_t^*. Consequently, it is of interest to examine which masters deviate significantly from the expected mean of zero.

6.1.2 T-testing

Using the scaled residuals, a two-sided t-test assuming equal variance is run on every master. The test observator t is given by:

$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}} \tag{6.3}$$

where $\bar{x}$ is the mean of the n observations for each master within the vessel segment, and μ is the assumed mean of all masters. Since the assumed mean of C_t^* is $\mu = 0$, we have $t = \frac{\bar{x}}{s_{\bar{x}}}$. The estimated standard error of the mean $s_{\bar{x}} = \frac{s}{\sqrt{n}}$ is based on the standard deviation s and size n of the sample. The test observator t is compared with the student's t-distribution with $n-1$ degrees of freedom, and a significance level of 1 %. Note that masters are still separated segment-wise, implying that a master could appear in more than one segment. An overview of masters with C_t^* significantly different from zero, along with the p-value and their respective average C_t^*, is displayed in table 6.1:

Table 6.1: Masters with significantly different consumption

	1 % significance level		
Class	Master	Avg. controlled consumption	p-value
Suezmax	602	-0.4628	0.0014
Medium Range	308	0.4625	0.0004
Medium Range	332	0.2038	0.0055

Some of the masters with low p-values are expected from the boxplots, such as masters 308 and 602. The results suggests that some masters, even after controlling for all known covariates, are still outliers in terms of consumption. Although the causality of the differences in fuel consumption is unknown, these masters have significantly different C_t^* in comparison to their colleagues during the period of observations. As a side note, when running the same test with log transformed variables, the exact same masters are significant. In appendix A5, results are shown for tests with a significance level of 5 %.

6.2 Analysis of chief engineers

Keeping in mind that masters have different C_t^* means, chief engineers will be compared against other chief engineers sailing under the same master. Vessel segments are still separated, but all observations of a master within the same segment are considered. The data is further filtered so that chief engineers sailing under the same master have a sample size of $n \geq 30$. This is to ensure that each chief engineer has a representative number of observations, as well as satisfying normality assumptions for testing in accordance with the central limit theorem (Kwak and Kim, 2017).

First, an ANOVA test is done to ascertain the presence of discrepancies in means between chief engineers of the same master. If the null hypothesis of equal means is rejected, the mean of each chief is tested using a two-sided t-test. However, instead of comparing the sample mean to zero, it is compared to the mean of their master. Thus, we have $\mu = \bar{x}$

compared to the mean of the chief $\bar{y}$, which gives the test observator:

$$t = \frac{\bar{y} - \bar{x}}{s_{\bar{y}}} \tag{6.4}$$

where t is tested for significance in the same way as before. If there are only one or two chief engineers with sufficient observations for a given master, the ANOVA test is skipped, and the chiefs are tested only using the described t-test. The chief engineers with average C_t^* significantly different from the respective master's mean are displayed in table 6.2, along with the p-value and difference in mean consumption, $\bar{y} - \bar{x}$.

Table 6.2: Chief engineers with significant differences in consumption

1 % significance level			
Class	Chief engineer	Difference from avg. consumption	p-value
Panamax	547	-0.5115	0.0012
Panamax	319	0.3360	0.0002
Suezmax	430	0.2624	4.03e-05
Suezmax	434	-0.4334	0.0004
Suezmax	430	-0.4891	1.36e-07
Suezmax	434	0.7057	1.41e-05
Suezmax	82	0.2626	0.0002
Suezmax	169	-0.2140	0.0078
Suezmax	392	-0.3553	0.0012
Suezmax	463	0.2790	0.0017
Suezmax	430	0.6205	6.16e-06
Suezmax	152	-0.2054	0.0013
Medium Range	223	0.5800	8.63e-05

For chief engineers, most of the significant differences appears in the Suezmax segment. Interestingly, engineers 434 and 430 show significant differences for two and three different masters, but with different signs. This is perhaps signaling that the effects of chief engineers could be largely due to unknown external factors and arbitrary variance. Log transforming variables leads to similar results, except that chief engineer 637 has a significant difference, while chief engineers 169 and 392 are no longer significant. Since chief engineers are only tested if the ANOVA is rejected, setting the significance level to 5 % leads to the identification of more chief engineers. This also applies to chief engineers with p-values lower than 1 %, as more ANOVAs are rejected. The chief engineers identified using a significance level of 5 % are presented in Appendix A5.

7 Results and discussion

In this section, the total impact of hull cleaning and crew members are quantified. Some of the prediction methods used in section 4.2 will be utilized again to see whether the inclusion of the thesis' findings can achieve higher prediction accuracy.

7.1 Regression output

To include the effects of crew in a fixed effects model, some assumptions are made. Adland et al. (2016) considered the matched effects of charterers and owners on freight rates using fixed effects models. Matched pair dummy variables for each buyer and seller were implemented to avoid *omitted variable bias*. In this case, with around 95 masters and 80 chief engineers, creating matched pairs would lead to exceptionally high fragmentation of data, as the number of dummies becomes extremely large. Such estimations would also be severely computationally demanding. A fixed effects OLS regression model building upon equation 4.1 and 4.5 is instead run with separate dummies for masters and chief engineers, and is given by:

$$lnC_{vt} = \alpha + \tau_k \cdot X_{vk} + \theta \cdot Y_{vt} + \eta_{xy} \cdot I_{xy} + \vartheta_v + \varepsilon_{vt} \tag{7.1}$$

where I_{xy} is a dummy variable equal to 1 for observations with master x and chief engineer y, and η_{xy} is the fixed effect coefficient for the same crew combination. $\theta \cdot Y_{vt}$ represents vessel characteristics for vessel v at time t. Note that the hull cleaning effects are no longer including interval specific time trends, meaning that $\tau_k = \alpha_k$. This is because estimating all of the different vessels' interval specific time trends would be nonsensical when individual timings differ. The results of the fixed effects model are displayed in tables 7.1 to 7.3, where different effects are added separately. Total effects are also estimated separately under ballast and laden conditions. Since some vessels and observations have been removed during the classification section, new regressions are fit with the updated data set without adding hull cleaning and crew performance as variables.

Table 7.1: Regression for different sets of regressors with fixed effects (FE) for Panamax vessels

	Dependent variable:						
	Log consumption						
	No FE	FE	Weather	HC	Crew	Ballast	Laden
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Speed	1.321***	1.260***	1.407***	1.410***	1.396***	1.467***	1.063***
	(0.028)	(0.026)	(0.023)	(0.023)	(0.024)	(0.036)	(0.037)
Draught	0.051***	0.055***	0.046***	0.045***	0.046***	0.013	0.038***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.008)	(0.003)
HC1				−0.096***	−0.134***	−0.221***	−0.064***
				(0.010)	(0.014)	(0.024)	(0.017)
HC2				−0.085***	−0.217***	−0.362***	−0.126***
				(0.015)	(0.024)	(0.040)	(0.029)
HC3				−0.118***	−0.427***	−0.968***	−0.143
				(0.023)	(0.084)	(0.205)	(0.089)
Since dry dock				0.001***	0.001***	0.002***	0.001***
				(0.0001)	(0.0002)	(0.0004)	(0.0003)
Constant	−0.424***	−0.255***	−0.752***	−0.883***	−1.044***	−1.037***	−0.026
	(0.069)	(0.065)	(0.063)	(0.064)	(0.107)	(0.169)	(0.138)
Observations	3,046	3,046	3,046	3,046	3,046	1,242	1,804
R^2	0.539	0.615	0.701	0.711	0.743	0.782	0.588
Adjusted R^2	0.539	0.614	0.699	0.709	0.736	0.767	0.568
Note:							*p<0.1; **p<0.05; ***p<0.01

Table 7.2: Regression for different sets of regressors with fixed effects (FE) for Suezmax vessels

	Dependent variable:						
	Log consumption						
	No FE	FE	Weather	HC	Crew	Ballast	Laden
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Speed	0.707***	0.704***	1.033***	1.132***	1.120***	1.267***	0.860***
	(0.036)	(0.035)	(0.035)	(0.032)	(0.032)	(0.058)	(0.035)
Draught	0.039***	0.040***	0.036***	0.036***	0.038***	−0.005	0.030***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.005)	(0.002)
HC1				−0.172***	−0.187***	−0.225***	−0.157***
				(0.010)	(0.012)	(0.026)	(0.012)
HC2				−0.303***	−0.334***	−0.317***	−0.284***
				(0.017)	(0.020)	(0.042)	(0.020)
HC3				−0.386***	−0.414***	−0.439***	−0.289***
				(0.024)	(0.029)	(0.055)	(0.033)
Since dry dock				0.004***	0.004***	0.003***	0.003***
				(0.0002)	(0.0002)	(0.0004)	(0.0002)
Constant	1.260***	1.288***	0.322***	−0.057	0.252	0.382*	0.857***
	(0.092)	(0.090)	(0.093)	(0.086)	(0.156)	(0.217)	(0.101)
Observations	2,717	2,717	2,717	2,717	2,717	904	1,813
R^2	0.383	0.417	0.526	0.613	0.659	0.629	0.576
Adjusted R^2	0.383	0.416	0.523	0.610	0.651	0.605	0.561
Note:							*p<0.1; **p<0.05; ***p<0.01

Table 7.3: Regression for different sets of regressors with fixed effects (FE) for Medium Range vessels

	Dependent variable:						
	Log consumption						
	No FE	FE	Weather	HC	Crew	Ballast	Laden
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Speed	0.963***	0.990***	1.392***	1.403***	1.340***	1.614***	0.804***
	(0.033)	(0.032)	(0.031)	(0.030)	(0.030)	(0.064)	(0.035)
Draught	0.082***	0.086***	0.076***	0.076***	0.077***	0.028**	0.051***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.011)	(0.002)
HC1				−0.093***	−0.104***	−0.116***	−0.099***
				(0.010)	(0.013)	(0.044)	(0.013)
HC2				−0.109***	−0.270***	−0.100	−0.250***
				(0.017)	(0.031)	(0.110)	(0.028)
HC3				−0.167***	−0.339***	−0.277**	−0.311***
				(0.022)	(0.041)	(0.131)	(0.041)
Since dry dock				0.001***	0.002***	0.001**	0.002***
				(0.0001)	(0.0002)	(0.001)	(0.0002)
Constant	−0.373***	−0.514***	−1.723***	−1.833***	−1.779***	−2.216***	0.007
	(0.084)	(0.081)	(0.083)	(0.084)	(0.093)	(0.237)	(0.108)
Observations	3,192	3,192	3,192	3,192	3,192	743	2,449
R^2	0.458	0.496	0.617	0.631	0.676	0.744	0.493
Adjusted R^2	0.458	0.495	0.615	0.629	0.666	0.711	0.473
Note:							*p<0.1; **p<0.05; ***p<0.01

In general, the size of the marginal effects[16] increases with each subsequent hull cleaning (Column 4), and especially in ballast condition (6). When pooled for both ballast and laden condition (4), the effect on Panamax and Medium Range vessels ranges from -8.8 % for the first hull cleaning to -15.3 % for the third hull cleaning. The effect is noticably higher for the Suezmax vessels ranging from -15.8 % to -32 %. The large effects are probably partly due to the substantially higher weekly increase in consumption, which is 0.4 % for Suezmax, compared to 0.1 % for both the Panamax and Medium Range segments. As the third hull cleaning is most likely performed during the latter part of the

[16]Marginal effects of logarithmic values are calculated in the same way as in section 4.1

three-year period, the effect of a higher weekly growth plays a bigger role in the marginal effect of later hull cleanings. The same logic is also applicable for Medium Range and Panamax segments in why later hull cleanings have a bigger marginal effect.

Adding the hull cleaning variable increases the adj. R^2 with 8.7 % for Suezmax, 1 % for Panamax and 1.4 % for Medium Range vessels, which is a major difference. Further, adding crew performance increases adj. R^2 with 4.1 % for Suezmax compared to 2.7 % for Panamax and 3.7 % for Medium Range. The difference between classes is considerably smaller for crew performance than for hull cleanings. The coefficients of masters and chief engineers are omitted from the table due to the high number of variables, while weather variables are omitted for visual purposes and because they have been previously discussed.

7.2 Variance decomposition

Variance decomposition is used to indicate the amount of information each variable contributes compared to the other variables. Similar to the regressions, such analyses are made for each class due to differences in vessel characteristics. This is to observe if hull cleaning or crew performance are prominent variables in explaining variance, and see to which extent they add extra information regarding fuel consumption. The results of ANOVA tests for each segment are displayed in tables 7.4 to 7.6. The percentage column is calculated as the individual sum of squares divided by the total sum of squares, excluding residuals. Thus, the column measures the share of explained variance that the respective variable captures, showing which variables provide the most in terms of explanatory power.

Table 7.4: ANOVA for Panamax vessels

	Df	Sum Sq	Mean Sq	Pr(>F)	Percentage
Log speed	1	89.15	89.15	0.0000	54.98
Draught	1	28.48	28.48	0.0000	17.56
Weather	9	19.28	2.14	0.0000	11.89
Vessel name	6	1.23	0.21	0.0000	0.76
Weeks since dry dock	1	5.12	5.12	0.0000	3.15
Hull cleaning	3	1.81	0.60	0.0000	1.11
Master	36	12.24	0.34	0.0000	7.55
Chief engineer	28	4.83	0.17	0.0000	2.98
Residuals	2960	55.96	0.02		

Table 7.5: ANOVA for Suezmax vessels

	Df	Sum Sq	Mean Sq	Pr(>F)	Percentage
Log speed	1	8.94	8.94	0.0000	10.41
Draught	1	41.04	41.04	0.0000	47.79
Weather	9	14.61	1.62	0.0000	17.01
Vessel name	3	1.66	0.55	0.0000	1.93
Weeks since dry dock	1	0.92	0.92	0.0000	1.07
Hull cleaning	3	3.59	1.20	0.0000	4.17
Master	23	6.94	0.30	0.0000	8.08
Chief engineer	21	8.17	0.39	0.0000	9.51
Residuals	2654	44.49	0.02		

Table 7.6: ANOVA for Medium Range vessels

	Df	Sum Sq	Mean Sq	Pr(>F)	Percentage
Log speed	1	33.09	33.09	0.0000	23.82
Draught	1	60.96	60.96	0.0000	43.88
Weather	9	25.17	2.80	0.0000	18.12
Vessel name	6	2.32	0.39	0.0000	1.67
Weeks since dry dock	1	6.02	6.02	0.0000	4.33
Hull cleaning	3	1.11	0.37	0.0000	0.79
Master	39	5.31	0.14	0.0000	3.82
Chief engineer	36	4.94	0.14	0.0000	3.55
Residuals	3095	66.45	0.02		

All variables have p-values less than 1^{-4}, which indicates that the inclusion of all variables is correct as they all provide additional information. In general, log speed, draught, and weather offer the highest explanatory power. The crew member variables actually have a bigger impact than fixed vessel effects, because the vessels are compared to almost identical sister ships. Since there are so many crew members, the possibility of overfitting is present, as the data is highly fragmented. Hence, it is possible that these variables capture a decent portion of unknown or arbitrary variation by chance. It is surprising that draught has a larger effect on variance than speed for two segments. However, this is likely because observations are pooled for both ballast and laden conditions. In appendix A6, ANOVA tests for vessels in ballast are displayed. The relationship between speed and consumption becomes more obvious when only considering ballast observations.

Since Panamax vessels are unique in terms of their dual-engines, speed has a greater impact on consumption, leading to competent crew playing an important role for vessel performance. This could explain why masters have a substantial effect on Panamax vessels.

The cause of the large effect of crew performance on Suezmax vessels remains unclear, with chief engineers being prominent in terms of variance. In section 6.2, most of the chief engineers with significant differences were serving on Suezmax vessels, and the number of chief engineers was even higher using a 5 % significance level. With this in mind, the results are not suprising as it is obvious that chief engineers have large variations on Suezmax vessels. One theory could be that the efficiency of operations cause larger fluctuations in performance for Suezmax vessels. Since chief engineers are responsible for submitting the noon reports, it is also possible that Suezmax vessels for some reason suffer more in terms of having correct, standardized noon reports, although this is hard to ascertain.

Suezmax vessels are also the largest in terms of deadweight tonnage, which means that draught varies to a greater extent. This is supported by the descriptive statistics in section 3.3.2, where the Suezmax segment has max observations of 17m in draught, compared to 12.7m and 12.6m for Panamax and Medium Range respectively. Therefore, it is logical that draught has a bigger impact in describing variance for the Suezmax class. The differences become minimal in ballast condition, where most of the variance is attributed to speed, weather and crew performance. Additionally, the draught of the Suezmax class may be a factor to why hull cleaning has a larger impact on these vessels, as more of the hull is exposed to frictional resistance compared to other classes. Medium Range vessels do a lot of part cargos and parcelling leading to less idle/waiting time, thus reducing the risk of biofouling. This could explain why hull cleaning makes up such a small portion of the variance for Medium Range vessels. The conclusions drawn from the results are based on knowledge provided by the ship owner.

7.3 Prediction accuracy

To see if adding hull cleaning and crew performance influence the accuracy of predicting fuel consumption, a selection of machine learning algorithms from section 4.2 are revisited. This includes three linear models (OLS, Lasso and Ridge) and three non-linear tree-based models (Extra Trees, Random Forest and xgBoost). Since some observations and vessels were removed in section 4, predictions of the trimmed data set are performed again to create fair comparisons. Again, predictions are divided into vessel segments because of the large differences in characteristics. The algorithms use the same validation set approach as before, splitting 80 % of observations into training data for model fitting and validation, and the remaining 20 % into a hold-out test set to generate unbiased results. All metrics are displayed from out-of-sample data. With the addition of the discussed variables, the effects on prediction accuracy are displayed in tables 7.7 to 7.9.

Table 7.7: Performance metrics for Panamax vessels

	Without HC and Crew			*All predictors*		
Panamax	RMSE	R2	MAE	RMSE	R2	MAE
OLS	0.1565	64.50%	0.1184	0.1501	67.49%	0.1118
Lasso	0.1564	64.50%	0.1183	0.1501	67.47%	0.1120
Ridge	0.1556	64.50%	0.1179	0.1501	66.98%	0.1127
Extra Trees	0.1365	72.84%	0.0903	0.1289	75.77%	0.0860
Random Forest	0.1315	74.62%	0.0897	0.1272	76.29%	0.0864
xGBoost	0.1334	74.53%	0.0913	0.1261	77.13%	0.0864

Table 7.8: Performance metrics for Suezmax vessels

	Without HC and Crew			*All predictors*		
Suezmax	RMSE	R2	MAE	RMSE	R2	MAE
OLS	0.1379	55.67%	0.1020	0.1248	63.75%	0.0928
Lasso	0.1379	55.64%	0.1020	0.1248	63.68%	0.0928
Ridge	0.1382	55.20%	0.1026	0.1292	60.86%	0.0962
Extra Trees	0.1065	73.60%	0.0738	0.0986	77.52%	0.0674
Random Forest	0.1036	75.64%	0.0721	0.0986	78.22%	0.0677
xGBoost	0.1019	75.71%	0.0707	0.1005	76.40%	0.0697

Table 7.9: Performance metrics for Medium Range vessels

	Without HC and Crew			*All predictors*		
Medium Range	RMSE	R2	MAE	RMSE	R2	MAE
OLS	0.1611	59.72%	0.1211	0.1535	63.57%	0.1145
Lasso	0.1610	59.76%	0.1209	0.1534	63.61%	0.1142
Ridge	0.1609	59.82%	0.1203	0.1546	62.87%	0.1154
Extra Trees	0.1299	73.85%	0.0874	0.1201	77.62%	0.0804
Random Forest	0.1272	75.10%	0.0869	0.1205	78.10%	0.0813
xGBoost	0.1306	74.85%	0.0872	0.1291	75.28%	0.0851

The results in tables 7.7 to 7.9 solidifies the evidence that adding hull cleaning and crew performance have a distinctly bigger impact on Suezmax tankers, as the R^2 of the OLS model increases by almost 8 %, compared to an increase in R^2 of 3 % and 4 % for Panamax and Medium Range respectively. For non-linear models, the improvement is more modest, with R^2 increasing by 2.5 %, 2.5 %, and 3 % in the best performing model for Panamax, Suezmax, and Medium Range vessels respectively. Analogous to section 4.2, more advanced algorithms perform better. Overall, Random Forest seems to be the best performing machine learning technique in terms of prediction accuracy.

The difference in performance between linear and non-linear models has however decreased markedly. Because of the low variance of linear models, the inclusion of more predictors leads to bigger improvements in accuracy relative to the high variance non-linear models. This is perhaps due to more advanced models having a higher tendency to overfit, as they use the available predictors to explain unknown variation to a larger extent. The inclusion of hull cleaning and crew performance as variables thus seem to explain a large portion of the differences in performance between OLS and other non-linear models. Of course, the possibility of overfitting remains potent for all methods due to the high fragmentation that occurs when including crew variables.

8 Limitations and further reasearch

The results regarding the impact of hull cleaning and the optimization of these, is based on results from the classification model in 4. All results derived from the ensuing analysis is consequently dependent on the accuracy of this model. If the detected hull cleanings stray far away from reality, the findings and interpretations could be inaccurate. However, as the comparison of fuel consumption profiles in section 4.5 show, the detected hull cleaning dates are more in line with expectations than the company-reported dates. The analysis is therefore assumed to be relatively trustworthy given the uncertainty in the data. This emphasizes the value of precise and reliable data.

It is hard to quantify the accuracy of the classification model, due to the lack of a training set with confirmed hull cleaning dates. Further research should aim to test the classification model in terms of prediction error. Although few classification models are designed to consider time series when recognizing events, it could be worth testing various models on training and test data to compare prediction errors. One method that seems to be a good fit for this type of time series classification problem, is a Recurrent Neural Network (RNN). RNNs take into account the order of observations in a time series and as such could perhaps recognize patterns preceding and following actual hull cleanings (James et al., 2013). Given a training set with sufficient observations and confirmed hull cleaning dates, further research should attempt implementing RNNs to identify hull cleaning dates. The authors believe this could be among the best prediction models for this classification problem.

The noon reports provided only cover three years, where most of the dry docks happened 1-2 years prior to the first observations. This leads to uncertainty in whether the first hull cleaning identified is actually the first hull cleaning of the dry dock interval. It is reasonable to assume that given the observed hull cleaning intervals, some hull cleanings have been performed in the period prior to the first observations. Hence, some caution is advised when quantifying and comparing the exact effect of the first, second and third hull cleaning, as the sequence number may be incorrect.

As some papers have previously discussed, including Idais et al. (2021), measuring the exact growth rate of biofouling is extremely complex. Further research could include additional variables that are sensitive in terms of being accurately quantified, such as water temperature, salinity, exact idle time, and time spent in areas with increased risk of biofouling. Given accurate measurements, the impact of external variables on friction could be correctly calculated, and enable more accurate predictions of bunker consumption.

The development of biofouling is not only dependent on exogenous variables affecting the vessel. The type of hull and choice of antifouling coating used also affects the rate of biofouling. This information is rarely available, but could have improved estimates of hull deterioration over time, and helped determine hull cleaning dates more precisely. The thickness of the antifouling coating applied during dry dock is also important, as it affects how long the coating remains effective. This is especially important for determining to what extent the coating is scrubbed off during underwater hull cleanings.

Although the optimizations in section 5 are accurate given the assumptions, the real-life conditions make these assumptions quite a stretch. Because of the several unknown or currently unattainable factors affecting biofouling, the optimization model will only go so far in regards to application. Further, several of the assumptions such as the size of hull cleaning shocks are unlikely to hold in real life, and would require extensive analysis in order to estimate accurately. Additionally, because some of the constraints in the theoretical optimization model is convex quadratic, linear solvers are not able to solve the optimization problem. Non-linear solvers were also attempted, but failed to deliver the desired results due to ignoring integrality constraints. The model was therefore implemented in R, and run as a brute force model in which the intervals that minimized consumption were selected. This leads to high computational times, which increases exponentially whenever parameters such as the minimum time interval m is reduced, or the number of hull cleanings h and time period n is increased.

To achieve an optimization model which delivers real-life application to a higher degree, another approach is likely required. As continuous monitoring of hull conditions becomes more widespread, the decision rules are to a larger extent based on real time features and thresholds. Further research could utilize simulations to find typical shipping routes, measure factors that are expected to affect biofouling, and find the optimal thresholds where the vessel is most likely to minimize consumption by performing a hull cleaning. This kind of model would probably move away from the realm of economical optimization, and instead be based on simulations and machine learning predictions. For instance, Bomholt and Thune (2020) used Extreme Gradient Boosting and RNNs with great results for shipping route simulations.

When back-transforming logarithmic variables, the estimates sometimes have a tendency to be biased. The regressions in this thesis are primarily estimated in log-log space to remove non-linear relationships between variables. However, the hull cleaning optimizations use actual fuel consumption for better interpretability and applicability. Hence, all estimated values in section 5 are back-transformed. Despite the arithmetic mean not necessarily deviating by much, this bias is important to keep in mind when back-transforming logarithmic values (Rothery, 1988).

There are several reasons to consider the results of the individual crew analysis under a high degree of caution. The models used to estimate the consumption net all other variables have limited explanatory power, as there are several known or unknown factors not available that could further explain consumption. If correlation between a specific master and high consumption exists, we cannot know whether this is caused by the master or by any number of external unknown factors, for instance mechanical issues. Further, it is likely a stretch to hold just one master and - to a lesser degree - one chief engineer, responsible for everything related to consumption. For instance, there are several other crew members in charge at various times during the 24 hours of a noon report.

The decision to not log transform variables before the statistical testing was made to increase inference validity, as tests with log-transformed variables are no longer operating with arithmetic means. Instead, the arithmetic mean of the log transformed consumption is actually the geometric mean of the non-transformed consumption. Although the t-test itself tests for differences in the provided mean with no differentiating, the results do not automatically translate to outside log space. It follows that using the log transformed consumption in testing would lead to different p-values, and therefore different conclusions. However, the estimation of external effects are likely worse when the power relationship between consumption and speed is not modeled as such. Hence there is a trade-off between a slightly worse fit for the model and a slightly more "correct" inference, although the conclusions remain closely related. In all cases, the results should not be taken for granted, as the causality of deviations from the mean consumption remains uncertain.

9 Conclusion

This thesis uses noon reports to identify the impact of hull cleaning and crew performance on bunker consumption under data uncertainty. To account for the uncertainty, a classification model to detect hull cleanings based on noon reports was proposed. Previous research suggests that hull cleanings lead to negative shocks in fuel consumption, with consumption between hull cleanings increasing over time mainly due the development of biofouling. While the hull cleaning dates provided by the data owner rendered largely nonsensical fuel consumption profiles, the hull cleaning dates detected by the classification model generated fuel consumption profiles which were substantially better in terms of fitting expectations. Although it is not possible to quantify the accuracy of the classification model without additional information, further research could test model accuracy and fine-tune model parameters on data sets with confirmed hull cleaning dates.

Economical optimizations of hull cleaning intervals were done by defining an optimization problem. Optimal hull cleaning intervals were discovered to be independent on both initial fuel consumption and its consumption increase. Thus, in theory, all that is needed to optimize hull cleaning intervals using the proposed model, are accurate estimates of the shock size relative to post dry dock levels. The optimization model was applied on actual vessels with two classified hull cleanings under strict assumptions. Using the best available estimate for the increase in frictional resistance over time, fuel savings of 0.3 % and 1.4 % were achieved over a three-year period by moving the classified hull cleanings to their respective optima. Adding an additional hull cleaning and moving the hull cleanings to their respective optima resulted in fuel savings of 2.1 % and 3.2 %. The magnitude of the estimated fuel savings are mainly dependent three factors; (i) the increase in consumption due to biofouling, (ii) the size of the hull cleaning shock, and (iii) by how much the vessels missed their respective optima.

The effects of individual crew members were explored using statistical tests. Results indicate that some masters consume significantly more or less fuel compared to their peers. Masters with significantly lower consumption incur fuel savings between 0.22 - 0.38 tons/day, while masters with significantly higher consumption have excess fuel usage of 0.37 - 0.59 tons/day. The tests further identifies which chief engineers that have significantly different consumption compared to their respective master's average. Though some chief engineers significantly differ from their colleagues, they do not appear to be consistent in whether they are better or worse than their colleagues. Despite significant differences in fuel consumption, there are no assumptions of causality. Although deviations could reflect the crew member's competence, they could also be due to unknown factors.

Identified hull cleanings and crew performance were included as variables to quantify their impact on consumption. Variance decompositions showed that draught and speed are the most important drivers of fuel consumption, explaining between 10-55 % and 17-48 % of variance respectively, depending on vessel segment. Identified hull cleanings exhibits an explanatory power of around 9 % for Suezmax vessels and 1 % for other classes. Considering that effects are based on classified hull cleanings, the results are dependent on classification accuracy. Crew performance explains between 2.7-4.1 % of bunker consumption, depending on vessel segment. In terms of prediction, OLS regression achieved explanatory powers between 63.5-67.5 %, up from 55.5-64.5 %. The best performing non-linear tree-based models reached explanatory powers of 77.1-78.2 %, up from 74.5-75.6 %. This indicates that more advanced non-linear models capture unknown variance even without the additional predictors, due to their high variance. In contrast, the linear models with lower variance benefit greatly from the added predictors, as they are less prone to overfitting.

Findings show that hull cleaning has a significant impact on bunker consumption, although the size of the effect vary by vessel segment. Optimization of hull cleaning intervals can lead to moderate fuel savings. Since the analyses involving hull cleanings are performed under data uncertainty, some caution is advised until the accuracy of the classification model is assessed. The results underline the value of having reliable data as a foundation for consumption modelling and decision-making to increase fuel efficiency. Further analysis reveals that crew performance has a noticeable impact on consumption, although the causality remains unclear. For shipping to reach the IMO emission goals, both the impact of hull cleaning and crew performance should be considered by ship owners. With large variation between individual crew members, and potential savings in the optimization of hull cleaning intervals, shipping companies can accommodate these findings in their daily operations to reduce fuel costs and emissions.

References

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. (2016). The influence of charterers and owners on bulk shipping freight rates. *Transportation Research Part E: Logistics and Transportation Review*, 86:69–82.

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.

Adland, R., Carriou, P., Wolff, F.-C., and Jia, H. (2017). Indirect ship hull condition monitoring using speed and fuel consumption analysis. Unpublished.

Benoit, K. (2011). Linear regression models with logarithmic transformations. *London School of Economics, London*, 22(1):23–36.

Bomholt, H. J. and Thune, T. S. (2020). Machine learning as a decision support system in capesize route optimization: predicting optimal route selection using recurrent neural networks and extreme gradient boosting. Master's thesis, NHH.

Bressy, C. and Lejars, M. (2014). Marine fouling : An overview. *Journal of Ocean Technology*, 9:19–28.

Copernicus (2021). Reducing ship's fuel consumption thanks to better marine coatings. Retrieved from: https://marine.copernicus.eu/services/use-cases/reducing-ships-fuel-consumption-thanks-better-marine-coatings.

CPH (2013). Difference-in-difference estimation. Retrieved from: https://www.publichealth.columbia.edu/research/population-health-methods/difference-difference-estimation.

Dafforn, K. A., Lewis, J. A., and Johnston, E. L. (2011). Antifouling strategies: history and regulation, ecological impacts and mitigation. *Marine Pollution Bulletin*, 62(3):453–465.

Dahly, D. L. (2017). Log log regression. Retrieved from: https://darrendahly.github.io/post/loglog/.

Farkas, A., Deguili, N., and Martić, Ivana, A. I. (2021). Energy savings potential of hull cleaning in a shipping industry. Unpublished.

Fortmann-Roe, S. (2012). Bias and variance. Retrieved from: http://scott.fortmann-roe.com/docs/BiasVariance.html.

GCaptain (2012). Technology solutions that keep your hull clean, and save you money. Retrieved from: https://gcaptain.com/technology-solutions-hull-clean/.

Gundermann, D. and Dirksen, T. (2016). A statistical study of propulsion performance of ships and the effect of dry dockings, hull cleanings and propeller polishes on performance. In *Proceedings of the 1st hull performance and insight conference (HullPIC'16), Castello di Pavone, Italy*, pages 282–291.

Hakim, M., Utama, I., Nugroho, B., Yusim, A., Baithal, M., and Suastika, I. (2017).

Review of correlation between marine fouling and fuel consumption on a ship. In *Proceeding of SENTA: 17th Conference on Marine Technology*, pages 122–129.

Hanck, C., Arnold, M., Gerber, A., and Schmelzer, M. (2019). Introduction to econometrics with r. *University of Duisburg-Essen.*

Hellio, C. and Yebra, D. (2009). *Advances in marine antifouling coatings and technologies.* Elsevier.

Hochkirch, K. and Bertram, V. (2010). Options for fuel saving for ships.

Idais, R. H., Abuhabib, A. A., and Hamzah, S. (2021). Recent advances in measuring and controlling biofouling of seawater reverse osmosis swro: A review. *Osmotically Driven Membrane Processes.*

IMO (2009). Second IMO GHG Study 2009. Retrieved from: https://wwwcdn.imo.org/localresources/en/OurWork/Environment/Documents/SecondIMOGHGStudy2009.pdf.

IMO (2011). A transparent and reliable hull and propeller performance standard. Retrieved from: https://jotunimages.azureedge.net/images/images/mepc-63-4-8_tcm214-10265.pdf.

IMO (2019). Initial IMO GHG strategy. Retrieved from: https://www.imo.org/en/MediaCentre/HotTopics/Pages/Reducing-greenhouse-gas-emissions-from-ships.aspx.

IMO (2020). Fourth IMO GHG Study 2020. Retrieved from: https://wwwcdn.imo.org/localresources/en/OurWork/Environment/Documents/Fourth%20IMO%20GHG%20Study%202020%20-%20Full%20report%20and%20annexes.pdf.

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

Kobayashi, E.-i., Yoneda, S., and Morita, A. (2014). Advanced route optimization in ship navigation. In *2014 4th International Conference On Simulation And Modeling Methodologies, Technologies And Applications (SIMULTECH)*, pages 572–577. IEEE.

Kojima, R., Kobayashi, S., Satuito, C. G. P., Katsuyama, I., Ando, H., Seki, Y., and Senda, T. (2016). A method for evaluating the efficacy of antifouling paints using mytilus galloprovincialis in the laboratory in a flow-through system. *Plos one*, 11(12):e0168172.

Kontovas, C. and Psaraftis, H. N. (2011). Reduction of emissions along the maritime intermodal container chain: operational models and policies. *Maritime Policy & Management*, 38(4):451–469.

Kwak, S. G. and Kim, J. H. (2017). Central limit theorem: the cornerstone of modern statistics. *Korean journal of anesthesiology*, (2):144.

Lande, P. B. (2017). Improving hull performance through iot. Retrieved from: https://www.wilhelmsen.com/media-news-and-events/industry-perspectives/2017/iot-hull-performance/.

Lindholdt, A., Dam-Johansen, K., Olsen, S., Yebra, D. M., and Kiil, S. (2015). Effects of biofouling development on drag forces of hull coatings for ocean-going ships: a review. *Journal of Coatings Technology and Research*, 12(3):415–444.

McDonald, J. (2014). Handbook of biological statistics. Retrieved from: http://www.biostathandbook.com/onewayanova.html", journal=One-way anova - Handbook of Biological Statistics.

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.

Moore, G. E. (1998). Cramming more components onto integrated ircuits. *Proceedings of the IEEE*, 86(1):82–85.

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

Rakke, S. G. (2016). Ship emissions calculation from ais. Master's thesis, NTNU.

RMS (2018). The beaufort scale: How is wind speed measured? Retrieved from: https://www.rmets.org/resource/beaufort-scale.

Rothery, P. (1988). A cautionary note on data transformation: bias in back-transformed means. *Bird Study*, 35(3):219–221.

Schultz, M. P. (2007). Effects of coating roughness and biofouling on ship resistance and powering. *Biofouling*, 23(5):331–341.

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

Stopford, M. (2008). *Maritime economics 3e*. Routledge.

StormGeo (2021). Vessel efficiency and climate change. Retrieved from: https://www.stormgeo.com/products/s-suite/s-insight/articles/vessel-efficiency-and-climate-change/.

Tripathi, R. (2016). Advances in antifouling coatings technology. Retrieved from: https://www.coatingsworld.com/issues/2016-10-01/view_features/advances-in-antifouling-coatings-technology/.

UN (2016). Maritime transport is 'backbone of global trade and the global economy', says secretary-general in message for international day | meetings coverage and press releases. Retrieved from: https://www.un.org/press/en/2016/sgsm18129.doc.htm.

UNCTAD (2018). Review of Maritime Transport 2018. Retrieved from: https://unctad.org/webflyer/review-maritime-transport-2018.

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

VPO (2018). The importance of crew in energy efficient shipping. Retrieved from: https://vpoglobal.com/2018/05/11/the-importance-of-crew-in-energy-efficient-shipping/.

Wang, S. and Meng, Q. (2012). Sailing speed optimization for container ships in a liner shipping network. *Transportation Research Part E: Logistics and Transportation Review*, 48(3):701–714.

Wankhede, A. (2021). What is noon report on ships and how is it prepared? Retrieved from: https://www.marineinsight.com/guidelines/what-is-noon-report-on-ships/.

Yang, J. W., Park, H., Chun, H. H., Ceccio, S. L., Perlin, M., and Lee, I. (2014). Development and performance at high reynolds number of a skin-friction reducing marine paint using polymer additives. *Ocean Engineering*, 84:183–193.

Yu, J.-W., Lee, C.-M., Seo, J.-H., Chun, H. H., Choi, J.-E., and Lee, I. (2021). Comparative study on the prediction of speed-power-rpm of the kvlcc2 in regular head waves using model tests. *International Journal of Naval Architecture and Ocean Engineering*, 13:24–34.

Appendix

A1 Geographic locations

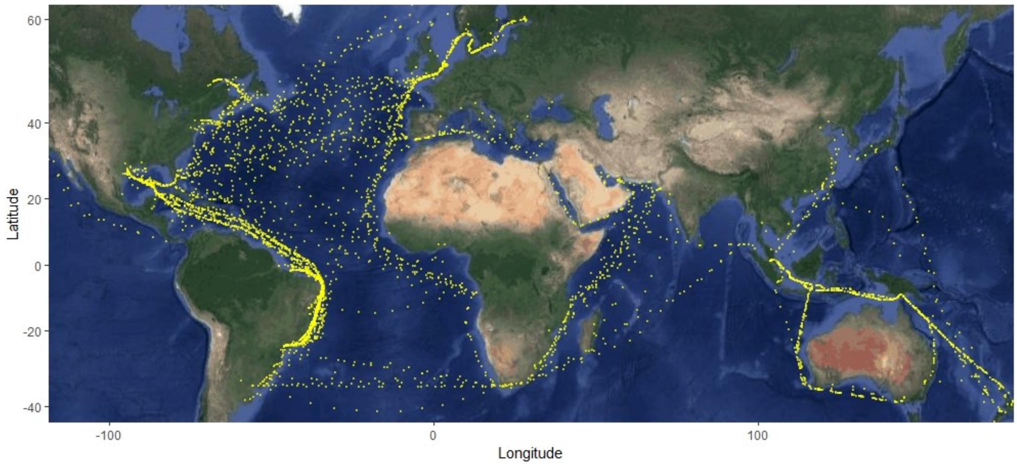


Figure A1.1: Trading routes for all Panamax vessels

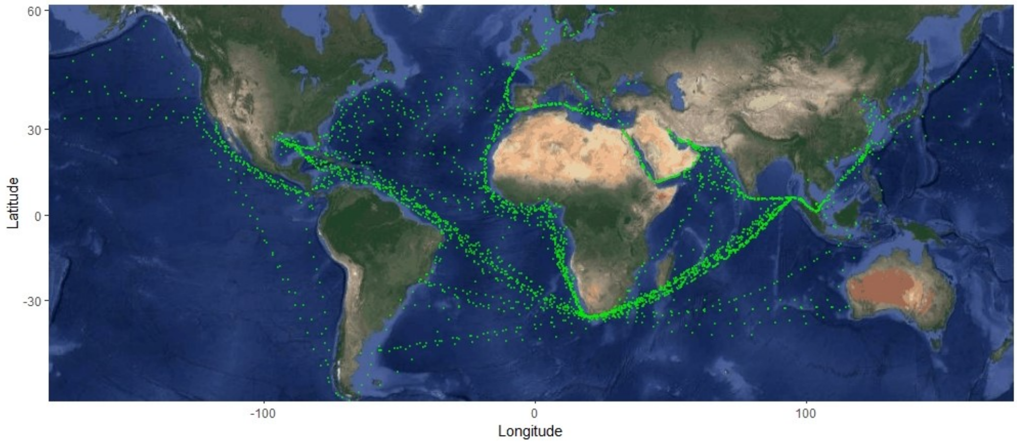


Figure A1.2: Trading routes for all Suezmax vessels

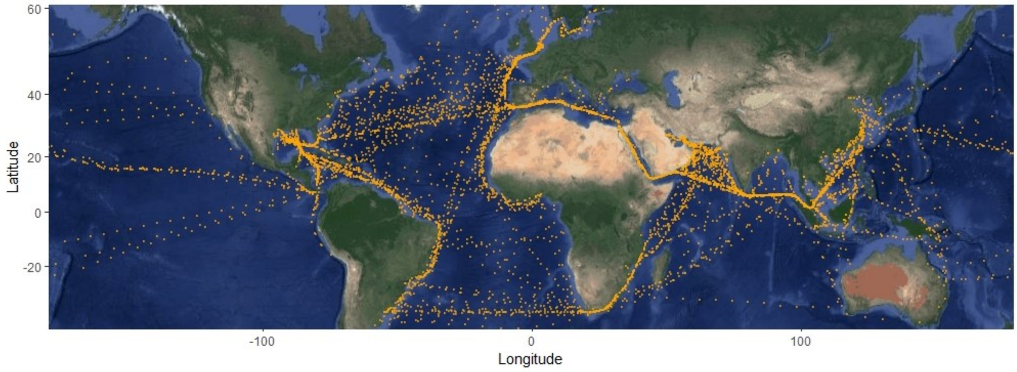
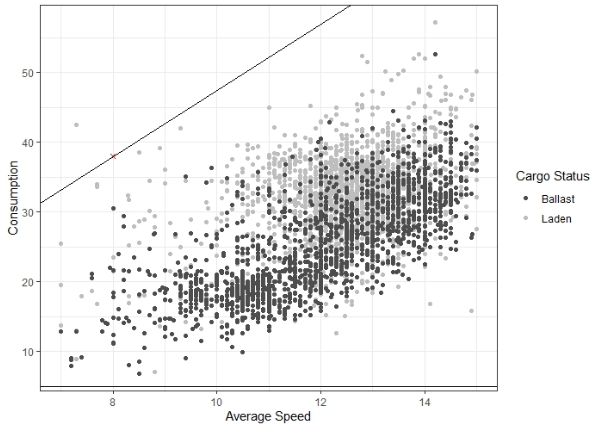
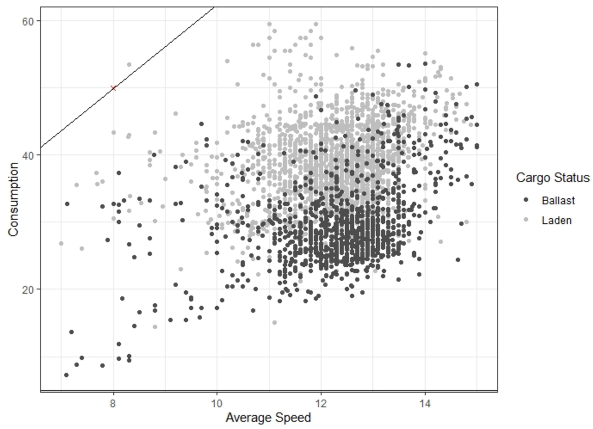


Figure A1.3: Trading routes for all Medium Range vessels

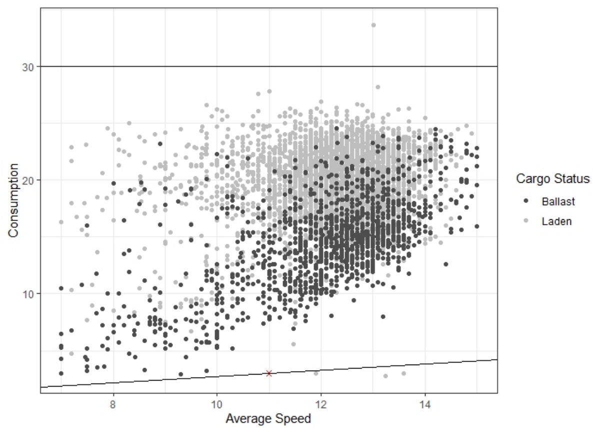
A2 Outliers in different vessels segments



(a) Panamax outliers



(b) Suezmax outliers

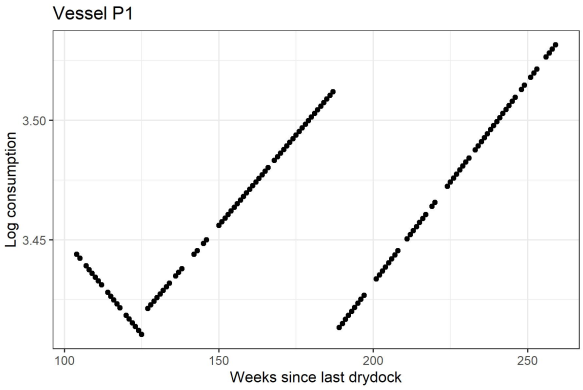


(c) Medium Range outliers

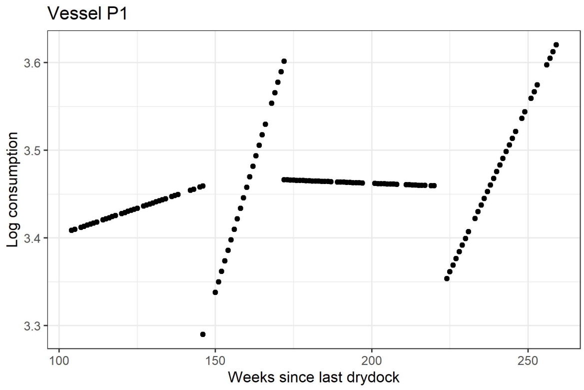
Figure A2.1: Cut-off lines for outliers for all segments

A3 Fuel consumption profiles

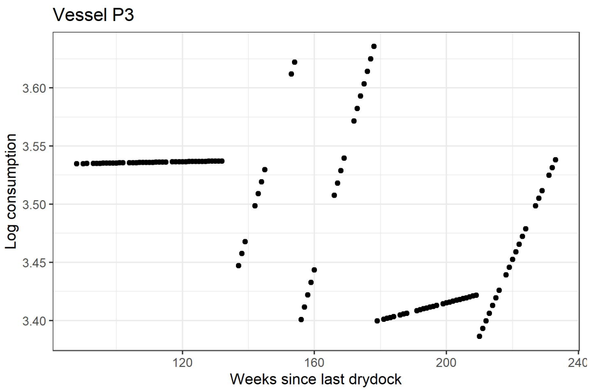
A3.1 Comparison of fuel consumption profiles



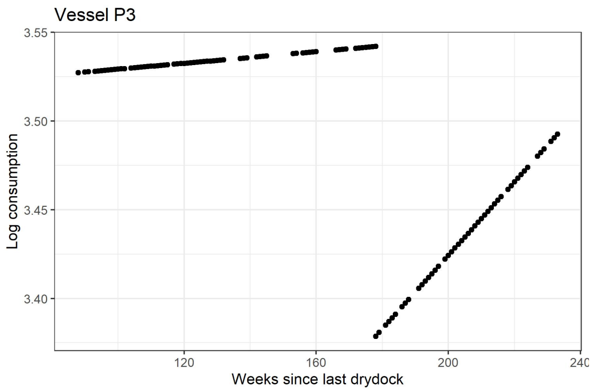
(a) Reported HC P1



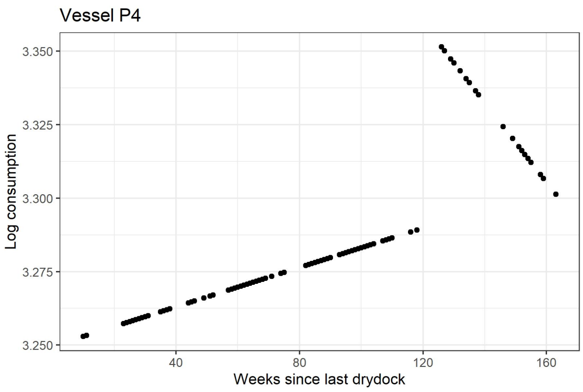
(b) Detected HC P1



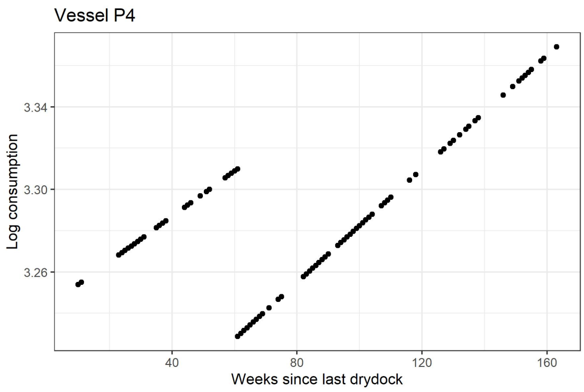
(c) Reported HC P3



(d) Detected HC P3

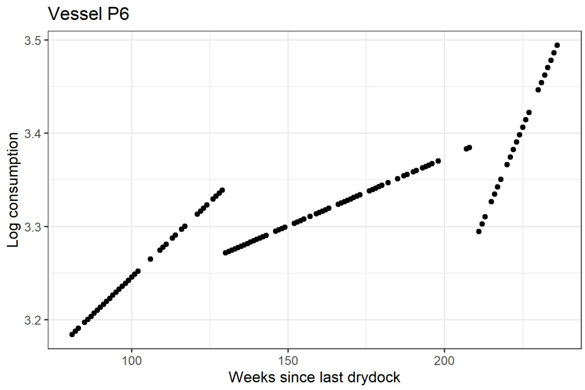


(e) Reported HC P4

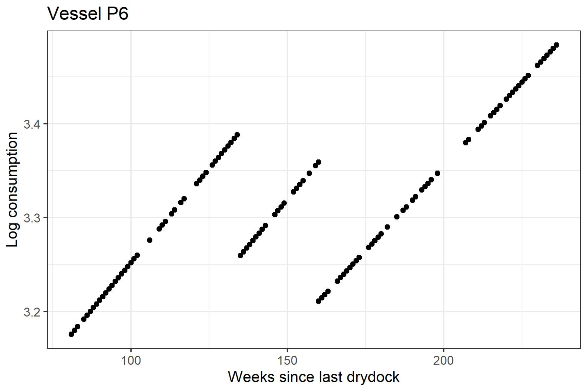


(f) Detected HC P4

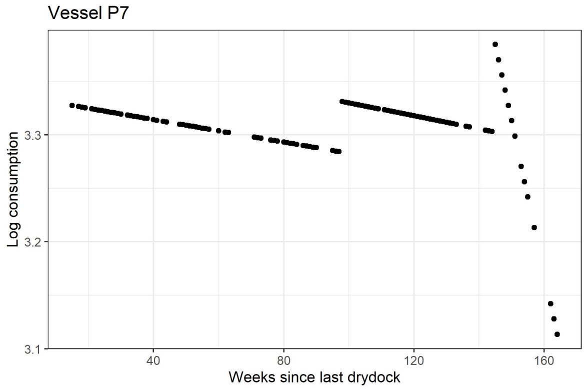
Figure A3.1: Comparison of fuel consumption profiles



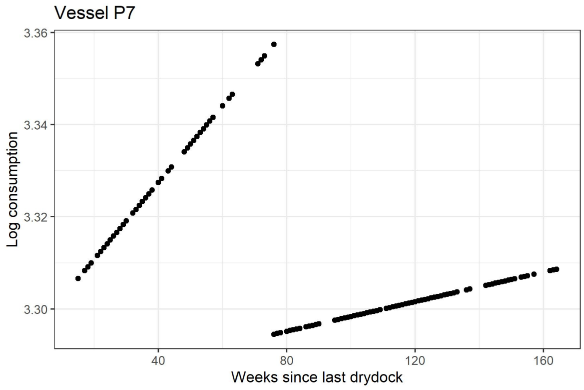
(g) Reported HC P6



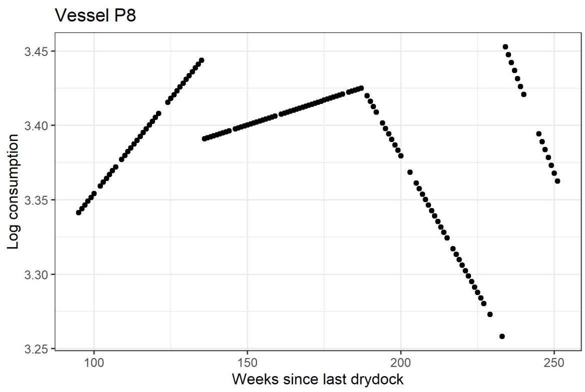
(h) Detected HC P6



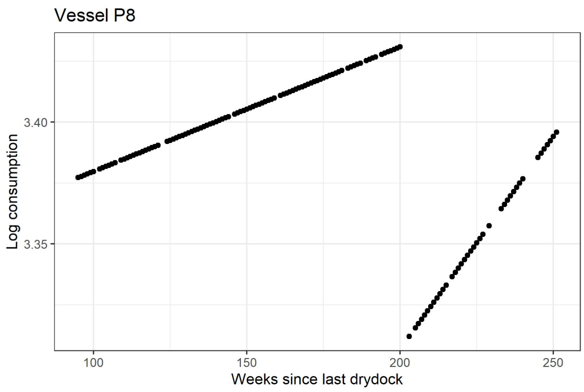
(i) Reported HC P7



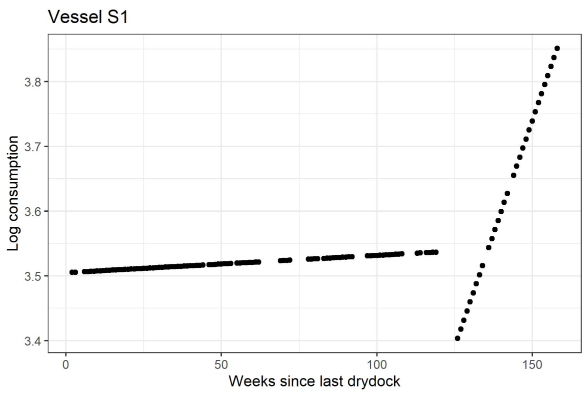
(j) Detected HC P7



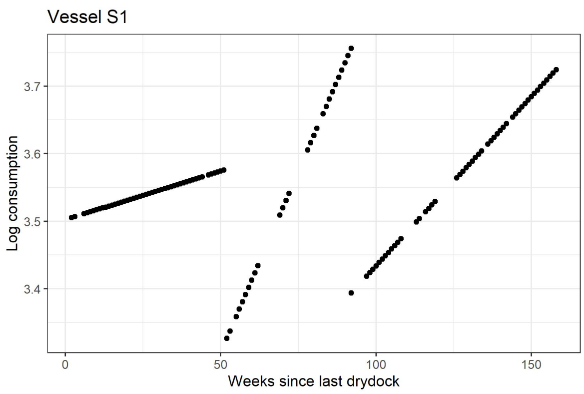
(k) Reported HC P8



(l) Detected HC P8

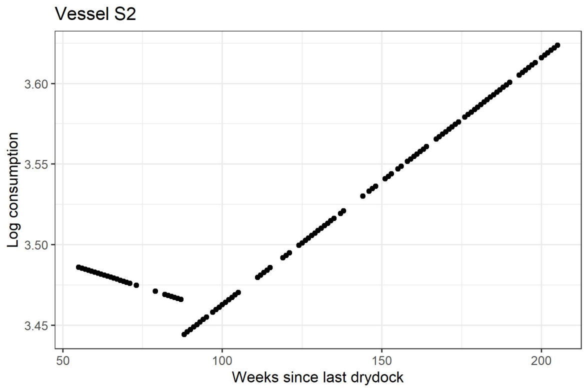


(m) Reported HC S1

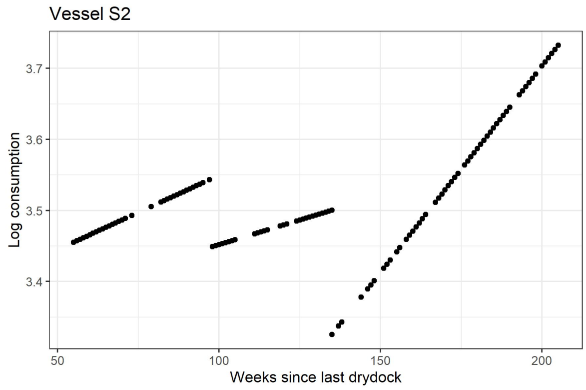


(n) Detected HC S1

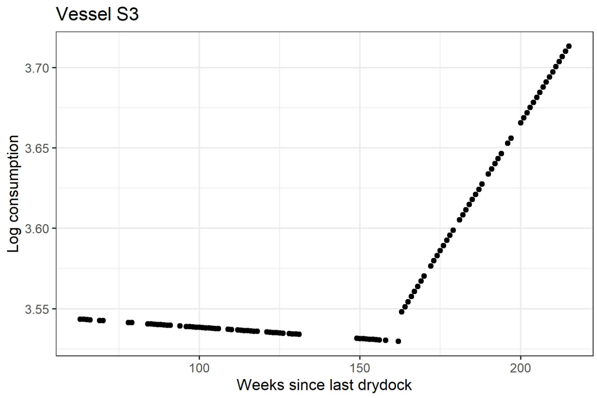
Figure A3.1: Comparison of fuel consumption profiles



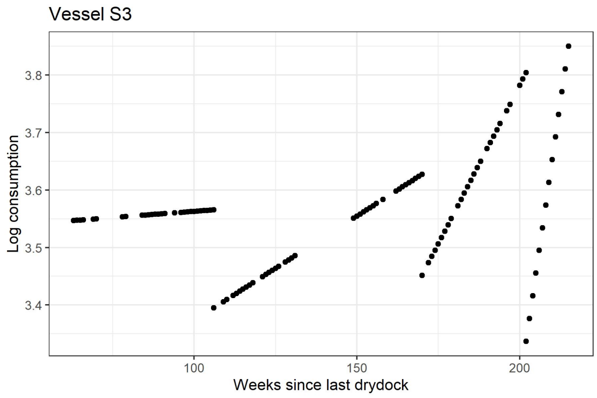
(o) Reported HC S2



(p) Detected HC S2



(q) Reported HC S3



(r) Detected HC S3

Figure A3.1: Comparison of fuel consumption profiles

A3.2 Classified hull cleaning dates for vessels with no basis of comparison

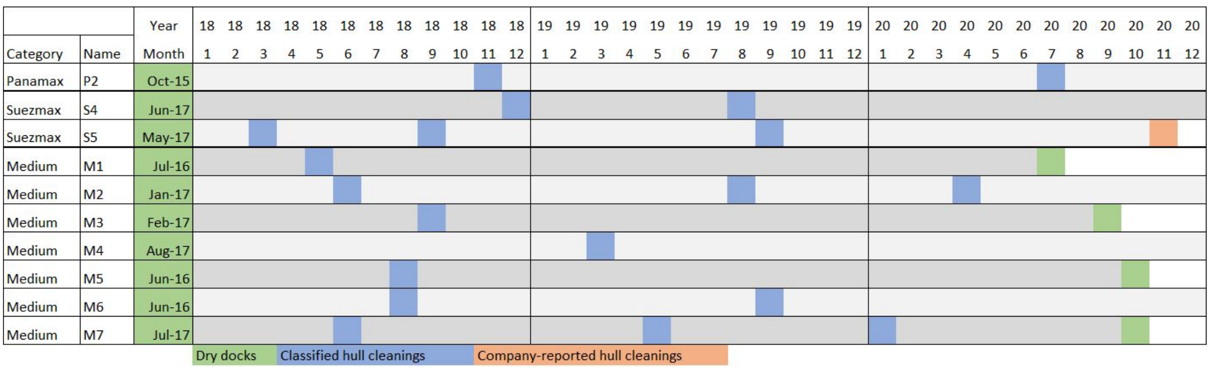
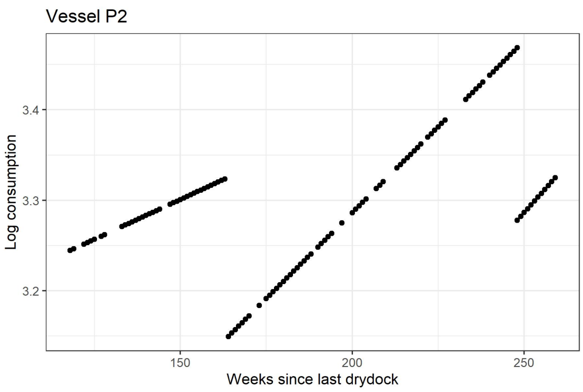
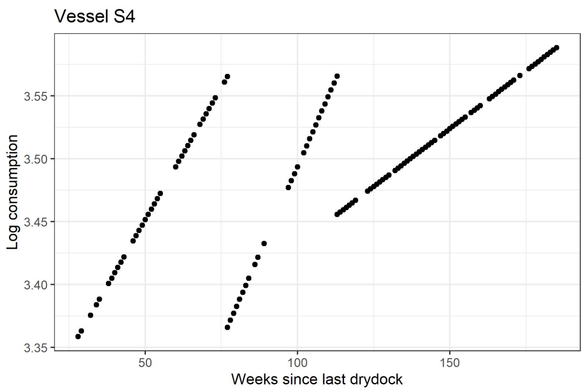


Figure A3.2: Classified hull cleaning dates for vessels with no basis of comparison

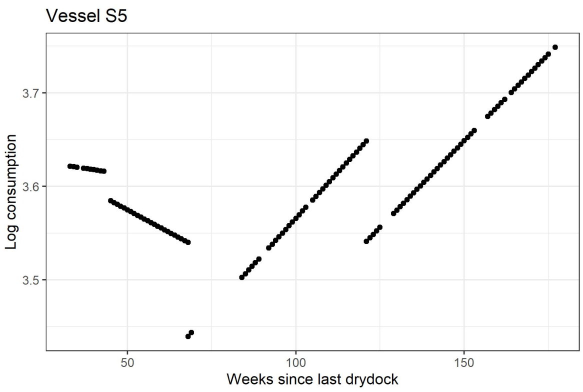
A3.3 Fuel consumption profiles for vessles with no basis of comparison



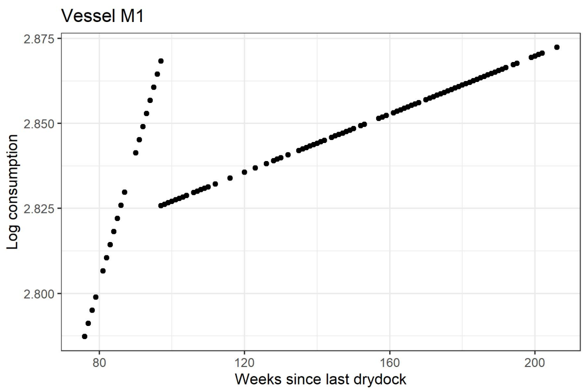
(a) Detected HC P2



(b) Detected HC S4

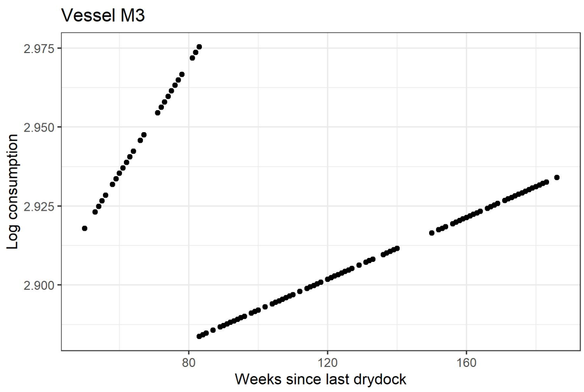


(c) Detected HC S5



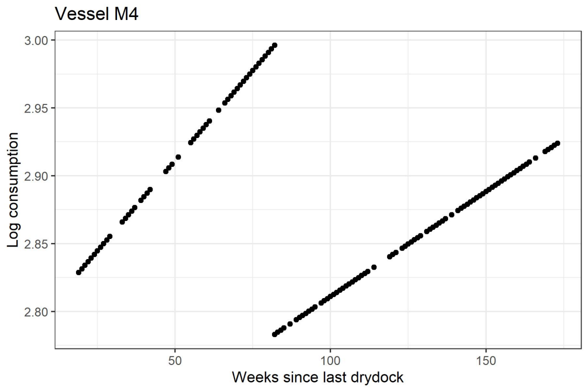
(d) Detected HC M1

(e) Detected HC M2

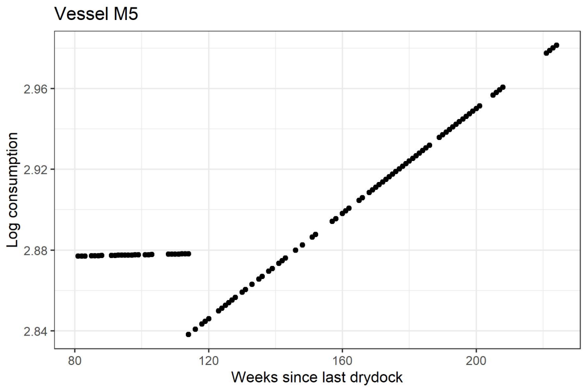


(f) Detected HC M3

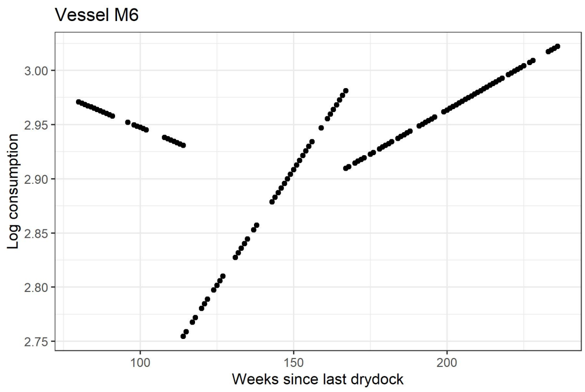
Figure A3.3: Classified fuel consumption profiles



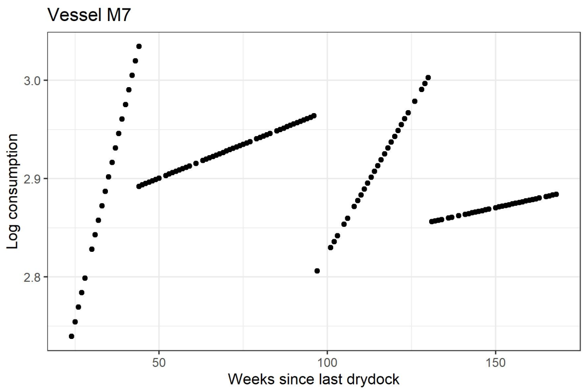
(g) Detected HC M4



(h) Detected HC M5



(i) Detected HC M6



(j) Detected HC M7

Figure A3.3: Classified fuel consumption profiles

A4 Sensitivity of parameters for optimization

A4.1 Interval sensitivity to shock size

Figure A4.1 shows the sensitivity of optimal timings subject to changes in shock size δ, with $0.3 < \delta < 0.9$.

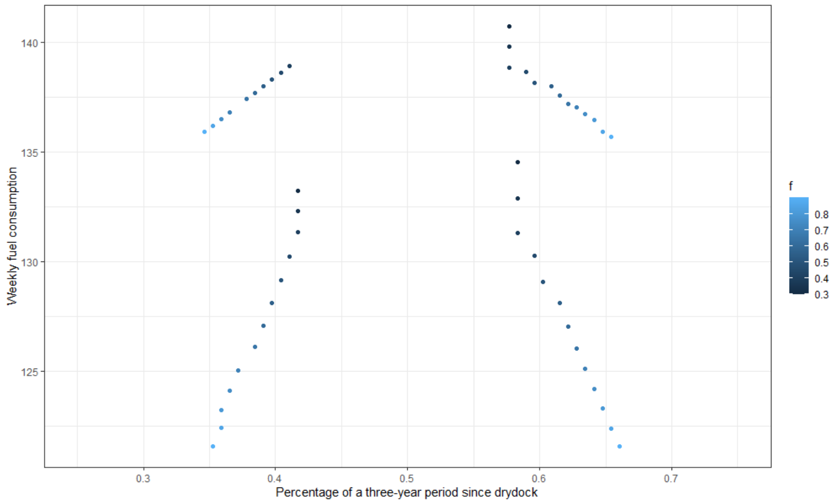


Figure A4.1: Optimal intervals for various shock sizes

The points correspond to the optimal timings given various shock sizes. An easier interpretation is shown in figure A4.2, where the example consumption profile has $\delta = 0.9$. Hull cleaning timings move closer to the middle when shock size decreases, until the minimum interval constraint m is binding. The shape of the graph in figure A4.1 is completely independent of the initial consumption C_0 and slope number l.

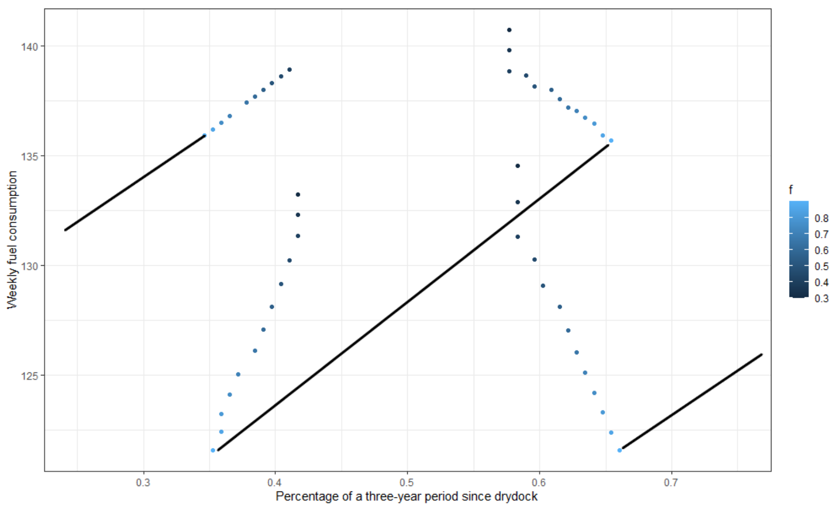


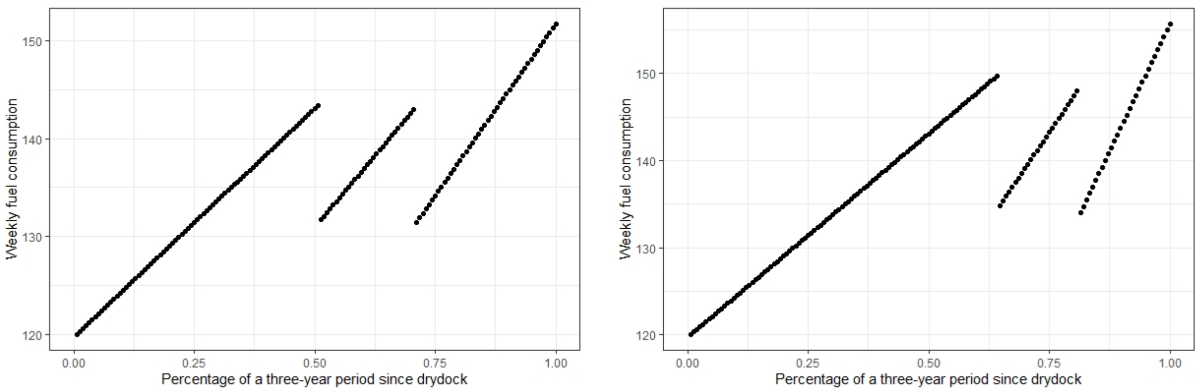
Figure A4.2: Optimal intervals for various shock sizes illustrated with $\delta = 0.9$

A4.2 Interval sensitivity with increasing slope numbers

Let l_k be the consumption increase after hull cleaning k, with l_0 being the initial consumption increase. Slope number l_k increases after each hull cleaning with the absolute value Δ, given by $\Delta = l_0 \cdot g$, where the growth factor g is a percentage of the initial consumption increase l_0. Thus, slope number is measured by $l_k = l_0 + k \cdot \Delta$, with a marginal percentage increase $\frac{l_k}{l_{k-1}} - 1 = \frac{g}{k}$. This extends equation 5.2 to account for interval-specific time trends l_k:

$$C_w = C_{w-1} + l_0 - (C_{w-1} - C_0) \cdot \delta \cdot X_{k,w} + int_{k,w} \cdot l_k \qquad \forall\, w \in W, k \in K \qquad (.1)$$

where $int_{k,w}$ is a dummy variable indicating 1 if the week is following hull cleaning k, and 0 otherwise. Figure A4.3 shows optimal intervals with $g = 0.25$ and $g = 0.75$, with $\delta = 0.5$.



(a) $g = 0.25$ **(b)** $g = 0.75$

Figure A4.3: Changing growth factor g

Figure A4.4 show optimal intervals with g ranging from 0 to 1.

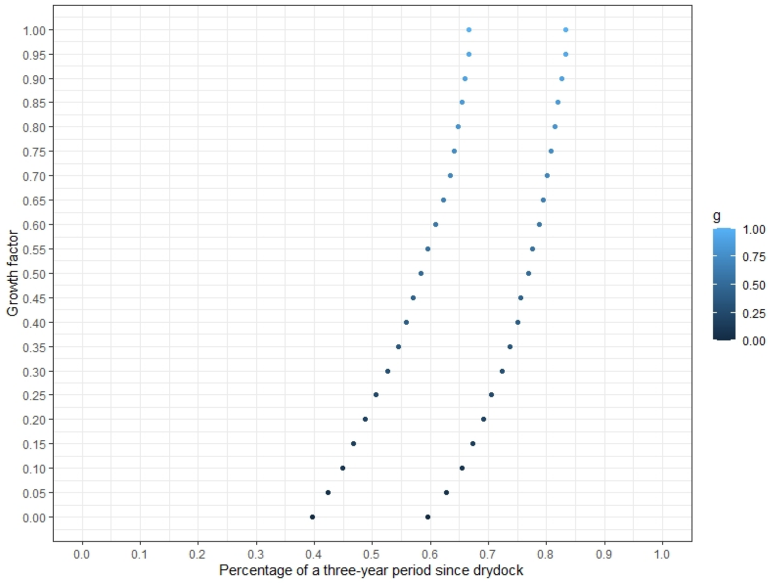
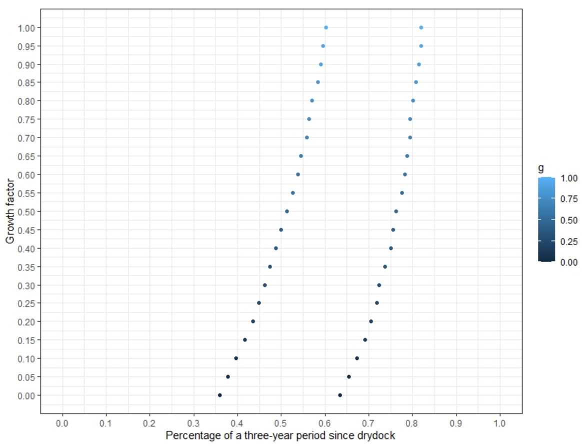


Figure A4.4: Optimal intervals for various growth factors

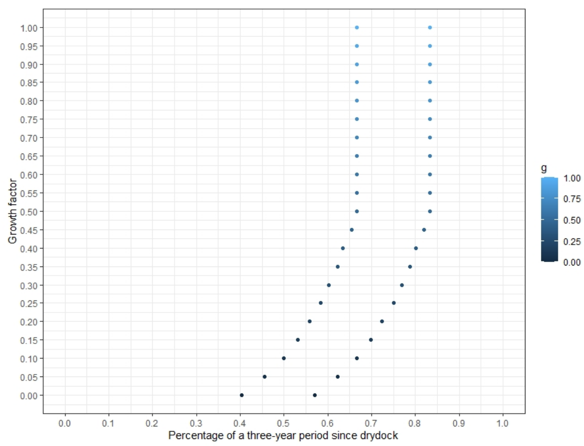
It is clear that both hull cleanings are pushed to higher percentiles when the growth factor increases. The graph is equal for all values for l_0 and C_0.

A4.3 Interval sensitivity changing shock size and growth factor

The graph of different growth factors with $\delta = 0.75$ in figure A4.4 is compared to the graphs in figure A4.5a and A4.5b with shock size parameters of $\delta = 0.75$ and $\delta = 0.25$ respectively.



(a) $\delta = 0.75$

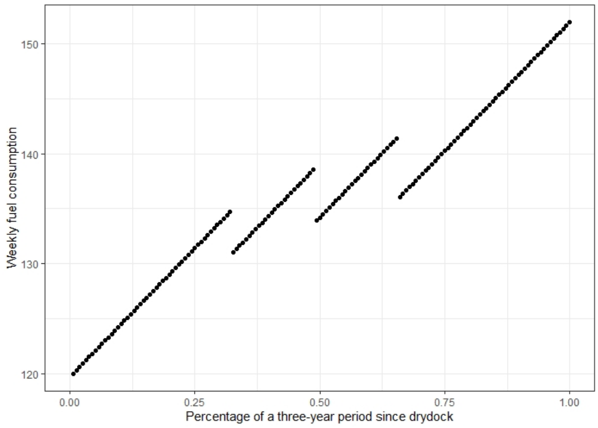


(b) $\delta = 0.25$

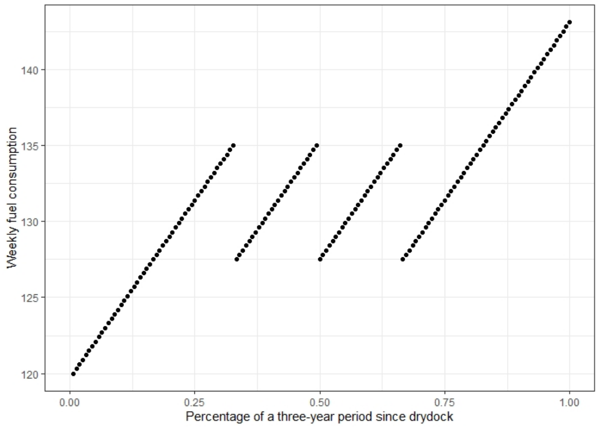
Figure A4.5: Comparison of growth factors with different shock sizes

It is clear that when lowering shock size, the later the optimal timing when growth rate increases, until the minimum interval constraint m is binding.

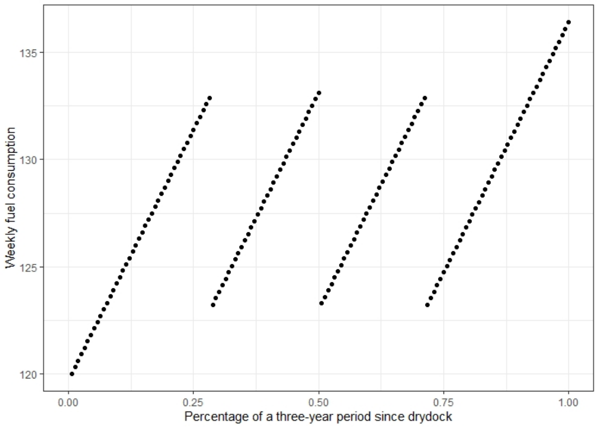
A4.4 Corresponding graphs with three performed hull cleanings



(a) $\delta = 0.25$



(b) $\delta = 0.5$



(c) $\delta = 0.75$

Figure A4.6: Fuel consumption slopes for different values of δ with three hull cleanings

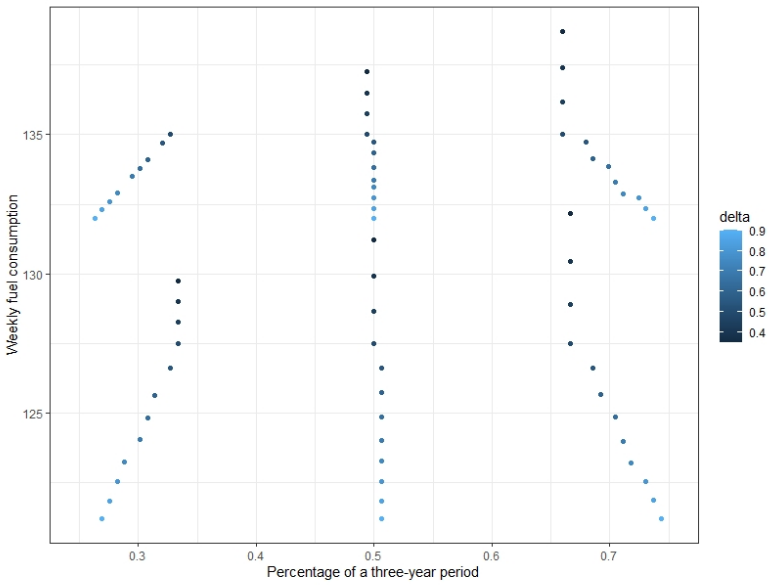
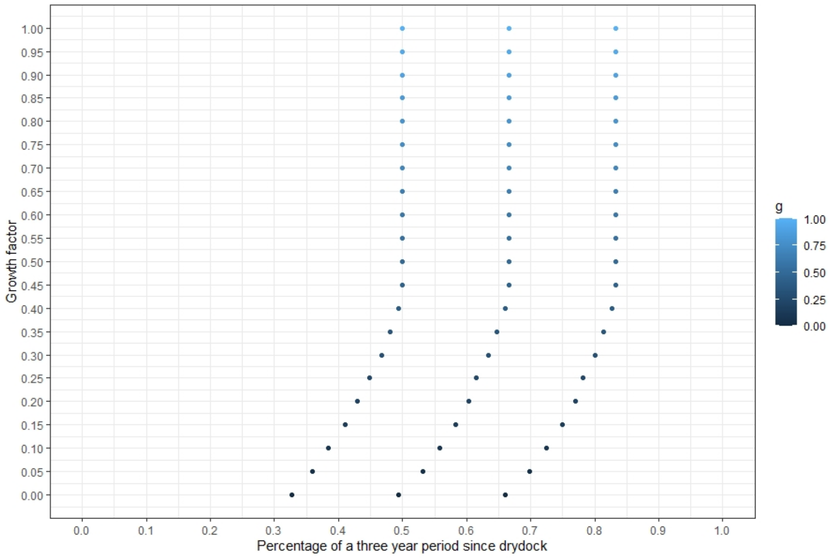
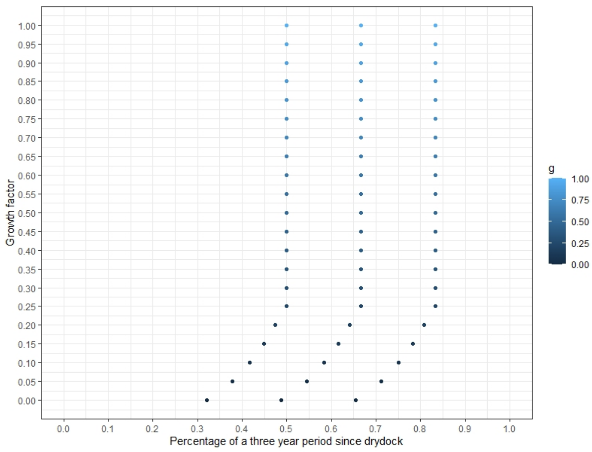


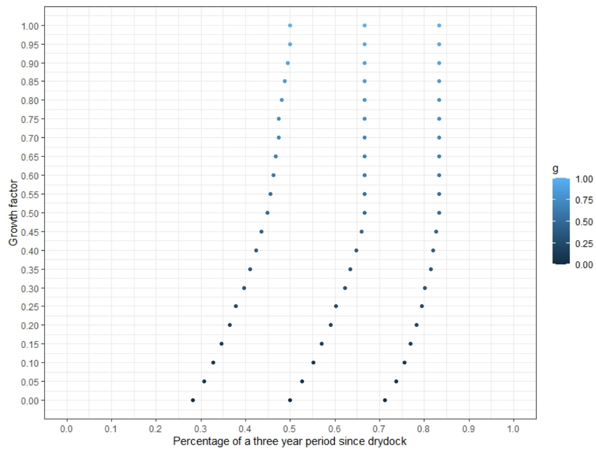
Figure A4.7: Optimal intervals for various shock sizes with three hull cleanings



(a) $g = 0.5$



(b) $g = 0.25$



(c) $g = 0.75$

Figure A4.8: Different growth factors with three hull cleanings

A5 Significant crew members with 5 % significance level

Table A5.1: Significant chief engineers with 5 % significance level

		5 % significance level	
Class	Master	Avg. Net Consumption	p-value
Suezmax	602	-0.4628	0.0014
Suezmax	620	-0.2302	0.0135
Suezmax	263	0.1436	0.0421
Medium Range	308	0.4625	0.0004
Medium Range	676	-0.2167	0.0485
Medium Range	64	-0.3431	0.0127
Medium Range	332	0.2038	0.0055

Table A5.2: Significant chief engineers with 5 % significance level

		5 % significance level	
Class	Chief engineer	Difference from Avg. Net Consumption	p-value
Panamax	547	-0.5115	0.0012
Panamax	319	0.3360	0.0002
Panamax	227	-0.3522	0.0227
Panamax	311	0.2699	0.0481
Panamax	813	0.4492	0.0203
Suezmax	430	0.2624	4.03e-05
Suezmax	434	-0.4334	0.0004
Suezmax	430	-0.4891	1.36e-07
Suezmax	434	0.7057	1.41e-05
Suezmax	82	0.2626	0.0002
Suezmax	169	-0.2140	0.0078
Suezmax	392	-0.3553	0.0012
Suezmax	463	0.2790	0.0017
Suezmax	430	0.6205	6.16e-06
Suezmax	152	-0.2054	0.0013
Suezmax	503	0.2672	0.0301
Suezmax	806	-0.4273	1.29e-06
Suezmax	637	0.0845	0.0170
Suezmax	90	0.2704	0.0191
Suezmax	82	-0.2590	0.0211
Medium Range	418	-0.0815	0.0492
Medium Range	351	-0.1272	0.0255
Medium Range	747	-0.3183	0.0107
Medium Range	223	0.5800	8.63e-05

A6 Variance decomposition for ballast observations

Table A6.1: Variance decomposition for Panamax vessels in ballast

	Df	Sum Sq	Mean Sq	Pr(>F)	Percentage
Log Speed	1	66.10	66.10	0.0000	77.11
Draught	1	0.27	0.27	0.0003	0.32
Weather	9	6.68	0.74	0.0000	7.78
Vessel name	5	0.99	0.20	0.0000	1.15
Weeks since drydocking	1	1.88	1.88	0.0000	2.19
Hull cleaning	3	1.30	0.43	0.0000	1.15
Master	34	6.38	0.19	0.0000	7.44
Chief engineer	28	2.12	0.08	0.0000	2.46
Residuals	1159	23.88	0.02		

Table A6.2: Variance decomposition for Suezmax vessels in ballast

	Df	Sum Sq	Mean Sq	Pr(>F)	Percentage
Log speed	1	9.11	9.11	0.0000	32.31
Draught	1	0.63	0.63	0.0000	2.23
Weather	9	8.62	0.96	0.0000	30.55
Vessel name	2	0.61	0.31	0.0000	2.17
Weeks since drydocking	1	0.22	0.22	0.0010	0.76
Hull cleaning	3	2.41	0.80	0.0000	8.56
Master	21	3.34	0.16	0.0000	11.85
Chief engineer	18	3.26	0.18	0.0000	11.54
Residuals	847	16.60	0.02		

Table A6.3: Variance decomposition for Medium Range vessels in ballast

	Df	Sum Sq	Mean Sq	Pr(>F)	Percentage
Log speed	1	30.42	30.42	0.0000	55.11
Draught	1	0.23	0.23	0.0052	0.41
Weather	9	11.80	1.31	0.0000	21.53
Vessel name	6	0.28	0.05	0.1405	0.51
Weeks since drydocking	1	2.88	2.88	0.0000	5.25
Hull cleaning	3	1.19	0.40	0.0000	2.17
Master	32	3.05	0.10	0.0000	5.56
Chief engineer	33	4.95	0.15	0.0000	9.03
Residuals	656	18.83	0.03		