

# **TOWARDS ENERGY EFFICIENT SHIPPING**

USING MACHINE LEARNING TO SUPPORT A SHIP'S CREW IN  
ENERGY EFFICIENT SAILING



Thesis for the degree MSc in Marine Technology in the specialization of Marine Engineering

# **TOWARDS ENERGY EFFICIENT SHIPPING**

**USING MACHINE LEARNING TO SUPPORT A SHIP'S CREW IN ENERGY EFFICIENT SAILING**

By

**Jelle VAN DER BOS**

Performed at

**Ministry of Defense , ROYAL NETHERLANDS NAVY**

This thesis SDPO.21.003.m. is classified as confidential in accordance with the general conditions for projects performed by the TUDelft.

**Monday February 15, 2021 at 2:00 PM**

## **Company supervisor**

Supervisor	Dr. W.W. Tiddens
E-mail	WW.Tiddens@mindef.nl

## **Thesis exam committee**

Chair/Responsible Professor	Ir. K. Visser
Staff Member	Dr. Ir. R.D. Geertsma
Staff Member	Dr. Ir. H.J. de Koning Gans
Staff Member	L. Cavalcante Siebert
Company Member	Dr. W.W. Tiddens

## **Author Details**

Studynumber	4454677
Author contact e-mail	jellevdbos@hotmail.com

Copyright © 2020 by J. van der Bos

An electronic version of this dissertation is available at  
<http://repository.tudelft.nl/>.



# CONTENTS

<b>Summary</b>	<b>vii</b>
<b>Samenvatting</b>	<b>ix</b>
<b>Preface</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	2
1.2 Literature Overview . . . . .	5
1.3 Research Questions . . . . .	9
1.4 Proposed Approach . . . . .	10
1.5 Thesis Outline . . . . .	11
<b>2 Background</b>	<b>15</b>
2.1 Design specifications . . . . .	15
2.2 Propulsion Layout . . . . .	16
2.3 Control System . . . . .	20
2.4 Monitoring System . . . . .	23
<b>3 Data preprocessing</b>	<b>25</b>
3.1 Data description . . . . .	26
3.2 Data exploration . . . . .	29
3.3 Construct required parameters . . . . .	30
3.4 Integration of data types . . . . .	40
3.5 Conclusions. . . . .	41
<b>4 Current operational performance</b>	<b>43</b>
4.1 Operational Profile . . . . .	43
4.2 Ship's resistance. . . . .	45
4.3 Fuel consumption . . . . .	48
4.4 Efficiency breakdown . . . . .	54
4.5 Diesel engine loading . . . . .	57
4.6 Conclusions. . . . .	58
<b>5 Data analysis method</b>	<b>61</b>
5.1 Multi-Layer Perceptron (MLP) Network . . . . .	61
5.1.1 Mathematical Formulation . . . . .	62
5.1.2 Hyper parameters . . . . .	64
5.2 Conclusions. . . . .	68

<b>6</b>	<b>Weather conditions from ship's data</b>	<b>71</b>
6.1	Feature Selection . . . . .	71
6.2	Data Preprocessing . . . . .	75
6.2.1	Feature Scaling. . . . .	77
6.2.2	Cross-validation . . . . .	77
6.3	Model Understanding. . . . .	77
6.4	Model training and validation. . . . .	78
6.4.1	Significant Wave Height . . . . .	80
6.4.2	Wave Period . . . . .	81
6.4.3	Wave Direction w.r.t. the ship . . . . .	81
6.5	Conclusions. . . . .	81
<b>7</b>	<b>Fuel Consumption Curves</b>	<b>83</b>
7.1	Target Understanding . . . . .	83
7.2	Model Description . . . . .	84
7.3	Feature Selection . . . . .	85
7.4	Train and test model . . . . .	86
7.5	Model output : Fuel Consumption Curves . . . . .	88
7.5.1	Comparison propulsion modes . . . . .	88
7.5.2	Comparison environmental conditions . . . . .	91
7.5.3	Comparison physical model . . . . .	92
7.6	Conclusions. . . . .	93
<b>8</b>	<b>Conclusions and Recommendations</b>	<b>95</b>
8.1	Discussion on results . . . . .	96
8.2	Recommendations for further work. . . . .	98
8.2.1	Data variety . . . . .	98
8.2.2	Combine with physical modelling . . . . .	99
8.2.3	Additional sensor measurements . . . . .	99
8.2.4	Accurate weather data . . . . .	99
8.2.5	Sensor calibration . . . . .	100
<b>A</b>	<b>Literature Search and Analysis</b>	<b>101</b>
<b>B</b>	<b>Weather data description</b>	<b>109</b>
<b>C</b>	<b>Vincenty's Formula</b>	<b>111</b>
	<b>References</b>	<b>113</b>

# SUMMARY

GLOBAL warming, caused by the greenhouse effect, is a much discussed term these days. Shipping plays a crucial role in the world economy, however responsible for about 2.5 % of global greenhouse gas emissions simultaneously. Therefore the current attention in the shipping industry is to reduce its environmental impact by using e.g. alternative fuels, power saving ship modifications or the power of wind. The state of the art alternative fuels are less power dense than the fuels currently used. To maintain the current endurance and capabilities of ships, in particular naval vessels, it is important to focus on how the energy available is used, now and in the future.

In this thesis the focus is on how the ship is operated by the crew. The crew of a ship is operating the ship and therefore their choices in speed and propulsion mode settings are highly affecting the energy efficiency of the ship. The ship in case is one of the Holland Class Oceangoing Patrol Vessels of the Royal Netherlands Navy. Those ships are acting in the lower spectrum of force, however should be able to perform there mission profile with an adequate reaction. During missions, their primary task is to complete the operation instead of reducing  $CO_2$  emissions. Therefore the aim of this thesis is to support a ship's crew in energy efficient sailing by taking into account the operational profile.

The ship's crew can be supported by fuel consumption curves which give the relation between the fuel consumption per mile and the speed over ground. From previous research we know that an optimal speed over ground exists, in fact a speed over ground with the least fuel consumption per mile. This optimal speed over ground is dependent on the propulsion configuration and environmental conditions. By giving real time insight to the ship's crew into the actual fuel consumption curves they are able to make well informed decision in  $CO_2$  dilemma's, e.g. saving fuel and being 10 minutes later at the destination, and they could maintain their mission profile.

Because of the lack of accurate data about the environmental conditions research has been down towards how we can infer environmental conditions from the ship's data. In this thesis a machine learning approach, multi-layer perceptron, is used to create models that are able to determine the actual environmental conditions from the ship's data only.

To establish fuel consumption curves, we developed algorithms able to determine the fuel consumption per mile by the ship's data and the inferred environmental conditions. By using the ship's data only it is possible to give the crew of the ship real-time insight into the fuel consumption curves for the different propulsion configuration corrected for the current environmental conditions.



# SAMENVATTING

**O**PWARMING van de aarde, veroorzaakt door het broeikaseffect, is een veel besproken onderwerp vandaag de dag. Scheepvaart speelt een grote rol in the wereld economie, maar tegelijkertijd is de sector verantwoordelijk voor 2.5 % van de totale broeikasgas uitstoot wereldwijd. Om deze reden is er veel aandacht voor om deze uitstoot te reduceren door bijvoorbeeld gebruik te maken van alternatieve brandstoffen, vermogensbesparende aanpassingen aan de schepen of door gebruik te maken van wind energie. Alternatieve brandstoffen hebben op dit moment een lage energie dichtheid in vergelijking met huidige fossiele brandstoffen. Daarom is het, nu en in the toekomst, belangrijk om te kijken naar hoe de energie aan boord van de schepen wordt gebruikt.

De aandacht in deze scriptie ligt daarom ook bij hoe het ship is gebruikt door de bemanning. De bemanning maakt namelijk keuzes over de snelheid en voortstuwingsmodus en daarmee de hoeveelheid brandstof die wordt verbruikt. Het schip waar dit onderzoek op is gebaseerd is een van de Holland Klasse 'Oceangoing Patrol Vessels' van de Koninklijke Nederlandse Marine. Deze schepen voeren missies uit in het lage geweldspectrum, maar moeten te allen tijde missies uit kunnen voeren met een hoog reactievermogen. Tijdens de een operatie is het hoofddoel om de missie te voltooien in plaats van het reduceren van emissies. Daarom is het doel van deze scriptie om de bemanning te ondersteunen in energie efficiënt opereren en daarbij aan hun missie profiel kunnen blijven voldoen.

De bemanning kan worden ondersteund door gebruik te maken van brandstofverbruiklijnen welke inzicht geven in de verbruikte brandstof per gevaren mijl tegen over de snelheid van het schip. De op dit moment uitgevoerde onderzoeken laten zien dat er een optimale snelheid voor het schip bestaat, in andere woorden een snelheid met het minste brandstof verbruik per gevaren mijl. Deze optimale snelheid is zeer afhankelijk van de gekozen voorstuwingsconfiguratie en de omgevingscondities waarin het schip opereert. Door deze brandstofverbruiklijnen realtime inzichtelijk te maken aan de bemanning van het schip zullen zij de mogelijkheid hebben om op elk moment een weloverwogen keuze te maken in snelheid en voortstuwings configuratie met het oog op reductie van emissies. Er kan bijvoorbeeld een keuze worden gemaakt om 10 minuten later op plaats van bestemming aan te komen middels een langzamere vaart, maar hiermee wel brandstof te besparen. Door het geven van realtime inzichten is het nog steeds mogelijk om te voldoen aan het missie profiel.

Accurate informatie over de omgevingscondities waarin het schip zich bevindt is niet beschikbaar. Daarom is onderzoek gedaan naar hoe deze omgevingscondities kunnen worden afgeleid van de scheepsdata, de metingen die constant aan boord van het schip worden gedaan. In deze scriptie wordt gebruik gemaakt van zelflerende algoritmes die het mogelijk maken om relaties tussen de scheepsdata en de omgevingscondities af te leiden. Deze modellen maken het mogelijk om met alleen de scheepsdata de omgevingscondities vast te stellen.

Het bepalen van brandstofverbruiklijnen volgt uit de algorithmes die zijn getraind om het brandstofverbruik per gevaren mijl te voorspellen voor elke voorstuwingsconfiguratie, afhankelijk van de huidige omgevingscondities. De methodologie zoals beschreven in deze scriptie maakt het mogelijk om brandstofverbruiklijnen op te stellen, met de huidige omgevingscondities in acht genomen. Deze worden opgesteld met alleen de scheepsdata, dit maak het mogelijk om ze op een realtime basis te presenteren aan de bemanning.

# PREFACE

**I**N September 2015 I started at the Technical University of Delft with the Bachelor Marine Technology. Thereafter I continued with the Master Marine Technology with the specialization Marine Engineering. During my Master I worked as a military working student for the department Maritime Systems at the Defense Material Organisation.

Throughout my period at the Defense Material Organisation and the Technical University I discovered the area of Life Cycle Management, Predictive and Condition Based Maintenance and the relevance of data. This brought me to the department Data For Maintenance in Den Helder at the Royal Netherlands Navy. In here I broaden my view about the importance of data and the enormous amount available. I was engaged in the project OceansX where we tried to use shared knowledge & technical innovations to start a movement towards  $CO_2$  reduction. During my military student job I provided the crew with insights into their energy efficiency and how they could improve it possibly. To create understanding of data they provided me with their practical experiences and together making steps into the reduction of  $CO_2$  emissions.

*"Coming together is a beginning,  
staying together is progress,  
and working together is success"*

**- Henry Ford -**

I could not have finished this project on my own, therefore I want to thank everyone who has helped me in this process. First of all, my supervisors Wieger Tiddens and Rinze Geertsma. Wieger, thanks for the inspiring conversations at the coffee corner in Den Helder and for the support of the writing process. Rinze thanks for the interesting conversations about the technical aspects and for the helpful reviews. Dennis Curves, and the colleagues at Data For Maintenance, thanks for the support and facility which made it even possible to perform data analysis.

Lastly I want to thank people who are very close to me. Special thanks to my girlfriend Iris, my parents Thijs and Martine, my sister Renske and friends and family for the support and the distraction in this process.

It was an experience I shall carry throughout the rest of my career and never forget!

*Jelle van der Bos  
Delft, February 2021*



# 1

## INTRODUCTION

*The maritime transport sector emits around 940 million tonnes of CO<sub>2</sub> annually and is responsible for about 2.5 % of global greenhouse gas (GHG) emissions, according to the GHG Study of the International Maritime Organisation (IMO), IMO (2014). Reduction of emissions from shipping and the impact on the environment is a well-known and much discussed topic these days because of the regulatory and financial pressure, [Viktorelius and Lundh \(2019\)](#). Therefore new technologies to reduce energy consumption arise such as the implementation of alternative fuels or substitution of power saving ship modifications. However, "Ships are operated by crews, often far away from the managers onshore, and are therefore responsible for many operational decisions affecting fuel consumption. Most of the suggested operational measures for energy efficiency are dependent on crew members' active participation and engagement." [Viktorelius and Lundh \(2019\)](#). Improvements in energy efficiency should therefore also be directed in supporting the crew by taking into account the everyday and ordinary activities.*

*This master thesis is performed at the department 'Data for Maintenance' at the Royal Netherlands Navy. This department together with the foundation 'OceansX' dedicate their efforts to the sustainability of the shipping sector by giving insights to operators of the ships into their fuel consumption based on data and inform them how to improve their ecological footprint.*

*This research supports a ship's crew in making optimal operational decisions regarding fuel consumption. Using operational ship data, insights can be given in the operational performance and the most energy efficient propulsion modes and speed settings. To be able to provide these insights data should be gathered, understood, prepared and analysed. Therefore, this thesis first proposes a methodology to process sources of raw data into a combined dataset that can be used for data analysis. Second, to establish weather conditions from the ship's data. Finally to give insight to the crew about their CO<sub>2</sub> performance and give advice to reduce their CO<sub>2</sub> emissions. All three objectives will contribute to innovations at the department 'Data for Maintenance' and the goals of the foundation*

'OceansX'.

*This chapter introduces the research performed in this master thesis; Section 1.1 gives the motivation for the research subject. In Section 1.2 we derived the gaps existing in the research that has been performed already. From these gaps the Research Questions are defined in Section 1.3. Then the proposed approach to answer the Research Questions is given in Section 1.4 and finally the thesis outline is defined in Section 1.5.*

## 1.1. MOTIVATION

GLOBAL warming is a much discussed term these days. Global warming is caused by the greenhouse effect, Callery and Bailey (2020b). This effect is a natural process where the earth's atmosphere is retaining some heat of the sun. This allows the Earth to host life. The biggest problem in global warming is that the human being is increasing the greenhouse effect by emitting greenhouse gases such as carbon dioxide and methane resulting in a temperature rise of the atmosphere even more.

There are numerous effects that will pop up due to this temperature rise, the most evident one is melting of the poles and so a rising sea level. Extreme weather phenomena are becoming more and more common, man could think of heavy rainfall, floods, drought that lead to enormous bushfires. Also the expectation that climate change will negatively affect nature and wildlife, extinction of many species, reduced diversity of ecosystems and damage in the ocean due to ocean acidification.

We as human being cannot avoid these climate changes and so completely avoid these negative effects, Callery and Bailey (2020a). However what we can do as human being is adapt to the consequences of these changes and to mitigate them as much as possible. Mitigation of these effects is a common interest and so also of the maritime industry.

The maritime transport sector emits around 940 million tonnes of  $CO_2$  annually and is responsible for about 2.5 % of global greenhouse gas (GHG) emissions, according to the GHG Study of the International Maritime Organisation IMO (2014). Furthermore, the shipping emission are projected to increase significantly between 50 % and 250 % in the period to 2050 under a business-as-usual. The growth of the  $CO_2$  emission can be mitigated by further action on efficiency and emissions, however the demand for maritime transport is projected to increase rapidly. If, as in the past, the ambition of the sector maritime transport continues to fall behind efforts in other sectors and if action to combat climate change is further postponed, the sectors'  $CO_2$  emissions share in global  $CO_2$  emissions may rise substantially to 17 % by 2050 ENVI (2015).

The projected increase in emission in the coming decades will undermine the objectives of the Paris Agreement UN (2015). The Paris Agreement builds upon the "United Nations Framework Convention on Climate Change" UN (1992), which has a near-universal membership with 197 parties to the convention. The Paris Agreement goal is to strengthen the global response to the threat of climate change by keeping a global temperature rise this century well below 2 degrees Celsius above pre-industrial levels and to pursue efforts to limit the temperature increase even further to 1.5 degrees Celsius.

To ensure that shipping will act according to the objectives of the Paris Agreement, the "International Maritime Organisation" (IMO) adopted regulations to address the emissions. First, in April 2018 IMO adopted an Initial IMO Strategy on reduction of GHG emissions from ships [IMO \(2018\)](#). The goal is to reduce GHG emissions from international shipping, more specific to reduce the total annual GHG emission by at least 50 % by 2050 compared to 2008. Secondly, a regulation consisting of energy-efficient measures as the "Energy Efficiency Design Index (EEDI)" and "Ship Energy Efficiency Management Plan (SEEMP)" to reduce emissions of shipping, under Annex VI of IMO's pollution prevention treaty [MEPC \(2011\)](#). The EEDI sets compulsory energy efficiency standards for new ships build after 2013 and the SEEMP requires ships to develop a plan to monitor and improve their energy efficiency.

Even though the new-build ships should take into account the required energy efficiency measures, the ships moving currently upon the sea continue there operations. Therefore, lessen the fuel consumption and taking into account there everyday and ordinary action is of importance to reduce fuel consumption and mitigate climate change.

### OPERATIONAL SHIP DATA

Data will play an increasing role in daily life as well as in shipping, for example the SEEMP. Measuring fuel consumption is an important component in energy efficiency management, for example the SEEMP. Also in the view of predictive maintenance, vessel performance monitoring or smart maintenance with regard to safety, reliability and emissions have a major influence in the desire for data. Data supported analyses give insight into the historical performance and could also give information about long term effects, for example the degradation of a diesel engine. Data can also be used in real-time, for example to support the actual operation.

The Oceangoing Patrol Vessels (OPV's) of the Royal Netherlands Navy are equipped with lots of sensors continuously measuring and storing data, mainly installed for monitoring and control purposes. The measurements consist of parameters that give insight in machine operation, ship motions, propulsion parameters, alarms, switches and many more. This data is made available and could be used to give insight into the ships' emissions during different operations, power configurations and weather conditions.

### OCEANGOING PATROL VESSEL

The Oceangoing Patrol Vessel is designed to perform missions in the lower spectrum of force. The primary tasks of the OPV are coast guard actions, maritime safety tasks and to support civil authorities. Also, the ship is designed to perform these actions in diverse areas, from subarctic to tropical. Design choices are made so that the ship is able to perform these various operations in a diverse environment and so it is only to a limited extent possible (during mid life updates) to reduce  $CO_2$  emissions by adjusting the design, the EEDI (Energy Efficiency Design Index) approach.

The ship's propulsion plant consist of two identical drive lines and has numerous of possible configurations. The drive lines can be driven both by their patrol electric motor (PEM) or main diesel engine, the latter requires a choice between the transit or manoeuvring mode. Also it can be chosen to have one drive line powered by the main diesel engine and the other drive line not powered, then the choice between the other

shaft trailing or blocked has to be made. The energy efficiency of the different propulsion modes is highly dependent on the loading conditions of the ship, the sailing speed, the weather circumstances and the technical condition of the components.

To satisfy the regulations as enforced by the IMO and to take responsibility for climate change, the energy efficiency should be improved. This could be reached by giving insights to the crew in the energy efficiency of the ship and how it could be improved. Giving insights to the crew of the ship about the energy efficiency of the different propulsion modes and speed settings makes it possible to make informed decisions about the usage of the ship with regard to minimizing fuel consumption.

## 1.2. LITERATURE OVERVIEW

ENERGY efficient sailing is a broad term and influenced by a lot of components inside and outside the vessel. Energy efficient sailing can be divided into two main components, the energy consumed for the propulsion of the ship and the energy consumed by the hotel consumers and other components. This thesis focuses on the energy used for the propulsion of the ship. The energy efficiency is defined as the kilogram of fuel consumed per sailed mile (FPM), see Equation 1.1 where  $\dot{m}_{f,tot}$  is the total consumed fuel for propulsion of the vessel in kilogram per hour and  $v_s$  is the ship speed over ground in knots.

$$FPM = \frac{\dot{m}_{f,tot}}{v_s} \quad (1.1)$$

The energy efficiency of the propulsion of the vessel is dependent on multiple factors as we can see in Figure 1.1. The yellow boxes in this Figure represent the operators influence on the energy efficiency of the ship, the speed setting, course and propulsion mode (e.g. electrical or mechanical propulsion). The green boxes represent the components that can not be influenced by the operator immediately.

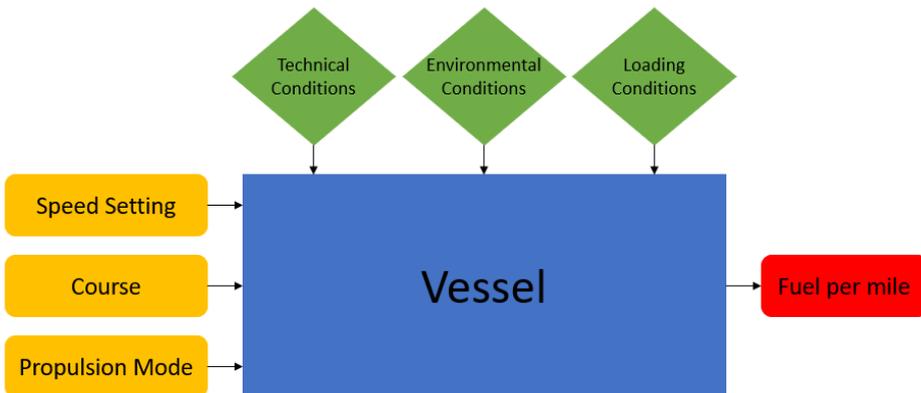


Figure 1.1: Components that determine the energy efficiency of the ship, divided into components that can be affected by the operator of the ship (yellow) and components that can not be affected by the operator of the ship immediately (green).

Traditionally, the fuel consumption was calculated by theoretical propulsion power for the actual condition using standard empirical resistance and propulsion methods, [Holtrop \(1984\)](#). These empirical methods are derived from model tests and sea trials, however a ship will encounter various conditions, e.g. load, trim or weather conditions. To have a more accurate model that gives insight into the fuel consumption real data is used more and more, because of the fact that this is a more precise approach towards real conditions.

The load condition of the vessel influences the fuel consumption of the vessel. A higher fuel efficiency can be reached by optimizing the trim of the ship. [Shivachev et al. \(2020\)](#) investigated the relation between trim and waves and propose a method that can determine an optimum trim. [Coraddu et al. \(2017\)](#) have been able to propose a trim optimization technique which exploits the predictive power of the proposed models for the online selection of the best configuration of the trim for reducing the fuel consumption.

Also, the technical conditions can influence the energy efficiency of the vessel. Propeller and hull roughness, marine fouling, are one of the components included into the technical conditions of the vessel. [Song et al. \(2020\)](#) developed an Unsteady Reynolds Averaged Navier-Stokes (URANS) based full-scale ship self-propulsion model to predict the effect of biofouling on the self-propulsion characteristics of the full-scale KRISO container ship (KCS). [Coraddu et al. \(2019a\)](#) focused on building effective Data-Driven Models to predict the hull state condition utilising the data collected from an automation system when the vessel was in operation. [Coraddu et al. \(2019b\)](#) build a digital-twin model to be tuned on data collected during a period of time where the marine fouling is not present and for a time period wide enough to observe the ship in many operational and environmental conditions. This digital-twin model is used to calculate the vessel's speed loss due to marine fouling.

As concern for the environment and the more data-driven regulations considering the protection of the environment, the Energy Efficiency Operation Index (EEOI) is a measure to monitor ship and fleet operational performance over time, provided by the [MEPC \(2009\)](#). EEOI can be calculated easily by operational ship data such as the fuel consumption during operation. However, the desire by shipyards about the EEOI of delivered ships and regulations that require monitoring the EEOI of all vessel with regard to the environmental protection is of importance only they have no access to the vessel operational data. Therefore much work is done to develop methodologies calculating the EEOI by using AIS data and leaving out the use of operational data. [Kim et al. \(2020\)](#) proposed a method to estimate EEOI without requiring the actual FOC, the verification using actual data shows that the proposed method can estimate EEOI from public data without actual FOC. More applications with AIS data are developed, [Feng et al. \(2020\)](#) demonstrates the value of deriving space-time trajectories from AIS data based on the concepts of time geography to assess time efficiency levels of maritime ports and monitor their performance over time, which offer useful information to both shipping lines and port authorities for operations such as efficient scheduling and logistic support. Also studies towards traffic behaviour, [Xiao et al. \(2015\)](#), route characterization and anomaly detection, [Rong et al. \(2020\)](#), based on AIS data are performed.

The use of noon reports as data source is widely used in the current prediction of fuel consumption. A noon report of a ship is a voyage report on a daily basis wherein the ship's sailing behaviour is noted. This behaviour is recorded by parameters such as the ship's geographical location, daily fuel consumption, distance travelled, cargo on board, engine speed and average propeller revolutions. Also the average weather of that day is reported, herein are parameters such as the wind direction, wind speed, wave height,

wave direction or sea current. [Yan et al. \(2020\)](#) developed a model that makes fuel and speed curves following from the noon reports of a handy-size dry bulk ship. [Isikli et al. \(2020\)](#) construct statistical models for fuel consumption with the noon report of a bulk carrier. The trend of future research in studies considering noon report data is that having high-frequent data can make the fuel consumption predictions more precise and accurate.

Fuel consumption is dependent on the weather conditions the ship is encountering. [Taskar and Andersen \(2020\)](#) studied the relationship between ship speed and fuel consumption with the use of monthly averaged weather data obtained from the *European Centre for Medium-Range Weather Forecasts* (ECMWF) with a resolution of 1 degree. They showed that the common assumption of cubic speed-power relation can cause a significant error in the estimation of bunker fuel consumption. Simulations in different seasons have revealed that fuel saving due to speed reduction are highly weather dependent. [Yan et al. \(2020\)](#) proposed a machine learning method performing regression task with high accuracy to make predictions on ship fuel consumption under different sailing speeds as well as cargo, weather and sea conditions. To implement the influence of weather the author includes weather data from the ECMWF that contains hourly data about the wave height and wind direction, with a resolution of 1 degree as said before.

Most research investigates the influence of weather on the fuel consumption of the vessel based on operational data make use of the weather data obtained from the ECMWF. This data has a resolution of 1 degree and all parameters are hourly averaged. As a result, the influence on the fuel consumption of the average weather can be investigated. The drawback of these methods is that the result is not that accurate, e.g. wind speeds could be 2 knots the first half of an hour and the second half 50 knots, having the same average as a continuous wind speed of 26 knots but a different impact on fuel consumption. Operational data that is measured with a higher frequency are influenced by these high and low wind speeds. Therefore, data with a different frequency of measurement frequencies can give a distorted view about what was intended to do.

Weather routing has seen a rapidly increasing attention from academics and practitioners in recent years, [Perera and Soares \(2017\)](#). [Zis et al. \(2020\)](#) performed a survey of weather routing and voyage optimization research in maritime transportation, explained the main methodological approaches, and the key disciplines that are dealing with this problem. Research towards weather routing is focused on two main aspects, the route optimization method and the vessels fuel consumption or movements due to different weather conditions. [Gkerekos et al. \(2019\)](#) developed a methodology that determines an optimal route decision support by historical sailing performance and the weather forecast. The fuel oil consumption (FOC) modelling based on 2066 data points out of the data results into a prediction of the FOC where 95.5 % of the predictions are within 10 % of their actual FOC. For the modelling of the FOC the author uses an artificial neural network (ANN). All applications or case studies in the field of weather routing are in the field of liner shipping, Ro-Ro, bulk or tanker vessels, in any case a vessel that has to go from A to B in a certain time. However, ships of the Royal Netherlands Navy have to perform

their missions with adequate reaction and therefore could alter their route throughout the mission, and therefore their route is not only from A to B.

Fuel consumption curves give the relation between the fuel consumed per sailed mile and the speed over ground. [Geertsma et al. \(2017\)](#) proposes a propulsion model that can be parameterised with publicly available manufacturer data and further calibrated with obligatory FAT measurements. This model predicts system performance within 5 % of actual measurements during FAT of the diesel engines and SAT of a case study navy ship. This study can help giving insight into the fuel consumption per mile, however this study gives only the relation in trial, design and off-design condition. [Bialystocki and Konovessis \(2016\)](#) established a ship fuel oil consumption curve prediction model in their research taken into account the ship's draft and displacement, weather force and direction and hull and propeller roughness. This research is intended for decision support and based on statistical analysis of 418 noon reports of a Pure Car and Truck Carrier. [Yan et al. \(2020\)](#) developed a machine learning model performing regression task with high accuracy to make prediction on ship fuel consumption under different sailing speeds as well as cargo, weather, and sea conditions. Based on this prediction the author proposed a speed optimization model. The data used in this study is the noon report of a handy-size dry bulk ship with in total 9 input variables and 242 useful entries. All models provide information about the fuel consumption per sailed mile against the speed over ground. However, again all models provide information based on noon reports or snapshots of the vessels performance. Real-time information seems to be difficult because of the fact that this requires real-time and high-frequent data.

The term Big Data is used to describe large and complex data sets that are difficult to process and analyse using traditional data processing techniques and application. [Zaman et al. \(2017\)](#) discusses features and risks associated with big data, how big data analytics will be turned into added value for the future maritime industry as well as the upcoming data oriented regulations. It covers the associated applications and obstacles to the implementation of big data in the industry. [Aiello et al. \(2020\)](#) proposed an analysis of the preparedness of the current shipping industry to face the fourth industrial revolution, while indicating the most significant open challenges and material insights. The result is that although it is true that on vessels, more and more information is gathered and stored as data, their involvement in the decision making process is still very limited, therefore such knowledge contributes only to a limited extent to strengthening of the value chain. A trend towards digitization is evident. Nevertheless, the current implementation level is uneven and fragmented, and outdated operational practices like manual data reporting, are still a common industrial practise.

An extended overview of all literature can be found in Appendix A. In this Appendix the search towards relevant research is described and the found literature that seems to be relevant is analysed by identifying the problem, key concepts and strengths and weaknesses.

### 1.3. RESEARCH QUESTIONS

THE literature research performed in section 1.2 identifies different gaps and difficulties into research performed in the field of ship fuel efficiency. One of these is the lack of high-frequency propulsion data and how to transform this data into useful measures about ship fuel efficiency. Also, a lot of studies investigate the influence of weather conditions using weather data obtained from outside sources that contain hourly averaged with a resolution of 1 degree or noon reports to with a resolution of 1 day. Accurate results about fuel consumption and the influence of weather conditions are not able to obtain, especially not for naval vessels. Therefore the following problem is answered in this thesis:

#### **How can $CO_2$ emissions be reduced given the operational profile of a naval vessel by giving real-time insights to the operator based on ship's data?**

The OPV of the Royal Netherlands Navy measures and stores thousands of parameters with a high resolution, therefore big data is a key concept in this thesis. [Zaman et al. \(2017\)](#) gives the application and obstacles to implementation of big data in the industry. [Aiello et al. \(2020\)](#) proposed an analysis of the preparedness of the current shipping industry to big data. The result is that although it is true that on vessels, more and more information is gathered and stored as data, the usage of data in the decision making process is still very limited, therefore such knowledge contributes only to a limited extent to strengthening of the value chain.

The focus of this thesis is to give ship operators and crews insights in their fuel consumption and optimal settings based on operational ship data. For this, analysis of raw vessel operational data requires in-depth analysis to acquire data free from disturbances, distortions and undesired physical effects, [Dalheim and Steen \(2020\)](#) presented a step wise method for preparation of data for ship operation and performance analysis. To deal with the difficulties of big data analysis and the use of raw data for ship performance analysis the first sub question in this thesis is:

#### **1 - How can $CO_2$ performance be established given the operational profile of a naval vessel with data?**

[Besikci et al. \(2016\)](#) predict ship fuel consumption for various operational condition through ANN and developed a decision support system employing ANN based fuel prediction model to be used on-board ships on a real time basis for energy efficient ship operations. This prediction model is build upon 'Noon Data' which is a daily report about the daily fuel consumption and average operational details filled by the crew. The decision support system gives support in the way of insight into the fuel consumption of the ship given the operational details at the moment observed by the crew. All research towards prediction of fuel consumption taking weather conditions into account require weather data from outside services or observation of the weather by the crew. Due to the fact that ships have to react rapidly onto changing conditions this information about weather conditions is not available.

Therefore, in this thesis research is done towards the possibilities of information

available during ship operations to establish the current weather conditions. Machine learning is used to investigate which parameters measured real-time, on change or every 3 seconds, by the ship will be representative to give information about the wave direction, period and the significant wave height. This leads to the second sub question in this thesis:

## **2 - How can real-time weather conditions be inferred from ship's data?**

Sailing speed is a highly determining factor in the fuel consumption of a ship and so it has a lot of impact on the emitted green house gasses. Therefore, due to focus on reducing greenhouse gasses a lot of research has been done about the optimal sailing speed. In this thesis the usage of the propulsion system is a highly determining factor and able to be influenced by the crew as well.

Ships of the Royal Netherlands Navy have to perform their adequate actions, and so their route is not only from A to B within a certain time, different from traditional weather routing problems. To achieve the lowest fuel consumption during their actions insight in their fuel consumption at the moment is required. Also, the current research done is focused on only one propulsion configuration the type of vessel is consisting of, liner shipping vessel mostly. The OPV of the Royal Netherlands Navy has multiple choices between propulsion modes, electrical driven, mechanical driven, one shaft trailing or one shaft blocked. Currently, there is no insight into the fuel consumption of those propulsion modes in different environmental conditions. Therefore, in this thesis research is conducted towards the fuel consumption per mile versus speed over ground for different propulsion modes and weather conditions based on the ship's data. This leads to the following research question:

## **3 - How can ship's data help the operator select optimal settings to reduce $CO_2$ emissions?**

### **1.4. PROPOSED APPROACH**

**T**HIS thesis first aims to propose a methodology to process sources of raw data into a combined dataset prepared for data analysis, second, to establish actual weather conditions from the ship's data only, and finally, to give insight to the ship's operator about their  $CO_2$  performance and give advice to the operator about  $CO_2$  emissions reduction.

In order to measure  $CO_2$  performance with data and answer Research Question 1, we first propose a methodology how the different sources of raw data can be combined into one dataset by taking into account the issues around accuracy. The geographical location of the ship considered in this thesis, a naval vessel, is unknown and therefore methods are applied to reconstruct the route and so weather data and ship's data can be combined. Also, by showing how  $CO_2$  performance can be measured we will make clear the potential of data regarding the use of the ship, modifications to the ship and new designs.

While a lot of research in the field of fuel efficiency affected by loading and weather conditions is based on hourly average and with a low resolution data, we propose a methodology to infer weather conditions from the ship's data and thus with a higher frequency and a higher resolution. In particular, the need for external weather data sources will disappear, research towards fuel efficiency affected by loading and weather conditions can be done based on the ship's data only and Research Question 2 is answered. The validation of the proposed method is done by transform the ship's data to the resolution of the weather data and calculate the mean absolute error between the true and predicted values. Further validation should be done.

The proposed methodology for the data processing is applied to the data sources and together with the methodology to establish weather conditions from the ship's data only we will establish fuel consumption curves depending on the weather conditions and the loading condition of the ship. The weather conditions established by the ship's data together with additional parameters from the ship's data will be used to develop a model able to determine the fuel per mile according to the given conditions which will be extended to a fuel consumption curves. These fuel consumption curves give the operator insight into the fuel consumption per mile as a function of speed over ground affected by the current weather and loading conditions. These curves give insight in optimal settings regarding the least  $CO_2$  emissions and thereby answering Research Question 3.

The proposed approach is visualized in Figure 1.2.

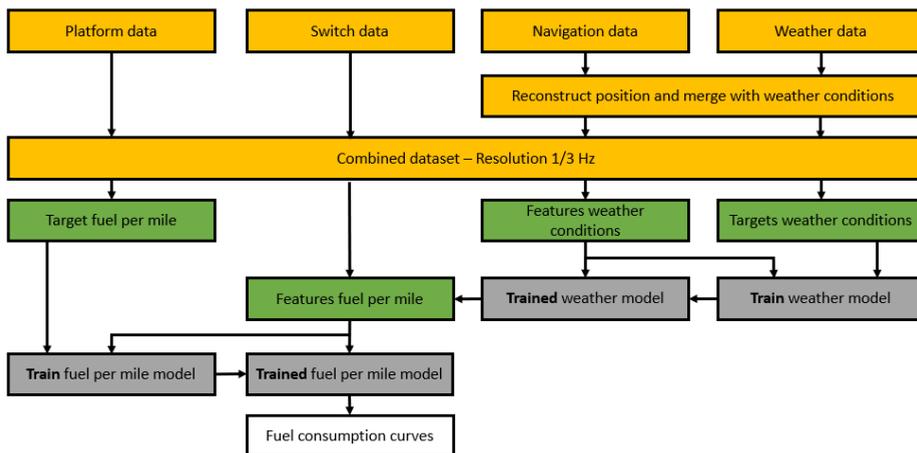


Figure 1.2: Schematic overview of the proposed approach. The yellow blocks are representing the databases, the gray blocks stand for the developed models and the green blocks are the target and feature variables from the database to train the models.

## 1.5. THESIS OUTLINE

THIS thesis is organised as follows and visualized in Figure 1.3.

- In **Chapter 2** background information is given to the work done in this thesis. The

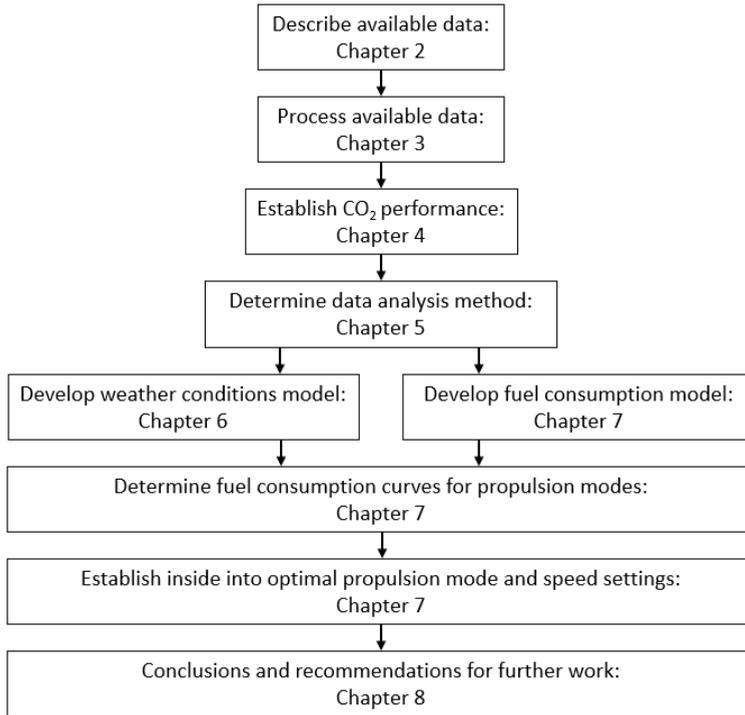


Figure 1.3: Schematic overview of the thesis outline given as major steps of the proposed methodology.

main working principles of the Ocean-Going Patrol Vessel of the Royal Netherlands Navy are set out, including the propulsion, control and monitoring systems.

- In **Chapter 3** required steps are gone through to prepare a dataset for data analysis. At first the four data sources used in this thesis are described. Second, a brief insight into the environmental conditions the ship has encounter in the time span of the dataset are visualised. Thereafter the data is enriched with parameters required for the research performed and finally the data sources will be merged to create a dataset prepared for analysis and modelling.
- In **Chapter 4** the operational performance of the ship derived from the dataset are described. At first we defined how the crew has operated the ship. Thereafter the resistance of the ship has been determined and it is described how  $CO_2$  emissions can be established. Finally, the different components in the propulsion system are quantified by their efficiency.
- In **Chapter 5** the data analysis method is explained that will be used in the next chapters. The mathematical formulation of this methodology is given as well as the different parameters that should be tuned during the training process of the model.

- In **Chapter 6** the data analysis method is applied to the data to infer ocean wave parameters from ship's data. The main components in this chapter are the feature selection for the machine learning model, hyper parameter tuning and the validation of the results.
- In **Chapter 7** proposes the methodology for the determination of fuel consumption curves. At first the target variable, the fuel per mile, as well as the features are describe upon which the model is trained are described. Thereafter a comparison between fuel consumption curves for the propulsion modes and environmental conditions is performed to give insight into optimal settings. Also, the model is validated and the performance of the model are investigated.
- In **Chapter 8** the main conclusions are given and the resulting fuel consumption curves are discussed. Finally, recommendations for future work are given.



# 2

## BACKGROUND

*The research done in this thesis is based on the Ocean Going Patrol Vessel of the Royal Netherlands Navy. This chapter provides the reader with background information to become familiar with the terminology used and the design information of the OPV that forms the basis of this thesis.*

*This chapter is organised as follows : Section 2.1 gives a brief overview of the initial mission and operational profile of the pilot ship. Thereafter in Section 2.2 the components of the propulsion system are described. Section 2.3 explains the different propulsion modes and the settings that can be chosen by the crew, in fact how the ship is controlled. Finally, in Section 2.4 a brief overview of the monitoring system for the sensors on board is given.*

### 2.1. DESIGN SPECIFICATIONS

**T**HIS section evaluates the initial design of the Oceangoing Patrol Vessel (OPV) of the Royal Netherlands Navy, see Figure 2.1. The Royal Netherlands Navy is consisting of a series of four OPV's which is referred as 'Holland-Class' ship types.

The OPV's are designed to perform missions in the lower spectrum of force. The type of missions the ship is designed for are coastguard tasks, maritime security and support of local authorities related. The coastguard tasks can be divided into enforcing laws and regulations and providing services. Those tasks should be performed in the Exclusive Economical Zones (EEZ) of The Netherlands as well as in the Netherlands Antilles and Aruba. The maritime security has to do with security, maritime presence and evacuation/disaster management which could have overlap with the tasks of support of local authorities also.

The operational profile of the ship is determined by those missions, the design profile is given in Table 2.1. The maximum speed of the OPV is 22 knots approximately, to intercept faster sailing vessel the ship is equipped with two FRISC interceptors and a NH90 helicopter. The FRISC's are used for embarkation on board of vessels that has to be inspected, to transport people and resources to and from the shore during relief- and/or



Figure 2.1: Holland Class Ocean-going Patrol Vessel of the Royal Netherlands Navy

rescue operations or infiltration of special forces. The helicopter can be used for image composition of the area and information gathering for boarding operations and 'Search and Rescue' operations.

Table 2.1: Designed operational profile of the OPV

Speed [knots]	Time [%]
0 - 5	20
5 - 10	10
10 - 15	40
>15	30

## 2.2. PROPULSION LAYOUT

THE OPV is designed with a hybrid propulsion system to satisfy the requirements associated with the defined missions and so the established operational profile. Hybrid propulsion is economically beneficial if the ship sails below 40 % of its top speed a significant amount of time. Also, hybrid propulsion is beneficial when the total electric load has a great spread over time and can improve availability and reduce noise, [Geertsma et al. \(2017\)](#). The propulsion system consist of two identical driving lines, those are both equipped with a main diesel engine (1), an electric motor (4), a gearbox (5) and a controllable pitch propeller (6), see [Figure 2.2](#).

The electric energy consumption is satisfied by the three diesel generator sets (3). The amount of diesel generator sets running depends on the amount of electric energy

consumption. The decision about which of the three diesel generator sets is running depends on the total running hours of the sets. The diesel generator set with the least amount of running hours will be chosen. The emergency diesel generator set (2) is used in situations where one or more of the diesel generator sets is in malfunction or in the maintenance.

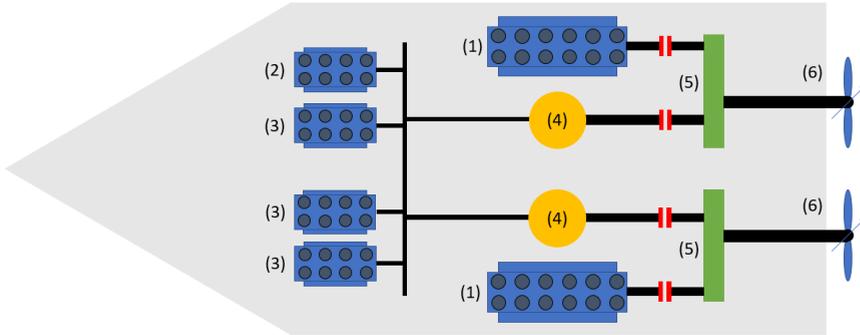


Figure 2.2: Propulsion lay-out of the OPV, wherein; (1) main diesel engine, (2) emergency diesel generator set, (3) diesel generator set, (4) electric motor, (5) gearbox, (6) controllably pitch propeller.

## MAIN DIESEL ENGINES

The mechanical propulsion in both drive lines of the ship is provided by the main diesel engine. The main diesel engines (12V28/33) are four-stroke V-engines with a stroke of 330 mm, a bore diameter of 280 mm and 12 cylinders that leads to a power output of 5400 kW at 1000 rpm. With a mean piston speed of 11 m/s these engines belong to the medium speed diesel engines portfolio and have a specific fuel consumption of 191 g/kWh at 100 % loading. The engine operating envelope of these diesel engines is shown in Figure 2.3. The minimum rotational speed of this engine is 40 % of the rotational speed at MCR which is equal to 400 rpm.

In this Figure the two load limits (Load Limit Range | and ||) that can be seen are defining areas of operation for the diesel engines. Range | is the operating range for continuous operation subjected to a light-running margin of 1.5 - 3 % where the lower value is to be aimed for. Range || is the operating range that is temporarily admissible during e.g. accelerations.

Low load operations of diesel engines are defined as operations at engine loads below 40 % of maximum continuous rating, [Tufté \(2014\)](#). This type of operation could be acceptable for a short period, however better to overcome. Namely, low load operation results in premature wear of pistons, piston rings, liners and crank case bearings and it comes with incomplete combustion resulting in increased emissions.

## DIESEL GENERATOR SETS

Electrical energy to provide all electric energy consumers in their consumption on board of the ship is generated by 1, 2 or 3 identical diesel generator sets. The number of diesel generator sets that are running is dependent on the amount of the demand for energy.

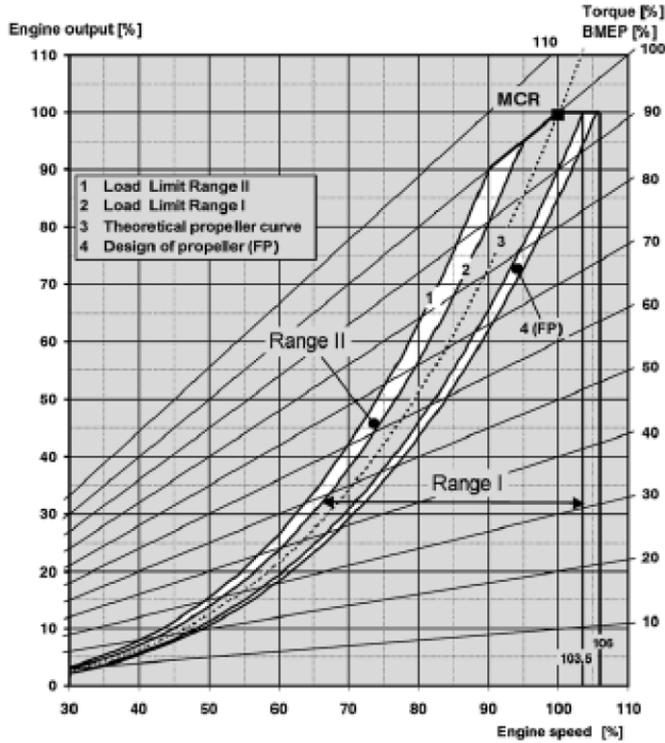


Figure 2.3: Power Envelope of the MAN 12V28/33 with different ranges of operations and constant torque and bmep lines. The power output is given as a percentage of the MCR power output which is 5400 kW. The engine speed is given as a percentage of the MCR engine speed which is 1000 rpm.

The emergency diesel generator set is there to generate electric energy in case of emergency, a total failure of the electric network or maintenance of the diesel generator sets. Each of the diesel generator sets is able to generate 920 kW at maximum load, Table 2.2 shows the characteristics of the diesel generator sets.

Table 2.2: The characteristics of the diesel generators on board of the OPV. The engine speed, Specific fuel consumption and electrical power are given for 4 different loading conditions.

Load [%]	25	50	75	100
Engine speed [rpm]	1800	1800	1800	1800
SFC [g/kWh]	255.8	222.7	210.4	204.0
Electrical Power [kW]	230	456	690	920

## CONTROLLABLE PITCH PROPELLERS

To match the required power by the propeller and the power output of the diesel engine without reaching the low load limit and prevent overloading of the diesel engine, a con-

trollable pitch propeller (CPP) is installed at both driving shafts. A CPP has a variable pitch that gives an additional degree of control, the power that is absorbed by the propeller from the diesel engine is not only dependent on the rotational speed of the shaft as by the traditional fixed pitch propeller (FPP), but also of the pitch of the propeller. When the pitch of the propeller is reduced at a certain shaft speed, the thrust delivered by the propeller is decreased and also the power that is absorbed from the diesel engine.

To make use of the extra degree of freedom there are predefined combinator curves. A combinator curve gives the relation between the rotational speed of the propeller and the pitch of the propeller. These combinator curves are designed in such a way that the diesel engines are prevented from low load or overloading operation, in fact these are chosen so that the diesel engine is running in its 'best point'. This point can be different for different ships, a choice can be to let the engine run in its lowest specific fuel consumption point or to run at an higher rotational speed to have higher manoeuvrability.

Also, a CPP utilises the ship with a higher manoeuvrability compared to a FPP. Namely, a controllable pitch propeller can give the vessel a forward or astern movement by control the pitch of the blades without reversing the rotational speed of the engine which requires more time.

The maximum rotational speed of the CPP installed on the OPV is 230 rpm and the minimum is 92, the diameter of the propeller is 3.2 metres.

## ELECTRIC MOTORS

During the design of the ship it was chosen to equip the ship with electric motors because of the mission profile, namely during patrolling the ship speed is most of the time below nine knots. Therefore for speeds up to nine knots a squirrel cage induction motor with an output power of 400 kW controlled by a variable speed drive is installed at both the drive lines. The nominal electric motor speed is 1788 rpm and the minimum rotational speed of the engine is 714 rpm because of the minimum rotational speed of the propeller. The nominal power of the electric motors is 400 kW therefore, two diesel generator sets are required to run when the electric motors are used.

## GEARBOXES

Both drive lines of the OPV are equipped with an 'two-input one-output' gearbox to deal with the different speed ranges of the main diesel engines and the electric motor. The gearbox configuration is a 'double helical' system, in this system the teeth of the gears are placed angled with respect to the tread. The greatest advantage of this type of gears is that the forces on the gear teeth are distributed over more than one tooth. This makes the gearbox more silent and the teeth encounter a more evenly distributed load which slows down degradation.

The gearbox has two different gearbox ratio's, this is the ratio between the rotational speed of the driving shaft and the driven shaft, see Equation 2.1. Herein  $i$  is the gearbox ratio,  $n_{primary}$  is the rotational speed of the driving shaft in rpm and  $n_{secondary}$  the rotational speed of the driven shaft in rpm. The gearbox ratio between the main diesel engine and the propeller shaft is 4.355 and for the electric motor 17.88.

$$i_{gb} = \frac{n_{primary}}{n_{secondary}} \quad (2.1)$$

Gearbox losses are typically divided into four main causes; friction in the bearings because of resistance in the bearings, gear teeth friction because of the friction between the teeth of the gears, swirl losses because of that the gears are rotating through a bath with oil and friction in the seals because the seals are there to keep the oil inside the gearbox, [Engelbrecht \(2015\)](#).

In general analysis, the power losses in gearbox is 1-2 % for single stage reduction gearbox and 3-5 % for complex gearbox, with two or three reduction stages. Also, when fully loaded a gearbox exhibit higher efficiency than when it is partially loaded, [Shi et al. \(2009\)](#), this is illustrated in Figure 2.4. In this thesis the gearbox losses are modelled as proposed by [de Waard \(2014\)](#).

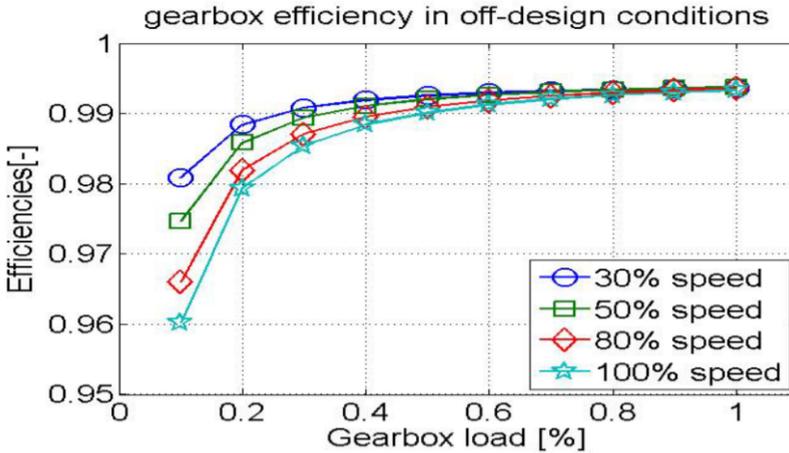


Figure 2.4: Gearbox efficiencies in design and off-design conditions, [Shi et al. \(2009\)](#)

### 2.3. CONTROL SYSTEM

**I**N this section the control strategies implemented on board of the OPV and how they can be used are explained. At first the control strategy by the 'virtual shaft speed' which is the speed setpoint that the crew sets is explained. Thereafter all different propulsion modes possible on the ship are explained.

#### VIRTUAL SHAFT SPEED

According to [Vrijdag et al. \(2014\)](#), the virtual shaft speed is a compound variable that contains both actual shaft speed ( $n_p$ ) and actual pitch ( $\theta_p$ ), see Equation 2.2. In this formula,  $\theta_{nom}$  and  $\theta_0$  are constants that stand for the nominal and the zero-thrust pitch angle of

the propeller. In static conditions, the virtual shaft speed is almost linearly related to ship speed, which allows for an intuitive use by the crew.

$$n_{virt} = \frac{\theta_p - \theta_0}{\theta_{nom} - \theta_0} \cdot n_p \quad (2.2)$$

Because of the different combinator curves in all propulsion and control modes, the linear relation in static condition between the virtual shaft speed and the speed over ground is different for all. Therefore, when the decision to sail a certain speed is made, the corresponding virtual shaft speed that has to be set should be looked up in the relation between the speed over ground and virtual shaft speed for the corresponding propulsion configuration. A change in propulsion configuration will result in a significant different virtual rotational speed of the propeller for a short time because of the difference in combinator curves.

### PROPULSION MODES

The ship has the possibility to sail into different propulsion modes, in fact to use different propulsion configurations. There are six different propulsion configurations; transit, manoeuvring, trailing shaft, blocked shaft, electrical driven and a stopped shaft. The choice for a certain propulsion mode is partly dependent on the crew's intuition and it is partly defined by a certain propulsion mode for a certain operation.

#### MANOEUVRING AND TRANSIT MODE

These two propulsion modes both using the propulsion configuration where the two propeller are driven by their main diesel engine. The difference between these two is the combinator curve. The manoeuvring mode combinator curve, see Figure 2.5, is chosen to ensure;

- sufficient available engine torque at low rotational speeds of the diesel engine for fast acceleration
- increased margin to the engine load limit in view of the higher acceleration of the ship
- smooth load increase during acceleration
- restricted engine load at maximum rpm to limit thermal engine load during acceleration

The transit mode combinator curve, see Figure 2.5, is chosen to ensure;

- sufficient margin to the engine load limit, see Figure 2.3, for continuous operation based on a propeller loading in sea state 4
- limited cavitation
- fuel economic operation

The propulsion mode transit has a lower speed in adjusting the rotational speed and pitch of the propeller which makes that the ship will have a lower maximum acceleration in comparison to the manoeuvring mode. They both have a different combinator curve as can be seen in Figure 2.5. The manoeuvring mode has a higher margin to the engine load limit because of high accelerations, this also makes the mode less fuel efficient compared to the transit mode, so sailing in transit is preferable when the manoeuvring mode is not necessary.

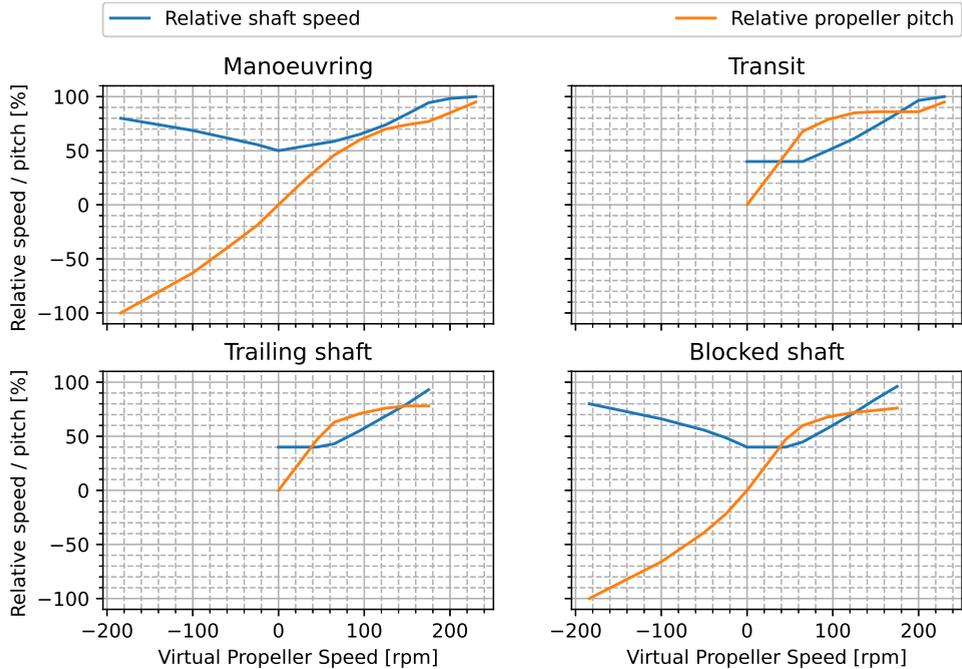


Figure 2.5: Combinator curves for the transit, manoeuvring, trailing shaft and blocked shaft mode, values of the pitch and shaft speed are relative to the maximum pitch and shaft speed and the virtual shaft speed is given in rpm.

### TRAILING SHAFT MODE

A trailing shaft is a propulsion mode that can be applied to the starboard or portside propulsion train at the time. In this mode one propulsion train is driven by the main diesel engine and the other shaft is disconnected to all power generators and so freely rotating with the maximum propeller pitch. The freely rotating shaft is driven by the water that flows through the propeller this is also called the 'windmill effect'. Settings such as engine speed adjustment rate and pitch adjustment rate are based on the transit mode. In this mode the driving propeller is required to maintain a certain ship speed to keep trailing the other shaft.

The trailing mode combinator curve, see Figure 2.5, is chosen to ensure;

- sufficient margin to the engine load limit, see Figure 2.3, for continuous operation based on a propeller loading in sea state 4 with one shaft trailing
- combinator curve and other settings as much as possible in accordance with the normal transit mode

#### BLOCKED SHAFT

A blocked shaft is a propulsion mode that can be applied to the starboard or the portside propulsion train at the time, as in the trailing mode. The combinator curve, see Figure 2.5, for this mode is chosen with the same requirements as for the trailing shaft mode that are listed above. The blocked shaft has no rotational speed and so the propeller act as an umbrella also called the 'umbrella-effect'. In this mode it is recommended to set the pitch of the blocked shaft to the maximum to lessen the 'umbrella-effect' of the propeller.

#### PEM

In this propulsion mode both propellers are driven by their patrol electric motor (PEM). It is only possible to sail with both electric motors, the propulsion modes blocked shaft or trailing shaft are not possible in combination with the electric motors. The PEM mode is design for patrolling operations, namely in this mode the maximum speed of the ship is around 9 knots.

The PEM combinator curve, see Figure 2.6, is chosen to ensure;

- operation at design pitch as long as possible for low noise and high efficiency
- a minimum shaft speed of 40 rpm to ensure sufficient lubrication of the outboard bearings

#### STOPPED SHAFT

The stopped shaft mode is applied on both propulsion trains and to the shaft itself. In this mode the shafts are physically blocked and are not able to rotate anymore. This mode is applied when the ship is at the shore and the ship is 'locked'. The relevance of this mode is for security and safety on board of the ship, in this thesis this mode is used to identify periods that the ship is in harbor and so a voyage is ended or a voyage will start.

## 2.4. MONITORING SYSTEM

**T**HE Integrated Platform Management System (IPMS) provides the remote monitoring and control services of a ship's platform by means of computerized equipment. This system allows even a single operator to monitor and control platform systems needed for power generation, propulsion, domestic services etc. The system measures lots of parameters (thousands), for example the rotational speed of the shaft but also if a specific door is closed or opened. The different types of parameters can be divided into three different categories;

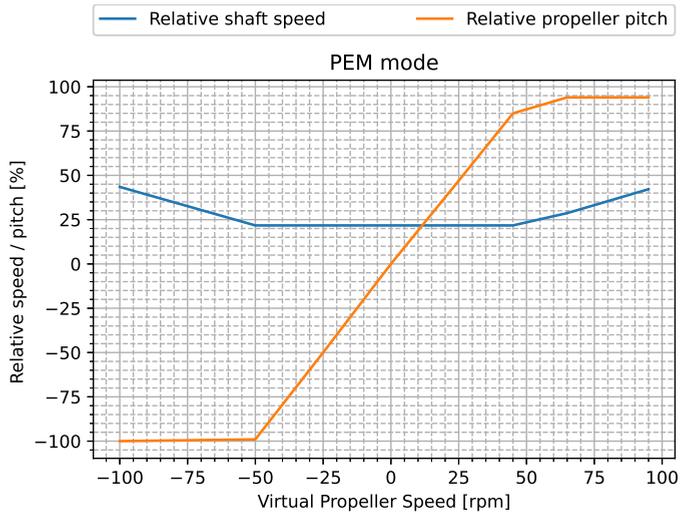


Figure 2.6: Combinator curves for the PEM mode, values of the pitch and shaft speed are relative to the maximum pitch and shaft speed and the virtual shaft speed is given in rpm.

- *Switch Parameters*, this type of parameters contains values like 0 and 1 where every number has his own meaning, so for example 0 when a door is opened and 1 when the door is closed.
- *Platform Parameters*, this type of parameters contains real values, so for example the rotational speed of a shaft or the temperature of a cylinder in the diesel engine.
- *Navigation Parameters*, this type of parameters contains also values, however these values are about the navigation of the ship, for example the speed of the ship or the relative wind speed the ship encounters.

The IPMS also stores all monitored parameters into a dataset for every type described above. The platform data is stored every three seconds, the values in between these three seconds are not stored.

The navigation data and switch data are stored on change and on the hour, this means that when a value changes the value is stored and every hour all values are stored. Due to this type of storage we will not miss any value, because it is evident that when there is no value the value is equal to the closest value back in time.

# 3

## DATA PREPROCESSING

*Data preprocessing is an important step in data science projects. The Cross-Industry Standard Process for Data Mining methodology (CRISP-DM) provides a structured approach to planning a science project, see Figure 3.1. It all starts with the business understanding, in fact this step is answering the question what you want to accomplish. Chapter 1 contains the content of this step, setting research questions and establish the proposed approach.*

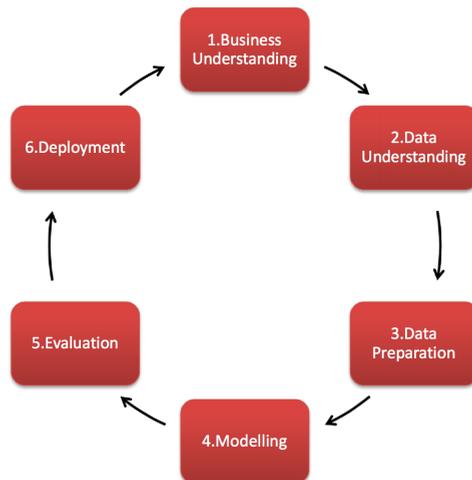


Figure 3.1: Visualization of Cross-Industry Standard Process for Data Mining methodology (CRISP-DM)

*This chapter describes the content of the next two steps: Data Understanding and Data Preparation. Data understanding consists of the description of the data that has been acquired and the exploration of the data. The data preparation step includes selecting the parameters from the data relevant for this thesis, those relevant parameters are called features. Also, it includes the construction of required data wherein new parameters will*

be derived from existing parameters in the data. The last step is the integration of all data sources. In this step all data sources are merged into one dataset prepared to be used in this thesis.

This chapter is organized as follows: Section 3.1 provides a description of the four data sources available and used in this thesis. Secondly, Section 3.2 gives insight into the environmental conditions the ship encountered during the time span of the data. In Section 3.3 the data is enriched by required parameters relevant in this thesis and required to integrate the data sources. Finally, Section 3.4 describes the integration of data sources and gives a description of the structure of the integrated dataset.

### 3.1. DATA DESCRIPTION

THE different data sources available in this thesis are:

- Platform data: value data from sensors from the integrated platform management system (IPMS) aboard the ship
- Switch data: switches indicating whether something is open/closed (e.g. a valve) or on/off (e.g. an engine)
- Navigation data: measuring all parameters related to the ship's navigation (e.g. speed and course) and environmental conditions (e.g. wind speed)
- Weather data: data from an external source containing oceans waves and atmosphere parameters for all oceans over the world.

All data as measured by the ship is known as the ship's data (platform, navigation and switch data). This section will describe the data sources introduced above in more detail.

#### PLATFORM DATA

The platform data is originating from parameters measured by the integrated platform management system (IPMS). The resolution of the platform data is one record stored every one, two, three, four, five or six seconds depending on the capacity and loading of the system. However, in this dataset, the resolution is generally 1/3 Hz. A record in the dataset includes the values of all parameters measured. The parameters are listed in Table 3.1 and are stored with three decimal places.

#### SWITCH DATA

The resolution of the switch data is varying because the values of the individual parameters are stored on change and every hour. This means that if the value of the measured parameter changes it will be stored at the time of change and that all measured parameters will be stored every hour on the hour. The values of the parameters are stored with one decimal place, therefore a value is stored at the moment that digit will change. Table 3.2 lists the different parameters in the switch data.

The switch data consists of six parameters, three for every propulsion train. The translation of these switches is given in Table 3.3. The translation in Table 3.3 is identical for both the portside and starboard propulsion train. The control mode switches

Table 3.1: Measured parameters from the integrated platform management system

Variable name	Unit
Rotational speed (propellers)	[rpm]
Rotational speed setting (VRA)	[rpm]
Rotational speed setting (main engines)	[rpm]
Rotational speed settings (patrol electric motors)	[rpm]
Fuel rack position (main engines)	[%]
Load (diesel generators)	[%]
Fuel consumption (main engines)	[kg/s]
Fuel consumption (diesel generators)	[kg/s]
Power (patrol electric motor)	[kW]
Actual pitch (propellers)	[%]
Torque (propellers)	[kNm]
Pitch	[degree]
Roll	[degree]
Ship draft (forward)	[m]
Ship draft (aft)	[m]

Table 3.2: Measured parameters from switches determining the propulsion mode of the ship.

Variable name	Unit
Propulsion mode portside	[-]
Propulsion mode starboard	[-]
Control mode transit portside	[-]
Control mode transit starboard	[-]
Control mode manoeuvring portside	[-]
Control mode manoeuvring starboard	[-]

contain information about the propulsion modes only when the ship is driven by the main diesel engines. The difference between the transit and manoeuvring mode are the used combinator curves and the propeller pitch adjustment speed. These switches give insight into which propulsion modes is chosen.

## NAVIGATION DATA

The resolution of the navigation data is varying because of the fact that the values of the individual parameters are stored on change and every hour (identical to the switch data). This means that if the value of the measured parameter changes (e.g. a small deviation in the course of the ship) it will be stored at the time of change and that all measured parameters will be stored every whole hour. The values of the parameters are stored with one decimal place, therefore a value is stored at the moment that digit will change. Table 3.4 lists the different parameters in the switch data.

Table 3.3: Translation table to gather information about the propulsion mode, this table is identical for both the starboard and portside propulsion train.

Propulsion Mode		Control Mode Transit		Control Mode Manoeuvring	
Value	Translation	Value	Translation	Value	Translation
1	-	1	No transit	1	No manoeuvring
17	Blocked shaft	17	Transit	17	Manoeuvring
33	Main diesel engine				
49	Patrol electric motor				
65	Trailing shaft				
97	Stopped				

Table 3.4: Measured parameters related to navigation and environmental conditions.

Variable name	Unit
Outside air temperature	[ °C ]
Relative humidity	[%]
Speed over ground (SOG)	[knots]
Speed through water (LOG)	[knots]
Course over ground (COG)	[degree]
Heading	[degree]
Relative wind speed	[km/h]
Relative wind angle	[degree]

### WEATHER DATA, HERSBACH ET AL. (2021)

Ocean waves and atmosphere parameters are obtained from *The Climate Data Store (CDS)*, a product of *The Copernicus Climate Change Service (C3S)*, [Hersbach et al. \(2021\)](#). C3S supports society by providing authoritative information about the past, present and future climate in Europe and the rest of the World.

*The Climate Data Store (CDS)* provides hourly averaged for a large number of atmospheric, ocean-wave and land-surface quantities. Data has been regridded to a regular lat-lon grid of 0.25 degrees for the atmosphere and 0.5 degrees for the ocean wave parameters. Degrees of latitude are parallel and so each degree remains constant, around 111 kilometers. A degree of longitude is widest at the equator, 111 kilometer, and gradually shrinks to zero at the poles. In the area where the ship operates a longitude degree is around 68 kilometer. Therefore the ocean wave parameter are defined in a lat-lon grid of around 56 x 34 kilometers and the atmosphere parameters in a lat-lon grid of around 28 x 17 kilometers.

The weather data contains over the hundred different parameters which describe the weather in each area over the world. The parameters from the *The Climate Data Store (CDS)* are summarised in Table 3.5 and described in detail in Appendix B.

Table 3.5: Parameters used from *The Climate Data Store* representing ocean waves and wind conditions.

Variable Name	Unit
10m u-component of wind	[m/s]
10m v-component of wind	[m/s]
Mean wave direction	[degree w.r.t. True North]
Mean wave period	[s]
Significant wave height	[m]

### 3.2. DATA EXPLORATION

**T**O gather insight into the different parameters in the datasets we will visualise them in this section. At first, the weather conditions the ship has encountered in the time frame of the data are evaluated.

#### WEATHER CONDITIONS

To gather information about the weather conditions in the dataset we establish the sea states the ship has encountered. The weather conditions are consisting of the significant wave height, wave period, wave direction, true wind speed and direction. The World Meteorological Organization (WMO) sea state code largely adopts the 'wind sea' definition of the Douglas Sea Scale, based on the significant wave height. The definitions of the sea states are summarised in Table 3.6.

Table 3.6: Sea State Codes as defined by the World Meteorological Organization (WMO) based on the definition of the Douglas Sea Scale.

WMO Sea State Code	Significant Wave Height
0	0 metres
1	0 to 0.1 metres
2	0.1 to 0.5 metres
3	0.5 to 1.25 metres
4	1.25 to 2.5 metres
5	2.5 to 4 metres
6	4 to 6 metres
7	6 to 9 metres
8	9 to 14 metres
9	Over 14 metres

The wind speed in the dataset is given as an u- and v-component wind speed in m/s, see Figure 3.2. The true wind speed is calculated by the root of the sum of the squares, see Equation 3.1.

$$TWS = \sqrt{u_{wind}^2 + v_{wind}^2} \quad (3.1)$$

The true wind angle with respect to True north (*TWA*) is also calculated by these components, see Equation 3.2.

$$TWA = \frac{\pi}{2} - \arctan\left(\frac{v_{wind}}{u_{wind}}\right) \quad \forall u_{wind} > 0$$

$$TWA = \frac{\pi}{2} - \arctan\left(\frac{v_{wind}}{u_{wind}}\right) + \pi \quad \forall u_{wind} \leq 0 \ \& \ v_{wind} \geq 0 \quad (3.2)$$

$$TWA = \frac{\pi}{2} - \arctan\left(\frac{v_{wind}}{u_{wind}}\right) - \pi \quad \forall u_{wind} \leq 0 \ \& \ v_{wind} < 0$$

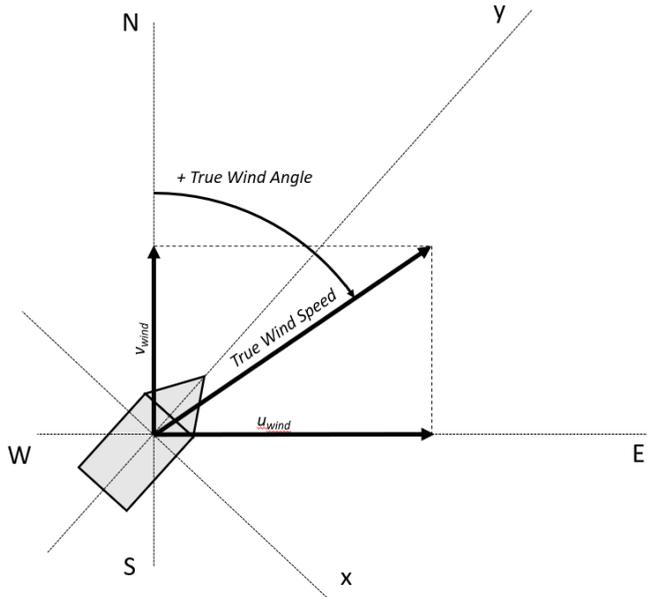


Figure 3.2: Visualisation of the calculation of the true wind speed and direction from the u- and v-component from the data.

Figure 3.3 summarises the weather conditions from the data. We can establish that the weather conditions in the data correspond with sea state 1 till 4/5 approximately which is in accordance with the characteristics from calm to moderate conditions. It can also be seen that the distribution of wind direction is almost equal with the distribution of the direction of waves.

### 3.3. CONSTRUCT REQUIRED PARAMETERS

THIS section gives an overview of the parameters that are added to the data and will be used further in this thesis. The dataset is enriched by the latitudinal and longitudinal position, propulsion modes, voyages, propeller thrust, set and drift, waves and wind w.r.t. the ship and the mean draft.

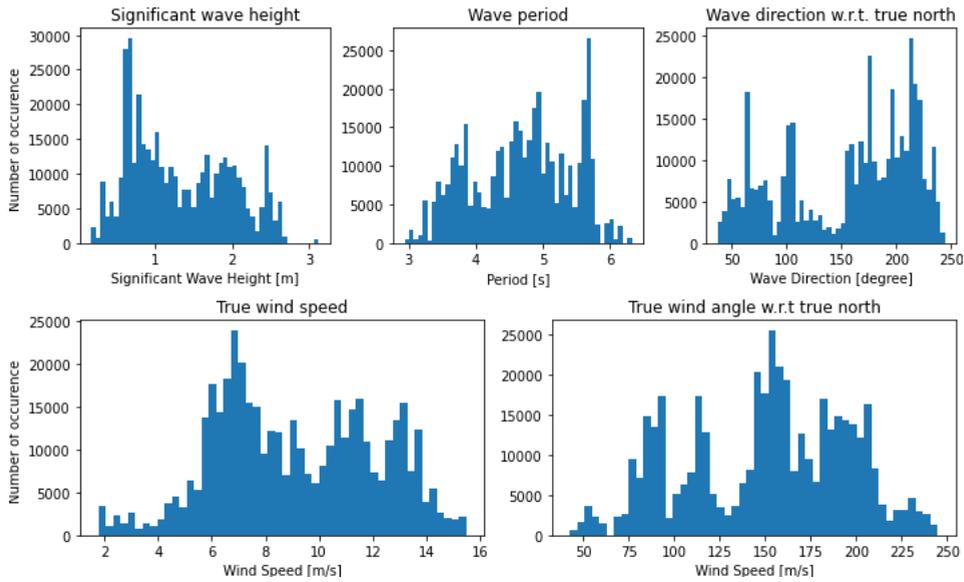


Figure 3.3: Exploration of the distribution of environmental conditions the ship encountered during the time span of the data.

### LATITUDE AND LONGITUDE

As can be noticed from Section 3.1 the latitude and longitude position of the vessel are not available in the ship's data. However, to merge the weather data with the ship's data the position of the ship expressed in latitude and longitude coordinates is required. The weather data contains the weather conditions at a certain time and geographical position, expressed in latitude and longitude coordinates. Therefore, we first have to reconstruct the position of the ship by making use of the speed over ground and course over ground from the navigation data.

To reconstruct the location of the ship, the distance sailed in a certain direction can be summed up from the initial position and so the route can be calculated, as the course over ground and speed over ground are available in the navigation data. The sailed distance,  $s$  in miles between the records is calculated by multiplying the ship's speed over ground,  $V_s$ , in knots and the difference in time between the records,  $\Delta t$ , in hours, see Equation 3.3. The direction of the sailed distance is given by the course over ground,  $COG$ , from the navigation data.

$$s = V_s \cdot \Delta t \quad (3.3)$$

If all distances with their belonging course over ground are simply cumulative summed up with the initial position, the spheroidal surface of the earth is not taken into account and the route is not accurate. Therefore we use Vincenty's formulae, it is an iterative method which takes into account the shape of the earth and is even more accurate than methods that assume a spherical Earth, such as great-circle distance. The direct formulae computes the location of a point that is a given distance and direction from another

point, the method is described in Appendix C.

To add the latitude and longitude position to the navigation data, it is required to define an initial position. With this initial position the position at the next record can be calculated by the method described where the starting point of the Vincenty's formulae is the defined initial position. Thereafter an iteration is made over all the records where the input of the first iteration is the course over ground in the first record and the distance travelled between the first and second record. Thereafter, the initial of the every iteration is the end position following from the previous iteration.

This method is verified by analysing the deviation in start and end position. For a certain voyage in the data we observe a deviation of 0.007 latitude degree and 0.003 longitude degree, corresponding to 0.78 and 0.33 kilometers. This deviation is far more smaller than the defined grids for ocean wave and atmosphere parameters. Therefore, this method is accurate enough to merge the environmental conditions to the position of the ship.

Now, every record in the dataset contains an geographical position expressed in a longitude and latitude position.

### PROPULSION MODES

To make a distinction between emissions or added resistance for the different propulsion modes these modes need to be defined in the data. The different propulsion modes can be defined by the parameters from the switch data. These parameters give information about the chosen propulsion mode for the starboard and portside propulsion train, see Table 3.3. To enrich the data with propulsion modes, we need to translate and combine these switches and determine which combination represents a certain mode.

To convert the data into the propulsion modes described in Section 2.3, the switch values of the two propulsion trains need to be combined. Therefore another translation table is created, see Table 3.7. In this table, both the propulsion modes trailing and blocked shaft, are divided into two modes. One where the starboard shaft is trailing or blocked and one where the portside shaft is trailing or blocked.

Table 3.7: This translation table gives the combination of switch values that lead to a certain propulsion mode.

Propulsion Mode	Value combination of Switches			
	Cont. Mode Sw. SB	Cont. Mode Sw. PS	Prop. Mode Sw. SB	Prop. Mode Sw. PS
Transit	1	1	33	33
Manoeuvring	17	17	33	33
Blocked Shaft SB	-	-	17	33
Blocked Shaft PS	-	-	33	17
Trailing Shaft SB	-	-	65	33
Trailing Shaft PS	-	-	33	65
PEM	-	-	49	49
Stopped	-	-	97	97

### VOYAGES

To recognize voyages the stopped shaft mode is used because this mode will indicate that the shaft is physically blocked, which is the protocol when the ship is in harbour.

So, voyages are defined by the periods between the periods where both propulsion trains are in the stopped mode.

At first we calculate whether the data set is starting in or not in a voyage and whether the dataset is ending in or not in a voyage. The data that covers the period from 01/11/2018 00:00:00 until 30/11/2018 23:59:59 begins in a voyage and ends when the ship is not in a voyage. The voyages that are taken into consideration are voyages with a start and an end, so from stopped shaft mode for portside and starboard to another mode and back to the stopped shaft mode. The resulting voyages are given in Table 3.8, it can be seen that the data contains 6 complete voyages.

	Start time	End time	Time travelled
Voyage 1	05/11/2018 06:31:09	07/11/2018 10:25:18	2 days 03:54:09
Voyage 2	07/11/2018 10:55:42	09/11/2018 10:47:24	1 days 23:51:42
Voyage 3	12/11/2018 06:35:33	13/11/2018 11:51:36	1 days 05:16:03
Voyage 4	14/11/2018 05:50:39	16/11/2018 12:08:48	2 days 06:18:09
Voyage 5	19/11/2018 06:39:18	21/11/2018 14:57:42	2 days 08:18:24
Voyage 6	26/11/2018 06:19:51	30/11/2018 14:26:12	4 days 08:06:21

Table 3.8: Definition of voyages in November 2018

### PROPELLER THRUST

The propeller thrust determines the force that is given by the propeller to the water to attain a forward or backward movement of the ship. The studied OPV is not equipped with a thrust measurement device. Therefore, another approach to calculate the propeller thrust is required.

The thrust and torque produced by the propeller are expressed in terms of non-dimensional characteristics, these characteristics are unique for every propeller geometry. These non-dimensional characteristic are established in a so called open water diagram, see Figure 3.4. An open water diagram is valid for every propeller with the same geometry. The characteristics of a propeller in an open water diagram are extrapolated from open water tests in a towing tank on model scale. The procedure of open water tests is highly optimized and therefore the reliability and accuracy of the measurement achieve a high level. However, propellers can degrade during operation, they could be damaged by cavitation or floating objects or covered by marine fouling. Both will affect the performance of the propeller and so the calculation of thrust by the open water diagram is less accurate.

The characteristics present in the diagram are the thrust coefficient, Equation 3.5, and the torque coefficient, Equation 3.6, given against the advanced coefficient, Equation 3.4. Also the efficiency of the propeller is given in the open water diagram, namely the open water efficiency, Equation 3.7. This efficiency gives the relation between the delivered thrust power and the absorbed torque power. The  $K_T$ ,  $K_Q$  and  $\eta_o$  lines versus  $J$  are given for different P/D ratio's. The pitch over diameter ratio defines the pitch of the propeller in meters per revolution. The pitch "P" is the distance the propeller "screws" itself forward through the water per revolution, provided that there is no slip. In these

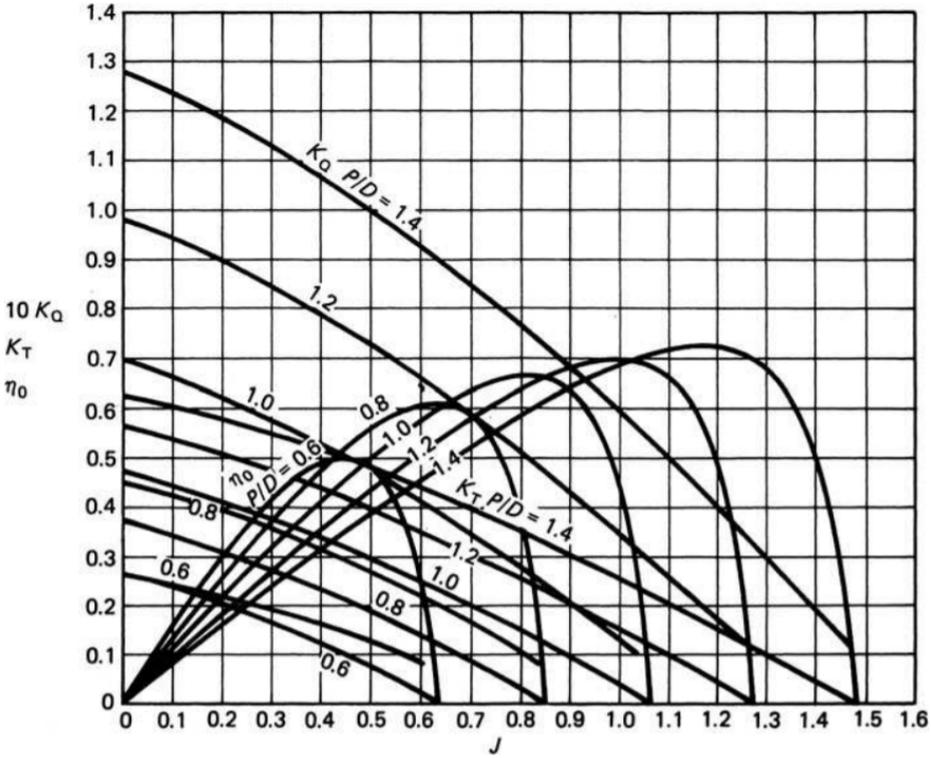


Figure 3.4: Open water diagram for Wageningen B5-75 screw series, [Oosterveld and van Oossanen \(1975\)](#)

equations,  $V_a$  is the speed at which the water is passing the propeller in m/s,  $n_p$  the rotational speed of the propeller in rotations per seconds,  $D_p$  the diameter of the propeller in m,  $T$  the thrust produced by the propeller in N and  $Q$  the torque produced by the propeller in Nm.

$$J = \frac{V_a}{n_p \cdot D_p} \quad (3.4)$$

$$K_T = \frac{T}{\rho \cdot n_p^2 \cdot D_p^4} \quad (3.5)$$

$$K_Q = \frac{Q}{\rho \cdot n_p^2 \cdot D_p^5} \quad (3.6)$$

$$\eta_o = \frac{T \cdot V_a}{2\pi \cdot n_p \cdot Q} = \frac{K_T \cdot J}{K_Q \cdot 2\pi} \quad (3.7)$$

To determine the thrust the non-dimensional characteristics are determined from the open water diagram, by calculation of the advanced ratio and the P/D ratio of the

propeller. At first, the advanced ratio  $J$  is calculated for the records in the data.  $n_p$  from the data is in rotations per minute and so this parameter is divided by 60 seconds,  $D_p$  is 3.2 meters and the propeller advanced velocity,  $V_a$ , needs to be calculated.

The advanced velocity can be calculated by taking into account the wake field of the ship, as this field is very dependent on the ship type. The wake field at the propulsor plane (propeller) arises from three principle causes: the streamline flow around the body, the growth of the boundary layer over the body and the influence of any wave-making components, [Carlton \(2007\)](#). From model tests of the OPV performed at *MARIN* in the design phase of the ship, the wake fraction dependent on the ship speed is determined already, see [Figure 3.5](#). With the wake fraction the advanced velocity at the propeller can be calculated with the ship's speed over ground, [Equation 3.8](#). Where  $V_w$  the ship's speed trough the water and so corrected for current in m/s and  $w(V_w)$  is the non-dimensional wake fraction dependent on the ship's speed through water.

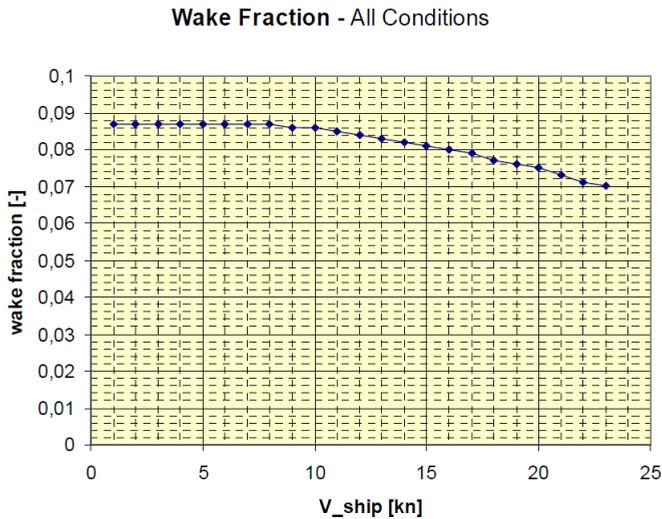


Figure 3.5: Wake fraction of the ship as a function of the ship speed in knots

$$V_a = (1 - w(V_w))V_w \quad (3.8)$$

With the advance ratio we can read up the corresponding thrust and torque coefficient for that particular advanced ratio in the open water diagram. The only thing that has to be done is to translate the P/D ratio from the data which is in percentages to the real P/D ratio of the propeller in meters per revolution. The thrust and torque coefficient lines are given for different P/D ratio in meters per revolution, so it is of importance to know which P/D ratio to look at. The translation of this ratio is according to the graph in [Figure 3.6](#).

To read up the non-dimensional coefficients of the propeller the open water diagram is extended by linear interpolation. There are created 200 lines between every known

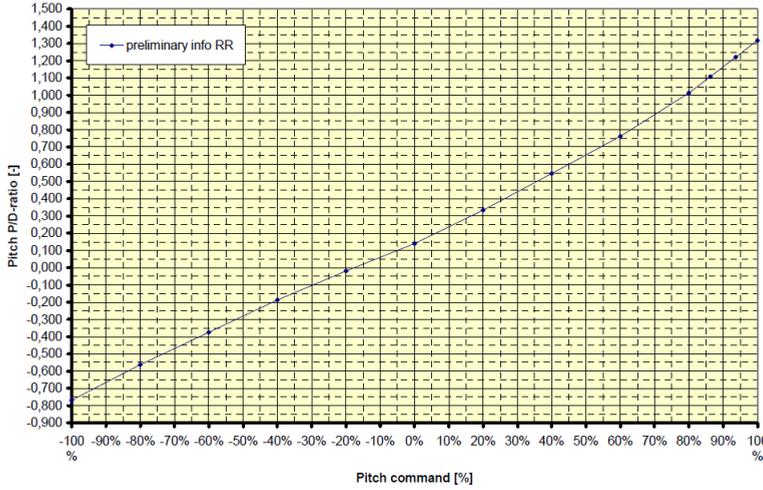


Figure 3.6: Translation of P/D command in percentages towards a P/D ratio in meters per revolution.

P/D ratio line in the open water diagram and so we created P/D ratio lines with a step size of 0.01 which makes the process of read up the coefficient more precise.

Now, the thrust and torque coefficient could be read from the open water diagram extended by the interpolated constant pitch-ratio lines at the calculated advanced ratio. With these coefficients, the rotational speed of the propeller, density of the water and the diameter of the propeller, the thrust and torque can be calculated, see Equation 3.9 and Equation 3.10.

$$Q = K_Q \cdot \rho \cdot n_p^2 \cdot D_p^5 \quad (3.9)$$

$$T = K_T \cdot \rho \cdot n_p^2 \cdot D_p^4 \quad (3.10)$$

### SET AND DRIFT

The terms set and drift are used to describe the external forces on the vessel during sailing, these forces will keep the vessel from her intended course. Set and drift are influenced by external forces, these forces include ocean currents, wind and waves. The set is in degrees and it represents the difference between the course over ground and the heading ( $\beta$ ) of the ship and the drift represents the magnitude of the oceans current, see Figure 3.7. Ignoring the effects of set and drift during the planning of navigation, there can be a navigational error of hundreds of nautical miles, dependent on the magnitude of drift and set and the length of the voyage.

Set and drift are expected to be highly determining factors for the amount of  $CO_2$  emissions. Data driven systems are not capable handling true degrees, a range of 0 to 360 degrees. These systems won't recognize that 0 and 360 are similar directions. For this reason the set and drift are translated into Cartesian coordinates, this helps making the

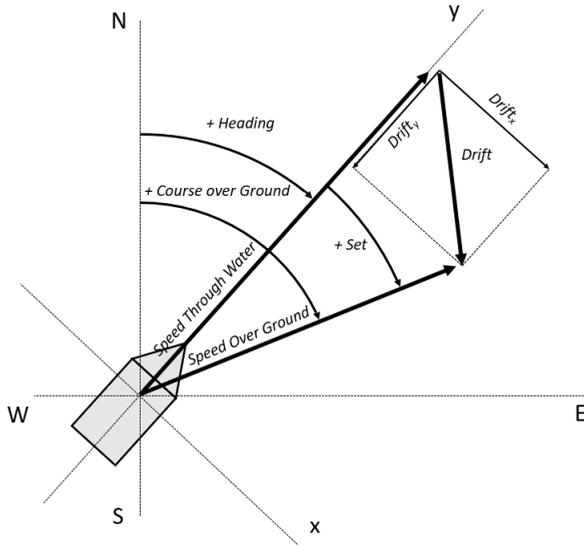


Figure 3.7: Illustration of the drift and set that can be calculated from the course over ground, heading, speed through the water and speed over ground.

algorithm more accurate, see Equation 3.11 and 3.12. The drift in  $y$  direction represents the current the ship encounters in the direction of movement. A negative  $D_y$  means that the currents is pointing into the direction of the ship's movement and so will have a positive effect on the  $CO_2$  emissions. The in  $x$  direction is the drift perpendicular to the drift in  $y$  direction.

$$D_x = V_s \cdot \sin(Set) = V_s \cdot \sin(COG - \beta) \quad (3.11)$$

$$D_y = V_w - (V_s \cdot \cos(Set)) = V_w - (SOG \cdot \cos(COG - \beta)) \quad (3.12)$$

### WAVES WITH RESPECT TO THE SHIP

The ocean waves and atmosphere parameters in the weather data, as describe in Section 3.1, are given in geographical coordinates. The wave direction is given with respect to True north (also called geodetic north) it is the direction along Earth's surface towards the geographic North Pole. The wind parameters are translated into the true wind speed and the true wind direction with respect to true north, see Equation 3.1 and 3.2. This paragraph describes the translation of the wave direction with respect to True North to the wave direction with respect to the ship.

The wave direction with respect to the heading of the ship can be calculated by making use of the heading of the ship namely, this parameter is with respect to True north also. The wave direction is calculated by subtracting the wave direction obtained from the data by the heading of the ship, see Equation 3.13 and Figure 3.8.

$$\mu = mwd - \beta \quad (3.13)$$

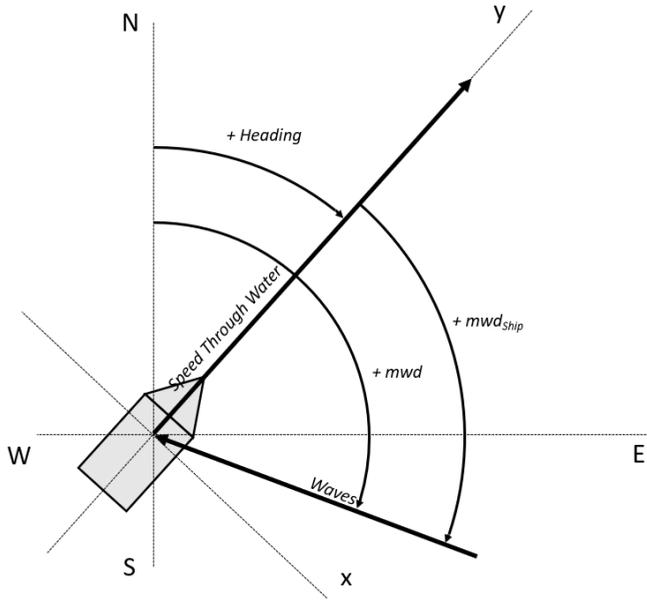


Figure 3.8: Illustration of mean wave direction from the data, the heading of the vessel and the wave direction with respect to the vessel.

### APPARENT TO TRUE WIND SPEED AND ANGLE

Even in windless weather, due to the fact that the ship is sailing, the ship has a relative speed with the air that results in a feeling of wind, [Xiong and Wang \(2019\)](#). This wind is called the apparent wind speed and angle and this parameters are measured by the sensors at the ship. The apparent wind speed and angle are dependent on the ship speed. Further in this research the true wind speed and angle, independent of the speed of the ship, are desired. Therefore the apparent wind translated into the true wind.

Figure 3.9 gives the sign conventions that are used in this thesis for the wind angles. The calculation of the apparent wind speed projected on the x and y axis of the ship can be calculated with the measured apparent wind speed,  $AWS$ , and the apparent wind angle,  $AWA$ , as shown in Equations 3.14 and 3.15. Taking into account the ship's speed over ground and the set, the true wind speed in x,  $TWS_{Ship,x}$ , and y direction,  $TWS_{Ship,y}$ , can be calculated as shown in Equations 3.16 and 3.17. With those, the true wind speed  $TWS$  can be calculated by the root of the sum of the squares of both, see Equation 3.18. The true wind angle with respect to the ship ( $TWA_{Ship}$ ) is calculated from the x and y projections of speeds as shown in Equation 3.19.

$$AWS_{Ship,x} = \sin(AWA_{Ship}) \cdot AWS \quad (3.14)$$

$$AWS_{Ship,y} = \cos(AWA_{Ship}) \cdot AWS \tag{3.15}$$

$$TWS_{Ship,x} = \sin(Set) \cdot SOG + AWS_{Ship,x} \tag{3.16}$$

$$TWS_{Ship,y} = \cos(Set) \cdot SOG - AWS_{Ship,y} \tag{3.17}$$

$$TWS = \sqrt{TWS_{Ship,x}^2 + TWS_{Ship,y}^2} \tag{3.18}$$

$$TWA_{Ship} = -\frac{\pi}{2} + \arctan\left(\frac{TWS_{Ship,y}}{TWS_{Ship,x}}\right) \quad \forall TWS_{Ship,x} > 0$$

$$TWA_{Ship} = -\frac{\pi}{2} + \arctan\left(\frac{TWS_{Ship,y}}{TWS_{Ship,x}}\right) + \pi \quad \forall TWS_{Ship,x} \leq 0 \ \& \ TWS_{Ship,y} > 0$$

$$TWA_{Ship} = -\frac{\pi}{2} + \arctan\left(\frac{TWS_{Ship,y}}{TWS_{Ship,x}}\right) - \pi \quad \forall TWS_{Ship,x} \leq 0 \ \& \ TWS_{Ship,y} < 0$$

(3.19)

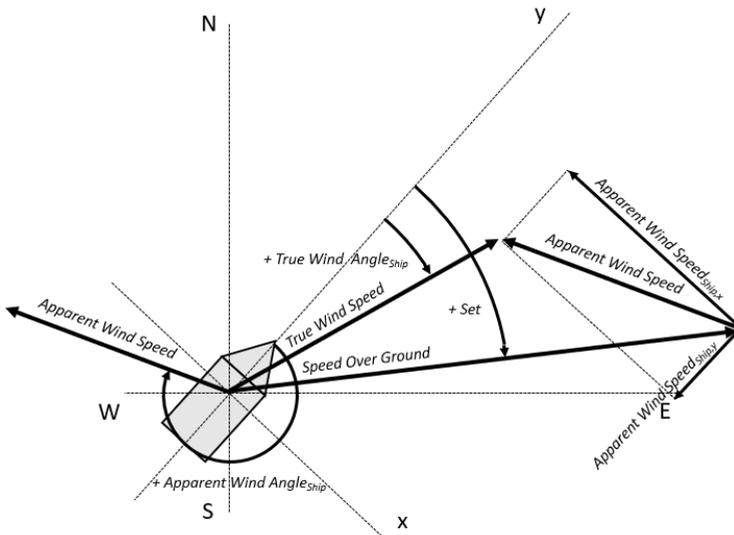


Figure 3.9: Illustration of wind direction from the data, the heading of the vessel and the wind direction with respect to the vessel.

### MEAN DRAFT

The main factors that are effecting the motions and responses of the vessel in weather circumstances are the size, shape and the lightweight and dead weight of the vessel. The size, shape and lightweight of the ship are considered as constants because of the fact that the ship has not the ability to alter these constants, the lightweight measures the weight of the ship without fuel, cargo or passengers. However, the dead weight of the ship measures the weight of the ship with her actual loading condition, in fact the displacement in any loading condition. The displacement can be calculated by multiplying the length, beam, draft and a constant called the block coefficient, approximately. The length, beam and the block coefficient are considered to be constants, however in fact they will alter with the draft of the ship.

To add the effect of loading conditions on the motions and responses of the vessel due to weather circumstances, the mean draft is expected to be influencing on  $CO_2$  emissions. This parameter will represent approximately the loading condition of the ship and so may be an important feature into the establishment of fuel consumption. The mean draft is given as in Equation 3.20, where  $T_f$  is the draft at the bow of the vessel in meters and  $T_a$  the draft at the stern of the vessel.

$$T_{mean} = \frac{T_f + T_a}{2} \quad (3.20)$$

Both  $T_f$  and  $T_a$  are influenced by waves. A peak at the location of the measurement will result in a smaller draft and a trough to a higher draft. Therefore, to implement the loading condition we will average  $T_{mean}$  over a period of time that we will determine later on in this thesis. The individual draft measurements will be used as well because they are expected to be determining factors to establish current weather conditions.

### 3.4. INTEGRATION OF DATA TYPES

THE ship's data consist, as said before, of platform, switch and navigation data. The resolution of the platform data is one record stored every three seconds chosen randomly by the monitoring system. A record includes the values of all parameters measured. All values in the platform data are stored with three decimal places.

The resolution of the switch and navigation data is varying because of the fact that they are stored on change and every hour. This means that if the value of the measured parameter changes it will be stored at the time of change and that all measured parameters will be stored every whole hour. The values measured for these data types are stored with one decimal place, therefore a parameters' value is stored at the moment that digit will change.

It is chosen to use the timestamps of the propulsion data as the timestamps of the combined dataset because of the fact that the values of the propulsion parameters are stored with a certain resolution and the values of these parameters is unknown in between the stored records. By re-sampling the navigation and switch data to data with a resolution of 1 seconds we can merge all data types and so create a combined dataset wherein all values are known at every record.

So, to merge the navigation and switch data with the timestamps of the propulsion data those two data types are re-sampled to a resolution of one second. As a result of

re-sampling, empty records will be created for the parameters. Those empty records are filled with the closest value back in time of that parameter, this can be done because of the fact that the values are stored on change.

Secondly, the timestamps in the up sampled navigation and switch are merged with their corresponding timestamp in the propulsion data, and so a combined data set is created with the original timestamps of the propulsion data. It can appear that for example a switch value changes every second. At the moment we merge the up sampled switch data with the propulsion data those changes are not included anymore because of the propulsion data resolution. However, the records that left over in the combined dataset contain the right values at that time and so this is not a major problem. Finally, the combined data set has the shape as represented in Table 3.9.

Table 3.9: Propulsion, navigation and switch data merged into one data set

Timestamp	<i>Propulsion</i>	<i>Propulsion</i>	<i>Navigation</i>	<i>Switch</i>
	<i>Parameter 1</i>	<i>Parameter 2 ...</i>	<i>Parameter 1 ...</i>	<i>Parameter 1 ...</i>
#1	Value	Values	Values	Values
#2	Value	Values	Values	Values
#3	Value	Values	Values	Values
...	....	....	....	....

To add weather conditions to every record the weather data can be used as a look up table namely, we have to compare the time, longitudinal and latitudinal position from the dataset and the weather data. When the combinations are equal, the weather parameters from the weather data belonging to this combination are added to the dataset. Because of the resolution of the weather data, the closest by point in terms of longitudinal and latitudinal position is chosen to pick the weather parameters from.

The combined dataset contains 433578 records which is in accordance with a time span of around 342 hours.

### 3.5. CONCLUSIONS

**I**N this chapter the data understanding and data preparation steps of the CRISP-DM methodology are performed. Those steps are crucial in understanding the business (how the ship is operated), describe the data (how it reflects the actual situation) and the creation of one dataset from the different data sources.

The parameters in the data sources are listed in Table 3.1 (Platform), Table 3.2 (Switch), Table 3.4 (Navigation) and the parameters in the weather data are described in Section 3.1.

All four data sources are coupled and one dataset prepared to use in this thesis is created. The combined dataset has a resolution of one record every three second. Also, every record consist of the actual value of all parameters from the data sources at the timestamp of the record. This makes it possible to compare and correlate the values with each other in every record and so to support further research in answering the re-

search questions.

To merge the weather data with the ship's data we reconstructed the latitude and the longitude position of the ship with the speed and course over ground and Vincenty's formula. The deviation in the start and end position for a certain voyage is 0.78 kilometers in latitude direction and 0.33 in longitude direction. The weather data is given on a lat-lon grid of 56 x 34 kilometers for oceans wave parameters and 28 x 17 for atmosphere parameters. Because of the high accuracy of the reconstruction of the ship's position compared to the resolution of the weather we succeed in coupling the weather data to the high resolution ship data.

The variety of environmental conditions the ship encountered derived from the weather data is given in Section 3.2. We established that the environmental conditions in the data correspond with sea state 1 till 4/5 approximately which is in accordance with the characteristics from calm to moderate conditions.

In Section 3.3 the data has been enriched by required parameters that are used in this thesis. From the switch data we derived the propulsion mode that was set by the crew and voyages made in the time span of the data are recognized.

We derived the propeller thrust from the open water diagram, the speed through water, rotational speed and pitch of the propeller. The precision of the calculated propeller thrust is dependent on the amount of marine fouling on the propeller and the technical condition of the propeller because this determines the validity of the open water diagram. This information is not available and therefore in Chapter 4 the propeller thrust will be evaluated further.

The set and drift are derived from the speed and course over ground and the heading and speed through the water. We will use these parameters further in this thesis as a measure for the actual current that the ship has encountered. The drift is split up into a component pointing in the direction of ship movement and perpendicular to that direction. Also, we translated the wind and wave direction obtained from the weather data to the direction with respect to the ship's coordinate system.

The accuracy of the derivation and calculation made in this chapter depend on the accuracy of the data. The accuracy of the data is determined by the accuracy of the measurements from the sensors on board of the ship. However, the accuracy of the sensors is not available and so the accuracy of calculations remains unknown.

# 4

## CURRENT OPERATIONAL PERFORMANCE

*Reducing CO<sub>2</sub> emissions by giving the crew insight into the energy efficiency of the ship so that they are able to make well informed decisions in CO<sub>2</sub> dilemma's, it is important to define the current operational performance.*

*Therefore, investigation is performed towards how the ship is used by the crew currently. Because of the lack of direct CO<sub>2</sub> emission measurement we should develop an alternative approach to establish those emissions. The CO<sub>2</sub> emissions are origination from multiple components, therefore the efficiency of those components is evaluated.*

*This chapter is organised as follows : In Section 4.1 the operating profile of the ship is visualised by the speed over ground distribution and the use of the propulsion modes. Section 4.2 the total resistance curve for the ship is established. The way how energy efficiency is defined in this thesis is covered in Section 4.3. Finally, Section 4.4 gives an overview of the different components in the propulsion train and quantifies the efficiency of the components.*

### 4.1. OPERATIONAL PROFILE

**T**O create an image about the how the ship is operated in the time frame that the data covers a look into the speed distribution, speed settings and use of propulsion modes is given in this section.

#### SPEED DISTRIBUTION

The ship's real speed over ground distribution from the dataset compared to the designed distribution is given in Figure 4.1. The most obvious differences between the design and real distribution is that the ship was designed to operate 30 % of operating time within a speed between 15 and 20 knots, however less than 10 % of time is operated between this speed. Also, obvious is the differences in the speed between 5 and 10 knots where the designed operating time is 10 % and during real operation this is around 30 %.

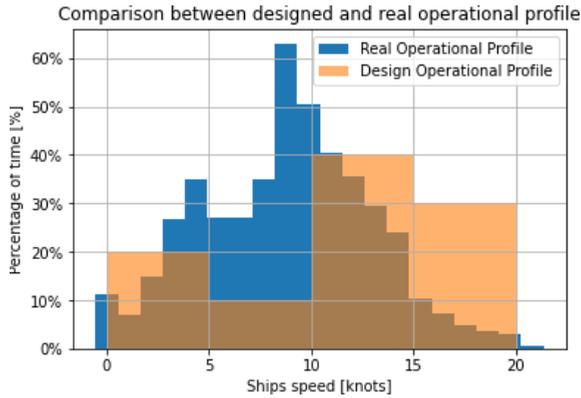


Figure 4.1: Illustration of wind direction from the data, the heading of the vessel and the wind direction with respect to the vessel.

4

## PROPULSION MODES

The ship has the ability to sail with different propulsion modes. Figure 4.2 gives an overview of the relative use of the those propulsion modes during the time frame that the dataset covers. It can be seen that the manoeuvring propulsion is chosen the greatest amount of time. During 146549 from the 433578 records the ship has sailed in the manoeuvring propulsion modes, this is in accordance with a total time of around 116 hours.

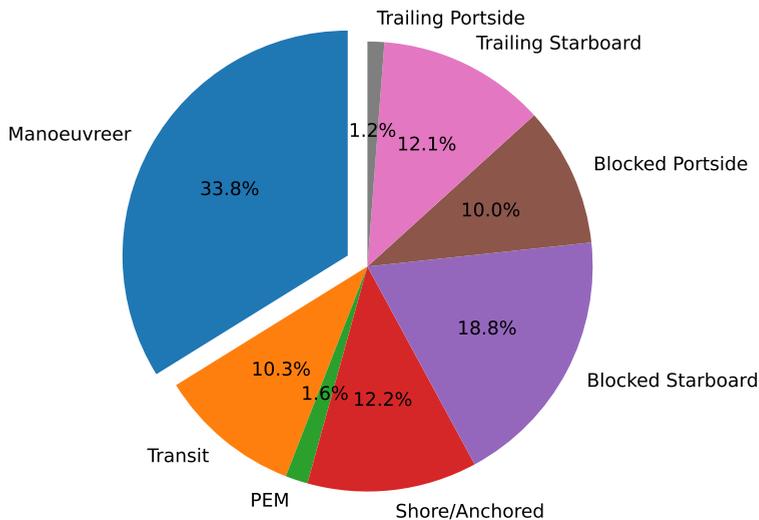


Figure 4.2: Pie chart to visualise the distribution of the usage of propulsion modes. It could be seen that the manoeuvring propulsion mode is used the most in the dataset, 33.8 % of the time span of the dataset.

## SPEED SETTINGS

The speed of the ship is determined by the virtual shaft speed setting by the crew. The virtual is there because it has, in static conditions, a linear relationship with the speed over ground and therefore to use intuitively by the crew. In Figure 4.3 a histogram is given for the virtual shaft speed settings by the crew for the starboard and portside propulsion train.

The peaks in the figure represent the predefined virtual shaft speed settings. Those predefined shaft speed settings are in accordance with codes on the telegraph. The telegraph code 'full speed ahead' is for example in accordance with a virtual shaft speed setting of 230 rotation per minute. Due to the predefined shaft speed settings those settings are used for a major part as speed setting by the crew. Therefore, the dataset is unbalanced which means that there are areas with a high density of measurement points (at the predefined settings). Later in this thesis we have to take into account the unbalanced dataset.

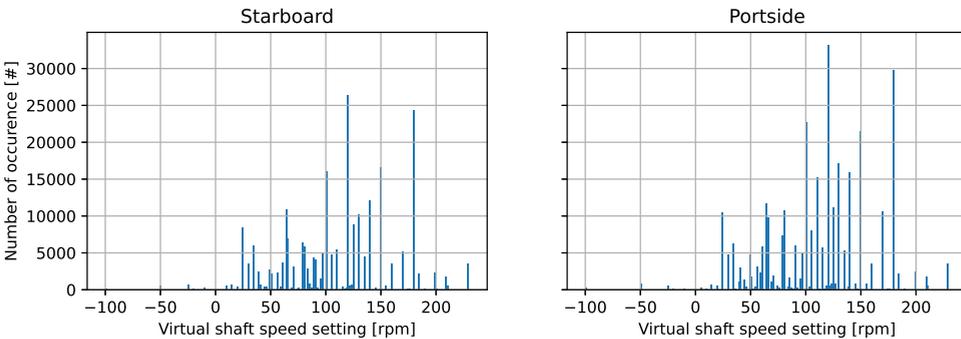


Figure 4.3: Histogram of the virtual shaft speed setting of the starboard and portside propulsion train. The high peaks represent the predefined shaft speed settings in accordance with the telegraph codes, e.g. full speed ahead).

## 4.2. SHIP'S RESISTANCE

**I**N this section the resistance curve of the ship is derived from the thrust calculated in Section 3.3. Also, the resistance curve is compared to resistance curves of the ship in different conditions obtained from towing tank tests performed by the 'Maritime Research Institute Netherlands' (MARIN). The thrust can give insight into the resistance of the ship and the performance of the ship.

The towing tank tests are performed by the 'Maritime Research Institute Netherlands' (MARIN) for three different sea states, sea states 0, 4 and 6. Sea state zero contains no correction for sea state and fouling only an additional resistance due to wind speeds of 5 knots maximal. Sea states 4 and 6 contain a correction for fouling corresponding to 6 months out of dock, additional resistance due to wind speeds of 21 and 47 knots respectively and a resistance correction for sea state 4 and 6 respectively.

The thrust force is not directing in the negative direction of the ship's movement only. Therefore the resistance is corrected by the thrust deduction factor. This factor describes

the suction of the propeller and can be used to calculate the resistance of the ship, the thrust delivered by the propeller is higher than the resistance of the ship usually. In Figure 4.4 the thrust deduction factor is given as a function of the ship speed over ground in knots obtained from towing tank tests. With the thrust calculated in 3.3 the total resistance of the ship can be calculated, see Equation 4.1.  $F_{acc}$  is the force due to acceleration in N and this force is determined later.

$$R_{tot} = (T_{SB} + T_{PS}) \cdot (1 - t) - F_{acc} \quad (4.1)$$

**Trust Deduction - All Conditions**

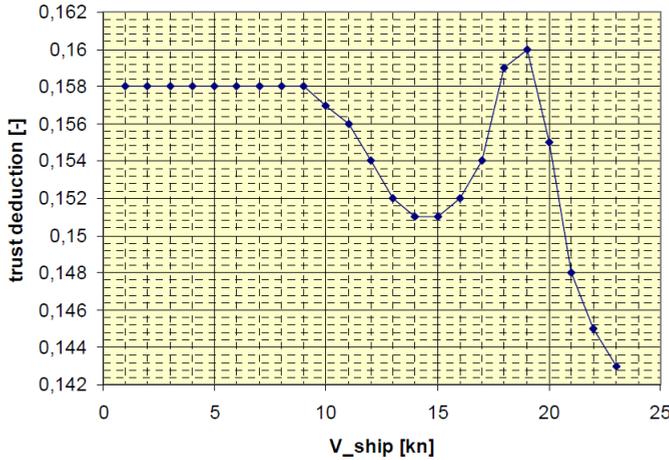


Figure 4.4: Thrust deduction factor of the ship as a function of the ship speed in knots

### SHIP'S ACCELERATIONS

The added resistance due to accelerations is approximated by *Newton's Second Law of Motion*. This law of motion can be formally stated as follows: The acceleration of an object as produced by a net force is directly proportional to the magnitude of the net force, in the same direction as the net force, and inversely proportional to the mass of the object,  $F = m \cdot a$ .

The acceleration of the vessel is subtracted from the speed through water difference over time as formulated in Equation 4.2. Where  $dV_w$  is the difference in vessel speed through between sequential timestamps and  $dt$  is the difference in seconds between these two timestamps.

$$a(t) = \frac{dV_w}{dt} = \frac{V_{w,t+1} - V_{w,t}}{t_{t+1} - t_t} \quad (4.2)$$

Due to the combination of the data sources in Section 3.4 and that the navigation data is stored with one decimal place it could be that at certain records the acceleration is zero. This acceleration is zero because there is no difference in the speed through

water between the records. As a result, the signal for the acceleration contains a lot of peaks, see Figure 4.5. Therefore we use a rolling average over the raw acceleration values of 18 seconds, 6 records, as well as over the speed through water. The high peak signal is flattened resulting in a more continuous acceleration, see *Continuous Acceleration* in Figure 4.5.

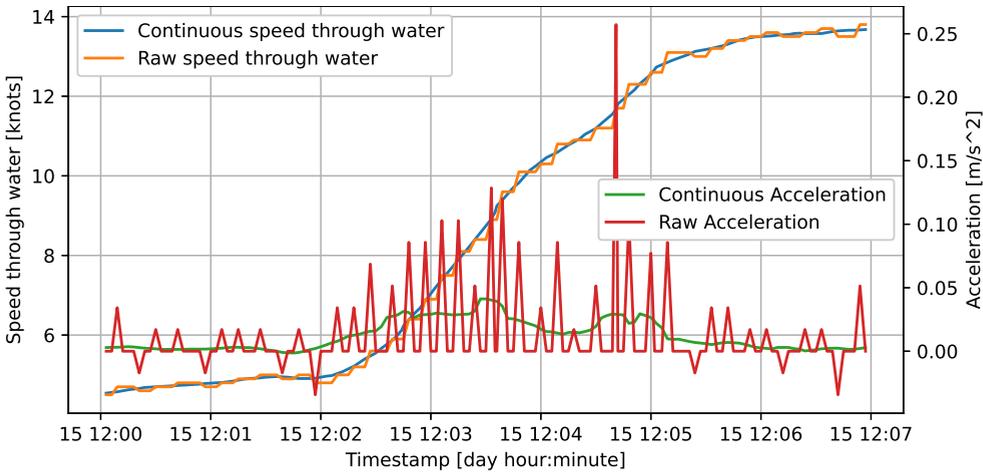


Figure 4.5: Visualisation of the raw and averaged acceleration and the speed through water during an acceleration. The speed through water is on the left axis and the acceleration on the right axis.

The mass of the ship is around the 3800 tons which can vary for the different loading conditions, however due to the lack of information we will not take the variability in mass into account. Because of the fact that the variability in mass is between approximately  $\pm 2\%$  a constant mass is assumed.

## TOTAL RESISTANCE

In Figure 4.6 both the resistance, scaled from towing tests and determined from the data, is given against the speed through water of the ship.

The thrust is calculated by the open water diagram of the propeller as described before. The results obtained could be discussed because of the fact that the open water diagram gives the characteristics of an ideal propeller without damage or fouling. From the figure it can be seen that most of the resistance points calculated from the data are in between the resistance curves obtained by model tests. Further investigation towards the degradation of the propeller compared to the open water diagram could be done. However, the degradation process considering marine fouling seems to take more time than only one month of data. Therefore, further investigation towards the degradation requires more data, for a longer time period. For this thesis the proposed resistance calculation is assumed to be sufficient.

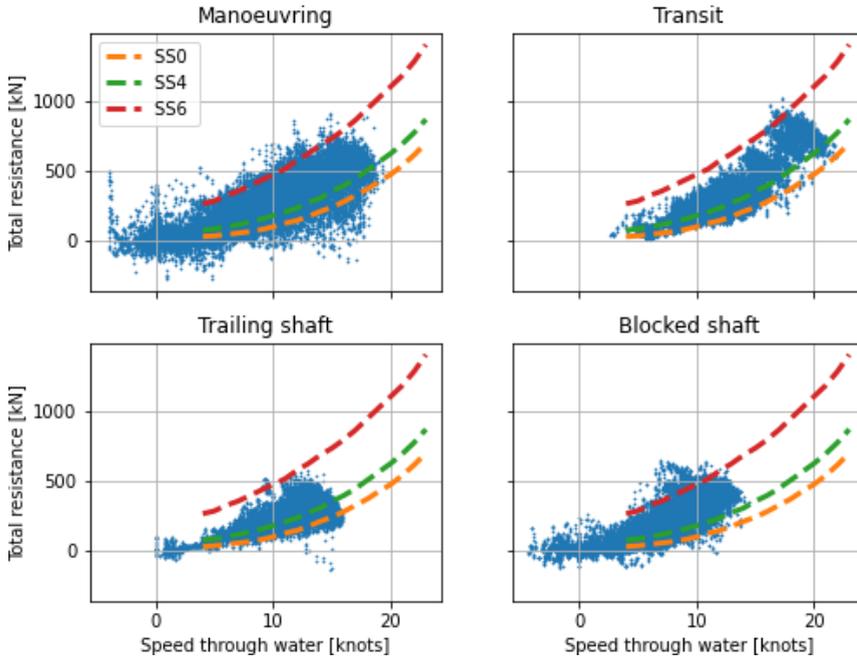


Figure 4.6: Total resistance of the ship calculated from the determined thrust of the propeller. The resistance is given for the manoeuvring, transit, trailing and blocked shaft mode.

### 4.3. FUEL CONSUMPTION

THE  $CO_2$  emission are not measured directly in the exhaust pipe of the ship and therefore we will derive these emissions from the fuel consumption. In this derivation the assumption is made that complete combustion occurs. We will make use of the fuel consumption of all fuel consuming components on board. The fuel consuming components on board of the OPV are the two main diesel engines for the delivery of mechanical power related to the propulsion and the three diesel generator sets for the delivery of electrical power related to the required hotel load and the electrical propulsion. All power suppliers are equipped with fuel flow sensors measuring the fuel flow in liter per hour for the diesel generator set and in kilogram per hour for the main diesel engines.

F-76 is the type of fuel that is consumed by the fuel consumers on board of the OPV. The lower heating value of the fuel F-76 is 42580 kJ/kg, [Tol and Linden \(2018\)](#), a density of  $847.4 \text{ kg/m}^3$  and a carbon content of 86.6%. To calculate the kilogram of  $CO_2$  emitted per kilogram of fuel consumed, assuming complete combustion, the carbon content of the fuel is multiplied by the ratio of molecular weight of  $CO_2$  (44) to the molecular weight of C (12), see Equation 4.3. Herein,  $CO_{2,fuel}$  is the kilogram of  $CO_2$  per kilogram fuel,  $C_{fuel}$  the carbon content in the fuel,  $M_{CO_2}$  the molecular weight of  $CO_2$  and  $M_C$  the molecular weight of carbon.

$$CO_{2,fuel} = C_{fuel} \cdot \frac{M_{CO_2}}{M_C} = 0.866 \cdot \frac{44}{12} \approx 3.175 \left[ \frac{\text{kg}_{CO_2}}{\text{kg}_{fuel}} \right] \quad (4.3)$$

### DIESEL GENERATOR SETS

The fuel consumption of the diesel generator sets are measured by fuel flow sensors and the load of the engine is measured as well. The generator sets operate at nearly constant frequency and so the load of the diesel engine is of interest. The load of the diesel engine is the ratio between the actual power output and the nominal power output of the diesel engine. The choice of the amount of diesel generator sets running is depending on the demand for electrical energy determined by the hotel load and load in case of electric propulsion.

From the data the conclusion can be made that approximately half of the time 2 diesel generators set are running together and half of the time 1 diesel generator is running. Figure 4.7 shows the distribution of the load of the diesel generators. The two peaks around 33 % and 51 % can be explained by the amount of diesel generators running. At the moment the diesel generator sets are loaded by 33 %, 2 sets are supplying electrical power and when loaded by 51 %, 1 set.

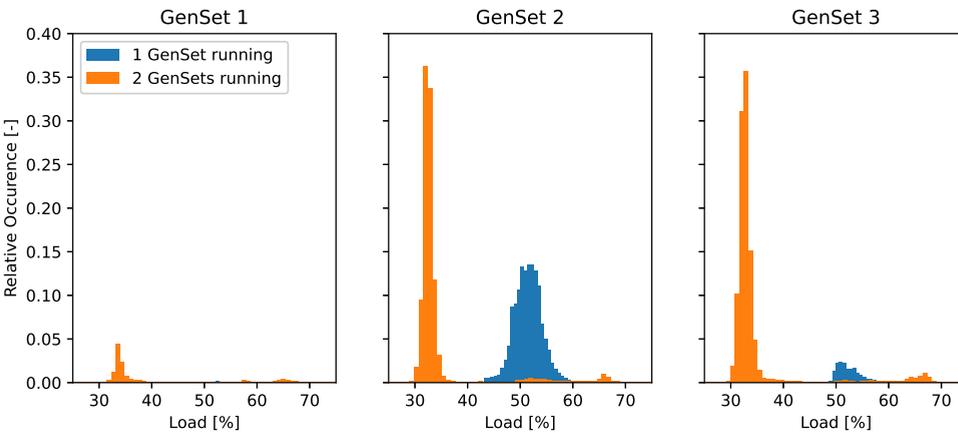


Figure 4.7: Distribution of the load of the 3 installed diesel generator sets. There is made a distinction between the number of diesel generator sets that are running.

The varying load between 33 % and 70 % is resulting from the period where the ship is electrical driven, in this propulsion mode two of the three diesel generators are providing electrical energy for the electrical energy hotel and propulsion consumption. It is not very clear from the figure, because of the little amount of records in the data where the ship is electrical driven.

The specific fuel consumption of the diesel engine part of the set is gathered from test data of the manufacturer. These tests consist of five operating points, in fact five measurements of the specific fuel consumption for different loads. These test results are given in Figure 4.8, it can be seen that the higher loading results in a lower specific fuel consumption.

To make the supply of electrical energy more energy efficient, the diesel generators have to be loaded as high as possible. To investigate a possible solution we have to explore the total electrical energy demand in different situations. Figure 4.9 shows the electrical energy supply of the diesel generators from the data. It can be seen that there

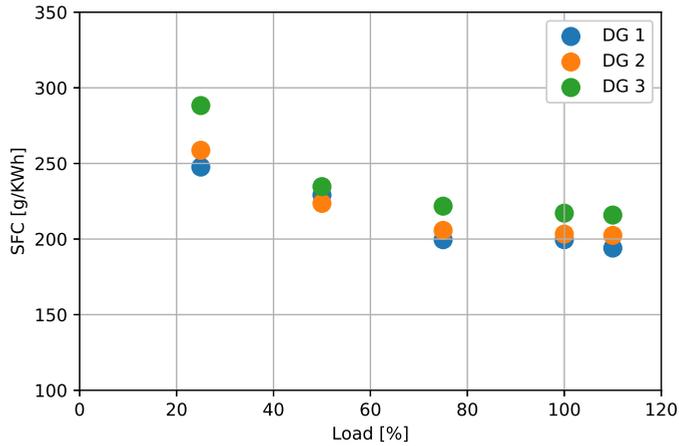


Figure 4.8: Specific fuel consumption of the diesel generator sets against the load, the data is obtained from manufacturer data.

are two different hotel loads, one moment the hotel load is around 475 kW and one moment around 600 kW. The second type of load is resulting from the electrical propulsion and the hotel load combined. It can be seen that this load is varying between the 600 and 1250 kW. A simple solution to decrease the fuel consumption of the diesel generators could be to supply electrical energy by one diesel generator set only when no electrical propulsion. Namely, the diesel generator sets have a nominal power output of 920 kW each which is by far enough to supply energy for the two types of hotel load.

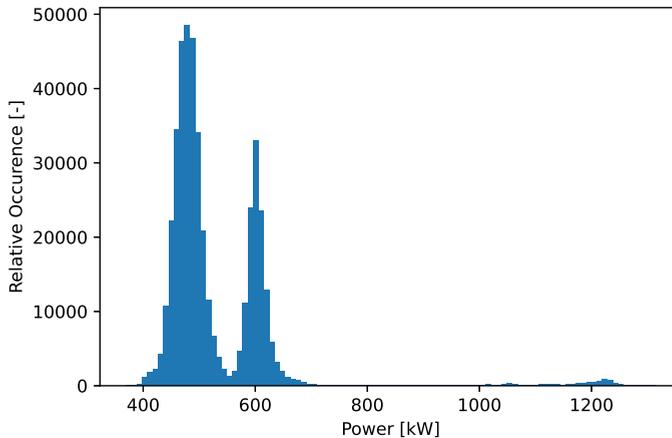


Figure 4.9: Histogram of the total power supply by the diesel generator sets, obvious are the two peaks of power supply.

### MAIN DIESEL ENGINES

The fuel flow measurements of the main diesel engines are components of the bigger fuel system and measuring the flow before the fuel enters the cylinder. Therefore, to measure the fuel consumption with higher accuracy we will make use of the fuel rack position of the diesel engine which determines the amount of fuel entering the cylinder directly in front of the cylinder and so more give a more accurate result.

The fuel rack position is given in percentages, therefore we will correlate the fuel rack position with the fuel flow measurements. Figure 4.10 shows relation between measurement of the fuel flow sensors of both diesel engines,  $\dot{m}_f$ , and the fuel rack position of both engines,  $X$ .

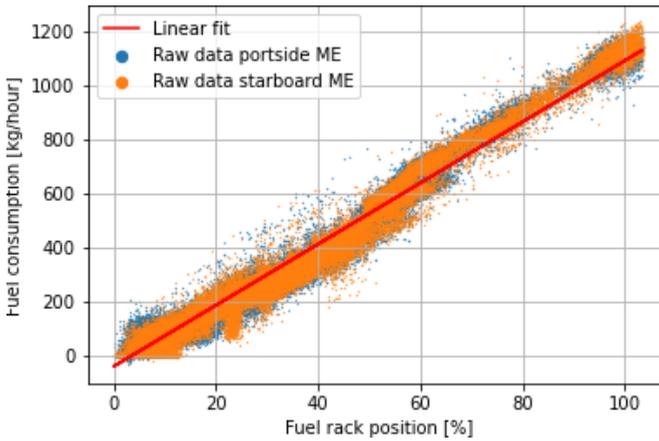


Figure 4.10: Relation between the fuel consumption of the main diesel engines in kg/hour and the fuel rack position in percentage of the maximum position.

The relationship between the fuel consumption as measured by the fuel flow sensors shows a highly linear relationship with the fuel rack position. Therefore, the coefficients  $a$  and  $b$  in the curve  $y = ax + b$  are derived from the raw data points. The optimized values for the parameters  $a$  and  $b$  are 11.51 and -48.87 respectively, see Equation 4.4

$$\dot{m}_{f,fit} = a \cdot x + b = 11.51 \cdot X - 48.87 \quad (4.4)$$

### FUEL CONSUMPTION PER MILE

Fuel consumption curves give the fuel consumption per mile (FPM) presented against the speed over ground. These curves could give insight into the an optimal ship speed regarding minimal fuel consumption. On board the OPV those curves are used to set up a navigation plan at the beginning of an operation. Specifically, the fuel consumption curves are used to calculate the required amount of bunker fuel given the route and the speed along this route. Geertsma et al. (2017) already established the fuel consumption curves from a complete ship model of the OPV for trial, design and off-design conditions, see Figure 4.11.

In this thesis we will calculate the fuel consumed per mile by Equation 4.5, where  $X$

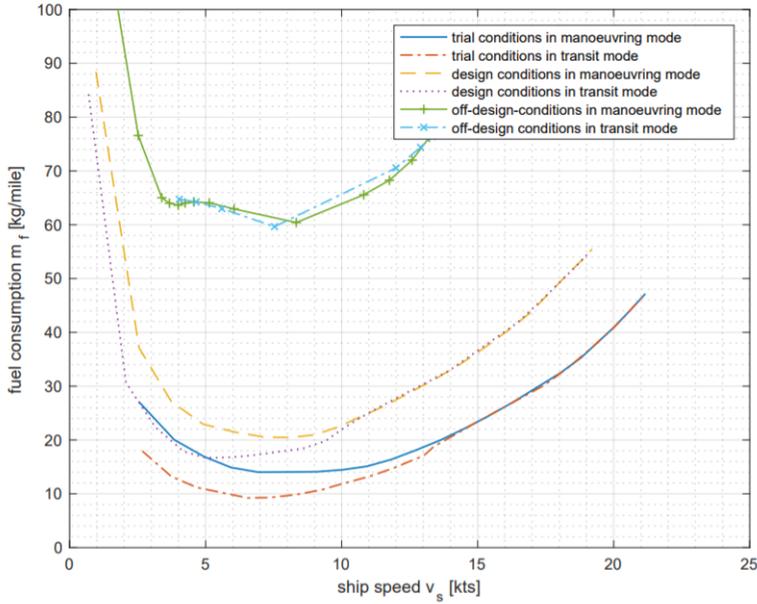


Figure 4.11: Fuel consumption plot as a function of ships speed for various conditions in manoeuvring and transit mode, [Geertsma et al. \(2017\)](#)

is the fuel rack position of the main diesel engines,  $V_s$  is the speed over ground in knots and the  $FPM$  is the fuel consumed per mile in kilogram per mile.

$$FPM = \frac{\dot{m}_f}{V_s} = \frac{11.51 \cdot X - 48.87}{V_s} \quad (4.5)$$

The FPM from ship's data shows high variety, this variety can be explained by e.g. the differences in weather conditions, accelerations and decelerations. To ignore the effect of accelerations and decelerations and only consider the influence of weather conditions. Accelerations and decelerations can be filtered out by determining the points where the speed setpoint of the vessel is changed.

To filter the continuous operation of the vessel, the moments of a change in the speed setpoint are determined. This is done by calculating the standard deviation over a predefined window/number of consecutive observations ( $N$ ) in the data, see Equation 4.6 where  $\bar{x}$  is the average over  $N$  observations and  $x_i$  the  $i$ th sample of  $N$  observations. The standard deviation measures the amount of variation or dispersion in the values, in this case the standard deviation is taken over the RPM-settings.

The standard deviation over a predefined window/number of consecutive observations ( $N$ ) is calculated at every timestamp, the number of consecutive observations is chosen to be 270 seconds. The moment that the standard deviation changes from 0 to a value higher than zero indicates that there is a change in speed setpoint. It is chosen to leave out the data within 180 seconds after a change of speed setpoint. Figure 4.12 shows

that this time window leaves out the effect of the decelerations or accelerations almost completely.

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \tag{4.6}$$

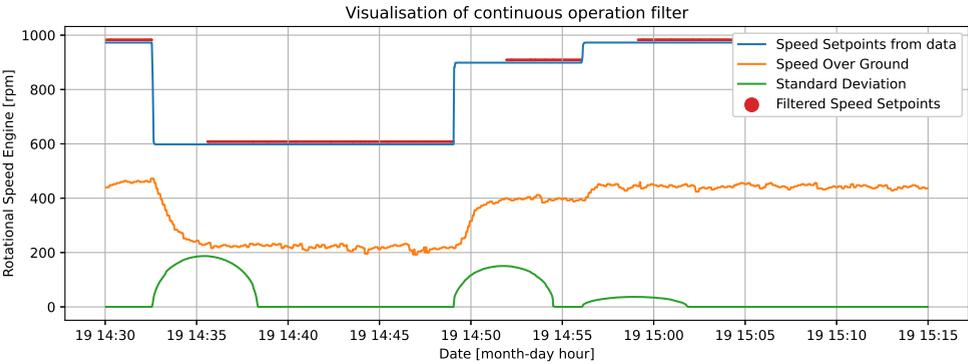


Figure 4.12: Visualisation of filtered data by continuous speed setting versus the raw data

Figure 4.13 gives the fuel consumed per mile, left out accelerations and decelerations, for the different propulsion modes against the speed over ground. It can be seen that for lower speeds over ground the fuel per mile increases significantly for the manoeuvring and blocked shaft mode. Also, for higher speeds over ground the fuel per mile increases.

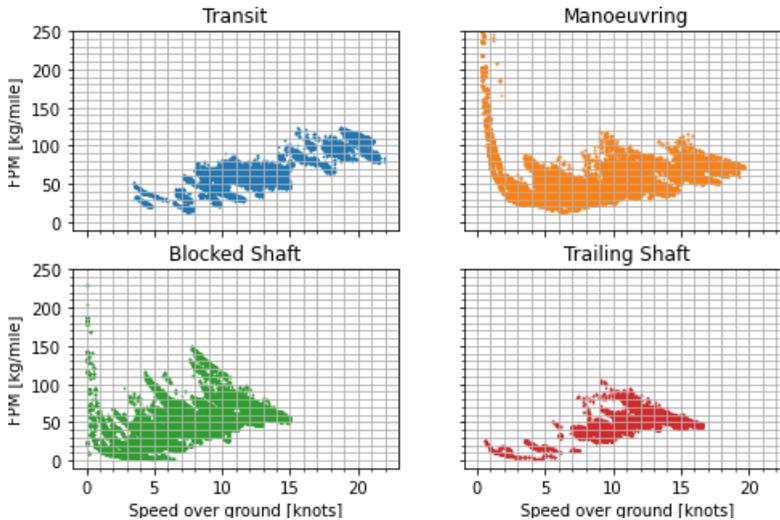


Figure 4.13: Fuel consumption per mile for the transit, manoeuvring, blocked and trailing shaft mode. The points visible in the figure are filtered for acceleration and decelerations.

#### 4.4. EFFICIENCY BREAKDOWN

IN the previous section we calculated the resistance of the ship and we have seen that the fuel consumption per mile increases significantly for low speeds. In this section we will evaluate the different efficiencies of the components in the propulsion train to gather insight into the influences of the components on the fuel per mile.

The thrust of the propeller is calculated by the open water diagram belonging to the propeller, the rotational speed, P/D ratio and advanced velocity of water. The power required to tow the ship at a certain ship speed  $V_s$  with resistance  $R$  is the effective power  $P_E$ , see Equation 4.7, Woud and Staperma (2002). This calculation of the effective power is valid in situations of calm water (no current), however in this thesis we will take into account the effect of current and therefore use  $V_w$  instead of  $V_s$ .

$$P_E = R \cdot V_s \quad (4.7)$$

The power as delivered by the propeller in water moving at velocity of advance  $V_a$  with the calculated thrust is the thrust power  $P_T$ , see Equation 4.8. The ratio between the effective power and the thrust power is called the hull efficiency,  $\eta_H$ . Due to the wake fraction and thrust deduction factor those two powers are not equal. The hull efficiency often is higher than unity because of the fact that the propellers have to deliver more thrust than the resistance of the ship and the velocity of advance,  $V_a$ , is lower than the ship's speed through water. Equation 4.9 determines the hull efficiency of the ship as a function of the speed through water.

$$P_T = T \cdot V_a = T \cdot V_w(1 - w(V_w)) \quad (4.8)$$

$$\eta_H = \frac{P_E}{P_{T,SB} + P_{T,PS}} = \frac{R \cdot V_w}{(T_{SB} + T_{PS}) \cdot V_w(1 - w(V_w))} \quad (4.9)$$

Together with Equation 4.1 it can be derived to Equation 4.10.

$$\eta_H = \frac{(T_{SB} + T_{PS}) \cdot (1 - t(V_w)) \cdot V_w}{(T_{SB} + T_{PS}) \cdot V_w(1 - w(V_w))} = \frac{(1 - t(V_w))}{(1 - w(V_w))} \quad (4.10)$$

#### OPEN WATER EFFICIENCY

Power is delivered to the propeller as torque,  $Q$ , and rotational speed,  $n_p$ . The open water power is defined as in Equation 4.11. Open water refers to the fact that propeller normally are tested in an open water tank or tunnel, the flow in front of the propeller is uniform. Because of that the flow in open seas is not uniform in front of the propeller the open water propeller efficiency is defined, the ratio between the thrust power and the open water power, see Equation 4.12

$$P_O = Q \cdot n_p \cdot 2\pi \quad (4.11)$$

$$\eta_O = \frac{P_T}{P_O} = \frac{T \cdot V_w(1 - w(V_w))}{Q \cdot n_p \cdot 2\pi} \quad (4.12)$$

### SHAFT EFFICIENCY

The torque in between the gearbox and the propeller is measured by a torque sensor and available in the data,  $M_s$ . The location on the shaft where the torque is measured is not known, however the most obvious location is at the output of the gearbox. And therefore we assume that frictional losses arising from the bearings can be determined by the ratio between the open water power,  $P_O$ , and the actual delivered power to the shaft,  $P_s$ , as shown in Equation 4.13. This ratio is called the shaft efficiency and given in Equation 4.14.

$$P_s = M_s \cdot n_p \cdot 2\pi \quad (4.13)$$

$$\eta_s = \frac{P_O}{P_s} = \frac{Q \cdot n_p \cdot 2\pi}{M_s \cdot n_p \cdot 2\pi} \quad (4.14)$$

The ratio between the thrust power,  $P_T$ , as delivered by the propeller and the shaft power,  $P_s$ , is called the propulsive efficiency,  $\eta_{prop}$ , see Equation 4.15.

$$\eta_{prop} = \eta_o \cdot \eta_s \quad (4.15)$$

### GEARBOX EFFICIENCY

The gearbox efficiency is defined by ratio between the power input and output of the gearbox. Because of that the power at the output of the gearbox is known only the gearbox efficiency will be determined. de Waard (2014) has fitted a fourth order polynomial on the torque losses of the gearbox based on information provided by the manufacturer. This function is dependent on the rotational speed of the gearbox and the nominal torque losses, see Equation 4.16.

$$\frac{M_{loss}}{M_{loss,nom}} = -1.1655 \left( \frac{n}{n_{nom}} \right)^4 + 2.8613 \left( \frac{n}{n_{nom}} \right)^3 - 2.6142 \left( \frac{n}{n_{nom}} \right)^2 + 1.9116 \left( \frac{n}{n_{nom}} \right) + 0.0048 \quad (4.16)$$

The nominal power loss of the gearbox is obtained from the information provided by the manufacturer also. The nominal torque loss is calculated by dividing the nominal power loss by the nominal speed and has a value of approximately 4.53 kNm. The gearbox efficiency can be calculated by the ratio between the torque at the output of the gearbox and the torque at the input. The torque at the output is known,  $M_s$ , and the torque at the input is modelled and gives as  $M_s + M_{loss}$ . This makes it possible to calculate the gearbox efficiency as in Equation 4.17.

$$\eta_{GB} = \frac{M_s}{M_s + M_{loss}} \quad (4.17)$$

With these efficiency the power losses of the gearbox are known and so the power as demanded from the diesel engine can be calculated. The demanded power from the diesel engine,  $P_B$ , is the delivered power to the shaft,  $P_s$ , plus the power losses in the gearbox. The brake power,  $P_B$ , is calculated as in Equation 4.18.

$$P_B = \frac{P_s}{\eta_{GB}} \quad (4.18)$$

### DIESEL ENGINE EFFICIENCY

In Section 4.3 the fuel consumption of the diesel engine has been determined. The diesel engine efficiency,  $\eta_{eng}$ , is the ratio of engine output power and the heat input. The heat input of the engine can be calculated by the fuel consumption times the lower heating value (LHV) of the fuel. The lower heating value of the fuel F-76 is 42580 kJ/kg, Tol and Linden (2018). Because of the engine power output is given by  $P_B$  the engine efficiency is given as in Equation 4.19.

$$\eta_{eng} = \frac{P_B}{LHV \cdot \dot{m}_f} \quad (4.19)$$

### TOTAL EFFICIENCY

The efficiency for the starboard and portside propulsion train from the heat input of the diesel engine to the thrust power of the propeller includes the propulsive, gearbox and diesel engine efficiency. Figure 4.14 shows the individual efficiency's including the total efficiency for the propulsion trains determined as in Equation 4.20 for the manoeuvring propulsion mode.

It can be seen that the efficiency in both the propulsion train are almost equal.

$$\eta_{tot,train} = \eta_{prop} \cdot \eta_{gb} \cdot \eta_{eng} \quad (4.20)$$

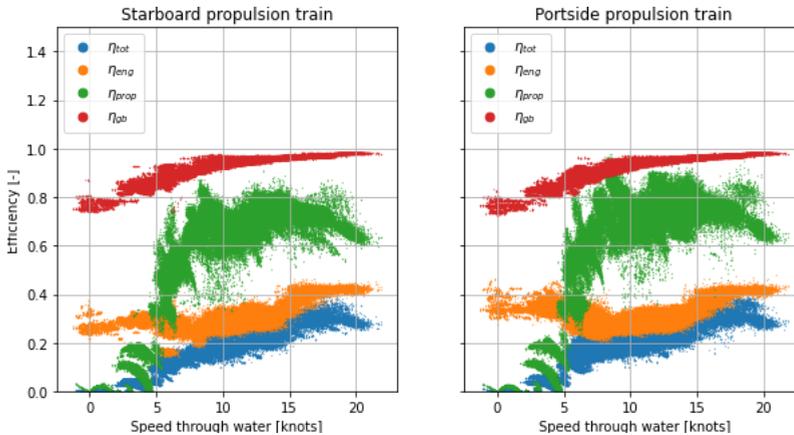


Figure 4.14: Visualisation of the efficiencies for the starboard and portside propulsion train during the manoeuvring propulsion mode versus the speed through water. The propulsive, gearbox and engine efficiency can be seen.

The total efficiency of the ship, see Figure 4.15, includes both propulsion trains, this efficiency gives the ratio between the effective towing power,  $P_E$ , and the heat input of

both diesel engines, see Equation 4.21.

Obvious is that the total efficiency drastically drops at lower speeds. Vasilikis (2020) defines the most energy effective and environmental friendly option is to sail at the lowest possible speed down to a minimum, though it is important to avoid a lower speed than this as energy effectiveness drops drastically. This statement is supported by the research done by Geertsma et al. (2017), see Figure 4.11 and Figure 4.13 where a drastic increase is visible in the fuel consumption per mile for lower speeds.

$$\eta_{tot} = \frac{P_E}{LHV \cdot (\dot{m}_{f,SB} + \dot{m}_{f,PS})} \quad (4.21)$$

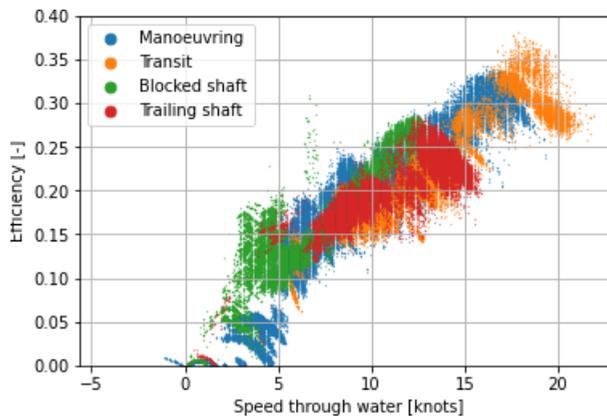


Figure 4.15: Visualisation of total efficiency for the propulsion modes versus the speed through water.

## 4.5. DIESEL ENGINE LOADING

The diesel engine loading is determined by the power demand of the propeller added with power losses in the gearbox and shaft bearings in between the propeller and the diesel engine. The power demanded of the propeller is function of the rotational speed, the power demand versus rotational speed line is called the propeller load curve. The total power demand from the diesel engine is equal to the brake power,  $P_B$ , see Equation 4.18.

The torque on the propeller is dependent on the combination of pitch and rotational speed of the propeller, established by a unique combinator curve for every propulsion mode.

In Figure 2.3 the power envelope of the installed diesel engines is represented. The power envelope of the diesel engine defines the amount of power the diesel engine can deliver at a certain rotational speed. Figure 4.16 gives the diesel engine loading for the propulsion modes on both the starboard and portside diesel engine. The different combinator curves are visible by the differences in propeller load curves.

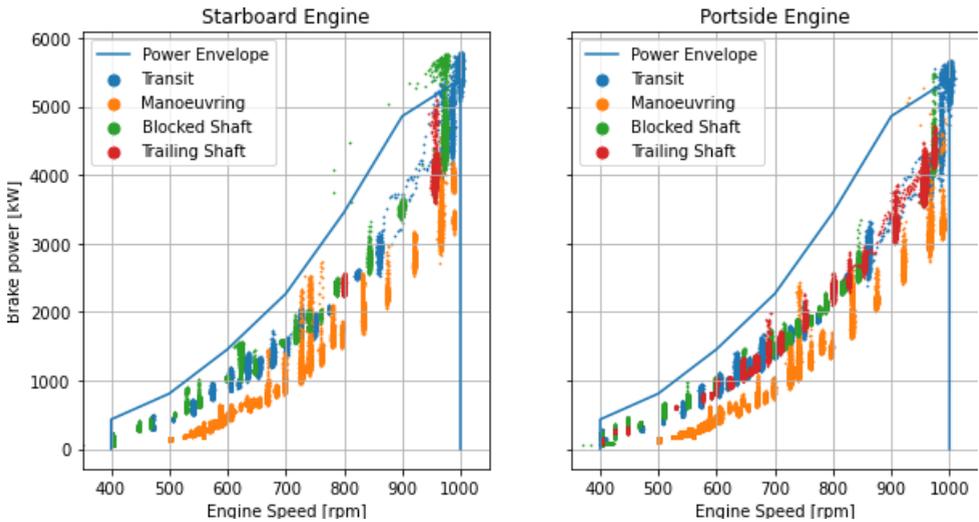


Figure 4.16: Visualisation of the diesel engine loading versus the rotational speed for the propulsion modes and filtered for accelerations and decelerations.

## 4.6. CONCLUSIONS

IN Chapter 3 the combined dataset is created with a resolution of one measurement every three seconds and enriched with various parameters. In this chapter the current operational performance of the ship are investigated by evaluating the operational profile during the time span of the combined dataset. With the research performed we can draw the following conclusions and answer Research Question 1:

*How can  $CO_2$  performance be established given the operational profile of a naval vessel with data?*

The  $CO_2$  emissions arising from the propulsion of the ship can be established by the fuel consumption of the main diesel engines when sailing in the propulsion modes were one of or both the main diesel engines drive the ship. When the ship is driven by the electric engines, the  $CO_2$  emissions from propulsion of the ship can be derived from the fuel consumption of the diesel generators taking into account the fuel consumption to satisfy the hotel load of the ship. However, because of the lack of sufficient records in the data for the PEM propulsion mode we left this out of consideration in this thesis.

The consumed fuel per main diesel engine is measured by a flow sensor, however this sensor is less accurate than the fuel rack position. The fuel rack position is directly measuring the amount of fuel that flows into the cylinders of the main diesel engine were the flow sensor measure the fuel flow far before the fuel enters the cylinder. Therefore, a linear correlation is determined between the fuel rack position and the fuel flow measurement to establish the fuel consumption by the fuel rack position.

From Figure 4.2 it can be seen that the combined dataset has the most records for the manoeuvring and starboard shaft blocked propulsion modes. This means that during the time span of the dataset a major part of time has been sailed into those propulsion modes. Also, the upward movement of the fuel consumption per mile for lower speeds, as expected from previous research, is not in the dataset for the transit and trailing shaft propulsion modes, see Figure 4.13.

The total resistance of the vessel can be determined by the parameters from the ship's data and the open water diagram of the propeller. The advanced ratio and the open water diagram leads to a thrust coefficient and this thrust coefficient together with the thrust deduction factor is used to determine the resistance. This resistance includes the resistance due to acceleration of the ship and therefore this resistance component is removed by making use of the derivative of the ship's speed and assuming a constant mass. From 4.6 it can be seen that the calculated resistance is roughly in accordance with the results obtained from model tests.

However, the precision of the thrust calculations is dependent on the precision and accuracy of sensor measurements and the technical condition of the propellers, we are using a static open water diagram. Because of the lack of insight into the technical conditions of the ship we assume them as constant in this thesis. Marine fouling on the propeller and hull are estimate to reduce the energy efficiency by 0.2 % per month, [Coraddu et al. \(2019a\)](#). We only have one month of data available in this thesis and therefore leaving out the impact of marine fouling in this thesis is valid assumption.

Further we will assume that the diesel engine are equal in performance during this month of operation and that degradation in the diesel engine as well as bearings and gearbox can be neglected because we are only taking into account one month of data.



# 5

## DATA ANALYSIS METHOD

*Machine learning is consisting of two main categories; Supervised Learning and Unsupervised Learning. Supervised learning algorithms try to model relationships between the target variable and the input parameters. Input features as well as the target variable are given to the computer and from the data the computer should determine the patterns and relationships. Common algorithms for supervised learning are; Nearest Neighbor, Naive Bayes, Decision Trees, Linear regression, Support Vector Machines and Neural Networks.*

*Unsupervised learning means that the computer will find relations and patterns in the data given a set of input parameters with the lack of a target variable. This means that the computer can give new insights from the data. Unsupervised learning is mostly used for problems where it is unknown where to look for in the data. Unsupervised learning is mostly applied in classification problem, for regression problems this is difficult to apply.*

*In this thesis the aim is to set up a model that is able to predict the fuel consumption per mile for varying technical, loading and environmental conditions, as shown in Figure 1.1. The fuel consumption per mile can be derived from the dataset, our target variable, and so in this thesis a supervised machine learning technique is applied.*

### 5.1. MULTI-LAYER PERCEPTRON (MLP) NETWORK

**R**EGRESSION analysis is a modelling technique that analyses the relation between the input variables and the target/output variable. In regression analysis the in- and output variables are continuous values, this is the biggest difference with classification. In classification the target variable is a discrete value and is representing a certain state. In this thesis we make only use of continuous values and so we make use of the regression based and supervised machine learning approach "Multi-Layer Perceptron".

MLP models are one of the most popular network architectures used in most of the research applications in engineering, mathematical modeling, [Yilmaz and Kaynar \(2011\)](#). They can be considered as most popular neural network subclass because of the high regression performances, [Le et al. \(2020\)](#), [Jeon et al. \(2018\)](#), [Tarelko and Rudzki \(2020\)](#).

A MLP network consist of an input layer, one or more hidden layers and an output layer, see Figure 5.1. Each layer is consisting of an amount of perceptrons which is known

as the width of the layer and the number of layers is called the depth of the neural network. MLP networks are feed forward neural networks, this means that the connection between different perceptrons do not form a cycle, the information moves in only one direction from the input layer to the output. Each layer of a MLP Neural Network has a number of perceptrons and each perceptron is fully interconnected with weighted connections to all perceptrons in the subsequent layer. The output of a perceptron is dependent on the activation function of that perceptron and the input values from the perceptrons in the subsequent layer. The size of the hidden layer is evaluated later in this section, the input layer has the size of the number of features that are used to predict the output. The size of the output is 1 in this thesis, so there is predicted only one value with one trained network with multiple inputs.

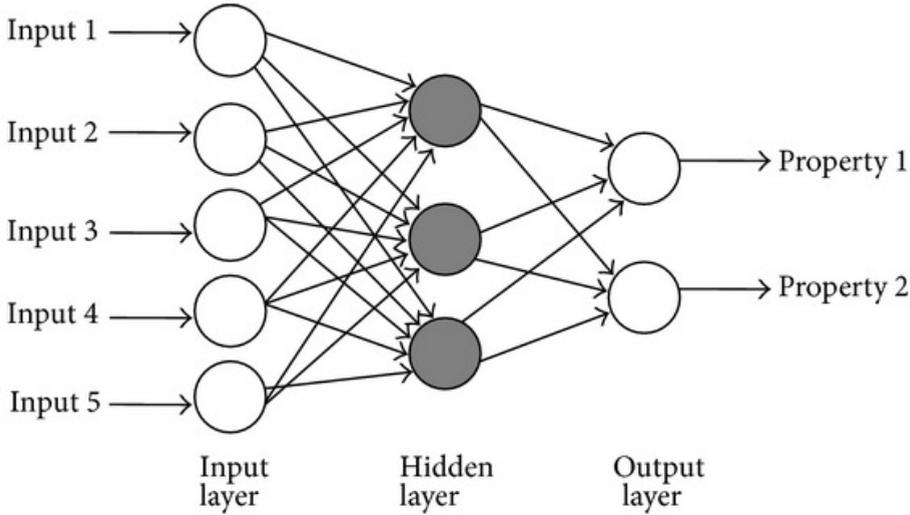


Figure 5.1: Example of MLP Neural Network structure with 5 input features, 1 hidden layer consisting of 3 perceptrons and one output layer consisting of two outputs.

### 5.1.1. MATHEMATICAL FORMULATION

At first the description of a single perceptron in the hidden layer will be explained, using Figure 5.2. Each perceptron has one input vector  $\mathbf{x}$  with  $n$  values,  $n$  is the number of values in the previous layer, and one output value  $y$ . The output value is determined by the activation function and the linear summation of the input values multiplied by their belonging weight,  $w_i$ , (i.e. the weighted sum) added with a bias,  $b_i$ , see Equation 5.1.

$$z = \sum_{i=1}^n w_i x_i + b_i = \mathbf{w}^T \cdot \mathbf{x} + \mathbf{b} \quad (5.1)$$

The output of the perceptron is determined by a certain activation function,  $f(z)$ , see Equation 5.2.

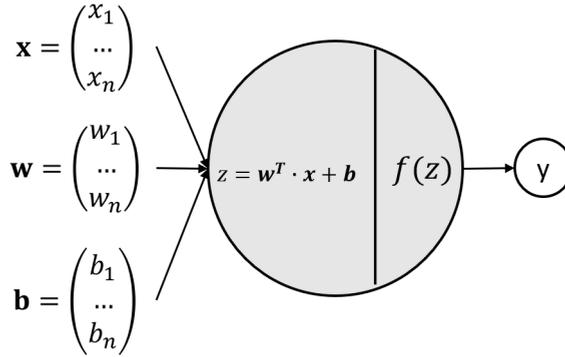


Figure 5.2: Mathematical explanation of a single perceptron in the hidden layer of a multi-layer perceptron model.

$$y = f(z) = f(\mathbf{w}^T \cdot \mathbf{x}) \quad (5.2)$$

5

An example of a complete multi-layer perceptron network is shown in Figure 5.3 and the belonging mathematical formulation in Equation 5.3. This example is consisting of an input layer consisting of 4 input values, 2 hidden layers with both 3 perceptrons and an output layer consisting of one output value. The input values are denoted by  $x_j$  where  $j$  is the number of inputs given to the network.  $h_i^{(l)}$  is the  $i$ th perceptron of the  $l$ th hidden layer, so  $l$  is the number of hidden layers that is used and  $i$  the number of perceptrons in the  $l$ th hidden layer.  $w_{ij}^{(l)}$  is the weight that is given to the connection between the  $i$ th perceptron of the  $l$ th hidden layer and the  $j$ th perceptron of the previous hidden layer or the input layer. Also,  $b_i^{(l)}$  is the bias for the  $i$ th perceptron of the  $l$ th hidden layer.  $\hat{y}$  is the calculated output by the model.

$$\begin{aligned} h_i^{(1)} &= f^{(1)} \left( \sum_j w_{ij}^{(1)} x_j + b_i^{(1)} \right) \\ h_i^{(2)} &= f^{(2)} \left( \sum_j w_{ij}^{(2)} h_j^{(1)} + b_i^{(2)} \right) \\ \hat{y} &= f^{(3)} \left( \sum_j w_{ij}^{(3)} h_j^{(2)} + b_i^{(3)} \right) \end{aligned} \quad (5.3)$$

In Equation 5.3 the  $f()$  represents the activation function of a certain node, we distinguish  $f^{(1)}$ ,  $f^{(2)}$  and  $f^{(3)}$ . It is possible to use a different activation function for the layer, for example if the output is binary. In this thesis we will use an unique activation function for every layer. The activation function determines the output of the node together with the sum of the weight multiplied by the inputs and the bias. In Section 5.1.2 multiple activation function are evaluated.

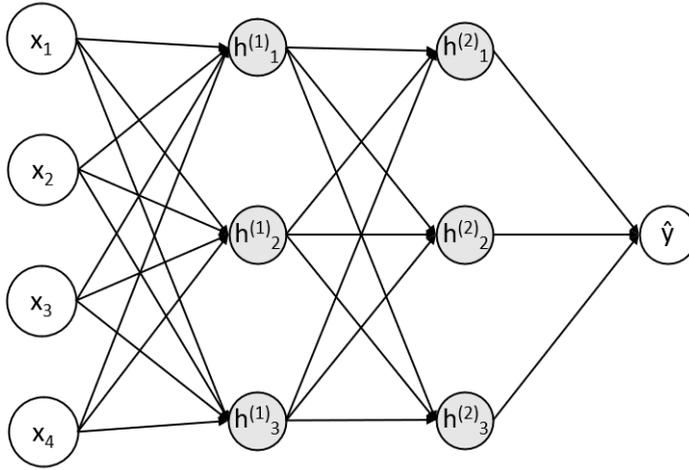


Figure 5.3: Example of MLP Neural Network structure with 4 input features, 2 hidden layers both consisting of 3 perceptrons and an output layer consisting of one output.

5

### 5.1.2. HYPER PARAMETERS

Training of the network means finding a set of weights and biases which make the loss function between the predicted and true output as small as possible. Hyper parameters are used to control and determine the training process and establish the functions in the neural network.

Training of the neural network will be based on backpropagation. At first feeding forward through the network (the forward pass) with initial weights and biases results in an output and so, an error between the output and the true value,  $E$ .

During the backward pass, feeding backward through the network, the weights and biases will be updated. This update is based on the partial derivative between the error and the weights, where the error is the difference between the true output  $y$  and the calculated output  $\hat{y}$ . With the contribution to the error in the output from the weights in every perceptron and every layer known, the weights and biases can be updated. Equation 5.4 shows the update of weights where  $\eta$  is the learning rate which controls the step size that is taken when updating the weights.

$$w \leftarrow w - \eta \frac{\partial \text{Error}}{\partial w} \quad (5.4)$$

#### ACTIVATION FUNCTION

In this thesis there are used identical activation functions for every layer, however it is possible to implement different activation functions for every layer, e.g. if the output is binary it can be a binary activation function for the output layer. The most common activation functions that evaluated are the hyperbolic tangent (Equation 5.5), logistic (Equation 5.6), rectified linear unit (Equation 5.7) and linear function (Equation 5.8).

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (5.5)$$

$$f(x) = \frac{1}{1 + e^{-x}} \quad (5.6)$$

$$f(x) = \max(0, x) \quad (5.7)$$

$$f(x) = x \quad (5.8)$$

In Figure 5.4 the different activation functions are visualised. The linear activation function takes the inputs multiplied by the weight for each input perceptron and creates an output signal proportional to the input. However, it is not possible to use backpropagation to train the model, the derivative of the function is a constant and so has no relation to the input. Also, the number of layers in the neural network makes no difference when a linear activation function is used. A linear combination of linear functions is still a linear function and so a neural network with a linear activation function is simply a linear regression model. A linear regression model has limited power and ability to handle complexity varying parameters of input data, therefore we also investigate non-linear activation functions.

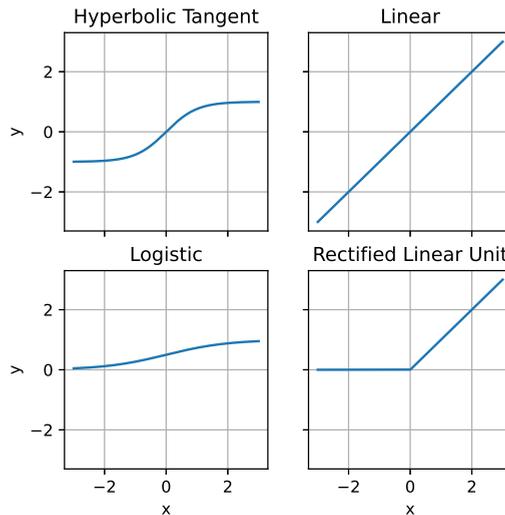


Figure 5.4: Visualisation of four different activation functions; Hyperbolic Tangent, Linear, Logistic and Rectified Linear Unit.

The rectified linear unit activation function allows the network to converge to an optimum very quickly and it allows for backpropagation. However, when input values approach zero, or are negative, the gradient of the function becomes zero. In this situation the network cannot perform backpropagation and so cannot learn. The logistic function

has a smooth gradient that prevents peaks in the output, however this function gives only outputs between 0 and 1. The hyperbolic tangent is zero centered and so it is easier to model inputs that have strongly negative, neutral and strongly positive values.

#### HIDDEN LAYER SIZE

The number of hidden layers has to be chosen to create the model architecture. [Heaton \(2015\)](#) states that problem that require more than two hidden layers were rare prior to deep learning. Two or fewer layer will often suffice with simple data sets. However, with complex datasets involving time-series or computer vision, additional layers can be helpful. He summarizes the capabilities of several common layer architectures in [Table 5.1](#).

Table 5.1: Determining the number of hidden layers [Heaton \(2015\)](#)

Number of Hidden Layers	Result
none	Only capable of representing linear separable functions of decisions.
1	Can approximately any function that contains a continuous mapping from one finite space to another.
2	Can represent an arbitrary decision boundary to arbitrary accuracy with rational activation functions and can approximate any smooth mapping to any accuracy.
>2	Additional layers can learn complex representations (sort of automatic feature engineering) for layer layers.

The number of perceptron in the hidden layer decides the overall neural network architecture and have an impact on the final network output. Too few perceptrons can lead to underfitting which means that the network is not able to detect all relations between the features and the target. On the other side, too many neural can lead to overfitting. In this case the many perceptrons cannot be trained because of the limited amount of information in the features relative to the number of perceptrons. Also, too many perceptrons can lead to very long training times and finally it is not possible to train the algorithm adequately even more.

Therefore [Heaton \(2015\)](#) summarized a few rule-of-thumb methods, that provides a good starting point;

- The number of hidden perceptrons should be between the size of the input layer and the size of the output layer.
- The number of hidden perceptron should be 2/3 the size of the input layer, plus the size of the output layer.
- The number of hidden perceptrons should be less than twice the size of the input layer.

Ultimately, the selection of an architecture will come down to trail and error, using the grid search method. So, to establish the optimum architecture of the model a grid search is performed with as starting point these rule of thumbs.

### OPTIMIZATION FUNCTIONS

As said before, backpropagation is an algorithm that is used to compute the gradients of the loss function with respect to the weights and the biases in the network and trains the network by updating them. Optimization functions are used to determine how the weights in the model are updated. The loss function used in this thesis and most commonly used for regression models is the least squares, see Equation 5.9. This approach minimizes the weighted sum of the squares of the error between the calculated output of the model and the true value. In this thesis we evaluated three optimization functions.

$$\min E = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2 \quad (5.9)$$

**SGD** The stochastic gradient descent is the most commonly used first-order optimization method. The idea of the stochastic gradient descent method is that variables update iteratively in the (opposite) direction of the gradients of the objective function. The learning rate  $\alpha$  determines the step size in each iteration, and thus influences the number of iterations to reach the optimal value.

Mathematical formulated, the stochastic gradient descent method update the weights and biases according to

$$w_{i+1} = w_i - \eta \frac{\partial E}{\partial w_i}$$

**Adam** Adam is a much used optimization algorithm too, it comes from *adaptive moments*. Compared to Stochastic Gradient Descent it is *adaptive* - it computes different learning rates for different parameters. It also make use of momentum, i.e it accumulates an exponential decaying average of previous gradients and therefore takes the history into consideration when calculating a new gradient. The exponential decaying average of the negative gradient is seen as a velocity from where the analogy to physics can be seen. The Adam optimizer uses two momentum terms, one first order and one second order, [Frost and Lavatt \(2020\)](#).

Mathematical formulated, the Adam method update the weights and biases according to

$$w_{i+1} = w_i - \eta \frac{s_i}{r_i}$$

where  $s$  and  $r$  are initialised to 0 and update according to

$$s_{i+1} = \frac{\beta_1}{1 - \beta_1^i} s_i + \frac{1 - \beta_1}{1 - \beta_1^i} \frac{\partial E}{\partial w_i}$$

$$r_{i+1} = \frac{\beta_2}{1 - \beta_2^t} s_i + \frac{1 - \beta_2}{1 - \beta_2^t} \frac{\partial E}{\partial w_i} \cdot \frac{\partial E}{\partial w_i}$$

The time step  $t$  is initialized to 0 and increased by one for each iteration. The exponential decay rates  $\beta_1$  and  $\beta_2$  can, same as learning rate  $\eta$ , be tuned/set by choice.

**L-BFGS** This optimization method is short for Limited memory-BFGS (Broyden-Fletcher-Goldfarb-Shanno algorithm). It is a second-order optimization algorithm with uses the second order derivatives of the loss function. The easiest method mathematical would be Newton's method, it is derived from the Taylor expansion around a local minimum. Newton's method make use of the Hessian, a symmetric matrix containing all second derivatives of the loss function evaluated at a specific parameter vector and updates according to, [Frost and Lavatt \(2020\)](#)

$$w_{i+1} = w_i - \eta (H_{xi})^{-1} \frac{\partial E}{\partial w_i}$$

5

## 5.2. CONCLUSIONS

To find the optimal model with respect to the model hyperparameters a grid search is performed. This way of finding an optimal hyper parameter subset is based on manually specify a subset of all hyperparameters that are desired to combine, the grid search trains the model for every combination of hyperparameters and gives the accuracy of the model. So, in the end we will chose the hyper parameter subset that gives the best accuracy. Using the grid search leads to accurate results and better prediction of MLP, [Harirchian et al. \(2020\)](#).

The accuracy of the model/error metric is defined by the relative mean absolute error,  $MAE\%$ , see Equation 5.11. This metric gives the mean absolute error  $MAE$ , see Equation 5.10 as a percentage of the target value range. The target value range is the difference between the real target values in the test dataset,  $Y_{max} - Y_{min}$ . The mean absolute error itself is given as the absolute difference between the predicted values,  $\hat{y}$ , and the real target value,  $Y_n$ , divided by the number of values in the test dataset,  $N$ .

$$MAE = \frac{\sum_{i=1}^N |Y_n - \hat{y}_n|}{N} \quad (5.10)$$

$$MAE\% = \frac{MAE}{Y_{max} - Y_{min}} \quad (5.11)$$

The combinations of hyperparameters as shown in Table 5.2 will be taken into account to find an optimal set of parameters. The investigation of the hidden layer size is based on the rule thumbs given by [Heaton \(2015\)](#). The example is given for 12 input features and one output.

Table 5.2: Overview of the hyperparameters used for the grid search.

<b>Hyper Parameter</b>	<b>Choices</b>
Activation Function	'Tanh', 'Linear', 'Logistic', 'ReLU'
Optimizer	'SGD', 'Adam', 'L-BFGS'
Hidden Layer Size	(24), (12), (9), (24, 24), (12, 12), (9,9), (24,24,24), (12,12,12), (9,9,9)
Learning Rate	0.1, 0.01, 0.001, 0.0001, 0.00001
$\beta_1$	0.1, 0.5, 0.9
$\beta_2$	0.1, 0.5, 0.9



# 6

## WEATHER CONDITIONS FROM SHIP'S DATA

*To establish fuel consumption curves to help the operator select optimal setting to reduce CO<sub>2</sub> emissions the current weather conditions are of importance. The weather conditions as measured by the ship are the wind speed and direction relative to the ship. Oceans wave parameters, the significant wave height, wave period and wave direction, are not available. Therefore, we reconstruct the route of the ship in Chapter 3 to add those ocean wave parameters to the dataset. However, these parameters are hourly averaged and given on a latitude-longitude grid of 0.5 °x 0.5 °which is in accordance with approximately a grid of 25 kilometer x 25 kilometer.*

*Obtaining accurate results for the influence of weather conditions on the fuel consumption will be difficult, because of that the ocean wave parameters are hourly averaged and with that resolution. This accuracy could be increased by determining the current weather conditions with a higher resolution. Weather conditions with a higher resolution will also cover the variability of weather such as gusts. In this chapter the relation between the ship's data and the oceans wave parameters is investigated and the prediction of those parameters by the ship's data is described, to answer Research Question 2: 'How can real-time weather conditions be inferred from ship's data?'*

*This chapter is organized as follows: In Section 6.1 the selected feature to infer the ocean wave parameters from the ship's data are described. Second, in Section 6.2 the dataset used to train the machine learning model is created and in Section 6.3 an overview of the model, features and target variables is given. Section 6.4 describes the hyper parameter tuning and the accuracy of the final models.*

### 6.1. FEATURE SELECTION

**T**HIS section explains the feature selection for the prediction of the ocean wave parameters. Feature selection is a very important concept in machine learning and it impacts the performance of the machine learning model hugely. Feature selection is the

procedure of selecting a subset (some out of all available) of the input variables that are most relevant to the target variable. Irrelevant features lead to inaccurate results of the model. A good feature subset reduces overfitting, improves the accuracy and reduces the training time.

Features can be characterized as, [Bell and Wang \(2000\)](#):

1. **Relevant** These are features which have an influence on the output and their role can not be assumed by the rest.
2. **Irrelevant** Irrelevant features are defined as those features not having any influence on the output, and whose values are generated at random for each example.
3. **Redundant** A redundancy exists whenever a feature can take the role of another (perhaps the simplest way to model redundancy).

The two main problems in machine learning are classification and regression problems. In a classification problem the goal is to categorize data into classes. For example, the available data in this thesis can be used to calculate in which propulsion mode (categories) the ship is sailing. In this chapter the goal is to predict ocean wave parameters which are discrete values from discrete values, the data, and so this is a regression problem.

For the prediction of the significant wave height, wave period and wave direction with respect to the ship features are selected based on domain knowledge. The features that could be relevant are;

### ROLL AND PITCH

Ship motions are expected to be well predicting features for the ocean wave parameters. Ship motions are described by six degrees of freedom, surge, sway, heave, roll, pitch and yaw, which define the ship motions due to wind and waves. We will only take into account the roll and pitch motion because of the fact that those are available in the dataset. The pitch gives the rotation about the side-to-side axis of the ship and the roll the rotation around the stern-to-bow axis.

The motions of a ship are arising from oceans waves and therefore it is expected that these parameters contain information about the ocean waves. In regular waves, the frequency of the motion and the frequency of the regular wave is always the same. However, oceans waves are almost irregular. Irregular waves can be viewed as the superposition of a number of regular waves with different frequencies and amplitudes. [Nakamura and Naito \(1975\)](#) has stated that the amplitude of heave, pitch and surge motions is proportional to the wave height of regular waves and the linear superposition method is valid for predicting the ship motions in irregular waves.

The pitch,  $\theta$ , and roll,  $\phi$ , motion of the ship in regular waves can be described by the amplitude of the motion,  $\phi_a$ , times cosine of the encounter frequency of the waves,  $\omega_e$ , plus a phase difference between the wave and the motion,  $\epsilon_{\phi\zeta}$ , see Equation 6.1 and 6.2, [Journée and Massie \(2001\)](#).

$$\phi = \phi_a \cdot \cos(\omega_e \cdot t + \epsilon_{\phi\zeta}) \quad (6.1)$$

$$\theta = \theta_a \cdot \cos(\omega_e \cdot t + \epsilon_{\theta\zeta}) \quad (6.2)$$

The static roll and pitch angles are given in the data at certain timestamps. The method to find the amplitude and the period of a signal is by searching for the value of peaks and the timestamp where these peaks occur. At first all local minimums and local maximums are located, where every local minimum and maximum is defined by a timestamp and a value.

At the moment the values of peaks are known, the local maximums can be minimized by their consecutive local minimum, the absolute value of this outcome results in the amplitude of the measurements at that moment. The period of the signal is defined by the difference between the timestamp of two consecutive minimums or maximums in seconds.

Figure 6.1 shows the roll angle from the data and the resulting minimums and maximums during a time period of 5 minutes from the measurements of the roll motion of the ship. On the right axis we can see ten times the amplitude and period as obtained from the method described above. The measurement of the pitch motion of the ship shows a similar movement, only the values are different.

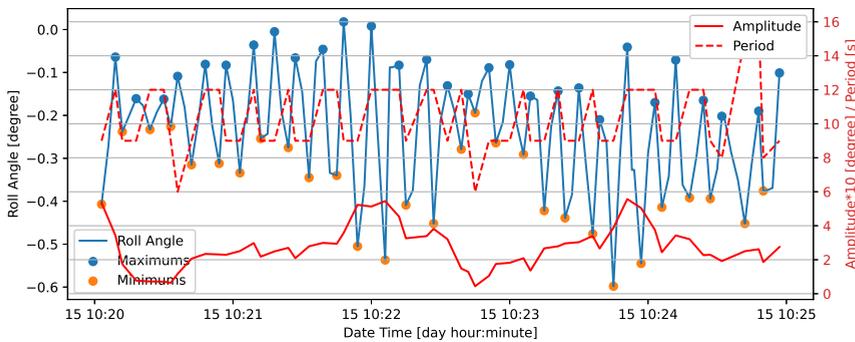


Figure 6.1: Visualisation of the roll angle for five minutes from the data. The peaks of the signal are indicated by the dots. The scale of the resulting amplitude and period are given on the right axis in the figure.

## FUEL RACK POSITION

The fuel rack determines the position of the fuel pump plungers which in fact determines the amount of fuel that enters the cylinder. The fuel rack position is controlled by the governor of the engine which continuously control the fuel rack position to maintain the engine speed set point. All effects that contribute to a fluctuation in the engine speed are defused by the engines governor and so the amount of fuel injected into the cylinder.

The produced thrust by the propeller is calculated in Section 3.3. The thrust is calculated with the non-dimensional thrust coefficient looked up from the open water diagram with the advanced ratio  $J$  and the actual propeller pitch. The advanced ratio  $J$  is dependent on the propeller inflow velocity  $v_a$ . Therefore the fluctuations in thrust produced by the propeller are dependent on the fluctuating propeller inflow velocity, [Nakamura and Naito \(1975\)](#).

This axial inflow velocity at the propeller is caused by the wave induced particle motion and surge motion of the ship demonstrated by Ueno et al. (2013). Also the pitch motion influences the inflow at the propeller, Faltinsen (1980) proposed that the wake velocities due to the pitching motion of the ship can be calculated assuming the bottom of the ship to be a flat plate. Taskar et al. (2016) combined these methods into a time varying total wake velocity in waves considering the mean increase as well as fluctuations. Taskar et al. (2016) also demonstrated the time varying fluctuation in engine torque due to the varying wake velocity.

The fluctuation in the fuel rack position is representative for torque fluctuations of the propeller resulting from oceans waves and ship motions. Therefore, to derive the ocean wave parameters we will extract this time varying fluctuation from the fuel rack position.

During continuous operation, so no acceleration and manoeuvring, the fuel rack position shows a highly continuous signal, see Figure 6.2.

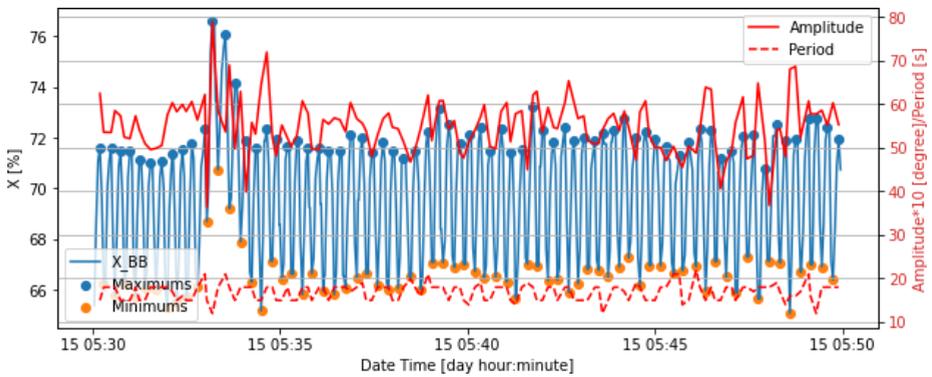


Figure 6.2: Visualisation of the fuel rack position for twenty minutes from the data. The peaks of the signal are indicated by the dots. The scale of the resulting amplitude and period are given on the right axis in the figure.

### DRAFT MEASUREMENTS

To add the effect of loading conditions on the motions and responses of the vessel due to oceans waves, the draft measurements at the bow and stern are assumed to be good features. These parameters represent approximately the loading condition of the ship and so may be an important feature into the prediction of oceans wave because the ship motions due to weather are influenced by the draft of the ship.

Also, the draft measurements are influenced by waves at the position of measurement, they measure when a wave peak or trough is passing the measurement device. Therefore, the draft at the bow and stern are also considered as good features to predict the oceans wave parameters.

### SET AND DRIFT

The drift is calculated by the data enrichment in Section 3.3, indirectly from the ship's data. The drift is a measure for the current and split up into a component pointing into

the direction of the ship's movement ( $Drift_y$ ) and perpendicular ( $Drift_x$ ) to this direction. The set gives the difference between the heading and the course over ground of the ship.

There are many factors, such as wind, current and oceans waves, that are working for or against the ship. These factors will have an impact on the set, because the ship has to compensate course to keep the desired course over ground. Therefore it is assumed that the set should have a relation with the ocean wave parameters. The current is affecting the propeller inflow velocity and therefore a factor that should be taken into account to infer ocean wave parameters.

#### RELATIVE WIND SPEED AND ANGLE

The relative wind speed and angle are measurements on board of the ship and existing in the ship's data. These parameters present the wind as felt by the ship as explained in Section 3.3. The speed of wind and the angle of wind are factors affecting the set of the ship. The ship has to compensate course to maintain the desired course over ground. Those factors have also an influence on the motions of the ship and therefore should be taken into account.

Instead of using the relative wind angle in degrees, the cosine and sin of this angle are chosen as features. The cosine of the relative wind angle is the fraction of relative wind speed that is directed onto the movement direction of the ship and the sine is the fraction of relative wind coming from aside.

#### SPEED OVER GROUND & THROUGH WATER

The speed over ground is assumed to be an important feature. To take into account difference into the wave encounter frequency. Namely, the ship's response due to incoming waves is different for waves coming from astern and waves coming from ahead. The encounter frequency of waves is given in Equation 6.3, Journée and Massie (2001).  $\omega$  is the circular wave frequency in rad/s,  $g$  the gravity constant in  $m^2/s$ ,  $V$  the speed relative to the water in  $m/s$  and  $\mu$  the direction of waves w.r.t. the ship.

$$\omega_e = \omega - \frac{\omega^2}{g} V \cos(\mu) \quad (6.3)$$

With the wave encounter frequency the wave elevation is determined and so the motions of the ship.

## 6.2. DATA PREPROCESSING

**D**ATA preprocessing is a major step in building in a machine learning model that is able to predict a target variable by features. In this section required steps are taken to prepare a dataset used to train and validate the created models.

At first it is chosen to create for every environmental parameter one model and we also for every unique propulsion mode. In Figure 4.2 we evaluated the amount of time the ship made use of the propulsion modes and clearly from this section is that the manoeuvring mode is used the greatest amount of time and so contains the most records. Therefore we will evaluate the performance of the models created by the manoeuvring

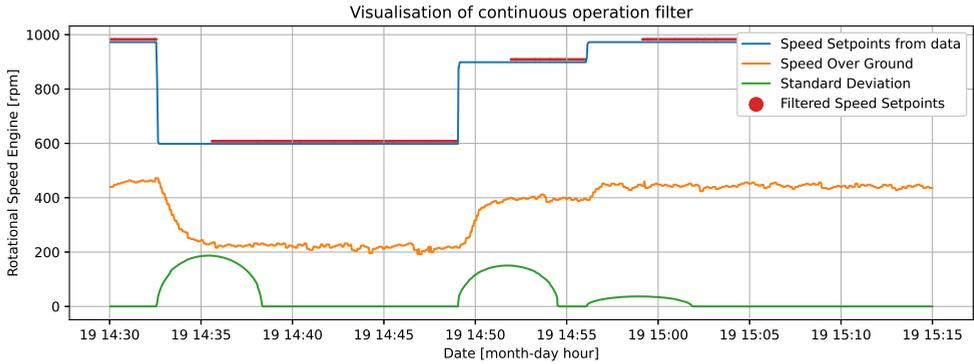


Figure 6.3: Visualisation of filtered data by continuous speed setting versus the raw data

propulsion mode. Also, by training the model the aim is that the model finds the physical relation between the features and the target variable. Those physical relations are different for the propulsion modes, the propulsion modes have their own combinator curve, and so for example the relation between the speed and fuel rack position is different. Because of the regression based nature of the model it is not able to link these differences in relation to the different modes.

Secondly, the ship's data shows high variety, this variety can be explained by e.g. the differences in environmental conditions, accelerations and decelerations. The aim is to establish the environmental conditions, therefore we filtered out a period of 180 seconds after the speed setpoint of the vessel is changed to ignore the effect of accelerations and decelerations.

To filter the continuous operation of the vessel, the moments of a change in the speed setpoint are determined. This is done by calculating the standard deviation over a predefined window/number of consecutive observations ( $N$ ) in the data, see Equation 6.4 where  $\bar{x}$  is the average over  $N$  observations and  $x_i$  the  $i$ th sample of  $N$  observations. The standard deviation measures the amount of variation or dispersion in the values, in this case the standard deviation is taken over the RPM-settings.

The standard deviation over a predefined window/number of consecutive observations ( $N$ ) is calculated at every timestamp, the number of consecutive observations is chosen to be 180 seconds. The moment that the standard deviation changes from 0 to a value higher than zero indicates that there is a change in speed setpoint. It is chosen to leave out the data within 180 seconds after a change of speed setpoint. Figure 6.3 shows that this time window leaves out the effect of the decelerations or accelerations almost completely.

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (6.4)$$

Also, the influence of a rolling mean over the dataset on the performance of the model is evaluated. The weather conditions that will be used are hourly averaged, however the ship's data has a resolution of one measurement every three seconds. To make

the model more accurate we will investigate further in this chapter if a rolling average over the dataset will give better results for the predicting capabilities of the model. A rolling mean means that we will add at every record the average of the values in front of the record in a predefined window. For example, at every record the average of the all the values five minutes back in time will be added.

Finally, the model has to be trained by the data where the ship is actually sailing. Therefore we will make use of the speed over ground. The values where the speed over ground is smaller than 0 knots will be left out the dataset.

### 6.2.1. FEATURE SCALING

Feature scaling is one of the most critical steps in different machine learning algorithms. Common scaling techniques in machine learning are standardization and normalization. Scaling is important because a machine learning algorithm is not able to recognize the meaning behind values. Without scaling, features with a high value range get great importance of the algorithm, however this is not necessarily true. Also, for some machine learning algorithms feature scaling helps decreasing the convergence time to a solution.

In this thesis the 'StandardScaler' is used as feature scaling method. This scaling method transform the dataset such that the resulting distribution's mean value is zero and the standard deviation is one. Equation 6.5 gives the transformed value of the feature,  $Z$ ,  $x$  is the original value,  $\mu$  is the mean of all values and  $\sigma$  is the standard deviation.

$$Z = \frac{x - \mu}{\sigma} \quad (6.5)$$

### 6.2.2. CROSS-VALIDATION

To be sure that the model is trained by all possible scenarios in the dataset and is trained by a dataset that do not miss important information cross-validation is used. The type of cross-validation that is used is K-fold cross validation, see Figure 6.4. This model guarantees that the score of the model is not depending on the way we chose the train and test dataset. This method randomly splits our training dataset into five folds, iterations. The model is trained and tested on every iteration and gives for every iteration a accuracy of the model. The average of the five resulting accuracy's is known as the cross-validation-score.

## 6.3. MODEL UNDERSTANDING

**T**HIS section gives an overview of the target variables and the model. Figure 6.5 gives an overview of input and output of the models that are trained in this chapter. An detailed description of the features used is given in Section 6.1.

Figure 6.6 shows the to be predicted targets, the significant wave height, wave direction and wave period. As can be seen, the dataset contains information about the significant wave height in the range of 0.3 to 2.6 meters, about the wave period in the range of 3.3 to 6.0 meters and about the wave direction w.r.t. to the ship in the range of 0 to 360 degree.



Figure 6.4: Visualisation of the K-Fold cross validation method.

The mean wave direction is predicted with two different models. The biggest challenge in the prediction of the wave direction with respect to the ship  $\mu$  is that the algorithm 'thinks' only in numbers. In this case, the algorithm interprets 0 and 360 degrees as totally different values where in fact these values are identical. To overcome this problem we make use of Cartesian coordinates, namely to the cosine and sine of the wave direction with respect to the ship. In fact, there is made a prediction of two variables, the cosine and sine of the wave direction, after these two separate predictions the two will be combined and the result is the wave direction, see Equation 6.6.

$$\mu = \begin{cases} \arctan\left(\frac{\sin(\mu)}{\cos(\mu)}\right), & \text{if } \cos(\mu) > 0, \\ \arctan\left(\frac{\sin(\mu)}{\cos(\mu)}\right) + \pi, & \text{if } \cos(\mu) < 0, \sin(\mu) \geq 0, \\ \arctan\left(\frac{\sin(\mu)}{\cos(\mu)}\right) - \pi, & \text{if } \cos(\mu) < 0, \sin(\mu) < 0, \\ +\frac{\pi}{2}, & \text{if } \cos(\mu) = 0, \sin(\mu) > 0, \\ -\frac{\pi}{2}, & \text{if } \cos(\mu) = 0, \sin(\mu) < 0, \\ \text{undefined}, & \text{if } \cos(\mu) = 0, \sin(\mu) = 0, \end{cases} \quad (6.6)$$

#### 6.4. MODEL TRAINING AND VALIDATION

FOR every ocean wave parameter a model will be trained with the previous described grid search and cross-validation method. At first performance of the significant wave height are described, thereafter the wave period and finally the wave direction with respect to the ship.

To chose the hyper parameters that give the highest accurate model performance an extensive grid search is performed. For all combinations of hyper parameter values,

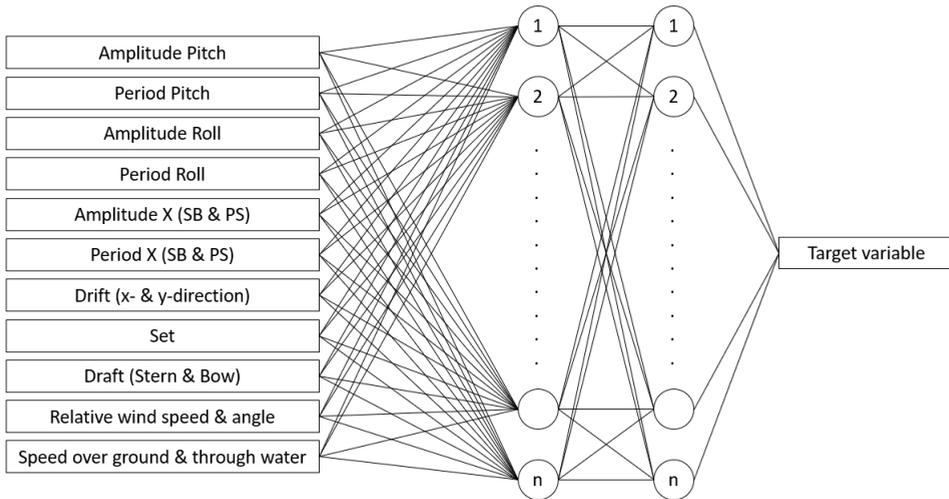


Figure 6.5: Description of the model for the prediction of ocean wave parameters, significant wave height, wave period and wave direction. Features are given in the rectangle boxes, the number of hidden layers is two and the number of neurons per hidden layer has to be determined.

listed in Table 5.2, the accuracy of the model is determined. The result of the extensive grid search shows that the combination for solver 'Adam', activation function 'ReLU', a hidden layer size of (18,18,18), a learning rate of 0.001, a  $\beta_1$  of 0.9 and a  $\beta_2$  of 0.5 give the most accurate results for all the models.

Figure 6.7 until Figure 6.12 visualises the model performance by plotting the true values from the dataset against the predicted values by the model. Figure 6.7 and 6.8 give the performance of the model for the prediction of the significant wave height with no rolling mean and a rolling mean of 3 minutes respectively, Figure 6.9 and 6.10 for the wave period and Figure 6.11 and 6.12 for the wave direction.

The accuracy's of the different models are summarised in Table 6.1. The target variables are predicted more accurate by taking a rolling mean over the data, except for the wave direction with respect to the ship. This can be explained by the fact that the wave direction with respect to the ship is derived from the hourly averaged wave direction with respect to True North and the heading of the ship. This makes that the wave direction with respect to the ship is varying with the heading of the ship which is origination from the high resolution dataset.

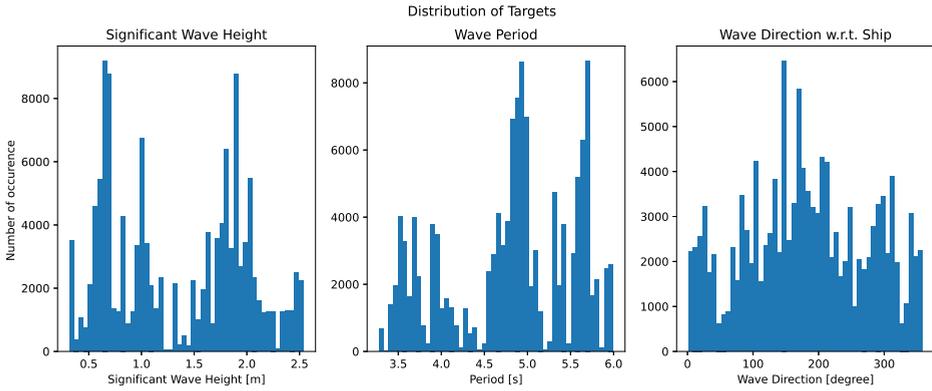


Figure 6.6: Visualisation of the distribution of the to be predicted ocean wave parameters, targets. The ocean wave parameters consist of the significant wave height, the wave period and the wave direction with respect to the ship.

Table 6.1: Accuracy of the models defined by the mean absolute error and the relative mean absolute error. The accuracy is given for the situation no rolling mean is applied and the situation a rolling mean of 3 minutes is applied to the dataset.

Targets	No rolling mean		Rolling mean of 3 minutes	
	MAE	MAE% [%]	MAE	MAE% [%]
Significant wave height	0.070 m	3.146	0.052 m	2.337
Wave period	0.087 sec	3.196	0.062 sec	2.278
Cosine wave direction	0.083	4.150	0.048	2.400
Sine wave direction	0.082	4.100	0.058	2.900
Wave direction	6.365 degree	1.770	9.160 degree	2.544

6

### 6.4.1. SIGNIFICANT WAVE HEIGHT

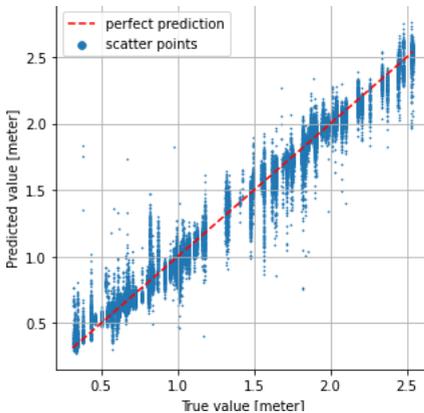


Figure 6.7: True values versus the predicted values for the significant wave height. The dataset that is used for training of this model is not provided with a rolling mean.

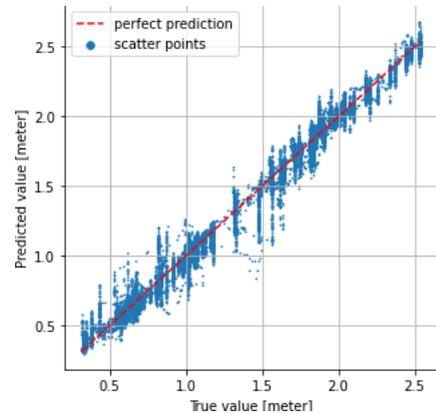


Figure 6.8: True values versus the predicted values for the significant wave height. The dataset that is used for training of this model is provided **with a rolling mean** of 3 minutes.

### 6.4.2. WAVE PERIOD

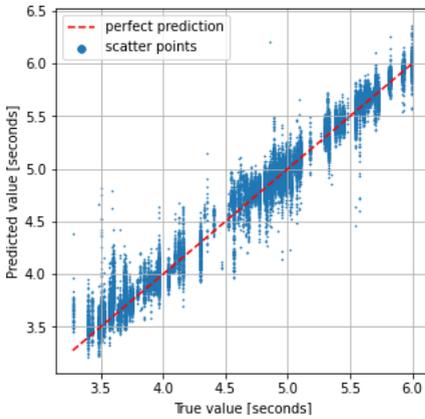


Figure 6.9: True values versus the predicted values for the wave period. The dataset that is used for training of this model is not provided with a rolling mean.

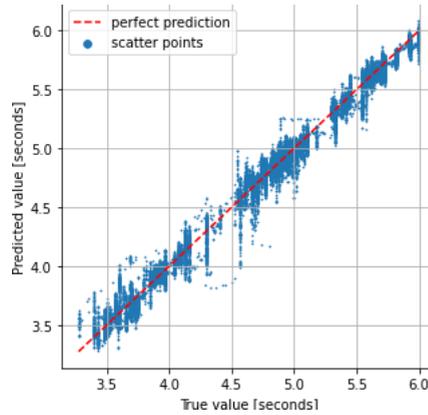


Figure 6.10: True values versus the predicted values for the wave period. The dataset that is used for training of this model is provided with a rolling mean of 3 minutes.

### 6.4.3. WAVE DIRECTION W.R.T. THE SHIP

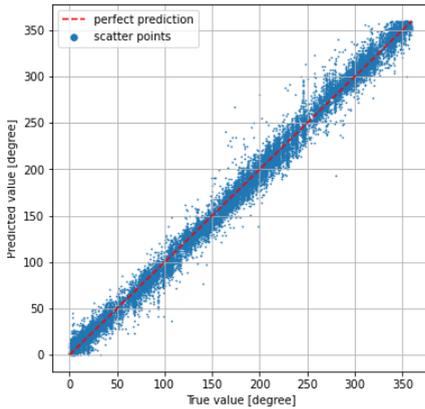


Figure 6.11: True values versus the predicted values for the wave direction w.r.t. the ship. The dataset that is used for training of this model is not provided with a rolling mean.

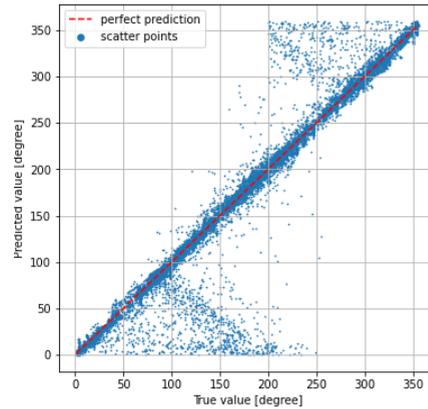


Figure 6.12: True values versus the predicted values for the wave direction w.r.t. the ship. The dataset that is used for training of this model is provided with a rolling mean of 3 minutes.

## 6.5. CONCLUSIONS

To give real-time insights to the crew in the  $CO_2$  performance of the ship, real-time information of weather conditions is required. The available data about weather conditions is historical, averaged over 1 hour and with a low resolution. Therefore this chapter has proposed a methodology to infer the significant wave height, wave period

and direction from the ship's data. The accuracy of the models, the selected features and the different steps to be taken to build the models are described and give an answer to Research Question 2:

*How can real-time weather conditions be inferred from ship's data?*

Real-time weather conditions can be inferred from ship's data with an accuracy of the trained algorithms listed in Table 6.1. The inaccurate and with a low resolution weather conditions can be predicted with a relative mean absolute error of less than 4.150 % for the wave period, wave direction and significant wave height. By using a rolling mean over the dataset the mean absolute error decreases more, except the of the wave direction.

The determined input features based on domain knowledge turn out to contain enough information about the ocean wave parameters. Also, the extensive grid search combined with the K-Fold cross validation method turns out to be an effective methodology to tune the hyper parameters.

However, it should be noted that the algorithms are trained with ocean wave parameters as target, hourly averaged and defined in a lat-lon grid of  $0.5^\circ \times 0.5^\circ$ . Results show that with a rolling mean of 3 minutes the mean absolute error decreases, except for the wave direction. This seems logical because with a rolling mean the features of for the model are averaged, as well as the targets, in other words the feature values are less fluctuating and more in accordance with the fluctuations in the target variables. The wave direction is derived from the heading of the ship and so contains fluctuations every 3 seconds because the heading is extracted from the ship's data.

When the ocean wave parameters are established with a higher frequency, the expectation is that the selected features will have higher predicting capabilities even more. Then, the resolution of information about ocean waves in the target variables will be more accordance with the resolution of the ship's data.

# 7

## FUEL CONSUMPTION CURVES

*In this chapter fuel consumption curves are established that give insight into the optimal speed setting and propulsion mode given the current weather conditions. Establishing these fuel consumption curves real-time we can support the crew to make informed decisions about their speed and propulsion mode setting and they still being able to satisfy their mission profile.*

*Geertsma et al. (2017) already proposed fuel consumption curves for the OPV in three different conditions and two propulsion modes, see Figure 4.11. This figure shows that the two propulsion modes have a varying optimal speed setpoint in the different conditions and the setpoint differs for the propulsion modes. This chapter proposes a methodology to establish the fuel consumption curves by the ship's data. To implement the influence of ocean wave conditions on a real-time base, we will establish the actual ocean wave conditions with the methodology proposed in Chapter 6.*

*This chapter is organised as follows : In Section 7.1 we establish the target variable that will be used to set up fuel consumption curves. Thereafter, in Section 7.2 a schematic overview and explanation of the model is given. In Section 7.3 the input parameters to predict the target variable are described and in Section 7.4 the model is developed able to predict the target with the input parameters. Finally, in Section 7.5 the performance and validation of the model outputs are evaluated.*

### 7.1. TARGET UNDERSTANDING

**T**HE target parameter, fuel per mile, is investigated to understand the performance and capabilities of the model. This is important to identify possible extrapolation difficulties or how the model deals with unbalanced data.

The fuel consumption per mile from the data for the different modes is given in 4.13. Figure 7.1 gives the fuel consumption against the speed over ground and the speed through water for the manoeuvring mode only. In this Figure the data is filtered for the continuous speed setting periods as described in Chapter 4. The colors in the figure represent the rotational speed of the propeller.

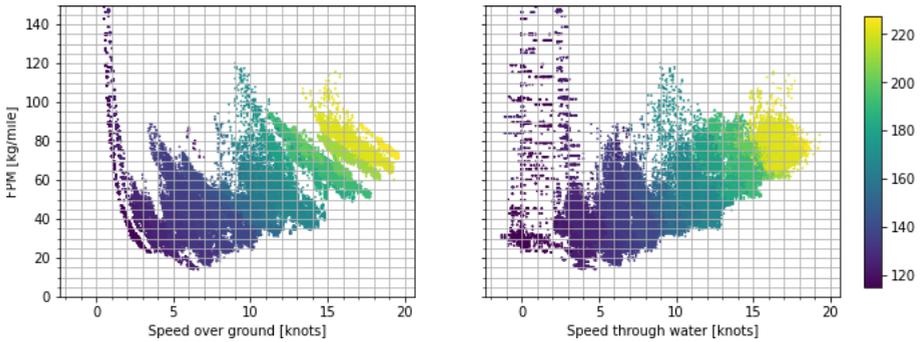


Figure 7.1: Fuel consumption per mile given against the speed over ground and the speed through water. The colors in the figure are representing the rotational speed of the propeller. The data is of the manoeuvring propulsion mode only and filtered for acceleration and decelerations.

In the figure where the FPM is given against  $V_s$  we can clearly see the continuous rotational speed settings of the propeller, indicated by the color of the points. This can be explained by the fact that the OPV has programmed speed settings, Figure 4.3, such as; full speed forward or medium speed forward which are coupled to a certain virtual shaft speed. The shape of the continuous rotational speed lines can be declared by the fact that the FPM is calculated by the  $V_s$ , see Equation 4.5. Therefore, the FPM is inversely proportional with the  $V_s$ .

Obvious is the enormous variation in speed over ground versus fuel consumption per mile for a continuous rotational speed of the propeller. A lower speed over ground results in a higher fuel consumption per mile for a constant rotational propeller speed.

The figure also gives the fuel consumption per mile against the speed through water. In this figure the continuous rotational speed of the propeller areas are visible too, however less clear. If we calculated the fuel consumption per mile with respect to the water we expect the inverse proportionate relation between the FPM and  $V_w$  too. In fact,  $V_s$  should be replaced by  $V_w$  in Equation 1.1. However, we will use the FPM calculated by the  $V_s$  as target variable to train the models. This because of fuel consumption curves dependent on the speed over ground will be more relevant for the crew. Namely, the navigation and bunker fuel plans are based on the speed over ground.

## 7.2. MODEL DESCRIPTION

To establish fuel consumption curves for the propulsion modes corrected for the current environmental conditions we will make use of the models created in Chapter 6. With those models it is able to establish the current environmental conditions by the ship's data only and therefore fuel consumption curves can be presented to the crew real time. Figure 7.2 summarises the model description.

At first the required features for the determination of the significant wave height, wave direction and period will be extracted from the ship's data. With these features the actual significant wave height, wave period and wave direction with respect to the ship can be established by making use of the models created in Chapter 6.

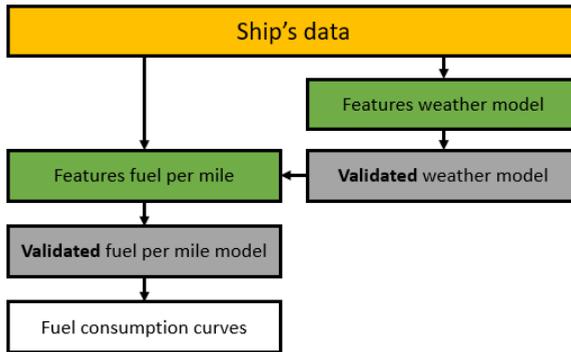


Figure 7.2: Description of the steps to establish fuel consumption curves from the ship's data influenced by weather conditions.

Second, the established ocean wave parameters will be used together with additional parameters from the ship's data to train the model used to establish the fuel consumption per mile. Further in this chapter the performance of this model and the additional parameters are evaluated.

Finally, with the additional features and the ocean wave parameters a model for every propulsion mode is trained able to establish the fuel consumption per mile. One of the input parameters is the speed over ground. The aim is to establish fuel consumption curves which are dependent on the speed over ground. Therefore it is desired to predict the fuel consumption per mile for a range of speed over grounds. To have a feature subset that is representing the reality, the features needs to be independent of the speed over ground. This makes it possible to establish the fuel consumption per mile for a range of speeds over ground keeping the other input parameters which are representing the environmental conditions as constants.

### 7.3. FEATURE SELECTION

**T**O set up fuel consumption curves we will train a model for every propulsion mode that is able to establish the fuel consumption per mile given the current environmental conditions. At the moment this model is trained we can set up the fuel consumption curves by holding the environmental conditions as constants and calculate the fuel per mile for different speeds over ground. To get accurate and reliable predictions, the features representing those conditions should be independent of the speed over ground.

The features that are representative for the environmental conditions and independent of the ship over ground chosen to train the model are;

#### OCEAN WAVE PARAMETERS

The ocean wave parameters inferred from the ship's data in Chapter 6 are used as features to establish the fuel consumption per mile.

The significant wave height and wave period are directly used, the wave direction is split up into the cosine and sine of the wave direction with respect to the ship. This

because of that the models do not recognize that 0 and 360 degree are equal values. However, the model will be able to evaluate the influence when for example the cosine is 1 (waves from ahead), the cosine is -1 (wave from astern) and the cosine is 0 (waves from aside).

#### SET AND DRIFT

The drift is coming from the data enrichment in Section 3.3, indirectly from the ship's data. The drift is a measure for the amount of current and split up into a component pointing into the direction of the ship's movement ( $D_y$ ) and perpendicular ( $D_x$ ) to this direction. A positive  $D_y$  means that the current is coming in reverse direction of the ship's movement.

#### TRUE WIND SPEED AND DIRECTION

The true wind speed is a feature coming from wind speed and direction measurement on board of the ship and so from the ship's data. The result of these measurements are the apparent wind speed and direction which are dependent on the speed over ground and the difference between the heading and course over ground (Set). Therefore we translate these apparent measurements into the true wind speeds in the direction of the ship's movement and perpendicular to this,  $TWS_y$  and  $TWS_x$  respectively. This translation is made in Section 3.3. A negative  $TWS_y$  means that the wind is coming in reverse direction of the ship's movement.

The wind force is dependent on the projected area on which the wind is directed. The projected area is dependent on the wind angle, if the wind is coming from astern this area is smaller than wind from aside.

#### MEAN DRAFT

To add the effect of loading conditions on fuel consumption per mile, the mean draft is assumed to be a good feature. The loading condition of the ship is expected to have major influence on the fuel consumption per mile. During a trip the loading condition changes because bunker fuel will be consumed and leaves the ship through the exhaust pipe as gasses. Also, a major difference in the loading conditions occurs at the moment that the helicopter takes off.

The draft is for a major part affecting the wetted surface of the hull and so the frictional resistance, an increase in draft means an increase of the frictional resistance and so more fuel consumption. The mean draft is given in Equation 3.20, where  $T_f$  is the draft at the bow of the vessel in meters and  $T_a$  the draft at the stern of the vessel.

#### TEMPERATURE AND HUMIDITY

The air temperature and humidity affect the charge air to the diesel engine and so the fuel consumption of the diesel engine. Therefore these two parameters should be considered as important features. The brake specific fuel consumption increase with both the temperature and humidity of charge air, [Lin and Jeng \(1996\)](#).

### 7.4. TRAIN AND TEST MODEL

**T**O train the multi-layer perceptron and tune the hyperparameters we will make use of the grid search combined with KFold cross validation described in Chapter 5. Figure

7.3 gives an overview of the input and output of the models that will be trained in this chapter. An detailed description of the features used is given in Section 7.3.

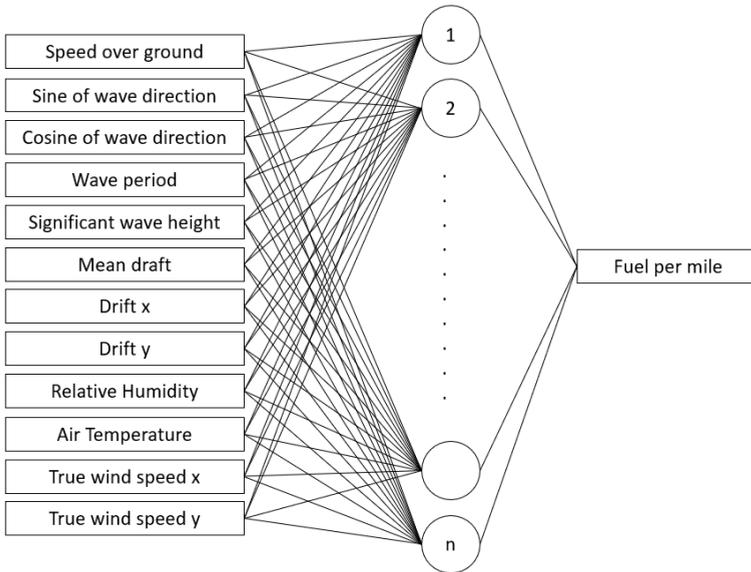


Figure 7.3: Description of the model for the prediction fuel per mile. Features are given in the rectangle boxes, the number of hidden layers is one and the number of neurons in the hidden layer equal the number of features.

For every propulsion mode a model is trained able to determine the fuel consumption per mile. From the extensive grid search it turned out that the combination of the stochastic gradient descent (SGD) solver, Rectified Linear Unit (ReLU) activation function, a learning rate of 0.01 and one hidden layer with the number of neurons equal to the number of features give the most accurate results for the modes. The resulting accuracy's of the models trained to establish the fuel consumption per mile given the input parameters are summarised in Table 7.1.

Table 7.1: Accuracy's of the trained models for the prediction of fuel consumption per mile for different propulsion modes.

	MAE [kg/mile]	Relative MAE [%]
<b>Manoeuvring</b>	3.631	1.551
<b>Transit</b>	2.888	2.664
<b>Blocked shaft SB</b>	3.294	1.359
<b>Trailing shaft SB</b>	2.101	2.080

The choice of hidden layer is investigated in more detail for the manoeuvring mode as shown in Figure 7.4. In this figure the influence of the hidden layer size on the fuel consumption curves for certain environmental conditions is visualised. Obvious is that increasing the number of hidden layers makes the model highly sensitive for feature

combinations not available in the training data. The models with 2 and 3 hidden layers seem to be over fitted, which means that the models are trained to rigidly on the combination of input and target values and therefore is not able to give accurate results for combinations of input values not in the training data. This can be enhanced by the fact that humps in the fuel consumption curves appears at the areas with less data, such as the areas in between the constant rotational speed of the propeller lines as shown in Figure 7.1. To overcome the risk of overfitting we will use one hidden layer.

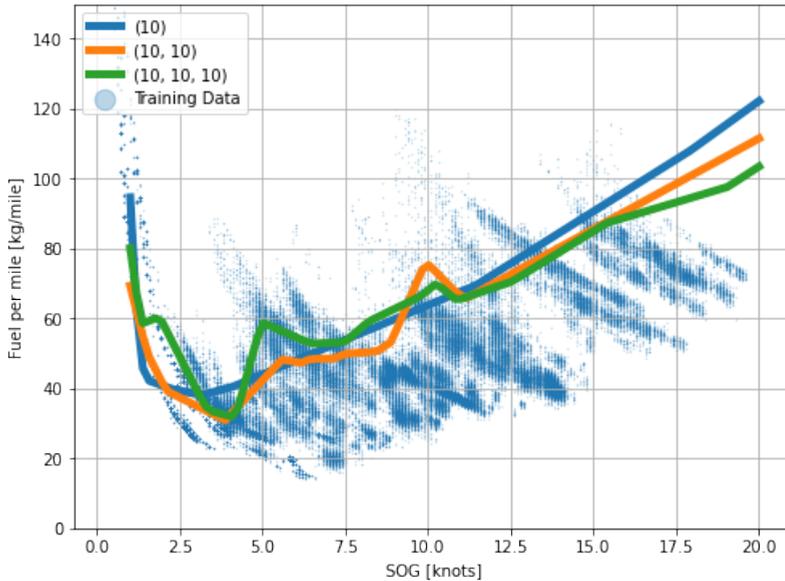


Figure 7.4: Establish the number of hidden layers by visualizing the fuel consumption curves obtained from models consisting of a different hidden layers size. Obvious is that the more hidden layers the higher sensitive model output.

## 7.5. MODEL OUTPUT : FUEL CONSUMPTION CURVES

**I**N this section the performances of the developed models are evaluated. At first the comparison is made between the resulting fuel consumption curves for the propulsion modes in varying environmental conditions. Thereafter we will visualise the impact on the fuel consumption curves for varying environmental conditions for the manoeuvring propulsion mode only. Finally, the fuel consumption curves as obtained from a complete physical ship model, [Geertsma et al. \(2017\)](#), are compared to the resulting fuel consumption curves from the created model.

### 7.5.1. COMPARISON PROPULSION MODES

A comparison between the fuel consumption curves belonging to the propulsion modes in varying environmental conditions is made to help the operator of the vessel select optimal settings regarding to the optimal speed and propulsion mode. The comparison

is made for 5 predefined environmental states, weather state 0, 1, 2, 3 and 4, these states are listed in Table 7.2. Those environmental states are based on the available data, namely the significant wave height is varying from 0.2 to 2.6 meters. In fact, the current is not related to the weather conditions however we will use the values as listed in the table for the comparison.

Table 7.2: Established median values of the weather parameters for the defined weather states.

WS [#]	swh [m]	Wave Period [s]	$T_{mean}$ [m]	TWS [km/h]	Current [knots]
<b>0</b>	0.5	3.0	4.44	5.0	0.0
<b>1</b>	1.0	3.7	4.44	10.0	1.0
<b>2</b>	1.5	4.4	4.44	15.0	2.0
<b>3</b>	2.0	5.1	4.44	20.0	3.0
<b>4</b>	2.5	5.8	4.44	25.0	4.0

The comparison between is made between the manoeuvring, transit, blocked shaft starboard and trailing shaft starboard propulsion modes. Because the available data is not consisting of sufficient records for the PEM propulsion mode, this mode is left out of consideration.

Figure 7.5 gives a complete overview of the fuel consumption curves created by the model for the propulsion modes. From left to right the direction of the weather is varying from ahead, aside and astern respectively. From up to down the weather state is increasing from weather state 0 to 4 respectively, as described in Table 7.2. The solid colored lines represent the fuel consumption curves for the different propulsion modes from the model given the input conditions. The colored scatter points in the background of the figure give the records from the dataset belonging to the environmental states in the figure and the propulsion mode. These colored scatter points represents the available data for the environmental state and propulsion mode whereby the models are trained.

The overall trend of the curves seems to be logical and can be explained. In the situation that the ship encounters wind, waves and current from ahead it can be observed that the fuel consumption per mile increases for heavier weather conditions. At the moment where wind, waves and current are coming from astern the fuel consumption curves show a downward movement for heavier weather conditions. This is logical because of the fact in this situation the current and wind are pointing into the positive direction of the ship's movement and so less thrust is required to reach the same speed over ground. Wind, waves and current from aside will lead to a higher fuel consumption per mile for heavier weather, however the increase in fuel consumption per mile is less than for the situation where the weather is coming from ahead. In this situation the ship has to compensate the weather from aside by increasing the set of the ship to maintain the course over ground. Therefore, the thrust delivered by the propeller is not directly pointing in the direction of the speed over ground and so more thrust is required to reach the desired speed over ground.

The scatter points in the background of the figure are the records from the data in accordance with the environmental state, weather direction and the propulsion mode on which the models are trained. It can be seen that, for example in weather state 4, the ship has not sailed with the trailing and blocked shaft propulsion mode. Therefore the

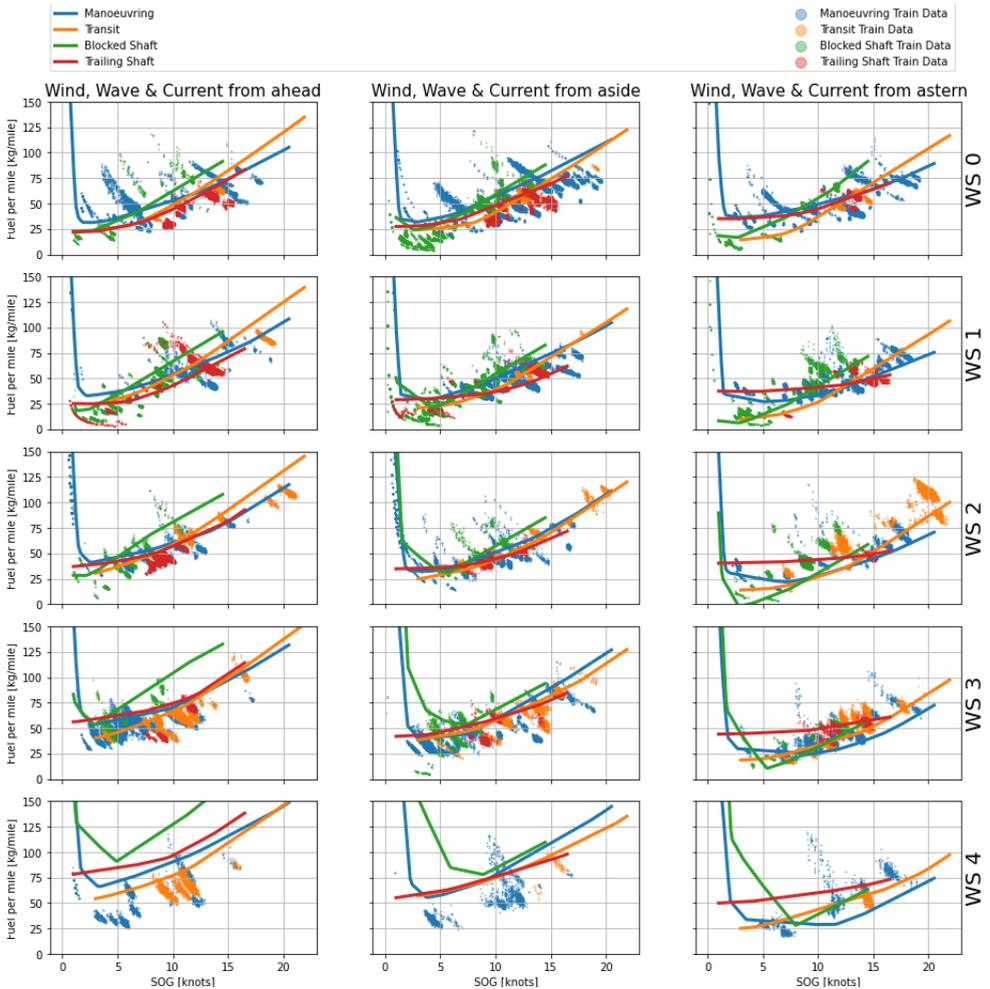


Figure 7.5: Comparison between the resulting fuel consumption curves from the models for varying environmental conditions and propulsion modes. From left to right the direction of the weather is varying from ahead, aside and astern respectively. From up to down the weather state is increasing from weather state 0 to 4 respectively as described in Table 7.2.

resulting fuel consumption curves are extrapolations created by the model. To validate these extrapolations it is recommended to acquire more data for every propulsion mode in every weather state and direction.

There are made various extrapolations by the model for the creation of fuel consumption curves. The fuel consumption curves that are established by the model for weather states and directions not available in the dataset are extrapolations and difficult to validate with the available dataset.

Also, the expected upward movement of the fuel consumption per mile at lower speeds is not visible for the transit and trailing shaft mode. This can be declared by the

fact that the data whereby the model is trained is not consisting of this upward movement, see Figure 4.13. To take into account this upward movement and so develop a model that is able to predict the fuel consumption per mile accurate in every weather state and direction and for every speed it is important that the data is consisting of all possible scenario's.

Despite the lack of validation of the extrapolated curves due to the amount of data we can see that the fuel consumption curves for the manoeuvring propulsion mode are more in line with the expectation. The records in the data belonging to the manoeuvring propulsion mode are consisting of the greatest amount of records and most varying conditions. Therefore, at the moment more data is acquired it is assumed that the fuel consumption curves for the other modes will be more accurate too. If accurate fuel consumption curves are available it should be possible to present the fuel consumption curves real time to the crew and so they could make well informed decision about the choice of propulsion mode and speed setting.

### 7.5.2. COMPARISON ENVIRONMENTAL CONDITIONS

The manoeuvring propulsion mode is used to investigate the performance of the model in more detail. As can be seen in Figure 4.2 and 7.5 the data contains the greatest amount of records and variety for the manoeuvring mode and therefore the model trained for this mode delivers the most accurate fuel consumption curves.

In Figure 7.6 the fuel consumption curves for the weather states are given on top of each other derived from Figure 7.5 for the manoeuvring mode only.

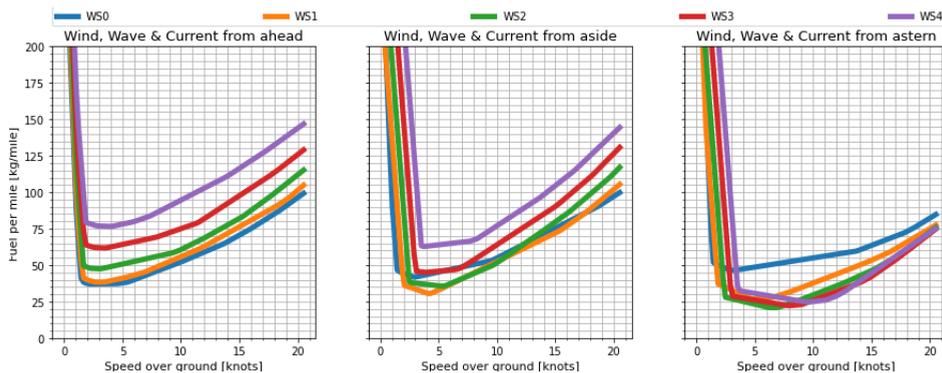


Figure 7.6: Comparison between the fuel consumption curves belonging to the manoeuvring propulsion mode. From left to right the weather direction is ahead, aside and astern respectively. The different weather states are listed in Table 7.2.

The optimal speed over ground, the speed with the least fuel consumption per mile, is varying for the environmental conditions and the direction of weather. In the scenario that wind, waves and current encounter the ship from ahead it can be seen that the optimal speed is between 3 and 5 knots and is not affected by the weather state. In the scenario that wind, waves and current encounter the ship from astern the optimal speed is highly affected by the weather state. For weather state 0, the optimal speed is around 3 knots and the increase in fuel per mile for higher speeds is lower than in higher weather

states. For weather state 4, the optimal speed is around 10 knots and the increase in fuel per mile is high for lower or higher speeds over ground compared to weather state 0.

### 7.5.3. COMPARISON PHYSICAL MODEL

For validation of model a comparison is made with the fuel consumption curves obtained from a complete physical ship model created by [Geertsma et al. \(2017\)](#), as shown in Figure 7.7. The fuel consumption curves from the physical model are given for the trail, design and off-design conditions for the manoeuvring propulsion mode. The trail conditions are defined as Sea State 0, wind speed of 3 m/s and no fouling. Design condition is fined as Sea State 4, wind speed of 11 m/s, head seas and wind and 6 months out of dock fouling. Off-design conditions is defined as Sea State 6, wind speed of 24 m/s, head seas and wind and 6 months out of dock fouling.

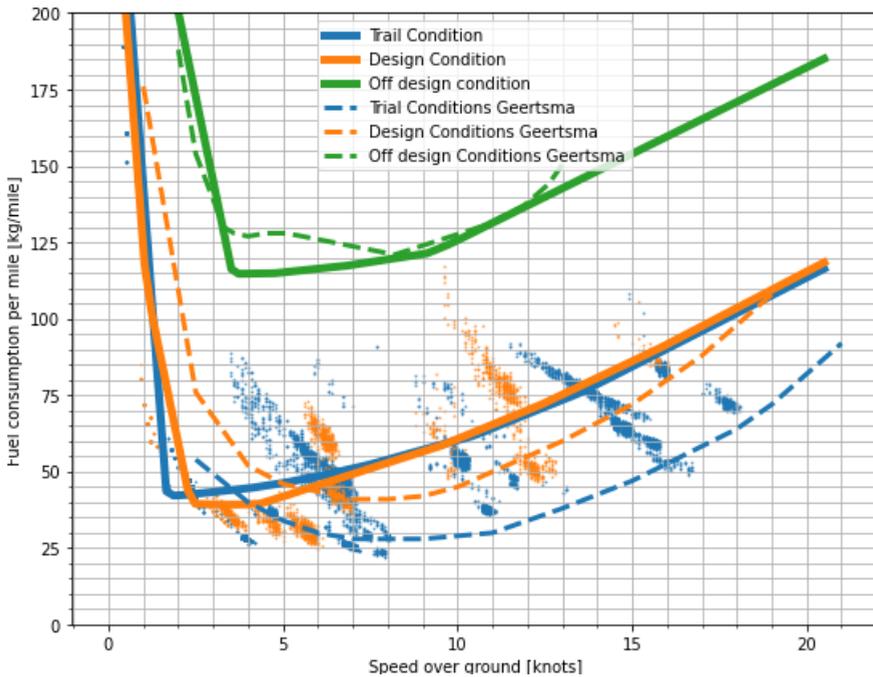


Figure 7.7: Comparison between the fuel consumption curves obtained from the physical ship model, [Geertsma et al. \(2017\)](#), and the fuel consumption curves resulting from the machine learning model for train, design and off-design conditions for the manoeuvring propulsion mode. The scatter points are representing the records from the data containing these conditions.

The solid lines in the figure represent the model outputs for the defined conditions, the striped lines give the fuel consumption curves from the physical ship model and the scatter points are representing the records from the data containing these conditions.

The fuel consumption curves determined by the machine learning model for the manoeuvring propulsion mode are approximately and in accordance with the points from the data for the trail and design condition. The off-design conditions is not present in the data and therefore the fuel consumption per mile for this condition is an extrapolation

established by the model. As can be seen in the figure, the optimal speed over ground for the trail and design condition is not in accordance with the optimal speed setting of the physical model. The fuel consumption curve for the off-design conditions is in fact an extrapolation made by the model, however it should be noted that this curve is roughly in accordance with the curve obtained from the physical model.

## 7.6. CONCLUSIONS

**I**N this chapter a methodology is proposed how fuel consumption curves can be established by the ship's data for the propulsion modes and current environmental conditions. The weather conditions are established by the ship's data with the models proposed in Chapter 6. The research performed in this Chapter leads to the answer on Research Question 3;

*How can ship's data help the operator select optimal settings to reduce CO<sub>2</sub> emissions by making use of machine learning models?*

In the Figures 7.5, 7.6 and 7.7 a speed setting in accordance with a minimum fuel consumption per mile is visible. By giving the operator/crew of the ship real-time insight into the fuel consumption curves for every propulsion mode corrected for the environmental conditions the operator can make well informed decisions regarding the speed setting and so reduce CO<sub>2</sub> emissions. However, the data currently available is not sufficient enough to set up fuel consumption curves for the entire spectrum of propulsion modes, speed over grounds and environmental conditions.

The balance/distribution of values in the dataset is highly influencing the performance of the machine learning model. As we can see in Figure 7.1 and 4.3 the ship is sailing for a major part with predefined RPM-settings resulting in areas with a high density of data points and areas with almost any points. Figure 7.4 shows the sensitivity of the model to this unbalanced data.

In Section 7.4 we investigated the performance of the model created to predict the fuel per mile by the determined weather conditions and additional parameters from the ship's data. The mean absolute error of the models is between 2.101 and 3.631 kg/mile and the relative mean absolute errors between 1.359 and 2.664 %. With these results we can conclude that the models are able to establish the fuel consumption per mile with high accuracy based on the selected features and the records in the data.

Thereafter, in Section 7.5 we evaluated the model output, the fuel consumption curves, for different environmental conditions and propulsion modes. Figure 7.5 visualises the fuel consumption curves obtained from the machine learning models for the propulsion modes and varying environmental conditions. It can be seen that the upward movement at low speeds is not present in the dataset as well as in the resulting fuel consumption curves for the trailing shaft and transit mode. The overall trend of the fuel consumption curves seems to be logical, see Figure 7.6. For the scenario where the environmental conditions are coming from ahead, the fuel consumption curves show an upward trend for an increasing weather state. For the scenario where the environmental conditions are coming from astern, the fuel consumption curves show a downward trend for an increasing weather state.

Therefore, we conclude that the proposed approach should be able to establish fuel consumption curves, however the dataset currently available is not sufficient to give a complete overview of fuel consumption curves in all possible scenarios the ship can encounter. To apply the proposed approach on board of ship's a dataset is required that contain all possible environmental conditions, in every propulsion mode and at every speed.

Also, validation of the sensors on which the fuel consumption per mile is determined should be performed. As can be seen in Figure 7.7, the fuel consumption curves are roughly in accordance with the results obtained from the physical model. However to make a justified statement about whether the measurements are a representation of the reality, it is recommended to perform sensor calibration.

# 8

## CONCLUSIONS AND RECOMMENDATIONS

The novelties in this thesis are: first, the extensive ship's data consisting of multiple data types is combined with environmental data by reconstructing the position of the ship. Thereafter, a model is constructed trained with weather data to establish the environmental conditions by the ship's data enriched with the amplitudes and periods of the roll and pitch motion and the fuel rack position of the diesel engines. Finally, we construct a model for the determination of the fuel consumption per mile and set up fuel consumption curves for different propulsion configurations accounted for the current environmental conditions. Because we only use the ship's data, those fuel consumption curves will be suitable to be used by the crew on a real time base to optimize the energy efficiency during the operation and thus reducing  $CO_2$  emissions. This research was guided by the following main Research Question:

*"How can  $CO_2$  emissions be reduced given the operational profile of a naval vessel by giving real-time insights to the operator based on ship's data?"*

This thesis has shown that  $CO_2$  emissions can be reduced by selecting the most energy efficient propulsion mode and speed for different environmental conditions. From Figure 7.5 we can conclude that the most energy efficient propulsion mode differs for varying environmental conditions and speeds. From Figure 7.6 we can conclude that the optimal speed, given a propulsion mode, is different for varying environmental conditions. Using a multi-layer perceptron model, the environmental conditions can be inferred from the ship's data with an accuracy of 4.150 % as shown in Table 6.1. This makes it possible to give the crew real-time insight into the fuel consumption curves for the various propulsion modes, without the use of external data sources. This helps the crew to make well informed decisions about the most energy efficient propulsion mode and speed setting while still being able to perform their operations.

However, due to the lack of a wide variety and amount of environmental conditions and propulsion modes in the dataset and the lack of sensor calibration, further research is recommended. A wide variety of environmental conditions for all propulsion modes and speed settings in the dataset and a more balanced dataset is required to improve and validate the fuel consumption curves for the entire spectrum of speeds, environmental conditions and propulsion modes. Calibration of the sensors is required to be sure that the measurements are the actual representation of the reality.

## 8.1. DISCUSSION ON RESULTS

THE ship's data consists of an enormous amount of parameters and records, as well as the weather data. To process the data with the goal to reduce  $CO_2$  emissions by giving the operator real-time insight in the  $CO_2$  performance of the ship we performed research in order to answer Research Question 1:

*How can  $CO_2$  performance be established given the operational profile of a naval vessel with data?*

From the research performed in Chapter 3 and 4 we can conclude that the  $CO_2$  performance of the ship can be established by making use of fuel consumption per mile derived from the ship's data. Because we do not measure  $CO_2$  directly, complete combustion is assumed. And so, the fuel consumption per mile gives insight into the  $CO_2$  emission per mile. The fuel consumption per mile is influenced by the technical, environmental and loading conditions of the ship and different for the propulsion modes as shown in Figure 4.13.

Using the ship's data to establish the  $CO_2$  performance is a step-wise development process and this thesis contributes to this. The Royal Netherlands Navy has a discreet policy to handle ship's data which means that the latitudinal and longitudinal position of the ship can not be disclosed. Therefore combining historical weather data and the ship's data by reconstruction the route with the speed over ground, course over ground and the initial position is a new concept. This new concept makes further research possible where influences of weather conditions plays a key role, for example research towards fatigue of the hull or the take off moment for the helicopter.

The fuel consumption per mile exhibit a drastic increase at lower speeds. This drastic increase arises from the decrease in the total efficiency of the propulsion system as demonstrated in Figure 4.15. The total efficiency, consisting of the propulsive, gearbox and engine efficiency, referred as the technical condition, of the propulsion system is affected by the efficiency of the individual components of the system. The technical conditions of the ship decay over time because of degradation in the components. The dataset is consisting of 6 voyages all undertaken in one month, and therefore we assumed to ignore the degradation of technical conditions over time.

When the ship is driven by one or both of her main diesel engines, the fuel consumption of the engines is the determining factor for the  $CO_2$  performance of the ship. The loading of the main diesel engines influences the fuel consumption given as the specific fuel consumption per kilowatt hour. Chapter 4 gives insight into the loading of the diesel engine by propeller load curves. These propeller load curves are established by

the torque and rotational speed measurements in the data together with the modelling of the gearbox efficiency. The results could be used to alter the combinator curves or to take this into account by designing the propulsion plant of new build ship's to reach a more  $CO_2$  friendly operation. Also, the information gathered is of importance for mechanical engineers responsible for the operation and maintenance of the main diesel engines. Identification of low load operation of the diesel engine is possible and low load operations is considered to be a major cause in diesel engine failure.

The resistance of the ship derived from the thrust of the propeller in Chapter 4 is approximately in accordance with the resistance as extrapolated from model tests. The thrust of the propeller is calculated by making use of the open water diagram and measurements of speed through water, rotational speed and pitch of the propeller as measured by sensors. The open water diagram represents the ideal performance of the propeller, this means unaffected by marine fouling or a damaged propeller. As a result, the determination of propeller thrust is depending on uncertainties in the state of marine fouling or damage and the accuracy of the sensors.

It can be concluded that the fuel consumption per mile is highly dependent on the weather conditions the ship encounters. Therefore, to determine the actual fuel consumption per mile corrected for these conditions real-time weather conditions are required. The ship's data is real-time and therefore, in Chapter 6 we investigated the relation between the ship's data and the weather conditions in order to answer Research Question 2:

*How can real-time weather conditions be inferred from ship's data?*

Chapter 6 demonstrated the determination of ocean wave parameters by making use of the ship's data. The creation of additional features, the amplitude and period of the roll, pitch and fuel rack position leads to accurate performance of the model.

The validation and testing process of the created algorithms show credible results, however we have to take into account the occurrence of overfitting. To validate whether the model is predicting the actual significant wave height, wave period and wave direction properly it is recommended to validate the model with observations/measurements of those parameters with a higher resolution.

By making use of the data analysis method described in Chapter 5 we can conclude that the real-time weather conditions can be established from the ship's data with an accuracy of around  $\pm 4\%$ . Now, real-time fuel consumption curves can be established corrected for those real-time weather conditions, and so the historical, with a low resolution and hourly averaged weather data is no longer required. Finally, we can perform research in order to answer Research Question 3 with the acquired results:

*How can ship's data help the operator select optimal settings to reduce  $CO_2$  emissions?*

From Chapter 7 we can conclude that the fuel consumption curves clearly show an optimal speed setting. With the research done previously we can give insight to the operator by showing the fuel consumption curves corrected for the actual weather conditions. The operator still has the control over the speed of the ship and so can perform

the required operations belonging to the operational profile of the ship. At the time the operator is not bounded to a certain speed over ground due to the operational profile the choice can be made to alter the speed settings and so reduce  $CO_2$  emissions.

However, as discovered in Chapter 7 the balance of the data is highly influencing the performance of the machine learning model. Unbalanced data will lead to a model that is highly focused on certain points, such as the predefined RPM-settings. Therefore we have chosen to use one hidden layer in the multi layer perceptron model.

The methodology to establish fuel consumption curves proposed in this thesis are of importance for weather routing problems also. Weather routing is an optimization problem with the lowest fuel consumption as optimization function in most cases. To reach this goal the influence of weather conditions on the fuel consumption of the ship is of importance. Most research performed in this area is applied to liner shipping as case study, the data used is low frequent and sufficient for the problem because of the highly continuous speed settings and operations of liner shipping, from A to B in a certain time. The methodology proposed in this thesis makes it possible to apply this problem to ship's with a more varying operational profile.

It should be noted that the work done in this thesis is applicable to the specific ship. However, the methodology described in this thesis is applicable to every ship type that is equipped with the required sensors used in this thesis. Therefore, this work can be easily extended and used by other types of ship's to attain insight into their  $CO_2$  performance and so being able to make reasoned decisions regarding  $CO_2$  emissions.

## 8.2. RECOMMENDATIONS FOR FURTHER WORK

**I**N this section recommendations for further work are given. These recommendations are consisting of increasing the data variety, making use of additional sensor measurements, make the combination with physical modelling, acquire accurate weather data and perform sensor calibration.

### 8.2.1. DATA VARIETY

The proposed methodology should be extended by using more variate data to set up fuel consumption curves for all propulsion modes in more various environmental conditions. The ship's crew is able to make well informed decisions about  $CO_2$  dilemma's only when the required information is available to them. The required information will be a complete overview of fuel consumption curves for the different propulsion modes corrected for the actual environmental conditions.

Currently, the weather circumstances in the dataset are around sea state 1 till 4 as shown in Section 3.2. Also, Figure 4.2 gives the distribution of records in the data for the different propulsion modes relative to the amount of records in the raw data. Together with the demonstrated filter to drop out the acceleration and decelerations of the vessel the available data is consisting of only 10300 records which is 24 % of the total dataset approximately.

To investigate the inter- and extrapolating capacities of the proposed machine learning model in more detail, data for validation is required. Therefore the recommendation for further work is to acquire data for every propulsion mode, in every propulsion mode

all possible environmental conditions and in every propulsion mode and all possible environmental conditions the whole range of speed settings.

### 8.2.2. COMBINE WITH PHYSICAL MODELLING

A combination with physical modelling could be made to investigate validate the inter- and extrapolating of the proposed machine learning model also. Physical modelling will have high inter- and extrapolating capacities. Also, a physical model is able to address the influences of the components inside the propulsion system. In this thesis a machine learning method is used where the relation between the input and output is difficult to interpret.

However, physical modelling is the development of a simplified representation of the reality where data from sensor measurements represent the actual reality. Therefore the recommendation is to combine these two approaches by creating a physical model based on relations in the data and tune the parameters with actual records from the data. In this way the inter- and extrapolating capacity and the clarity of the physical model are maintained and the physical model is a representation of the actual reality.

### 8.2.3. ADDITIONAL SENSOR MEASUREMENTS

Making use of additional sensors can give insight into the technical conditions of the ship, required at the moment when the methodology will be used with a dataset consisting of a longer time span.

A thrust measurement sensor directly measures the thrust produced by the propeller and so a more accurate resistance calculation can be made. Uncertainties arising from the calculation of thrust from the static open water diagram and the measured rotational speed and speed through water will disappear. With accurate thrust measurements it is capable to quantify fluctuations in thrust which is assumed to be suitable to infer the environmental conditions more accurate and a decrease in propeller performance over time can be used to identify propeller fouling or a damaged propeller. Also, a more accurate calculation of the ship's resistance can be made and over time this can give insight into the state of marine fouling on the hull. And so, thrust measurements are highly relevant in setting up more accurate propeller and hull cleaning policies or identification of a damaged propeller.

Also, torque measurements on multiple places in the propulsion train can give insight into the different torque/power losses of specific components. With a torque measurement at the output shaft of the engine the brake power is measured unaffected by gearbox losses and so give more accurate insights into the diesel engine loading. Also, with torque measurements at both sides of the gearbox and bearings insight can be obtained into the actual losses and degradation of these components.

### 8.2.4. ACCURATE WEATHER DATA

The models developed in Chapter 6 have a credible predictive capacity for the weather data averaged over one hour by the ship's data. However, to gather insight into the influence on the fuel consumption per mile by the actual weather conditions the ship encounters, the models need to be trained on more accurate and higher resolution weather data. To reach this, more accurate and higher resolution weather data needs to be present

in the data.

Possible ways to acquire this can be by operating close to a wave buoy and obtaining high frequent measurements or by using additional radars. When the more accurate weather conditions are established the steps described in this thesis can be performed and the prediction of the accurate weather conditions could be done.

### **8.2.5. SENSOR CALIBRATION**

In this thesis we did not take into account sensor calibration, because no information is available about the calibration. Therefore, it is recommended to perform sensor calibration in further research to make a justified statement about the representation of reality of measurements.

With sensor calibration is meant that the accuracy and the precision of the sensors used are investigated. The accuracy of a sensor gives an indication of the difference between the mean of the measured values by the sensor and the true value. The precision of a sensor is an indication of the fluctuation in the differences between the measured values and the mean of measured values.



# LITERATURE SEARCH AND ANALYSIS

THE goal of this section is to find literature relevant to make the research question more precise and to define the sub questions. At first, the search towards relevant research is done based on various relevant search terms. Various databases consisting of journals, articles, books, dissertations and master theses are applied to search for relevant research, these are listed below. Secondly, the found literature that seems to be relevant is analysed by identifying the problem, key concepts and strengths and weaknesses.

- *IEEE Xplore*, this database is the leading academic database in the field of engineering and computer science. It's not only journal articles, but also conference papers, standards and books that can be search for.
- *ScienceDirect*, this database contains over the 12 million contents of scientific and medical research published by Elsevier out of scientific magazines and e-books.
- *Scopus*, this database uniquely combines a comprehensive, curated abstract and citation database with enriched data and linked scholarly content.
- *Repository TU Delft*, this data base contains all scientific documents that are produced by TU Delft researchers and students.

## RELEVANT SEARCH TERMS

To find the relevant literature in the mentioned databases we will set up different search terms. The main research question as given in the introduction is split into key concepts that are used to search for relevant research. The key concepts from the main question are; operational ship data, machine learning and reduced  $CO_2$  emissions. During the search to relevant literature these key concepts are modified multiple times to end up with more relevant literature. The key concept machine learning is added with the term

'shipping' because of the fact that machine learning is an application in multiple domain. Also, the key concept reduced  $CO_2$  emissions is added with this term and this key concept is formulated as 'reduce fuel consumption shipping' as well.

### RELEVANT RESEARCH

In this section the found literature by the relevant search terms that seems to be relevant are analysed. This analysis is performed to summarize this literature by their problem, key concepts and strengths and weaknesses. The next section is based upon this literature and compare them to identify the different gaps in this literature.

**Probabilistic modeling of ship powering performance using full-scale operational data** [Yoo and Kim \(2019\)](#) developed regression models that can be used to predict the ship speed and engine power under different operational settings and weather conditions. In their study, they employed machine learning techniques to estimate the powering performance of a full-scale ship by constructing regression models using the ship's operational data. In order to minimize the risk of over fitting in the regression process, domain knowledge based on physical principles is combined into the regression models. The domain knowledge used is about the basic ship propulsion principles consisting of the propeller thrust and torque from the thrust and torque coefficients in combination with the propeller speed and the engine power as a result from the torque of the propeller. The normalized root mean squared error of the proposed models are in between 2.22 ad 3.29 %.

**Propulsion performance of large-scale ship model in real sea environment,** [yu Guo et al. \(2020\)](#) studied the influence of waves on propeller performance on a large-scale model, using equal thrust and equal torque methods. They built a model of a bulk carrier with a length of 25 m, and conducted a free-running test in coastal waters. The independent research and development of large-scale high-precision broken-shaft self-propulsion instruments is an important technological breakthrough in measuring propeller propulsion performance. The ultimate purpose of collected data from the tests is to explore the scale effect and conversion relationship on propulsion performance between the large-scale model and full-scale ship, and then guide the performance design of full-scale ship. However, this research work is limited to the propulsion performance of the large-scale model because of the lack of full-scale ship's measured data.

**Comparative analysis between different methods for calculating on-board ship's emissions and energy consumption based on operational data,** [Moreno-Gutiérrez et al. \(2019\)](#) compares four existing methods for calculating energy consumption and emissions, and presents a more realistic method, based on a case study. The purpose is to examine the differences between all of these methods, in order to propose the most suitable method of obtaining the data needed for better energy management, and a method that can be applied to any type of ship. The application of this method does not, in itself, reduce fuel use or improve efficiency, but it should be the necessary first step to establish uniform operational measures that will improve the management of energy on board ship and monitor accurately the performance of the fleet. By applying the method of the au-

thor the energy consumption and emissions produced for each operational phase of the ship can be calculated in real time by its operator. This information give operators better control, and improve management of energy used on board their ships. This method is based on data from the noon-reports of a Ro-Pax and no AIS data was used because the AIS system has many uncertainties.

**Estimation of ship operational efficiency from AIS data using big data technology**, [Kim et al. \(2020\)](#) propose a method to estimate the Energy Efficiency Operational Indicator (EEOI) without requiring the actual Fuel Oil Consumption (FOC). The Automatic Identification System (AIS) data, ship static data, and environment data that can be publicly obtained are used to calculate the EEOI. Since the public data are of large capacity, big data technologies, specifically Hadoop and Spark, are used. They verify the proposed method using actual data, and the result shows that the proposed method can estimate EEOI from public data without actual FOC.

**Speed optimization of a container ship on a given route considering voluntary speed loss and emissions**, [Li et al. \(2020\)](#) propose a speed optimization model of a 4250 TEU container ship in which the influence of sea states and voluntary speed loss on ship sailing are considered with the objective of minimizing the main engine fuel consumption and the ship operating costs. This study provides a practical and feasible method for ship speed optimization, which can be used as a reference by shipping companies. However, due to the confidentiality of shipping company data, this study lacks ship hull lines and actual navigation data, so the analysis presented herein is limited by the lack of detailed hydrodynamic calculations and ship motion calculations in wind and waves. Speed optimizations of multiple ships and multiple routes in which hydrodynamics and ship motion theory are used in the optimization model to calculate involuntary and voluntary speed loss is recommended as the focus of future research.

**Fuel usage data analysis for efficient shipping operations** [Trodden et al. \(2015\)](#) describes a methodology for isolating from a continuous data-stream periods associated with steady-state free-running condition, and transient free-running conditions. Further analysis in this research is performed to determine the fuel efficient performance of the ship. Analysis of the steady-state free-running condition can be used to assess degradation in performance over time.

**Grey-box modeling of an ocean vessel for operational optimization** [Leifsson et al. \(2008\)](#) developed a grey-box modeling approach for the simulation of ocean vessels. The modeling approach combines conventional analysis models based on physical principles (a white-box model) with a feed forward neural-network (a black-box model). The results of simulating several trips of a medium sized container vessel show that the grey-box modeling approach, both serial and parallel approaches, can improve the prediction of the vessel fuel consumption significantly compared to a white-box model.

**Ship efficiency forecast based on sensors data collection: Improving numerical models through data analytics** [Coraddu et al. \(2015\)](#) investigates the problem of prediction

the fuel consumption of a vessel in real scenario based on data measured by the on board automation systems. They exploited three different approaches: White, Black and Grey Box models. The White-Box model neglects some aspect of the phenomena that are not easy to take into account, e.g. weather conditions. Black-Box modeling show a remarkably improvement, this method needs a large amount of historical data. The authors proposed a Gray-Box model able to exploit both mechanistic knowledge of the underlying physical principles and available measurements. Results shows that the Gray-Box model can achieve level of performance that are close to the one of the Black-Box model but requiring less historical data.

**Vessels fuel consumption forecast and trim optimisation: A data analytics perspective** Coraddu et al. (2017) investigates the problems of predicting the fuel consumption and of providing the best value for the trim of a vessel in real operations based on data measured by the on board automation systems. They have shown how data driven, or black-box, models can outperform state-of-the-art numerical, or white-box, models which exploits the physical knowledge of the system in the task of predicting fuel consumption. Based on these models new approaches for modelling the system have been developed, namely the grey-box models, which are able to exploit the advantages of two philosophy: gray-box models are able to obtain the same performance of the black-box but requiring less historical data thanks to the knowledge embedded in the white-box models. The authors have been able to propose a trim optimisation technique which exploits the predictive power of the proposed models for the online selection of the best configuration of the trim for reducing the fuel consumption.

**Ship operational performance modelling for voyage optimization through fuel consumption minimization** Chaal (2018) compared in this thesis different black-box models to predict ship fuel consumption, which depends on the ship specific and the navigational input parameters. The research is about finding the best predictive model to use in a decision support system for energy efficient ship operation, it can be confidently concluded that the developed methodology is a promising direction.

**The development of a ship performance model in varying operating conditions based on ANN and regression techniques** Farag and Ölcer (2020) developed a ship performance model in varying operating conditions based on ANN and regression techniques, an estimation of the B.S.F.C. and the ship's power under varying sea conditions. To the best of the authors' knowledge, highly frequent data was used one of the first time to model and predict the ship's performance. The input parameters for the ANN model are; Ship's speed, Seawater depth, Apparent wind speed on ship's direction, waver parameters (3), swell parameters (3) and the sea current (2). The performance prediction result of the proposed ANN-MR model shows a high total fuel consumption precision with 99.6 % accuracy for the overall voyage estimation.

**Machine learning models for predicting ship main engine Fuel Oil Consumption : A comparative study** Gkerekos et al. (2019) performed a study that presents a comparison of data-driven, multiple regression algorithms for predicting ship main engine FOC

considering two different data acquisition strategies, noon-reports and Automated Data Logging & Monitoring systems. The effectiveness of different regression algorithms is investigated based on the mean and median average error and the coefficient of determination ( $R^2$ ) including Support Vector Machines (SVMs), Random Forest Regressors (RFRs), Extra Tree Regressors (ETRs), Artificial Neural Networks (ANNs), and ensemble methods. ETR and RFR models were found to perform best in both cases, whilst the existence of an ADLM system increased accuracy by 7 % and reduced the required period for data collection by 90 %.

**An artificial neural network based decision support system for efficient operations** [Besikci et al. \(2016\)](#) developed an prediction method for the ship's fuel consumption for various operational conditions through an Artificial Neural Network and a decision support system employing ANN-based fuel prediction model to be used on-board ships on a real time basis for energy efficient operations. The fuel prediction model uses operating data - 'Noon Data' - which provides information on a ship's daily fuel consumption, the performance of the ANN is compared with multiple regression analysis and its superiority is confirmed.

**Development of a two-stage ship fuel consumption prediction and reduction model for a dry bulk ship** [Yan et al. \(2020\)](#) proposed a machine learning method performing regression task with high accuracy to make predictions on ship fuel consumption under different sailing speeds as well as cargo, weather and sea conditions. The second stage is a sailing speed optimization model based on the prediction results of the first stage. The data used in this work are provided by noon reports with a time resolution of one record every 24 h. The author described a possible way to make the fuel consumption prediction more precise by incorporating more data features from other data sources such as weather data and ship sensory data with a higher resolution.

**Estimating Fuel Consumption in Maritime Transport** [Isikli et al. \(2020\)](#) aim to construct a statistical model for fuel consumption rate in maritime transport considering several factors, the speed, rotational speed of the propeller, distance travelled, draught, cargo load, wind and sea waves. The author constructed several models for the fuel consumption rate and chose the best one considering several theoretical and practical issues. The study aims to help interested parties to find and efficient fuel economy with the control and alteration of some related factors. One of the limitations of such predictive models lies within the fact that any type of fuel consumption does not stay constant over time since it depends on the main engine load. Future research should also consider overcoming this problem using real-time data or high frequent data.

**Impact of trim on added resistance of KRISO container ship (KCS) in head waves: An experimental and numerical study** [Shivachev et al. \(2020\)](#) employs EFD and CFD and potential theory based methods to investigate ship motions and added resistance in regular head waves at six different trim angles. Numerical computations of ship motions and added resistance were validated against model experiments. The author's proposes to move from single point optimisation to multi point optimisation so that resis-

tance and added resistance at different trim angles at different loading conditions are taken into consideration. This can improve fuel efficiency and so reduced gas emissions throughout the vessel's operational life. Future study should be extended to include a rotating propeller to investigate the effect of trim on propulsive performance in seaway. Another interesting future study will be the extension of the numerical study to full scale in order to investigate scale effects on the optimum trim angle.

**Benefit of speed reduction for ships in different weather conditions** Taskar and Andersen (2020) studied the relationship between ship speed and fuel consumption. Also, the effect of speed reduction in different weather conditions has been studied. The aim is to assess whether mandatory speed reduction will reduce shipping emissions even after increasing the number of ships to keep transport work constant. The weather data used in this thesis is obtained from ECMWF Hersbach et al. (2021) and monthly averaged. They find that potential fuel savings by reducing the speed are highly dependent on weather conditions. However, it was observed that the additional fuel consumption for a voyage due to waves does not depend on ship speed.

**Penalty of hull and propeller fouling on ship self-propulsion performance** Song et al. (2020) investigated the effect of biofouling on the self-propulsion characteristics by applying CFD models to the self-propulsion characteristics of the full-scale KRISO container ship. The result suggests that hull fouling increases total resistance coefficients up to 52 % with the most severe surface condition, whereas the effect of propeller fouling on the total resistance coefficients is negligible. This study has provided several important findings including the effect of biofouling on the ship self-propulsion parameters and propulsive efficiencies, as well as flow characteristics around the hull. However, it is still challenging to correlate these specific hull and propeller fouling conditions with real ships operating in the oceans.

**A novelty detection approach to diagnosing hull and propeller fouling** Coraddu et al. (2019a) focused on building effective Data-Driven Models to predict the hull state condition utilising the data collected from an automation system when the vessel was in operation. The results show that with just a few samples from data collected during sea trials in the nominal state for the vessel, where the performances are unaffected by degradation due to hull and propeller fouling, it is possible to fine tune this methodology to achieve satisfactory performances in prediction of the hull state. Data obtained from Research Vessel (RV) The Princess Royal, a twin hull catamaran, operating mainly off the coast of Blyth, UK has been exploited to show the effectiveness of the proposed approaches and to benchmark them in a realistic maritime application.

**Data-driven ship digital twin for estimating the speed loss caused by the marine fouling** Coraddu et al. (2019b) build a digital-twin model to be tuned on data collected during a period of time where the marine fouling is not present and for a time period wide enough to observe the ship in many operational and environmental conditions. This digital-twin model is used to calculate the vessel's speed loss due to marine fouling. A comparison between the proposed method and the ISO 19030 on real-world data coming from two Handymax chemical/product tankers has been carried out and is presented

in this work.



# B

## WEATHER DATA DESCRIPTION

**10m u-component of wind** -  $ms^{-1}$ , This parameter is the eastward component of the 10m wind. It is the horizontal speed of air moving towards the east, at a height of ten metres above the surface of the Earth, in metres per second. Care should be taken when comparing this parameter with observations, because wind observations vary on small space and time scales and are affected by the local terrain, vegetation and buildings that are represented only on average in the ECMWF Integrated Forecasting System (IFS). This parameter can be combined with the V-component of 10m wind to give the speed and direction of the horizontal 10m wind.

**10m v-component of wind** -  $ms^{-1}$ , This parameter is the northward component of the 10m wind. It is the horizontal speed of air moving towards the north, at a height of ten metres above the surface of the Earth, in metres per second. Care should be taken when comparing this parameter with observations, because wind observations vary on small space and time scales and are affected by the local terrain, vegetation and buildings that are represented only on average in the ECMWF Integrated Forecasting System (IFS). This parameter can be combined with the U component of 10m wind to give the speed and direction of the horizontal 10m wind.

**Mean wave direction** - *degree*, This parameter is the mean direction of ocean/sea surface waves. The ocean/sea surface wave field consists of a combination of waves with different heights, lengths and directions (known as the two-dimensional wave spectrum). This parameter is a mean over all frequencies and directions of the two-dimensional wave spectrum. The wave spectrum can be decomposed into wind-sea waves, which are directly affected by local winds, and swell, the waves that were generated by the wind at a different location and time. This parameter takes account of both. This parameter can be used to assess sea state and swell. For example, engineers use this type of wave information when designing structures in the open ocean, such as oil platforms, or in coastal applications. The units are degree true which means the direction relative to the geographic location of the north pole. Zero means "coming from the north" and 90

"coming from the east".

**B**

**Mean wave period** -  $s$ , This parameter is the average time it takes for two consecutive wave crests, on the surface of the ocean/sea, to pass through a fixed point. The ocean/sea surface wave field consists of a combination of waves with different heights, lengths and directions (known as the two-dimensional wave spectrum). This parameter is a mean over all frequencies and directions of the two-dimensional wave spectrum. The wave spectrum can be decomposed into wind-sea waves, which are directly affected by local winds, and swell, the waves that were generated by the wind at a different location and time. This parameter takes account of both. This parameter can be used to assess sea state and swell. For example, engineers use such wave information when designing structures in the open ocean, such as oil platforms, or in coastal applications.

**Significant height of combined wind waves and swell** -  $m$ , This parameter represents the average height of the highest third of surface ocean/sea waves generated by wind and swell. It represents the vertical distance between the wave crest and the wave trough. The ocean/sea surface wave field consists of a combination of waves with different heights, lengths and directions (known as the two-dimensional wave spectrum). The wave spectrum can be decomposed into wind-sea waves, which are directly affected by local winds, and swell, the waves that were generated by the wind at a different location and time. This parameter takes account of both. More strictly, this parameter is four times the square root of the integral over all directions and all frequencies of the two-dimensional wave spectrum. This parameter can be used to assess sea state and swell. For example, engineers use significant wave height to calculate the load on structures in the open ocean, such as oil platforms, or in coastal applications.

# C

## VINCENTY'S FORMULA

THE constant inputs for this method are the length of the semi-major axis of the ellipsoid,  $a_{earth} = 6378137 \text{ m}$ , also the flattening,  $f$ , of the ellipsoid is required which is  $1/298.257223563$  for the Earth. From the initial point  $(\phi_1, L_1)$  and the initial azimuth/course over ground,  $\alpha_1$ , and distance,  $s$ , this method will calculate the end point  $(\phi_2, L_2)$ .

**Vincenty's Formula** Start by calculating the following with the starting point,  $(\phi_1, L_1)$  in degree, the distance  $s$  in meters and the course over ground  $\alpha_1$  in degree as inputs;

$$U_1 = \arctan[(1 - f) \tan(\phi_1)] \quad (\text{C.1})$$

$$\sigma_1 = \arctan 2(\tan(U_1), \cos(\alpha_1)) \quad (\text{C.2})$$

$$\sin(\alpha) = \cos(U_1) \sin(\alpha_1) \quad (\text{C.3})$$

$$b = (1 - f) a \quad (\text{C.4})$$

$$u^2 = (\sin^2(\alpha) - 1) \left( \frac{a^2 - b^2}{b^2} \right) \quad (\text{C.5})$$

$$A = 1 + \frac{u^2}{16384} (4096 + u^2 [-768 + u^2 (320 - 175u^2)]) \quad (\text{C.6})$$

$$B = \frac{u^2}{1024} (256 + u^2 [-128 + u^2 (74 - 47^2)]) \quad (\text{C.7})$$

Then, using an initial value  $\sigma = \frac{s}{bA}$ , iterate the following equations until there is no significant change in  $\sigma$ :

$$2\sigma_m = 2\sigma_1 + \sigma \quad (\text{C.8})$$

$$\Delta\sigma = B \sin(\sigma) \left( \cos(2\sigma_m) + \frac{B}{4} \left( \cos(\sigma)[-1 + 2\cos^2(2\sigma_m)] - \frac{B}{6} \cos[2\sigma_m][-3 + 4\sin^2(\sigma)][-3 + 4\cos^2[2\sigma_m]] \right) \right) \quad (\text{C.9})$$

$$\sigma = \frac{s}{bA} + \Delta\sigma \quad (\text{C.10})$$

Once  $\sigma$  is obtained to sufficient accuracy evaluate:

$$\phi_2 = \arctan 2 \left( \frac{\sin(U_1) \cos(\sigma) + \cos(U_1) \sin(\sigma) \cos(\alpha_1)}{(1-f) \sqrt{\sin^2(\alpha) + (\sin(U_1) \sin(\sigma) - \cos(U_1) \cos(\sigma) \cos(\alpha_1))^2}} \right) \quad (\text{C.11})$$

$$\lambda = \arctan 2 (\sin(\sigma) \sin(\alpha_1), \cos(U_1) \cos(\sigma) - \sin(U_1) \sin(\sigma) \cos(\alpha_1)) \quad (\text{C.12})$$

$$C = \frac{f}{16} \cos^2(\alpha) [4 + f(4 - 3\cos^2(\alpha))] \quad (\text{C.13})$$

$$L = \lambda - (1 - C) f \sin(\alpha) (\sigma + C \sin(\sigma) (\cos[2\sigma_m] + C \cos(\sigma)[-1 + 2\cos^2(2\sigma_m)]) \quad (\text{C.14})$$

$$L_2 = L + L_1 \quad (\text{C.15})$$

The calculated end position is given by the point  $(\phi_2, L_2)$ .

# REFERENCES

- Aiello, G., Giallanza, A., and Mascarella, G. (2020). Towards shipping 4.0. a preliminary gap analysis. *Procedia Manufacturing*.
- Bell, D. and Wang, H. (2000). A formalism for relevance and its application in feature subset selection. *Machine Learning*.
- Besikci, E. B., Arslan, O., Turan, O., and Olcer, A. (2016). An artificial neural network based decision support system for energyefficient ship operations. *Computers & Operations Research*, 66:393–401.
- Bialystocki, N. and Konovessis, D. (2016). On the estimation of ship's fuel consumption and speed curve:a statistical approach. *Journal of Ocean Engineering and Science*.
- Callery, S. and Bailey, D. (2020a). Is it too late to prevent climate change? Technical report, NASA.
- Callery, S. and Bailey, D. (2020b). What is the greenhouse effect? Technical report, NASA.
- Carlton, J. S. (2007). *Marine Propellers and Propulsion*. Butterworth-Heinemann.
- Chaal, M. (2018). *Ship Operational Performance Modelling for Voyage Optimization through Fuel Consumption Minimization*. PhD thesis, Aalto University.
- Coraddu, A., Lim, S., Oneto, L., Pazouki, K., Norman, R., and Murphy, A. J. (2019a). A novelty detection approach to diagnosing hull and propeller fouling. *Ocean Engineering*.
- Coraddu, A., Oneto, L., Anguita, D., and Baldi, F. (2015). Ship efficiency forecast based on sensors data collection: Improving numerical models through data analytics. In *IEEE Discovering Sustainable Ocean Energy for a New World*. IEEE.
- Coraddu, A., Oneto, L., Anguita, D., and Baldi, F. (2017). Vessels fuel consumption forecast and trim optimisation: A data analyticsperspective. *Ocean Engineering*.
- Coraddu, A., Oneto, L., Baldi, F., Atlar, M., and Savio, S. (2019b). Data-driven ship digital twin for estimating the speed loss caused by the marine fouling. *Ocean Engineering*, 186.
- Dalheim, O. O. and Steen, S. (2020). Preparation of in-service measurement data for ship operation andperformance analysis. *Ocean Engineering*.
- de Waard, D. (2014). Parameterization of ship propulsion drives. Master's thesis, Technical University Delft.

- Engelbrecht, J. (2015). Het gebruik van de power take in als powertake out op de holland klasse schepen. Master's thesis, Netherlands Defense Academy.
- ENVI (2015). Emission reduction targets for international aviation and shipping. Technical report, Policy Department A for the Committee on Environment, Public Health and Food Safety.
- Faltinsen, O. M. (1980). Prediction of resistance and propulsion of a ship in a seaway. *Proceedings of the 13th Symposium on Naval Hydrodynamics*.
- Farag, Y. B. A. and Ölcer, A. I. (2020). The development of a ship performance model in varying operating conditions based on ann and regression techniques. *Ocean Engineering*, 198.
- Feng, M., Shaw, S.-L., Peng, G., and Fang, Z. (2020). Time efficiency assessment of ship movements in maritime ports: A case study of two ports based on ais data. *Journal of Transport Geography*.
- Frost, J. and Lavatt, R. (2020). Comparison of second order optimization algorithms in neural networks applied on large-scale problems. Master's thesis, KTH Royal Institute of Technology; School of Engineering Sciences, Stockholm, Sweden.
- Geertsma, R., Negenborn, R., Visser, K., Loonstijn, M., and Hopman, J. (2017). Pitch control for ships with diesel mechanical and hybrid propulsion: modelling, validation and performance quantification. *Applied Energy*, 206.
- Gkerekos, C., Lazakis, I., and Theotokatos, G. (2019). Machine learning models for predicting ship main engine fuel oil consumption: A comparative study. *Ocean Engineering*, 188.
- Harirchian, E., Lahmer, T., and Rasulzade, S. (2020). Earthquake hazard safety assessment of existing buildings using optimized multi-layer perceptron neural network. *Energies*, (13):2060.
- Heaton, J. (2015). *Artificial Intelligence for Humans, Volume 3: Deep Learning and Neural Networks*. Createspace Independent Publishing Platform, 2015.
- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Sabater, J. M., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N. (2021). Era5 hourly data on single levels from 1979 to present. Technical report, ECMWF.
- Holtrop, J. (1984). A statistical re-analyses of resistance and propulsion data. *International Shipbuilding Progress*.
- IMO (2014). Key findings from the third imo ghg study 2014. Technical report, International Maritime Organisation (IMO).
- IMO (2018). Adoption of the initial imo strategy on reduction of ghg emissions from ships and existing imo activity related to reducing ghg emissions in the shipping sector. Technical report, International Maritime Organisation.

- Isikli, E., Aydin, N., Bilgili, L., and Toprak, A. (2020). Estimating fuel consumption in maritime transport. *Journal of Cleaner Production*.
- Jeon, M., Noh, Y., Shin, Y., Lim, O.-K., Lee, I., and Cho, D. (2018). Prediction of ship fuel consumption by using an artificial neural network. *Journal of Mechanical Science and Technology*.
- Journée, J. and Massie, W. (2001). *Offshore Hydrodynamics*. Delft University of Technology.
- Kim, S.-H., Roh, M.-I., Oh, M.-J., Park, S.-W., and Kim, I.-I. (2020). Estimation of ship operational efficiency from ais data using big datatechnology. *International Journal of Naval Architecture and Ocean Engineering*.
- Le, L. T., Lee, G., Park, K.-S., and Kim, H. (2020). Neural network-based fuel consumption estimation for container ships in korea. *Maritime Policy & Management*.
- Leifsson, L. P., Savarsdóttir, H., Sigurosson, S. P., and Vésteinsson, A. (2008). Grey-box modeling of an ocean vessel for operational optimization. *Simulation Modelling Practice and Theory*.
- Li, X., Sun, B., Guo, C., Du, W., and Li, Y. (2020). Speed optimization of a container ship on a given route considering voluntary speed loss and emissions. *Applied Ocean Engineering*.
- Lin, C.-Y. and Jeng, Y.-L. (1996). Influences of charge air humidity and temperature on the performance and emission characteristics of diesel engines. *Journal of Ship Research*, 40(2):172–177.
- MEPC (2009). Guidelines for voluntary use of the ship energy efficiency operational indicator (eeoi). *IMO Tier 5*.
- MEPC (2011). Amendments to the annex of the protocol of 1997 to amend the international convention for the prevention of pollution from ships, 1973, as modified by the protocol of 1978 relating thereto. Technical report, International Maritime Organisation.
- Moreno-Gutiérrez, J., Pájaro-Velázquez, E., Amado-Sánchez, Y., Rodríguez-Moreno, R., Calderay-Cayetano, F., and Durán-Grados, V. (2019). Comparative analysis between different methods for calculating on-board ship's emissions and energy consumption based on operational data. *Science of the Total Environment*.
- Nakamura, S. and Naito, S. (1975). Propulsive performance of a container ship in waves. *J.S.N.A. Kansai Japan*, (159).
- Oosterveld, M. and van Oossanen, P. (1975). Further computer-analyzed data of the wageningen b-screw series. *International Shipbuilding Progress*.
- Perera, L. P. and Soares, C. G. (2017). Weather routing and safe ship handling in the future of shipping. *Ocean Engineering*, (130):684 – 695.

- Rong, H., Teixeira, A., and Guedes-Soares, C. (2020). Data mining approach to shipping route characterization and anomaly detection based on ais data. *Ocean Engineering*.
- Shi, W., Stapersma, D., and Grimmelius, H. (2009). Analysis of energy conversion in ship-propulsion system in off-design operation conditions. *Transactions on Ecology and the Environment*.
- Shivachev, E., Khorasanchi, M., Day, S., and Turan, O. (2020). Impact of trim on added resistance of kriso container ship (kcs) in headwaves: An experimental and numerical study. *Ocean Engineering*.
- Song, S., Demirel, Y. K., Muscat-Fenech, C. D. M., Tezdogan, T., and Atlar, M. (2020). Fouling effect on the resistance of different ship types. *Ocean Engineering*.
- Tarelko, W. and Rudzki, K. (2020). Applying artificial neural networks for modelling ship speed and fuel consumption. *Neural Computing and Applications*.
- Taskar, B. and Andersen, P. (2020). Benefit of speed reduction for ships in different weather conditions. *Transportation Research Part D*.
- Taskar, B., Yum, K. K., Steen, S., and Pedersen, E. (2016). The effect of waves on engine-propeller dynamics and propulsion performance of ships. *Ocean Engineering*, (122):262–277.
- Tol, R. and Linden, Y. (2018). Effect of gas-to-liquid (gtl) fuels on marine diesel engines compared to f-76. *International Naval Engineering Conference*.
- Trodden, D., Murphy, A., Pazouki, K., and Sargeant, J. (2015). Fuel usage data analysis for efficient shipping operations. *Ocean Engineering*.
- Tufte, E. D. (2014). Impacts of low load operation of modern four-stroke diesel engines in generator configuration. Master's thesis, Norwegian University of Science and Technology - Department of Marine Technology.
- Ueno, M., Tsukada, Y., and Tanizawa, K. (2013). Estimation and prediction of effective inflow velocity to propeller in waves. *Journal Maritime Science Technology*, (18):339–348.
- UN (1992). United nations framework convention on climate change. Technical report, United Nations.
- UN (2015). Paris agreement. Technical report, United Nations.
- Vasilikis, N. I. (2020). Operational data-driven energy efficiency and effectiveness assessment of a hybrid propulsion equipped naval vessel. *International Naval Engineering Conference & Exhibition*, (15).
- Viktorelius, M. and Lundh, M. (2019). Energy efficiency at sea: An activity theoretical perspective on operational energy efficiency in maritime transport. *Energy Research and Social Science*.

- Vrijdag, A., Stapersma, D., and van Terwisga, T. (2014). Control of propeller cavitation in operational conditions. *Journal of Marine Engineering & Technology*.
- Woud, H. K. and Staperma, D. (2002). *Design of Propulsion and Electric Power Generation Systems*. IMarEST.
- Xiao, F., Ligteringen, H., van Gulijk, C., and Ale, B. (2015). Comparison study on ais data of ship traffic behavior. *Ocean Engineering*.
- Xiong, H. and Wang, X. (2019). Effect of leeway and drift angle of ship navigation and determination method. *MATEC Web of Conferences*.
- Yan, R., Wang, S., and Du, Y. (2020). Development of a two-stage ship fuel consumption prediction and reduction model for a dry bulk ship. *Transportation Research Part E*.
- Yilmaz, I. and Kaynar, O. (2011). Multiple regression, ann (rbf, mlp) and anfis models for prediction of swell potential of clayey soils. *Expert Systems with Applications*.
- Yoo, B. and Kim, J. (2019). Probabilistic modeling of ship powering performance using full-scale operational data. *Applied Ocean Research*.
- yu Guo, C., hai Zhong, X., gang Zhao, D., Wang, C., feng Lin, J., and wei Song, K. (2020). Propulsion performance of large-scale ship model in real sea environment. *Ocean Engineering*.
- Zaman, I., Pazouki, K., Norman, R., Younessi, S., and Coleman, S. (2017). Challenges and opportunities of big data analytics for upcoming regulations and future transformation of the shipping industry. *Procedia Engineering*, 194:537–544.
- Zis, T. P., Psaraftis, H. N., and Ding, L. (2020). Ship weather routing: A taxonomy and survey. *Ocean Engineering*.



# NOMENCLATURE

## Greek Symbols

$\alpha$	forward rotation of the geodesic [deg]
$\beta$	heading [deg]
$\beta_{1,2}$	exponential decay rates
$\epsilon_{\phi\zeta}$	phase difference between wave and motion [rad]
$\eta$	learning rate
$\eta_o$	open water efficiency propeller
$\eta_{eng}$	engine efficiency
$\mu$	wave direction w.r.t. ship [deg]
$\omega_e$	wave encounter frequency [rad/s]
$\phi$	latitude of position
$\phi$	pitch motion ship [degree]
$\phi_a$	amplitude of pitch motion [m]
$\rho$	density of seawater [ $\text{kg}/\text{m}^3$ ]
$\sigma$	angular separation between positions
$\theta$	roll motion ship [degree]
$\theta_0$	pitch at which zero thrust is achieved [m/rev]
$\theta_a$	amplitude of roll motion [m]
$\theta_p$	actual propeller pitch [m/rev]
$\theta_{nom}$	design pitch [m/rev]

## Roman Symbols

$\Delta t$	time difference between records [s]
$\dot{m}_f$	fuel consumption [kg/s]
$\hat{y}$	predicted variable
$a$	length of semi-major axis of the ellipsoid [m]
$b$	length of semi-minor axis of the ellipsoid [m]
$C_{fuel}$	carbon content in fuel
$COG$	course over ground [degree]
$D_p$	propeller diameter
$D_x$	current perpendicular to $D_y$ [knots]
$D_y$	current in ship's movement direction [knots]
$f$	flattening of the ellipsoid
$F_{acc}$	acceleration force [kN]
$FPM$	fuel consumption per mile [kg/mile]
$i_{gb}$	Gearbox reduction ratio
$J$	advanced ratio
$K_Q$	non-dimensional torque coefficient propeller
$K_T$	non-dimensional thrust coefficient propeller
$L$	longitude of position

$M$	molecular weight [kg/mol]
$m$	ship's mass [kg]
$M_s$	torque at shaft [kNm]
$MAE$	mean absolute error
$MAE_{\%}$	relative mean absolute error [%]
$mwd$	wave direction w.r.t. true north [deg]
$mwp$	wave period [s]
$n_e$	rotational speed main diesel engine [rpm]
$n_p$	propeller rotational speed [rpm]
$n_{PEM}$	rotational speed patrol electric motor [rpm]
$n_{virt}$	virtual shaft speed [rpm]
$P$	power [kW]
$Q$	propeller torque [kNm]
$R_{tot}$	total resistance [kN]
$RH$	relative humidity [%]
$RWA$	relative wind angle [deg]
$RWS$	relative wind speed [knots]
$s$	ellipsoidal distance between two points [m]
$swh$	significant wave height [m]
$T$	propeller thrust [kN]
$t$	thrust deduction factor
$T_a$	draft at the stern [m]
$T_f$	draft at the bow [m]
$T_{amb}$	outside air temperature [°C]
$TWA$	true wind angle [degree]
$TWS$	true wind speed [knots]
$U$	reduced latitude of position
$u_{wind}$	10m u-component of wind [m/s]
$V_a$	advanced velocity of water [m/s]
$V_s$	speed over ground [knots]
$V_w$	speed through water [knots]
$v_{wind}$	10m v-component of wind [m/s]
$w$	wake fraction
$X$	fuel rack position main diesel engine [%]

### Subscripts

$eng$	engine
$GB$	gearbox
$max$	maximum
$min$	minimum
$PS$	portside
$SB$	starboard
$Ship$	with respect to the ship
$tot$	total
$x$	projected perpendicular to the ship's movement direction
$y$	projected on the ship's movement direction