**NTNU Trondheim**
**Norwegian University of Science and Technology**
*Department of Marine Technology*

# MASTER THESIS IN MARINE TECHNOLOGY

## SPRING 2019

## FOR

## Jonas Munch Wahl

## Prediction of fuel consumption of a ship in transit using machine learning

The fuel consumption is of importance for the economy of ship operation, as well as for environmental considerations, logistics etc. By using the increasing number of onboard measurements in combination with machine learning methods, it is foreseen that a prediction model for fuel consumption that accounts for not only the speed, but also weather and environment variables, loading condition and other aspects.

The objective of the thesis is to develop and validate at least one (partly) machine-learning based prediction model for fuel consumption. The thesis shall include a detailed account of the in-service data utilized and the processing of that data, as well as the methodology applied to develop the prediction model. It is beneficial to compare different methods using different approaches to the problem.

In the thesis the candidate shall present his personal contribution to the resolution of problem within the scope of the thesis work.

Theories and conclusions shall be based on mathematical derivations and/or logic reasoning identifying the various steps in the deduction.

The thesis work shall be based on the current state of knowledge in the field of study. The current state of knowledge shall be established through a thorough literature study, the results of this study shall be written into the thesis. The candidate should utilize the existing possibilities for obtaining relevant literature.

The thesis shall be organized in a rational manner to give a clear exposition of results, assessments, and conclusions. The text should be brief and to the point, with a clear language. Telegraphic language should be avoided.

The thesis shall contain the following elements: A text defining the scope, preface, list of contents, summary, main body of thesis, conclusions with recommendations for further work, list of symbols and acronyms, reference and (optional) appendices. All figures, tables and equations shall be numerated.

The supervisor may require that the candidate, in an early stage of the work, present a written plan for the completion of the work. The plan shall include a budget for the use of laboratory or other resources that will be charged to the department. Overruns shall be reported to the supervisor.

The original contribution of the candidate and material taken from other sources shall be clearly defined. Work from other sources shall be properly referenced using an acknowledged referencing system.

The thesis shall be submitted electronically (pdf) in Inspera:

**NTNU Trondheim**
**Norwegian University of Science and Technology**
*Department of Marine Technology*

- Signed by the candidate
- The text defining the scope (this text) (signed by the supervisor) included

The candidate will receive a printed copy of the thesis.

Supervisor : Professor Sverre Steen
Advisor : Sigbjørn Rudå
Start : 15.01.2019
Deadline : 11.06.2019

Trondheim, 15.01.2019

Sverre Steen
Supervisor

# Abstract

This thesis presents a novel technique for fuel prediction of ships in transit considering external environmental factors such as current, wind and waves. Among the many performance monitoring systems and methods, less common are performance forecasting grounded in available monitored data. A new methodology for accurate prediction of fuel consumption by combine well established physical and empirical methods with state of the art machine learning algorithms and artificial neural networks forms the foundation of this thesis .

In order to apply data-driven methods, a framework for in-service operational data were developed to identify and process measurements. By use of hindcast climate data, unreliable measurements were replaced by simulated entries from ECMWF and Tidetech to ensure high quality data for the prediction model. The result of the data analysis are discussed in extent to illustrated pitfalls, bias and random errors.

By isolating wind and frictional resistance acting on a ship in transit with empirical models, a set of methods were proposed for determination of the residual resistance. Attempting to decouple the calm water wave resistance and added resistance due to waves, and considered the residual resistance as one were among the methods explored. Shallow artificial neural network and Gaussian process regression showed an impressive precision with a mean deviation of 2.5% in the prediction of fuel consumption.

Bootstrapping showed a model behavior that reflects the involved physics in the system and confirmed that these type of prediction models are suitable for fuel prediction. As a result, the models are expected to predict even more accurate as the amount of training data increase.

# Sammendrag

Denne masteravhandlingen presenterer en ny teknikk for drivstoffsprediksjon for skip i transitt som tar hensyn til eksterne miljømessige faktorer som strøm, vind og bølger. Blant de mange ytelsesovervåkingssystemene og metodene, er det mindre vanlig med ytelsesprediksjoner basert på tilgjengelig måledata. En ny metode for nøyaktig prediksjon av drivstofforbruk ved å kombinere veletablerte fysiske og empiriske metoder med det siste innen maskinlæringsalgoritmer og kunstige nevrale nettverk danner grunnlaget for denne avhandlingen.

For å anvende de datadrevne metodene, ble et rammeverk utviklet for å identifisere avvik, samt å prosessere målerverdier fra driften av skip. Ved å benytte historisk værdata kunne unøyaktige og upålitelige målepunkter erstattes av simulerte verdier fra ECMWF og Tidetech for å forsikre høykvalitets data for trening av prediksjonsmodellen. Resultatet av denne analysen er diskutert i sin helhet for å illustrere fallgruver, systematiske og tilfeldige feil i målesystemet.

Ved å isolere vind og friksjonsmotstanden som virker på et skip i transitt med empiriske modeller, ble et sett metoder foreslått for å bestemme restmotstanden. Forsøk på å analysere koblet og frikoblet bølgemotstand i stille vann og økt motstand grunnet bølger, samt å betrakte disse samlet var blant de foreslåtte metodene. Grunne nevrale nettverk og Gaussian prosess regresjon viste imponerende presisjon med et gjennomsnittlig avvik på kun 2.5% i prediksjonen av drivstofforbruk.

Bootstrapping viste en modelloppførsel som reflekterer den involverte fysikken i systemet og bekreftet dermed at disse type prediksjonsmodeller er godt egnet for drivstoffprediksjon. Dermed kan det forventes av disse modellene å predikere enda mer nøyaktig i takt med at mengden treningsdata øker.

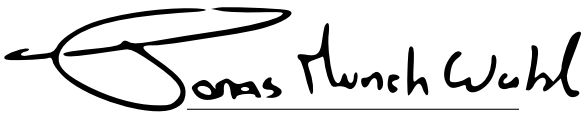
# Preface

This thesis is written by Jonas Munch Wahl as a part of a Master's degree at the Maritime Technology Department at the Norwegian University of Science and Technology (NTNU) in Trondheim. The topic was motivated and carried out in collaboration with Wilhelmsen Ship Management.

During the autumn of 2018, a pre-study was conducted as preparatory work for the master's thesis. The primary focus of this study was to familiarize with machine learning and how it is used for prediction purposes within ship resistance. Parallel with the study of machine learning, a data acquisition process of in-service data was initiated and a preliminary analysis of the data quality was conducted.

An important part of the master thesis has been to conduct a detailed analysis of the in-service operational data by evaluating different preprocessing approaches and methods. Further, with the available data explore the various state of the art methods for performance monitoring and predictions. Development and testing of a new prediction method form a major part of the work conducted.

This thesis is the result of a major within the field of hydrodynamics even though it explores fields of study that usually is considered as statistic and elements from computer science. It shows how these disciplines are beneficial and applicable across traditional conventions. Precisely this symbiosis of disciplines have been challenging to carry out due to the lack of experience, but have in return given a substantial learning outcome.



Jonas Munch Wahl, 11.06.2019

# Acknowledgments

Most of all I would like to express my gratitude to Professor Sverre Steen, Head of Department of Marine Technology, for valuable guidance in the field of hydrodynamics, as well as for keeping me on a steady course by frequent supervising sessions. His door was always open and he would gladly share his knowledge, discuss and explain findings.

I would further like to thank Ph.D. Candidates Øyvind Øksnes Dalheim and Prateek Gupta for sharing and providing valuable experience, guidance, and routines.

Additionally, I would like to thank Sigbjørn Rudaa at RaaLabs for pitching the idea, providing data and guidance on short notice. Sigbjørn has throughout the work emphasized the commercial value of a promising result in the end which have been of great motivation along the way.

Lastly, I would like to thank all of my fellow office colleagues for their patience and support throughout the final year at NTNU and Marine Technology. Together, we created a great environment for academic discussion parallel with a lot of fun.

J.M.W

# Contents

# List of Figures

# List of Tables

# Nomenclature

**Acronyms**

| | |
|---|---|
| ANN | Artificial neural network |
| APE | Absolute percentage error |
| API | Application programming interface |
| BR | Bayesian regularization |
| DG | Diesel generator |
| GPR | Gaussian process regression |
| HFO | Heavy fuel oil |
| IMO | International Maritime Organization |
| LM | Levenberg-Marquardt |
| MAE | Mean absolute error |
| MAPE | Mean absolute percentage error |
| MCR | Maximum continuous rating |
| MGO | Marine gas oil |
| ML | Machine learning |
| MSE | Mean squared error |
| NaN | Not a number |

NMSE Normalized mean squared error

PCA Principle component analysis

PPMCC Pearson product-moment correlation coefficient

PS Port side

REP Relative error percentage

RMS Root mean square

RPM Revolutions per minute

SB Starboard side

SFC Specific fuel consumption

SOG Speed over ground

STW Speed through water

UTC Universal time coordinated

**Greek Symbols**

$\alpha$ Learning rate

$\beta_w$ Wind direction

$\beta_{rw}$ Relative wave angle

$\eta_m$ Mechanical efficiency

$\gamma_w$ Wind angle to bow

$\gamma_{rw}$ Relative wind angle to bow

$\lambda$ Scale ratio

$\mu$ Kinematic viscosity

$\psi$ Heading

$\rho_a$ Density air

$\rho_w$ Density sea water

**Superscripts**

$Z$ Number of blades

$A_p$ Projected area

B Breadth

$C_{AA}$ Air resistance coefficient

$C_B$ Block coefficient

$C_{DB}$ Transom stern resistance coefficient

$C_D$ Drag coefficient

$C_f$ Friction resistance coefficient

$C_r$ Residual resistance coefficient

$C_T$ Total resistance coefficient

D Propeller diameter

$F_n$ Froude number

H Hull roughness

$J_a$ Advance ratio

J* Propulsion point

$K_Q$ Torque coefficient

$K_T$ Thrust coefficient

$L_{OA}$ Length over all

$L_{PP}$ Length between perpendicular

$L_{WL}$ Length of waterline

P Motor power

$P_a$ Auxiliary power

$P_D$ Engine power

Q Torque

$R_{nco}$ Local Reynolds number

$R_N$ Reynolds number

S Wetted surface area

$S_{DB}$ Transom stern area

T Thrust

$T_{AP}$ Draft aft perpendicular

$T_{FP}$ Draft front perpendicular

$V_a$ Velocity of advance

$V_s$ Ship velocity

$V_a$ Velocity of advance

$V_{rel}$ Relative velocity

$V_{rw}$ Relative wind velocity

$V_w$ Wind velocity

**Subscripts**

$\dot{m}_f$ Mass flow of fuel

$w$ Wake

c Chord length

| | |
|---|---|
| k | Form factor |
| $k_p$ | Blade roughness |
| n | Rate of revolutions |
| t | Maximum thickness of chord/blade |
| t | Thrust deduction |

# Introduction

## Background and Motivation

Fuel consumption of a ship in service is a measurement of the amount of energy consumed in order to obtain a certain operational state. These states are defined by short term parameters such as speed, trim, loading- and weather condition and long term parameters as fouling and erosion. Over a ship's lifespan, the rate of consumed fueled is expected to increase as the performance decrease due to the long term parameters.

It is of the ship owners and operators interest to get a sufficient overview of both the short and long term trends in the performance as it is directly related to the fuel consumption. A performance deterioration of 5% on a cruise ship that daily consuming an average of 50 tonnes of the heavy fuel oil at the price of 450 USD/t and another 10 tonnes to marine gas oil at the price of 600 USD/t an annual penalty of approximately 500 000 USD could be expected.

The short term parameters are equally important to describe the total fuel consumption of a ship in transit. Maruo (1957) showed that additional resistance due to waves are proportional to the square of the wave height whereas the wind and air resistance is proportional to the squared of projected superstructure area. Kim et al. (2017) showed that an added resistance due to wind and waves in a Beaufort 6 (MetOffice 2016) environment would lead to a speed loss of approximately 5%. A thorough understanding of how environmental factors influence the fuel consumption allow the owners and operators to optimize route planing, better budgeting and most important satisfy the stricter emission regulatory framings stressed by the International Maritime Organization, IMO.

With increasing number of measured parameters and more frequent sampling, the complexity of the performance monitoring increase with the same rate; an operational state can be described as a function of numerous sensor inputs. Conventional analytic and empirical methods are not taking advantage of the valuable information in these inputs. Hansen et al. (2011) showed promising results in a more advanced performance monitoring analysis using the bond graph method and Pedersen et al. (2014) illustrated that artificial neural networks

and Gaussian process regression are well suited for prediction of performance at a long term perspective.

Data-driven models, also known as machine learning models are being used in an increasing number of fields and there are several machine learning models available and their areas of applications seems to expand rapidly. Due to their natural flexibility and ability to mimic both physical and unphysical processes solidly based on the the data, its magnitude and correlations, these models should also be able to predict the short term effect of current, wind and waves on a ship in transit.

## Objective

The main objective of this thesis is to develop and validate a machine learning based prediction model for fuel consumption for a ship in transit that consider consider the short term parameters. This includes a detailed study of the in-service operational data as well as a detailed description of the developed methods where logic reasoning is communicated clearly. For quantification of prediction performance it is beneficial to compare different methods using different approaches to the problem.

## Scope of Work and Limitations

To accomplish the objective it is expected to develop a framework for preprocessing of ship monitoring data. The framework is expected to identify faults, bias and random errors with assistance and interpretation for better decision making. The machine learning models involved are expected to be validated and their performance quantified, and it is beneficial to explore different algorithms for the purpose of prediction.

It is not within the scope of hydrodynamics and thereby this masters thesis to establish a framework for route generation and optimization for testing. Benchmarking of the precision should therefore be performed with a fraction of the existing data.

## Structure of Thesis

The following is an explanation of the flow of the thesis:

**Part I: Theory** review the necessary literature required for a complete understanding of the argumentation in the results. Chapter 1 summarize aspects of naval hydrodynamics related to ship resistance and propeller theory. Chapter 2 gives an introduction to the fundamentals within machine learning and represents most of the literature study conducted in this thesis. Following, chapter 3 is a review of data preprocessing methods and techniques which are utilized to obtain the results.

**Part II: Methodology** review how the theory should be applied to achieve the overall objective. Chapter 4 gives an overview of the available data sources in this thesis and complements the next chapter. Chapter 5 gives a detailed description of the stepwise process from preprocessing of data, to how the resistance is decomposed and isolated, and how machine learning should be applied for this purpose. Followed by how the predicted resistance will be correlated with fuel consumption.

**Part III: Results** presents the findings in this thesis. Chapter 6 review a thorough data analysis of the in-service operational data including a continuous discussion of the quality. Recommendations and conclusions to which parameters are reliable and not are found consecutively. Chapter 7 presents the results from the resistance and fuel predictions from one vessel, followed by a brief discussion of the overall performance, including results from the study of a second vessel. Part III finishes with a formulation of a conclusion to the objective and recommendations for further work in chapter 8.

# Part I

# Theory

# Chapter 1

# Naval Hydrodynamics

## 1.1 Ship Resistance

The objective of this chapter is to give an understanding of what ship resistance is being caused by, how it can be more easily understood and calculated by dividing into different components.

A prerequisite in the analysis of ship resistance is fluid mechanics. To fully understand different aspects of the analysis, chapter 5 and chapter 7 in White (2011) are recommended literature. These chapters introduce how and why dimensionless properties are useful in fluid mechanics, as well as an introduction to viscous flow.

### 1.1.1 Decomposing Resistance

A fundamental classification of the resistance consist of

- Pressure resistance - acting normal to the surface of the hull
- Frictional resistance - acting tangential to the surface of the hull

Frictional resistance is caused by viscous effects. In ideal fluids, where potential theory is valid such viscous effects are not present, hence the use of potential theory to describe the flow field around the hull will only consider pressure resistance.

Further, in potential theory there is no resistance on a deeply submerged body - this phenomenon is known as d'Alembert's paradox - meaning that the pressure resistance in potential theory is caused by free surface effects. In the case of a ship, the free surface effects are the generation of waves, hence the pressure resistance is often known as wave resistance. The

observant reader will note that waves change the flow field at the surface and thereby affect the frictional contribution, but this is an effect that normally is neglected in ship resistance calculations, meaning that frictional resistance and wave resistance is calculated independently.

However, this assumption does not reflect the whole nature of pressure resistance since it gives rise to viscous effects, such as flow separation - a viscous pressure component. Viscous pressure is difficult to quantify in a general matter, but add a minor contribution to the total resistance. Hence simplified methods have been developed to account the effect.

A common division of resistance components, suitable for calculation of the total resistance consist of

- Frictional resistance
- Viscous pressure resistance
- Wave resistance
- Wind resistance
- Appendage resistance

### 1.1.2 Frictional Resistance

Whereas dry friction is dependent on the pressure, fluid friction is described by shear forces occurring when fluid layers are moving relative to each other. One could, therefore, assume that the frictional resistance on a ship is described by a flow on a flat plate with the same surface area and the same Reynolds number as the submerged part of the hull.

Expressing the frictional resistance in a dimensionless manner, the frictional coefficient occur as

$$C_F = \frac{R_F}{\frac{1}{2}\rho V_s^2 S} = f(R_N), \quad R_N = \frac{V_s L_{WL}}{\nu} \tag{1.1.1}$$

The magnitude of the resistance force dependents on whether the flow is laminar or turbulent, which is described by the Reynolds number. For laminar flow, the frictional coefficient is expressed by Blasius formula

$$C_F = \frac{1.327}{\sqrt{R_N}} \tag{1.1.2}$$

and for turbulent flow, the ITTC'57 correction line are

$$C_F = \frac{0.075}{(\log_{10}(R_N) - 2)^2} \tag{1.1.3}$$

There exist several other friction lines for turbulent flow, but the differences between these lines are not significant, such that the ITTC'57 line will be used in this thesis. Describing the flow field around a ship will in all practical cases involve turbulent flow, as the Reynolds number is in order of $10^9$.

#### Form factor

Even though the pressure differences are neglected when calculating the frictional resistance, there exists a pressure distribution around the hull with the coherent velocity distribution - in line with Bernoulli's equation. In the bow and stern, there are lower local velocities due to the local higher pressure, while on the longitudinal sides and bottom the velocities are relatively higher. The increasing velocities give rise to an increasing frictional resistance - the coefficient is proportional to the velocity squared, equation (**??**).

However, expressing the velocity distribution as a function of the geometry is not convenient for practical use. By introducing a factor based on the geometry this effect can be accounted for. Determining this factor - known as the form factor - can be done in multiple waves, e.g, empirical methods, Prohaska's method or by towing a model in low speed where wave resistance is negligible. MARINTEK's formula is fully based on the geometry and loading conditions

$$k = 0.6\phi + 145\phi^{3.5}, \quad \phi = \frac{C_B}{L_{WL}}\sqrt{(T_{AP} + T_{FP})B} \tag{1.1.4}$$

### 1.1.3 Roughness and Fouling

An important assumption when calculating the frictional resistance by either Blasius or ITTC'57 is that it assumes smooth surfaces, which not reflect the true nature of a ship hull. Welds, flaking paint or fouling are properties that will increase the friction and should, therefore, be taken into account.

For newly painted or docked vessels, the roughness is in the area of 50-150 $\mu m$, while for vessels exposed to fouling, the roughness could be significantly higher. There exists an empirical method to account for this effect, and is given as

$$\Delta C_F = [110(HV_s)^{0.21} - 403]C_F^2 \tag{1.1.5}$$

where H is the roughness in $\mu m$ and V is the velocity in m/s.

### 1.1.4 Viscous Pressure Resistance

As indicated, viscous pressure resistance appears mainly as a result of flow separation. When the flow separates from the body, a low-pressure field occur causing a drag force in the opposite direction of the course. This phenomenon can occur at propeller shafts, shaft brackets, stabilizer fins, tunnel thruster openings, bilge keels, rudders and behind transom stern.

When the transom stern is wetted, the resistance coefficient can be expressed as

$$C_{DB} = 0.029 \frac{(S_{DB}/S)^{3/2}}{\sqrt{C_F}} \tag{1.1.6}$$

In cases where the transom stern is dry, the resultant force will be more significant and could be expressed by

For high-speed vessels with a transom stern, the water separates at the crossing between the transom, the sides and the bottom leaving the transom exposed to atmospheric pressure. The hydrostatic pressure that acts on all other parts of the hull below the waterline results in a resultant force.

$$C_{\text{transom}} = \frac{\rho g \int_0^H zB(z)dz}{\frac{1}{2}\rho V_s^2 S} = 2\frac{g}{V_s^2}\int_0^H B(z)\frac{z}{S}dz \tag{1.1.7}$$

### 1.1.5 Air Resistance

The presence of the superstructure of a ship will result in a resistance component from the relative wind (section A.3). The wind resistance is found by use of a drag coefficient, $C_D$

$$R_{AA} = \frac{1}{2}\rho_a V_{rel}^2 C_D A_p \tag{1.1.8}$$

In order to relate the wind resistance to ship resistance, the coefficient should be related to the wet surface to the hull, such that

$$C_{AA} = \frac{\rho_a C_D A_p}{\rho_w S} \tag{1.1.9}$$

As seen from equation (1.1.8), the resistance is proportional to the relative velocity squared, meaning that for low-speed vessels with relatively small projected areas exposed to wind, the wind resistance can be safely neglected. However, for high-speed vessels and vessels with significant superstructures, like cruise, container and cargo vessels the wind resistance can potentially contribute significantly to the total resistance.

The drag resistance coefficient in equation (1.1.8) is not trivial to determine, as it highly depends on the shape of the superstructure. There exist methods and standards with a reference value for different ship types, such as ISO (2015, Annex C) which includes a method to estimate the added resistance due to the wind. Usage of wind tunnel tests the drag coefficients is another method and enables the opportunity to find the coefficients for all angles. Such a test would be more suitable for the specific ship than the generalized standards and other literature suggests.

### 1.1.6 Wave Resistance

The displacement of the vessel changes the velocities along the hull, as mentioned in section 1.1.2. In the bow, the water will be forces outwards and on the stern, the water will return to the centerline of the vessel. With the increasing velocity in the bow, Bernoulli (1.1.10) states that the pressure decrease accordingly, however, at the free surface the pressure is equal to the atmospheric pressure. Hence, an elevation of the free surface ($\zeta$) ensure equilibrium, i.e a wave is generated.

$$\frac{1}{2}\rho V_s^2 + \rho g\zeta + p = \text{constant} \tag{1.1.10}$$

The same applies to the stern, where the relative velocity decrease and a wave system is generated. Generation of waves consumes energy and should, therefore, be considered as a resistance component. However, quantifying and isolating the wave resistance is not trivial. A method using in resistance scaling methods is to subtract the known and measured resistance components from the total resistance and then assume that the remaining resistance is the wave resistance (also known as the residual resistance), i.e

$$C_r = C_T - (C_F + \Delta C_F)(1 + k) - C_{AA} - C_{DB} \tag{1.1.11}$$

#### Analytic Estimation of Added Resistance of Ships in Waves

Faltinsen (1993) derived a formula for added resistance of ships in waves based on direct pressure integration (integrate pressure field from Bernoulli's equation). Valid for small Froude numbers, i.e. $F_n < \approx 0.2$ and head sea

$$\frac{\bar{F}_1}{\zeta_a^2} = \frac{1}{2}\rho g(1 + \frac{2\omega_0 U}{g}) \int_{L_1} \sin^2(\theta) n_1 dl \tag{1.1.12}$$

where $\omega_0$ is the circular frequency of oscillation of the waves and $U$ is the forward speed of the ship. $L_1$ is the exposed surface to the waves.

Gerritsma & Beukelman (1972) have derived a formula based on strip-theory approximation where the added resistance is quantified as

$$R_{AW} = \frac{k}{2\omega_e} \int_L (B_{33}^{(2D)} + U\frac{d}{dx} A_{33}^{(2D)}) V_{za}^2(x) dx \tag{1.1.13}$$

which is an integral along the length of the ship and $V_{az}(x)$ is the amplitude of the relative vertical velocity between the ship and the waves. $k$ is the wave number and $\omega_e$ is the encounter frequency.

### Effect of Bulb

A bow wave system and an independent stern wave system gives rise to the use of bulbs. By manipulation the effective length of the waterline, there should be possible to set up a bow wave system that is equalized by the stern system. In such a case the resulting wave resistance is decreased since the bow wave crest is canceled by the wave trough from the stern. Opposite effect if the bow crest is amplified by the stern crest making the contributing range of bulbs limited.

When the relations between the ship velocity and wave velocity are

$$V_s = \sqrt{\frac{\beta\lambda g}{n\pi}} = \begin{cases} \text{Increased resistance} & \text{for } n = 1, 3, 5... \\ \text{Reduced resistance} & \text{for } n = 2, 4, 6... \end{cases} \tag{1.1.14}$$

or in a more general way

$$F_N = \frac{V_s}{\sqrt{gL_{WL}}} = \sqrt{\frac{\beta}{n\pi}}, \quad 0.8 \leq \beta \leq 1.0 \tag{1.1.15}$$

where $\beta$ is the wave-making length of the ship.

## 1.2 Propeller Theory

### 1.2.1 Propeller Geometry

A propeller is a complex object where a three-dimensional model is required to give a complete description of its shape and geometry. However, for practical and applicational purposes, there is a need to describe the same geometry with a limited set of parameters. Figure 1.1 defines the main dimensions, such as the propeller diameter $D$, radius $R$, chord length $c(r)$, as well as boss diameter $d$ and boss radius $r_B$.

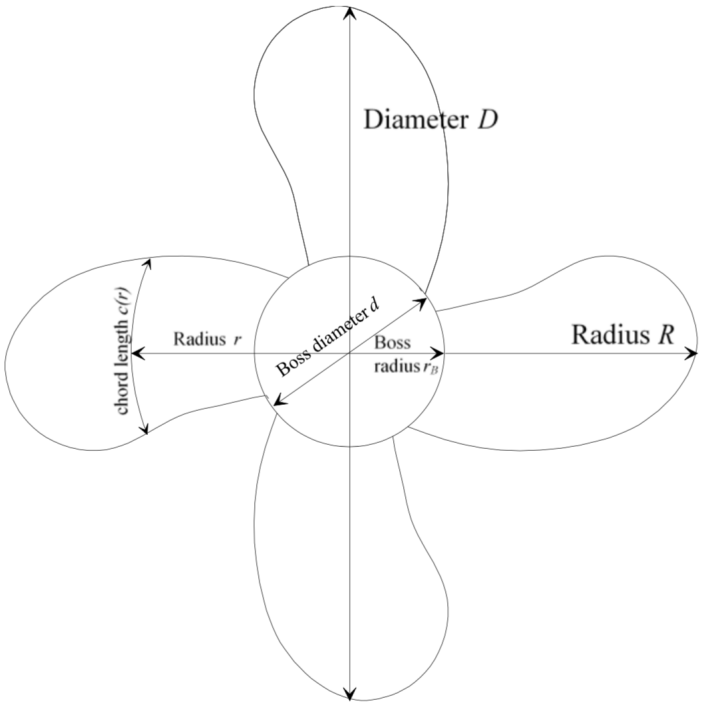



Figure 1.1: Propeller main dimensions, from Steen & Minsaas (2014)

Each vessel has its own uniquely designed propeller. Even though figure 1.1 shows a propeller with four blades there exists propellers with fewer and more blades and controllable pitch.

### 1.2.2 Performance and Efficiency

When measuring a propeller's performance and efficiency, it is based on the ability to produce thrust $T$. Mechanical torque $Q$ provided through the shaft for a given velocity of advance $V_a$

and the shaft's rotational speed $n$ is input to the efficiency analysis. The velocity of advance should not be confused with the velocity of the ship, but as

$$V_a = V_s(1 - w) \tag{1.2.1}$$

where $V_s$ is the speed through water and $w$ the wake.

The vessel's total resistance $R_T$ together with the speed through water reflect the effective power

$$P_E = V_s R_T \tag{1.2.2}$$

Together with the shaft power the propulsive efficiency can be defined as

$$\eta_D = \frac{P_E}{P_D} \quad P_D = 2\pi n Q \tag{1.2.3}$$

where $n$ is the rate of revolutions of the shaft. There are several efficiency components reflected in the propulsive efficiency which can be decomposed into

$$\eta_D = \eta_H \eta_R \eta_0 \tag{1.2.4}$$

where $\eta_H$ is the hull efficiency, $\eta_R$ the relative rotative efficiency and $\eta_0$ is the propeller efficiency. The relative rotative efficiency accounts for the effect of a non-homogenous wake field in front of the propeller. By separating the relative rotative efficiency from the propeller efficiency, the propeller efficiency could be measured in a more controllable environment. An open water environment could be replicated in a towing tank or cavitation tunnel where the performance of the propeller is only caused by the geometrical properties.

### 1.2.3 Open Water Diagram

When adding up all resistance components the total resistance the vessel will face for a given speed is known. At this constant, given speed equilibrium implies that the thrust of the propeller is equal to the resistance. Hence the propeller's performance is measured in the ability to produce thrust. For a given geometry this ability is measured in an open water test, where the presence of the hull is neglected. A model scale open water test reveals the characteristics of the propeller.

When conducting an open water test the propeller thrust, torque, rate of revolutions and velocity of advanced is measured. In a cavitation tunnel, the velocity of advance is represented

by the velocity of the uniform flow, while in a towing tank the velocity of the carriage is utilized for the same purpose. The measured parameters are made dimensionless by

$$J_a = \frac{V_a}{nD} : \quad \text{Advance Ratio} \tag{1.2.5}$$

$$K_T = \frac{T}{\rho n^2 D^4} : \quad \text{Thrust Coefficient} \tag{1.2.6}$$

$$K_Q = \frac{Q}{\rho n^2 D^5} : \quad \text{Torque Coefficient} \tag{1.2.7}$$

The torque is a parameter of which represents the power consumption by the system while the thrust represents the power produced, hence the efficiency of the propeller is

$$\eta_0 = \frac{TV_a}{2\pi nQ} = \frac{J_a}{2\pi}\frac{K_T}{K_Q} \tag{1.2.8}$$

These dimensionless parameters are presented in an open water diagram, where the thrust and torque coefficients are plotted as a function of the advance ratio align with the efficiency, as seen on figure 1.2.

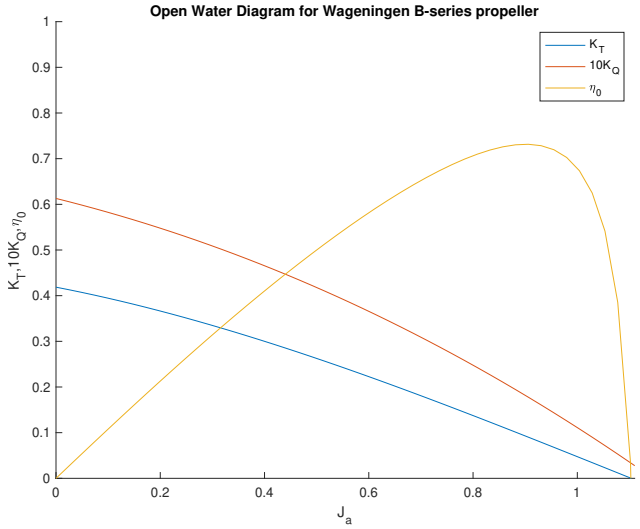



Figure 1.2: Wageningen B-series open water diagram

Figure 1.2 illustrates a propeller operating under a wide range of loading conditions; represented by the advance ratio. An advance ratio of zero implies that the velocity of advance is zero, resulting in the consumed and produced power while the vessel in standing still.

Negative thrust and torque coefficients are unphysical or would imply that the vessel is being towed.

Maximum utilization of the propeller's design is found at the propulsion point where the efficiency is maximum.

### Scaling of Propeller Characteristics

The open water diagram is the result of an open water test, which is a model scales version of the propeller. Scaling effects occur and the model scale experiment does not truly replicate the environment the full-scale propeller will be operating in. Whereas geometrical similarity of ensures length scaling is correct, kinematic similarity ensures that the ratio between velocities are satisfied

$$\lambda = \frac{L_F}{L_M}: \quad \text{Geometrical similarity} \tag{1.2.9}$$

$$\frac{V_F}{n_F D_F} = \frac{V_M}{n_M D_M} \Rightarrow J_F = J_M: \quad \text{Kinematic similarity} \tag{1.2.10}$$

Dynamic similarity ensures that the ratio between forces is the same for model and full scale. From decomposition of the total resistance of the vessel, the same sources of forces occur at the propeller: inertia, viscous, gravitational and pressure forces are present and dependent on the parameters found in geometrical and kinematic similarity, i.e.

$$\text{Inertia: } F_i \propto \rho V^2 L \tag{1.2.11}$$

$$\text{Viscous: } F_v \propto \mu V L \tag{1.2.12}$$

$$\text{Gravitational: } F_g \propto \rho g L^3 \tag{1.2.13}$$

These similarities address an issue with scaling from model to full size; ensuring geometrical, kinematic and dynamic similarity at the same time is not practically possible. The conflict occurs if trying to ensure that both inertia and viscous forces are properly scaled and since both these force contributions are significant for a propeller a scaling error occurs. The ITTC

method for correction of the thrust and torque coefficients are

$$K_{TS} = K_{TM} - \Delta K_T, \quad \Delta K_T = -\Delta C_D 0.3 \frac{PcZ}{D^2} \tag{1.2.14}$$

$$K_{QS} = K_{QM} - \Delta K_Q, \quad \Delta K_Q = -\Delta C_D 0.25 \frac{cZ}{D} \tag{1.2.15}$$

where $\Delta C_D = C_{DM} - C_{DS}$ and $P/D$ is the pitch ratio.

$$C_{CM} = 2(1 + 2\frac{t}{c})[\frac{0.044}{R_{nco}^{\frac{1}{6}}} - \frac{5}{R_{nco}^{\frac{2}{3}}}] \tag{1.2.16}$$

$$C_{DC} = 2(1 + 2\frac{t}{c})(1.89 + 1.62 \log \frac{c}{k_p})^{-2.5} \tag{1.2.17}$$

$$R_{nco} = \frac{c\sqrt{(V(1-w))^2 + (0.75\pi/J)^2}}{\nu} \tag{1.2.18}$$

### 1.2.4 Powering Prediction

To predict the required power delivered to the propeller shafts to obtain a certain speed, the link between the propeller characteristics and the vessel's resistance (section 1.1) must be established. Since equilibrium require that the total resistance is equal to the produced thrust a relation can found to be

$$\frac{K_T}{J^2} = \frac{R_T}{\rho_w (1-t) D^2 V_s^2 (1-w)^2} \tag{1.2.19}$$

where the full-scale propulsion point $J^*$ can be found by converting the open water diagram to a $K_T/J^2 - J$ - curve. These propulsion points for each velocity are used to find the coherent thrust and torque coefficients. Following, the rate of revolution and power are found from

$$RPM = \frac{60(1-w)}{D} \frac{V_s}{J^*} \tag{1.2.20}$$

$$P_D(kW) = \frac{2\pi}{1000} \rho_s D^5 (\frac{RPM}{60})^3 \frac{K_Q^*}{\eta_R} \tag{1.2.21}$$

# Chapter 2

# Machine Learning

This chapter will look into the fundamental of machine learning, the different segments and areas of utilization. The chapter first considers ordinary regression with a detailed description of the methods and techniques. Further, classification is exemplified with logistic regression before unsupervised clustering is discussed. These three branches in both supervised and unsupervised learning are building block for the neural network discussion, hence the thorough introduction to these methods within machine learning. At the end of this chapter, it should be clear that artificial neural networks are a clever composition of familiar mathematical models.

## 2.1 Introduction to Machine Learning

There is no specific and uniform definition of machine learning. However, a common informal definition was proposed by Mitchell (1997) stating that algorithms should improve their prediction performance based on experiences gained on performing a task, i.e

> A computer program is said to learn from experience $E$ with respect to some class of tasks $T$ and performance measured $P$, if its performance at tasks in $T$, as measured by $P$, improves with experience $E$.

There are different ways to train a program to perform task $T$ with a better performance P. One approach is to tell the algorithm what to learn by feeding it with some past data containing both input features and an output target. This is known as supervised learning. Unsupervised learning, on the other hand, is an approach whereas the algorithm itself try to find hidden structures and patterns without knowing explicitly what to look for.

Figure 2.1 illustrates the major branches within the field of machine learning.

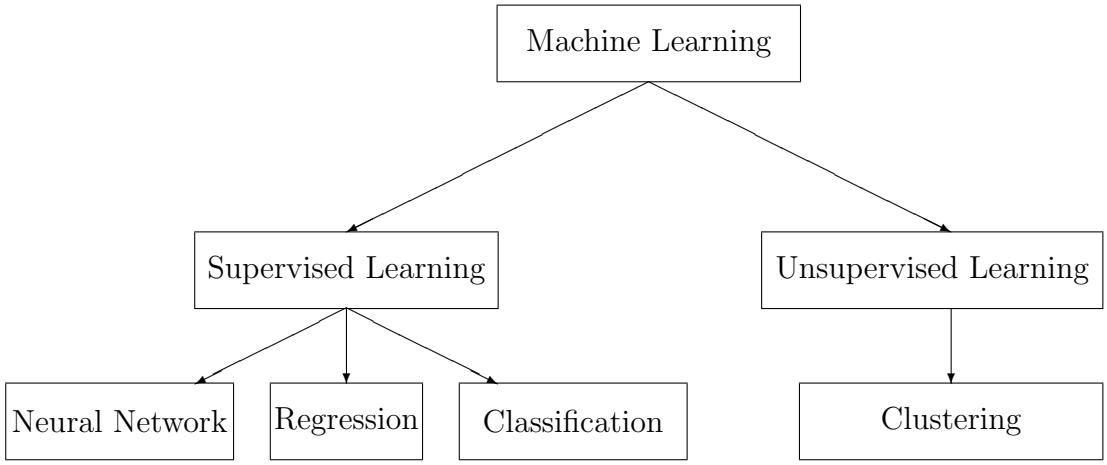



Figure 2.1: Major branches within machine learning

For the purpose of prediction fuel consumption for a ship in transit where the past operational data are available, both supervised and unsupervised learning are relevant. Supervised learning since the true fuel consumption is known, however, due to the numerous input variables and their dependency of each other, unsupervised learning could reveal hidden patterns in the data and will, therefore, be of interest in this project.

### 2.1.1 Supervised Learning

As indicated in figure 2.1, supervised learning is one of the major branches within machine learning. The term itself reveals the idea behind the technique; teach the computer to perform a task given an input scenario and the corresponding outcome. This is done by splitting a complete set of data into a training set and a validation set, utilize the training set to train the algorithm and then validate the performance with the validation set.

Within supervised learning, there exist other sub-branches that utilized the concept of training and validation set in different ways; in regression tuning of parameters such that the continuous hypothesis function makes the best fit with the sample data, and classification is trying to map the input variables into a set of discrete output categories.

### 2.1.2 Unsupervised Learning

In contrast to supervised learning, the concept of unsupervised learning is to find a structure among the unlabeled data. Instead of establishing a hypothesis function and make the best

possible fit to the measured data, unsupervised learning aims to categorize data without knowing in advance what to look for. In unsupervised learning the data sets are not divided into training and validation sets; the whole set is utilized at all time.

Clustering is the most common approach within unsupervised learning and there are multiple approaches and algorithms available, such as $K$-means clustering, hierarchical clustering and even unsupervised neural networks that self-organize the data (most probably in a structure similar to what $K$-means or hierarchical clustering do).

## 2.2 Regression

As indicated in figure 2.1, regression is one of the branches within supervised learning. In regression, the objective is to tune parameters such that a continuous hypothesis function makes the best possible fit to the sample data. In this section, the procedure of regression will be described.

### 2.2.1 Training a Model

When training a model in general, the objective is to find the algorithm and mapping function which fits the training set in best possible way. A training set for a supervised learning model consisting of $N$ samples with $M$ dimensions, the set of inputs variables (features) are given as $\mathrm{x}^{(\mathrm{i})} = (x_1, x_2, ..., x_M)^T$ where $i = 1, 2, ..., N$, such that all input variables are given as $\mathrm{x} = (\mathrm{x}^{(1)}, \mathrm{x}^{(2)}, ..., \mathrm{x}^{(N)})^T$ (be aware that the superscript has nothing to do with exponentiation). For each set of features $\mathrm{x}^{(i)}$, there exist an output target $\mathrm{y}^{(i)}$, such that $\mathrm{y}^{(\mathrm{i})} = (y_1, y_2, ..., y_M)^T$, and a pair of features and targets are $(\mathrm{x}^{(i)}, \mathrm{y}^{(i)})$ are known as a training example. The mapping of input feature to output target is done by an unknown function $f$, such that

$$\mathrm{y} = f(\mathrm{x}) + \epsilon \tag{2.2.1}$$

where $\epsilon$ is noise. The function $f$ is unknown and the idea in supervised learning is to come up with a hypothesis function $h$ that perform the mapping. The concept of establishing the hypothesis function $h$ is best described with linear regression.

### Linear Regression

The simplest version of regression is univariate linear regression (regression with one variable) where the input is mapped to the output through a relation on the form of $h_\theta(\mathrm{x}) = \theta_0 + \theta_1 \mathrm{x}$, where $\theta_0$ and $\theta_1$ are unknown tuning parameters.

The most common way of determining the tuning parameters are with the method of squared error function, where the idea is to minimize the difference between the hypothesis function and the sample data measurement, i.e., $\underset{\theta_0,\theta_1}{\text{minimise}}\ J(\theta_0, \theta_1)$, where

$$J(\theta_0, \theta_1) = \frac{1}{2N} \sum_{i=1}^{N} (h_\theta(x^{(i)}) - y^{(i)})^2 = \frac{1}{2N} \sum_{i=1}^{N} (\theta_0 + \theta_1 x^{(i)} - y^{(i)})^2 \tag{2.2.2}$$

and $J$ is known as the squared error function or the cost function.

Further interpretation of the cost functions reveals that the combination of tuning parameters that provide the minimum cost function value implies that the difference between the hypothesis and the true values are minimal. Hence the hypothesis function $h$ maps the input features to the output target with the best possible performance if $J(\theta_0 = a, \theta_1 = b) = 0$, i.e a perfect hypothesis function.

### Gradient Descent

A technique to establish the unknown tuning parameters is the gradient descent technique which is an iterative method for finding the minimum of a function. The idea of the method is to evaluate the derivative at a given point along with the function in order to determine in which direction the function value decrease by the highest rate. With the direction of the steepest increment established, the size of the step is defined by the learning rate parameter $\alpha$.

Figure 2.2 illustrate the surface elevation of irregular waves at a given time for a Jonswap spectrum. In the analogy of regression, the surface elevation represents the cost function $J$ with coherent tuning parameters. The scatter dots represent the iterative process of the gradient descent method. Hence the distance between the scatters are defined by the learning rate.

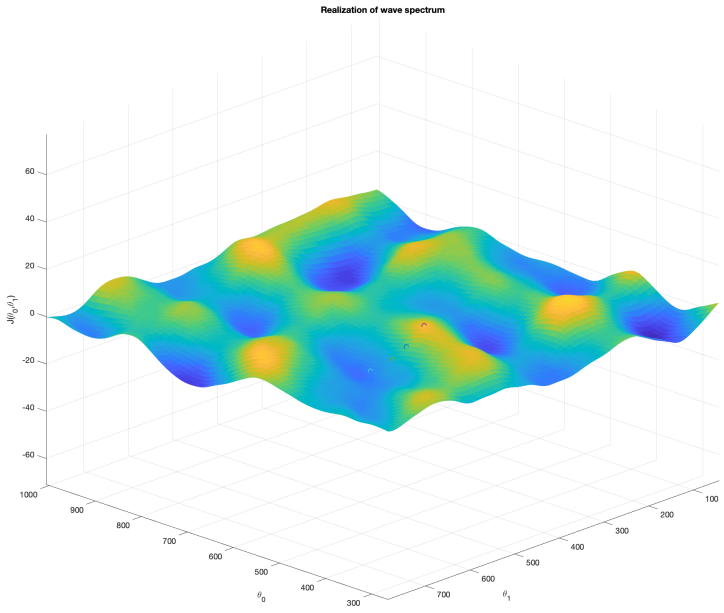



Figure 2.2: Irregular wave from Jonswap spectrum

Initial values for the tuning parameters are important for the final results, as the gradient method not guarantee that a global minimum is found. Just as important in terms of computational power is the choice of the learning rate, which determines the size of step from one iteration to the next, since the learning rate is proportional to the convergence rate (or divergence rate if chosen inappropriate).

The iterative algorithm for the gradient method is as follow

$$\theta_j = \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, ..., \theta_n) \tag{2.2.3}$$

where the partial derivatives are included. As a result, the cost function must be differentiable.

Establishing the cost function and then differentiate it the respect to all the tuning parameters is a computational expensive activity. This is illustrated by the cost function in figure 2.2, which consist of only two tuning parameters, but with a discrete increment of 0.1 the matrix containing the cost function is a 10 000 $\times$ 10 000 matrix. Each element in this matrix is a result of a sum over 10 000 evaluations (see equation (2.2.2)), resulting in a $\mathcal{O}(n^3)$ time complexity problem. Adding more complexity by introduction more tuning parameters and

higher order hypothesis increases the computational cost with a high rate.

### 2.2.2 Validate the Model

With the tuning parameters established, the hypothesis function and thereby the regression model is also known. These parameters are found based on the training set; a portion of the total available data set. The input features X of the validation set are tested on the regression model and the predicted outcome targets Y are evaluated with the true outcomes from the validation set.

A parameter describing how accurate the regression model mapped the input features to the output is the coefficient of determination or the $R^2$-value. The coefficient describe the ratio of the variance of the dependent variable $(\mathrm{X}, \mathrm{Y})$ and the independent variables $(\mathrm{Y}, \bar{\mathrm{Y}})$ where $\bar{\mathrm{Y}}$ is the mean of Y, i.e.

$$R^2 = 1 - \frac{\sum_i^{\hat{N}} (f^{(i)} - \bar{\mathrm{Y}})^2}{\sum_i^{\hat{N}} (\mathrm{y}^{(\mathrm{i})} - \bar{\mathrm{Y}})^2} \tag{2.2.4}$$

where $\hat{N}$ is the number of samples in the validation set.

Observe that the numerator in equation (2.2.4) is the same as the cost function established in section 2.2.1. However, whereas the objective with establishing the cost function was to minimize it to get the best possible fit represented by a $J = 0$, the coefficient of determination quantifies the degree of correlation between $\mathrm{Y}_{pred}$ and $\mathrm{Y}_{obs}$. If scatter occur in the validation set, meaning that $y_i$ fluctuate for similar $x_i$´s and otherwise, the coefficient of determination will be influenced in a way the gives a lower $R^2$-value. Hence the coefficient of determination is not an appropriate parameter to evaluate the quality of a fit but should be used as a guideline

### 2.2.3 Adding More Complexity

To this point the linear regression model consisting of two tuning parameters, $(\theta_0, \theta_1)$, and one feature $x^{(i)} = x_1^{(i)}$ have been illustrated. However, for a multivariate process a linear model may not be sufficient, hence the need for a more complex model. By assuming that the process is better described by a second-order polynomial, an input feature space would be fully described by

$$\mathrm{x} = [1, x_1, x_2, x_1^2, x_1x_2, x_2^2, x_1^2x_2, x_1x_2^2, x_1^2x_2^2] \tag{2.2.5}$$

which would require a hypothesis function $h(x_1, x_2)$. Determining the order of the hypothesis function is not trivial with a multivariate input feature. A too complex hypothesis function does not necessarily describe the behavior of the features and a too simple hypothesis will not see the details. These two phenomena are known as over- and underfitting within machine learning and could potentially compromise the prediction model if not taken care of.

Bishop (2013) exemplifies these phenomenon by generating a synthetic sample set with a sinusoidal function and some noice, i.e $f(x) = \sin(2\pi x) + \epsilon$. The scatter dots in figure 2.3 are the training examples, the green line is the underlying function and the red line is the polynomial hypothesis function of order M = 0,1,3,9. The order of the hypothesis is given on the form $h(x) = w_0 + w_1 x + w_1 x^2 + ... + w_M x^M$, where $w_i, i = 0, 1, ..., M$ is the tuning parameters.

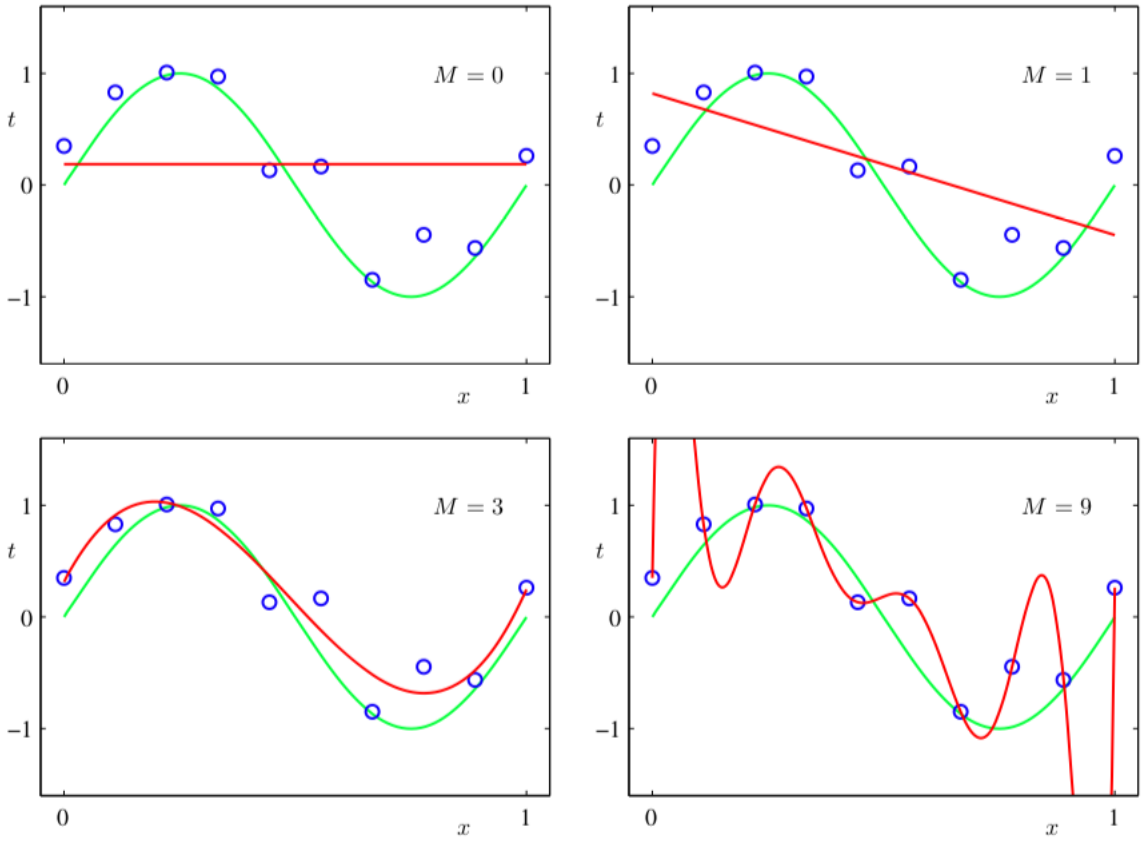



Figure 2.3: Illustration of over- and underfitting by Bishop (2013)

The constant and linear hypothesis functions (M = 0,1) are examples of underfitting; these do not represent the oscillating behavior of the system. The highest order (M = 9) is an example of overfitting. Even though the hypothesis polynomial passes through all training

points, the curve oscillates between the points in a non-sinusoidal way. The third order (M = 3) polynomial represents the sinusoidal function best among these four polynomials, even though minimization of the cost function not would state otherwise.

Figure 2.3 is an example of why the coefficient of determination (discussed in section 2.2.2) not should be used to determine the accuracy of the hypothesis. For M = 9, $R^2 = 1$ i.e. perfect fit. This is not by coincidence since the polynomial contain 10 degrees of freedom ($w$); exactly the same as the number of points in the data set so there exists a unique solution. However, the tuning coefficients are in this case finely tuned in order to fit all points, and the oscillating behavior is a result of these tuning parameters.

These faults reveal why over- and underfitting can cause inaccurate regression models even though the hypothesis seems to be a good fit to the data points. Since the power series of a sinusoidal function contain terms of all polynomial orders, one should imagine that a higher order polynomial would represent the true function better. In addition, the higher order polynomial contains all the lower order polynomials (the third order polynomial could be produced from the ninth order by eliminating the higher order components by $w = 0$), which is another argument in favor of the higher orders. However, the root of the problem is not in polynomials itself, it is the principle of minimizing the cost function that causes the problems. By introducing a regularization term to the cost function, the overfitting problem can be solved.

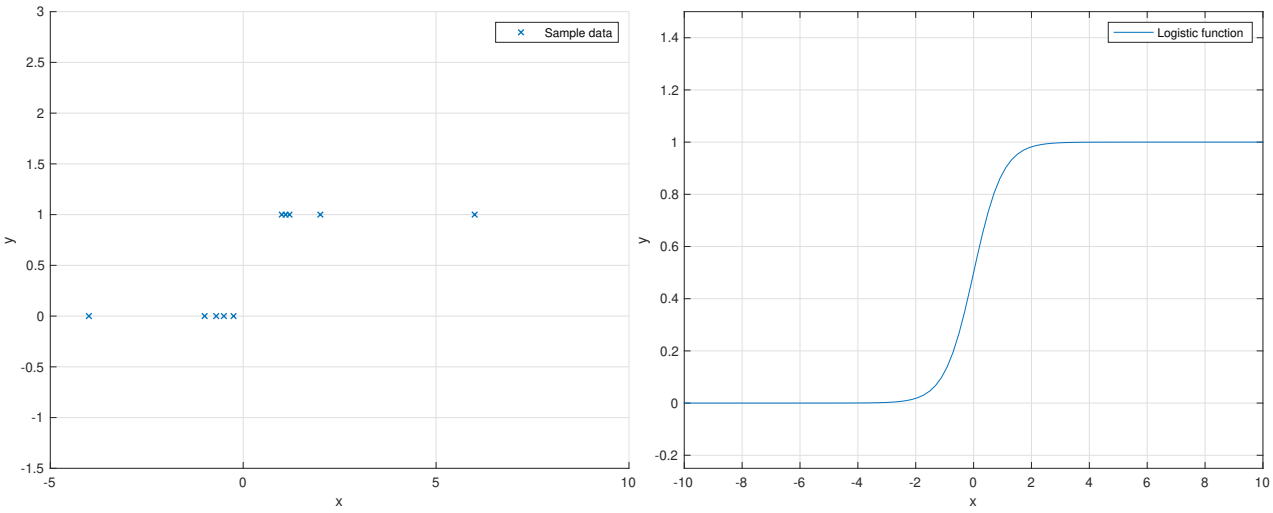
$$J = \frac{1}{2N}\sum_{i=1}^{N}(h_\theta(x^{(i)}) - y^{(i)})^2 + \frac{\lambda}{2}h_\theta(x^{(i)}) \tag{2.2.6}$$

where $\lambda$ is a regularization coefficient that ensures that the tuning parameters decay towards zero if they are not supported by a data point, also called weight decay.

## 2.3 Classification

Just as in regression, the principle of classification is mapping an input feature to an output target. Even though regression analysis is appropriate for some cases, it is not necessarily appropriate for all situations.

Figure 2.4a illustrates some input feature x and output target y. A linear regression model on these sample data turns out to be inappropriate as the input feature $x \to \pm\infty$ since that would give $y \to \pm\infty$.




(a) Sample data (b) Logistic function

Figure 2.4: Motivation for classisfication

### 2.3.1 Logistic Regression

Figure 2.4 illustrate that a new type of hypothesis and cost function are needed to make a mathematical model that fits the sample data in figure 2.4a. In logistic regression, the hypothesis is on the form

$$h_\theta = g(\Theta^T x), \quad g(z) = \frac{1}{1+e^{-z}} \tag{2.3.1}$$

The logistic function $g$ have the property of $0 \leq g(z) \leq 1$ which is more appropriate the for the sample data in figure 2.4a. Just as in regression, the tuning parameters $\theta$ occur, now in a matrix form $\Theta$ and the need to be determined. The same procedure is used in logistic regression as in ordinary regression; minimize a cost function with respect to the tuning parameters.

#### Logistic Regression Cost Function

The cost function for the logistic regression model is on the form

$$\text{Cost}(h_\theta(x^{(i)}), y^{(i)}) = \begin{cases} -\log(h_\theta(x)) & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases} \tag{2.3.2}$$

Summaries all contributions to establish the cost function

$$J = \frac{1}{m}\sum_{i=1}^{m} \text{Cost}(h_\theta(x^{(i)}), y^{(i)}) \tag{2.3.3}$$

Determining if tuning parameters is done by minimizing the cost function, just as in ordinary regression. Gradient descent method could be used for such a purpose, but was mentioned in section 2.2.1 computational costly. Other algorithms such as conjugate gradient, BFGS, and L-BFGS have the advantage of determining the learning rate $\alpha$ automatically and usually converge faster. However, they are more complex and is therefore not further considered in this project.

## 2.4 *K*-means Clustering

Whereas least error and logistic regression are examples of supervised learning, $K$-means clustering is an example of unsupervised learning. The output target y is in this case not known and the data sets consist of unlabeled attributes. Hence the intention for a clustering algorithm is not to predict an output target given a set of input features, but to find hidden structures and relationships between the attribute's values.

Suppose a given data set with $N$ samples and $M$ with unlabeled variables for each sample, such that the data sets are described by x = $(x^{(1)}, x^{(2)}, ..., x^{(N)})$ and each sample by $x^{(i)} = (x_1^{(i)}, x_2^{(i)}, ..., x_M^{(i)})$. The objective in the $K$-means clustering algorithm is to initially distribute $K$ clusters or centroids in the $M$-dimensional sample space and then assign each cluster to a group of data points where the inter-point distances are small.

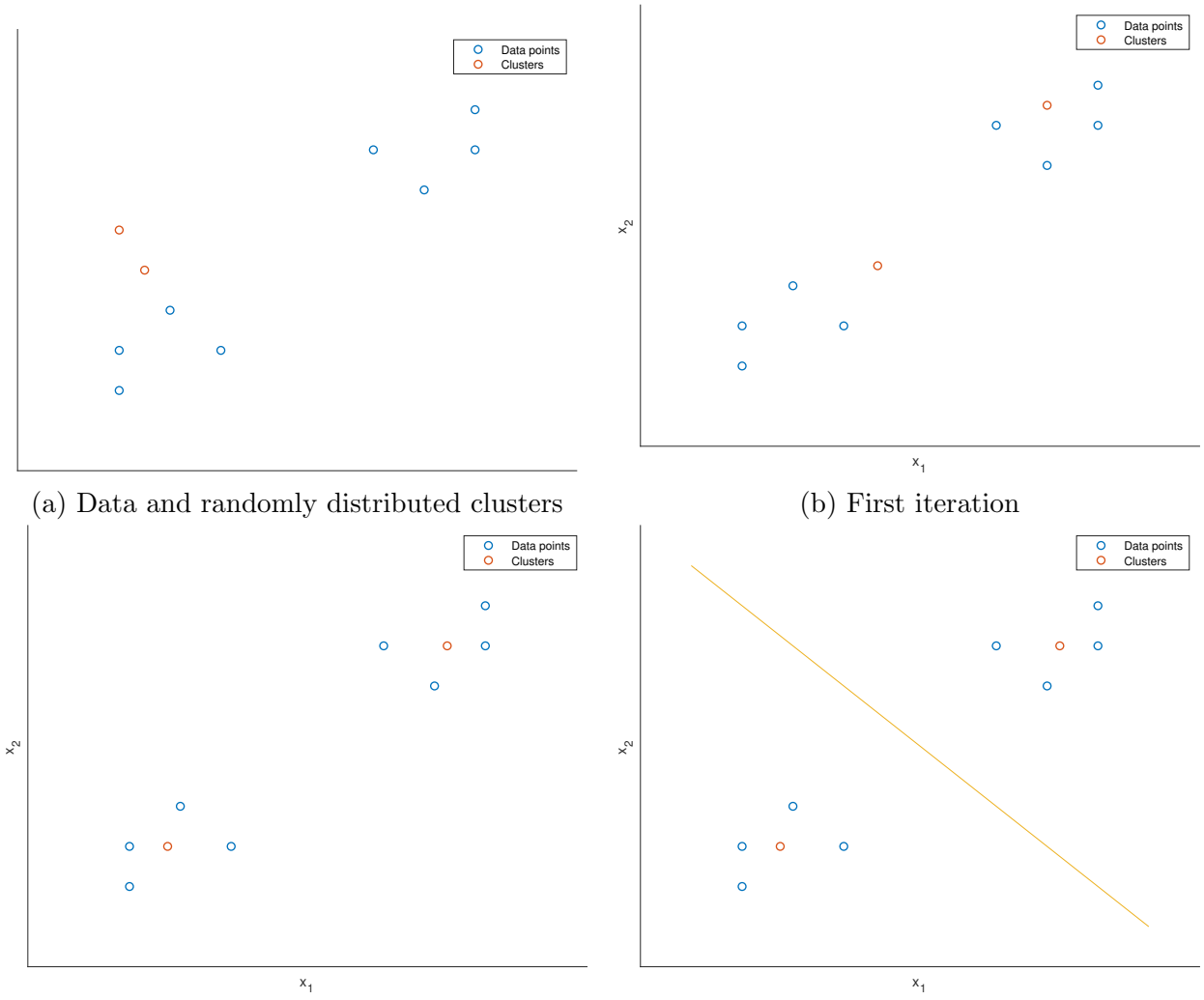
The iterative process first finds the special distance between each point and all clusters by

$$d_{nk} = \sum_{n=1}^{N}\sum_{k=1}^{K}\sqrt{(x^{(n)} - \mu_k)^2} \tag{2.4.1}$$

where $\mu_k, k = 1, ..., K$ is the clusters features in the sample space. Next, each cluster is assigning to the data points with the minimum distance between and then finally the features of each cluster is recalculated by the mean of the features from the assign data points.

By repeating this algorithm until the attribute values of the cluster converge the clusters will eventually gather in special areas where the input features are concentrated. Application of such an algorithm expresses data sets with thousands of measurements and multiple attributes with just a few clusters.

The whole algorithm is illustrated on figure 2.5, where figure 2.5a shows the eight data points (in blue) and two clusters (in red). It seems like there is a relation between the features $(x_1, x_2)$ forming two groups of data points where the inter-point distances are significantly lower among the two groups. On figure 2.5b the clusters features are updated by the mean of the assign data points. On figure 2.5c the cluster features are updated once again and finally on figure 2.5d the cluster features are unchanged and the convergence has been achieved.




(a) Data and randomly distributed clusters

(b) First iteration

(c) Second iteration

(d) Third iteration

Figure 2.5: Concept of $K$-means clustering

### 2.4.1 Elbow Method

How many clusters that should be used was from visual interpretation of figure 2.5a obvious. However, when the number of features for each sample increase, a visual interpretation is not possible.

A method to determine the optimal number of clusters is the elbow method; evaluate the sum of squared distances between each data point and the cluster. When the rate of change in sum approach to zero, it indicates that additional clusters do not represent a new group of related points. Figure 2.6 shows how the rate of change in the sum of squared distance decrease rapidly as the number of clusters increases, i.e. from k = 4 to k = 5 the change is marginal. Hence four clusters are sufficient in this synthetic data set.

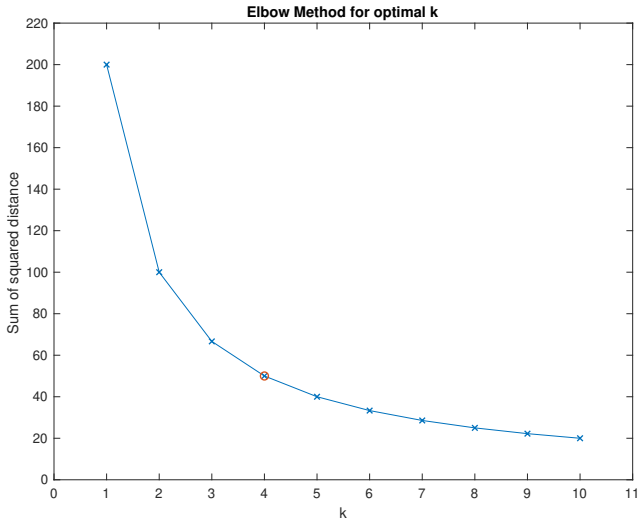



Figure 2.6: Elbow method

## 2.5 Principle Component Analysis - PCA

Similar to $K$-means clustering, the objective of a principal component analysis is to examine the variability in among the attributes in the data set. Instead of finding the hierarchy among the attributes, the idea in PCA is to convert a correlated variable into linear uncorrelated (independent) variables known as principal components. As a result, the dimension space of the data set could be reduced by either uses the principal components for further analysis or extract a few of the correlated variable to represent the variation in the data set.

Exemplified by an imaginary data set consisting of typical measurements on a vessel where the engine power, RPM, shaft torque, speed over ground, air temperature and heading are

available. It is reasonable to believe that the engine power, RPM, torque, and SOG are closely correlated, while a change in air temperature or heading not could be expected to be transparent in the engine power measurements. Hence a PCA analysis would most likely show that two principal components would be sufficient to describe the majority of variation on the data set. A PCA would, in this case, reduce the dimension space for the data set from six to two.

Dimension reduction is a powerful tool in data analytics and useful in machine learning, as training algorithms and prediction methods are complex and computer wise expensive in terms of time complexity. For further details, mathematics, and implementation, the reader is referred to Jolliffe (2011), Wold et al. (1987)

## 2.6 Artificial Neural Networks

Artificial neural networks are the final block from figure 2.1 within supervised learning. ANNs is inspired by the brain and aims to mimic how neurons communicate in order to process inputs into an output. From a biological point of view, neurons make up a network that responds to impulses from other neurons in the network. Neurons in the input layer triggers neurons in the hidden layers in which perform a calculation depending on the inputs which again trigger the output layer where the response is the result.

### 2.6.1 Simple Neural Network with Logistic Activation Function

A neural network is assembled by one or more neurons, so for exemplification, a simple network a neuron with logistic activation function (see section 2.3.1) is used. In fact, the model used in this example is replicate exactly the logistic regression example.

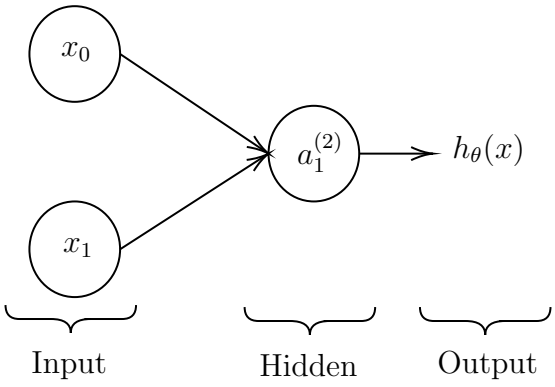



Figure 2.7: Simple neural model

Figure 2.7 shows a neuron receiving a bias unit $x_0$ and an input feature $x_1$. The neuron perform a calculation from its activation function $a_1^{(2)}$, which in this case is chosen to be the logistic function, as described by equation (2.6.1).

$$a_1^{(2)} = g(\mathbf{x}) = \frac{1}{1 + e^{\Theta^T \mathbf{x}}}, \quad \mathbf{x} = [x_0 \ x_1]^T \tag{2.6.1}$$

where $\Theta$ is a vector containing weights (or tuning parameters). Finally, the hypothesis function is established through a second activation function $h_\Theta(x) = a_1^{(3)} = g(a_1^{(2)})$ (recall the hypothesis function from section 2.3.1). The superscript refers to the layer whereas the subscript refers to the unit in that layer.

Activation functions (sigmoid functions) are function that transform any input $x$ such that $-1 \leq g(x) \leq 1$, which is an important feature in a neural network (Nielsen 2015). Other activation functions with the same properties are the hyperbolic tangent function $\tanh(x)$ and the error function $erf(x)$. A common feature for these functions is that they have derivatives, which will be shown to be important when training the network.

### 2.6.2 Adding More Neurons

The previous section showed the simplest neural model, but by combining multiple neurons in the hidden layer, more complex networks are possible. Each neuron can have a unique bias and weight, even unique activation functions are possible.

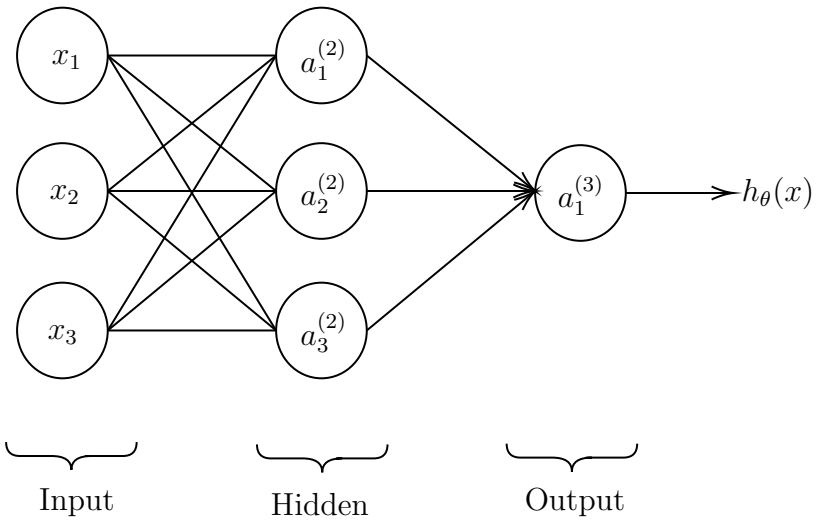



Figure 2.8: Neural model with multiple neurons

Figure 2.8 shows a neural model with three neurons in the hidden layer. All input features are parameters in each node and their biases and weights evaluated individually in the hidden layer. In the output layer, the outputs from the hidden layer are evaluated before the prediction is complete. The equation system corresponding to figure 2.8 is

$$\begin{aligned} a_1^{(2)} &= g(\Theta_{10}^{(1)}x_0 + \Theta_{11}^{(1)}x_1 + \Theta_{12}^{(1)}x_2 + \Theta_{13}^{(1)}x_3) \\ a_2^{(2)} &= g(\Theta_{20}^{(1)}x_0 + \Theta_{21}^{(1)}x_1 + \Theta_{22}^{(1)}x_2 + \Theta_{23}^{(1)}x_3) \\ a_3^{(2)} &= g(\Theta_{30}^{(1)}x_0 + \Theta_{31}^{(1)}x_1 + \Theta_{32}^{(1)}x_2 + \Theta_{33}^{(1)}x_3) \\ h_\Theta(x) = a_1^{(3)} &= g(\Theta_{10}^{(2)}a_0^{(2)} + \Theta_{11}^{(2)}a_1^{(2)} + \Theta_{12}^{(2)}a_2^{(2)} + \Theta_{13}^{(2)}a_3^{(2)}) \end{aligned}$$

where $\Theta_{i0}$ is the weight of the bias unit $x_0$. This algorithm starting from the input layer, evaluating the weights and biases in the hidden layer and eventually find a predictive value is known as forward propagation. Since the hypothesis is a key element in the cost function, forward propagation can be interpreted as an important tool when establishing the cost function.

### 2.6.3 Training a Neural Network

Training a neural network follows the same procedure as for linear and logistic regression. Establish a cost function that represents the difference between a measured sample point and the hypothesis function and minimizes this difference be changing the weights and biases (tuning parameters). For neutral network this logistic activation functions the cost function is

$$J = -\frac{1}{m}[\sum_{i=1}^{m}\sum_{k=1}^{K} y_k^{(i)} \log h_\theta(x^{(i)})_k + (1-y_k^{(i)}) \log(1-h_\theta(x^{(i)})_k))] + \frac{\lambda}{2m}\sum_{l=1}^{L-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_{l+1}}(\Theta_{ji}^{(l)})^2 \quad (2.6.2)$$

where $L$ is the total number of layers in the network, $s_l$ is the number of neurons in layer $l$, $K$ the number of output targets and $m$ the number of input feature. Note the regularization term known from section 2.2.3. Further interpretation of equation 2.6.2 illustrates the concept of forward propagation through the indices of the sums which starts with the input layer and ends up at the output layer

**Backward Propagation**

The cost function itself is not sufficient to train the network; finding the partial derivatives are equally important. In ordinary regression and logistic regression finding these partial derivatives involved differentiating a set of functions independently and update the tuning parameters until convergence (see equation (2.2.3)). However, finding these derivatives are not as trivial due to the multiple layers in the network.

Backward propagation is an algorithm that finds these derivatives and involves some advanced mathematical proofs which not will be elaborated further in this project. The algorithm is as follows

1. Evaluate the error in the output layer with the true sample result: $\delta^{(L)} = a^{(L)} - y$.
2. Back propagate and find the error in the second last layer $\delta^{(L-1)} = (\Theta^{(T)})\delta^{(L)} \cdot g'(z^{(L-1)})$.
3. Initialise $\Delta_{ij}^{(l)} = 0$.
4. For each layer and neuron, find $\Delta_{ij}^{(l)} = \Delta_{ij}^{(l)} + a_j^{(l)}\delta^{l+1}$.

By the mathematical proof, the partial derivatives $D_{ij}^{(l)}$ are

$$D_{ij}^{(l)} = \begin{cases} \frac{1}{m}\Delta_{ij}^{(l)} + \lambda\Theta_{ij}^{(l)} & \text{if } j \neq 0 \\ \frac{1}{m}\Delta_{ij}^{(l)} & \text{if: } j = 0 \end{cases} \tag{2.6.3}$$

With the partial derivatives known the training of the model is complete when the minimum of the cost function is found. The results are a set of biases and weights that are components in each neuron as it evaluates input and produces an output, either as an input to another neuron or as a predictive outcome.

### 2.6.4 Adding More Complexity

When training a neural network (or any other model) to describe a phenomenon, the objective is to find a function that fits the sample data in the best possible way by changing the biases and weights in the functions that perform this mapping. When the network becomes more complex, consisting of multiple layers in serial, parallel and possibly recursive the number of weights and biases increase significantly. In order to train all neurons to get an accurate prediction, a large number of training data is necessary. As seen from techniques as gradient descent and backward propagation, the computational complexity means that these processes are also computational time-consuming.

Imagining a neural network as a set of neurons that receives and sends impulses of different intensity to the neighboring neurons as a way of transferring information, the same analogy can be replicated in mathematics. The mathematical version of the neural network is a set of matrices where the transferred information is reviled by multiplication. For a network with multiple layers (deep learning) and where the number of neurons per layer decrease, the matrices describing the system will consist of multiple "blanks". As matrices $(n \times n)$ directly results in time and space complexity of $\mathcal{O}(n^2)$, these blanks are computationally expensive. That being said, step 2 in backward propagation consists of vector multiplication, making this technique a very efficient learning scheme.

When designing a neural network the same challenges as when deciding number of clusters (in $K$-means clustering section 2.4) arise; how many neurons makes the best prediction and which sigmoid function should are better. A *deep neural* network is a multilayered network capable of replicating more advanced phenomenon. However, just as more clusters not necessarily performed better than a smaller number of clusters, a deep neural network will not necessarily outperform a single layer network. The structure of the network should coincide with the complexity of the phenomenon.

Another issue with deep networks is the vanishing gradient problem (Nielsen 2015) which is a direct result of the increasing number of layers. In the backward propagation scheme, the partial derivatives of the cost function were evaluated, or more precisely the gradient of the activation functions. These derivatives are usually small and when multiplied with a weight $-1 \leq \theta \leq 1$ they get closer to zero. These small values are inputs to the next layer, making the result even closer to zero. Eventually, they gradient vanish, making the training of a deep network an issue.

# Chapter 3

# Preprocessing of Data

The following chapter describes the challenges of data preprocessing and how to detect measurement errors. This includes an overview of the different categorize of errors and available techniques to handle these errors.

## 3.1 General Methods

Systems measuring multiple variables at a high frequency over a long period of time generates large data sample sets. The quality of these data sets depends on a number of parameters, but the source or measuring device is the crucial factor (Frank et al. 2009). Despite the effort of reducing the deviation from the measured parameter and its true value errors do still occur.

Preprocessing of measurements after being logged is the next step in the attempt of obtaining high quality data sets. Two general methods that can be utilized for error detection is

1. **Statistical:** Simple statistical parameters such as mean and standard deviation could reveal outliers and other errors in a time series of measurements. Fitting the measurements to a probability model and evaluation of the measurements with a confidence interval would uncover the most likely and most extreme measurements. For initial studies a statistical interpretation are suitable for univariate analysis, whereas for multivariate analysis error propagation could be difficult to handle.

2. **Clustering:** Unsupervised machine learning is applicable for the purpose of the detecting errors in data sets. The same techniques used to categorize groups of measurements could be utilized to determine the inter-point distance. Measurements with a significant Euclidian distance from a cluster could therefore be considered as an outlier. As

indicated in section 2.4 this method is computationally complex and expensive.

Both methods have individually limitations, however, common for both are that the should be applied on a steady state part of the time series. Mean and standard deviation of a ship velocity over a year's time would not produce a logical result. By isolating time periods where e.g. the velocity measurements are fairly constant these methods are applicable. During acceleration of the vessel the machinery is exposed to heavy load and extreme measurements could occur and still describe the true nature the measured parameter.

## 3.2 Outliers

Outliers are measurements that deviates significantly from the other measurements in the same steady state period. It does not necessarily implies that the observation is wrong, but that there are some irregularities in the nature of what is measured. Telling the difference between an extreme value and an error in the measurement system could be difficult, especially if the system is multivariate and thereby dependent on a number of inputs.

A reliable measurement quantify the true nature of the phenomenon of interest such that other observations must be wrong. Technical issues when transmitting the data from the measuring device itself to the measurement system, or the device miss out an occurrence would compromise the overall quality and integrity of the data.

Abnormal situations occur from time to time and should be present in the data sets. Waves, wind and current are elements that could influence the operational profile both over a short and long period of time. These measurements are important to capture the true nature of the circumstances and should be evaluated with care before filtered out.

### 3.2.1 Statistical Outliers

Statistical properties from a time series of a measured parameter can be utilized to reveal outliers. For a measurement that deviates a given number of standard deviations from the mean could be classified as an outlier. More formal; if a measurement $\mathrm{x}_i \geq \mu_\mathrm{x} + \epsilon\sigma_\mathrm{x}$ or $\mathrm{x}_i \leq \mu_\mathrm{x} - \epsilon\sigma_\mathrm{x}$, where $\mu_\mathrm{x}$ is the mean and $\sigma_\mathrm{x}$ is the standard deviation of theo time series of x. $\epsilon$ is the number of standard deviations allowed before a measurement is defined as an outlier.

Using statistical properties blindly on a time series should not be considered best practice. Transient states in the time series where the measure system is changing its nature, a statistical cleaning scheme is inappropriate if it labels measurements as outliers solidly on a predetermined deviation ($\epsilon$) from the mean.

### 3.2.2 Smoothing of Data

Fluctuating data could occur due to the presence of environmental disturbance and result in a time series including noise and thereby a misleading representation of the physical measured system. Before mining the data and feature extraction, it could be beneficial to smooth the time series to remove spikes while at the same time preserve the information in the series; sharp transitions and gradients should be retained if these are intended.

Running mean or moving average (Gu & Zhou 2010) algorithm is useful for reducing fluctuations in short term data series and defined as

$$\hat{y}_k = \frac{y_k + y_{k-1} + ... + y_{k-(n-1)}}{n} = \frac{1}{n}\sum_{i=0}^{n-1} y_{k-i} \tag{3.2.1}$$

Denote that when calculating the successive values, the previous mean value occur in the sum such that an efficient algorithm ($\mathcal{O}(n)$) is

$$\hat{y}_i = \hat{y}_{i-1} + \frac{y_i - y_{i-n}}{n} \tag{3.2.2}$$

Running or moving median (Pitas & Venetsanopoulos 2013) is another algorithm that is useful for the purpose of smoothing the data series with the presence of outliers. Whereas a running mean would use an outlier value, a running median filter numerically sort $\mathcal{O}(\log(n))$ the entries within the window and find the median among these. The algorithm is therefor slower $\mathcal{O}(n\log(n))$ then the running mean, but have different capabilities.

$$\hat{y}_i = \text{median}(y_i, y_{i-1}, ..., y_{i-(n-1)}) \tag{3.2.3}$$

## 3.3 Missing Values

Missing values often occur in big data sets and can be difficult to handle. Depending on the measurement system and how it preprocess the data before publishing, missing values can appear as a frozen or constant value, or explicit set to zero or NaN. The two latter are easier to observe, while frozen or constant measurements can compromise the quality of the data if not taken care of.

For missing values, Frank et al. (2009) describe four techniques on how they should be treated.

1. **Deleting cases with missing values.** Delete all attributes from the same time instant in the time series. Whether this method is applicable depends on the size of the data set and how frequent missing values occur.

2. **Assign the most common value.** The missing values are replaced with either the mean of the existing values or the most probable value if a probability model is fitted to the existing values.

3. **Assign all possible values to the attribute.** For an attribute with boolean values (0 or 1) a duplicate of all attributes at the time instant of the missing value is created and the missing value is assigned with 0 or 1. For a continuous attribute values this method is not practical applicable.

4. **Interpolate between the known values.** Linear interpolation between the start and stop of the drop-out period. Applicable for continuous processes if the size of the data set is limited.

## 3.4 Drifting

Unlike outliers related to error in the measuring device or system, drifting is a phenomenon that develops over time. Instead of varying randomly about a mean, the mean of the observations change over time, i.e. $\mu_{x(t_0,t_1)} \neq \mu_{x(t_2,t_3)}$.

## 3.5 Visual Interpretation

Visual interpretation of the time series could uncover and reveal both systematic and random errors (Keaveny et al. 1997) such as outliers, missing values and drifting. Identification of outliers and missing values from mono variate analysis by statistical analysis is suitable. For multivariate analysis a scatter plot analysis illustrates drifting, unphysical relations and allows a subjective evaluation grounded in theoretical principles.

Experience and domain knowledge are key factors in a visual interpretation. Understanding underlying theories and assumptions in the phenomenon under investigation are equally important as well developed data cleaning software.

## 3.6 Normalization

The range and magnitude of the data varies from attribute to attribute in the data set. For quantification of physical properties the magnitude of each entry combined with its unit is necessary to describe the behavior of the system. However, for a learning algorithm, the magnitude itself can slow down the learning rate if not normalized (Ioffe & Szegedy 2015).

Discrepancy between two variables can lead to faulty interpretation of the variable relation. When K-means clustering (section 2.4) identifies the Euclidian distance between the entries, the variables with smaller magnitudes may be masked. Therefore, normalization of the input data is necessary so that all variables are in the same range and thereby weight. Two common methods (and suggested by Kotsiantis et al. (2006)) for scaling are

$$z_i = \frac{(z_{max} - z_{min})(x_i - x_{min})}{x_{max} - x_{min}} + z_{min} \tag{3.6.1}$$

$$z_i = \frac{x_i - \mu_x}{\sigma_x} \tag{3.6.2}$$

where $z_{max}$ and $z_{min}$ defines the range of the desired output, typically $-1 \leq z_i \leq 1$ or $0 \leq z_i \leq 1$. $\mu_x$ and $\sigma_x$ are respectively the mean and standard deviation of $x$.

## 3.7 Cleaning Methodology

Combined all the listed methods and techniques represents a general procedure for data preprocessing. Depending on the system or process under investigation the preprocessing methodology must customized to be applicable. For a continuous process such as an for an operating ship, a proposed procedure would be

1. Identify steady states on the time series. By evaluating the change or variance in moving window along the time series, steady states can be identified.

2. Data cleaning and smoothing for detection of outliers, signal spikes and noice. Running median filtering for smoothing the time series preserve sharp transitions and gradients in the signal, and filter out single outliers (Arce 2005). The running window passes through entry by entry and replacing the entry itself with the median of the neighboring entries. The size of the filter window should be adjusted to preserve the properties of the signal.

3. Linear interpolation of missing values. Alternatively, remove all attributes on the same time instant if it not reduce the size the data set by significance.
4. Visual interpretation of a multivariate analysis where known physical relations are evaluated. E.g observe if the measured power produced by the engine is coherent with the measured torque on the shaft. Such a multivariate analysis could reveal irregularities in either the measurement of power, torque or rate of revolution on the shaft, whereas the univariate analysis of the individual attributes would miss out any faults.
5. For data sets consisting of multiple variables, whereas some are expressed by a large and others by a small magnitude, normalization of each variable ensure that the learning routines not interpreter all variables with the same initial weight.

# Part II

# Method

# Chapter 4

# Data Description

This chapter describe the available data used in for this thesis, including their source, format and amount of data. The previous chapters have been general and an introduction to relevant theory in terms of data driven prediction and data preprocessing, as well as domain based theory within naval hydrodynamics.

It is primarily used two different data sets in this thesis; ship monitoring data from four sister cruise ships and weather data from different sources. The ship monitoring data are measurements from an on-board system on the vessels, whereas the weather data are simulated and interpolated data from the European Centre for Medium-Range Weather Forecasts (ECMWF), Norkyst800 and Tidetech. All weather data are collected from APIs where the time and location of the vessels are input features, such that these data sets include the current weather at the time and place of the vessels.

## 4.1 Ship Monitoring Data

The four sister cruise ships are equipped with a system which monitors a wide range of attributes. These attributes are continuously sampling and stores the measurements in a cloud based system. By use of an API these measurements are available on demand in real time.

All vessels are equipped the a diesel-electric propulsion system, meaning that diesel engines and generators produce electricity to the motors. In addition, the auxiliary propulsion system and hotel operations consume power provided by the generations. Despite being an important and significant consumer of power and thereby fuel, the auxiliary systems and hotel operations are not a part of the performance analysis; measurements and attributes related to the main propulsion system are of interest in this thesis.

Table 4.1 summarize the relevant attributes from the ship operational data

Table 4.1: Relevant attributes from ship monitoring system

| Variable name | Unit/Format | Comment |
|---|---|---|
| Time | YYYY-MM-DD HH:MM:SS | - |
| Heading | [°] | - |
| Latitude | [°] | - |
| Longitude | [°] | - |
| GPS speed | [kn] | Speed over ground |
| LOG speed | [kn] | Speed through water |
| Rate of revolution | [RPM] | Port and starboard |
| Thrust | [N] | Port and starboard |
| Torque | [kNm] | Port and starboard |
| Power motors | [MW] | Port and starboard |
| Power engines[1] | [MW] | - |
| Fuel consumption | [L/h] | - |
| Relative wind direction | [°] | Relative to heading |
| Relative wind speed | [kn] | Relative to speed over ground |

Other key features about the attributes and samplings are

- Measurements are averaged other a period of 1, 5 and 30 minutes, i.e. there are three data sets available for each of the four vessels containing all attributes from table 4.1.
- Continuously measurements from January 1st 2018 to February 22th 2019.
- None of the vessels are dry docked during this period, but hull and propeller cleaning are performed at an unknown frequency and time.
- For confidential reasons the vessels are named Vessel A, Vessel B, Vessel C and Vessel D in this thesis.

### 4.1.1 Operational Profile

All four vessels continuously measured all attributes during the thirteen and a half month which makes out the data sets. However, for the purpose of prediction fuel consumption for ships in transit, operational states as maneuvering and when in port are less relevant. Figure 4.1 shows the distribution of time the four vessels spend in each state; blue shows time in

[1] $2 \times 6.54$ MW and $2 \times 4.90$ MW.

port, orange shows time spend on maneuvering and yellow shows time spend in transit state. The states are defined by the speed over ground criteria given in table 4.2.

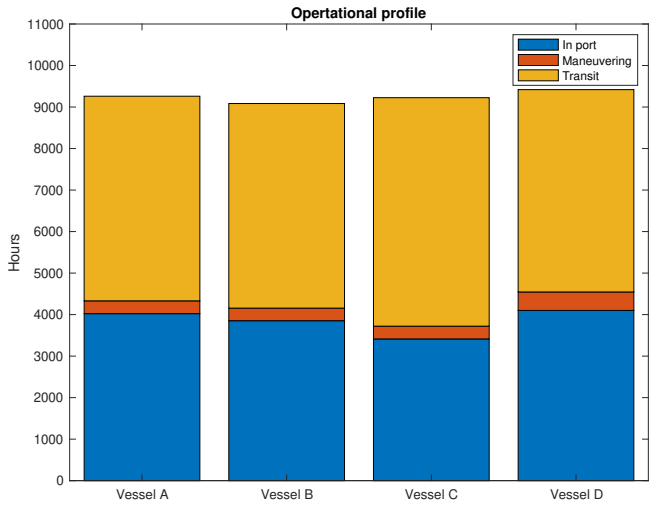



Figure 4.1: Operational profile from January 1st 2018 to February 22th 2019

Table 4.2: Speed over ground criteria for operational state

| **State** | **Limits [kn]** |
|---|---|
| In port | SOG $\leq 0.2$ |
| Maneuvering | $0.2 <$ SOG $\leq 5$ |
| Transit | $5 <$ SOG |

The observant reader notice that the height of each column varies from vessel to vessel, indicating that there are missing values among the speed over ground measurements.

## 4.2 Climate Data

In addition to the onboard measurement system describing internal processes the environmental conditions are in interest in order to quantify the performance. The presence of currents, waves and wind will contribute to either increase or reduce the necessary power to obtain a certain speed, hence the need for a precise representation of the environmental conditions.

### 4.2.1 Norkyst800

The Norkyst800 model (Albretsen et al., 2011) from Havforskningsinstituttet simulates ocean currents, salinity and temperature over a grid resolution of 800x800 meters on different water depths in the water column. At each grid point for multiple depths the current velocity is decomposed in a north ($u$) and east ($v$) component with units of m/s for each timestamp in a serial date format. To reduce the size of the data sets an averaged value of the current velocities at the surface and on 3 meters depth from July 1st 2018 to August 31th 2018 was utilized. Among the ship monitoring data time, latitude and longitude positions were available, hence the size of the data sets were reduced to only contain data coinciding with the time and location of the vessels.

Conversion of current velocities in global north and east components to ship longitudinal current velocity was done as described in appendix A.3.1. The resulting attributes of the data set was at the format shown in table 4.3.

Table 4.3: Format of Norkyst800 data set

| **Serial Date** | **Latitude** | **Longitude** | **Longitudinal** | **Longitudinal** | **North** | **East** |
|---|---|---|---|---|---|---|
| [s] | [°] | [°] | [m/s] | [kn] | [m/s] | [m/s] |
| (...) | (...) | (...) | (...) | (...) | (...) | (...) |

The model only contain data close to the Norwegian shore, whereas the ships are sailing in other national and international waters. Hence the need for a second source of weather data.

### 4.2.2 European Centre for Medium-Range Weather Forecasts - ECMWF

A secondary source of climate data describing the is obtained from ECMWF. ECMWF provide analysis and reanalysis for weather for more or less all offshore areas in the world. Based on observation and measurements from different key locations they simulate the atmospheric and oceanographic conditions. The trade-off is that these analysis are made on a courser resolution, i.e. 80x80km on a 6-hourly basis. Attempting to increase the resolution would be a result of an interpolation in both time and space.

The upside is that ECMWF provide both atmospheric and oceanographic analysis, whereas Norkyst800 is limited to current only. With that said the oceanographic analysis include properties of the waves and not current. Thus, the data sets from ECMWF is on the format as shown in table 4.4.

Table 4.4: Format of ECMWF data set

| Variable name | Unit/Format | Comment |
|---|---|---|
| Time | [s] | Serial date |
| U wind component | [m/s] | Eastward wind component at 10m above sea level |
| V wind component | [m/s] | Northward wind component at 10m above sea level |
| Mean wave direction | [°] | Spectral mean direction. Zero degree means waves propagation from north to south. 90 degrees means wave propagating from east to west |
| Mean wave period | [s] | Spectral mean wave period |
| Significant wave height | [m] | Defined as 4 times the square root of the area under the wave spectrum or the mean of the 1/3 highest waves |

### 4.2.3 Tidetech

A third source of data describing the climate is the data sets provided by Tidetech. In fact, Tidetech provides a wide range of simulated weather data obtain from different sources as well as their own product. Wave data are supplied by the ECEP WW3 model, ocean models from the Mercator (NEMO) model, atmospheric data from numerous sources (GFS, ECMWF and CMC) and tidal elevation and currents using in-house developed methods. In addition, Tidetech provide supplementary simulations such as on ice, temperatures in ocean and air, humidity and atmospheric pressure, but these are not utilized in this thesis.

Hence the Tidetech data set provide a more detailed product than Norkyst800 and ECMWF in terms of climate data. There are coinciding data in Tidetech with other sources and an unknown interpolation in time and space must be conducted to get the full data set. The format of Tidetech data set is as shown in table 4.5.

Table 4.5: Format of Tidetech data set

| **Variable name** | **Unit/Format** | **Comment** |
|---|---|---|
| Time | [s] | Serial date |
| U wind component | [m/s] | Eastward wind component at 10m above sea level |
| V wind component | [m/s] | Northward wind component at 10m above sea level |
| Mean wave direction | [°] | Spectral mean direction. Zero degree means waves propagation from north to south. 90 degrees means wave propagating from east to west |
| Mean wave period | [s] | Spectral mean wave period |
| Significant wave height | [m] | Defined as 4 times the square root of the area under the wave spectrum or the mean of the 1/3 highest waves |
| Current magnitude | [kn] | |
| Current direction | [°] | Relative to north. 0 degrees means from north to south, 90 degrees means from east to west. |

# Chapter 5

# Methodology

This chapter reflects the methodology and the steps involved to predict the fuel consumption for vessels in transit. It should be by the end of this chapter clear for the reader how the different fields of studies in the previous chapters are relevant, and how they contribute to the solution.

## 5.1 Preprocessing of Data

Chapter 3 introduced and suggested a general procedure on data preprocessing and chapter 4 presented the available attributes both from the operating ship and the climate data. Quantifying the quality and evaluate their reliability through uni- and multivariate analysis are essential for trustworthy result in the end.

Hence a univariate analysis of the attributes in table 5.1 are necessary. To exclude extreme values from the time series, the steady states are identified from analysis of the rate of revolution on the propeller's shaft.

Whereas a univariate analysis could exclude extreme values and missing values from the time series, it could be difficult to evaluate whether their values are unphysical or bias. A multivariate analysis on the other hand could reveal such phenomenons.

Chapter 1 introduced the open water diagram as summary of the characteristics of the propeller where speed through water, thrust, shaft torque and rate of revolutions were input variables (density of the water and propeller diameter are constants). This summary would therefore reveal whether any of these variables deviates from their expected values relative to the other variables.

Table 5.1: Summary of the preprocessed univariate attributes

| Variable name | Unit/Format | Comment |
|---|---|---|
| Speed | [kn] | Speed through water and speed over ground |
| Wind | [kn] | Attribute analysis of wind direction and speed from operational data, ECMWF and Tidetech |
| Fuel | [L/h] | Fuel consumption |
| Power | [MW] | Power produced by the engines/generators |
| Power | [MW] | Power consumed by the motors |
| Torque | [kNm] | Torque measurements on the shaft |
| Thrust | [kN] | Thrust measurement |

From mechanical engineering the relation between fuel consumption and provided power from a diesel engine is familiar. Studies of the efficiency of such engine evaluate the ratio between the delivered power (or torque) and the potential energy in the fuel. As the open water diagram summary the characteristics of the propeller, a specific fuel consumption curve summarize the efficiency of a combustion engine as a function of the engine thrust.

From a hydrodynamic point of view, the speed through water are usual more relevant for the study than the speed over ground. Whereas the speed over ground are measured by an accurate and well-developed GPS system, the speed through water measurement devices are considered less reliable. In theory, the difference between speed through water and speed over ground should reflect the current, both ocean and tidal current. A comparison between the measured current on the vessels and the simulated currents would to a certain degree reflect whether the speed through water measurements are reliable. For deep water voyages[1] the simulated current and measured current are expected to coincide more than closer to shore where local differences and lack of resolutions in the simulations could compromise the solution. Both the Norkyst800 and Tidetech model are suitable for the purpose of validation of the speed through water measurements.

Cross validation of the climate data and the coinciding attributes are of interest. As mentioned in chapter 4, the Tidetech data sets are a merge of different data sources and simulations where conventions are different. Conventions and reference frames are defined in appendix A.3.1.

[1] Deep water in this context means voyages far away from shore and coastlines

Table 5.2: Summary of the multivariate analysis

| Variable name | Involved variables |
|---|---|
| Current | Speed through water and speed over ground. Comparison with Norkyst800 and Tidetech |
| Propeller characteristics | Open water analysis: thrust, torque, rate of revolutions and speed through water |
| SFC | Specific fuel consumption: fuel consumption and engine thrust |

## 5.2 Decomposing Resistance

With the available attributes described in chapter 4, different methods for prediction of fuel consumption appears and are excluded. Suitable methods are also dependent on the preprocessing results and the quality of the measurements. The benefit of having reliable thrust measurements allows a resistance decomposition on the form

1. Frictional resistance: $C_f = f(R_N), \quad \Delta C_f = f(V_s, H, C_f)$
2. Air and wind resistance: $C_{AA} = f(V_s, \gamma_{rw})$
3. Residual resistance: $C_r = f(V_s, H_s, T_p, \ \psi, \beta_w, T_{AP}, T_{FP})$
4. Total resistance: $C_T = f(T, V_s)$

The third point is however challenging to find analytically or empirically. A method is to rearrange the problem and find the residual resistance in the same manner as in scaling from model to full size resistance. Equation (1.1.11) is applicable for the same purpose, now in a different scenario.

As indicated in the same point, the residual resistance is a function of several parameters, all of them with different level of accuracy, randomness and variation. Different methods and approaches on how to describe the variation and randomness are given in table 5.3.

Method 1 takes advantage of the ISO15016 annex D (STAWAVE1) procedure for determining the added resistance due to waves. The procedure assume limited wave heights and relative wave direction as described in table 5.3. Note that the procedure accounts for the added resistance due to waves, such that the wave generation from the vessel itself is not included. Hence the need for a residual component that reflects the resistance from the wave system initiated by the vessel, $C_{r0}$.

Method 2 is an extend of method 1. Whereas method 1 is limited by the wave height and direction, method 2 exploit the excluded sea states from method 1 by use of a data driven machine learning algorithm.

Method 3 do not include the ISO-standard like the two first methods did. The residual resistance is still assumed to be a sum of the calm water wave resistance and the added resistance due to waves, but the contribution from waves are solidly described by a data driven machine learning model.

Method 4 tends to describe the residual resistance by the use of machine learning. In contrast from the first three methods, method 4 do not separate the calm water wave resistance from the resistance added resistance due to waves.

Method 5 is an extend of method 4, and exploit the fact that the speed through water could be a to significant weight in a learning algorithm. This method use a data driven machine learning model to describe the error between the predicted and true total resistance as a function of the sea state. Hence method 5 implicitly assume that the residual resistance components (as in method 1-3) should be considered independently.

Table 5.3: Different approaches for determining the residual resistance

| | **Approach**[2] |
|---|---|
| Method 1 | $C_r = C_{r0} + C_{r1}$<br>$C_{r0} = f(H_s \to 0, T_p \to 0, \text{stw})$<br>$C_{r1} = \text{ISO15016} = f(H_s < 3\text{m}, \beta_{rw} < 45°)$ |
| Method 2 | $C_r = C_{r0} + C_{r1} + C_{r2}$<br>$C_{r0} = f(H_s \to 0, T_p \to 0, \text{stw})$<br>$C_{r1} = \text{ISO15016} = f(H_s < 3\text{m}, \beta_{rw} < 45°)$<br>$C_{r2} = ML(H_s > 3\text{m}, \beta_{rw} > 45°)$ |
| Method 3 | $C_r = C_{r0} + C_{r1}$<br>$C_{r0} = f(H_s \to 0, T_p \to 0, \text{stw})$<br>$C_{r1} = ML(H_s, T_p, \beta_{rw}, \text{stw})$ |
| Method 4 | $C_r = ML(H_s, T_p, \beta_{rw}, \text{stw})$ |
| Method 5 | $C_r = ML(H_s, T_p, \beta_{rw}, \text{stw})$<br>$C_r^e = C_T^{\text{pred}} - C_T^{\text{true}}$<br>$C_r^e = ML(H_s, T_p, \beta_{rw})$ |

[2]ML: Machine learning

## 5.3 Predicting Resistance

By assuming at this stage that a sufficient model exist for the residual resistance component as a function of the parameters listed in the previous section, predicting the total resistance will consist of adding the different resistance components together. From the operational data, the speed through water and heading are the only two attributes required, while from the climate data waves and wind parameters are required.

At this point it should be clear for the reader how the analogy with training set (chapter 2) coincide with decomposing the resistance to estimate the residual resistance and composing the resistance components with the test data sets.

### 5.3.1 Measuring Performance

There are multiple methods to quantify the performance of the trained model; visual interpretation of the time series of the predicted and true target and by statistical properties such as MAE, MAPE, MSE, NMSE and REP Ghelardoni et al. (2013), Elattar et al. (2010) i.e

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{5.3.1}$$

$$\text{MAPE} = 100\frac{1}{N}\sum_{i=1}^{N}|\frac{y_i - \hat{y}_i}{y_i}| \tag{5.3.2}$$

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{5.3.3}$$

$$\text{NMSE} = \frac{1}{\Delta^2 N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad \Delta^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{y})^2 \tag{5.3.4}$$

$$\text{REP} = 100\sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} y_i^2}} \tag{5.3.5}$$

$$\text{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{5.3.6}$$

### 5.3.2 Bootstrapping of Model

Another method to evaluate the performance is to perform a so called bootstrapping of the model. The idea is to generate synthetic data and observe how the model responds to the data. This is somehow similar to Monte Carlo simulations in terms of generating data to test a model, despite that in bootstrapping some attributes are predetermined whereas others are their true value.

For the purpose testing the models in method 2-5, bootstrapping is performed by observing how the models responds when e.g the wave height in the testing set is overwritten to a fixed value. Hence the variation in residual resistance should be a result of varying wave period, relative wave angle and speed through water. This method could reveal whether the models responds as expected to a change in the inputs.

## 5.4 Thrust to Fuel Consumption

At this stage it is assumed that the resistance is sufficiently described by empirical and data driven methods such that the remaining part is to relate the predicted thrust to fuel consumption. There are multiple methods and approaches to do this conversion; follow the energy transformation from thrust, through the shaft torque, to the power from the motor, to the power from the engine and the coherent fuel consumption for the given engine power based on the specific fuel consumption. All these steps require accurate and reliable measurement in order to reduce the propagation in variation.

### 5.4.1 Thrust to Motor Power

Relating provided thrust from the propeller to the torque on the shaft is dependent on the characteristics of the propeller. These characteristics are described by the open water diagram where speed through water and rate of revolutions are input variables.

Preceding from torque, motor power is the following measurement of the energy transformation. Determining the motor power from torque follows known relations, either from (1.2.21) or by the more general torque - power relation[3]

$$P = \frac{2\pi n Q}{\eta_m} \tag{5.4.1}$$

[3]Note: motor power $P$ refers to the consumed power by the motor, hence the efficiency term $\eta_m$.

### 5.4.2 Motor Power to Engine Power

For these diesel-electric cruise vessels, the motor power ($P$) and engine power ($P_D$) would deviate significantly due to the consumed power from the hotel and other auxiliary power consuming processes. However, quantifying the auxiliary power consume by physical relations is not suitable nor accurate, hence a mean difference between the motor and engine power would represent this consume, i.e.

$$P_a = \overline{P_D - P} \tag{5.4.2}$$

Major varying auxiliary power consuming processes on a cruise vessels should be accounted for when using equation (5.4.2) to reduce spread and randomness in the prediction. Such processes could be production of fresh water (reverse osmosis) and air-conditioning in warm climate, which are dependent on external factors related to the geographical location.

### 5.4.3 Engine Power to Fuel

From section 5.1 the specific fuel consumption relations were established with the intension of serving the purpose of relate engine power to fuel consumption. Depending on the result from the multivariate analysis the SFC curve could be established from operational data, or from documentation from the producer of the engines.

# Part III

# Results

# Chapter 6

# Data Analysis

This chapter summarize the results of a thorough data analysis of measured attributes on vessel B with the 5 minutes average data set. Simulated entries from ECMWF, Norkyst800 and Tidetech are used for verification purposes.

The structure and format of this chapter is similar to the equivalent chapter in the project thesis (Wahl, 2018), but the content and analysis is conducted again with new techniques and thereby some new conclusions.

## 6.1 Data Cleaning, Smoothing and Steady State Identification

Ensuring data integrity before mining and variable analysis, filtering the data for noice and error and smooth out transitions are beneficial. Each variable in the data set where treated according to point 1-3 in section 3.7; a running median filter with window length of 5 remove spikes and noice, while a running variance window of length of 5 identify steady states (RPM-attribute were used for this purpose). Missing values within each steady state was linearly interpolated between the neighboring entries when the number of consecutive missing values not exceeded 3 entries (15 minutes). For missing values over a longer period of time, all measurements at these time stamps were excluded.

As a result, the RPM time series on figure 6.1 shows the filtered signal, including the steady states regions. Note how the time series miss out some values after noon May 18th where a consecutive period of NaN values occurred and were removed from the data set.

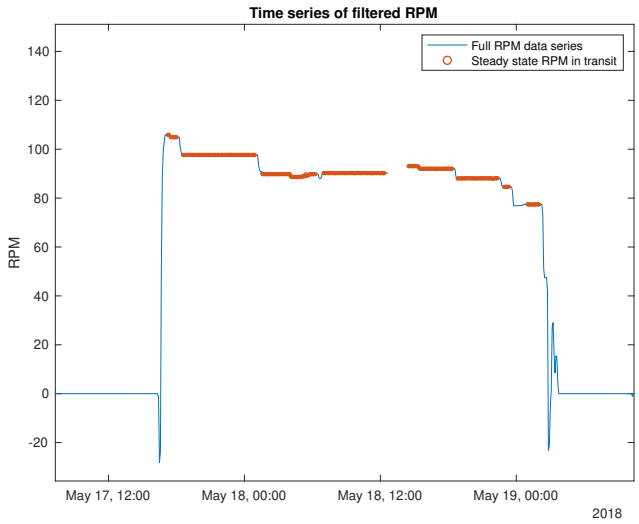



Figure 6.1: Identification of steady states along filtered RPM time series

## 6.2 Univariate Attribute Analysis

### 6.2.1 Speed

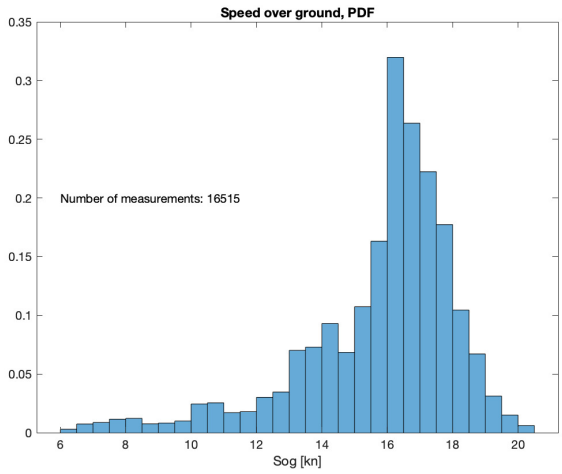
Measurements of speed through water and speed over ground are analyzed separately and in combinations. The vessel measure speed over ground using an onboard GPS system and is initially assumed to a sufficiently accurate for the purpose of this thesis. Measuring the speed through water on the other hand is traditionally considered more difficult to perform with a high level of accuracy.

Figure 6.2a and figure 6.2b shows the distribution of respectively speed over ground and speed through water after filtering and cleaning. If any in the raw data, the figures shows that outliers are removed or smoothed out in the time series. Among the remaining steady state data, the figures indicate that the most common cruising speed is in the range of 15-18 knots (SOG).

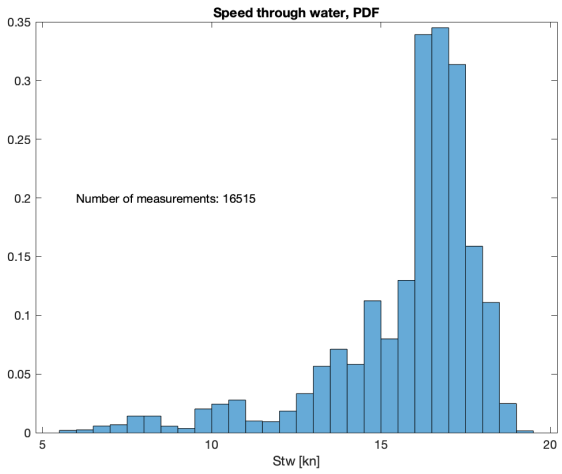
However, it seems like there are differences between the distributions in figure 6.2; speed through water measurements shows a denser and more narrow distribution around 17 knots than speed over ground, and the speed over ground measurements exceed 20 knots. Hence a more informative relation is shown in figure 6.3, where the filtered (figure 6.3a) and raw (figure 6.3b) speed over ground is illustrated in a scatter plot agains speed through water.

The difference between speed through water and speed over ground should reflect the current[1], which is shown in figure 6.3 by the color index of the scatters. In addition to the scatter plots,

[1] Current = STW - SOG




(a) PDF of speed over ground




(b) PDF of speed through water

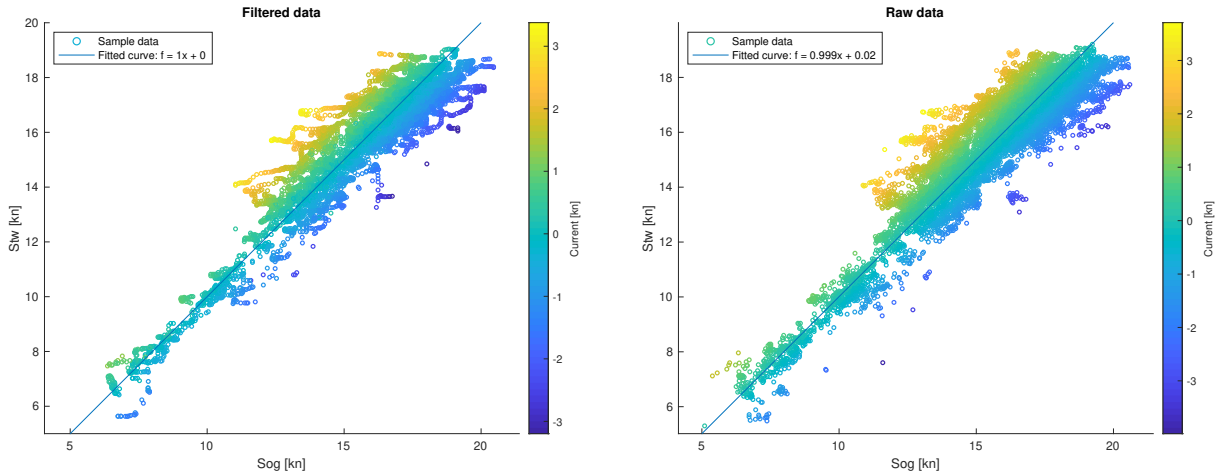
Figure 6.2: Velocity distribution

a linear regression line is to be found in both figure 6.3a and figure 6.3b. When operation in the absence of current, the speed over ground and speed through water measurements should be equal and upstream operation should indicate a positive current. The regression line indicate that the vessel operates equally (based on the regression coefficients) in an upstream and in a downstream environment. Based on this analysis there are no reason to suspect any bias error.

Another aspect that emerge from figure 6.3a is the alleged increasing current with increasing speed measurements, i.e. the spread between SOG and STW are significantly larger for SOG $>$ 10 knots. The most extreme current measurements indicate both up- and downstream current at $\sim$ 3 knots, which should be further evaluated by comparison with independent sources at the time and location of interest.

### Norkyst800

When comparing the measured current with the simulated current in the Norkyst800 model, the result do not show a strong correlation between the entries, as figure 6.4 indicate. The Norkyst800 measurements are (as described in section 4.2.1) converted relative to the heading such that the longitudinal current is plotted in figure 6.4 against the measured current. Hence the transverse current component is neglected.




(a) Filtered speed over ground and speed through water

(b) Raw speed over ground and speed through water

Figure 6.3: Scatter plot of SOG and STW

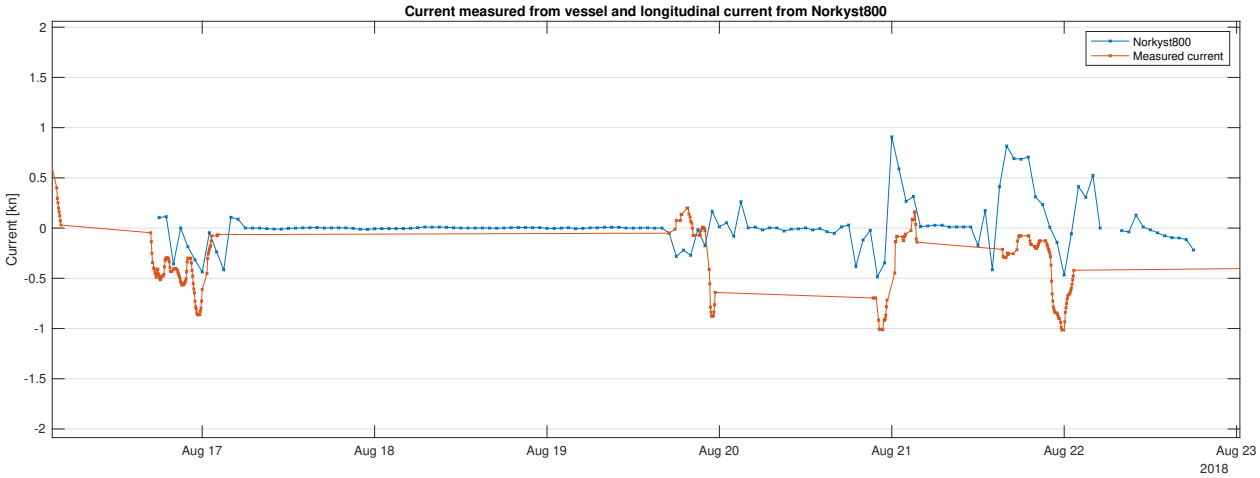



Figure 6.4: Measured current from vessel and longitudinal current from Norkyst800

The available data from Norkyst800 is limited to less than a week (the time the vessel were within the range of the model), and spend during that week a significant amount of hours in port. However, what is not shown in figure 6.4 is that in the time period around August 17th, the vessel sailed in open water between the British Islands and the west coast of Norway. In this period of time, the measurements and simulated entries are fairly similar in terms of magnitude and direction (up- or downstream). Whereas from August 20th to August 22th the vessel sailed in sheltered water along the Norwegian coast and the differences are more frequent and significant. Local differences in e.g. tidal water is a possible explanation to the

deviating current measurements.

As a result, the Norkyst800 model could neither verify nor contradict whether the speed through water measurements are reliable and further analysis is required.

### Tidetech

The second data set containing simulations regarding current is the Tidetech data set. Unlike the Norkyst800 model, Tidetech provide simulations all over the world and is therefore more applicable for purpose of comparing the measurements from a world wide sailing vessel.

Figure 6.5 shows the measured current from the vessel and the simulated longitudinal current from Tidetech over a two week time period. The results are over this particular period more coinciding than for the Norkyst800 model. Short term differences occur at all time, while the longer trends are more similar. Note how the extreme values ( $> 3$ kn) measured from the vessel are of significantly larger magnitude than the simulated values at the same time and place.

The measured current seems to be more fluctuating and could benefit of being smoothed through a stricter low-pass filter than the running median filter with length of 5 entries. However, all entries in the data set are a mean value over the previous 5 minutes, hence applying another stricter filter be obtain a better fit with the Tidetech model could potentially compromise the integrity of measurements.

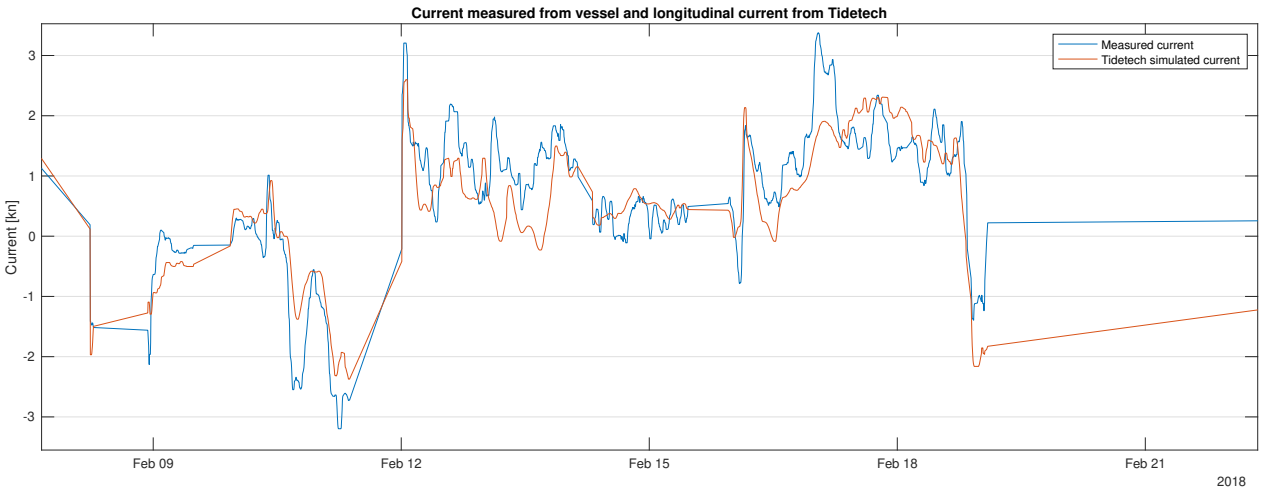
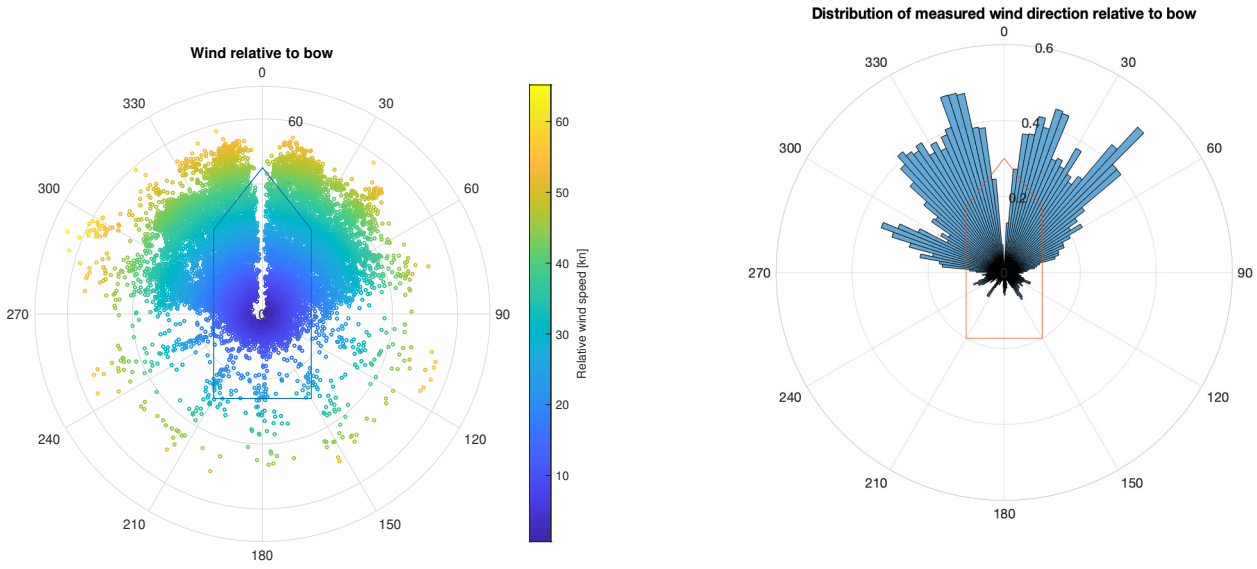



Figure 6.5: Measured current from vessel and longitudinal current from Norkyst800

As a result, there are reasons to believe that the speed through water measurements are somewhat reliable. Further analysis including the speed through water measurements should also evaluate the sensitivity of using the measured or simulated values.

### 6.2.2 Wind

The wind is quantified by the direction and magnitude by an onboard measuring device. Hence the measured properties are the relative direction relative to the bow and relative to the speed over ground. Both these features are illustrated in figure 6.6a where the angular position describe the relative angle to the bow, and the radial position (and scatter color) describe the relative wind speed. The angular position refers to the source of the wind, meaning that a relative wind direction of 90 degrees implies that the wind is approaching from the starboard side.




(a) Relative wind direction and speed

(b) Measured wind direction relative to bow

Figure 6.6: Measured wind properties

Figure 6.6b illustrate an important aspect from the wind direction measurements; it seems like head wind in the range of $\pm 5$ degrees rarely occur, which is questionable. Outside this blind zone there are frequent occurrences for all angular positions, indicating that measuring device have the required resolution to measure head wind.

A more likely explanation is the location of the measuring device itself. Ideally the device should have been mounted such that the presence of other physical objects not disturb the wind flow, e.g in the top of a mast. However, figure 6.6b indicate that this is not the case for this device. Consequently, there are reasons to believe that the measured wind direction are compromised and the wind properties should be described by either ECMWF or Tidetech.

### Wind from ECMWF and Tidetech

By comparison with the hindcast data from ECMWF and Tidetech the wind assumption of unreliable wind measurements were to a degree verified. After conversion to the same reference frame and sign convention the wind hindcast from the two sources are showed in figure 6.7 and figure 6.8 for January 2018. The figures show a strong correlation between the two hindcasts, both with respect to the direction and speed, whereas the measured wind properties varies frequently and randomly.

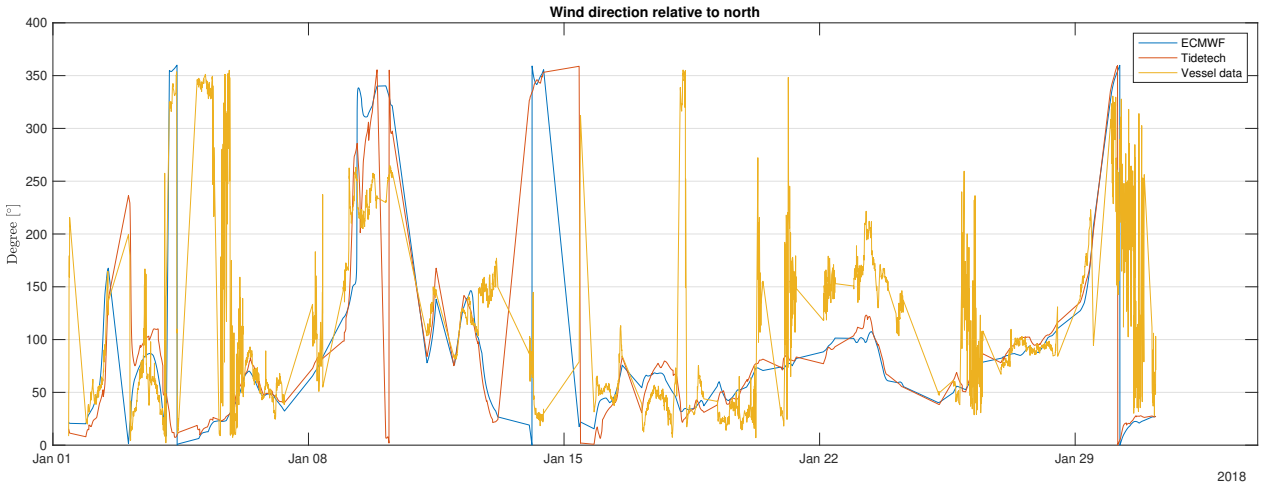



Figure 6.7: True wind direction from ECMWF, Tidetech and vessel

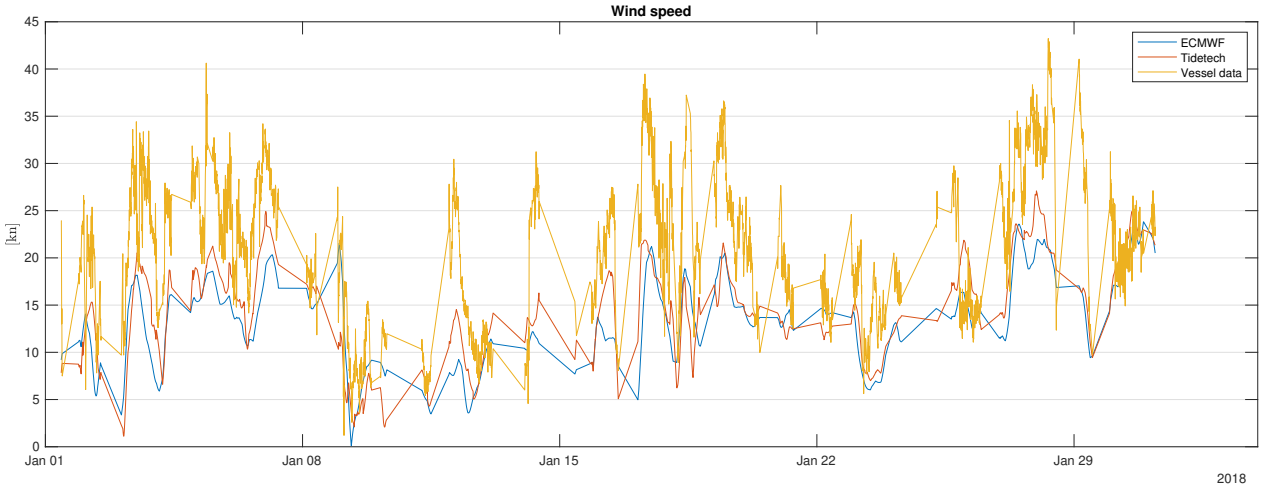



Figure 6.8: True wind speed from ECMWF, Tidetech and vessel

Whether ECMWF or Tidetech are more accurate is difficult to determine with the given

data. The important result is that when wind properties are involved in further analysis the hindcast data are the preferred data source.

### 6.2.3 Thrust

Equivalent to the total resistance experienced by the vessel is the thrust provided by the propeller. Thus, a reliable measurements of the thrust is valuable for quantification of hydrodynamic performance.

The vessel with its twin screwed propulsion configuration measure the thrust on both port and starboard side which is plotted against each other in figure 6.9. The color index on the scatter represents the rate of revolutions and the fitted curve shows the best linear regression between port and starboard thrust.

While the vessel sails on a steady course and heading and not need to compensate for any environmental external forces, the thrust on port and starboard side should be similar (under the assumption that the hull is symmetrical and no other geometrical properties cause a turning moment).

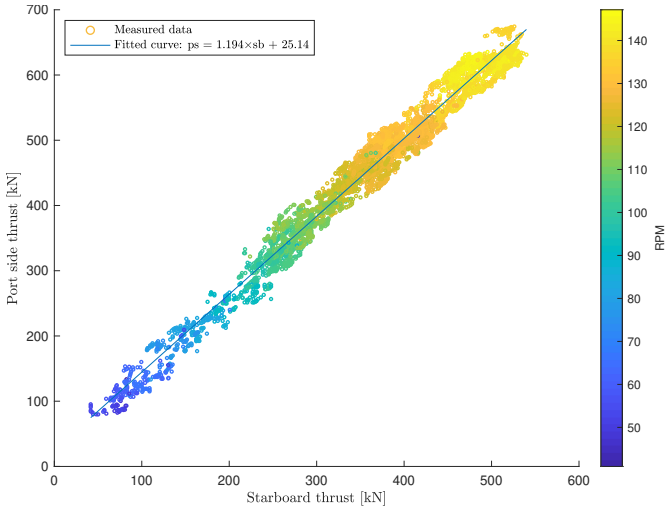
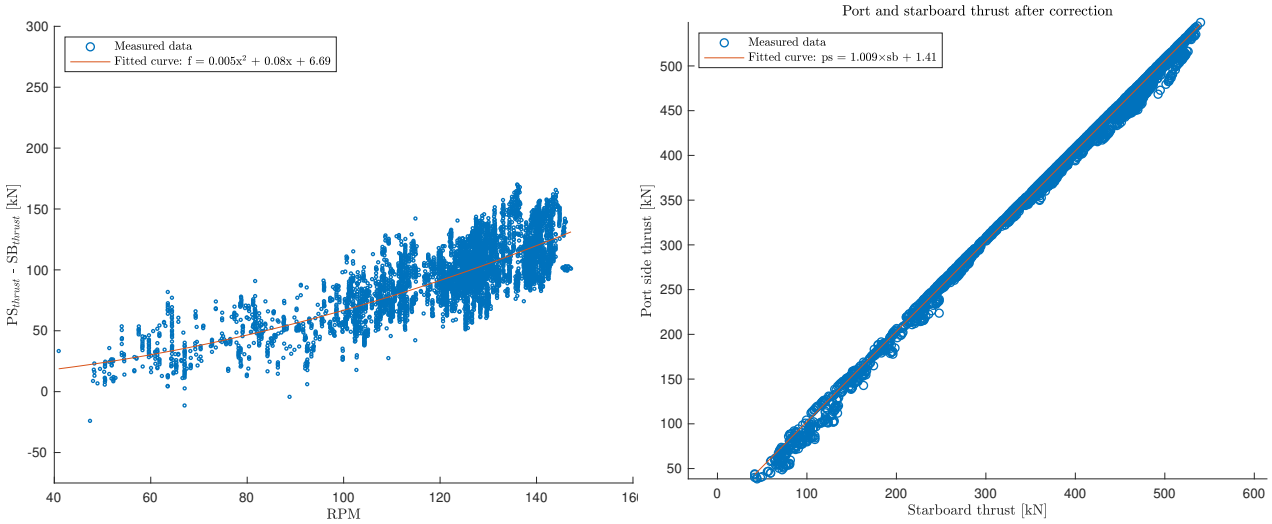



Figure 6.9: Port and starboard thrust scatter plot with regression line: $PS_{thrust} = 1.194 \times SB_{thrust} + 25.14$

Figure 6.9 shows that there are irregularities between the thrust measurements where the port side measurements are consistently of larger magnitude than for the starboard side. Further, the regression coefficients shows that extrapolation to zero thrust at starboard side gives a port side thrust of 25.15 kN. In addition to the deviation in the zero point extrapolation, the port side thrust is increasing with a rate of 19.4% faster than the starboard side.

Evaluation of how the port and starboard thrust measurements deviates shows a correlation with the rate of revolution[2], as illustrated in figure 6.10a. The figure includes a least square polynomial regression line of second order with regression coefficients as given in table 6.1 where x = RPM.

By correction of the port side thrust according to the regression coefficients in table 6.1 the measurements shows in figure 6.10b a more linear correlation with a minor bias in the extrapolation.




(a) Deviation between port and starboard thrust as a function of RPM

(b) Port and starboard thrust after correction according to regression line in figure 6.10a

Figure 6.10: Error analysis of thrust measurements

Table 6.1: Regression coefficients from deviation analysis of thrust measurements

| **Format** | **$ax^2 + bx + c$** |
|---|---|
| a | 0.005 |
| b | 0.08 |
| c | 6.69 |

Figure 6.10 and the analysis of the thrust deviation assume that the port side measurements are too larger and that starboard side is more reliable. There are no aspects in this analysis that support this assumption, hence the port side could be the benchmark for thrust measurements. Further analysis is required (section 6.3.1) and depending on the findings, the error analysis could be performed with respect to the port side measurements if necessary.

[2]The same correlation is found with the speed through water, torque, motor power etc.

### 6.2.4 Torque

Next after thrust, torque is the second measure of energy conversion from speed through water to hydrocarbons in the fuel.

Figure 6.11 shows the correlation between port and starboard side torque including a linear regression line. Based on these regression coefficients, there are no reason to believe that the measurements have some bias or other random error.

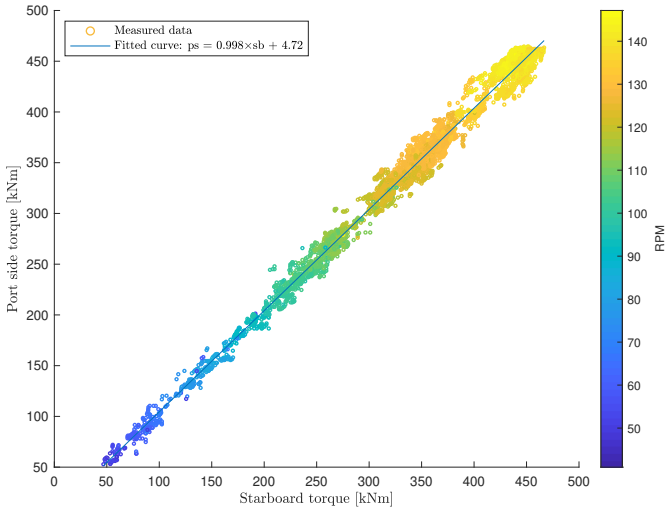



Figure 6.11: Port and starboard torque scatter plot with regression line: $PS_{torque} = 0.998 \times SB_{torque} + 4.72$

### 6.2.5 Motor Power

Motor power and shaft torque are closely related as both are a measure of how the energy transform over the motor from electrical power to mechanical torque. The consumed power of the port and starboard motor are shown in figure 6.12. Note how the linear regression line indicate a strong correlation between port and starboard side, with a minor extrapolated bias and spread.

The vertical spread from the regression line could either be related to faults in the measurements, or to physical properties as maneuvering, external environmental forces or even a non-symmetric hull. Figure 6.13 shows how the torque and power on port and starboard side respectively are correlated through relation (5.4.1).

Figure 6.13 shows significantly less vertical spread than when evaluating the power consume from port and starboard side at the same time instant. Thus, it is reasonable to believe that

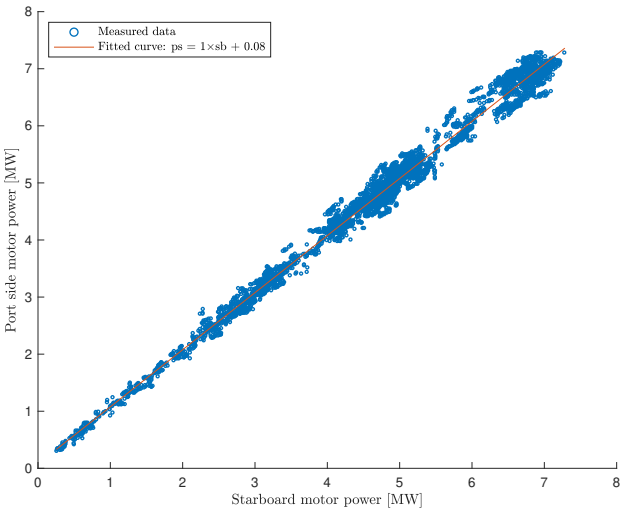



Figure 6.12: Port and starboard motor power scatter plot with regression line: $PS_P = 1.00 \times SB_P + 0.08$

both the power, torque and rate of revolution measurements are all reliable. The figures also illustrate that further preprocessing with respect to detection of outliers are unnecessary for the involved attributes, even thou a handful of entries deviates substantial. The minor deviation from the mean line illustrate that the measurements are in fact measured, not calculated from known physical relations.

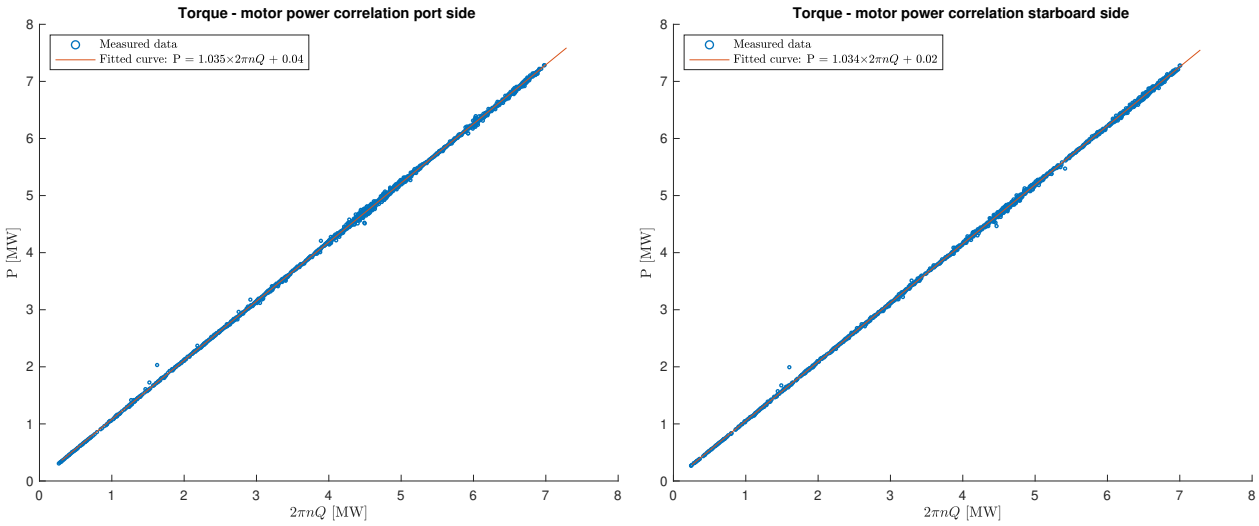
The slope of the regression line in figure 6.13a and figure 6.13b illustrates the efficiency of the motors; $2\pi nQ/P \approx 0.96$. An efficiency in the range of $\eta_m = 0.96$ for an electrical motor is not unlikely (Ådnanes 2003).

As a result, both torque and motor power measurements are considered reliable due to the correlation between the attributes when separating port and starboard side. Hence, the vertical spread in figure 6.11 and figure 6.12 were explained by external factors and not faults in the measurement system.

### 6.2.6 Engine Power

Unlike the motor power which produce power for propulsion purposes, the main engines produce power for all systems and operations on the vessel. This include the propulsion as well as operation of the hotel, auxiliary propulsion systems, fresh water production, air conditioning etc. Hence the engine power could be analyzed relative to the motor power by equation (5.4.2).

Figure 6.14 shows how the engine power varies relative to the motor power. The scatter




(a) Port side (b) Starboard side

Figure 6.13: Torque - motor power correlation

indicate that there is an underlying change happening when the engine power increase while the motor power stays the same, hence there are two linear parallel correlation between the power measurements. When adding the sea temperature represented by the color index, another correlation is illustrated. It seems like the increase in produced power by the engine is partly described by the increase in environmental temperature.

By separating the high ($\geq 17.5°C$) and low temperatures, linear regression shows the additional power consumed as a function of environmental temperature. When sailing in hotter climate, the hotel and other auxiliary processes consume 4.11 MW, while in colder climate 3.53 MW is the mean consume. Thus, there are no reason to believe that the engine power measurements are unreliable, but the fraction of produced power spend for propulsive purposes are dependent on environmental factors.

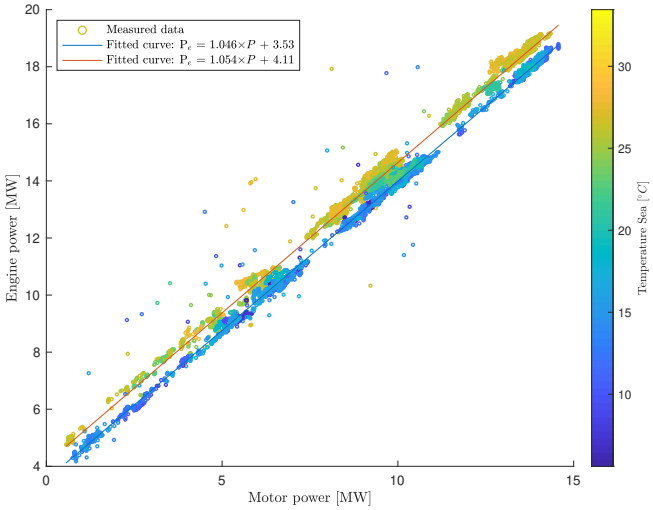



Figure 6.14: Engine power relative to motor power

## 6.3 Multivariate Analysis

### 6.3.1 Propeller Characteristics

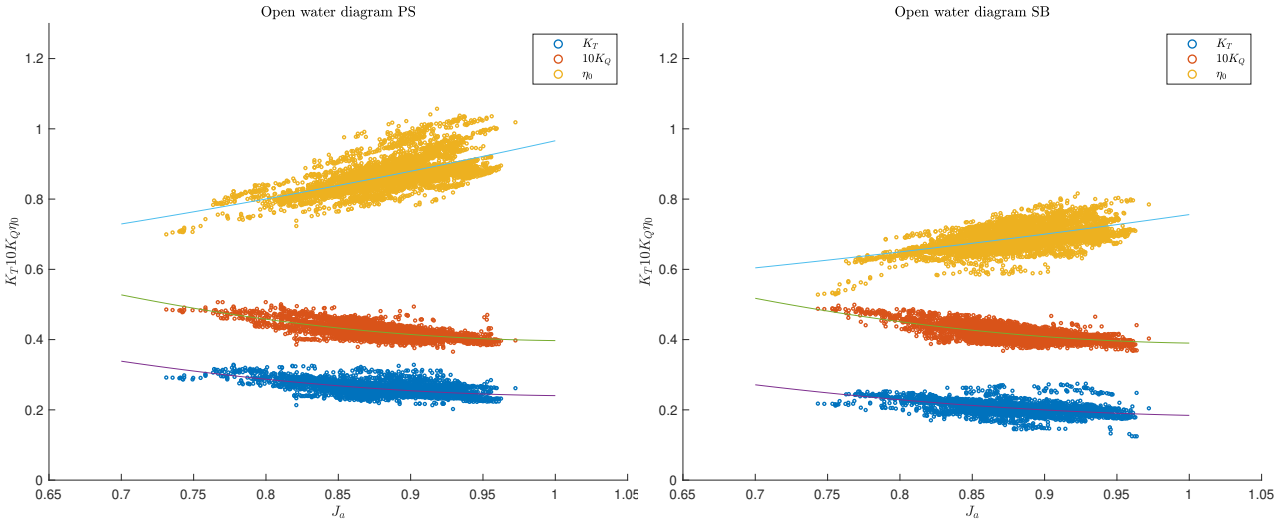
As seen in section 1.2.3, an open water diagram is a result of combination of multiple variables related to the propeller, i.e. thrust, torque, rate of revolutions and speed through water.

Among these variables the preliminary analysis of thrust revealed some irregularities in terms of whether the port or starboard had a bias offset. Based on the thrust measurements itself, it was not concluded which measurement were more likely to represent the most accurate thrust. However, since the torque measurements were found to be reliable, the same relations as for the open water diagram allows a more thorough study of the measured characteristics.

Figure 6.15 shows the open water diagram for respectively the port and starboard propulsion system based on the operational data points. Unlike the experimentally established open water diagram in figure 1.2, the operational diagram shows a significantly narrower range of advance ratio along the horizontal axis, i.e. the ratio between speed through water and rate of revolutions do not change significantly.

There is an noticeable difference between figure 6.15a and figure 6.15b and that is the magnitude of the efficiency $\eta_0$. The underlying factor that cause this difference is the thrust, which has been discussed previously. Observe how the efficiency in figure 6.15a are in range of 0.8 and seems to be greater than 1 at certain high advance ratios. Compared to Wageningen B-series open water diagram (figure 1.2) where the maximum efficiency is 0.8, the starboard side operational open water diagram (6.15b) is more similar. As a result, the starboard side

thrust measurements are considered more reliable, and the port side measurements should be corrected after the result of error analysis in section 6.2.3.




(a) Port side (b) Starboard side

Figure 6.15: Open water diagram from operational data

The vertical spread in thrust and torque coefficients in figure 6.15 is of interest, as a change in these coefficients at the same advance ratio should indicate that the thrust or torque varies independently of the rate of revolutions. For a controllable pitch propeller this would be natural, but that is not the case for these vessels. However, as illustrated on figure 6.16 the vertical spread is a result of a minor part of the measurements; the mean and standard deviation are $\mu_{K_T} = 0.203, \sigma_{K_T} = 0.013$. Thus, it is more likely that the spread is caused by some noice and uncertainness in the measurements.

### 6.3.2 Specific Fuel Consumption

Since the engine power was found to be reliable, the credibility of the fuel consumption measurements could be verified by the use of a specific fuel consumption curve. Plotting these curves from the operational data and comparing with reference curves from the literature (Mollenhauer & Tschoke 2010) allows such a verification.

Figure 6.17 show the SFC curves for each of the engines. On the vertical axis the is ratio between mass flow of fuel and produced power ($\dot{m}_f/P$ ) and horizontal axis if the percentage of maximum engine load[3]. The majority of sampling points on DG1, DG3 and DG4 follows

[3]DG1 and DG4 are of equal size while DG2 and DG3 are equal.

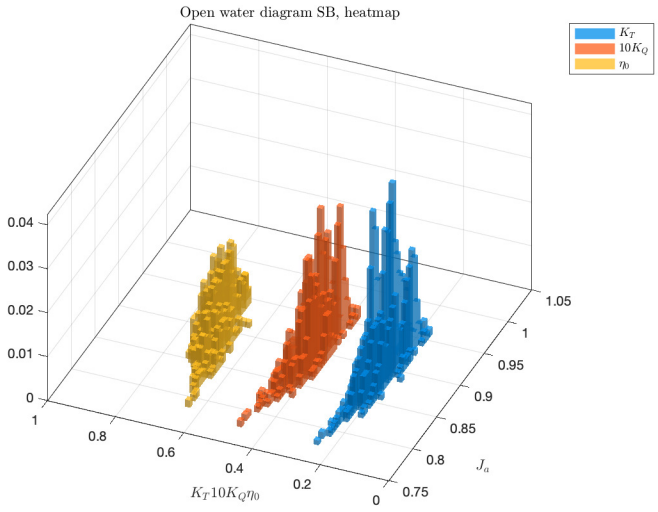



Figure 6.16: Open water diagram histogram for starboard side

what seems to be a trend line similar to what is found in the literature, with some scatter, while DG2 on the other hand show less convincing reliability. On $\eta = 0.8$ DG1 and DG4 shows a SFC = 220 kg/MWh and DG3 a SFC = 195 kg/MWh.

As a result, these SFC values on $\eta = 0.8$ represents the fuel consumption to a satisfactory degree and will be used for further analysis. Hence the fuel consumption for propulsive purposes can be found by evaluating the consumption based on motor power and varying hotel and auxiliary process consumption can be neglected.

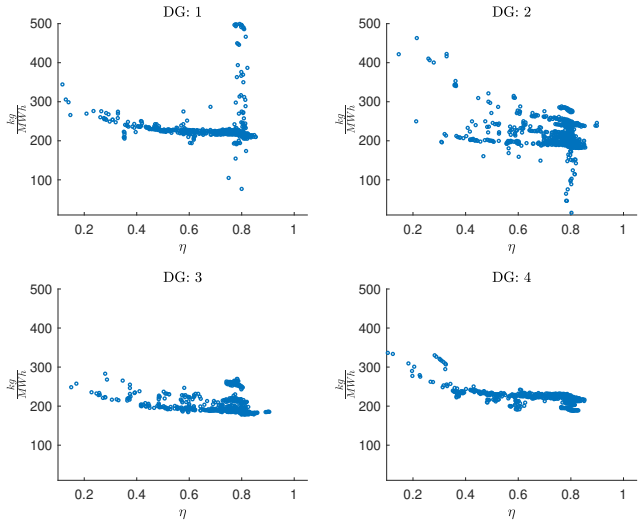



Figure 6.17: Specific fuel consumption curves for each of the four diesel generators

## 6.4 Summary of Data Analysis

The qualitative study of the data shows promising results and form a solid basis for further data mining and applications.

Using a running median filter demonstrated powerful properties during cleaning and smoothing of time series. Measurements of the speed through water were concluded reliable after an extensive data acquisition including both Norkyst800 and Tidetech. From a hydrodynamic point of view reliable speed through water measurements are valuable as it represents the operational environment more precise than speed over ground measurements.

Wind measurements from the vessel were found to be unreliable. Apparently the vessel never experienced head wind in the range of ±5 degrees which seemed odd. After comparison with simulated wind data from ECMWF and Tidetech it was concluded that the simulated data should be used to represent the wind instead of the measured parameters.

A bias error was found in the thrust measurements. A correction as a function of the RPM was suggested such that the port and starboard measurements were consistent. By establishing the open water diagram from the operational data the starboard side was concluded the most reliable such that for further use the port side should be corrected, alternatively the total thrust assumed to be two times the starboard measurement.

Torque, motor and engine measurements were all found to be reliable, both on port and starboard side. Hotel and auxiliary power consume were uncovered by evaluation of the difference between motor and engine power, where an environmental temperature dependency was found.

Fuel consume measurements showed a tendency to be fluctuating and less reliable. After a study of the operational SFC curves it was concluded that a constant should be used for further analysis, which allows a study of fuel consumption for propulsive purposes only.

# Chapter 7

# Resistance and Fuel Prediction

## 7.1 Frictional Resistance

As discussed and assumed in section 1.1.2 the frictional resistance is a function of the Reynolds number, meaning the speed through water, length between the perpendiculars and the kinematic viscosity. In addition, the roughness on the hull paint add contribute to some added resistance, which were discussed in section 1.1.3. This also implies that the frictional resistance is angular independent, such that it contribute with a resistance component against the heading.

From the ITTC friction line, equation (1.1.3), the frictional coefficient decrease with increasing Reynolds number and is illustrated in figure 7.1. With increasing Reynolds number the effect of roughness occur and is illustrated by the orange line in the same figure. This is in the same range at for the vessels, meaning that the frictional resistance is fairly constant at $C_f + \Delta C_f = 1.6 \cdot 10^{-3}$.

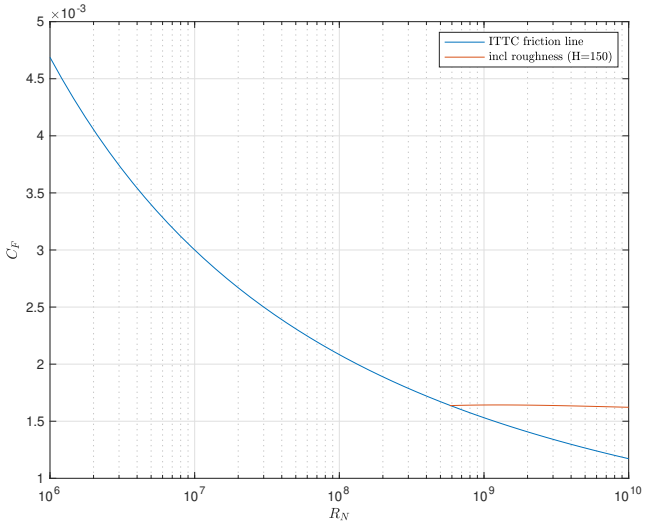



Figure 7.1: ITTC'57 friction line

## 7.2 Air and Wind Resistance

As the onboard wind measurements were found to be unreliable, it was concluded that the wind properties must be described by either the ECMWF or Tidetech data set.

After conversion from true to relative reference frame the resistance from caused by air and wind is presented in figure 7.2. It is assumed that the resistance is symmetric along the center line, such that the figure illustrate the resistance in the range from 0 to 180 degrees. The radial position represents the resistance in newton, while the color index represents the relative wind velocity. Further, the figure indicate that the effect of wind decrease significantly for relative angles larger than 30 degrees.

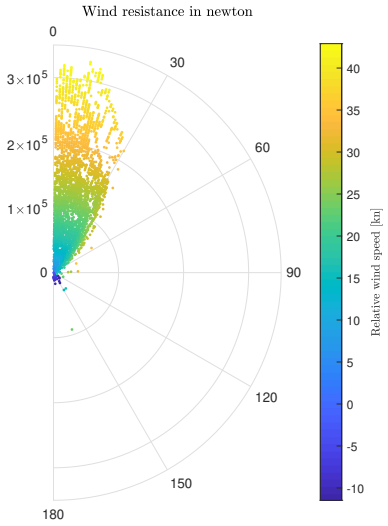



Figure 7.2: Wind resistance as a function of relative wind

## 7.3 Residual Resistance

With an estimate of the frictional resistance and the wind resistance, residual resistance represent a major contributor to the total resistance that should be quantified. However, as indicated in chapter 1, there are no generic accurate method to predict the residual part other than isolating the remaining resistance from the total when frictional and wind resistance are subtracted.

Chapter 5 described five different approaches on how to quantify the residual resistance. The first three approaches assume that the residual resistance is best quantified by isolating resistance added by the wave system generated by the vessel and the added resistance due to incoming waves.

### 7.3.1 Method 1

Figure 7.3 shows the resulting residual resistance coefficient as a function of the significant wave height $H_s$ and the mean wave period $T_p$. The surface in the same figure is a third order polynomial best fitting surface from the measured coefficients. The red line at $C_r(H_s = 0, T_p) = 2.0 \cdot 10^{-3}$ is a mean value established from visual interpretation where the surface intersect with $(H_s = 0, T_p)$. Thus, $C_{r0} = 2.0 \cdot 10^{-3}$ represents the resistance due to waves initiated by the vessel, or the wave resistance in calm water.

In theory, the calm water wave resistance should converge to the same value independently of whether $H_s \to 0$ or $T_p \to 0$ since both represents calm water. However, from the fitted

surface in figure 7.3 this is not the situation where intersection $C_r(H_s, T_p = 0)$ seems to vary a lot. The surface do not reflect the variation of $C_r$, but is an unfortunate result of fitting the surface to $\min(T_p) \approx 3s$ where extrapolation not is suitable.

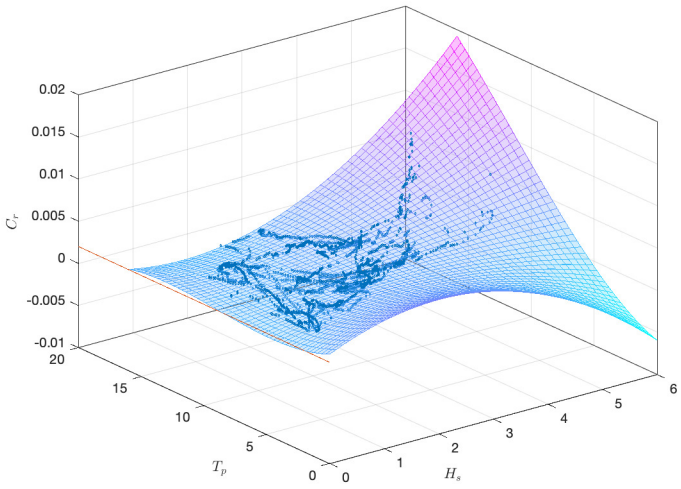



Figure 7.3: Regression of residual resistance as a function of $H_s$ and $T_p$

With the calm water established at $C_{r0} = 2.0 \cdot 10^{-3}$, method 1 assume that the added resistance due to waves is calculated accordingly to the ISO15016 annex D standard (STAWAVE1). As described in table 5.3 the procedure is limited by relative wave angle and significant wave height. Thus, the test cases are limited by these criteria.

Figure 7.4 shows how method 1 performed in calculating the total resistance relative to the true thrust. By the color index there is clear that the effect of waves are present, as the predicted and true resistance increase with increasing wave height. However, the vertical spread (difference between the predicted and true resistance) is significant and varies randomly.

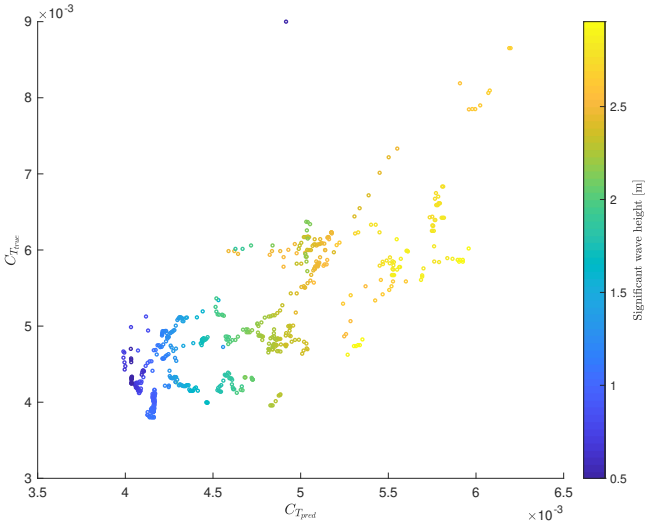



Figure 7.4: Testing results from method 1

### 7.3.2 Method 2

Method 2 is an extend of method 1, where a machine learning model is trained to predict the added resistance due to waves beyond the limitations of method 1. The input features to the training model are the significant wave height, return wave period, relative wave angle and speed through water, whereas the target is the total measured resistance corrected for friction and wind resistance. By some trail and error the network that performed best had the properties as given in table 7.1. The results of the trail and error are shown in figure 7.5 where the mean squared error between predicted and true error is plotted as a function of the total number of neurons in the network.

Table 7.1: Properties of neural network, method 2

| Property | Value/Format |
|---|---|
| # input layers | 1 |
| # input neurons | 4 |
| # hidden layers | 2 |
| # hidden neurons | 31 |
| Kernel function | Sigmoid |
| Training algorithm | Levenberg-Marquardt[1] |
| # rows in training set | 606 (70%) |
| # rows in validation set | 130 (15%) |
| # rows in test set | 130 (15%) |

[1]The algorithm use a early stop method to prevent overfitting

The training of the neural network is quantified by how the MSE develop as a function of total number of neurons in the network in figure 7.5. When exceeding 20 neurons, additional neurons do not seem to reduce the error as it has converges to a MSE $\approx 0.05 \cdot 10^{-7}$. Figure 7.6 shows how the MSE develops as a function of total number of neurons in the network when predicting with the testing data set. The error increase for all networks with more than 10 neurons.

The available training data is given in appendix B, which consist of a limited number of observations ($\approx 2000$). Another aspect of the training set is the variation among the samples i.e. the sample space is limited. This could be unfortunate if the model would have to extrapolate beyond the training sample space.

As a result, method 1 and 2 do not show convincing prediction precision, but both STAWAVE1 and neural networks shows potential in terms of detecting the added resistance do to waves, but are not sufficiently accurate. The performance parameters for method 2 are found in table 7.2

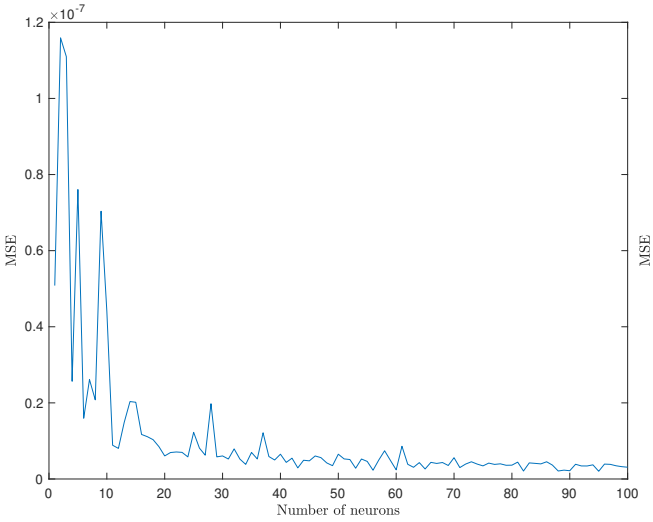



Figure 7.5: MSE as a function of neurons in network for training set, method 2

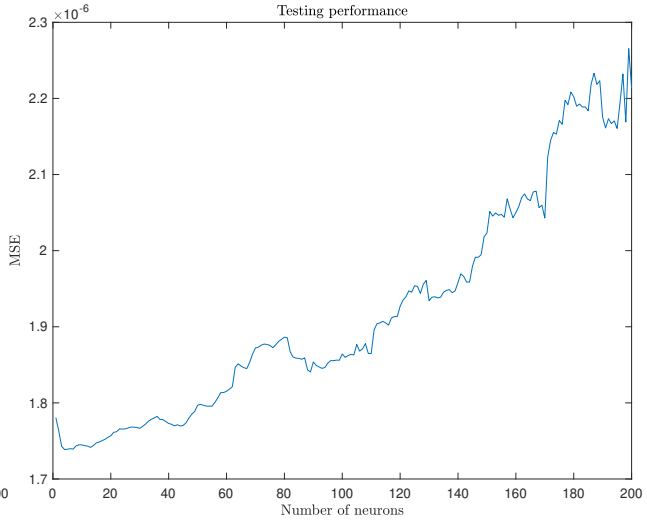



Figure 7.6: MSE as a function of neurons in network for testing set, method 2

### 7.3.3 Method 3

From the findings in the previous attempts on estimating the residual resistance, the use of neural networks will exploited further. By use of an alternative training algorithms, bayesian learning algorithm (Burden & Winkler 2009, MacKay 1992, Foresee & Hagan n.d.), the prediction could improve. In addition, a new prediction model is introduced, namely

Table 7.2: Performance coefficients, method 2

| | **Neural network** Levenberg-Marquardt |
|---|---|
| MSE ($10^{-6}$) | 1.72 |
| MAPE (%) | 15.15 |
| NMSE | 7.48 |
| RMS ($10^{-3}$) | 1.3 |
| REP (%) | 31.21 |

a Gaussian process regression which instead of fitting a weight and bias to a neuron fits a probability distribution to relate the features to the target. This means that the three machine learning models used in method 3 are as given in table 7.3

Table 7.3: Properties of machine learning algorithms, method 3

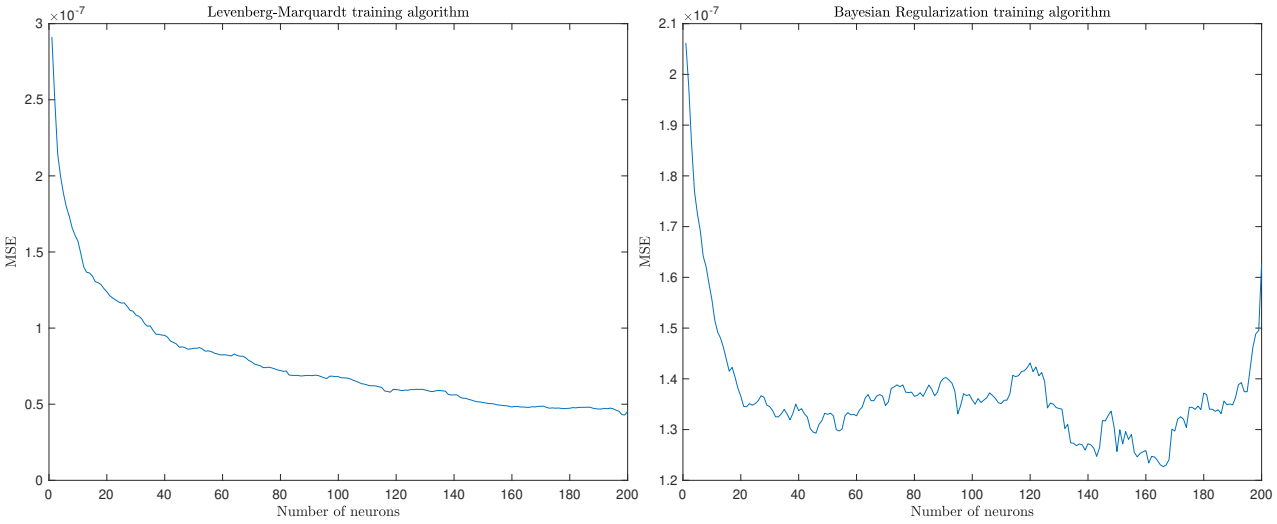
| | **Neural network** | **Neural network** | **Gaussian process regression** |
|---|---|---|---|
| Training algorithm/kernel function | Levenberg-Marquardt | Bayesian regularization | Exponential |
| Training/Validation/Test ratio [%] | 70/15/15 | 70/15/15 | 70/15/15 |

Development of the neural networks are the same as in method 2; the total number of neurons are increased until the MSE on the training set have converged. Figure 7.7a shows how the precision increase by the increasing number of neurons in the network for the Levenberg-Marquardt training algorithm, and it seems like there is no significant increased precision when exceeding 140 neurons. Figure 7.7b on the other hand shows another behavior for increasing number of neurons; more than 40 neurons to not influence the MSE significantly. Despite a more stabile development of MSE in for the Bayesian regularization algorithm the MSE value have converge to a larger value then the Levenberg-Marquardt algorithm.

### Testing method 3

As seen in chapter 2, measuring the goodness of a regression model only on the mean squared error value could be dangerous if a better fit is caused by overfitting. Even though figure 7.7a indicate that more neurons are better the models should be tested with the testing data set.

Figure 7.8 shows a clear tendency of overfitting for the LM algorithm and substantiates the assumption that 200 neurons in a network with four input features and one output is too complex. The figure shows that approximately 40 neurons are suitable - independently of training algorithm - for the purpose of describing the variation in residual resistance. More neurons increase the mean squared error.




(a) Levenberg-Marquardt training algorithm (b) Bayesian regularization training algorithm

Figure 7.7: MSE as a function of neurons in network for training set, method 3

Table 7.4 summarize the performance parameters from method 3 where the bold numbers indicate the best observation per parameter. The neural network with a Levenberg-Marquardt training algorithm perform better at most points, while the difference between GPR and Bayesian regularization algorithm is neglishable, but with an edge to the neural network.

Table 7.4: Performance coefficients, method 3

| | **Neural network** Levenberg-Marquardt | **Neural network** Bayesian regularization | **Gaussian process regression** Exponential |
|---|---|---|---|
| MSE ($10^{-7}$) | **5.62** | 8.65 | 9.00 |
| MAPE (%) | 11.28 | 10.99 | **10.57** |
| NMSE | **0.71** | 2.66 | 2.66 |
| RMS ($10^{-4}$) | **7.49** | 9.30 | 9.48 |
| REP (%) | **16.19** | 20.69 | 21.09 |

### 7.3.4 Method 4

Instead of separating the residual resistance into calm water wave resistance and added resistance due to waves, method 4 treat the residual resistance as one. For practical purposes, this means that $C_{r0}$ from method 1 is neglected and data-driven methods are left to describe the residual resistance by $C_r = f(H_s, T_p, \beta_{rw}, \text{stw})$.

The change in the initial assumptions means that the models from method 3 not are applicable

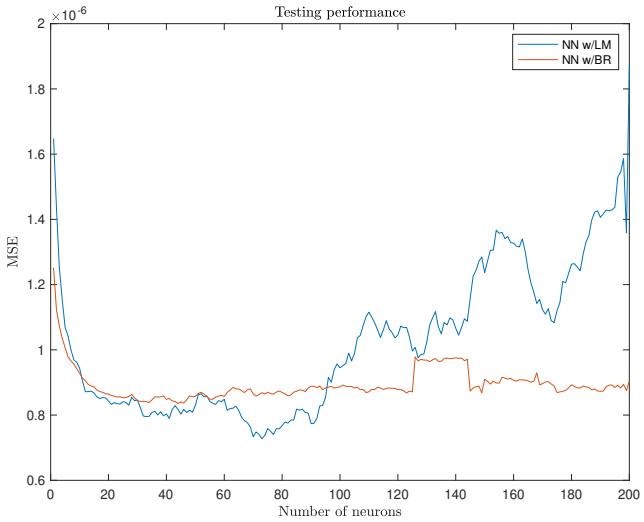



Figure 7.8: MSE as a function of neurons in network for testing set, method 3

and they must be retrained. The mean squared errors as a function of the total number of neurons in the network are plotted in figure 7.9. The shapes of the plots are similar to what in seen in figure 7.7 (MSE for method 3) and the accuracy increase with increasing number of neurons for LM algorithm and is more stabile when using BR.

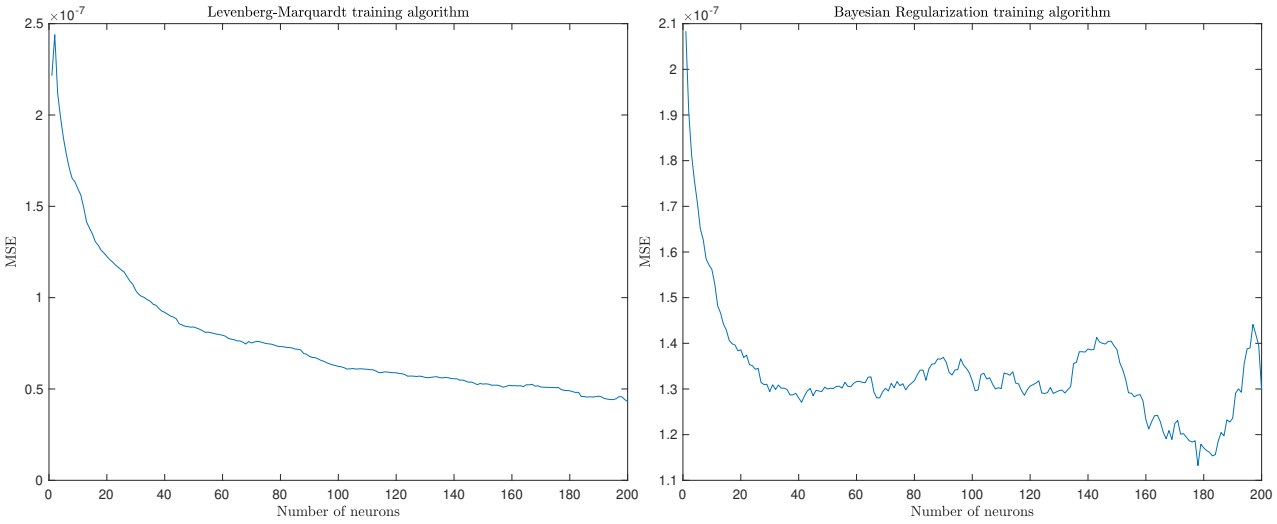
### Testing method 4

Training networks and evaluating their performance and accuracy by MSE have been shown unfortunate and testing is crucial. Figure 7.10 illustrates the same behavior as seen in method 3; overfitting occur when the network exceed approximately 60 neurons. At 40 neurons the testing set perform at its best, MSE $\approx 0.8 \cdot 10^{-6}$, which is more or less the same as in method 3.

Table 7.5 summarize the performance parameters from method 4 where the bold numbers indicate the best observation per parameter. The Levenberg-Marquardt learning algorithm had the most accurate prediction at the measured parameters, with a small margin. The Gaussian process regression model proves to by less applicable for this particular purpose[2].

Based on the performance parameters in method 3 and 4, there are for all practical purposes no difference between the two. Anyhow, neural networks proved to outperform GPR in this particular study.

[2]Training a model depends on initial seed numbers, such that these results and thereby conclusion could change if training was repeated (Lera & Pinzolas 2002). Hence a series of training sessions with different seed numbers are recommended.




(a) Levenberg-Marquardt training algorithm (b) Bayesian regularization training algorithm

Figure 7.9: MSE as a function of neurons in network for training set, method 4

Table 7.5: Performance coefficients, method 4

| | **Neural network** Levenberg-Marquardt | **Neural network** Bayesian regularization | **Gaussian process regression** Exponential |
|---|---|---|---|
| MSE ($10^{-7}$) | **6.27** | 7.47 | 9.06 |
| MAPE (%) | 12.65 | **10.44** | 10.62 |
| NMSE | **0.45** | 1.86 | 2.75 |
| RMS ($10^{-4}$) | **7.91** | 8.64 | 9.51 |
| REP (%) | **17.02** | 19.03 | 21.19 |

### 7.3.5 Method 5

Promising results by the machine learning models in the previous methods substantiates the assumption that such models are applicable for the purpose of prediction resistance. Even though good regression models are established, method 5 assume that the error in predictions could be corrected for by reducing the number of input variables and use a second model to reduce the prediction error.

Table 7.6 summarize the performance coefficients from method 5. The same models from method 4 were utilized at first and then a new set of model were trained with $H_s, T_p, \beta_{rw}$ as features and the prediction error ($C_T^{\text{pred}} - C_T^{\text{true}}$) as target. The table shows decreased performance compared to method 4 and table 7.5, meaning that the new models could not find a systematic pattern among the features that resulted in a predictive error.

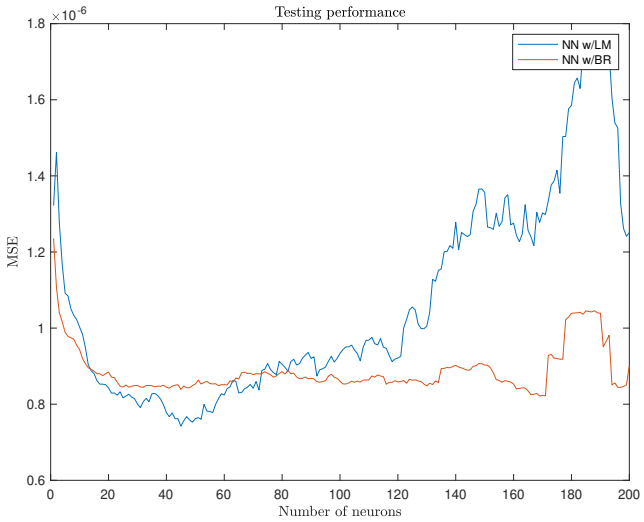



Figure 7.10: MSE as a function of neurons in network for testing set, method 4

Table 7.6: Performance coefficients, method 5

| $f(H_s, T_p, \beta_{rw}, \text{stw})$ | **Neural network** Levenberg-Marquardt | **Gaussian process regression** Exponential |
|---|---|---|
| $f(H_s, T_p, \beta_{rw})$ | **Gaussian process regression** Exponential | **Neural network** Levenberg-Marquardt |
| MSE ($10^{-7}$) | **8.95** | 9.39 |
| MAPE (%) | 11.64 | **10.93** |
| NMSE | **1.77** | 2.72 |
| RMS ($10^{-4}$) | **9.46** | 9.69 |
| REP (%) | **20.67** | 21.66 |

## 7.4 Bootstrapping of Prediction Model

Based on the performance coefficients it was method 4 that predicted the most accurate resistance. As an extend to validate the model, bootstrapping allows a study of whether the model responds as expected to a change in the inputs.

Figure 7.11 shows how the predicted $C_r$ responds to varying wave height when the wave period, speed through water and relative wave angle is fixed. On the left vertical axis and in blue is the predicted residual coefficient and on the right vertical axis and in orange is the wave height. From the figure it is clear that the wave resistance is closely correlated to the wave as the resistance increase with increasing wave height. Observe how the wave resistance decrease at the sample points around 700 when the waves are higher than 4 meters. This

could be a result of limited number of occurrences of this particular sea state in the training set and the model i is extrapolating. As a result it is recommended to be aware of inaccurate predictions in height head waves.


Figure 7.11: Bootstrapping with fixed $T_p = 1s$, STW $= 15$kn, $\beta_{rw} = 10°$ and varying $H_s$

Similar, in figure 7.12 where the relative wave angle are plotted together with the residual coefficient the response is as expected. Head waves results in a higher wave resistance whereas following waves reduce the resistance.

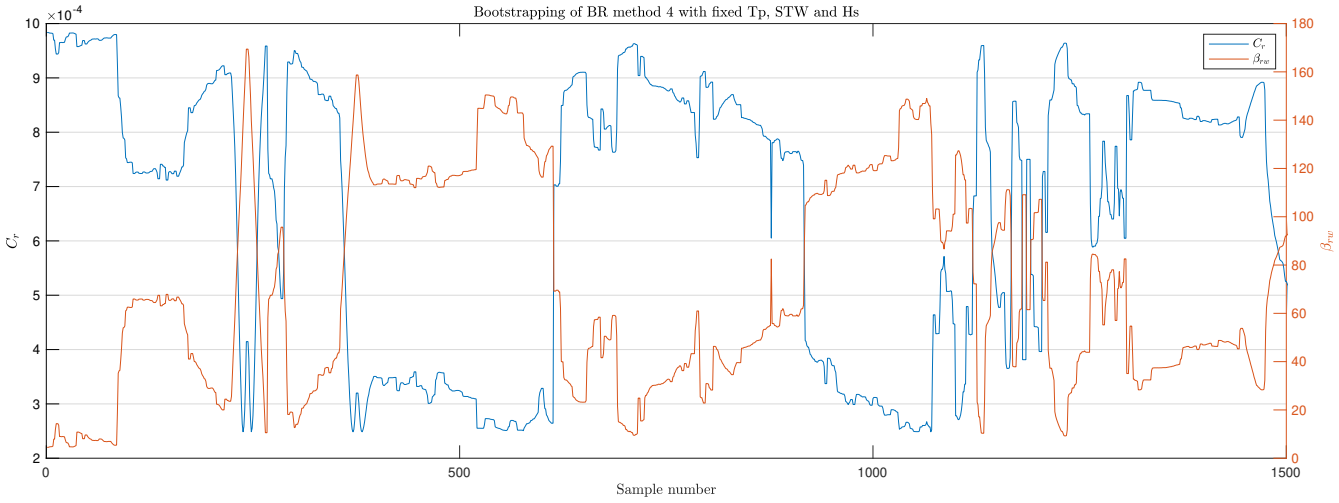



Figure 7.12: Bootstrapping with fixed $T_p = 1s$, STW $= 15$kn, $H_s = 1m$ and varying $\beta_{rw}$

As a result, it is reasonable to conclude that the neural network replicates the physical behavior of the vessel in waves. However, the magnitude of the residual resistance predictions

are neither confirmed nor denied from this analysis, only the trends relative to the free variable should be interpreted.

## 7.5 Thrust to Fuel Consume

The data analysis showed an almost perfect correlation between torque and motor power (equation (5.4.1) and figure 6.13). Thus, relating thrust to torque by use of open water diagrams and then find the coherent power are equivalent to relate the thrust directly to motor power via a simple regression. This will introduce some uncertainty into the prediction of fuel due to error propagation from thrust to power, and the performance coefficients on power prediction are expected to show a reduction on precision. Further, fuel for propulsive purposes is obtained from a SFC = 200 kg/MWh.

Table 7.7 summarize the performance parameters for motor power prediction after conversion from thrust to motor power. MAPE (mean absolute percentage error) and REP (relative percentage error) are comparable with the performance coefficients from method 4 and as expected the precision have decreased with 1 and 1.5 percentage points respectively.

Table 7.7: Performance parameters, power prediction [MW]

| **Method 4** | **Neural network** Levenberg-Marquardt |
|---|---|
| MSE | 3.82 |
| MAPE (%) | 13.91 |
| NMSE | 0.81 |
| RMS | 1.95 |
| REP (%) | 18.51 |

## 7.6 Method Summary and Discussion

The overall objective of this thesis has been to predict the fuel consumption of a ship in transit by including machine learning models in the predictions. The solution to this task included development of a data preprocessing framework for operational monitoring data which utilize external climate data for a better insight the data quality. These data were input to an approach where the total ship resistance were decomposed into frictional, wind and wave resistance. For better performance and faster learning rate for the machine learning models, frictional and wind resistance were isolated using well established empirical models,

whereas the wave resistance were to be determine by the models. Eventually, the predicted resistance were correlated to the required motor power and finally to fuel consumption.

Ensuring data integrity were an essential part for further mining and feature extraction. Filtering, smoothing, statistical analysis and visual interpretation facilitated efficient training of the data-driven models. The results presented in chapter 6 and 7 summarize the analysis of vessel B, which showed more promising data quality in a preliminary analysis.

Development of the framework for decomposing the resistance such that the machine learning models could be trained was the second part of the solution. Friction and wind resistance were isolated by empirical methods, whereas 5 methods were proposed for determination of the wave resistance. During training and testing of these methods, the concept and pitfall of overfitting occurred and illustrated the danger of blindly trust a machine learning algorithms performance. As for the framework for preprocessing, this framework are applicable for vessel A, C and D as well.

Among the different suggested approaches to predict the residual resistance method 3 and 4 were found to be more or less equally accurate for vessel B. However, when correlating the predicted resistance to power and fuel, it was shown in table 7.7 a RMS = 1.95 MW, which with a SFC = 200 kg/MWh results in a prediction error of 390 kg/h. In context, if the vessel on average consume 5 metric tons per hour in transit, the predictions are on average off by $\approx 8$ %.

In method 5 where the prediction errors where analyzed by two machine learning algorithms, no systematic pattern were discovered when the wave properties were input. A similar study was conducted with the wind attributes as well, without any systematic pattern and thereby no increase in performance. Hence, among the external environmental parameters, current is not directly involved in the prediction, but should be reflected in the speed through water measurements. Varying quality in these measurements is a possible source of errors. Another, and more likely explanation to the uncertain prediction is the limited sample space the training data set. Extrapolation beyond the training sample space could be unfortunate, especially if the network work is overfitted and extrapolated values diverge.

This hypothesis was substantiated when applying both frameworks to vessel A (results are shown in appendix C). Predictions using method 4 gave a RMS = 0.61 on motor power predictions, or an average error of $\approx 2.5$ % in fuel consumption per hour. These results indicate that the approach and methods used and developed in this thesis are promising for this particular purpose. It could be argued whether the results from vessel A are due to a limited testing sample space or more reliable thrust or speed through water measurements, but the prepossessing results did not indicate difference of significance compared to vessel B.

Applying the prediction model from vessel A on vessel B did not result in any increase in performance. Nor did merging the training set and retrain the models and make predictions

on based on a more extensive training sample space.

# Chapter 8

# Conclusion and Recommendations

## 8.1 Conclusion

The research goal in this thesis has been to apply state of the art machine learning to predict the fuel consumption of a ship in transit. An extensive literature study within naval hydrodynamics, machine learning and data cleaning were conducted in order to establish the theoretical basis for further application and analysis. These fields of studies were combined in the development and validation of a framework for preprocessing of monitoring data and a second framework for decomposition of the ship resistance, training and validation machine learning models.

Preprocessing of the data were considered as an essential part of the required work and were therefore conducted to an extent. Moving median filter showed itself powerful to eliminate outliers, to smooth the time series and at the same time preserve the gradients and sharp transitions in the time series of measurements. Irregularities were found in some attributes and replaced with coherent simulated data from ECMWF and Tidetech to ensure high quality and reliable data for further data mining and feature extraction.

Prediction of fuel consumption consisted of isolating wind and frictional resistance and estimate the residual resistance by use of machine learning models. Five different methods on how to predict the residual resistance were proposed with a varying degree of machine learning dependency. Among the ML models tested where two shallow neural networks with Levenberg-Marquardt and Bayesian regularization training algorithms, and a Gaussian process regression model with an exponential kernel function. During validation a text-book example of overfitting were reveled and illustrated the importance of a thorough understanding of the mechanisms involved when using these black-box models. Due to the extensive literature study within machine learning the overfitting problem were avoided.

The precision of the prediction were found to be marginal better when the models were trained without any interference from other routines (method 4). A mean deviation of 2.5% from the true consume were found for vessel A as the best prediction which substantiate that these models are applicable for the purpose of predicting ship resistance and fuel consumption. Further evaluation on how the model responds to changing inputs are promising. Higher waves causes an increase in residual resistance and following waves cause a reduction.

Thus, predicting fuel consumption of a ship in transit using machine learning have showed itself both possible and promising.

## 8.2 Recommendations for Further Work

- The thrust measurements showed itself to be less reliable than other measurements of energy transformation from fuel to thrust. By correlating e.g. torque to the resistance the scatter in which the machine learning models are training could be reduced. Chances of overfitting could potentially be reduced and more accurate predictions obtained.
- An important attribute that not is included in the analysis is the trim. The effect of trim should and could potentially be accounted for
- Optimization of operational profile. The diesel engines are not operating on maximum efficiency at all time (possibly out of the scope of hydrodynamics)
- This thesis assume more accurate prediction by isolating known and well established physics and let machine learning account for other aspects. A study of whether a machine learning model with all the same inputs would outperform this thesis approach is interesting (possibly out of the scope of hydrodynamics)

# Bibliography


Ådnanes, A. K. (2003), *Maritime electrical installations and diesel electric propulsion*, ABB.

Albretsen, J., Sperrevik, A., Staalstrøm, A., Sandvik, A., F., V. & Asplin, L. (2011), 'Norkyst-800 report no. 1: User manual and technical descriptions'.

Arce, G. R. (2005), *Nonlinear signal processing: a statistical approach*, John Wiley & Sons.

Bishop, C. M. (2013), *Pattern recognition and machine learning*.

Burden, F. & Winkler, D. (2009), *Bayesian Regularization of Neural Networks*.

Elattar, E. E., Goulermas, J. & Wu, Q. H. (2010), *Electric Load Forecasting Based on Locally Weighted Support Vector Regression*.

Faltinsen, O. (1993), *Sea loads on ships and offshore structures*, Cambridge university press.

Foresee, F. D. & Hagan, M. T. (n.d.), Gauss-newton approximation to bayesian learning, *in* 'Proceedings of International Conference on Neural Networks (ICNN'97)'.

Fossen, T. I. (2011), *Handbook of Marine Craft Hydrodynamics and Motion Control*, John Wiley & Sons.

Frank, E., Hall, M., Holmes, G., Kirkby, R., Pfahringer, B., Witten, I. H. & Trigg, L. (2009), *Data mining and knowledge discovery handbook*.

Gerritsma, J. & Beukelman, W. (1972), *Analysis of the resistance increase in waves of a fast cargo ship*.

Ghelardoni, L., Ghio, A. & Anguita, D. (2013), *Energy Load Forecasting Using Empirical Mode Decomposition and Support Vector Regression*.

Gu, G.-F. & Zhou, W.-X. (2010), *Detrending moving average algorithm for multifractals.*

Hansen, S. V., Petersen, J., Jensen, J. & Lützen, M. (2011), *Performance monitoring of ships*, DTU Mechanical Engineering.

Ioffe, S. & Szegedy, C. (2015), *Batch normalization: Accelerating deep network training by reducing internal covariate shift.*

ISO (2015), Ships and marine technology - guidelines for the assessment of speed and power performance by analysis of speed trail data, Standard, International Organization for Standardization.

Jolliffe, I. (2011), *Principal component analysis*, Springer.

Keaveny, T. M., Pinilla, T. P., Crawford, R. P., Kopperdahl, D. L. & Lou, A. (1997), *Systematic and random errors in compression testing of trabecular bone*, Wiley Online Library.

Kim, M., Hizir, O., Turan, O., Day, S. & Incecik, A. (2017), 'Estimation of added resistance and ship speed loss in a seaway'.

Kotsiantis, S., Kanellopoulos, D. & Pintelas, P. (2006), *Data preprocessing for supervised leaning*, Citeseer.

Lera, G. & Pinzolas, M. (2002), *Neighborhood based Levenberg-Marquardt algorithm for neural network training.*

MacKay, D. J. (1992), *Bayesian interpolation*, MIT Press.

Maruo, H. (1957), 'The excess resistance of a ship in rough seas', *International Shipbuilding Progress* .

MetOffice (2016), 'Beaufort wind force scale'.
**URL:** *https://www.metoffice.gov.uk/guide/weather/marine/beaufort-scale*

Mitchell, T. M. (1997), *Machine learning*, McGraw-Hill series in computer science, Artificial intelligence.

Mollenhauer, K. & Tschoke, H. (2010), *Handbook of diesel engines*, Vol. 1, Springer.

Nielsen, M. A. (2015), *Neural networks and deep learning.*

Pedersen, B., Andersen, P., Larsen, J. & Sinding, P. (2014), Data-driven Vessel Performance Monitoring, PhD thesis.

Pitas, I. & Venetsanopoulos, A. N. (2013), *Nonlinear digital filters: principles and applications*, Springer Science & Business Media.

Steen, S. & Minsaas, K. (2014), *TMR4220 Naval Hydrodynamics compendium.*

Wahl, J. M. (2018), Prediction of fuel consumption of a ship in transit using machine learning, Project thesis, Norwegian University of Science and Technology (NTNU).

White, F. M. (2011), *Fluid mechanics*, McGraw Hill.

Wold, S., Esbensen, K. & Geladi, P. (1987), *Principal component analysis.*

# Part IV

# Appendices

# Appendix A

# Reference Frames

The purpose of this chapter is to provide the reader with the sign and direction conventions for this thesis.

## A.1 Vessel Reference Frame

A vessel have the ability to move in six degrees of freedom (DOF) where three are in translation and three in rotation, figure A.1. Surge, sway and heave are refer to the translation in x, y and z-direction, while roll, pitch and yaw represents rotation around x, y and z-axis. The origin coincide with the midship at the waterline.

The heading of a vessel refers to the relative angle between north and the bow, whereas the course refers to the cardinal direction the vessel is to be steered. When there is no lateral drift in sway direction heading and course coincide, but ocean currents or wind could cause drift. Hence the course is to be distinguished from the heading.

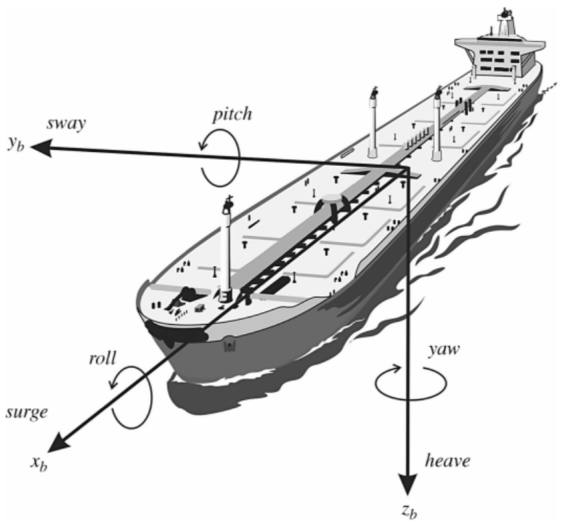



Figure A.1: Six degree of motion of a vessel (Fossen 2011)

## A.2 Geographic Reference Frame

When analyzing a vessel at sea, it is convenient to define a reference frame to describe a specific location on the Earth surface, as indicated in figure A.2. Latitude describes a specific north-south position ranging from 0° at the Equator to ±90° at the North and South Pole. Longitude specifies the east-west position ranging from 0° namely the Prime Meridian which passes through Greenwich, England west- and eastwards by ±180°. Longitude together with latitude is used to specify a precise location on the surface on the Earth.

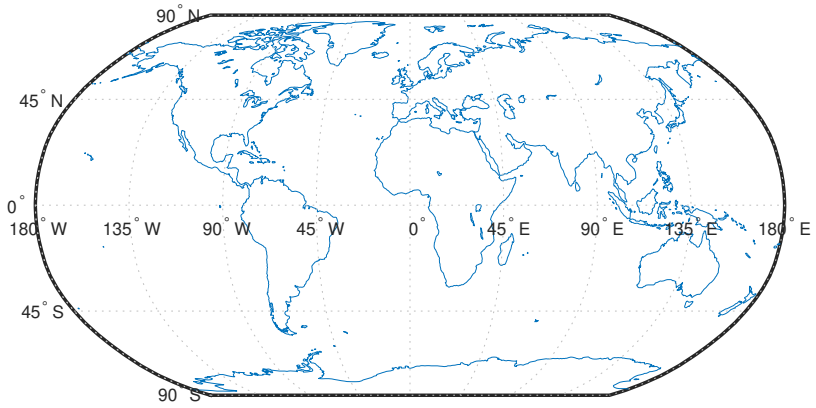



Figure A.2: Geographic reference frame

There exists other conventions and reference frames where latitude range from 0° at the South Pole to 180° at the North Pole, whereas longitude range from 0° to 360°.

Hence there are valid latitude and longitude values for both reference frames, but they would refer to different location on the surface of the Earth. When working with different systems, caution should be made to the describe a precise location by use of latitude and longitude coordinates.

## A.3 Wind Reference Frame

While wind in general is defined as the movement of air relative to the surface of the earth, a mathematical description of the wind is quantified with vectors with magnitude and direction. Figure A.3 refers to these vectors and how it is related to a vessel. Wind from west to east is described by an angle $\beta_w = 90$, while a vessel heading south is described by $\psi = 180$.

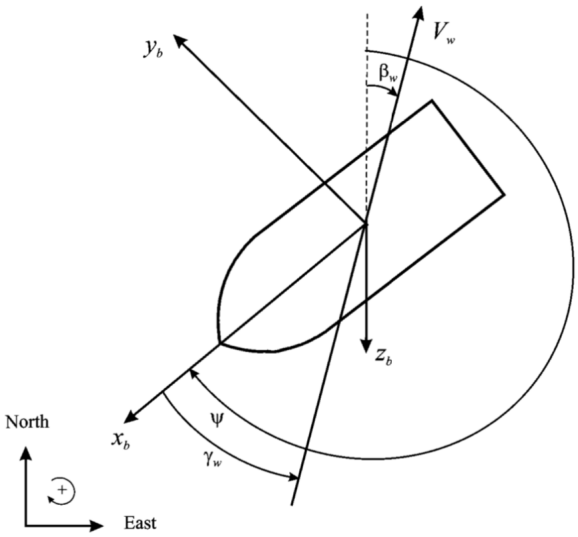



Figure A.3: Definition of wind angles relative to vessel, (Fossen 2011)

### A.3.1 Conversion from True to Relative Direction and Speed

For a vessel in motion exposed wind change as a function of the speed over ground and the heading. The relative wind velocity and angle are defined as

$$V_{rw} = \sqrt{u_{rw}^2 + v_{rw}^2} \tag{A.3.1}$$

$$\gamma_{rw} = -\text{atan2}(v_{rw}, u_{rw}) \tag{A.3.2}$$

where

$$u_{rw} = u - u_w, \quad u_w = V_w \cos(\beta_w - \psi) \tag{A.3.3}$$

$$v_{rw} = v - v_w, \quad v_w = V_w \sin(\beta_w - \psi) \tag{A.3.4}$$

and $u, v$ are the velocities of the vessel in surge and sway.

# Appendix B

# Training Data

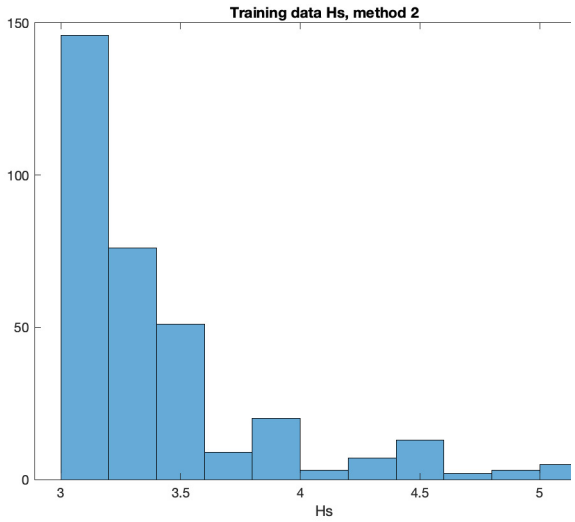



Figure B.1: Distribution of significant wave height

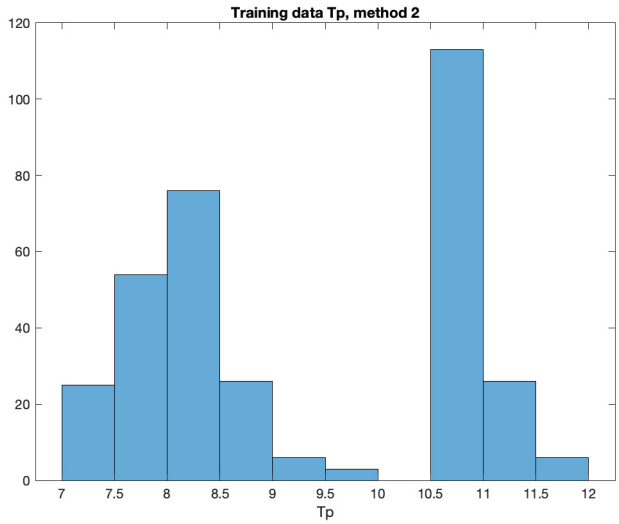



Figure B.2: Distribution of return wave period

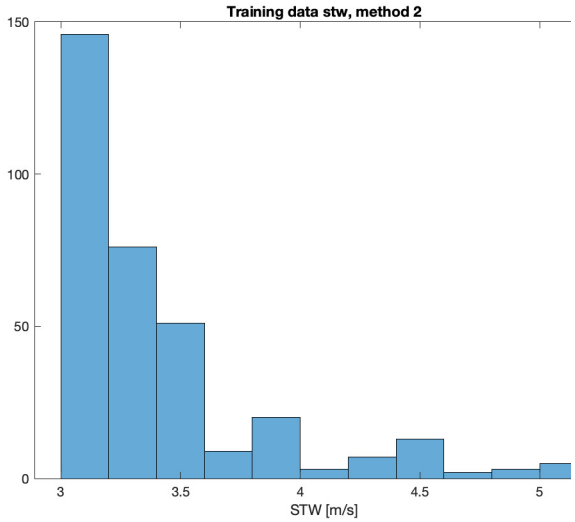



Figure B.3: Distribution of speed through water

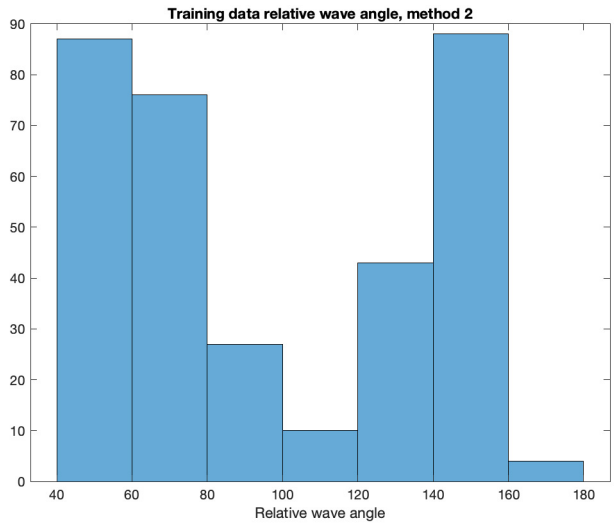



Figure B.4: Distribution of relative wave angle

# Appendix C

# Resistance and Fuel Prediction for Vessel A

Table C.1 summarize the performance parameters from method 4 with vessel A. The Bayesian regularization algorithm performed better in this particular test. Development of MSE as a function of neurons in network are shown in figure C.1 where a different behavior occur; as few neurons as possible seems to be the better. Note how the MSE for BR converge at a significantly smaller value ($< 3 \cdot 10^{-7}$) for vessel A whereas vessel B converged to ($> 8 \cdot 10^{-7}$).

Observe how the REP decreased in the transition from prediction of resistance to power. The regression between resistance and power appears to be more suitable for vessel A then for vessel B. Gaussian process regression with exponential kernel function did perform the most accurate prediction of fuel consumption of a ship in transit.

Table C.1: Performance coefficients on resistance predictions, method 4 for vessel A

| | **Neural network** Levenberg-Marquardt | **Neural network** Bayesian regularization | **Gaussian process regression** Exponential |
|---|---|---|---|
| MSE ($10^{-7}$) | 3.24 | **3.08** | 3.12 |
| MAPE (%) | 11.99 | **11.47** | 11.85 |
| NMSE | **1.33** | 1.86 | 1.43 |
| RMS ($10^{-4}$) | 5.69 | **5.55** | 5.58 |
| REP (%) | 15.25 | **15.01** | 15.15 |

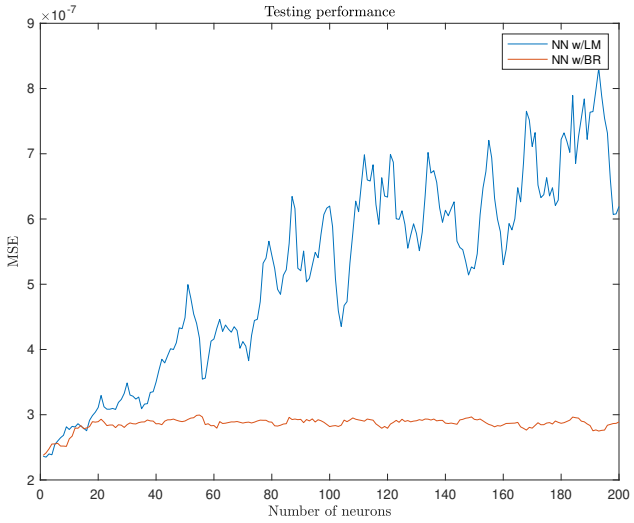



Figure C.1: MSE as a function of neurons in network for testing set, method 4 vessel A

Table C.2: Performance parameters on power prediction [MW], method 4 for vessel A

| | **Neural network** Levenberg-Marquardt | **Neural network** Bayesian regularization | **Gaussian process regression** Exponential |
|---|---|---|---|
| MSE | 1.63 | 0.42 | **0.37** |
| MAPE (%) | 19.57 | **11.36** | 11.53 |
| NMSE | 0.12 | **0.03** | **0.03** |
| RMS | 1.27 | 0.64 | **0.61** |
| REP (%) | 17.77 | 9.45 | **8.85** |