Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Engineering
Department of Marine Technology

Simon Krogstad Munkeby

On application of digital twin in ship operation and performance

Master's thesis in Marine Technology
Supervisor: Amir R. Nejad
Co-supervisor: Etienne Purcell
June 2022

NTNU
Norwegian University of Science and Technology

Simon Krogstad Munkeby

On application of digital twin in ship operation and performance

Master’s thesis in Marine Technology
Supervisor: Amir R. Nejad
Co-supervisor: Etienne Purcell
June 2022

Norwegian University of Science and Technology
Faculty of Engineering
Department of Marine Technology

NTNU
Norwegian University of
Science and Technology

Abstract

Shipping is a growing industry and is today responsible for over 90% of world trade[65]. Tight schedules while in port and long travelling distances are common aspects of transportation, and hence efficiency is very important to maintain maximum profit. One of the most important factors to maintain profitability is to reduce operational downtime. Methods for predicting these events can be developed to prevent unwanted repairs and shutdowns. Condition monitoring is a strategy which involves monitoring the condition of physical components located on e.g. a vessel. If irregularities in vibration data or temperature are discovered, condition-based maintenance can be performed on the relevant components *before* the failure happens. This reduces the maintenance cost as well as increases operational up-time. To predict future failures, three different modelling methods are typically used. Data-driven modelling uses datasets and machine learning technology to develop a model by either using classification or regression algorithms. Physics-based modelling is a mathematical representation of an asset which uses common physical characteristics to simulate its behaviour. An approach combining these two condition-based methods is called hybrid modelling. A model gathering real-time data for monitoring the condition of a physical asset is called a Digital Twin. A Digital Twin is a digital representation of an object. This technology is gradually getting more attention within several industries. In this thesis, a data-driven model is developed to predict fuel oil consumption per day for marine vessels, which directly correlates to the power output, a relevant parameter in the context of machinery conditions. Several regression models have been trained based on data from four of Gearbulk's vessels. Data from a fifth vessel has been used as test data to validate the accuracy of the model. A ranking of the ten best algorithms showed a low Root Mean Squared Error (RMSE) for both the training and test dataset. The results showed a clear correlation between fuel oil consumption per day and the predictor parameters. The accuracy was higher on the test dataset which indicates a potential overfitted model. This can however be solved by adjusting the data handling and training method. The use of condition monitoring systems and digital twins are increasing in popularity within the maritime industry, but the technology is not yet flawless. It is important to understand how the technology works before it is implemented to maximise its potential.

Sammendrag

Shipping er en voksende industri og står for over 90% av verdenshandelen i dag[65]. Stramme tidsskjemaer i havn og lange reiseavstander er typisk innen transportering, og derfor er effektivitet svært viktig for å opprettholde maksimal fortjeneste. En av de viktigste faktorene for å opprettholde lønnsomheten er å redusere operasjonell nedetid. For å forhindre uønskede reparasjoner og plutselig maskineristans, finnes det metoder for å forutse disse hendelsene. Tilstandsovervåking er en strategi som innebærer å overvåke tilstanden av fysiske komponenter plassert på et fartøy. Hvis uregelmessigheter i vibrasjonsdata eller temperatur oppdages, kan tilstandsbasert vedlikehold utføres på relevante komponenter før feilen skjer. Dette reduserer også vedlikeholdskostnadene og øker driftstiden. Tre forskjellige modelleringsmetoder blir ofte brukt for å forutse fremtidige feil. Datadrevet modellering bruker datasett og maskinlæringsteknologi for å utvikle en modell enten ved bruk av klassifiserings- eller regresjonsalgoritmer. Fysikkbasert modellering er en matematisk representasjon av en gjenstand som bruker typiske fysiske egenskaper for å simulere dens oppførsel. En kombinasjon av disse to tilstandsbaserte metodene kalles hybrid modellering. En modell som samler sanntidsdata for overvåking av tilstanden til en fysisk gjenstand kalles en Digital Tvilling. En Digital Tvilling er en digital representasjon av et fysisk objekt. I dag har denne teknologien har fått stadig mer oppmerksomhet innenfor flere bransjer. I denne oppgaven er det blitt utviklet en datadrevet modell for å forutsi fyringsoljeforbruk per dag for marine fartøyer, som direkte korrelerer med kraftuttaket, en relevant parameter i sammenheng med maskinens tilstand. Flere regresjonsmodeller har blitt trent basert på data fra fire av Gearbulk's fartøyer. Data fra et femte fartøy har blitt brukt som test-data for å validere nøyaktigheten til modellen. En rangering av de ti beste algoritmene viste en lav Root Mean Squared Error (RMSE) for både trening og test datasett. Resultatene indikerte en klar sammenheng mellom drivstofforbruk per dag og prediktor parametrene. Nøyaktigheten var høyere på testdatasettet, noe som indikerer en potensiell overtilpasset modell. Dette kan imidlertid løses ved å justere datahåndteringen og treningsmetoden. Tilstandsovervåking og digitale tvillinger har økt i popularitet innen den maritime industrien, men teknologien er enda ikke feilfri. Det er viktig å forstå hvordan teknologien fungerer før den implementeres for å maksimere potensialet.

Preface

This thesis was written as a part of my master's degree program at the Institute of Marine Technology at the Norwegian University of Science and Technology. The idea behind the theme of the thesis was developed in the autumn of 2021 in the course TMR4560 - Marine Systems Design, Specialisation Project. This course worked as a preparation for the master thesis.

A significant part of my thesis was the analysis of vessel data for the development of my model. A surprisingly difficult part was actually to get real data. The first part of the semester involved contacting companies and asking for data. This resulted in a delay in my thesis as I could not start analysing the data at an early stage.

I did not know a whole lot about digital twins before writing the thesis, but the technology behind really made me interested and I wanted to learn more. Throughout the thesis, I have learned much about modelling including machine learning methods. This is information I will take with me further on and which might become useful someday.

Simon Krogstad Munkeby,
Trondheim, 29th June 2022

Acknowledgements

There are several people I want to thank for helping me during the writing of this thesis. First of all, I want to thank my supervisor prof. Amir Nejad for great supervision and informative conversations throughout the semester. I will also thank PhD Candidate Etienne Purcell for helping me to understand the application ”Regression Learner” and machine learning in general. The collaboration with Torkel Soma from Sayfr AS helped me get in touch with various companies within the marine industry to prepare for interviews. I would like to thank Bjørn Egil Skogvoll and Joachim Grindheim from Gearbulk for providing data from their vessels. Lastly, I want to thank Sharat Valluri from ABS and Jarle Fosen and Bjarne Augestad from Gard for contributing to the interview process.

Table of Contents

List of Figures **viii**

List of Tables **xi**

1 Introduction **1**

1.1 Marine industry 1

1.2 State-Of-The-Art - Maintenance Strategies 1

1.3 Condition monitoring in the maritime industry 4

1.4 Condition based methods 5

1.5 Concept of a digital twin 7

1.5.1 Digital twins in maritime applications 8

1.6 Previous incidents 10

1.6.1 Viking Sky 10

1.6.2 MS Kong Harald 11

1.7 Research Motivation 12

1.8 Thesis objective 13

1.9 Thesis outline 13

2 Methodology **15**

2.1 Machine learning 15

2.1.1 Unsupervised machine learning 16

2.1.2 Supervised machine learning 17

2.1.3 Importance of model fitting 17

2.1.4 K-fold Cross Validation 19

2.1.5 Root Mean Square Error 20

2.2 Machine learning algorithms 21
2.2.1 Linear Regression 21
2.2.2 Gaussian Process Regression 22
2.2.3 Support Vector Machine 23
2.2.4 Artificial Neural Networks 26
2.3 Case Study 27
2.3.1 Machinery 28
2.3.2 Geographical areas 29
2.3.3 Weather conditions 30
2.3.4 Modelling approach 30

3 Data collection **33**

3.1 Parameters 33
3.1.1 Sailing time in days and miles 34
3.1.2 Fuel oil consumption per day 34
3.1.3 Miles per metric ton fuel 34
3.1.4 RPM 34
3.1.5 Average draft and trim 34
3.1.6 Wind and sea force 35
3.1.7 Wind and sea direction 35
3.2 Data preparation 36

4 Results **37**

4.1 Training results 37
4.2 Model 38

5 Discussion **42**

5.1 Industry view on digital twins 44

6 Conclusion **46**

6.1 Contribution and further work 47

Appendix **I**

A Engine data of case study vessels I

B Plot of significant wave heights from 1979-2014 IV

C Training results of 2nd to 10th best model V

List of Figures

1 Different failure patterns[1] 2

2 Development in maintenance strategies[43] 5

3 Illustration of different model approaches[6] 6

4 Main steps of digital twin development[36] 7

5 Fictional illustration of a digital twin[56] 8

6 Digital twin of a wind turbine[81] 10

7 Viking Sky in rough weather conditions[31] 11

8 MS Kong Harald after anchoring[39] 12

9 K-means algorithm[59] 16

10 Example of supervised learning with geometrical figures[37] 17

11 Examples of different fitted regression models 18

12 Figure of underfitted, fitted and an overfitted classification model[28] . . . 19

13 The procedure of K-fold Cross Validation[71] 20

14 Figure of simple linear regression[54] 22

15 Illustration of a linear SVC and a non-linear SVR model[52] 24

16 Illustration showing high and low c and gamma parameter[67][51] 25

17 Example of a shallow neural network[4] 26

18 Mathematical model of an artificial neuron[22] 27

19 Plover Arrow[77] 27

20 Heat map showing port activity by the vessels from January 2021 to May 2022 29

21 Number of port visits sorted by country between January 2021 and May 2022 30

22 Figure of the modelling process 32

23 Figure showing a vessel with a trim[73] 35
24 Wind and sea directions 36
25 Predicted versus true fuel oil consumption per day, with Matern 5/2 GPR Model 39
26 Validation residuals for true response 40
27 Predicted versus true fuel oil consumption per day, with Matern 5/2 GPR with test data 41
28 Validation residuals for true response, for the test set 41
29 Figure showing an example of boundaries for the model with boundary error of ~ 14 42
30 Avocet Arrow engine data I
31 Corella Arrow engine data II
32 Plover Arrow engine data II
33 Swift Arrow engine data III
34 Weaver Arrow engine data III
35 Historical data of the average significant wave heights from year 1979-2014 [70] IV
36 Predicted versus actual data, Cubic SVM V
37 Training data of Cubic SVM V
38 Predicted versus actual data, Trilayered Neural Network VI
39 Training data of Trilayered Neural Network VI
40 Predicted versus actual data, Wide Neural Network VII
41 Training data of Wide Neural Network VII
42 Predicted versus actual data, Bilayered Neural Network VIII
43 Training data of Bilayered Neural Network VIII
44 Predicted versus actual data, Quadric SVM IX
45 Training data of Quadric SVM IX

46 Predicted versus actual data, Rational Quadric GPR X
47 Training data of Rational Quadric GPR X
48 Predicted versus actual data, Squared Exponential GPR XI
49 Training data of Squared Exponential GPR XI
50 Predicted versus actual data, Medium Neural Network XII
51 Training data of Medium Neural Network XII
52 Predicted versus actual data, Exponential GPR XIII
53 Training data of Exponential GPR XIII

List of Tables

1 A comparison between four commonly used maintenance strategies . . . 4

2 Different covariance functions for Gaussian process modelling[79] 23

3 List of vessels used for model development. Data are received from Gearbulk's homepage and MarineTraffic 28

4 Table describing the wind force and sea state values, according to the Beaufort scale[49] 35

5 Result of the 10 best algorithms sorted by RMSE 37

6 Result of the 10 best algorithms for test data, sorted by RMSE 38

Acronyms

H_s Significant wave height. 30

r^2 Coefficient of determination. 17, 32

ANN Artificial Neural Networks. 26

CbM Condition-based Maintenance. 2, 3, 5, 10, 13, 42

CM Corrective maintenance. 2, 3

DdM Data-driven Modelling. 5–7, 15, 42

FODay Fuel oil per day. 37, 39

GPR Gaussian Process Regression. 17, 22, 23, 26, 31, 43

HFO Heavy Fuel Oil. 28

IoT Internet of Things. 8

MDO Marine Diesel Oil. 28

ML Machine Learning. 15, 17

MoU Memorandum of Understanding. 9

MT Metric Tons. 33, 34

MTBF Mean Time Before Failure. 3

MTBM Mean Time Between Maintenance. 3

OSP The Open Simulation Platform Joint Industry Project. 9

PbM Physics-based Modelling. 5, 6, 15, 42

PdM Predictive maintenance. 2–4

PM Preventive maintenance. 2–4

RMSE Root Mean Squared Error. i, ii, 20, 32, 43

RUL Remaining Useful Life. 4, 8, 9, 15

SVC Support Vector Classification. 23, 24

SVM Support Vector Machine. 23–26, 31, 43

SVR Support Vector Regression. 23–25

1 Introduction

1.1 Marine industry

In the last few decades, the merchant fleet has seen a large growth in dead-weight tons from 750M dwt in the year 2000, to around 2.1B dwt in 2021[75]. The main contributor to this increase is oil tankers and bulk carriers. The cruise industry has also seen an increase of over 700% since the year 1990 before Covid-19[78]. The industry took a big hit in 2020 during Covid-19 but has gradually been growing since. As a result of the ever-increasing number of maritime vessels in operation, measures for increasing the safety of personnel, customers and ship owners have become more important.

1.2 State-Of-The-Art - Maintenance Strategies

When working in the maritime industry, avoiding system failures is important. Vessels like cruise ships often carry a large number of people which therefore increases the risk of loss of human lives if there should be an accident. Examples are the passenger ferry *MV Doña Paz* in 1987 which is often referred to as the deadliest peacetime maritime disaster in history with 4386 deaths, and the cruise ferry *MS Estonia* which sank in the Baltic Sea in 1994 and claimed 852 lives[58][12]. It is not just the cruise industry which suffers when exposed to a system failure. The shipping industry often relies on tight schedules, and delays as a result of a failure can result in extensive expenses. Vessels such as container vessels, bulk carriers and crude oil tankers often have cargo which can be more valuable than the ship itself. Maintaining strategies to ensure a safe and low-risk environment at sea is therefore substantially important in the maritime industry.

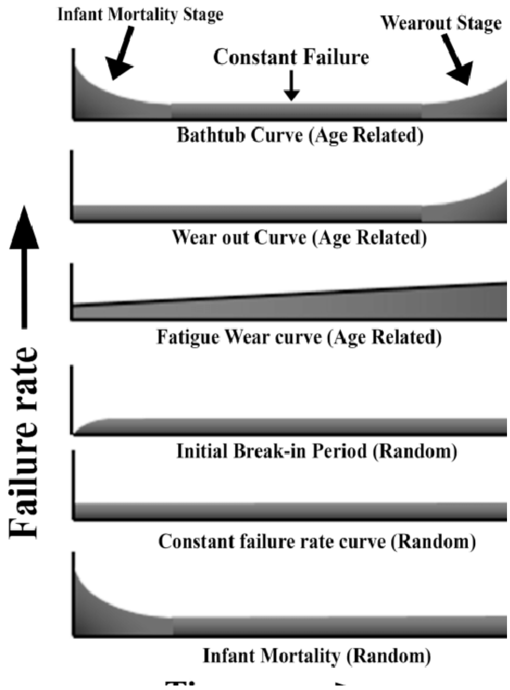


Figure 1: Different failure patterns[1]

To prevent future incidents, different maintenance strategies are often implemented. The reason behind this is to make sure the system keeps functioning well, while not being interrupted in operation. The choice of strategy often depends on the equipment. This is because some equipment can be replaced easily without much cost, and other elements of a vessel can be very expensive to replace if a fault has been detected. It is therefore important to know the system well before implementing a specific maintenance strategy. Although there are several approaches, this section will mainly focus on three methods commonly used today, Corrective maintenance (CM), Preventive maintenance (PM) and Condition-based Maintenance (CbM) which is a type of Predictive maintenance (PdM). The failure rate of components often follow different patterns such as the *Bathtub Curve*, *Wear out Curve*, *Fatigue Wear curve* etc.. The most relevant maintenance strategy is hence often dependent on the failure patterns. An overview of the different patterns is shown in Figure 1.

Corrective maintenance, also known as *Reactive maintenance* or *Run-To-Failure maintenance* is a reactive approach where maintenance is only performed after equipment has failed[27]. CM can in many situations be superior compared to other methods. Unlike PM and PdM, CM is mainly used when the cost of equipment is substantially low relative to the maintenance cost. If the probability of failure for a component does not increase over time, such as the constant failure rate curve and Infant Mortality curve in Figure 1, CM can be a suitable approach where maximising the lifetime of a component is economically beneficial. This often applies to components with low criticality and does not directly affect the performance or safety of the operation. In ship operation, CM can often be non-practical since this requires spare parts on-board which takes up space for potential cargo or other economical-gaining factors. Maintenance on light bulbs is a common example where CM is used.

Preventive maintenance is a proactive, systematic- and routine-based strategy where the goal is to lower the maintenance cost, increase the lifespan and reduce the downtime of an asset. This schedule-based approach is one of the most common practices in the shipping industry today[66]. The maintenance itself consists of regular inspections performed by the personnel. In addition to inspections, activities such as lubrication, cleaning and reassembling of equipment also play a big part in this strategy. The time between maintenance is determined based on the manufacturers' recommendations, reliability- and Mean Time Before Failure (MTBF) analysis, and previous experience. This approach is not always time fixed as various assets have different characteristics regarding failure rate. Components which follow an age-related failure pattern such as the wear-out-, bathtub- and fatigue wear curve from Figure 1 are often most benefited when performing PM. The reason is that components are dealt with before the probability of failure becomes too high.

The last maintenance strategy is predictive maintenance. This proactive approach is a relatively new concept where data from sensors are used to monitor the condition of assets. The data are supplied in real-time and used to accurately *predict* when an asset will require maintenance, in addition to prevent equipment failure[76]. CbM which is a type of PdM also uses real-time data gathered from sensors placed on real assets in a system. The purpose is to get information on when the system needs maintenance and then fix the problem *before* it happens, which helps reduce unplanned downtime and labour hours[20]. There are several advantages related to implementing CbM. The Mean Time Between Maintenance (MTBM) increases since the maintenance only takes place when it is needed, unlike traditional PM. This reduces resources used on unnecessary work, saves time and thus reduces cost. Since monitoring of the asset happens simultaneously as the asset is working, interruptions in production are lowered as a result of not having to perform a shutdown for inspection. An obvious disadvantage to such a system is the immense technical knowledge required to operate it.

Table 1: A comparison between four commonly used maintenance strategies

Maintenance strategy	Pros	Cons
Corrective maintenance (CM)	- Less planning required - Lower short-term costs - Simplified process - Asset is not critical	- Highly unpredictable - Downtime due to interrupted operations - Increased cost of maintenance - Reduces safety
Preventive maintenance (PM)	- Less unplanned downtime - Increased equipment lifespan - Efficient with well-trained personnel - Safer work environment	- More labor-intensive (need more staff) - Expensive upfront cost - Potential for over-maintenance through unnecessary maintenance
Condition based maintenace (CbM)	- Maintenance work is only performed when needed - Fewer unplanned downtime events - Improved prioritization of maintenance time	- High cost of installation, training and maintenance - Sometimes difficult to choose correct sensor equipment
Predictive maintenance (PdM)	- Maintenance needed is predicted - Reduces maximum amount of downtime - Improved automation of maintenance tasks	- Expensive to implement and maintain - Time- intensive

1.3 Condition monitoring in the maritime industry

In the last decades, there has been a transition from typically corrective maintenance, where maintenance is done when a system fails or brakes down, to predictive condition monitoring [Figure 2]. PdM defines a lifetime of a component or system, and assumes an increased failure rate after it's "worn out". Estimating the Remaining Useful Life (RUL) can sometimes be challenging since failure sometimes can be random, and not strictly time-based. The reason is often different elements which can affect the asset such as temperatures, external forces etc. This can typically be observed in PM where performing scheduled maintenance can result in unnecessary expenses or high costs. This is mostly due to maintenance being done either too early or too late. A study done by Nowlan and Heap on aircraft equipment showed a significant difference in the amount of age-related and random failures with 89% being random[1]. Failures encountered on marine vessels, more specifically submariners, followed a similar distribution at 71% for random events, according to the SUBMEPP study done by the US Navy[5]. Since preventive maintenance often relies on age-related failures, a condition-based approach would be the best alternative. Predicting when a system most likely fails will save significant costs and prevent unplanned maintenance. Based on the statistics mentioned above, predictive and condition-based monitoring have a huge potential in the marine industry.

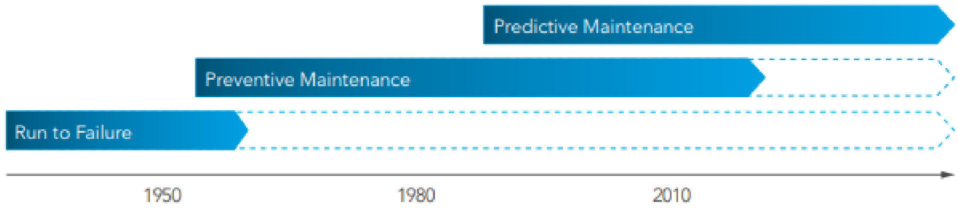


Figure 2: Development in maintenance strategies[43]

The current development of CbM within the marine industry varies greatly. Factors such as the age of the vessel, available equipment and ship type often correlate to the technical monitoring system found on board. Implementation of CbM on older ships can often become quite costly since various parts most likely will require adjustments/replacements to better function with sensors.

It is important to know which parameters to monitor on an asset when implementing CbM. Since the implementation of such a system can be relatively expensive, choosing the relevant assets which have a *critical* impact on the system is crucial. Sensors can measure parameters such as vibration, temperature, pressure and strain.

The main contributor to maritime-related incidents is machinery failure, which often leads to propulsion failure. Components such as bearings, gears and propellers are often monitored because of their criticality to the propulsion system. Condition-based monitoring such as vibration analysis on machinery is one of the most common methods used today on ships. This involves among other things misalignment and imbalance detection and is usually applied for rotating machinery. Oil analysis for wear detection, overheating and contamination of marine machinery is also a common monitoring technique. Data monitoring is usually processed locally on vessels, but systems which allow remote monitoring of sensor data from shore are getting more and more common. This allows the data to be handled by the ship owner or CbM service providers. The advantage of sending real-time data to shore is the opportunity to process data which requires high computational power. The downside is however the increase in latency[43].

1.4 Condition based methods

There are primarily two model approaches used in CbM today; Physics-based Modelling (PbM) and Data-driven Modelling (DdM). PbM, also known as a model-based approach is a mathematical representation of an asset which uses physical characteristics such as mass, torque, momentum and energy equations to simulate its behaviour[41]. Based on the correlation between the simulation and the actual physical measurements, changing the model accordingly will make it possible to diagnose the component. The advantage of using PbM is the opportunity to predict outside the range of existing data and predict future events. Physics-based monitoring is therefore a good way to model when lacking data, and hence suitable for e.g. newly developed equipment[16]. Physics such as fluid mechanics calculations are often highly computationally intensive and therefore often expensive to run. Calculations requiring much computational power can often be slow and dependent on verification methods to be adjusted to fit real-time data. Another limitation concerning PbM is the detailed physical knowledge required of the components. Physics-based modelling of bearings and gear fault propagation are common within the maritime industry[61]. These faults are often a result of a crack growth which contributes to the degradation of the object. Measuring the crack size in a piece of operating machinery can be difficult, thus research has proposed various methods to relate this size to measurable vibration parameters.

DdM is based on computer science, statistics and data science and uses data gathered from previous failures to recognise and learn failure patterns using methods such as machine learning, clustering and neural networks. The quality of the model heavily depends on the quantity and quality of the gathered data. Machine learning algorithms in condition monitoring can e.g. be used to predict future machinery breakdowns through learning from condition parameter patterns gathered from previous faults. An article published in 2010 showed how it was possible to develop a simulation model of a crankshaft, using limited amount of data to make it agree with measurements, and then train a neural network based on responses simulated for different levels of combustion fault in different cylinders. The trained neural network was 100% accurate when recognising actual combustion faults, and was also able to identify faulty cylinders as well as a good estimate of the fault severity[24].

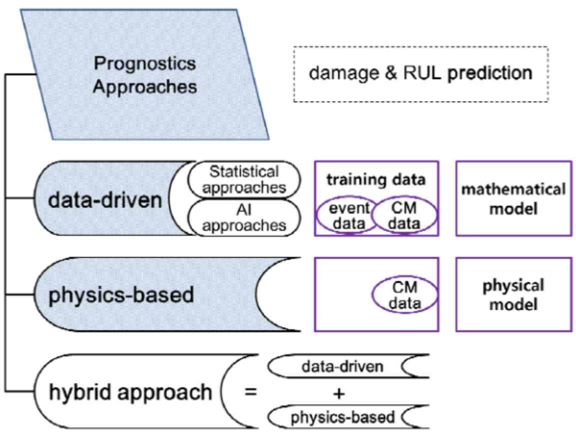


Figure 3: Illustration of different model approaches[6]

A relatively new and popular method is combining both physics-based and data-driven modelling to a so-called *Hybrid model* or *Hybrid machine learning*. This approach uses advanced knowledge about the physical world combined with data science[25]. Data-driven modelling is used to speed up the processing time experienced in physics-based simulation, but lacks accuracy and is heavily dependent on high quality data. Hybrid ML solves this problem by improving the accuracy of DdM by using gathered synthetic data from physics-based simulation for training. Hybrid models can also be used to create what is called a *Digital Twin*. This is further discussed in Section 1.5. An illustration of the different condition monitoring approaches is shown in Figure 3.

1.5 Concept of a digital twin

A brief definition of a digital twin, according to IBM:

"A digital twin is a virtual representation of an object or system that spans its lifecycle, is updated from real-time data, and uses simulation, machine learning and reasoning to help decision-making."[8]

This means that the main goal when implementing a digital twin is to gain knowledge about a physical object or system through a highly complex virtual model (the twin). The model is built on real-time data gathered from a real asset for accurate testing and measurements. An example of a digital twin could be a digital representation of an offshore oil platform. Through information and data from sensors and other physical assets, engineers and operators are not only able to understand how the product is performing, but also how it will perform in the future through different scenarios and conditions. This is crucial since the information can help reduce future expenses connected to major accidents and other undesirable scenarios.

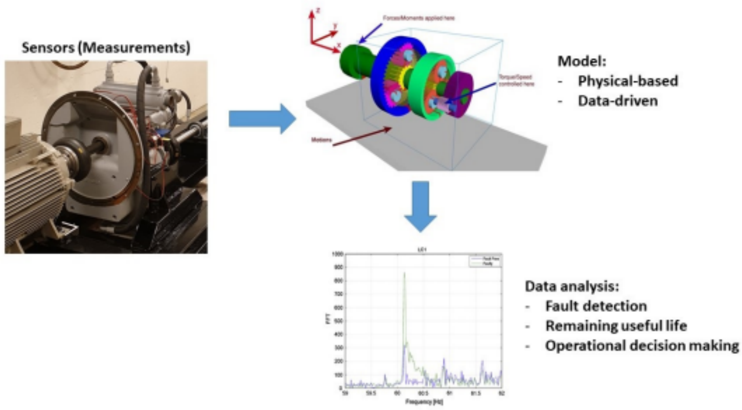


Figure 4: Main steps of digital twin development[36]

There are mainly three steps when developing a digital twin. The first step is to measure data from a physical asset. This is done by first installing sensors on the asset to be monitored[36]. The sensor has to be calibrated, efficient, and placed in optimal areas on the object. The second step is to choose an optimal model for analysing the gathered sensor data. This can either be based on physical-based modelling or data-driven modelling. A combination of these two called hybrid modelling is also possible. The theory and practices behind each approach are presented in Section 1.3. The third and final step evaluates the output data to determine the RUL and fault prediction. Analysing the data requires a comprehensive understanding of the entire system. An illustration of the main steps is shown in Figure 4.

A good example of an industry which has come a long way regarding digital twin technology is aviation. Since carrying passengers at high velocity and altitude is a risk with potentially severe consequences, safety management plays a central role. Engineers and scientists are using Internet of Things (IoT), location and sensor data to simulate how an aeroplane would behave and react in different weather conditions and scenarios[44].

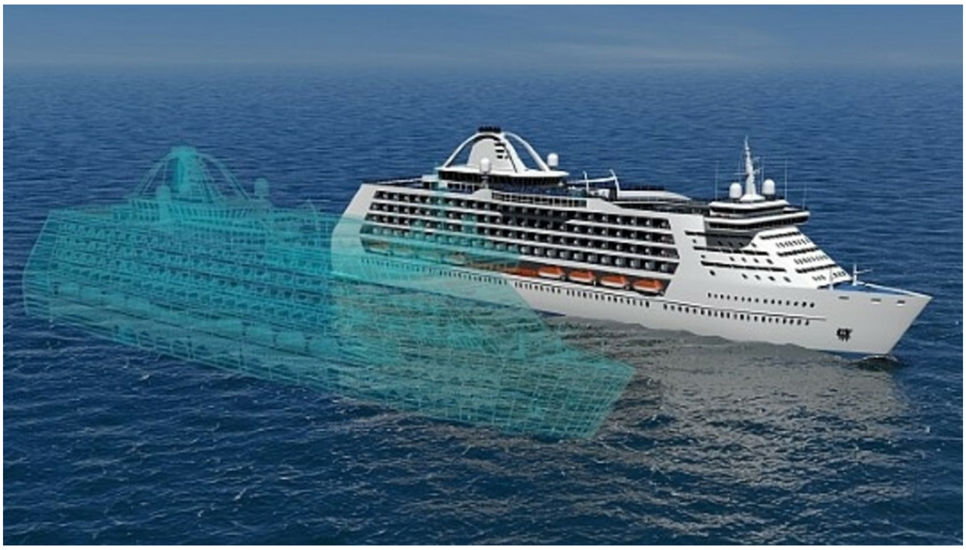


Figure 5: Fictional illustration of a digital twin[56]

1.5.1 Digital twins in maritime applications

Today, the world of technology and engineering is moving towards a more digitalised world where tools such as IoT and artificial intelligence are getting more common when solving problems. As mentioned earlier, the state-of-the-art concerning digital twin implementation varies highly between different industries. The marine industry has a lot to learn from aviation, but several companies such as DNV-GL, MPA Shipping and SINTEF have invested time and resources in condition monitoring through digital twin technology[26]. A project called The Open Simulation Platform Joint Industry Project (OSP) was founded in 2018 and aimed to create an open-source initiative for co-simulation of marine equipment, systems and entire ships[57][55]. The project was developed as a collaboration between DNV GL, Kongsberg Maritime, the Norwegian University of Science and Technology (NTNU) and SINTEF Ocean. The OSP uses digital twins based on large sets of interconnected models and components.

MPA Singapore has recently joined the digital twin trend[18]. They signed a Memorandum of Understanding (MoU) with Keppel Marine and Deepwater Technology (KM-DTech) and TCOMS to develop a digital twin of a tugboat to simulate its behaviour in several scenarios as well as use data analytical tools for improving its control and response. An ongoing study called HealthProp aims to improve safety in Arctic and Antarctic operations by developing a digital twin for intelligent predictive monitoring of the propulsion system drive line of ships. This project, founded by MarTERA, started in 2020 and is scheduled for completion by August 2023[53][10]. The objects monitored include power generation, propeller and components such as shafts, bearings and gears.

There are several applications where a digital twin can help maritime safety and increase both efficiency and operability. Within shipping, areas such as fleet management, port efficiency and optimisation of the end-to-end supply chain are likely to be benefited. As shipping companies often serve multiple customers simultaneously, optimisation of the fleet in terms of e.g. cargo carrying capacity serves an important part[60]. Historical data and predictions of business transactions can help decision-making through the detection of trade patterns. Fleet optimisation could also help make decisions based on unpredictable factors such as weather conditions. A port optimisation model can address questions such as how many berths are needed for the port to be most effective.

Offshore wind is today a rapidly growing industry, and several companies including Equinor have invested in floating offshore wind farms[30]. The main reason for downtime within the offshore wind industry is related to the drivetrain[40]. As a result, condition monitoring and digital twin implementation on the wind turbine saves maintenance costs and optimises operational uptime. This can typically be done by modelling gear loads and load responses. The turbine model can monitor stress and fatigue in real-time, detect degradation and further estimate RUL. Fedem Technology is a company which since 2016 has developed a digital twin of a wind turbine located in northern Norway[81]. Figure 6 illustrates how the real wind turbine interacts with its digital twin.

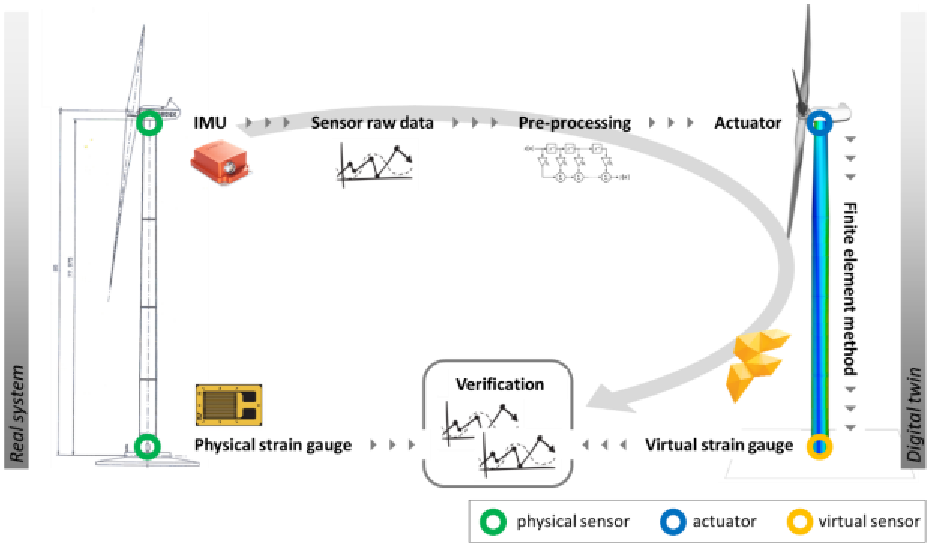


Figure 6: Digital twin of a wind turbine[81]

1.6 Previous incidents

This section will cover two cases of previous incidents within the cruise industry which share several similarities. The reasoning behind this is to discuss if condition monitoring would have been relevant for preventing these incidents from happening. A little summary of each incident, as well as possible CbM solutions, are presented in the following subchapters.

1.6.1 Viking Sky

On March 23, 2019, the cruise ship Viking Sky suffered an engine shutdown in Hustadvika between Kristiansund and Molde, just outside the coast of Norway, and was forced to make an emergency stop[17]. The potential consequences were high as the vessel containing over 1300 people almost ran aground. A helicopter rescue mission was launched which lasted for several hours in extreme weather conditions. This incident quickly got attention from the media worldwide and the question about what caused it arose. Investigations have concluded that there were several causes why the engine experienced a shutdown. Sjøfartsdirektoratet published a press release explaining the main findings in the cause of accident investigation[68]. The engine shutdown was a direct result of low oil pressure. The big waves most likely caused the oil tank to move in such a way which stopped the supply to the lubricating oil pump. The amount of lube oil in the tank was within the limit before the voyage, but this did not account for extreme weather conditions.

Condition monitoring technology could most likely have predicted the outcome of this disaster. If sensors were installed within the oil tank to measure the oil level in different sea states, a trained model could have warned the captain about the result of exposing the cruise ship to harsh weather conditions. This is only one example of how this technology can help decision-making.

Figure 7: Viking Sky in rough weather conditions[31]

1.6.2 MS Kong Harald

MS Kong Harald suffered a similar incident in the same area in August 2021. The Hurtigruten ship, containing 236 passengers and 70 crew members, was scheduled to sail from Kristiansund to Molde through Hustadvika but suffered a power shutdown of the ship's port engine shortly after departure. As repairs were underway, the starboard engine also suffered a shutdown which meant that the ship lost all propulsion power[64]. The vessel began drifting towards the coast but was fortunately stopped by deploying an anchor. This gave the crew enough time to restore the engine power. The initial shutdown was caused by a leak of cooling fluid to the engine, according to officials. It was also discovered that the reason behind the second shutdown was because of a failure of its fuel supply system, mainly caused by wear and tear.

As mentioned in Section 1.3, installing sensors which can measure parameters such as vibration, temperature, pressure and strain is crucial when monitoring the condition of an asset. Since the second shutdown was caused by wear and tear, predictive maintenance could have most likely prevented the incident from happening. These incidents are a good reminder of the potential danger when operating in rough weather conditions often found in such areas.

Figure 8: MS Kong Harald after anchoring[39]

1.7 Research Motivation

With regards to the operation of a vessel, maintaining up-time is crucial when trying to be competitive in the industry. This is especially important in the cruise and shipping industry due to strict schedules and deadlines. As mentioned in Section 1.3 condition monitoring contributes to decreasing downtime and is gradually getting more attention within the maritime industry. Today downtime is often connected to failures or routine-based maintenance.

Irregularities and abnormal conditions in the machinery are often the deciding factors when talking about the cause of a system shutdown or a basic mechanical failure. The key is to track these irregularities *before* it happens, through condition monitoring. As discussed in Section 1.4 there are mainly three ways to analyse and predict these parameters. Data-driven modelling uses data to train and develop a mathematical model for predicting values. Another approach is to develop a physical-driven model which uses physical characteristics to simulate its behaviour. The last model combines these two into a hybrid model.

Vessels are often operating in various conditions. Bad weather at sea is not something unusual, and commercial ships are usually built to sustain a broad spectre of environmental conditions such as wave heights and wind speeds. Since these parameters highly vary over time, the condition of equipment onboard will also change correspondingly. Developing data-driven models based on data collected from different vessels in operation will help predict early faults by observing and discovering unusual behaviour in data plots when testing new data with the constructed model. A research paper published in 2011 showed a direct correlation between fuel oil consumption and engine power[11]. Since irregularities in power output can be a highly possible factor for a system shutdown, the indication of abnormality in fuel consumption can be used to predict if a failure will soon be happening.

Various studies within the condition monitoring field have mentioned how environmental conditions affect machinery. Although, developing a model to predict machinery behaviour as a function of fuel consumption anomalies on marine bulk carriers is yet to be done. Implementation of a data-driven model like this could be useful for future marine operations to forecast the condition of a vessel and improve decision making like e.g. deciding if a vessel should be taking on a voyage in rough sea conditions.

As condition monitoring and digital twin technology gradually increase in popularity, rules and guidelines become important to define. The reason for this is to make sure that every stakeholder in the industry, such as different shipping companies, insurance companies, manufacturers and others has the same standard and understanding of what digital twin technology is and how it works. Some of the views and thoughts on DT technology and condition monitoring among maritime companies are discussed in Section 5.

1.8 Thesis objective

The objective of this thesis is to study failures related to ship accidents and incidents. A data-driven model will further be trained and implemented, using machine learning technology and data from vessels. The relevant parameters used in the model to predict certain behaviours and characteristics of the ship will be defined in Section 3. The main goal of the thesis is to answer the following question;

Is it possible to predict accidents based on anomalies in machinery measurements?

This question is going to be answered through modelling and analysis of ship behaviour. The analysis involves developing a standard model to identify irregularities in the condition of the vessel through data-driven modelling. The model is going to use historical data from five different cargo vessels to be able to predict fuel oil consumption. Fuel oil consumption as mentioned earlier has a direct correlation to power output. Irregularities in power output can indicate malfunction and failures in the machinery on a ship.

1.9 Thesis outline

In the introduction chapter, a study on different types of maintenance strategies is done. The State-Of-The-Art of maintenance strategies are presented and CbM is introduced as the future of maintenance. Typical condition monitoring strategies in the maritime industry and how it helps reduce incidents are discussed. In Section 1.4 three different condition based methods are introduced. The characteristics behind data-driven, physics-based and hybrid-based modelling are also discussed.

In Section 1.5 the concept of a digital twin is presented. The typical steps required for developing a digital twin are shown, and examples of where it is already in use within various industries today are also discussed. The state-of-the-art of DT-technology within the maritime industry is presented and examples of ongoing projects which use this technology intending to improve safety are also mentioned. Short summaries of the Viking Sky incident as well as the MS Kong Harald incident are presented to give the reader an indication of typical failures where condition monitoring could have been a deciding factor. The motivation for this thesis is described under Section 1.7.

Section 2 introduces the fundamentals of supervised and unsupervised machine learning. The chapter will include some examples of common algorithms as well as a discussion about the pros and cons of each method. This machine learning theory will work as a basis for the result chapter in Section 4. Knowing the theory and technical aspects of data-driven machine learning is important to understand the background behind the result. The chapter called Case Study introduces the vessels which the data used in the thesis are retrieved from. This includes five general cargo ships owned by the shipping company Gearbulk. Data of the machinery installed onboard are shortly defined. The most common operational areas are illustrated and typical weather conditions at sea are mentioned. The modelling approach which describes the approach step by step will also be presented in this chapter.

In Section 3 each parameter retrieved from the dataset as well as the parameters used as predictors in the model will be defined. Section 4 contains the machine learning results and compares all the different algorithms in terms of accuracy and efficiency. Plots of predicted versus true responses as well as residuals are shown for both the training and test datasets.

Section 5 is where the results and relevance of the modelling are being discussed. How the model performed and the potential for improvements is discussed. A few interviews were done beforehand and includes their perception of digital twin technology from their point of view. This is also presented in this chapter.

Section 6 is the last chapter and concludes the thesis with final thoughts and suggestions. The contribution to the research community as well as factors for improvement for further work is presented.

2 Methodology

As mentioned in Section 1.4 there are mainly three types of modelling strategies; Physics-based Modelling, Data-driven Modelling and hybrid modelling which is a combination of these. A physics-based approach can for example be used on rotating machinery for condition monitoring purposes to predict the Remaining Useful Life. This is done by reviewing relevant failure modes and their degradation mechanisms. Typically failure modes on gears are crack detection and length estimation and tooth breakage[21]. Scuffing and pitting fatigue on bearing are also possible failure modes. Although physics-based modelling is a useful strategy when modelling the condition of a vessel or machinery, the method is not as relevant for the problem described in the thesis. Data-driven modelling performs best when huge data sets are available for analysis. Since the objective is to predict failures based on abnormality and irregularities found in past data, this thesis focuses on modelling using a data-driven approach.

2.1 Machine learning

Section 4 describes the development and training of a model with the purpose to predict incidents based on abnormal behaviour in data. Since the model is data-driven, the approach is based on computational intelligence and machine-learning methods[69]. This section will cover the fundamentals of machine learning. The goal is to give the reader a general understanding.

Machine Learning (ML) is a term commonly used in computer science. Arthur Samuel, a pioneer within machine learning, famously defined machine learning in 1959 as

"The field of study that gives computers the ability to learn without explicitly being programmed"[63].

ML is usually separated in two subcategories,

- Unsupervised machine learning
- Supervised machine learning

What differentiates one from the other is based on how learning is received and how feedback on the learning is given to the system being developed[72]. Generally, models which use supervised learning are trained with *labelled* data, which means that the model will learn and become more precise over time. On the other hand, unsupervised learning uses pattern recognition to find and recognise patterns in data which are *unlabelled*. Each of the methods is described in detail below.

2.1.1 Unsupervised machine learning

Unsupervised machine learning is where models are not supervised using a training dataset. Instead of using labelled data, the model itself discovers hidden patterns and insights from the given data[38]. The goal is to find the underlying structure of a dataset and gather the data in groups based on similarities and common characteristics. The dataset is then presented in a compressed format. Unsupervised learning algorithms can be grouped into two categories: clustering and association. Some common and popular unsupervised learning algorithms are K-means clustering, hierarchical clustering and neural networks. This machine learning strategy is often preferred when solving more complex tasks and when labelled data is difficult to gather. The drawback is the potentially less accurate result as input data is not labelled. The K-means algorithm is illustrated in Figure 9. In this illustration, K=2 clusters are chosen and two different centroids are randomly placed on two data points. All the other data points are assigned to the closest cluster centroid. The centroids are then placed based in the centre of their assigned data points. These steps repeat until the two centroids have successfully divided the dataset into two groups (clusters).

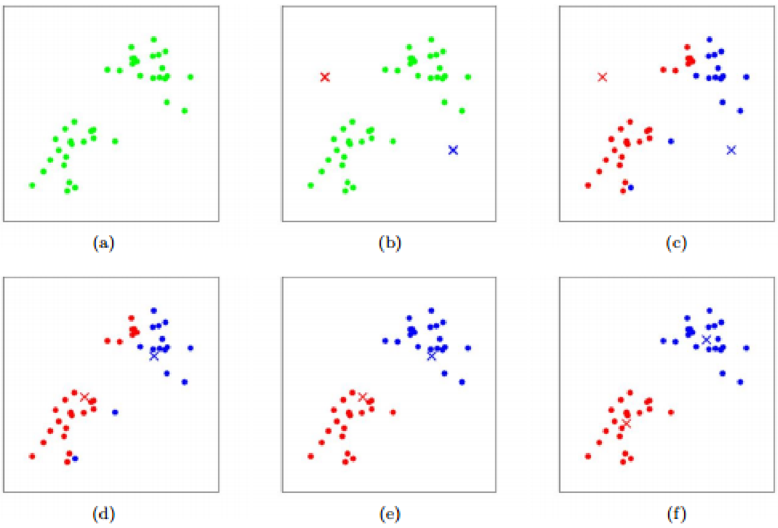


Figure 9: K-means algorithm[59]

2.1.2 Supervised machine learning

Supervised ML uses labelled datasets to train the model for a specific output. The system is adjusting existing data to predict the outcome. The purpose is to compare the actual output with the output learned by the algorithm to find errors and modify the model accordingly. An example of supervised ML is training a model for animal recognition. The strategy is done by showing the algorithm a data set of various animal pictures and labelling them correctly. The model will then find the patterns for the different labels and use this as a basis when deciding if an illustration is for example a cat or a dog. After the model is trained, a new set of pictures the model has not previously observed are inserted for validation to calculate the accuracy. If the animal suggestion is false, the model is then modified. A basic illustration of supervised learning concerning geometrical figures is shown in Figure 10.

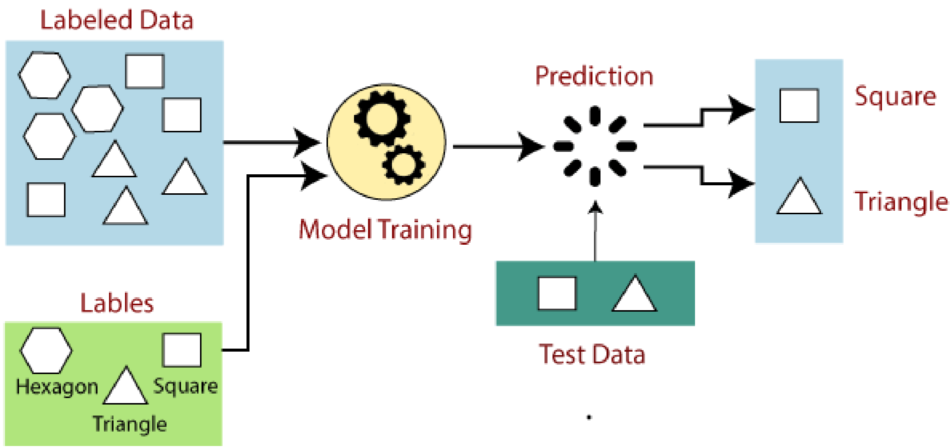


Figure 10: Example of supervised learning with geometrical figures[37]

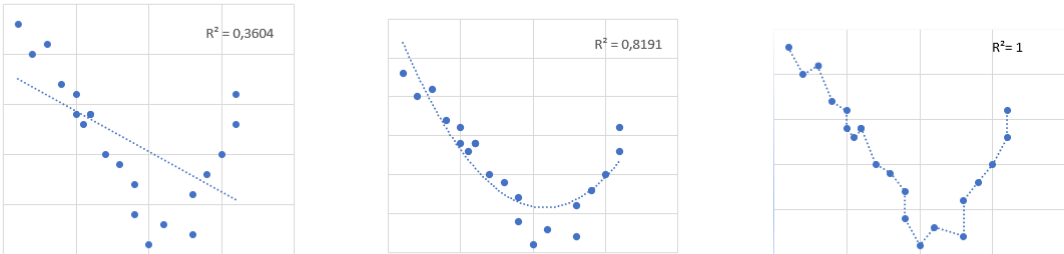
There are several approaches within supervised learning strategy. *K-nearest neighbour*, *Gaussian Process Regression (GPR)*, *decision trees*, *linear regression* and *neural networks* are some of the most common algorithms. Some of the more relevant methods will be further described and discussed in Section 2.2.

2.1.3 Importance of model fitting

Several different factors determine whether a model is good or bad. The accuracy and versatility of a model are not only reflected in a single correlation score such as the Coefficient of determination (r^2), but also in the ability to accurately predict new data which is fed into the model, also called test data. A model which correlates perfectly with current data but fails when new data are inserted is called an *overfitted* model. The definition of overfitting is defined by Oxford as;

"The production of an analysis which corresponds too closely or exactly to a particular set of data, and may therefore fail to fit additional data or predict future observations reliably."[47].

To counter overfitting, separating data from the original dataset before training the model is a good approach. After the model has been trained, the separated data can then be inserted into the model to test its ability to adapt to new data. Figure 11 illustrates three different regression models on the same dataset. The graph in the middle has a good correlation and is an example of a nice fitted model.



(a) Underfitted model (b) Good fitted model (c) Overfitted model

Figure 11: Examples of different fitted regression models

The figure to the left has almost no correlation and is underfitted. An underfitted model is often unable to accurately capture the relationship between input and output parameters. The error rate is high in the training set, as well as the unseen test set. Underfitting often occurs when the model is oversimplified and the training time is low. Simple solutions to avoid underfitting can be to decrease regularisation. Regularisation is normally used to decrease the variance with a model by applying a penalty to the input parameters with the larger coefficients[29]. Reducing regularisation results in higher complexity and variation in the model which means more successful training.

Another way to avoid underfitting is to increase the duration of model training. Introducing more features to the model will also increase the complexity, hence yielding better training results. Lastly removing noise from the data will also help avoid underfitting. Figure 12 is fairly similar to Figure 11 but illustrates classification rather than regression.

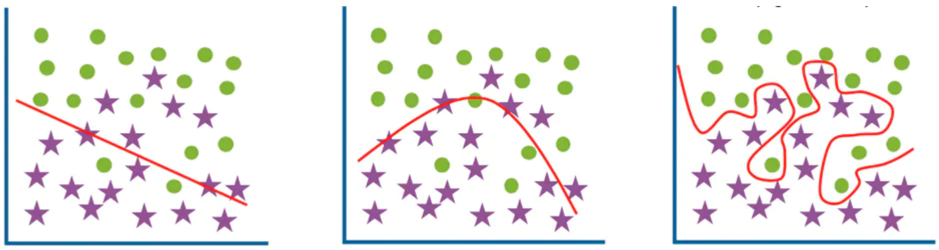


Figure 12: Figure of underfitted, fitted and an overfitted classification model[28]

The last graph shows what an overfitted model could look like. Every single data point is located on the dotted line and external test data will almost certainly differ from the line, which makes it a less ideal model. There are some strategies to avoid overfitting when training a model. Increasing the amount of training data can increase the model accuracy by giving more opportunities to analyse the dominant relationship between the input and output variables. This works best when the data inserted into the model is clean, which means a well-filtered dataset that is optimised to create a good model. Unlike underfitting, decreasing the number of features will simplify the model and further reduce the risk of overfitting. Inserting data noise can also help avoid overfitting. Data noise is meaningless data which makes the data harder to fit, thus harder to overfit.

2.1.4 K-fold Cross Validation

K-fold Cross Validation is a powerful measure for the prevention of overfitting. The *k* stands for the number of groups the given data sample is divided into. The procedure for this strategy is rather simple[14];

1. Mix the training set randomly
2. Divide the training set into *k* parts
3. (a) Select one of the k parts as the test data set
   (b) Use the rest of the parts as the training data set
   (c) Insert the test data into the trained model and evaluate the result
   (d) Give the model a score based on the result
4. Repeat the third step until every part has been used as the test data set
5. Calculate the average score of the k scores acquired
6. The average score is now the accuracy of the model

This method is popular in the world of machine learning. The reason is that it is simple to understand and it generally results in a less biased or less optimistic estimate of the model, compared to other methods. Choosing a good k-value should be done carefully as a poor k-value may result in a high model score variance and therefore a misconception of the performance of the model. A common strategy for choosing the value of k is to make sure that each group has enough data to be representative of the whole dataset. The most popular k-value in machine learning is 10 as it has generally shown to be relatively accurate at a low computational expense. A value of 5 is also commonly used, but this will of course depend on the dataset as there is not a fixed correct answer to this decision. Figure 13 describes how the procedure is performed.

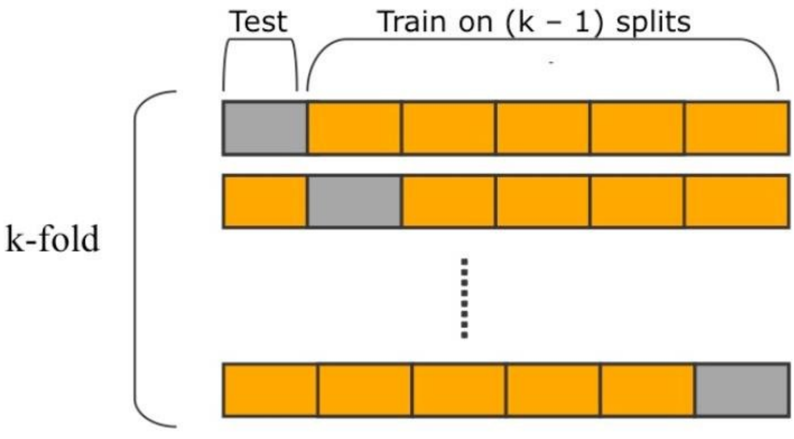


Figure 13: The procedure of K-fold Cross Validation[71]

2.1.5 Root Mean Square Error

The Root Mean Squared Error (RMSE) is one of the most popular methods within supervised learning applications and is used to evaluate the quality and accuracy of the predictions done by the model. The RMSE value is calculated using Euclidean distance and describes how far the predictions fall from measured true values[15]. To calculate the RMSE, the difference between the predicted and observed value for the i^{th} observation on the dataset is calculated for each data point. The absolute value of the error is squared and added together for each i and divided by the total sample size. The square root of this value is the RMSE. The formula for RMSE is defined in Equation 1.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \|y(i) - \hat{y}(i)\|^2}{N}} \tag{1}$$

When experimenting with numerous different machine learning algorithms to find the best fit for a model, having the RMSE value makes it substantially easier to choose the preferred model. Choosing the best model based on e.g. RMSE will be further discussed in Section 4.

2.2 Machine learning algorithms

Various machine learning algorithms can be used when developing a model. The best performing algorithms often vary and are often highly dependent on the dataset used. In this chapter different popular supervised machine learning algorithms are introduced and discussed.

2.2.1 Linear Regression

Linear regression is a simple way to identify the relationship between a dependent variable, also called the *response*, and one or more independent variables, commonly called *predictors*[35]. Each type of linear regression seeks to plot a straight line of best fit. A process using multiple predictors is referred to as *multiple linear regression*, and a process with only one predictor is referred to as *simple linear regression*. These models are often fitted using the least-squares approach. A mathematical representation of a simple linear regression algorithm is shown in Equation 2 and a multi-linear regression algorithm is shown in Equation 3.

$$y = \beta_0 + \beta_1 X_1 + \varepsilon \tag{2}$$

$$y = \beta_0 + \beta_1 X_1 + ... + \beta_n X_n + \varepsilon \tag{3}$$

The variable y represents the predicted value of the dependent variable. β_0 is the value of y when all other parameters are equal to 0. The term $\beta_1 X_1$ represents the regression coefficient of the first independent variable X_1. In other words, β_1 is the *weight* of the variable X_1. n is the number of independent variables and ε represents the error in the model. This is the variation of the estimated y.

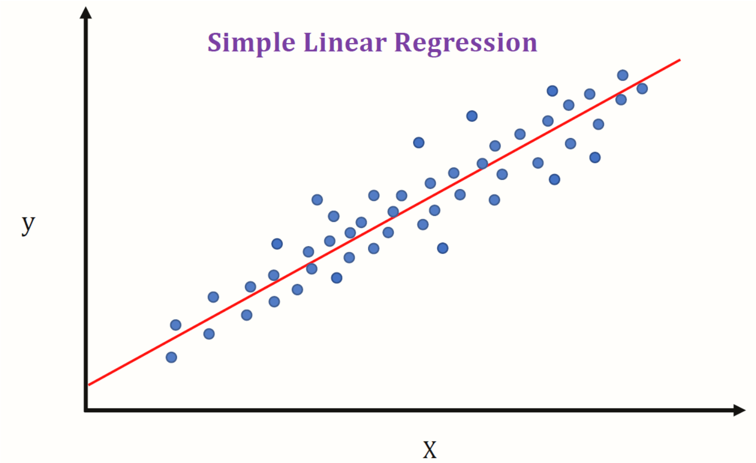


Figure 14: Figure of simple linear regression[54]

The advantage of linear regression modelling is its simplicity. The algorithm is computationally efficient with fast speed and generally has a relatively interpretable output. Despite this, it can become too simple for more complex modelling problems. The assumptions for linearity do make it sensitive to outliers and homoscedasticity is assumed[42].

2.2.2 Gaussian Process Regression

Gaussian Process Regression (GPR) is a nonparametric, Bayesian approach to regression and is becoming increasingly more popular and common in the machine learning society[62]. The algorithm uses probability distributions over all possible values rather than learning exact values for every parameter in a function, which differs from many other supervised learning methods. The algorithm is nonparametric which means not limited by a specific functional form. Instead of probability distribution calculation of parameters from a function, GPR calculates the probability distribution of all permissible functions which fit the data.

The GPR models are kernel-based probabilistic models with a finite collection of random variables with a multivariate distribution[80]. A kernel (or covariance function) defines the behaviour of the function which is modelled, and thus important to choose wisely to get an accurate model. The covariance function is a measure for the correlation of two states, x and x'[79]. This is shown in Equation 4.

$$k(x,x') = k(||x-x'||) = k(r) \tag{4}$$

An overview of some of the different kernels used in GPR and their expression are shown in Table 2 with further explanations of variables.

Table 2: Different covariance functions for Gaussian process modelling[79]

Kernel	**Mathematical expression**
Constant	$k = {\sigma_0}^2$
Linear	$k_{lin}(x,x') = x^T x' + c$
Polynomal	$k_{poly}(x,x') = (x^T x' + {\sigma_0}^2)^p$
Squared exponential	$k_{SE}(r) = exp(-\frac{r^2}{2l^2})$
γ - exponential	$k_\gamma(r) = exp(-(\frac{r}{l})^\gamma)$ for $0 < \gamma \leq 2$
Rational quadric	$k_{RQ}(r) = (1 + \frac{r^2}{2\alpha l^2})^{-\alpha}$
Power	$k_p(r) = -r^p$
Matern - 5/2	$k\left(x_i, x_j\right) = \sigma_f^2\left(1 + \frac{\sqrt{5}r}{\sigma_l} + \frac{5r^2}{3\sigma_l^2}\right)\exp\left(-\frac{\sqrt{5}r}{\sigma_l}\right)$

- σ_0^2 = hyperparameter
- c = constant
- l = lengthscale
- p = polynomial degree
- α = relative weighting of large-scale and small-scale variations

Gaussian processes are fully specified by its mean $m(x_i)$ and the covariance function $k(x_i, x_j)$, and hence defined in Equation 5.

$$f(x_i) \sim GP(m(x_i), k(x_i, x_j)) \tag{5}$$

GPR is often a preferred model strategy when dealing with smaller datasets because of their well-tuned smoothing and remail computationally affordable. Gaussian processes can also be optimised exactly, given the values of the hyperparameters, which are also valid for other kernel methods. With the use of kernels, the method can be highly versatile in handling various data structures. A drawback is often a high computational time compared to the alternatives.

2.2.3 Support Vector Machine

Support Vector Machine (SVM) is a supervised machine learning technique mainly divided in two categories, Support Vector Classification (SVC) and Support Vector Regression (SVR)[32][9]. It was first developed by Vapnik in 1995 and is based on statistical learning theory[19]. Just like Gaussian Process Regression, SVM also operates with kernels. The algorithm is used for recognising subtle patterns in complex datasets. It uses a high-dimensional feature space. The main method in SVC is finding an optimal hyperplane which separates two classes. A hyperplane is generally a linear decision boundary that differentiates classes in SVM. This operates as a line in two dimensions and as a plane in three dimensions.

To find the desired hyperplane, the norms from the vector w need to be minimised. A norm is referred to the length of a vector. This measure is the same as maximising the margin between two classes, which is the distance between the hyperplane and the optimal hyperplane and normally never contains any data points. The closest data points to the hyperplane are called support vectors[50].

As for SVR the method is to construct a hyperplane which is positioned as close to most of the data points as possible. This means choosing a hyperplane with small norms while also minimising the sum of the distance between the hyperplane and the data points[74]. Figure 15 shows an illustration of the two different SVM methods. The figure to the left shows a non-linear SVR model and the right-hand model is an example of SVC. The dotted line represents the hyperplane and the green lines show the tolerable error.

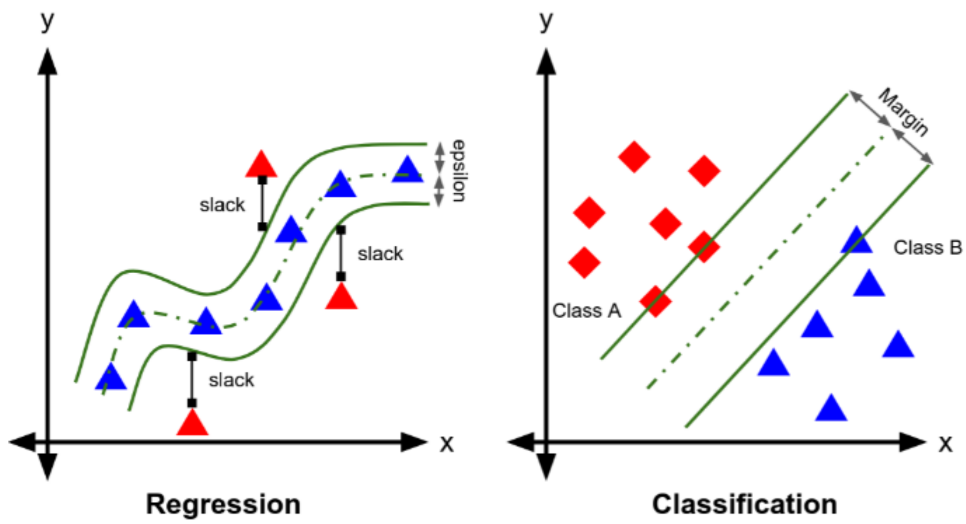
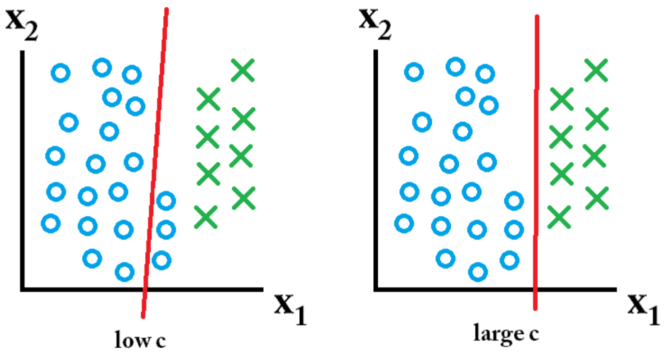


Figure 15: Illustration of a linear SVC and a non-linear SVR model[52]

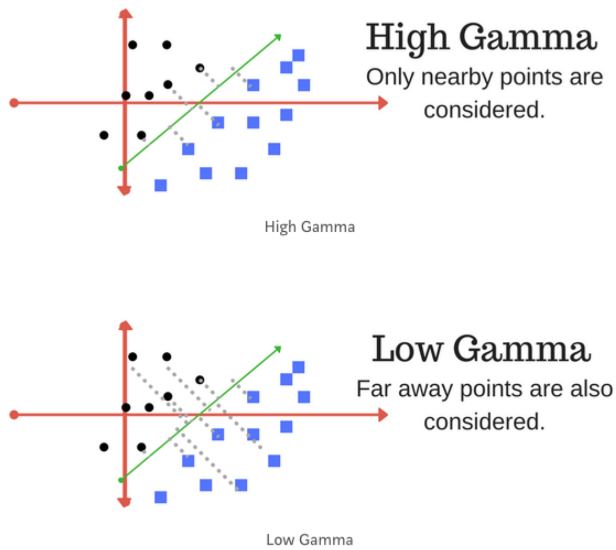
To achieve a good model in SVM, tuning parameters are important. One of the relevant parameters are the *regularisation* (also referred as c) parameter. This tells how important avoiding misclassifications on each training set are. It is the relation between model complexity and empirical error. A large value often results in an overfitted model, and when the value is small, the risk of underfitting the SVM increases.

Another parameter which is usually tuned is the *gamma* parameter. The parameter defines how far the influence of a single training example reaches. With a high gamma, only points close to the separation line are considered. Increasing the gamma often results in overfitting the SVM. A small gamma means including points far away from the separation line. Instead of overfitting, decreasing the gamma leads to an underfitted model as more points are considered. Figure 16 shows how low and high c and gamma parameters affect the model.

The last relevant tuning parameter is the *epsilon* parameter. This parameter defines the loss function that ignores the error. In SVR, epsilon gives a maximum allowed error for the regression model.



(a) Difference between high and low c



(b) Difference between high and low gamma

Figure 16: Illustration showing high and low c and gamma parameter[67][51]

When comparing SVM to logistic regression, SVM is often preferable if the dataset is small and complex. It even performs well with unstructured data. Implementation of different kernel solution functions makes the model quite versatile when solving complex problems. SVM also have good scaling of high dimensional data. A drawback of having a lot of potential kernel solution functions is having to choose as there can sometimes be more than one good model. Similar to GPR time spent training can be relatively time-consuming when using large datasets[3].

2.2.4 Artificial Neural Networks

Just like SVM and GPR, Artificial Neural Networks (ANN) are used for solving complex non-linear relationships between features and targets. This method can be used in both supervised and unsupervised learning. ANN is a biologically inspired computational model and consists of hundreds of single units called *artificial neurons* connected with coefficients, or weights, which make up the neural structure[2]. It is a parameterised system and does not gather its knowledge through programming, but instead detects patterns in data and is trained through experience. A network of functions is used to translate a set of data inputs into the desired output. The components of the model can be simplified into three parts; an input layer, one or more hidden layers, and lastly an output layer. Figure 17 illustrates an example of a single-layered artificial neural network.

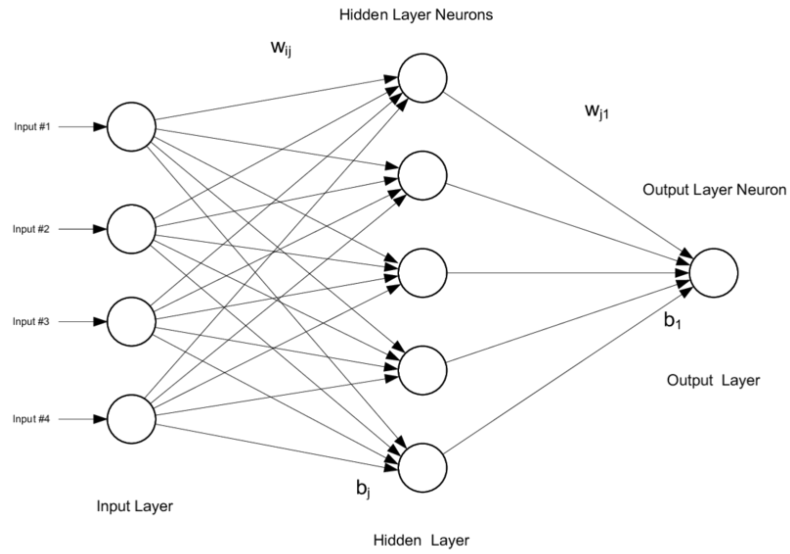


Figure 17: Example of a shallow neural network[4]

The input layer consists of nodes of information which is data used for model learning. The hidden layer consists of a set of neurons and is the place where all computational calculations on the input data happen. The output layer is where the conclusions and results of the model's calculations lie. It is possible to have more than one output node if the classification problem is non-binary[23]. A network consisting of only one hidden layer is called a *shallow neural network*. This is the simplest form of a neural network. Examples of shallow networks are *Narrow NN, Medium NN* and *Wide NN*. Two or more hidden layers, such as *bilayered NN* and *trilayered NN* are often called deep neural networks. How a neuron works concerning input and outputs is illustrated in Figure 18.

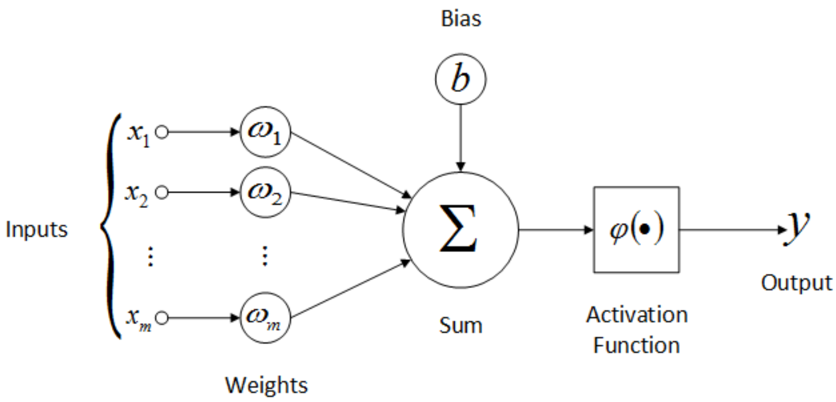


Figure 18: Mathematical model of an artificial neuron[22]

The mathematical expression is shown in Equation 6. x_m is the input value from the input layer, ω_m is the weight of the input, ϕ is the activation function and b is a constant added for better model fit. y represents the output.

$$\mathrm{y} = \varphi\left(\sum_{\mathrm{m}}(\omega_{\mathrm{m}}\mathrm{x}_{\mathrm{m}}) + \mathrm{b}\right) \tag{6}$$

2.3 Case Study

Figure 19: Plover Arrow[77]

Vessel data from the company Gearbulk are processed and used when developing a data-driven condition-based model. Gearbulk is an international shipping company, and together with its related companies, operates the world's largest fleet of open hatch vessels[33][34]. Their fleet consists of around 70 open hatch and other specialised and conventional vessels. Gearbulk has a joint venture with G2 Ocean which operates the fleet. The data used in the developing process of the model in this study is received from five different vessels. The data are chosen over a 16-month period starting from the 1st of January 2021 to the 4th of May 2022. An overview of each vessel is shown in Table 3.

Table 3: List of vessels used for model development. Data are received from Gearbulk's homepage and MarineTraffic

Name	**Ship type**	**LOA [m]**	**Breadth [m]**	**Engine [kW]**	**DWT**	**Built year**
Plover Arrow	OPEN HATCH GANTRY CRANE (OHGC)	199.7	32.2	11 520	55 459 mt	1997
Swift Arrow	TOTALLY ENCLOSED FORESTRY CARRIERS (TEFC)	185	30.4	9 378	42 276 mt	1992
Corella Arrow	OPEN HATCH GANTRY CRANE (OHGC)	225	32.27	12 577	72 863 mt	2009
Weaver Arrow	OPEN HATCH GANTRY CRANE (OHGC)	199.7	32.2	11 520	55 402 mt	1998
Avocet Arrow	OPEN HATCH JIB CRANE (OHJC)	199.98	32.26	7 730	62 841 mt	2015

2.3.1 Machinery

All the vessels in the study use a two-stroke marine diesel engine with an in-line cylinder configuration. Diesel engines are often sorted in three categories, slow ($<$ 300 RPM), medium (300 - 900 RPM) and high speed ($>$ 900 RPM)[46]. The rotational speed at the maximum continuous output power of all the machinery used in this thesis has a range between 90 to 130 RPM. They are therefore considered slow-speed engines. There are several advantages of having a two-stroke main propulsion engine compared to a four-stroke. The running cost of operating a vessel is often reduced as a result of the ability to burn low-grade fuel oil[7]. The thermal and engine efficiency of two-stroke engines are often much better. This type also has a higher power-to-weight ratio and is, therefore, able to carry more weight and cargo. Some other factors are the increased reliability while operating and less required maintenance. The fuel type used in the vessels is either Heavy Fuel Oil (HFO) or Marine Diesel Oil (MDO). The machinery data from each vessel are shown in Section A.

2.3.2 Geographical areas

The vessels mentioned in Table 3 are used for bulk transport and are in service all around the world. The dataset used in the model contains data from all around the world, which gives the model a relatively general and versatile field of application. As Figure 20 illustrates, the majority of the ports visited are located in Asia, Oceania and South America, but various other ports are also included. The heatmap is generated from *Maply* which is an application for visualising and analysing location-based data[48]. This map is based on the port destinations found in the dataset, independently of the specific vessel. Each port has been given a weighting based on the number of visits.

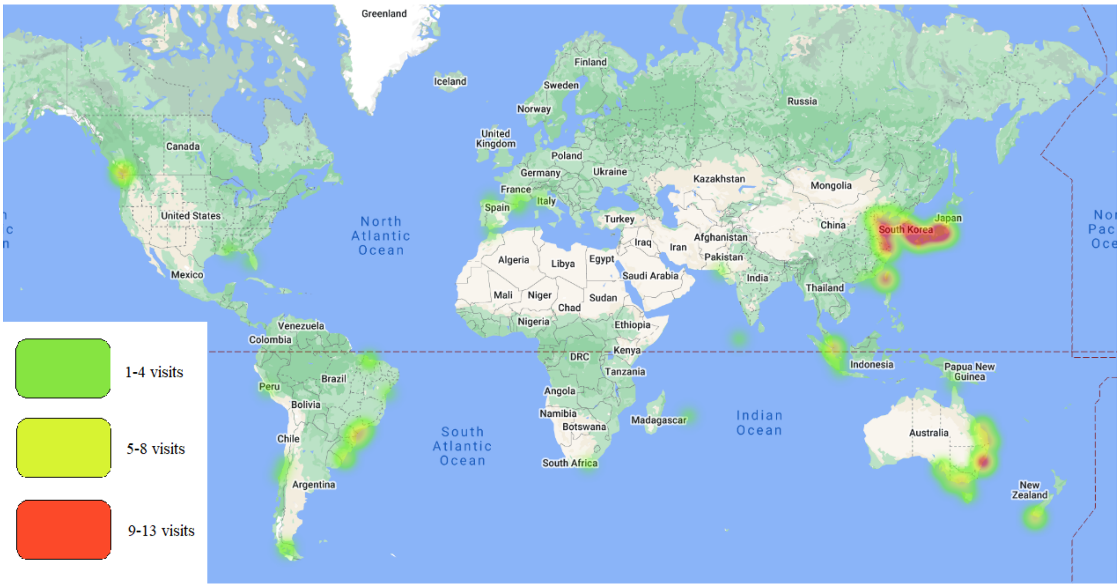


Figure 20: Heat map showing port activity by the vessels from January 2021 to May 2022

Figure 21 illustrates the frequency of port visits for each country. As observed, Japan, Australia, China, South Korea and Brazil represent the majority of the visits.

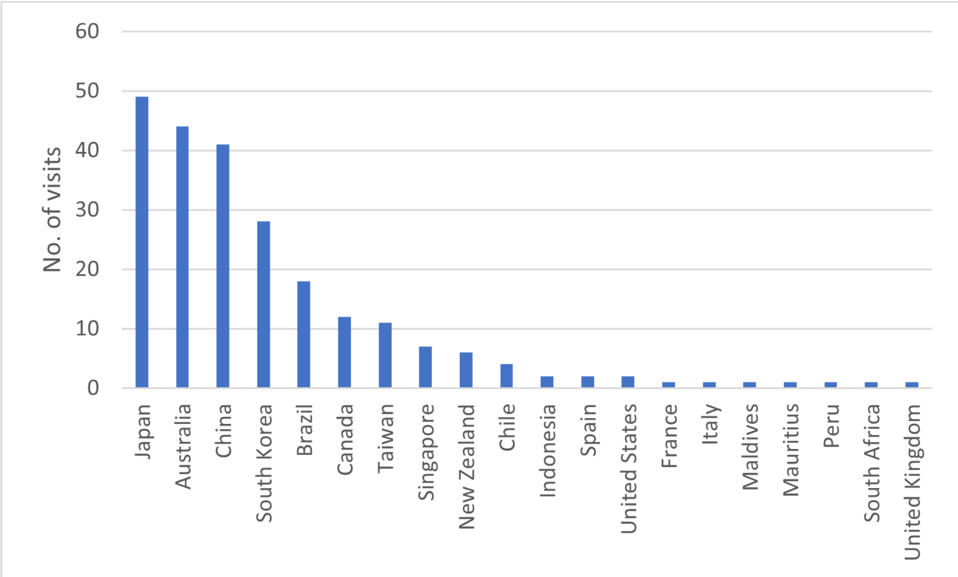


Figure 21: Number of port visits sorted by country between January 2021 and May 2022

2.3.3 Weather conditions

As mentioned, the vessels operate in all kinds of sea conditions and are built to sustain rough weather conditions. A study from 2020 on historical wave data showed a Significant wave height (H_s) ranging between 2-3.5 meters in the Pacific ocean and 1-2.5 meters in the Atlantic ocean, between the year 1979 and 2014[70]. Big waves and high wind speeds are often direct factors to wear and tear on offshore ships and structures. The weather data collected from the dataset in this study are based on the wind and waves measured while sailing. Different wave heights and wind speeds are categorised according to the Beaufort scale. The scale ranges from 0 to 12 for wind force, 0 to 9 for the sea force and describes the weather condition in form of wind speed and wave height. The Beaufort scale is defined in Table 4 under Section 3.1.

2.3.4 Modelling approach

The goal of the thesis is to make a model which can predict machinery failure. There are various measurable parameters which can indicate a probability of failure, and two of the most relevant variables are main engine power and auxiliary power. As mentioned in Section 1.7, there is a direct correlation between power and fuel oil consumption[11]. As a result, the model will be constructed to predict fuel oil consumption per day since the dataset provided by Gearbulk does not include power consumption.

A dataset containing *four* of the five vessels is then exported to Matlab as a formatted excel file. Data cleaning is performed to remove data which have a negative influence on the model. Handling missing data values are also important as not having data sometimes counts as a value of zero, which can heavily influence the model. To assure a versatile and well-fitted model, the dataset from vessel no. 5 is excluded from the model training and used for later testing.

The training of the dataset is performed by an in-built application in Matlab called *Regression Learner*. This is a tool which lets you explore data, select features, specify validation schemes, optimise hyperparameters and so on without writing any code. After importing the training dataset to Regression Learner, the response parameter is chosen from the dataset which in this setting is fuel oil consumption per day. As mentioned in Section 2.1.4, avoiding overfitting during model training is important. K-fold is performed with five cross-validation folds. The K-fold strategy is not to be confused with excluding vessel no. 5, as K-fold is only for testing the model made from vessels 1 to 4.

A set of relevant parameters from the dataset are chosen as *predictors*. The method used for determining the best predictors is called *Feature Selection*. The goal of this method is to mainly remove non-informative or redundant predictors from the model[45]. Feature selection can either be performed using algorithms such as *MRMR* and *F Test* or simply iterate between several combinations of predictors from the dataset and compare the results to find the best group. The last method is the approach used in this thesis. When all relevant features are selected, plots representing the fuel oil consumption per day as a function of each feature are generated.

To train the model, the user has to choose a fitting algorithm. There are over 20 algorithms in Regression Learner, and the main methods include Linear Regression Models, Regression Trees, Support Vector Machine (SVM), Gaussian Process Regression (GPR), Kernel Approximation Regression, Ensembles of Trees and Neural Network. The chosen approach was to train all the algorithms simultaneously and further choose a fitting algorithm for testing. An overview of all algorithms used for training is listed below.

- Linear Regression Models
  - Linear
  - Interactions linear
  - Robust linear
  - Stepwise linear
- Regression Trees
  - Fine tree
  - Medium tree
  - Coarse tree
- Support Vector Machines
  - Linear SVM
  - Quadratic SVM
  - Cubic SVM
  - Fine Gaussian SVM
  - Medium Gaussian SVM
  - Coarse Gaussian SVM

Gaussian Process Regression

- Squared Exponential GPR
- Matern 5/2 GPR
- Exponential GPR
- Rational Quadratic GPR

Kernel Approximation Regression

- SVM Kernel
- Least Squares Regression Kernel

Ensembles of Trees

- Boosted Trees
- Bagged Trees

Neural Network

- Narrow Neural Network
- Medium Neural Network
- Wide Neural Network
- Bilayered Neural Network
- Trilayered Neural Network

After the training is done, a score which is used to check the fitting of a model is generated and is usually called Root Mean Squared Error. The math behind RMSE is explained in Section 2.1.5. This indicates the difference between the actual data and the regression line and is referred to as the standard deviation of the prediction errors, also commonly called residuals. A plot of predicted versus actual data, as well as residuals are made for each algorithm. The training results include parameters such as Coefficient of determination, predicting speed and model training time.

Lastly, the model is tested by inserting data from vessel number 5 which has not been exposed to the model before. The results of the modelling are shown and discussed in Section 4. A figure illustrating the approach explained in this section is shown in Figure 22.

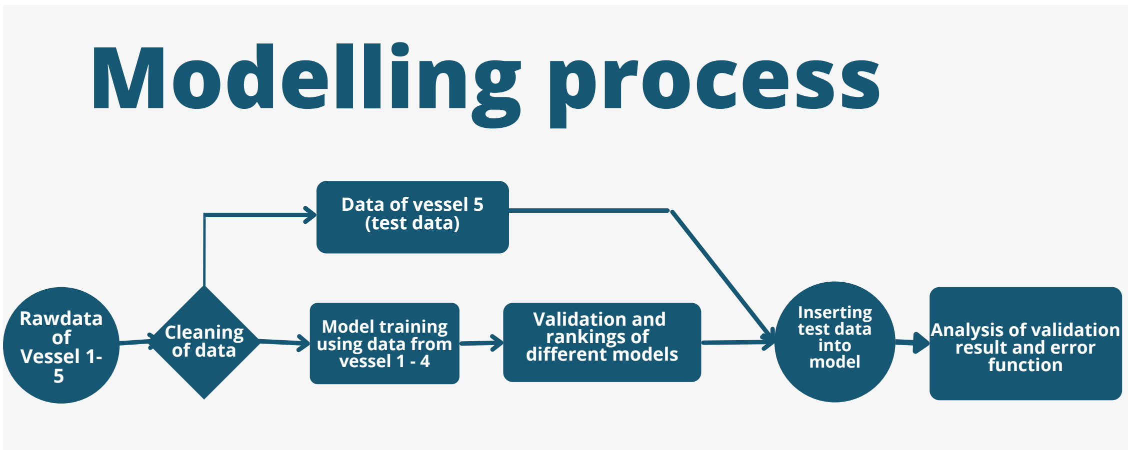


Figure 22: Figure of the modelling process

3 Data collection

The data used in this thesis are collected from the shipping company Gearbulk and involve data from five of their vessels. As mentioned in Section 2.3, the period of the data is from January 2021 to May 2022. The data are tracked by sensors and observations. A report containing all the data is referred to as a Noon report and is a data sheet prepared by the ship's chief engineer once each day at noon local time.

The dataset used in this thesis contains over 20 parameters and has over 1400 data points/measurements quite evenly spread between all five vessels. A list of the most important parameters are shown below, and their value is valid for the specific interval between the measurements.

- Vessel name
- Destination
- Sailing time [days]
- Speed [knots]
- Miles
- Fuel oil consumption
- Fuel oil consumption per day
- Miles per Metric Tons (MT) fuel
- RPM
- Average draft
- Trim of vessel
- Sea direction
- Sea force
- Wind direction
- Wind force
- Report date
- % Sulphur emissions
- Cargo in MT

When choosing the most relevant parameters, it is important to exclude data with no relation to the response parameter. If the model tries to adjust itself based on irrelevant data, it can result in a wrong correlation basis and ruin the purpose of the relevant data. The parameters used to develop the model are defined in Section 3.1 and based on the feature selection method described in Section 2.3.4.

3.1 Parameters

This chapter will introduce the different parameters used, as well as their relevance and impact on the results of the model.

3.1.1 Sailing time in days and miles

This parameter is defined as the amount of time between the data measurements where the vessel is in operation. The value generally spans between 0 and 1 day, but some of the values are 1.04 and 0.94 days which is a deviation of one hour due to changes in time zones. This happens when the vessel either sails westbound or eastbound. The *miles* parameter is defined as the distance in miles which have been sailed between each measurement.

3.1.2 Fuel oil consumption per day

Fuel oil consumption per day is the amount of fuel oil consumed divided by the sailing time. This is the parameter that the model will try to predict as it is closely related to power output, which can indicate if there is something wrong with the machinery.

3.1.3 Miles per metric ton fuel

The number of miles covered per Metric Tons fuel is related to fuel efficiency. Large and unusual fluctuations can indicate a malfunction in the machinery or other fuel-related components on the vessel.

3.1.4 RPM

The RPM, also known as revolutions per minute, describes the speed of rotation in the machinery. As mentioned earlier, the vessels in this case study operate with slow-speed two-stroke engines and have maximum RPM values around 100 RPM $\pm$ 20 RPM.

3.1.5 Average draft and trim

The average draft of the vessel is directly correlated to the amount of cargo on board. The trim of the vessel is defined as the difference between the forward and the aft draft. Figure 23 shows a vessel with a difference in forward and aft draft, resulting in a trim. Different trims have an influence on the fuel consumption since viscous and drag resistance are dependent on the underwater surface area.

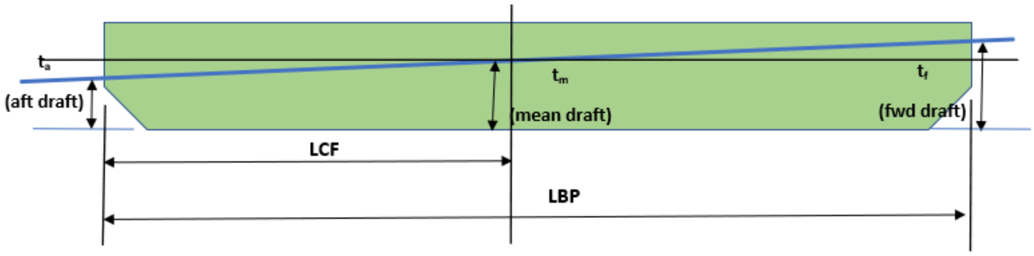


Figure 23: Figure showing a vessel with a trim[73]

3.1.6 Wind and sea force

Each vessel is exposed to different weather conditions in the form of wind and waves when sailing across the oceans. The data describing the conditions are given in *wind and sea forces* instead of knots and wave heights. The wind force ranges from 0 to 12 and the sea force ranges from 0 to 9 on the Beaufort scale. Table 4 defines the different values for the wind and sea forces[49].

Table 4: Table describing the wind force and sea state values, according to the Beaufort scale[49]

Wind force	Description	Wind speed [knots]	Probable wave height [m]	Sea force
0	Calm	<1	<0.1	0
1	Light Air	1 - 3	0.1	1
2	Light breeze	4 - 6	0.2	2
3	Gentle breeze	7 - 10	0.6	3
4	Moderate breeze	11 - 16	1.0	3 - 4
5	Fresh breeze	17 - 21	2.0	4
6	Strong breeze	22 - 27	3.0	5
7	Near gale	28 - 33	4.0	5 - 6
8	Gale	34 - 40	5.5	6 - 7
9	Strong gale	41 - 47	7.0	7
10	Storm	48 - 55	9.0	8
11	Violent storm	56 - 63	11.5	8
12	Hurricane	>63	>14	9

3.1.7 Wind and sea direction

The wind and sea direction parameters are defined as values in a coordinate system and range from 1 to 8. The values are relative to the vessels cruising direction. Figure 24 shows the different values relative to the direction of the vessel.

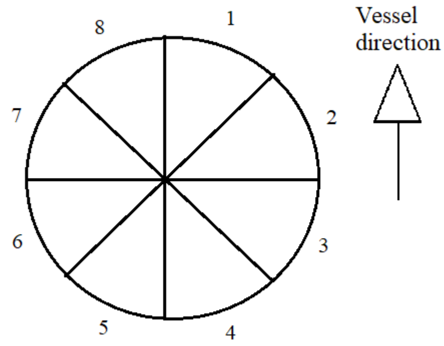


Figure 24: Wind and sea directions

3.2 Data preparation

As mentioned in Section 2.1, clean data are important when preparing a dataset for training. If there are values which are either too big, too small or physically impossible, they should be removed before the training algorithms are running. If the goal is to develop a data-driven model based on a dataset, it is important to maximise the potential of the data. If the goal is a well-performing and accurate model, handling the data correctly and structured is a deciding factor. The process of data cleaning involves preparing a dataset for analysis by filtering and removing data which is irrelevant to the model and/or has a negative impact on the results.

Data cleaning as a process can sometimes be divided into three stages: screening, diagnosis and editing[13]. The screening phase involves removing four types of abnormalities: lack of data, outliers, strange patterns and unexpected analysis results. The goal is to minimise the noise while also maximising the signal in the dataset by identifying and removing errors. The diagnostic phase has a purpose of diagnose each data point for incorrectness, missing data, true extreme, true normal or idiopathic (no explanation found, but still suspect). It is also important to be aware of data which are either biological or physically impossible. After the diagnostic phase is done, editing the remaining data should be done. Each abnormal observation should be either corrected, removed or remained unchanged. Impossible values should always either be corrected if a solution is found, or removed.

4 Results

This section will cover the results of the training and test dataset. The algorithms used for training are compared, and plots showing the result of the best-fitted model are illustrated.

4.1 Training results

After the algorithms presented in Section 2.3.4 were trained, a list of all the results was created. This list contains important data about each model and is used when determining the best model for predicting Fuel oil per day (FODay). An overview of the top 10 algorithms sorted by RMSE-value is shown in Table 5.

Table 5: Result of the 10 best algorithms sorted by RMSE

Algorithm name	**RMSE**	**R-Squared**	**Training time [s]**
Matern 5/2 GPR	**0.95256**	0.98	55.04
Cubic SVM	**0.96515**	0.98	7.36
Trilayered Neural Network	**1.0054**	0.98	83.41
Wide Neural Network	**1.085**	0.98	73.51
Bilayered Neural Network	**1.0924**	0.97	75.43
Quadratic SVM	**1.1471**	0.97	4.14
Rational Quadratic GPR	**1.1758**	0.97	122.77
Squared Exponential GPR	**1.3267**	0.96	61.09
Medium Neural Network	**1.3433**	0.96	65.83
Exponential GPR	**1.3623**	0.96	48.12

The results are relatively close to each other with only about 0.4 difference in RMSE value. Of all algorithms, the Matern 5/2 GPR seems to give the best result. This is also the case in the R-squared value and all values are also substantially close to each other with only 0.02 in difference from the highest to lowest value.

That being said, interesting information lies in the training time. The results show a big difference in the time used for training. The Rational Quadratic GPR have the longest training time of 122.77 seconds which is almost 30 times longer than the Quadratic SVM. Whether a good model is dependent on efficiency, indicated by a low training time, is debatable, but this thesis has evaluated the best models purely by the Root Mean Squared Error. If the dataset contains a huge amount of data, a more efficient model would probably be preferred.

The result showing how the models perform when exposed to unseen data from vessel 5, is presented in Table 6. The rankings differ from the previous table, but the Matern 5/2

GPR still stays on top. When comparing the RMSE from the previous table to Table 6, the values are lower.

Table 6: Result of the 10 best algorithms for test data, sorted by RMSE

Algorithm name	**RMSE (test data)**	**R-Squared (test data)**
Matern 5/2 GPR	**0.28486**	0.98
Rational Quadratic GPR	**0.35728**	0.97
Bilayered Neural Network	**0.42382**	0.96
Trilayered Neural Network	**0.44178**	0.96
Cubic SVM	**0.50402**	0.95
Wide Neural Network	**0.52356**	0.94
Quadratic SVM	**0.62718**	0.92
Squared Exponential GPR	**0.69372**	0.90
Exponential GPR	**0.69601**	0.90
Medium Neural Network	**0.77239**	0.87

4.2 Model

The Matern 5/2 GPR has the best result in terms of fit. A common method to show the fit of a model is a *Predicted response versus the True response* plot. A plot showing the correlation for the Matern 5/2 GPR algorithm is shown in Figure 25. Since the purpose of the model is to predict fuel consumption per day, this parameter is going to be defined and referred to as the *response*.

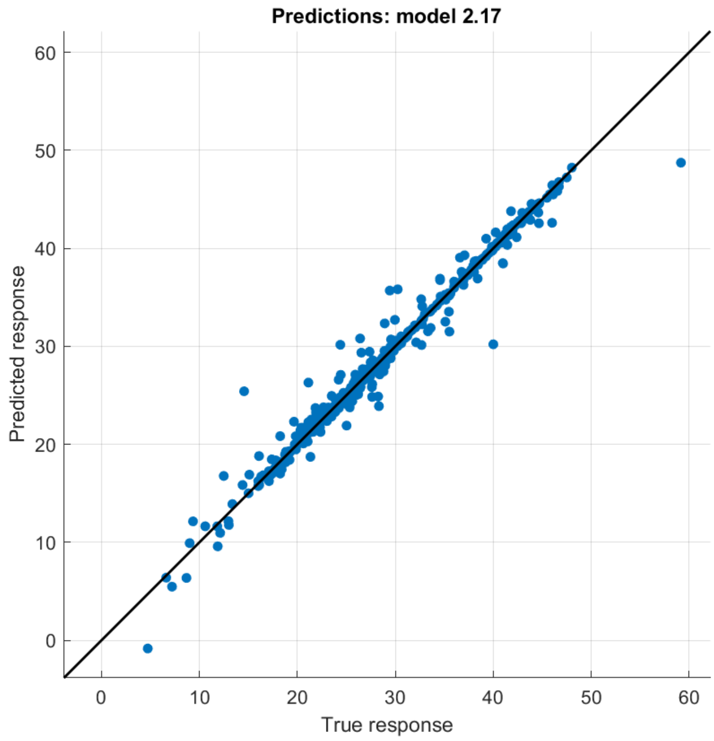


Figure 25: Predicted versus true fuel oil consumption per day, with Matern 5/2 GPR Model

The straight line in the plot represents the optimum value where $x = y$, in other words where the predicted is equal to the true response. Of all predicted plots, the highest error value is 10.85 which is an error of 74%. Otherwise, the model has a generally good fit.

Figure 26 shows the residuals (errors) for the predicted and true responses separately. Residuals are defined as Residual=Observed–Predicted for a specific response value. This figure gives a better visualisation of the error values. A negative value means a prediction value higher than the true value and vice versa. FODay is defined as fuel oil per day.

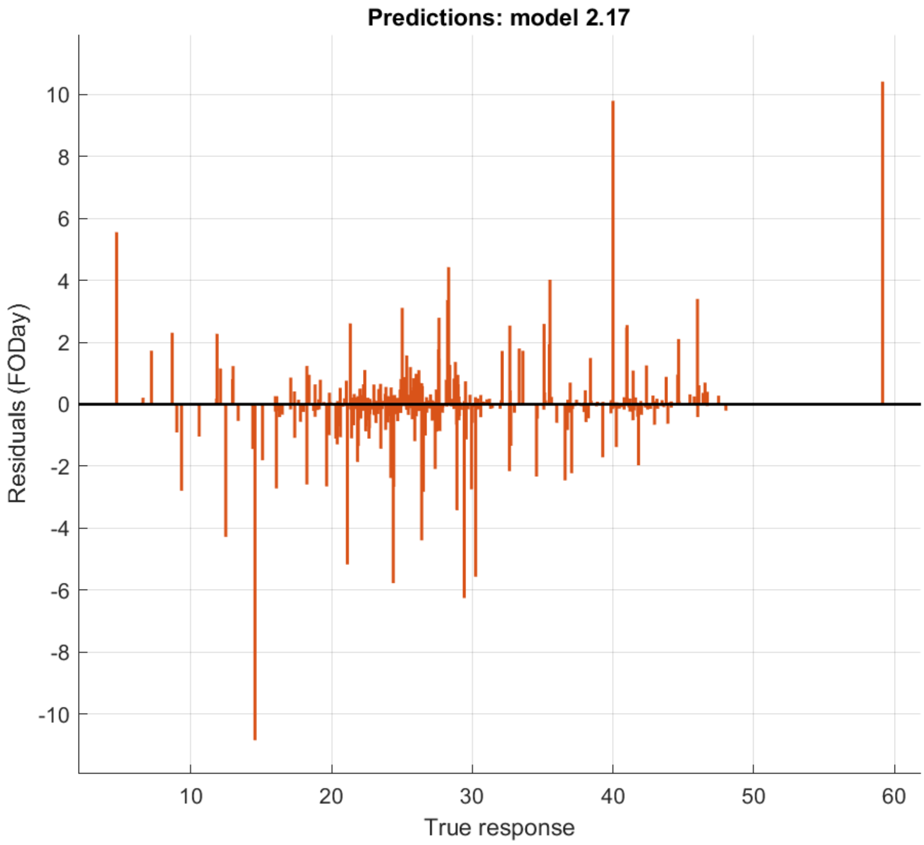


Figure 26: Validation residuals for true response

When testing the model with data from vessel 5, the Matern 5/2 GPR performs well. The errors are almost negligible and lower than the training dataset. This is rather abnormal as test errors often are close to the training error. This will further be discussed in Section 5.

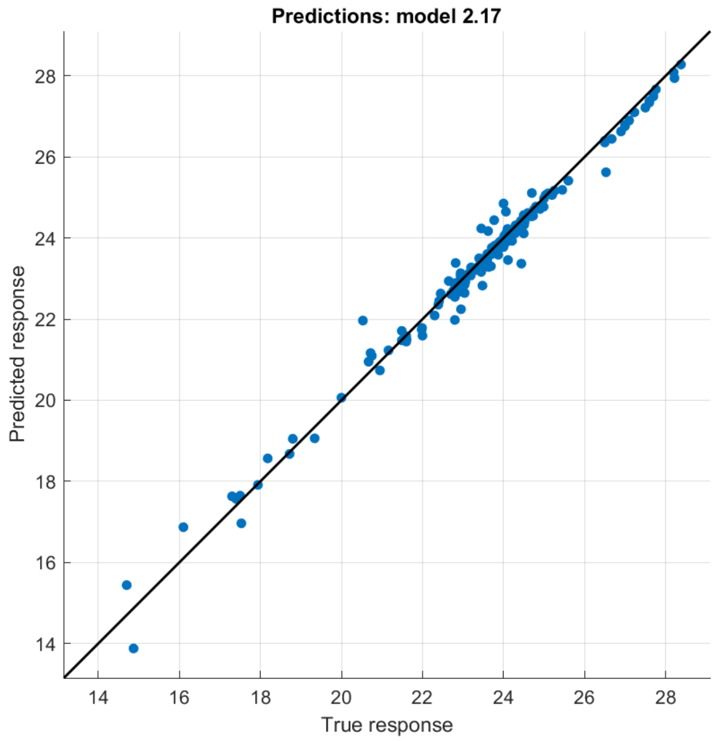


Figure 27: Predicted versus true fuel oil consumption per day, with Matern 5/2 GPR with test data

Figure 27 shows the plot of the predicted response versus the true response of the test data. Figure 28 is the error for each true response.

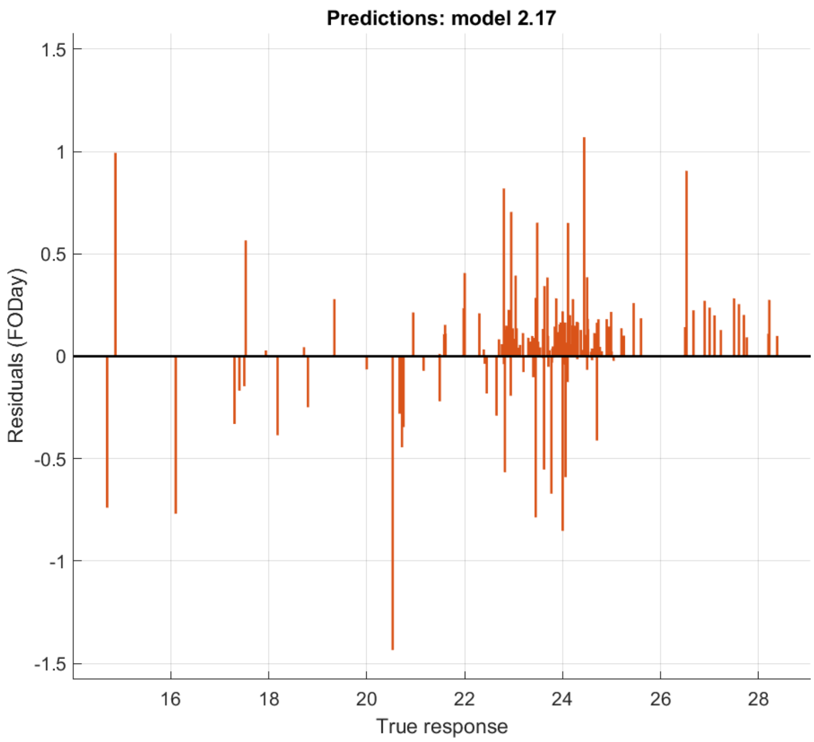


Figure 28: Validation residuals for true response, for the test set

5 Discussion

When Condition-based Maintenance is performed, live data from sensors are often fed into a model to give any indication if maintenance is needed. The key to hybrid modelling is to utilise the advantages from both Data-driven Modelling and Physics-based Modelling. The model developed in this thesis has only been using a data-driven machine learning modelling approach. One of the advantages of DdM is that knowledge about the underlying process is not needed. The model relies purely on data from measurements. This approach is only viable if the data used is clean and at an acceptable quality level. Good knowledge and understanding of the data are therefore required to separate relevant from bad data.

In the model, abnormal data are used as an indication when predicting potential failures. The purpose of the model is not to use data from failures in the machine learning training phase, but rather to create a standard of how the machinery behaves under normal conditions. When a correlation is found between the predictors and the response, in this case fuel consumption per day, it is possible to define a safety boundary where data within the boundary are considered normal. If data from another vessel are inserted into the model and shows several abnormal errors *outside* the boundary, it could indicate unusual power output which would most likely mean a fault in the machinery. Figure 29 is an example of a boundary error set to a value of ~ 14.

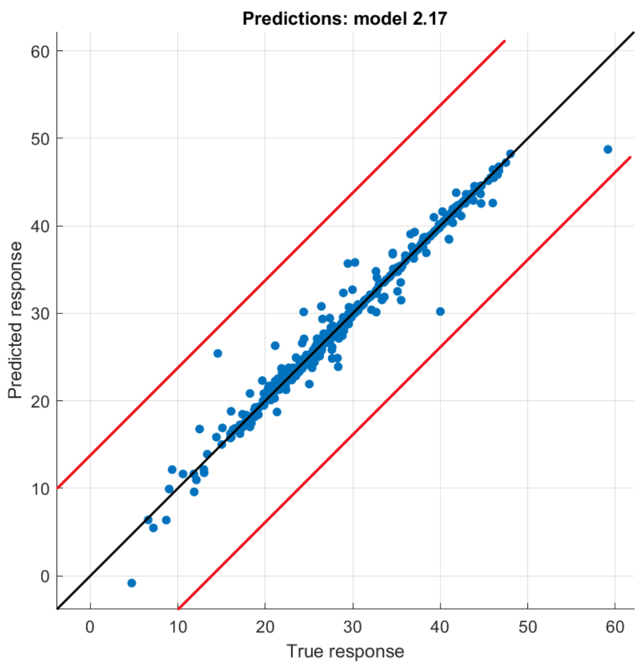


Figure 29: Figure showing an example of boundaries for the model with boundary error of ~ 14

The results of the model presented in Section 4 show a very good fit and a low RMSE which often defines the quality of a model. As mentioned, training time did not affect the development phase in this case since the dataset was not at a very large scale. This is something which should be considered if operating with more data points. Generally, the Support Vector Machine algorithms were superior in training efficiency with only short of six seconds on average. GPR and Neural Networks on the other hand were not as efficient, with ~ 72 and ~ 74 seconds respectively in average training time.

Even though the RMSE score is within acceptable values, the score of the test dataset for the best model is substantially lower at ~ 0.285 compared to ~ 0.953, which is also the case for all other algorithms. This looks like a good result but means that the model is better at predicting unknown data rather than known data. In a good model, the training error is often slightly lower than the test error. A lower test error relative to the training error can also be caused by an unreasonably high degree of regularisation. To resolve this problem, reducing the number of free variables is a potential solution. Another reason why the test error could be lower is the lack of hard-to-predict points in the test dataset, which means a test set that only contains easy cases. This can be solved by collecting more data from more vessels. Cross-validation (K-fold) was done on the training dataset but did not include data from the test dataset (vessel 5). If the K-fold method were performed on a dataset containing data from all vessels with a random portion being the test dataset, this result could probably have been avoided. Changing the model hyperparameters to better fit the data is also a potential solution.

The model is primarily fitted for cargo vessels. Vessels of other dimensions and different machinery such as cruise ships will naturally react differently in different weather conditions. Making a model which fits every vessel is almost impossible. Several models should therefore be developed if the goal is to predict failures on different vessel types.

Since the feature parameter of this model is fuel oil per day, the results can also be used for other applications. Reducing CO_2 emissions within the shipping industry is an ongoing topic, and as new rules and regulations are implemented, the race toward a greener maritime industry is currently in progress. A model which can predict fuel oil consumption can e.g. be used to find the most optimal sailing path concerning emissions. The operator of the model can use weather conditions as input parameters to choose between different routes to minimise resistance and also save time.

5.1 Industry view on digital twins

Condition monitoring and digital twin implementation are gradually getting more popular and are used by several companies within the maritime industry today. Since the *digital twin* term often is used in different settings, the understanding and definition of this new technology vary. This chapter discusses the view of digital twins from an industry point of view and is based on interviews.

ABS (American Bureau of Shipping) is a class society which focuses primarily on the safety of people, property and the environment within the maritime industry. According to them the use of digital twins mostly depend on what you use it for and how it is used. How digital twins are used is very dependent on the targeted outcome and therefore has become a buzzword because there is no clear definition of what a digital twin is. A twin does not need to be a typical finite element digital model, but can also be used in e.g. data filing or documents because almost everything can have a digital copy of itself. ABS has a goal to implement some kind of digital twin on all vessels within their fleet in the future.

One of the main benefits of transferring from calendar-based maintenance (typically preventive maintenance) to condition-based maintenance is the time saved. This is because focusing on specific and critical parts rather than having to pay attention to the whole vessel becomes possible. The time saved can instead be used on maintenance planning instead of a sudden shutdown. Unplanned maintenance is often expensive and time-consuming. While there are many benefits to digital twin implementation, it is also important to be aware of the potential challenges. Installing sensors for real-time data streaming is not cheap, and according to ABS, minimising the number of sensors without losing accuracy is a deciding factor. It is important to not overdo it by installing an unnecessary amount. Digital twins are often based on machine learning models, and indications of problems are not always clearly discovered. Therefore, the reliance on machine learning models should be tempered with empirical knowledge of experienced engineers. Having an over-reliance on a technology still under development can potentially have huge consequences. Moving in a step-by-step manner and understanding the consequences involved when going into the digital age are important.

Gard is a marine insurance company which provides liability, property, and income insurance to shipowners and operators. They have a rather conservative view of digital twin technology compared to e.g. class societies. The cost of insurance for ship owners does not depend on whether condition monitoring or a digital twin is already implemented, but rather the performance of the ship. This is because having a digital twin does not in itself guarantee good performance since the models vary and often need adjustments and care. According to Gard, a clear drawback of digital twins is the lack of shared standards between companies. The result is different definitions and understandings of the term which makes it difficult for insurance companies to operate. A benefit of remotely monitoring the condition of vessels from shore is the reduced risk as a result of having less crew onboard. If data are available for more people to analyse, the increased expertise in ship operation and maintenance will most likely benefit most companies.

6 Conclusion

Knowing the condition of the machinery and the vessel in general becomes important to avoid unwanted situations like unexpected expenses and downtime becomes important. Condition monitoring and digital twins might be the future of maritime operations and could likely be the standard on all vessels in a few decades. The problem does not lie in the technology itself, but rather in how we use it. The technological potential is substantial, but it is very important to understand how this technology work to be able to enjoy benefits and extract useful information.

Developing models to predict a ship's behaviour is highly beneficial, but can also be quite challenging. This thesis had the aim of developing a model for predicting potential failures in the machinery based on anomalies in fuel consumption per day. A dataset containing five somewhat different vessels owned by the shipping company Gearbulk was used in the model development process. The data measurements ranged from January 2021 to May 2022 and contained weather conditions from all around the world, which prevented the model from being too specific and less versatile.

When choosing the optimal modelling strategy, data-driven modelling was chosen since the available dataset contained useful and relevant parameters for creating a model. Since the objective of the model was to *predict* data, a supervised machine learning strategy was preferable compared to an unsupervised approach. Constructing a clean dataset was somewhat difficult as not all data points included every parameter needed in the model. When deciding which machine learning algorithm to use, a regression analysis of the data was done for over 20 different algorithms to find the best-fitted one. The Matern 5/2 GPR model proved to have the best fit in both the training and test dataset.

The model showed a lower training accuracy compared to the test accuracy, which means that the model needs further tweaks and adjustments before it can be fully used for fault detection purposes. The origin of the error most likely comes from the test data being another vessel, rather than training the model on data from all vessels. Although the model did not produce the result desired, it still shows how data extracted from vessels can be used to find a correlation between environmental conditions, fuel oil consumption and machinery data.

The world is currently experiencing a digital revolution and technologies such as digital twins and condition monitoring becomes increasingly popular within various industries. The maritime industry has a lot to learn from aviation. However, a lot of research and ongoing projects within the maritime community are currently in process. One of the main problems with the implementation of digital twin technology in the industry is the lack of a common understanding of what a digital twin is and a consensus of standards. Standardised quality assurance of digital twins is a step in the right direction if the industries want to fully take advantage of the technology.

6.1 Contribution and further work

Even though the model is not based on hybrid modelling, it can be used as a basis for further development toward a digital twin. As discussed in Section 5 the model showed a clear correlation between ship parameters and fuel oil consumption. If the model is adjusted as suggested, the result would most likely improve and the model could be used as a fault detection tool. To develop a more accurate model, more data gathered over a longer period should be included. More research towards feature selection would also probably result in a better model fit as some predictors could be unnecessary and create more harm than good

As mentioned in Section 5 a model predicting fuel oil consumption per day can also be used for purposes other than fault detection. Fuel oil consumption has a direct correlation to the amount of emissions, and can therefore be used to choose a more environmentally friendly sailing route based on environmental condition data.

References

[1] Muhammad Abid et al. 'Reliability Centered Maintenance Plan for the Utility Section of a Fertilizer Industry: A Case Study'. In: *International Journal of Science and Advanced Technology* 4 (Apr. 2014), pp. 2221–8386.

[2] S Agatonovic-Kustrin and R Beresford. 'Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research'. In: *Journal of Pharmaceutical and Biomedical Analysis* 22.5 (2000), pp. 717–727. ISSN: 0731-7085. DOI: https://doi.org/10.1016/S0731-7085(99)00272-1. URL: https://www.sciencedirect.com/science/article/pii/S0731708599002721.

[3] Berke Akkaya and Nurdan Çolakoğlu. 'Comparison of Multi-class Classification Algorithms on Early Diagnosis of Heart Diseases'. In: Sept. 2019.

[4] Mahamad Alam. 'Codes in MATLAB for Training Artificial Neural Network using Particle Swarm Optimization'. In: (Aug. 2016). DOI: 10.13140/RG.2.1.2579.3524.

[5] Timothy M. Allen. 'U . S . Navy Analysis of Submarine Maintenance Data and the Development of Age and Reliability Profiles ABSTRACT'. In: (2005).

[6] Dawn An, Joo Choi and Nam Kim. 'Options for Prognostics Methods: A review of data-driven and physics-based prognostics'. In: *Collection of Technical Papers - AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference* (Apr. 2013). DOI: 10.2514/6.2013-1940.

[7] Anish. *Why 2-stroke Engines are Used More commonly than 4-stroke on Ships?* URL: https://www.marineinsight.com/main-engine/why-2-stroke-engines-are-used-more-commonly-than-4-stroke-on-ships/ (visited on 10/06/2022).

[8] Maggie Mae Armstrong. *Cheat sheet: What is Digital Twin?* URL: https://www.ibm.com/blogs/internet-of-things/iot-cheat-sheet-digital-twin/ (visited on 14/12/2021).

[9] Debasish Basak, Srimanta Pal and Dipak Patranabis. 'Support Vector Regression'. In: *Neural Information Processing – Letters and Reviews* 11 (Nov. 2007).

[10] Anriëtte Bekker, Ruediger Ulrich Franz Von Bock und Polach and Amir R. Nejad. 'HEALTHPROP - Life prediction and health monitoring of marine propulsion system under ice impact'. In: (). URL: https://www.researchgate.net/project/HEALTHPROP-Life-prediction-and-health-monitoring-of-marine-propulsion-system-under-ice-impact (visited on 15/12/2021).

[11] Tadeusz Borkowski, Lech Kasyk and Przemyslaw Kowalak. 'Assessment of ship's engine effective power, fuel consumption and emission using the vessel speed'. In: *Journal of KONES* 18 (Jan. 2011).

[12] Kristina Brandänge and Petter Gustavsson. 'After the MV Estonia ferry disaster A Swedish nationwide survey of the relatives of the MV Estonia victims'. In: *Dialogues in clinical neuroscience* 2 (Mar. 2000), pp. 63–9.

[13] Jan Broeck et al. 'Data Cleaning: Detecting, Diagnosing, and Editing Data Abnormalities'. In: *PLoS medicine* 2 (Nov. 2005), e267. DOI: 10.1371/journal.pmed.0020267.

[14] Jason Brownlee. *A Gentle Introduction to k-fold Cross-Validation*. URL: https://machinelearningmastery.com/k-fold-cross-validation/ (visited on 12/06/2022).

[15] c3.ai. *Root Mean Square Error (RMSE)*. URL: https://c3.ai/glossary/data-science/root-mean-square-error-rmse/#:~:text=Why%5C%20is%5C%20Root%5C%20Mean%5C%20Square,widely%5C%20used%5C%20measures%5C%20for%5C%20this. (visited on 12/06/2022).

[16] Alexandre CAPONE. *What if Physics-based modeling/simulation is replace by Ai prediction model?* URL: https://www.linkedin.com/pulse/what-physics-based-modelingsimulation-replace-ai-model-capone (visited on 12/12/2021).

[17] Norwegian Directorate for Civil Protection. *Assessment of the Viking Sky Incident*. URL: https://www.dsb.no/globalassets/dokumenter/rapporter/assessment_of_the_viking_sky_incident.pdf (visited on 31/05/2022).

[18] Keppel Corporation. *Keppel to develop its first autonomous vessel for operations in Singapore*. URL: https://www.kepcorp.com/en/media/media-releases-sgx-filings/keppel-to-develop-its-first-autonomous-vessel-for-operations-in-singapore-6564/ (visited on 14/12/2021).

[19] C Cortes and V Vapnik. 'Support-vector networks'. In: *Mach Learn* 20 (1995), pp. 273–297. DOI: https://doi.org/10.1007/BF00994018.

[20] Marc Cousineau. *How to make condition-based maintenance more effective*. URL: https://www.fiixsoftware.com/blog/effective-condition-based-maintenance/ (visited on 04/12/2021).

[21] Adrian Cubillo, Suresh Perinpanayagam and Manuel Esperon Miguez. 'A review of physics-based models in prognostics: Application to gears and bearings of rotating machinery'. In: *Advances in Mechanical Engineering* 8 (Aug. 2016). DOI: 10.1177/1687814016664660.

[22] Rodrigo De Oliveira et al. 'A System Based on Artificial Neural Networks for Automatic Classification of Hydro-generator Stator Windings Partial Discharges'. In: *Journal of Microwaves, Optoelectronics and Electromagnetic Applications* 16 (Sept. 2017), pp. 628–645. DOI: 10.1590/2179-10742017v16i3854.

[23] Deepanshi. *Beginners Guide to Artificial Neural Network*. URL: https://www.analyticsvidhya.com/blog/2021/05/beginners-guide-to-artificial-neural-network/ (visited on 16/06/2022).

[24] Mathieu Desbazeille et al. 'Model-based diagnosis of large diesel engines based on angular speed variations of the crankshaft'. In: *Mechanical Systems and Signal Processing* 24 (July 2010), pp. 1529–1541. DOI: 10.1016/j.ymssp.2009.12.004.

[25] Kongsberg Digital. *BEYOND THE VISUALS*. URL: https://www.kongsberg.com/digital/solutions/kognitwin-energy/beyond-the-visual/ (visited on 12/12/2021).

[26] Marine Digital. *5 Benefits Of Digital Twin In The Shipping Industry*. URL: https://marine-digital.com/article_5benefits_of_digital_twin (visited on 14/12/2021).

[27] Jeffrey Dutschke. *Run to failure: make it part of your maintenance planning*. URL: https://www.fiixsoftware.com/blog/run-failure-make-part-maintenance-planning/ (visited on 04/12/2021).

[28] IBM Cloud Education. *Overfitting*. URL: https://www.ibm.com/cloud/learn/overfitting (visited on 12/06/2022).

[29] IBM Cloud Education. *Underfitting*. URL: https://www.ibm.com/cloud/learn/underfitting%5C#%5C:%5C~%5C:text%5C=Underfitting%5C%20is%5C%20a%5C%20scenario%5C%20in,training%5C%20set%5C%20and%5C%20unseen%5C%20data. (visited on 12/06/2022).

[30] Equinor. *The future of offshore wind is afloat*. URL: https://www.equinor.com/en/what-we-do/floating-wind.html (visited on 15/12/2021).

[31] Hickey law firm. *Low Engine Oil to Blame for Viking Sky Ordeal*. URL: https://www.hickeylawfirm.com/blog/low-engine-oil-to-blame-for-viking-sky-ordeal (visited on 01/06/2022).

[32] Esperanza García-Gonzalo et al. 'Hard-Rock Stability Analysis for Span Design in Entry-Type Excavations with Learning Classifiers'. In: *Materials* 9 (June 2016), p. 531. DOI: 10.3390/ma9070531.

[33] Gearbulk. *About Gearbulk*. URL: https://www.linkedin.com/company/gearbulk/about/ (visited on 01/06/2022).

[34] Gearbulk. *About Gearbulk*. URL: https://www.gearbulk.com/about/ (visited on 01/06/2022).

[35] IBM. *Supervised learning*. URL: https://www.ibm.com/cloud/learn/supervised-learning (visited on 02/06/2022).

[36] Michaela Ibrion, Nicola Paltrinieri and Amir R. Nejad. 'On Risk of Digital Twin Implementation in Marine Industry: Learning from Aviation Industry'. In: *Journal of Physics: Conference Series* 1357 (Oct. 2019), p. 012009. DOI: 10.1088/1742-6596/1357/1/012009.

[37] JavaTpoint. *Supervised Machine Learning*. URL: https://www.javatpoint.com/supervised-machine-learning (visited on 02/06/2022).

[38] Javatpoint. *Unsupervised Machine Learning*. URL: https://www.javatpoint.com/unsupervised-machine-learning (visited on 06/06/2022).

[39] jCaptain. *Hurtigruten Kong Harald in trouble*. URL: https://forum.gcaptain.com/t/hurtigruten-kong-harald-in-trouble/60443 (visited on 01/06/2022).

[40] Sigrid Johansen and Amir R. Nejad. 'On Digital Twin Condition Monitoring Approach for Drivetrains in Marine Applications'. In: (June 2019). DOI: 10.1115/OMAE2019-95152.

[41] Ioannis Kakadiaris. 'Physics-Based Modeling, Analysis and Animation'. In: *Technical Reports (CIS)* (Apr. 1993).

[42] Sandali Khare et al. 'Real Estate Cost Estimation Through Data Mining Techniques'. In: *IOP Conference Series: Materials Science and Engineering* 1099 (Mar. 2021), p. 012053. DOI: 10.1088/1757-899X/1099/1/012053.

[43] Knut Erik Knutsen, Gabriele Manno and Bjørn Vartdal. 'Beyond condition monitoring in the maritime industry'. In: (Jan. 2014). DOI: 10.13140/RG.2.1.1807.8562.

[44] Kontakt.io. *Digital Twins in Aviation and 8 Ways They Can Transform Workflows*. URL: https://kontakt.io/blog/digital-twins-in-aviation-and-8-ways-they-can-transform-workflows/ (visited on 27/10/2021).

[45] Max Kuhn and Kjell Johnson. *Applied Predictive Modeling*. Jan. 2013, p. 488. ISBN: 978-1-4614-6848-6. DOI: 10.1007/978-1-4614-6849-3.

[46] Martin Leduc. *The marine Diesel engine, Part one: The two stroke*. URL: https://www.dieselduck.info/machine/01%5C%20prime%5C%20movers/diesel_engine/diesel_engine.01.htm#:~:text=Slow%5C%20speed%5C%20are%5C%20considered%5C%20to,generators%5C%20and%5C%20%5C%2F%5C%20or%5C%20the%5C%20propeller. (visited on 10/06/2022).

[47] Lexico. *Meaning of overfitting*. URL: https://www.lexico.com/definition/overfitting (visited on 10/06/2022).

[48] Maply. *Heatmap Maply*. URL: https://maply.com/ (visited on 09/06/2022).

[49] MetMatters. *The Beaufort Scale*. URL: https://www.rmets.org/metmatters/beaufort-scale (visited on 08/06/2022).

[50] MLMath.io. *Math behind SVM (Support Vector Machine)*. URL: https://ankitnitjsr13.medium.com/math-behind-support-vector-machine-svm-5e7376d0ee4d (visited on 15/06/2022).

[51] MLMath.io. *Support Vector Machine*. URL: https://sgao323.gitbooks.io/artificial-intelligence-projects/content/ml-basics/svm.html (visited on 15/06/2022).

[52] MLMath.io. *Support Vector Machine: Regression*. URL: https://medium.com/it-paragon/support-vector-machine-regression-cf65348b6345 (visited on 15/06/2022).

[53] Amir R. Nejad. *About HealthProp*. URL: https://www.ntnu.edu/imt/healthprop/ (visited on 15/12/2021).

[54] Sai Chandra Nerella. *Linear Regression*. URL: https://saichandra1199.medium.com/linear-regression-1e279814e2bb (visited on 14/06/2022).

[55] Hellenic Shipping News. *Open Simulation Platform – the next generation of digital twins*. URL: https://www.hellenicshippingnews.com/open-simulation-platform-the-next-generation-of-digital-twins/ (visited on 15/12/2021).

[56] Riviera Newsletters. *Digital twins optimise vessel design, maintenance and performance*. URL: https://www.rivieramm.com/opinion/opinion/digital-twins-optimise-vessel-design-maintenance-and-performance-41073 (visited on 14/12/2021).

[57] OpenSimulationPlatform. *OPEN SIMULATION PLATFORM - Towards a maritime ecosystem for efficient co-simulation*. URL: https://opensimulationplatform.com/ (visited on 15/12/2021).

[58] Anthony Perez, Carl Antonio and Rafael Consunji. ‘The Sinking of the MV Doña Paz – I. An analysis of the event’. In: *Acta medica Philippina* 45 (Mar. 2011).

[59] Chris Piech. *K Means*. URL: https://stanford.edu/~cpiech/cs221/handouts/kmeans.html (visited on 06/06/2022).

[60] Naida Hakirevic Prevljak. *Digital twins for the maritime sector*. URL: https://www.offshore-energy.biz/digital-twins-for-the-maritime-sector/ (visited on 14/12/2021).

[61] Robert Bond Randall. ‘Vibration-based Condition Monitoring: Industrial, Aerospace and Automotive Applications’. In: (2010), pp. 229–248.

[62] C. E. Rasmussen and C. K. I. Williams. ‘Gaussian Processes for Machine Learning’. In: (2006), pp. 7–30.

[63] A. L. Samuel. ‘Some studies in machine learning using the game of checkers’. In: *IBM Journal of Research and Development* 44.1.2 (2000), pp. 206–226. DOI: 10.1147/rd.441.0206.

[64] Mike Schuler. *Hurtigruten Coastal Ship Loses Power Off Norway*. URL: https://gcaptain.com/hurtigruten-coastal-ship-loses-power-off-norway/ (visited on 01/06/2022).

[65] International Chamber of Shipping. *Shipping and World Trade: World Seaborne Trade*. URL: https://www.ics-shipping.org/shipping-fact/shipping-and-world-trade-world-seaborne-trade/#:~:text=The%5C%20international%5C%20shipping%5C%20industry%5C%20is,would%5C%20simply%5C%20not%5C%20be%5C%20possible. (visited on 25/06/2022).

[66] Ships$_{B}$*usiness*. *Ship Maintenance Requirement and Energy Efficiency Measures*. URL: http://shipsbusiness.com/pollution-by-overboard-maintenance.html (visited on 04/12/2021).

[67] Marc Shivers. *Influence of C in SVM*. URL: https://stats.stackexchange.com/questions/31066/what-is-the-influence-of-c-in-svms-with-linear-kernel (visited on 15/06/2022).

[68] Sjøfartsdirektoratet. *Press release: Viking Sky*. URL: https://www.sdir.no/en/news/news-from-the-nma/press-release-viking-sky// (visited on 31/05/2022).

[69] Dimitri Solomatine, Linda See and R.J. Abrahart. 'Data-Driven Modelling: Concepts, Approaches and Experiences'. In: vol. 68. Jan. 2008, pp. 17–30. ISBN: 978-3-540-79880-4. DOI: 10.1007/978-3-540-79881-1_2.

[70] Zhenya Song et al. 'Centuries of monthly and 3-hourly global ocean wave data for past, present, and future climate research'. In: *Scientific Data* 7 (July 2020), p. 226. DOI: 10.1038/s41597-020-0566-8.

[71] Safae Sossi Alaoui, Yousef Farhaoui and B. Aksasse. 'Classification algorithms in Data Mining'. In: *International Journal of Tomography and Simulation* 31 (Aug. 2018), pp. 34–44.

[72] Lisa Tagliaferri. *An Introduction to Machine Learning*. URL: https://www.digitalocean.com/community/tutorials/an-introduction-to-machine-learning (visited on 02/06/2022).

[73] Team TheNavalArch. *How to use a ship's hydrostatics to calculate its draft and trim*. URL: https://thenavalarch.com/how-to-use-a-ships-hydrostatics-to-calculate-its-draft-and-trim/ (visited on 13/06/2022).

[74] Theodore Trafalis and Huseyin Ince. 'Support Vector Machine for Regression and Applications to Financial Forecasting.' In: vol. 6. Jan. 2000, pp. 348–353. DOI: 10.1109/IJCNN.2000.859420.

[75] UNCTAD. *Merchant fleet*. URL: https://hbs.unctad.org/merchant-fleet/ (visited on 15/04/2022).

[76] UpKeep. *Predictive Maintenance: What is it and what are the Benefits?* URL: https://www.onupkeep.com/learning/maintenance-types/predictive-maintenance (visited on 04/12/2021).

[77] VesselFinder. *Plover Arrow*. URL: https://www.vesselfinder.com/ship-photos/683625 (visited on 01/06/2022).

[78] Cruise Market Watch. *Growth of the Ocean Cruise Line Industry*. URL: https://cruisemarketwatch.com/growth/ (visited on 15/04/2022).

[79] Jesús Zambrano. 'Gaussian Process Regression - method and applications'. In: (Nov. 2017). DOI: 10.13140/RG.2.2.15890.56002.

[80] Nian Zhang et al. 'Gaussian Process Regression Method for Classification for High-Dimensional Data with Limited Samples'. In: (June 2018), pp. 358–363. DOI: 10.1109/ICIST.2018.8426077.

[81] Daniel Zwick et al. *Technology for a real-time simulation-based system monitoring of wind turbines*. URL: https://www.sintef.no/globalassets/project/eera-deepwind2017/presentasjoner/d2_zwick.pdf (visited on 15/12/2021).

Appendix

A Engine data of case study vessels

Propulsion engine

Function
Code: 411.1
Name: Conventional propulsion line driving
Onboard name:

Component
Code: C101
Name: Reciprocating internal combustion engine[6S50ME-C8, stroke 2000]
Assigned:

Parameters

Cylinders pressure maximum	NOT SET	bar
Rotational speed at reduced power	NOT SET	rpm
NOx reduction device	Not applicable	
Product certificate no.	NOT SET	
Cylinders pressure mean effective (MEP)	NOT SET	bar
Engine type	Two stroke	
Serial number	6308	
Fuel type	Heavy fuel oil; Marine diesel oil	
Manufacturer	MAN Energy Solutions, branch of MAN Energy Solutions SE, Germany	
Power, reduced speed	NOT SET	kW
Cylinder configuration	In-line engine	
Type approval certificate no.	TAM00000SV	
Crankshaft torsional stress in way of crankpin, maximum allowable	NOT SET	
Rotational speed at maximum continuous output power	108	rpm
Power, maximum continuous output	7730	kW
Designer	MAN Energy Solutions, branch of MAN Energy Solutions SE, Germany	

Figure 30: Avocet Arrow engine data

Parameters

Product certificate no.	KOB-08-767
Cylinders pressure mean effective (MEP)	NOT SET bar
Fuel type	HFO; MDO
Rotational speed at reduced power	NOT SET rpm
Manufacturer	Kawasaki Heavy Industries, Ltd. Kobe Works
Power, maximum continuous output	12577 kW
Power, reduced speed	NOT SET kW
Type approval certificate no.	NOT SET
Designer	Kawasaki Heavy Industries
Cylinder configuration	In-line engine
Crankshaft torsional stress in way of crankpin, maximum allowable	NOT SET
Rotational speed at maximum continuous output power	93 rpm
Cylinders pressure maximum	160 bar
Engine type	Two stroke
Serial number	8211
NOx reduction device	Not applicable

Figure 31: Corella Arrow engine data

Function

Code: 411.1
Name: Conventional propulsion line driving
Onboard name:

Component

Code: C101
Name: Reciprocating internal combustion engine[6L60MC]
Assigned:

Parameters

Cylinders pressure mean effective (MEP)	NOT SET bar
Rotational speed at maximum continuous output power	123 rpm
Engine type	Two stroke
Type approval certificate no.	No
Rotational speed at reduced power	NOT SET rpm
Crankshaft torsional stress in way of crankpin, maximum allowable	NOT SET
Product certificate no.	NOT SET
Fuel type	HFO
Power, reduced speed	NOT SET kW
Manufacturer	Dalian Marine Diesel Co., Ltd.
Designer	B&W
Serial number	NOT SET
Power, maximum continuous output	11520 kW
NOx reduction device	Not applicable
Cylinder configuration	In-line engine
Cylinders pressure maximum	NOT SET bar

Figure 32: Plover Arrow engine data

Propulsion engine

Function
Code: 411.1
Name: Conventional propulsion line driving
Onboard name:

Component
Code: C101
Name: Reciprocating internal combustion engine[5S60MC]
Assigned:

Parameters

Cylinders pressure maximum	NOT SET	bar
NOx reduction device	Not applicable	
Fuel type	NOT SET	
Rotational speed at reduced power	NOT SET	rpm
Power, reduced speed	NOT SET	kW
Designer	B&W	
Cylinder configuration	In-line engine	
Product certificate no.	NOT SET	
Serial number	NOT SET	
Crankshaft torsional stress in way of crankpin, maximum allowable	NOT SET	
Cylinders pressure mean effective (MEP)	NOT SET	bar
Type approval certificate no.	NOT SET	
Engine type	Two stroke	
Power, maximum continuous output	9378	kW
Manufacturer	NOT SET	
Rotational speed at maximum continuous output power	102	rpm

Figure 33: Swift Arrow engine data

Propulsion engine

Function
Code: 411.1
Name: Conventional propulsion line driving
Onboard name:

Component
Code: C101
Name: Reciprocating internal combustion engine[6L60MC]
Assigned:

Parameters

Rotational speed at maximum continuous output power	123	rpm
Engine type	Two stroke	
NOx reduction device	Not applicable	
Manufacturer	Dalian Marine Diesel Co., Ltd.	
Cylinder configuration	In-line engine	
Power, reduced speed	NOT SET	kW
Crankshaft torsional stress in way of crankpin, maximum allowable	NOT SET	
Fuel type	HFO	
Type approval certificate no.	No	
Cylinders pressure maximum	NOT SET	bar
Designer	B&W	
Serial number	NOT SET	
Cylinders pressure mean effective (MEP)	NOT SET	bar
Product certificate no.	NOT SET	
Rotational speed at reduced power	NOT SET	rpm
Power, maximum continuous output	11520	kW

Figure 34: Weaver Arrow engine data

B Plot of significant wave heights from 1979-2014

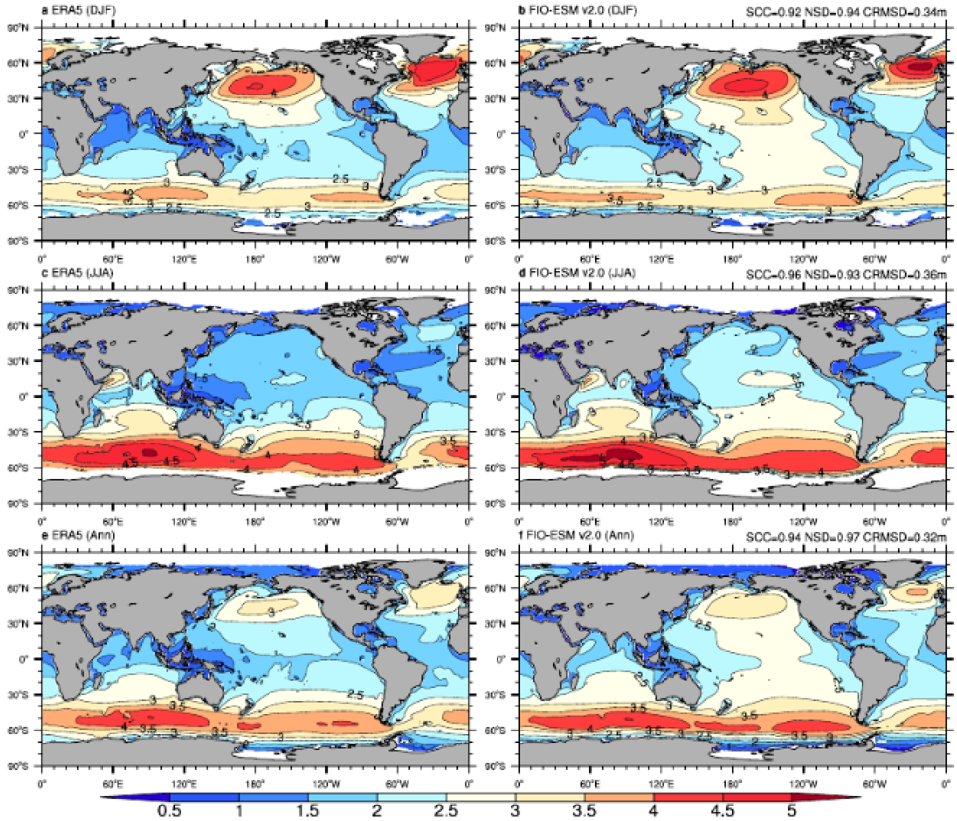


Fig. 3 Climatological distributions of the significant wave height from monthly mean data of ERA5 (left column) and FIO-ESM v2.0 (right column). (**a–f**) are boreal winter (December-January-February), boreal summer (June-July-August), and annual mean results, respectively. The averaged period is from 1979 to 2014. SCC, NSD, and CRMSD represent the spatial correlation coefficient, the normalized standard deviation, and the centered-root-mean-square difference, respectively.

Figure 35: Historical data of the average significant wave heights from year 1979-2014 [70]

C Training results of 2nd to 10th best model

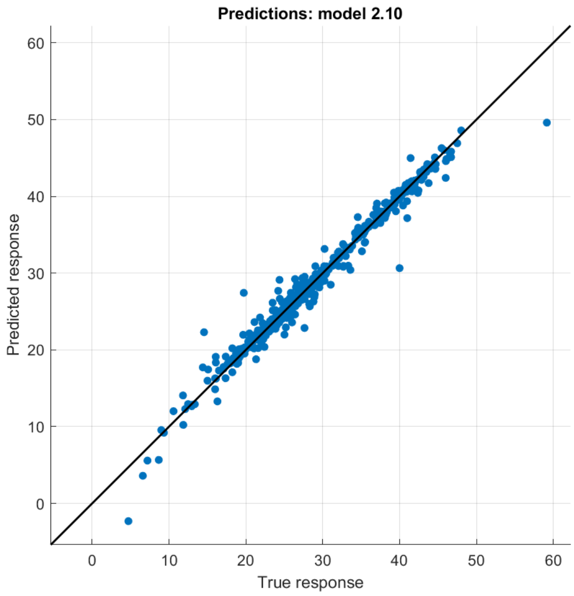
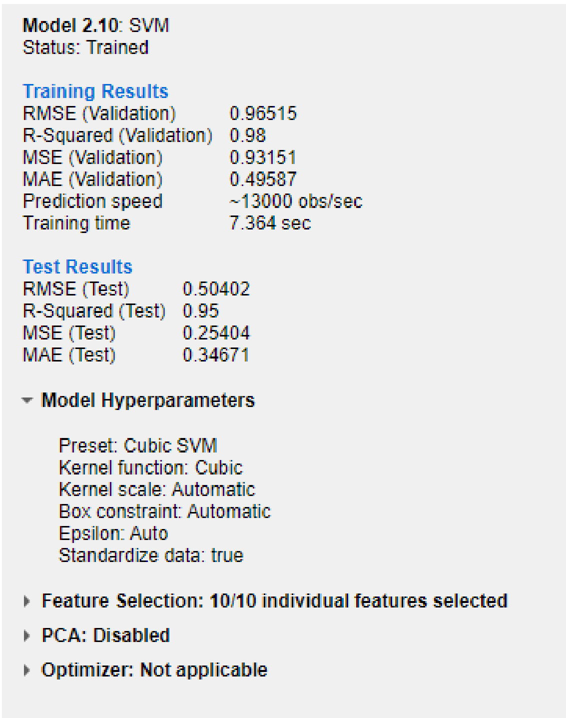


Figure 36: Predicted versus actual data, Cubic SVM



Model 2.10: SVM
Status: Trained

Training Results

RMSE (Validation)	0.96515
R-Squared (Validation)	0.98
MSE (Validation)	0.93151
MAE (Validation)	0.49587
Prediction speed	~13000 obs/sec
Training time	7.364 sec

Test Results

RMSE (Test)	0.50402
R-Squared (Test)	0.95
MSE (Test)	0.25404
MAE (Test)	0.34671

Model Hyperparameters

Preset: Cubic SVM
Kernel function: Cubic
Kernel scale: Automatic
Box constraint: Automatic
Epsilon: Auto
Standardize data: true

Feature Selection: 10/10 individual features selected

PCA: Disabled

Optimizer: Not applicable

Figure 37: Training data of Cubic SVM

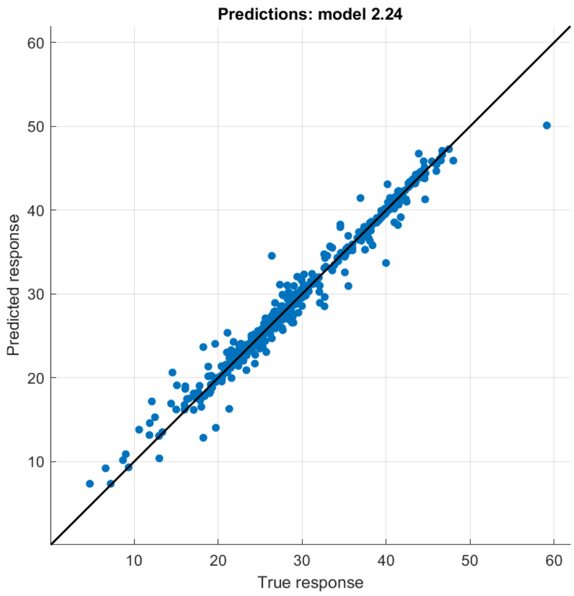
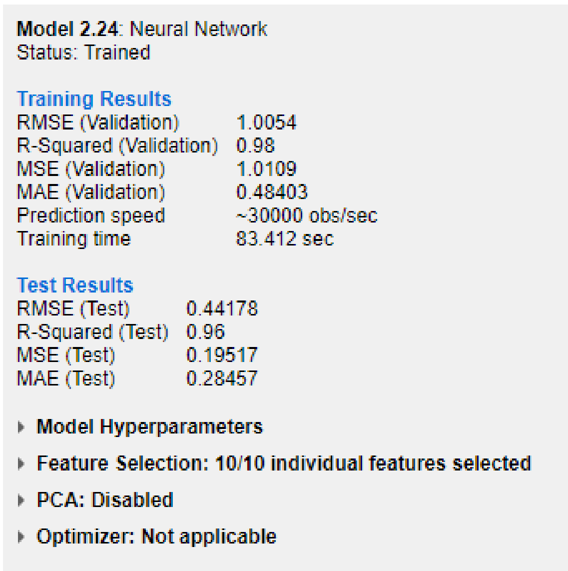


Figure 38: Predicted versus actual data, Trilayered Neural Network



Model 2.24: Neural Network
Status: Trained

Training Results

RMSE (Validation)	1.0054
R-Squared (Validation)	0.98
MSE (Validation)	1.0109
MAE (Validation)	0.48403
Prediction speed	~30000 obs/sec
Training time	83.412 sec

Test Results

RMSE (Test)	0.44178
R-Squared (Test)	0.96
MSE (Test)	0.19517
MAE (Test)	0.28457

- **Model Hyperparameters**
- **Feature Selection: 10/10 individual features selected**
- **PCA: Disabled**
- **Optimizer: Not applicable**

Figure 39: Training data of Trilayered Neural Network

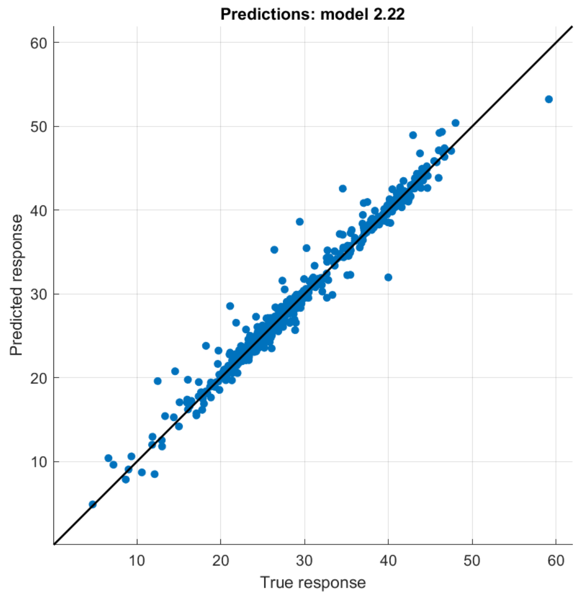


Figure 40: Predicted versus actual data, Wide Neural Network

Model 2.22: Neural Network
Status: Trained

Training Results

RMSE (Validation)	1.085
R-Squared (Validation)	0.98
MSE (Validation)	1.1772
MAE (Validation)	0.51977
Prediction speed	~35000 obs/sec
Training time	73.51 sec

Test Results

RMSE (Test)	0.52356
R-Squared (Test)	0.94
MSE (Test)	0.27412
MAE (Test)	0.26163

Model Hyperparameters

Preset: Wide Neural Network
Number of fully connected layers: 1
First layer size: 100
Activation: ReLU
Iteration limit: 1000
Regularization strength (Lambda): 0
Standardize data: Yes

Feature Selection: 10/10 individual features selected

PCA: Disabled

Optimizer: Not applicable

Figure 41: Training data of Wide Neural Network

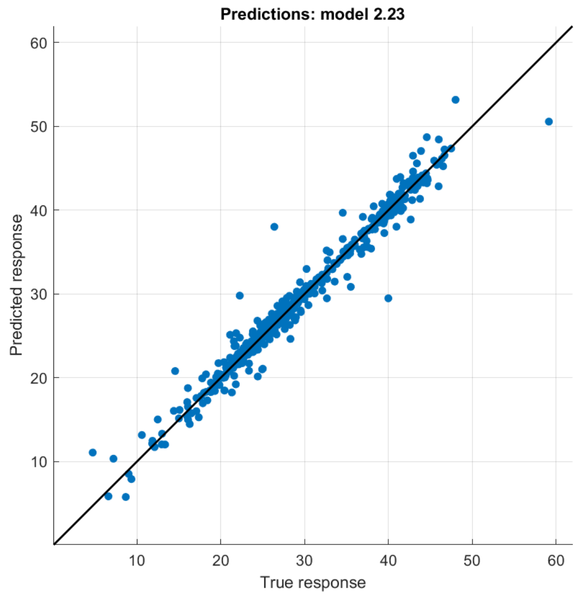


Figure 42: Predicted versus actual data, Bilayered Neural Network

Model 2.23: Neural Network
Status: Trained

Training Results

RMSE (Validation)	1.0924
R-Squared (Validation)	0.97
MSE (Validation)	1.1932
MAE (Validation)	0.52147
Prediction speed	~8300 obs/sec
Training time	75.427 sec

Test Results

RMSE (Test)	0.42382
R-Squared (Test)	0.96
MSE (Test)	0.17962
MAE (Test)	0.25301

Model Hyperparameters

Preset: Bilayered Neural Network
Number of fully connected layers: 2
First layer size: 10
Second layer size: 10
Activation: ReLU
Iteration limit: 1000
Regularization strength (Lambda): 0
Standardize data: Yes

Feature Selection: 10/10 individual features selected

PCA: Disabled

Optimizer: Not applicable

Figure 43: Training data of Bilayered Neural Network

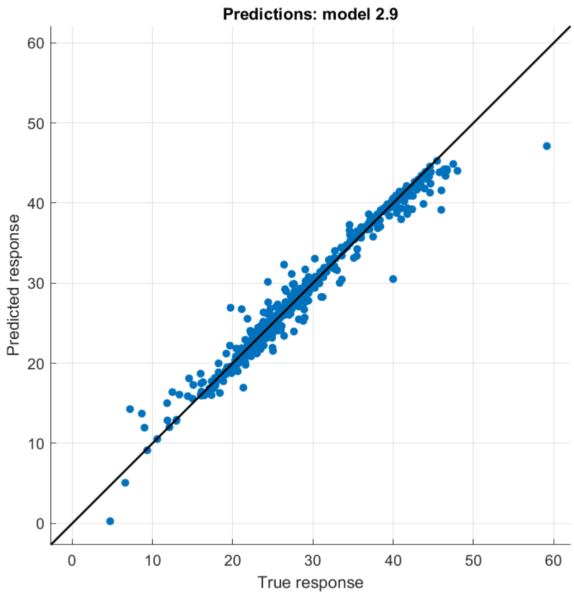


Figure 44: Predicted versus actual data, Quadric SVM

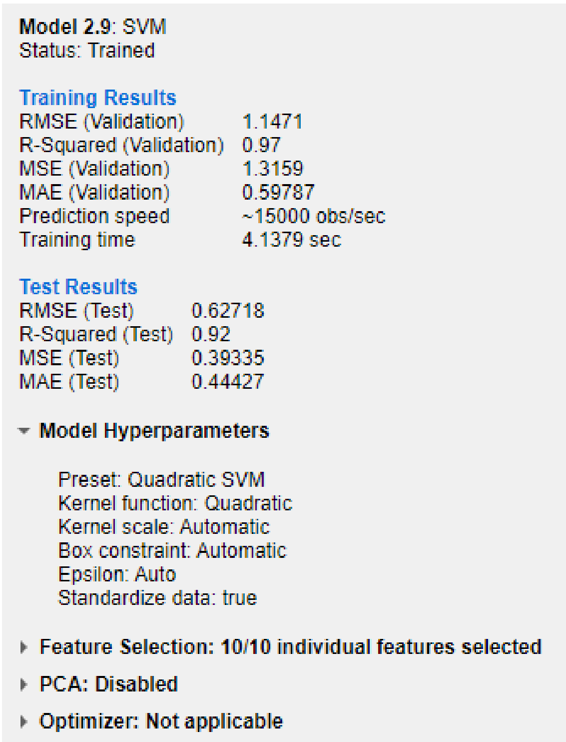


Figure 45: Training data of Quadric SVM

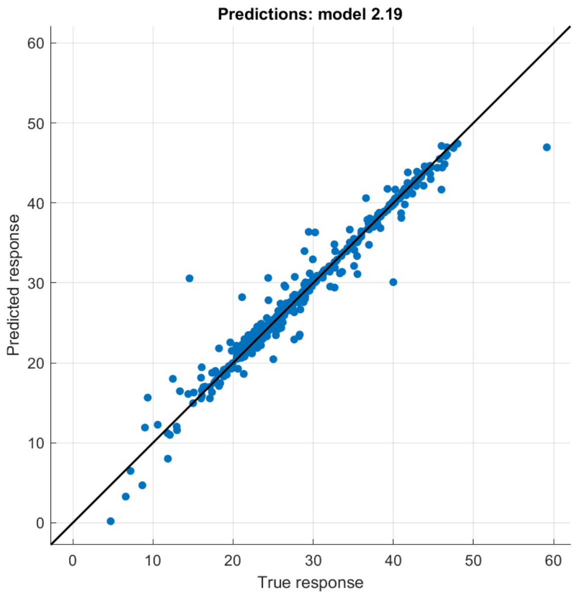
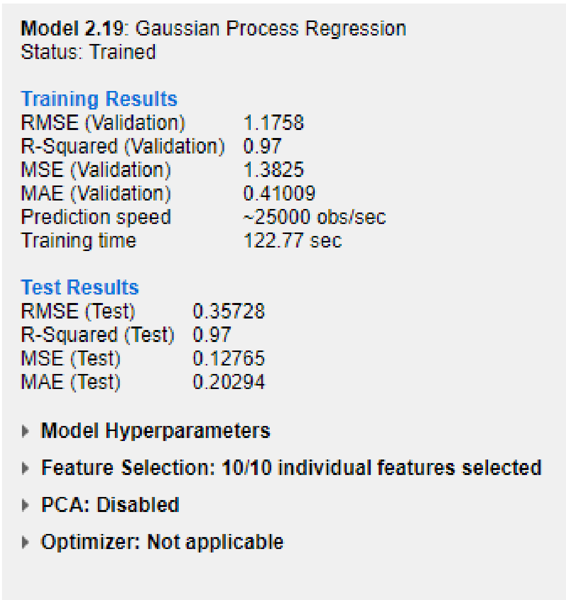


Figure 46: Predicted versus actual data, Rational Quadric GPR



Model 2.19: Gaussian Process Regression
Status: Trained

Training Results

RMSE (Validation)	1.1758
R-Squared (Validation)	0.97
MSE (Validation)	1.3825
MAE (Validation)	0.41009
Prediction speed	~25000 obs/sec
Training time	122.77 sec

Test Results

RMSE (Test)	0.35728
R-Squared (Test)	0.97
MSE (Test)	0.12765
MAE (Test)	0.20294

- **Model Hyperparameters**
- **Feature Selection: 10/10 individual features selected**
- **PCA: Disabled**
- **Optimizer: Not applicable**

Figure 47: Training data of Rational Quadric GPR

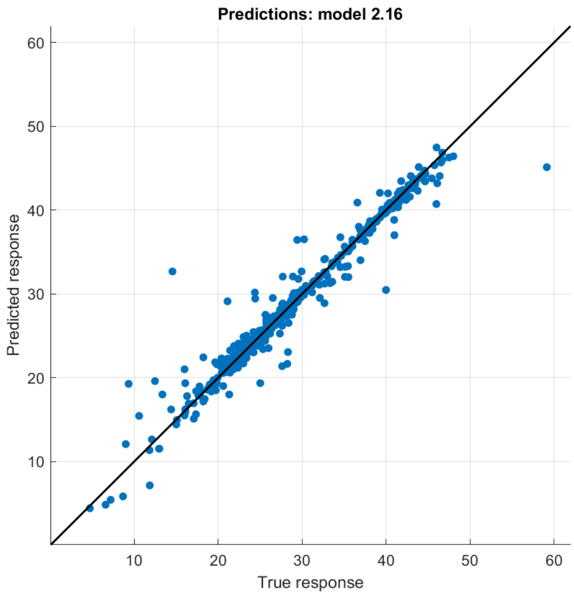
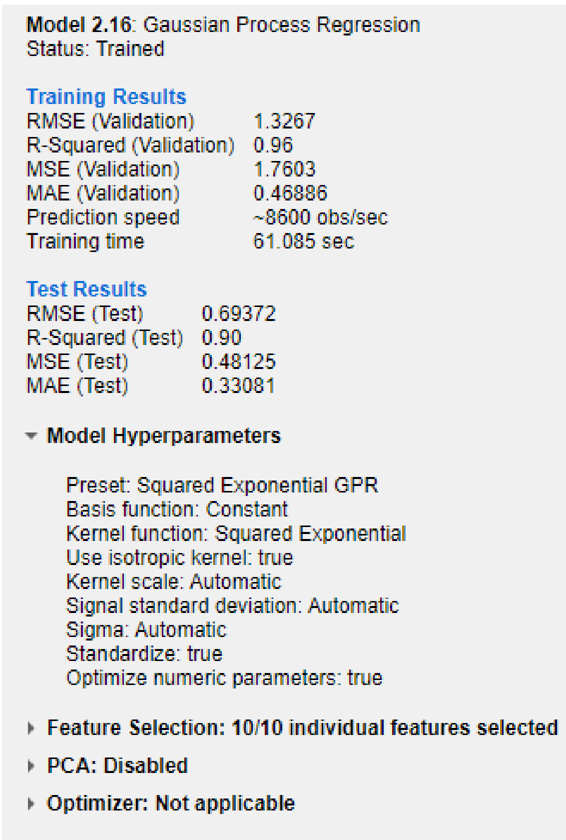


Figure 48: Predicted versus actual data, Squared Exponential GPR



Model 2.16: Gaussian Process Regression
Status: Trained

Training Results

RMSE (Validation)	1.3267
R-Squared (Validation)	0.96
MSE (Validation)	1.7603
MAE (Validation)	0.46886
Prediction speed	~8600 obs/sec
Training time	61.085 sec

Test Results

RMSE (Test)	0.69372
R-Squared (Test)	0.90
MSE (Test)	0.48125
MAE (Test)	0.33081

Model Hyperparameters

Preset: Squared Exponential GPR
Basis function: Constant
Kernel function: Squared Exponential
Use isotropic kernel: true
Kernel scale: Automatic
Signal standard deviation: Automatic
Sigma: Automatic
Standardize: true
Optimize numeric parameters: true

Feature Selection: 10/10 individual features selected

PCA: Disabled

Optimizer: Not applicable

Figure 49: Training data of Squared Exponential GPR

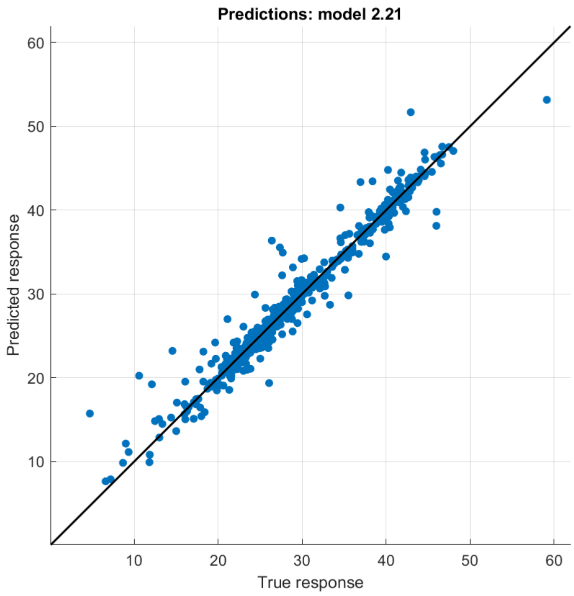


Figure 50: Predicted versus actual data, Medium Neural Network

Model 2.21: Neural Network
Status: Trained

Training Results

RMSE (Validation)	1.3433
R-Squared (Validation)	0.96
MSE (Validation)	1.8046
MAE (Validation)	0.63804
Prediction speed	~29000 obs/sec
Training time	65.834 sec

Test Results

RMSE (Test)	0.77239
R-Squared (Test)	0.87
MSE (Test)	0.59658
MAE (Test)	0.51005

Model Hyperparameters

Preset: Medium Neural Network
Number of fully connected layers: 1
First layer size: 25
Activation: ReLU
Iteration limit: 1000
Regularization strength (Lambda): 0
Standardize data: Yes

Feature Selection: 10/10 individual features selected

PCA: Disabled

Optimizer: Not applicable

Figure 51: Training data of Medium Neural Network

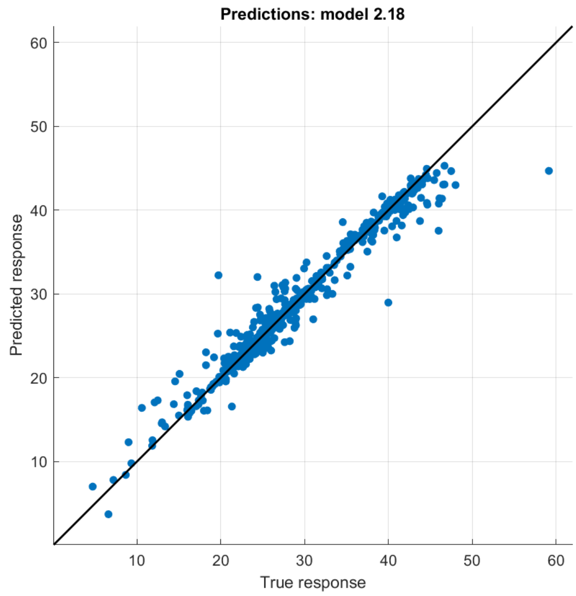


Figure 52: Predicted versus actual data, Exponential GPR

Model 2.18: Gaussian Process Regression
Status: Trained

Training Results

RMSE (Validation)	1.3623
R-Squared (Validation)	0.96
MSE (Validation)	1.8559
MAE (Validation)	0.62605
Prediction speed	~11000 obs/sec
Training time	48.124 sec

Test Results

RMSE (Test)	0.69601
R-Squared (Test)	0.90
MSE (Test)	0.48444
MAE (Test)	0.4586

- **Model Hyperparameters**
- **Feature Selection: 10/10 individual features selected**
- **PCA: Disabled**
- **Optimizer: Not applicable**

Figure 53: Training data of Exponential GPR

NTNU
Norwegian University of
Science and Technology