

# Physics-based, data-driven, and hybrid methods to predict the energy requirements for cargo heating: a case study for chemical tankers

B.K. van Blokland





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

# Physics-based, data-driven, and hybrid methods to predict the energy requirements for cargo heating: a case study for chemical tankers

by

B.K. van Blokland

Performed at

Stolt Tankers BV

This thesis (MT.22/23.031.M) is classified as confidential in accordance with the general conditions for projects performed by the TUDelft.

May 16, 2023

**Company supervisors**

Responsible supervisor: ir. J. Bogaard

**Thesis exam committee:**

Chair/Responsible professor Dr. ir. A. Coraddu

Staff Member: Dr. ir. L. van Biert

Company member: ir. J. Bogaard

**Author Details**

Studynumber 4464176

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



# Preface

This master thesis is the final part of my master Marine Technology. In this thesis, many different topics from the master came together. As this thesis was done in cooperation with Stolt Tankers, I learned a lot about how a shipping company operates. In the day to day operations I recognized quite some topics that were taught as part of the master. Nevertheless, I have to admit that as this master thesis was the first major project it was quite a challenge to combine multiple disciplines into this project while practice deviated slightly from theoretical models.

During this thesis I had two supervisors, whom I would like to thank: Andrea Coraddu as an academic supervisor and Jordi Bogaard as the supervisor from the company. During the whole process, Andrea made himself available and as such kept the focus on the academic relevance of the subject matter. He always managed to support me and assist me in resolving issues despite his busy schedule. I would like to thank Jordi for assisting me in defining the topic of my thesis and for introducing me to the company Stolt Tankers. Jordi was always very open to my ideas for a thesis subject and helped connect me not just within, but also to the company. Furthermore, Jordi ensured that I was well involved in all team activities at Stolt Tankers and gave me the freedom to work on my own project. Thank you Andrea and Jordi for accompanying me on this journey.

Besides my two supervisors, I would like to thank the TAM team from Stolt Tankers to welcome me as one of their team members. I really enjoyed all the lunches where the question was not if we would go for a walk, but where the walk was heading. Lunchtime created an excellent opportunity to address on the one hand topics in an informal manner and go over the items that needed to be resolved, while also sharing ideas for upcoming weekends and holidays.

Furthermore, I would like to thank my family and girlfriend Johanna for their support during this thesis. Thanks to their support it enabled me to accomplish this project and finish my thesis. They all have been supportive and helpful in the goal of reaching the end. Johanna showed patience in listening to the stories I shared with her about my day to day work at Stolt Tankers and the challenges to resolve the gaps I experienced in my model design.

Last but not least I would like to thank my friends for studying together. Being together at 3mE made studying a lot easier, talking about the issues you encountered during a cup of coffee. These coffee breaks were always good fun and I will miss the amount of them in the future!

*B.K. van Blokland  
Delft, April 2023*



# Abstract

Chemicals within a chemical tanker need heating to keep the cargo above a desired temperature. This research focuses on the prediction of the fuel consumption used to heat the cargo. There are three methods discussed in this research: a physics-based model (PBM), a data-driven model (DDM) and a hybrid model (HM). The PBM is a model based on the physical relations defined in the theory of heat transfer. The DDM is a model based on data gathered by a chemical shipping company (Stolt Tankers). Here an algorithm is trained and tested to acquire the model. The last method is a hybrid model. This HM is the combination of both the PBM and the DDM to use the best of both worlds.

The PBM uses ordinary differential equations to model the fuel consumption and uses a numerical approximation. This method is first compared with a simulation in COMSOL Multiphysics and then scaled to an existing vessel. The PBM is able to predict the fuel consumption over a trip where the deviation in fuel consumption is 0.043 tons of fuel per day (5.6% deviation). The temperature deviation averaged over all tanks is 0.68 °C. The DDM compares two algorithms: the linear algorithm based on the least squares method and an algorithm based on support vector regression (SVR). The model using SVR acquires the most accurate result, with a mean absolute error (MAE) of  $0.0014 \pm 1.7880 * 10^{-4}$  and a mean absolute percentage error (MAPE) of  $5.18616 \pm 1.05567$ . When the PBM and the DDM are combined into one model the HM is formed. The HM is able to predict the fuel consumption of the vessel with a MAE of  $0.0001441 \pm 0.00010$  and a MAPE of  $0.8929 \pm 1.0955$ .

The PBM could also be used to calculate the reduction in fuel consumption when insulation is applied. Different types of insulation are simulated, with different thicknesses and thermal conductivity properties. Insulation of the cargo tanks could potentially reduce the cargo heating consumption by 64.8%. This reduction makes it worth looking into the insulation of vessels to reduce the fuel consumption and emissions. The DDM and HM are also tested to see if these models are able to predict values outside the range for the trained data. The DDM is not able to do so, whilst the HM is able to calculate scenarios where the average cargo tank temperature is outside the range for trained data.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Cargo Heating . . . . .	2
1.2	Temperature and fuel consumption estimations for chemical cargo tanks. . . . .	2
1.3	Stolt Tankers . . . . .	3
1.4	Heating system . . . . .	4
1.5	Research questions . . . . .	4
1.6	Outline . . . . .	6
<b>2</b>	<b>Methodology</b>	<b>7</b>
2.1	Physics-based model. . . . .	8
2.1.1	Theory of heat transfer . . . . .	8
2.1.2	Cooling down model . . . . .	13
2.1.3	Chemical tanker model. . . . .	14
2.1.4	Verification . . . . .	16
2.1.5	Application of physics-based model: Insulation. . . . .	16
2.2	Data-driven model . . . . .	16
2.2.1	Algorithm selection . . . . .	16
2.2.2	Hyperparameter selection . . . . .	19
2.2.3	Extrapolation . . . . .	20
2.3	Hybrid model . . . . .	21
2.4	Error Estimation data-driven model and hybrid model . . . . .	21
2.4.1	Interval of confidence. . . . .	22
<b>3</b>	<b>Feature description</b>	<b>23</b>
3.1	Features physics-based model . . . . .	23
3.1.1	Layout of reference vessel . . . . .	23
3.1.2	Conversion of fuel consumption . . . . .	24
3.2	Features data-driven model . . . . .	24
3.2.1	Dataset acquisition . . . . .	24
3.2.2	Preliminary Data Exploration. . . . .	25
3.2.3	Data selection . . . . .	26
<b>4</b>	<b>Simulation setup</b>	<b>29</b>
4.1	Physics-based model. . . . .	29
4.1.1	Cooling down model . . . . .	29
4.1.2	Input for model . . . . .	29
4.1.3	Heat Transfer Model for a Chemical Tanker. . . . .	30
4.1.4	Insulation . . . . .	30
4.2	Data-driven approach . . . . .	32
4.2.1	Merging data . . . . .	32
4.2.2	Filtering data . . . . .	32
4.2.3	Data exploration . . . . .	32
4.2.4	Data selection for simulation. . . . .	33
4.2.5	Normalize data . . . . .	33
4.2.6	Hyperparameter selection . . . . .	34
4.3	Hybrid model . . . . .	35

<b>5</b>	<b>Results</b>	<b>37</b>
5.1	Physics-based model . . . . .	37
5.1.1	Result cooling down model . . . . .	37
5.1.2	Mesh refinement and timestep refinement . . . . .	38
5.1.3	Results Matlab versus COMSOL Multiphysics . . . . .	38
5.1.4	Chemical tanker model results . . . . .	41
5.1.5	Chemical tanker insulation . . . . .	42
5.2	Data-driven model . . . . .	43
5.2.1	Linear model . . . . .	43
5.2.2	Non-linear model . . . . .	43
5.2.3	Extrapolation . . . . .	44
5.3	Hybrid model . . . . .	44
5.3.1	Extrapolation . . . . .	45
5.4	Combination of results . . . . .	45
<b>6</b>	<b>Conclusion</b>	<b>47</b>
6.1	Physics-based model . . . . .	47
6.2	Data-driven model . . . . .	47
6.3	Hybrid model . . . . .	48
6.4	Research question . . . . .	48
<b>7</b>	<b>Discussion and further recommendations</b>	<b>51</b>
7.1	Physics-based model . . . . .	51
7.2	Data-driven model . . . . .	52
7.3	Hybrid model . . . . .	52
7.4	Further recommendations . . . . .	52
	<b>Bibliography</b>	<b>53</b>
<b>A</b>	<b>Available data parameters</b>	<b>59</b>
<b>B</b>	<b>Data visualisation</b>	<b>61</b>
<b>C</b>	<b>Cargo specifications</b>	<b>65</b>
<b>D</b>	<b>Thermodynamic Model Results</b>	<b>67</b>
<b>E</b>	<b>Insulation tank layout</b>	<b>73</b>

# List of Figures

1.1	Fuel consumption distribution of chemical tankers, Stolt Tankers internal data . . . . .	2
1.2	Temperature limits, actual scenario and improved scenario. Lower cargo temperature is better due to less additional heat, own figure . . . . .	3
1.3	Heating configuration with an example of heating coils on a chemical vessel . . . . .	5
2.1	Methods for fuel prediction . . . . .	7
2.2	Conduction . . . . .	9
2.3	Convection in inclined volume . . . . .	10
2.4	Heat transfer through wall, based on source [1] . . . . .	12
2.5	Matlab flow CDM . . . . .	14
2.6	Matlab flow CTM . . . . .	15
2.7	Build up of algorithm, own figure . . . . .	17
2.8	Schematic overview of SVR, based on Smola 2004 [2] . . . . .	18
2.9	Influence of $\epsilon$ , from [2] . . . . .	19
2.10	K-fold cross validation, own figure . . . . .	20
2.12	HM schematics . . . . .	21
2.11	Extrapolation method . . . . .	21
3.1	Schematic overview Stolt Concept and Stolt Effort . . . . .	24
3.2	Schematic overview Stolt Concept and Stolt Effort, numbered tanks . . . . .	24
3.3	Comparison between fuel consumption for cargo heating and air/sea temperature . . . . .	25
3.4	Correlation matrix with NR data, temperature data and temperature upper and lower limits . . . . .	27
3.5	From input features to desired output of model . . . . .	28
4.1	Overview of CDM . . . . .	30
4.2	Sea and air temperature during the simulated voyage . . . . .	31
4.3	Correlation Matrix . . . . .	34
4.4	HM dataset creation . . . . .	35
5.1	Matlab result CDM . . . . .	37
5.2	Timestep and interval show . . . . .	38
5.3	Mesh comparison COMSOL . . . . .	39
5.4	Results cooling down model Matlab and COMSOL . . . . .	39
5.5	Absolute difference between Matlab and COMSOL . . . . .	40
5.6	Relative difference between Matlab and COMSOL . . . . .	41
5.7	Predicted and reported fuel consumption . . . . .	42
5.8	Comparison: simulation without insulation and insulation with Charcoat CTI with a thickness of 4 [mm] . . . . .	42
5.9	Results DDM RidgeCV model . . . . .	43
5.10	Results DDM SVR model . . . . .	44
5.11	Results hybrid model . . . . .	45
5.12	Fuel consumption prediction over over time; PBM, DDM and HM . . . . .	46
B.1	Histograms of Unfiltered Noon Report Data . . . . .	61
B.2	Histograms of Unfiltered Temperature Data . . . . .	62
B.3	Histograms of Unfiltered Pressure data . . . . .	62
B.4	Histograms of Unfiltered Minimum Cargo Temperature Requirements . . . . .	63
B.5	Histograms of Unfiltered Maximum Cargo Temperature Requirements . . . . .	63
D.1	Temperature profile tanks 1-15 . . . . .	68

---

D.2	Temperature profile tanks 16-30 . . . . .	69
D.3	Temperature profile tanks 31-44 . . . . .	70
D.4	Difference between reported and predicted temperature per tank [degC] . . . . .	71
E.1	General overview of the location of insulation of the tanks . . . . .	74

# List of Tables

1.1	Fleetlist Stolt Tankers [3]	4
1.2	Fleet size overview	4
2.1	Typical values for thermal conductivity, from [4], p. 781	9
3.1	Principle dimensions Stolt Concept and Stolt Effort	23
3.2	Datatypes and location	25
4.1	Test setup CDM	29
4.2	Location of walls	30
4.3	Chemical properties of used materials in CDM	31
4.4	Different insulation materials	31
4.5	Filter values data	33
4.6	Chosen parameters for data-driven model	34
4.7	Hyperparameter range	34
5.1	COMSOL computational time	38
5.2	Different insulation materials	43
5.3	MAE and MAPE per model	43
5.4	Extrapolation of certain parameters	44
5.5	Results SVR hybrid model	45
5.6	Extrapolation of certain parameters for the hybrid model	45
6.1	Comparison between models	48
A.1	Cargo information and sensor information	59
A.2	Noon report data	60
C.1	Cargo specifications	65



# Acronyms

- BBM** black box model. 5, 7, 8, 21
- CDM** cooling down model. viii, ix, xi, 8, 14, 16, 29–31, 37
- CI** Confidence Interval. 22
- CII** carbon intensity indicator. 1
- CTM** chemical tanker model. ix, 8, 15, 29
- DDM** data-driven model. v, vii, ix, 7, 8, 16, 21, 23–25, 27, 29, 32, 34, 35, 37, 43–49, 51, 52
- DWT** deadweight tonnage. 4, 23
- ETS** emissions trading system. 1
- GBM** grey box model. 7, 8
- GHG** greenhouse gasses. 1
- HFO** heavy fuel oil. 24
- HM** hybrid model. v, vii, ix, xi, 7, 8, 21, 23, 24, 29, 35, 37, 44–49, 51, 52
- IMO** International Maritime Organization. 1
- MAE** mean absolute error. v, xi, 21, 43–45
- MAPE** mean absolute percentage error. v, xi, 21, 43–45, 48, 49
- MDO** marine diesel oil. 24
- MSE** mean square error. 21
- NN** neural network. 16
- NR** noon report. ix, 8, 24, 26, 27, 32, 33, 41, 52
- OLS** ordinary least squares. 17, 18
- PBM** physics-based model. v, vii, ix, 7, 8, 15, 16, 21, 23, 29, 30, 35, 37, 41, 45–49, 51, 52
- RBF** radial basis function. 18
- RFR** random forest regressor. 16
- RMSE** root mean square error. 21
- SVM** support vector machine. 18
- SVR** support vector regression. v, ix, 16, 18, 19, 34, 43–45, 52
- WBM** white box model. 7, 21



# Nomenclature

$\bar{Q}$	Heat flow rate
$\beta$	Thermal expansion coefficient
$\kappa$	Thermal conductivity
$\mu$	Dynamic Viscosity
$\nu$	Kinematic Viscosity
$\sigma$	Electrical conductivity
$\sigma$	Stefan Boltzmann constant
$\tau$	Relaxation time
$A$	Area
$C$	Total heat capacity
$c$	Heat capacity
$c_p$	Specific heat
$l$	Particle free path
$L_0$	Lorentz number
$n$	Particle density
$q_x, q_y, q_z$	Heat flux
$T$	Temperature
$v$	Velocity
$F$	Shape factor
$FC$	Fuel consumption
$g$	Gravitational acceleration
$g$	gravitational constant
$Gr$	Grashof Number
$h$	Heat transfer coefficient
$L$	length
$m$	Mass
$Nu$	Nusselt Number
$Pr$	Prandtl Number
$RaL$	Rayleigh Number
$Re$	Reynolds Number



# 1

## Introduction

Recent research by the International Maritime Organization (IMO) has shown that the shipping industry is responsible for 2.89% of the global emissions of greenhouse gasses (GHG) [5]. In 2018, the shipping industry emitted 1076 million tonnes of GHG, of which 1056 million tonnes were  $CO_2$ . To reduce the total amount of GHG, the IMO has determined that the global GHG by vessels should be reduced by 50% by 2050 [6]. How this could be done is still unknown, but in the meantime, some institutions, e.g. the European Union, have started the process in which emissions are monitored and billed [7]. The European Union aims at including shipping companies within the emissions trading system (ETS) [8]. The ETS is a trading system where emissions are charged. Besides the regular fuel costs, shipping companies become more and more penalized for their GHG-emissions, including the heating of their cargo. This is the first reason why shipping companies would like to reduce the fuel consumption. The second aspect of why a shipping company wants to be as efficient as possible is because of the carbon intensity indicator (CII) rating [9]. Each year each vessel receives a score from A-E. Here A is considered the best possible score, and E is regarded as the worst score. When a vessel receives an E score one year or a D score for three consecutive years, it has to come up with a plan to reduce emissions for the following year. In general, the CII is a measurement for the  $CO_2$  emitted per year divided by the distance travelled and the cargo carrying capacity [10]. This means that fuel consumption must be reduced in order to improve the CII rating. This requires a better understanding of the mechanics that impact fuel consumption related to cargo heating. One of the questions that arises is the question of how beneficial it is to insulate the cargo tanks.

This thesis is performed in cooperation with Stolt Tankers, which is an operator of chemical tankers. Figure 1.1 shows the average fuel consumption distribution for chemical tankers owned by Stolt Tankers. The largest fuel consumer is the propulsion system of a vessel (72%). The second largest consumer is the generator (21%), which is used for steering and the electronics. The third aspect, illustrated in Figure 1.1, is the cargo heating consumption, which makes up for about 5% of the total fuel consumption of a vessel. The last aspect of the chemical tankers is the tank cleaning consumption (2%). Tanks have to be cleaned after usage to prevent contamination of chemicals.

Multiple studies have been performed to improve the fuel consumption for propulsion [11], [12], [13]. The generator for steering is analysed in [14] and [15]. There is almost no literature found about the fuel consumption prediction for cargo heating systems on board of chemical tankers. Besides this, it becomes increasingly interesting for shipping companies to specify how much fuel is used for which reasons. Vessel operators would like to start to bill the costs for fuel consumption and emissions to their customers. One part of this is that cargo that requires heat could be charged with a higher price than cargo that does not require heat. To be able to bill these costs to their customers, it must be clear how much energy is consumed to heat the cargo.

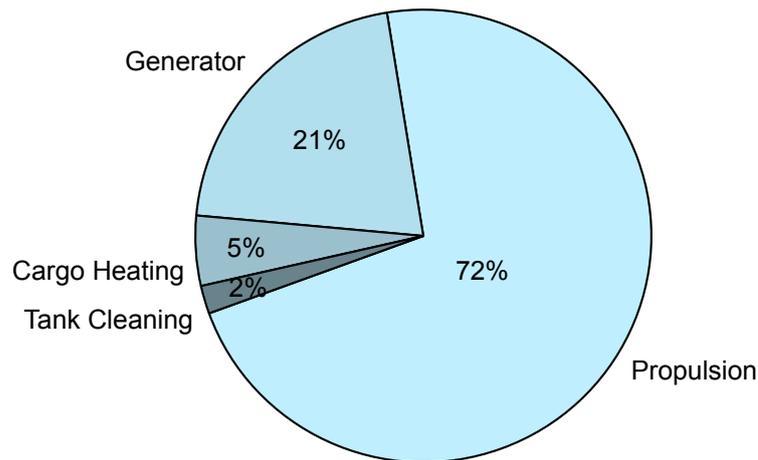


Figure 1.1: Fuel consumption distribution of chemical tankers, Stolt Tankers internal data

## 1.1. Cargo Heating

Cargo heating is an important aspect of the transportation of chemicals. The various transported chemicals all have a specific temperature window at which they must be transported to keep the chemicals stable. The owner of the chemicals defines this range, and the operator of the chemical tankers must keep the chemicals in this range. This means the vessel operator must be able to control the temperature of the chemicals and therefore, the operator has the ability to heat the cargo in a tank with the help of heating coils. The temperature range is important because firstly, a chemical that becomes too hot or cold can react with itself. This can mean that the cargo changes in chemical properties and thus is destroyed. Secondly, besides the possible loss of the cargo, safety issues can occur when a cargo reacts with itself. Cargo above the stated limit can start an exothermic reaction. This means that cargo will become warmer and warmer, and in the end, explosions might occur. The third reason to consider the heating and heat exchange between cargo tanks is for economic reasons.

Figure 1.1 shows that the fuel consumption for cargo heating is 5% of the total energy consumption. However, limited research is available for the cargo heating of chemical tankers. Bureau Veritas studied crude oil tankers [16]. The vessels that are considered in this research are Aframax size. This research shows that cargo heating optimization for crude oil tankers could reduce fuel consumption by 30% if the voyage is longer than 30 days. When a voyage is longer, there is more time to optimise the cargo heating consumption. Crude oil has a certain discharge temperature which is above the transportation temperature. A longer voyage results in more time to decrease the temperature and gradually increase it to reach the discharge temperature. The main fuel consumption reduction can be achieved by reducing the cargo temperature within the specified limits. The crude oil temperature is allowed to drop from 120 °C to 38 °C, and then heated to the discharge temperature again. Figure 1.2 shows the current operational cargo temperature profile and the aim for a better operational cargo temperature profile. The aim of a shipping company is to add as little as possible heat to the cargo tanks whilst keeping the cargo within the specified range. In the current situation, it often occurs that cargo is heated while the temperature is not close to the lower limit. This means that there is room to reduce the fuel consumption for cargo heating by having lower temperatures in the cargo tanks. The operator must keep a focus on the temperature limit, but fuel could be saved there.

## 1.2. Temperature and fuel consumption estimations for chemical cargo tanks

Within the chemical tanker industry, there have been accidents caused by cargo being above the maximum temperature limit. There have been cases where this was a result of cargo heating neighbouring cargo above the maximum allowed temperature. To prevent this, there is a need to control the heat flow and thus, temperature estimations for each individual tank.

There are several methods to be able to calculate the heat transfer between tanks. The first method

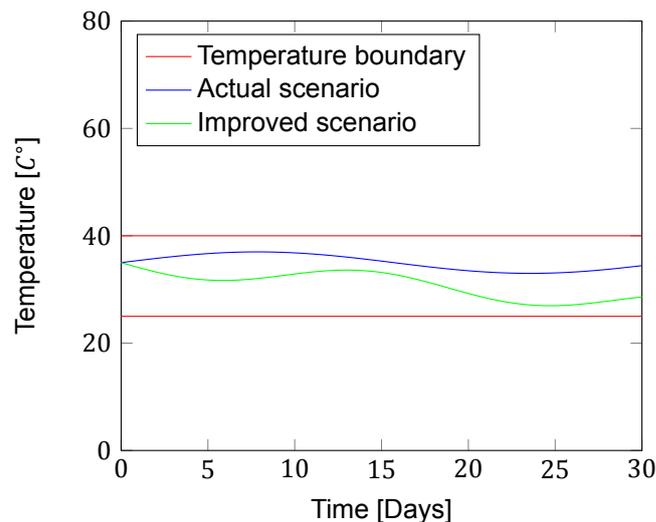


Figure 1.2: Temperature limits, actual scenario and improved scenario. Lower cargo temperature is better due to less additional heat, own figure

is to model the tanks based on thermodynamic heat transfer. There are many authors who have written books about the fundamentals of heat transfer: for example, the books of Lienhard [17], Dass [18] and Incropera [19]. These books however are not specifically focused on chemical cargo tanks within vessels.

There has been some research performed on the thermodynamic behaviour in cargo tanks. Based on the thermodynamics and energy flows, Baldi [20] has studied how energy is used on board of chemical tankers. The chemical tanker considered by Baldi, however, does not require cargo heating, but the tanks that carried the fuel for the vessel needed heating. This already makes up for 0.9% of the total fuel consumption. Gao [21] performed a numerical simulation with the help of COMSOL for oil tankers in the Arctic region. The simulation time Gao considered was 600s. The study of Gao gives insights into the disappearance of energy. The study by Gao shows that the energy dissipates due to natural convection. Three physical properties are needed for this research: oil density, specific heat and thermal conductivity. Bochenski [22] analysed the cargo heat demand. Bochenski did this by analysing two vessels: a container vessel and a general cargo vessel. The result of the performed study was that the heat demand was dependent on the water temperature and the air temperature. When the water and air temperature was higher, less energy was required to control the cargo temperature. Sari [23] studied the energy balance in a chemical tanker. The result of this research was that this chemical tanker used 85% of the energy for propulsion, 11% was used for electricity and 2% was used for heating. All other demands were 2% of the total energy consumption. Shen [24] studied the heating of fuel tanks and showed that there was a difference of  $1.5\text{ }^{\circ}\text{C}$  between different measurement points. The first point was located at 0.1m above the coils, and the highest measurement point was located 0.4m above the coils. Both numerical simulations and experiments show that the temperature rises proportionally to the heating time.

Besides the research based on thermodynamic principles, studies have focused on the prediction of certain aspects within the maritime industry based on operational data. Only limited research about the heating of cargo within a chemical tanker is available. However operational data give the opportunity to monitor fuel consumption and to predict the fuel consumption for future trips. There is research done on the fuel oil consumption prediction [25] and monitoring [26], speed optimization [27] and maintenance optimization [28], [29] based on operational data. However, there is no literature found where operational data is used within the maritime industry to predict fuel consumption for cargo heating.

### 1.3. Stolt Tankers

This thesis is made in cooperation with Stolt Tankers. Stolt Tankers operates several fleets, sometimes in cooperation with another company. Some of the vessels within a fleet are owned by Stolt Tankers, sometimes Stolt Tankers is in a joint venture with another company and sometimes Stolt Tankers is only

accountable for the technical management of the vessels. Each fleet consists of multiple vessels. The fleets Stolt Tankers operates are shown in Table 1.1. The main fleet that Stolt Tankers operates is the STJS fleet. This fleet contains 56 vessels and is owned by Stolt Tankers. Stolt Tankers operate about 5.8% of the chemical parcel tankers, and the chemical tankers are about 2.3% of the world merchant fleet based on deadweight tonnage (DWT) (see Table 1.2).

Table 1.1: Fleetlist Stolt Tankers [3]

Fleet name	Vessels owned by Stolt Tankers [Y/N]	Vessel Amount
Stolt Tankers Joint Services (STJS)	Y	56
NYK Stolt Tankers	N	9
Tufton Investment Management	N	8
Hassel Shipping 4 AS	N	8
E&S Tankers	Partially	14
Stolt NYK Asia Pacific Services (SNAPS)	N	11
Stolt-Nielsen Inter-Asia Service (SNIAS)	Y	2
Stolt-Nielsen Inter-Caribbean Service (SNICS)	Y	8
Stolt-Nielsen Inland Tanker Service (SNITS)	N	36
US Gulf Barging	N	3
SC-Stolt Shipping	N	9
<b>Total</b>		<b>164</b>

Table 1.2: Fleet size overview

	World fleet	Chemical Tankers	Stolt Tankers BV fleet
DWT	2.2E9 [30]	49956358 [30]	2880141 [30]
GT	1551E6 [31]	32399162 [30]	1876296 [30]

## 1.4. Heating system

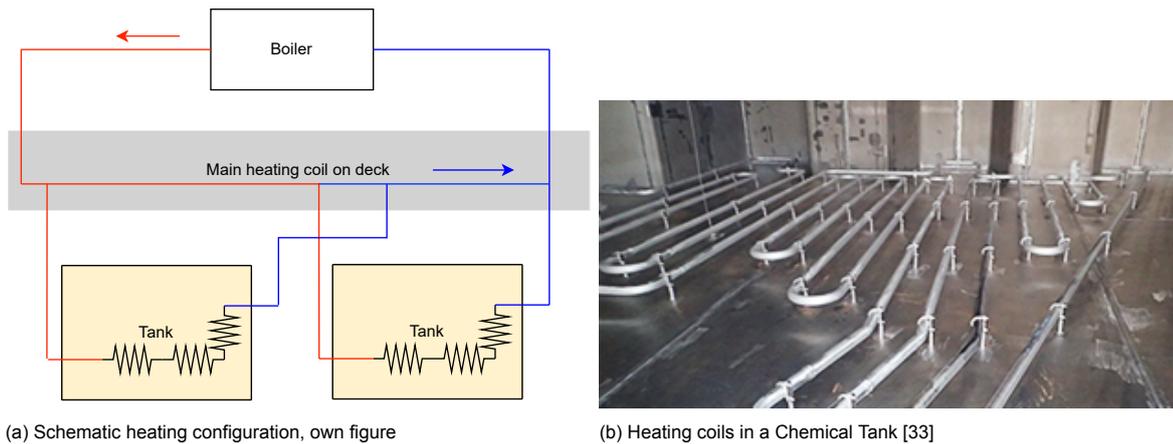
Each vessel at Stolt Tankers is equipped with a cargo heating system. The cargo heating system is a system where in each cargo tank heating coils are fitted. This system is in place to enable tanks to be heated if necessary. Figure 1.3a shows a schematic overview of the heating system. On deck, there is a main heating coil, which is connected to all tanks. These coils are located at the bottom of the tank and/or at the walls. An example of these heating coils at the bottom is displayed in Figure 1.3b. These heating coils are filled with steam or thermal oil based on the cargo specifications [32]. The inflow of steam and/or oil is manually controlled by valve adjustment. The main heating coil on deck is sometimes insulated, but that is vessel dependent. The reason to insulate these heating coils is to reduce heat losses. The insulation, however, is able to react with seawater or rain, since it is open on deck. This leads to a deterioration of the insulation at the deck and is sometimes a reason to not install or remove the insulation.

## 1.5. Research questions

As discussed in the previous sections, limited research is available on this topic. This created an opportunity to research this topic in this thesis. The theory of heat transfer is widely known, but not yet applied in the chemical tanker industry. The goal of this thesis is to find a method to predict the fuel consumption for cargo heating. This is stated in the research question below.

**Research question:** What models can be used to predict the fuel consumption for heating of cargo for a given vessel and specific voyage conditions, taking operational use into account?

The research question can be decomposed into several elements to focus on in this research. "Cargo": chemical tankers have the ability to carry different cargo types. This research does not only focus on crude oil, which is previously researched [16], but aims at predicting the fuel consumption for



(a) Schematic heating configuration, own figure

(b) Heating coils in a Chemical Tank [33]

Figure 1.3: Heating configuration with an example of heating coils on a chemical vessel

different types of cargo. A *"given vessel"* means that an existing vessel is taken for a reference case. This means that the predictions could be compared against an actual vessel, but the research depends on the vessel geometry and layout. The third element is the *"specific voyage conditions"*. Each voyage encounters different circumstances and this changes the fuel consumption for cargo heating (e.g. wave height, current, water temperature). Besides this, the sentence starts with *"How is it possible to predict"* this is because there are potentially different possibilities to predicting the cargo heating consumption. The part *taking operational use into account* ensures that the research keeps in mind that it should be a usable approach for a chemical shipping company. To obtain an answer to this research question, the following subquestions must be answered.

**I:** Which physical phenomena influence the fuel consumption of the heating of cargo?

The start of this research focuses on the principles behind the heating of cargo. This question focuses on the theoretical background. This knowledge is later on used to determine the models to predict the cargo heating consumption.

**II:** To what extent are ordinary differential equation models, based on available physical data, solved with numerical Euler methods able to capture the fuel consumption of the heating of cargo per trip?

The second subquestion makes use of the knowledge gathered in subquestion 1. Subquestion 1 aims to find the theory of how heat transfer works and subquestion 2 determines a model to calculate the heat transfer between the cargo tanks. This subquestion takes the computational effort into account, since the ordinary differential equations can be solved by Euler methods. The aim is to build a model that runs within a day, otherwise, it will not be of practical use. The book by Lienhard [17] explains that the heat transfer equations are based on ordinary differential equations. These ordinary differential equations have to be solved. There are multiple options to approximate these equations in time and this will be further investigated during this research. Furthermore, the geometries have to be modelled. These geometries are modelled based on available physical data.

**III:** How can operational data be exploited for the consumption prediction of heating of cargo for a single trip?

Shipping companies have started to gather data. There are multiple papers within the maritime industry where a black box model (BBM) is used to determine the fuel consumption of a vessel. These papers are not specifically for the fuel consumption for cargo heating, but this thesis will use a BBM to see if this is a suitable method to predict the fuel consumption for cargo heating.

## **1.6. Outline**

After the introduction, the report continues with chapter 2, where the methodology for this report is explained. Chapter 3 describes all the features and data sources needed in this research. Chapter 4 continues with the setup of the simulations and calculations. The results of these simulations are presented in chapter 5. This is followed by conclusions about these results in chapter 6. The report finishes with chapter 7, where there is a discussion of the results and recommendations for further research are made.

# 2

## Methodology

The goal of this research is to predict the fuel consumption for cargo heating for a vessel during a specific trip. There are several possibilities to achieve this goal. For example, it is possible to take the average of the boiler consumption during a specific year, or take the average boiler consumption while considering the amount of cargo taken into the vessel. The research question states that the fuel consumption must be predicted based on a specific voyage and vessel. Literature suggests that there are three different types of models possible: white box model (WBM), black box model (BBM) and grey box model (GBM) [34], [35].

The first approach used in this research is a physics-based model (PBM), a so-called WBM. The second approach used in this thesis is a BBM. This approach uses the operational data to create a model to calculate the fuel consumption. The third approach is to combine these two methods into one method, called a hybrid model (HM), or a GBM. The first approach gives a clear indication of the optimal solution whilst the second approach gives insights into the actual performance. The first approach is called the physics-based model (PBM) and the second approach is called the data-driven model (DDM). The third approach in this research is called a hybrid model (HM). This is a combination of both the PBM and the DDM. The overview is displayed in Figure 2.1.

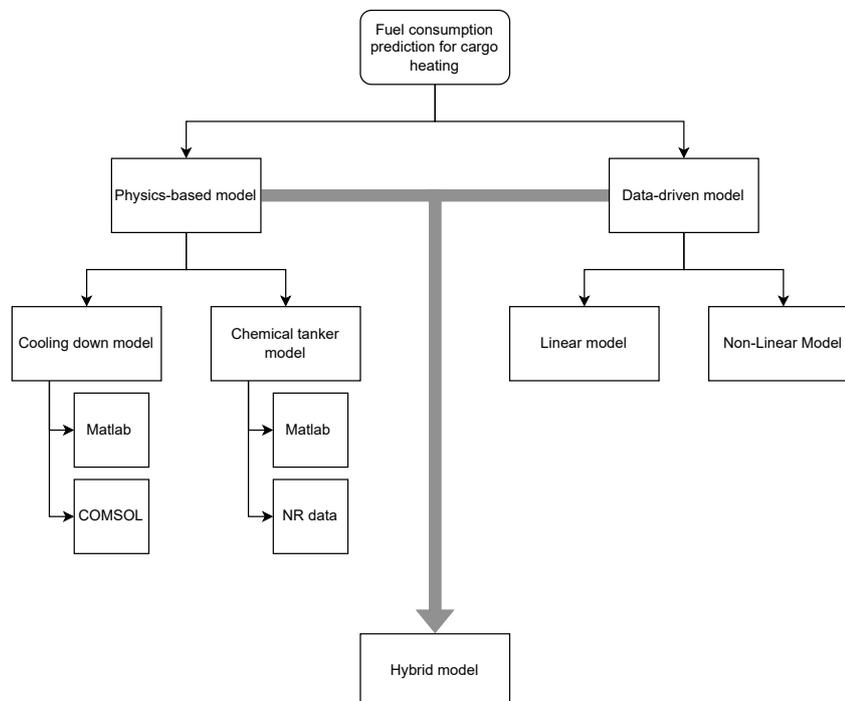


Figure 2.1: Methods for fuel prediction

The PBM consists of two parts. The first part is a *cooling down model (CDM)* and the second part is the *chemical tanker model (CTM)*. The CDM is a simplification of reality, to be able to verify the thermodynamic equations used in the model. The result is compared to COMSOL Multiphysics to see if the same input generates an equal output. The next step is to model a trip of a chemical tanker vessel and compare the output against the actual data. The actual data is the available noon report (NR) data. The output of the CTM is a graph of fuel consumption over time given a certain input. This approach is explained in section 2.1.

The DDM uses the available operational data. Two different possible algorithms are investigated. The first algorithm is a linear model and the second algorithm that will be used is a nonlinear model. The data is processed in Python, where the Scikit-Learn toolboxes are used for the algorithms [36]. This can be seen as a BBM. The methodology for the DDM is explained in section 2.2.

The HM is a combination of the PBM and the DDM and is called a grey box model (GBM). The DDM is based solely on data, whilst the PBM is based solely on thermodynamic calculations. The HM uses a dataset containing the information of both methods to learn a model of the relations to be able to benefit from the two methods. The HM is explained in section 2.3.

## 2.1. Physics-based model

This section explains the methodology for the PBM. First, the theory is explained in subsection 2.1.1, followed by the CDM in subsection 2.1.2. This method is then scaled to a trip for a vessel in subsection 2.1.3 and concludes with the method for the verification in subsection 2.1.4.

### 2.1.1. Theory of heat transfer

This section explains the theory of heat transfer. Energy has three possibilities to transfer. These are conduction, convection and radiation. Conduction is heat transferred due to the motions of electrons and is first explained. Convection is a fluid flow that occurs due to a change in temperature. This change in temperature causes the medium to have a different specific weight and, thus a different gravity. This leads to a movement of heat flow and will be further explained. The third aspect of heat transfer is radiation. Radiation is when the movement is transferred into electromagnetic radiation.

**Conduction** Conduction is heat transferred due to the motions of electrons within a material. When a material is heated, the material will be warmer close to the heat source than far away from the heat source. In 1878 Fourier published a book called 'The Analytical Theory of Heat', where the 'Fourier Law of Conduction' was first described. Three factors influence conduction: fluid motion, fluid nature and surface geometry [37]. Conduction is visualised in Figure 2.2. A heat flux  $q_1$  goes through a solid material with thickness  $\Delta x$ . The heat flux  $q_2$  is a result of this. It can be seen that there are two temperatures involved in the equation, namely  $T_1$  and  $T_2$ . The Fourier Law of Conduction is stated in Equation 2.1. In this equation,  $q_x$ ,  $q_y$  and  $q_z$  are the heat fluxes,  $\kappa$  expresses the thermal conductivity and  $\frac{\partial T}{\partial x}$  is the gradient of temperature in the solid. Equation 2.1 shows that the heat flows from the higher temperature towards the lower temperature.

$$q_x = -k \frac{\partial T}{\partial x}, \quad q_y = -k \frac{\partial T}{\partial y}, \quad q_z = -k \frac{\partial T}{\partial z} \quad (2.1)$$

To be able to use the Fourier law of conduction, the thermal conductivity ( $\kappa$ ) should be known. The definition of thermal conductivity is the heat flow rate vector ( $\vec{Q}$ ) divided by the change in temperature ( $\vec{\nabla}T$ ). This is shown in Equation 2.2. The ability of a material to conduct heat is called thermal conductivity and is the sum of the internal energy carried by electrons. Each electron contains a certain amount of energy from excitation [38]. The total thermal conductivity ( $\kappa$ ) is the sum of all components of thermal conductivity from excitation. This is expressed in Equation 2.3, where  $\alpha$  the excitation is. Equation 2.3 shows the summation of all thermal conductivity.

$$\kappa = -\frac{\vec{Q}}{\vec{\nabla}T} \quad (2.2)$$

$$\kappa = \sum_{\alpha} \kappa_{\alpha} \quad (2.3)$$

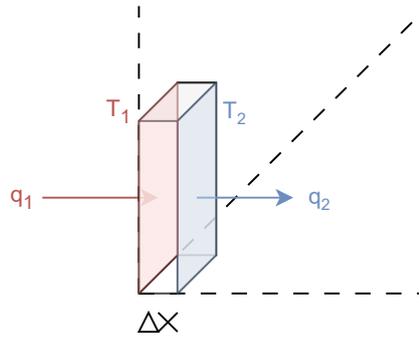


Figure 2.2: Conduction

The average heat flow rate  $\vec{Q}$  can be calculated to sum all particles [38]. This is stated in Equation 2.4. The symbols  $n$ ,  $c$ ,  $\tau$ ,  $\vec{v}$ , and  $\Delta T$  stands for the particle density, the heat capacity, the relaxation time, the velocity and the temperature difference.

$$\vec{Q} = -nc\tau\langle\vec{v} \cdot \vec{v}\rangle\vec{\nabla}T = -\frac{1}{3}nc\tau v^2\vec{\nabla}T \quad (2.4)$$

When Equation 2.2 and Equation 2.4 are merged into one equation, this results into Equation 2.5. As shown in Equation 2.3, this can be summated over  $\alpha$ , leading to the general Equation 2.6. The coefficient  $C$  is the total heat capacity and  $l$  is the particle mean free path.

$$\kappa = \frac{1}{3}nc\tau v^2 = \frac{1}{3}Cvl \quad (2.5)$$

$$\kappa = \frac{1}{3} \sum_{\alpha} C_{\alpha} v_{\alpha} l_{\alpha} \quad (2.6)$$

However, in most cases the thermal conductivity cannot be found exactly. Table 2.1 displays standard values for the thermal conductivity. In general, the conduction could be calculated according to Equation 2.7. Here, the Fourier law of conduction is multiplied with the area to get the total heat flow rate vector.

$$Q_{net,conduction} = q_{net,conduction}A = -kA \frac{\partial^2 T}{\partial x^2} \partial x \quad (2.7)$$

Material	Thermal conductivity [W/mK]
Silver	410
Copper	385
Aluminium	225
Cast-iron	55-65
Steel	20-45
Concrete	1.2
Glass	0.75
Glass wool	0.03
Water	0.55-0.77

Table 2.1: Typical values for thermal conductivity, from [4], p. 781

**Convection** Convection is the heat transfer caused by the movement of a fluid. A hot surface will influence a fluid's temperature. The fluid could start moving due to the change in density. This is called convection [37] and is displayed in Figure 2.3. Convection can be calculated according to Equation 2.9. The amount of convection is influenced by the temperature of the body and the temperature of the fluid. The difference between these two temperatures ( $T$ ), multiplied by the heat transfer coefficient ( $h$ ),

leads to the heat flux for the convection as shown in Equation 2.8. The total amount of heat due to convection can be calculated by multiplying Equation 2.8 with the area (A) which is considered. This results in Equation 2.9.

$$q_{convection} = h(T_S - T_{\infty}) \quad (2.8)$$

$$Q_{convection} = -hA(T_S - T_{\infty}) \quad (2.9)$$

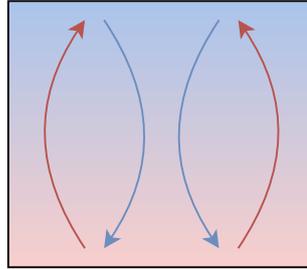


Figure 2.3: Convection in inclined volume

Nusselt derived a dimensionless coefficient which is the ratio between the convective and conductive heat transfer. This number can be calculated according to Equation 2.10. In this equation,  $h$  is the heat transfer coefficient,  $L$  is the surface length and  $k$  is the thermal conductivity. When this equation is rearranged, the heat transfer can be calculated (Equation 2.11).

Equation 2.8 and Equation 2.9 are equations to calculate the heat transfer when convection occurs. There are two types of convection, natural convection and forced convection. In Equation 2.8 and Equation 2.9, the  $h$  is the heat transfer coefficient,  $T_S$  is the temperature of the surface,  $T_{\infty}$  is the fluid temperature and  $A$  is the area. The difference between natural convection and forced convection is that with forced convection the fluid is moving not only by the speed caused by the convection, but is also moved by an external factor. This external factor is for example the forced movement of water along the hull of a vessel, wind flowing over a deck or movement of the liquid in the cargo tanks caused by the movement of the vessel.

$$Nu_x = \frac{hL}{k} = f(Re, Pr, Gr, Ec) \quad (2.10)$$

$$h = \frac{k}{L} Nu_L \quad (2.11)$$

The next step to calculate the heat transfer coefficient is the Nusselt number. The Nusselt number, however, depends on the kind of convection. There is a distinction between natural convection and forced convection. First, natural convection is explained and after that, the differences with forced convection are explained.

#### *Natural convection*

Churchill has derived an empirical function to determine the Nusselt number as a function of both the Rayleigh and Prandtl functions [39] for natural convection. This research shows two equations to calculate the Nusselt number as a function of Rayleigh and Prandtl. The equations for Rayleigh and Prandtl are shown in Equation 2.15 and Equation 2.14 [19]. The Rayleigh number is the Grashof number multiplied by the Prandtl number.

When equations Equation 2.10 and Equation 2.12 are combined, the heat transfer coefficient can be determined. The Rayleigh number is a number that is associated with the buoyancy driven force and the Grashof number is a dimensionless number that expresses the ratio between the buoyancy and viscous forces on a liquid. When these are known, the heat transfer coefficient can be determined as stated in Equation 2.16. In the Grashof number, the  $g$  is the gravitational acceleration,  $\beta$  is the thermal expansion coefficient,  $T_S$  is the surface temperature,  $T_{\infty}$  is the fluid temperature,  $L$  is the length

and  $\nu$  is the kinematic viscosity. In Equation 2.14  $c_p$  is the specific heat,  $\mu$  is the dynamic viscosity and  $k$  is the thermal conductivity.

$$Nu_{natural} = \frac{hL}{k} = \begin{cases} 0.68 + \frac{0.67Ra_L^{1/4}}{[1 + (0.492/Pr)^{9/16}]^{4/9}} & Ra_L \leq 10^9 \\ (0.825 + \frac{0.387Ra_L^{1/6}}{[1 + (0.492/Pr)^{9/16}]^{8/27}})^2 & Ra_L > 10^9 \end{cases} \quad (2.12)$$

$$Gr = \frac{g\beta(T_s - T_\infty)L^3}{\nu^2} \quad (2.13)$$

$$Pr = \frac{c_p\mu}{k} \quad (2.14)$$

$$Ra_L = Gr \times Pr \quad (2.15)$$

$$h = \begin{cases} \frac{k}{L} \left( 0.68 + \frac{0.67Ra_L^{1/4}}{[1 + (0.492/Pr)^{9/16}]^{4/9}} \right) & Ra_L \leq 10^9 \\ \frac{k}{L} \left( 0.825 + \frac{0.387Ra_L^{1/6}}{[1 + (0.492/Pr)^{9/16}]^{8/27}} \right)^2 & Ra_L > 10^9 \end{cases} \quad (2.16)$$

#### Forced convection

In the case of forced convection, the movement of the fluid is not only caused by the thermal gradients at the surface [40]. The Nusselt number changes. This becomes Equation 2.17 [41]. In this equation,  $C$ ,  $m$  and  $n$  are constants dependent on the geometry of the plates and the friction in the boundary layer. These constants vary for each geometry and are empirically determined. Furthermore, in calculations of forced convection, the Nusselt number becomes dependent on the Reynolds number instead of the Rayleigh number. The Reynolds number is a dimensionless ratio between inertial and viscous forces. The equation for the Reynolds number is given in Equation 2.18. Here,  $u$  is the flow speed,  $L$  is the length and  $\nu$  is the kinematic viscosity.

$$Nu_{forced} = CRe^m Pr^n \quad (2.17)$$

$$Re = \frac{uL}{\nu} \quad (2.18)$$

**Radiation** Radiation is the transfer of energy which is emitted through an electromagnetic process [17]. A body emits energy according to Equation 2.19. This equation is equal to the Stefan Boltzmann law, with the exception of the addition of  $\varepsilon$ . The Stefan Boltzmann law describes the energy that a black body emits, however, not all bodies are black bodies.  $\varepsilon$  is the relation between the real body emissive power and the black body emissive power, and can be calculated according to Equation 2.20.  $\sigma$  is the Stefan-Boltzmann constant. The value of  $\sigma \approx 5.67 \times 10^{-8} \text{ Wm}^{-2}\text{K}^{-4}$  [42]. Kirchoff states that when a body is in thermal equilibrium, it emits the same amount of energy in each wavelength and direction as it absorbs. It is important to mention that objects do not need to be in direct physical contact for radiation to take place [43]. Furthermore, the larger the temperature difference, the more radiation occurs.

$$e_b = \varepsilon\lambda\sigma T^4 \quad (2.19)$$

$$\varepsilon_\lambda = \frac{e_\lambda(\lambda, T)}{e_{\lambda_b}(\lambda, T)} \quad (2.20)$$

This all results in a net heat flux that can be calculated according to Equation 2.21. In this equation,  $T_1$  is the higher temperature and  $T_2$  is the lower temperature. However, this equation shows the heat

flux. To calculate the total amount of heat that transfers, the area has to be considered. Furthermore, a shape factor  $F_{1-2}$  comes into play. This results in Equation 2.22 [17]. In this equation, it is assumed that all energy that leaves body 1 is received by body 2. That is the reason that only area 1 has to be considered.

$$q_{net,radiation} = \varepsilon\sigma(T_1^4 - T_2^4) \quad (2.21)$$

$$Q_{net} = A_1F_{1-2}\sigma(T_1^4 - T_2^4) \quad (2.22)$$

**Total heat interaction** The total heat interaction is an addition of the conduction, convection and radiation that takes place at the same time. Conduction can be calculated according to Equation 2.7, convection can be calculated according to Equation 2.9, and radiation can be calculated according to Equation 2.22. The first law of thermodynamics (Equation 2.26) describes a method that can be used to calculate the change in temperature of the object. All four equations are shown again in Equations 2.23, 2.24, 2.25, 2.26. Zhang states that radiation in both liquid and solid material is negligible compared to conduction and convection [13]. This is because of the large density and because the radiation is only absorbed by a thin layer close to the surface. Thus radiation is not taken into account in the model.

$$Q_{net,conduction} = -kA \frac{\partial^2 T}{\partial x^2} \partial x \quad (2.23)$$

$$Q_{net,convection} = -hA(T - T_\infty) \quad (2.24)$$

$$Q_{net,radiation} = A_1F_{1-2}\sigma(T_1^4 - T_2^4) \quad (2.25)$$

$$Q_{net,firstlaw} = mc \frac{dT}{dt} \quad (2.26)$$

Figure 2.4 shows the heat transfer between two tanks without the influence of the surroundings. Tank 1 has a higher initial temperature than tank 2, which means that the energy flows from tank 1 towards tank 2. In this figure, both convection and conduction take place. Convection takes place from the middle of the tanks to the wall of the tanks and conduction takes place within the wall of the tank.

1.  $T_{\infty 1} \rightarrow T_1$  Convection
2.  $T_1 \rightarrow T_3$  Conduction
3.  $T_3 \rightarrow T_2$  Conduction
4.  $T_2 \rightarrow T_{\infty 2}$  Convection

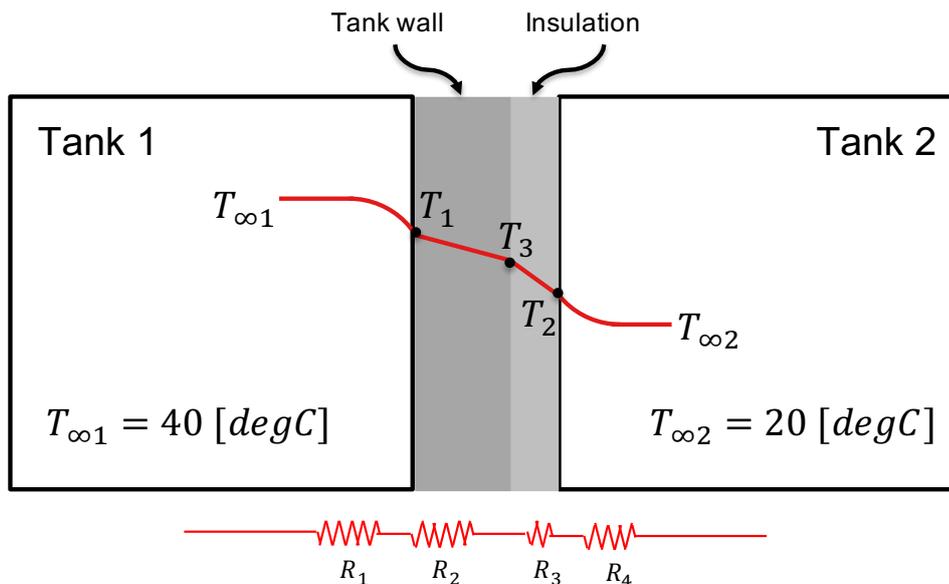


Figure 2.4: Heat transfer through wall, based on source [1]

According to Incropera [19], the heat transfer of a combined system should be modelled as resistances. In Figure 2.4 the resistances  $R_{1-4}$  are shown in red. In Equation 2.27 and in 2.28,  $h_i$  is the heat transfer coefficient from tank  $i$ ,  $A_i$  is the wall area,  $x_i$  is the thickness of the wall or insulation,  $k_i$  is the thermal conductivity. The heat transfer coefficient is dependent on both the liquid inside the tank and the wall geometry. The thermal conductivity depends on the material properties.

$$R_{convective} = \frac{1}{h_i A_i} \quad (2.27)$$

$$R_{conductive} = \frac{x_i}{k_i A_i} \quad (2.28)$$

The total heat transfer between tank 1 and tank 2 can be calculated according to Equation 2.29.

$$\vec{q} = \frac{T_{\infty 1} - T_{\infty 2}}{\frac{1}{h_1 A_1} + \frac{x_2}{k_2 A_2} + \frac{x_3}{k_3 A_3} + \frac{1}{h_4 A_4}} \quad (2.29)$$

### 2.1.2. Cooling down model

**Matlab model setup** Previous sections of this chapter described the theoretical framework. Upon this basis, the resulting application in the model set-up will now be described in this section. The schematics of the model are shown in Figure 2.5. Step 1 is to determine the settings. This part of the script determines the duration of the simulation and the interval between points to be calculated. The script uses a Forward Euler method to solve the equations. The Forward Euler method is displayed in Equation 2.30. In this equation,  $y_n$  is the value for the solution at step  $n$ ,  $y'_n$  is the first derivative of the system and  $\Delta T$  is the timestep.

$$y_{n+1} = y_n + y'_n \Delta T \quad (2.30)$$

The settings of the timesteps are used to create the right initial parameters for the model. The characteristics are divided into three parts: the tank characteristics, the material characteristics and the environmental characteristics. The tank characteristics are the physical dimensions of the tanks, the material characteristics are to assign cargo and the steel properties of the plates to each tank. The environmental characteristics are the weather conditions. The heat calculations first derive the heat transfer coefficient for each wall, using Equation 2.16. These coefficients can be used to calculate the total resistance of the wall according to Equation 2.27. The characteristics of the walls enable the model to calculate the resistance of the walls according to Equation 2.28. The combination of these two equations leads to Equation 2.29. For each individual timestep, the heat transfer coefficient has to be determined again since it is dependent on the temperature difference. This leads to different resistances over time and thus, a different heat flow. The last step of the model is the visualisation of the temperatures and heat flows over time.

$$h = \begin{cases} \frac{k}{L} \left( 0.68 + \frac{0.67 Ra_L^{1/4}}{[1 + (0.492/Pr)^{9/16}]^{4/9}} \right) & Ra_L \leq 10^9 \\ \frac{k}{L} \left( 0.825 + \frac{0.387 Ra_L^{1/6}}{[1 + (0.492/Pr)^{9/16}]^{8/27}} \right)^2 & Ra_L > 10^9 \end{cases} \quad (2.16)$$

$$q = mc \frac{dT}{dt} \quad (2.31)$$

**COMSOL Multiphysics** COMSOL Multiphysics is simulation software to give engineering solutions. In this thesis, the heat transfer module is used to compare the results from COMSOL Multiphysics with the suggested Matlab model. COMSOL Multiphysics is software based on numerical methods and is computationally more demanding than Matlab for the same simulation time. The basis for this module of heat transfer is the law of conservation of energy. The manual of COMSOL states that this is called the 'Heat Balance Equation': Equation 2.32 [44].  $E_{\Omega}$  is the internal energy,  $Q_{exch}$  is the exchanged heat rate and  $P_{str}$  is the stress power, converted into heat by dissipation. Furthermore, COMSOL takes the conservation of mass, linear momentum and angular momentum into consideration.

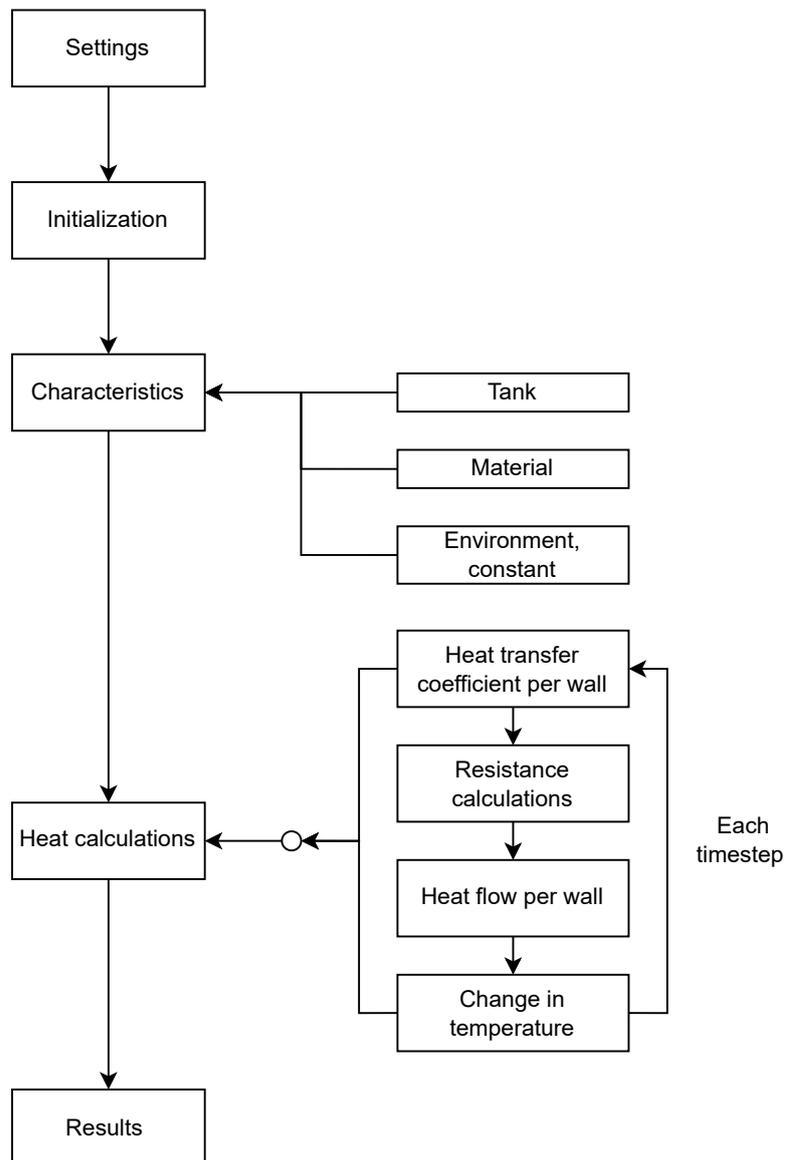


Figure 2.5: Matlab flow CDM

$$\frac{dE_{\Omega}}{dt} = P_{str} + Q_{exch} \quad (2.32)$$

### 2.1.3. Chemical tanker model

This section is the expansion of the CDM towards a trip that has been sailed with a chemical tanker vessel. During this trip, data is recorded, leading to a comparison between the simulated trip and the reported trip. The same method is used as explained in Figure 2.5, with some small adjustments. This can be seen in Figure 2.6. There are two coloured boxes, corresponding to the two modifications in the model. The first modification is that the environmental temperature is not considered to be constant, but changes over time. Furthermore, the vessel is surrounded by two different materials. The materials surrounding a vessel are the sea water and the air. The second aspect is that the tanks do not only cool down to the environmental temperature, but must be maintained at a certain temperature. For each timestep, the cargo tank temperatures are checked against the minimum allowed temperature. Most tanks have a higher temperature than the sea and air temperature, so cooling down occurs by transmitting energy (and thus reducing the temperature) to the sea and air. The heating spirals in the tanks

heat up the cargo if necessary. When the cargo tank temperature drops below the minimum temperature for a tank, the model turns on the boiler and energy is added to the tank for that specific moment. The energy is then transferred towards a fuel consumption. This results in the fuel consumption for the cargo heating. When this is done for each timestep in the simulation, the total fuel consumption for the trip can be calculated. The PBM could also be used to calculate the influence of insulation during a trip. Insulation could be modelled as shown in Figure 2.4.

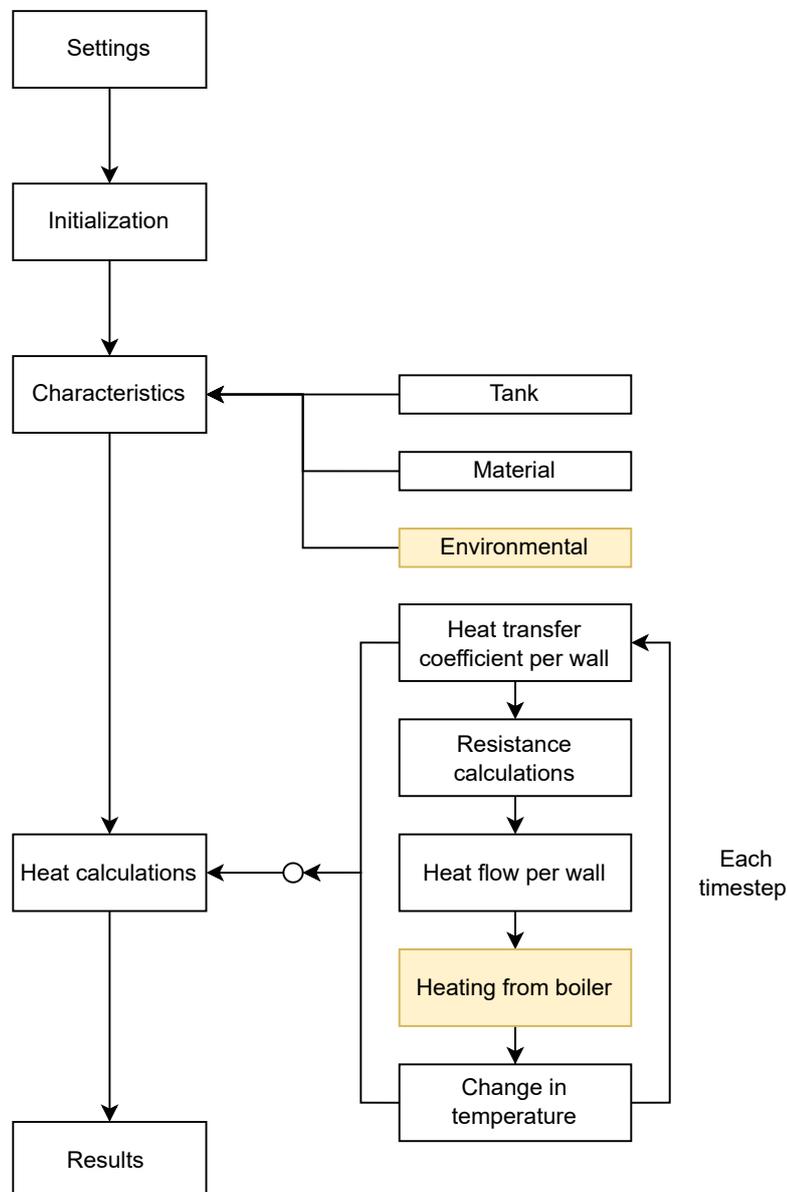


Figure 2.6: Matlab flow CTM

Magazinovic [45] performed a CFD analysis on a cargo tank for crude oil. The article states that the fluid inside a cargo tank is moving due to the movement of the vessel. The CFD analysis shows that the heat transfer is higher than only the result of natural convection, and a substantial part is caused by forced convection. Magazinovic analysed six hours of the temperatures within one cargo tank. This simulation had a duration of 28 hours while using an 8 cores 2.4 GHz processor. Wu [46] analysed a tanker trip during a voyage as well. The vessel during this trip was loaded with crude oil. Here, it was possible to analyse the forced convection. Wu solved the issue of forced convection by correcting all the Nusselt numbers.

To simplify this method, and since not all geometries are fully known, an approximation of the cor-

rection factor of the heat transfer coefficient is used ( $\alpha$ , Equation 2.33). This is suggested by the book of Jiji [37]. Jiji states that the difference between natural and forced convection is a factor between 2.5-30. Jiji states that the heat transfer coefficient normally is a big uncertainty, even when calculated correctly, causing errors up to 30% in the outcome. There are two items to check the correctness of the heat transfer coefficient: the behaviour of the temperature inside the tanks and the used fuel consumption. This leads to Equation 2.33. For this reason, convection will be modified during this thesis.

$$Q_{net,convection} = \alpha * h * A * (T - T_{\infty}) \quad (2.33)$$

#### 2.1.4. Verification

The first step in the verification of the model is to compare the CDM with a numerical simulation from COMSOL Multiphysics. These two methods should give a similar result, since these models are calculating the same cooling down behaviour of the box. Small deviations could occur, since both models are numerical methods. When the behaviour is similar, the Matlab method can be scaled to a chemical vessel model. The second step is thus to compare the simulation results for the trip with the reported fuel consumption during the trip. Furthermore, it is possible to compare the calculated tank temperatures with the reported actual tank temperatures.

#### 2.1.5. Application of physics-based model: Insulation

An important aspect of a physics-based model is that the effect of insulation can be calculated. The geometry of the vessel has to be modelled in the PBM. This gives the freedom to change the geometry of the vessel in order to see what effect this has. This is especially useful when insulation is considered. Figure 2.4 shows that insulation could be added as a wall with different properties than the existing steel wall. By calculating the effect of insulation, it is possible to build a business case to calculate the savings in US dollars to see if it is beneficial to insulate tanks. Furthermore, the PBM could be used when designing new vessels to see if the savings due to the insulation outweighs the cost of installing the insulation. Overall, the ability to model the effect of insulation is a key strength of the physics-based model approach. It provides designers with the flexibility to explore different design options and optimize vessel performance for improved energy efficiency and safety.

## 2.2. Data-driven model

As explained earlier, the second method is a DDM. First, the learning models are explained, where the algorithms used are explained in more detail. This is followed by the selection of the hyperparameter. After the selection of the hyperparameter, the setup for extrapolation is explained.

### 2.2.1. Algorithm selection

**Learning models** Ayodele [47] explained that there are six common algorithm types within machine learning. These types are: (1) supervised learning, (2) unsupervised learning, (3) semi-supervised learning, (4) reinforcement learning, (5) transduction and (6) learning to learn. The aim of supervised learning models is to create a function that is able to map the input to the output. This research uses the toolbox from Scikit-Learn [48] for the algorithms. Scikit-Learn is a toolbox where commonly known algorithms are programmed and can be loaded into Python. Multiple algorithms from the Scikit-Learn toolbox are suited to train a model to predict the fuel consumption. Among the possibilities are the least squares method, support vector regression (SVR), neural network (NN) and random forest regressor (RFR). Each model has certain benefits and disadvantages. The least squares method is suitable for handling big data, and is able to create both linear and non-linear models. SVR searches for an approximation function [49] and a NN needs training data, test data and validation data to ensure that the NN is able to predict the outcome with reasonable accuracy (normally around 5%) [50]. The RFR builds multiple decision trees on different subsamples of the entire dataset, and then averages these results [51]. Hence, the least squares method and the support vector regression were chosen out of these possibilities.

During this study, a linear and nonlinear algorithm are used and compared. The chosen linear algorithm is called RigdeCV, and the nonlinear algorithm is SVR from the Scikit-Learn toolbox. RigdeCV is first further explained and followed by an explanation of SVR. Both algorithms have the same order of processing. First, the range for error coefficient has to be determined. Then, the model is trained

on a subset of the data with the help of k-fold. The next step is to determine the error coefficients. With these coefficients, a model can be determined and applied onto a test set to be able to see the accuracy of the model.

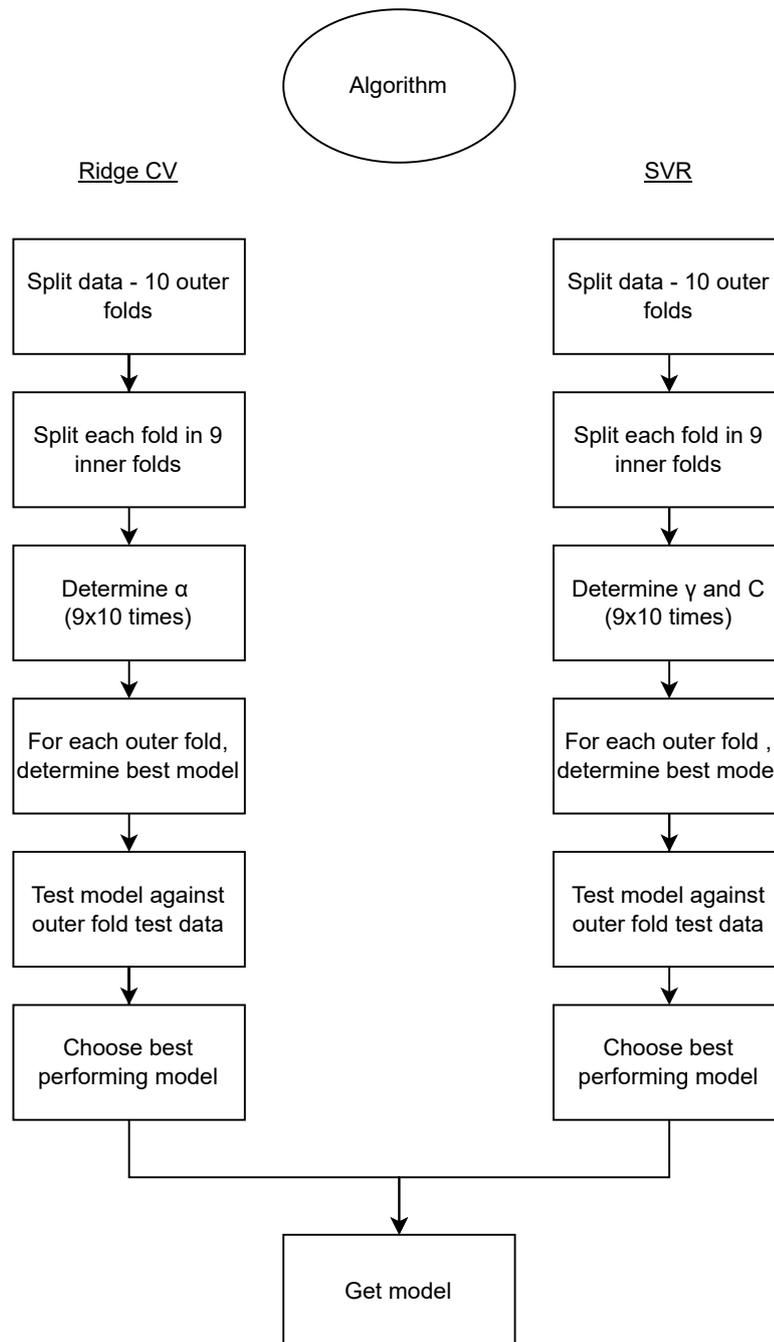


Figure 2.7: Build up of algorithm, own figure

### Linear Model - RidgeCV

The linear model used during this study is in the form of Equation 2.34. In this equation,  $\hat{y}$  is the predicted value,  $\vec{w} = (w_1, \dots, w_n)$  are the coefficients for this model, and the input parameters are  $\vec{x} = (x_1, \dots, x_n)$ . With the help of Equation 2.35, the error will be minimized and the coefficients  $\vec{w}$  are determined. This equation shows the method of ordinary least squares (OLS). In this equation,  $\alpha$  is

the complexity parameter. This parameter determines the tradeoff between matching the training data and the test data. On one side, the model aims at fitting the training data. On the other hand, a model aims to predict a value which is not in the training data. Matching the training data too precisely causes overfitting. In all cases the following statement must be true:  $\alpha \geq 0$ . When  $\alpha$  is very small, the solution goes towards OLS. When  $\alpha$  is very large, the regularization effect dominates the loss function. The implemented algorithm RidgeCV from Sklearn [52] determines the best performing  $\alpha$  for the model.

$$\hat{y}(w, x) = w_0 + w_1x_1 + \dots + w_nx_n \quad (2.34)$$

$$\min_w \|Xw - y\|_2^2 + \alpha \|w\|_2^2 \quad (2.35)$$

### Non Linear Model - Support Vector Regression

Support vector regression (SVR) is based on support vector machine (SVM). The SVR is stated as an optimization problem [53]. SVR is a method that is well suited to predict an outcome based on certain parameters. These parameters create a dataset. This dataset is used to train the model. Drucker [54] has researched that SVR has advantages over regression trees and ridge regression. The main advantage is that SVR optimization is not dependent on the dimensionality of the input space. Vapnik [55] explains that SVR uses a kernel basis function to train the model. The kernel is the mapping of low dimensional data into higher dimensional data [56]. The kernel basis function RBF is stated in Equation 2.36. The objective function and boundary functions for SVR can be seen in Equation 2.37 [53]. The schematic architecture of an SVR can be seen in Figure 2.8.

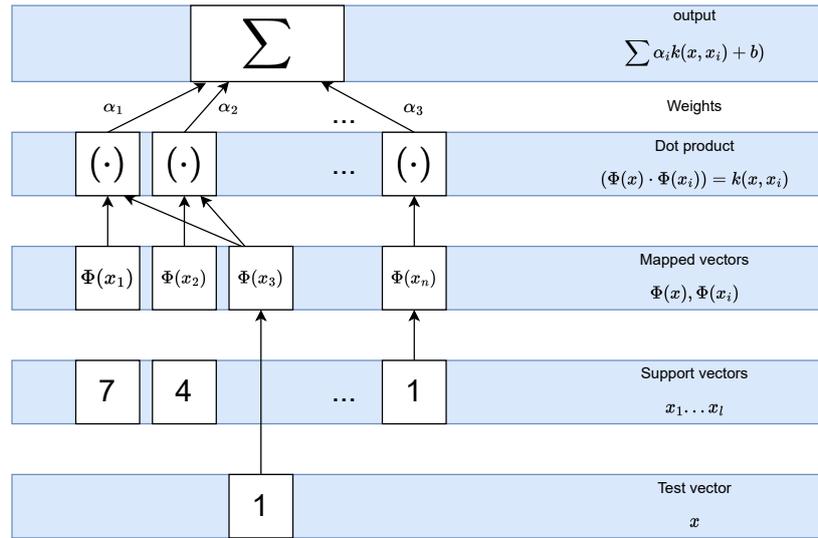
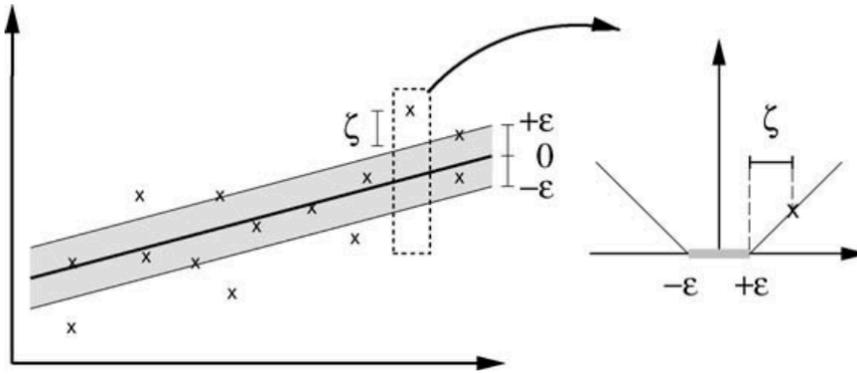


Figure 2.8: Schematic overview of SVR, based on Smola 2004 [2]

$$K_\gamma = \exp(-\gamma|x - x_i|^2) \quad (2.36)$$

$$\begin{aligned} \text{Objective: Minimize} \quad & \min_{w, b, \zeta, \zeta^*} \frac{1}{2} w^T w + C \sum_{i=1}^n (\zeta_i + \zeta_i^*) \\ \text{Subject to} \quad & \begin{cases} y_i - \langle w, x_i \rangle - b \leq \epsilon + \zeta_i \\ \langle w, x_i \rangle + b - y_i \leq \epsilon + \zeta_i^* \\ \zeta_i, \zeta_i^* \geq 0 \end{cases} \end{aligned} \quad (2.37)$$

Figure 2.9: Influence of  $\epsilon$ , from [2]

With the optimization problem and boundaries defined, it is now possible to construct an Lagrange function based on Equation 2.37. The result of this is a dual optimization problem, stated in Equation 2.38.

$$\begin{aligned} \min_{\alpha, \alpha^*} \quad & \frac{1}{2}(\alpha - \alpha^*)^T Q(\alpha - \alpha^*) + \epsilon e^T(\alpha + \alpha^*) - y^T(\alpha - \alpha^*) \\ \text{subject to} \quad & e^T(\alpha - \alpha^*) = 0 \\ & 0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, \dots, n \end{aligned} \quad (2.38)$$

Equation 2.38 can be reformulated using a kernel trick. The result is stated in Equation 2.39, where  $\alpha_i$  and  $\alpha_i^*$  are the Lagrange Multipliers and  $K$  refers to the kernel function.

$$\sum_{i \in SV} (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (2.39)$$

The book written by Vapnik [55] explains the fundamentals of SVR in detail. There are two hyperparameters which have to be chosen:  $C$  and  $\epsilon$ .

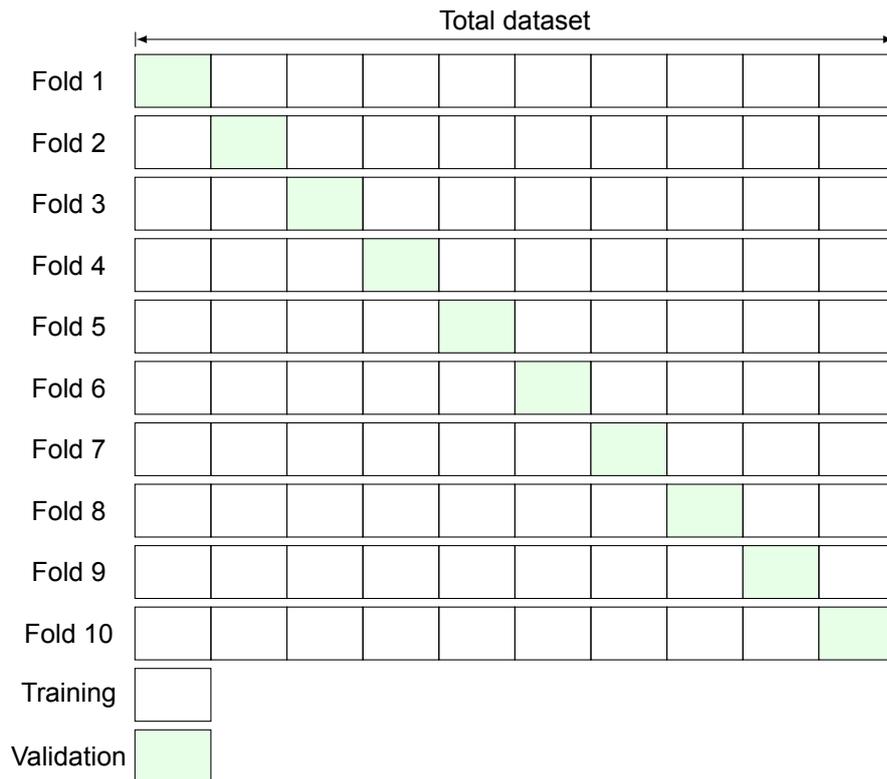
The parameter  $C$  determines the tradeoff between the misclassification of the training examples and the correctness of the fit in the training dataset. A larger  $C$  leads to an increase of support vectors and therefore more complexity into the model. A smaller  $C$  causes less support vectors and thus increases the flatness of the model [57]. The parameter  $\epsilon$  is an indication for the influence of a single value [2]. The influence is visualised in Figure 2.9. A larger value for  $\epsilon$  increases the flatness of the model, because only a small number of support vectors are determined. A smaller value for  $\epsilon$  increases the number of support vectors [58]. How these parameters are selected, is explained in subsection 2.2.2.

### 2.2.2. Hyperparameter selection

Cross validation of a model is in place to be able to measure the performance of a prediction model. The prediction model should not only perform well on the data it is trained against, but also on new data. This causes the need for cross validation. This research uses a K-fold cross validation due to limited data availability. K-fold cross validation is a method to train a model against all data and test the model against all data. The idea of K-fold cross validation is that a dataset is split in  $k$  subsets, as explained by Berrar [59]. Every time the model uses  $k - 1$  subsets of the data to train the model and 1 subset of the data is used to test the model. The K-fold cross validation method is visualised in Figure 2.10. Simon [60] explains that a 10-fold cross validation is a special case of the K-fold cross validation. Figure 2.10 shows the 10 folds. Each fold contains 9 subsets of the training data. Both chosen models (*RidgeCV* and *SVR*) require a search to select the hyperparameter. This search for the hyperparameters is done in each individual fold. In the fold, 9 models will be made and the best model with corresponding hyperparameters is chosen to test against the validation set. This means that the inner loop is always trained with 80% of the data. The result is that there are 10 models selected, one for each inner fold. These models are tested against the validation data and the best performing model

is the chosen model. Molinaro [61] studied heavily biased datasets and researched the results when a K-fold method was applied. The K-fold method gives k-amount of scores, namely each fold has a model with a score. The average of the scores gives a good estimation for the total score states the research of Molinaro.

Figure 2.10: K-fold cross validation, own figure



### 2.2.3. Extrapolation

The K-fold method enables the model to be trained and tested against the data multiple times. When the result of this model is sufficient, it means that the model should be able to interpolate between these values. It could be beneficial to test the performance of the model when it has to calculate extreme values. This could be beneficial, for example, when the average cargo temperature is higher than usual or the sea or air temperature is higher or lower than what has ever been experienced. Figure 2.11 shows how this could be done. The top 10% for one of the parameters is taken out of the dataset. The model is trained with a subset of the data and then, the model is checked against the extrapolation dataset.

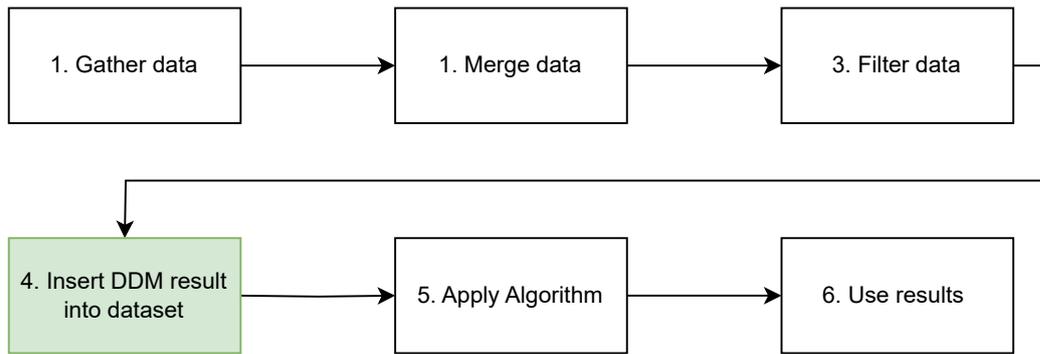


Figure 2.12: HM schematics

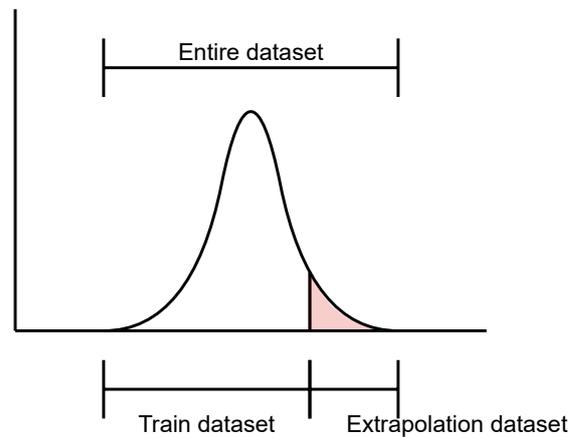


Figure 2.11: Extrapolation method

## 2.3. Hybrid model

The idea of the HM is that it uses the knowledge of both the BBM and the WBM and merges it into one model. This results in a model that benefits from both the DDM and the knowledge from the PBM. This results in a model which should be more accurate, and should be able to be based on less data [62], [63]. The steps to be able to create this model are displayed in Figure 2.12. The algorithm selected by the HM is the same best performing algorithm from the models of the data-driven model. The dataset from the DDM is combined with data from the PBM. This results in a new dataset. This dataset is trained and tested in a same manner as the DDM.

## 2.4. Error Estimation data-driven model and hybrid model

A study by Botchkarev [64] shows the opportunities for error estimation in machine learning regression algorithms. Botchkarev mentions that there are over 54 metrics for error estimation. The most common metrics for error estimation are: mean absolute error (MAE), mean absolute percentage error (MAPE), mean square error (MSE), root mean square error (RMSE) among others. The formulae for these error estimations are Equations 2.40, 2.41, 2.42 and 2.43. In these equations,  $\hat{y}_i$  is the predicted value for item  $i$ ,  $y_i$  is the actual value for item  $i$  and  $n$  is the amount of the predictions. The MAE and RMSE are in the same order of magnitude as the target variable, while the MSE is the difference squared. The MAPE is a percentage. A study by Fildes [65] shows that the two most used metrics for estimations are the MAE and the MAPE. For this reason, this research uses the MAE and MAPE to be able to compare it with most available research.

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (2.40)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{y_i} \quad (2.41)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|^2 \quad (2.42)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|^2} \quad (2.43)$$

### 2.4.1. Interval of confidence

As explained in subsection 2.2.2, each fold generates a Confidence Interval (CI). The CI is the interval in which a certain amount of the answers are. For example, a CI of 95% means that the answer is for 95% within this range. The CI can be calculated according to Equation 2.44 [66]. In this equation,  $z_{\alpha/2}$  is the normal distribution. The value for 95% confidence can be read from a t-distribution table and the according value is 1.96.  $\sigma^2$  is the variance, thus  $\sigma$  is the standard deviation and  $n$  is the sample size. The paper by Šimundić [67] explains in more detail how to calculate the CI. The results have the form of mean result  $\pm$  CI.

$$CI = (\bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}) \quad (2.44)$$

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2} \quad (2.45)$$

# 3

## Feature description

This chapter describes all the features needed to run the simulations. Chapter 2 explained the three different methods: PBM, DDM and HM. Section 3.1 explains the features that are needed for the PBM. These features are the layout of the reference vessel and the conversion of fuel consumption. Section 3.2 explain the features that are needed for the DDM. The HM used the information of both the PBM and the DDM, so no additional features are needed for this model.

### 3.1. Features physics-based model

As stated in the research question, a specific voyage and vessel have to be considered to calculate the fuel consumption and temperature behaviour in a cargo tank. The trip considered was from Port Kelang, through the Suez canal and towards Rotterdam. This trip had a duration of 29 days, where no cargo was loaded or discharged. The sea temperature during this trip was between 12 to 30 °C and the air temperature was between -3 to 27 °C.

During the voyage, there were 40 tanks filled with cargo. The four deck tanks were empty, and were accordingly not taken into consideration in the model. The cargo specifications can be found in Appendix C.

#### 3.1.1. Layout of reference vessel

As explained in subsection 2.1.1, the physical dimensions are essential to calculate the heat transfer between cargo tanks. The schematics of these two vessels are displayed in Figure 3.1. This schematic overview contains multiple spaces. There are cargo tanks numbered from 1 to 12 with the additions of P (Port), S (Starboard), C (Center), CA (Center Aft) and CF (Center Forward). On deck are four deck tanks. On the outsides are 14 ballast tanks, starting with a W and numbered from 1 to 5. There are two tanks dedicated to compensate for heeling of the vessel. The tanks starting with a C are the spaces between cofferdams. The main particulars of the Stolt Concept and Stolt Effort are mentioned in Table 3.1. These tank names are transferred to numbers corresponding to each tank, shown in Figure 3.2. These numbers will be used in the thesis to name the tanks.

Table 3.1: Principle dimensions Stolt Concept and Stolt Effort

Length (LOA)	176.75	[m]
Breadth moulded	31.00	[m]
Depth moulded	15.60	[m]
Draught moulded	11.80	[m]
Displacement	49400	[t]
DWT	37000	[t]
Service speed	16.2	[kts]
Cargo tanks	44	[-]
Boiler efficiency	83	[%]

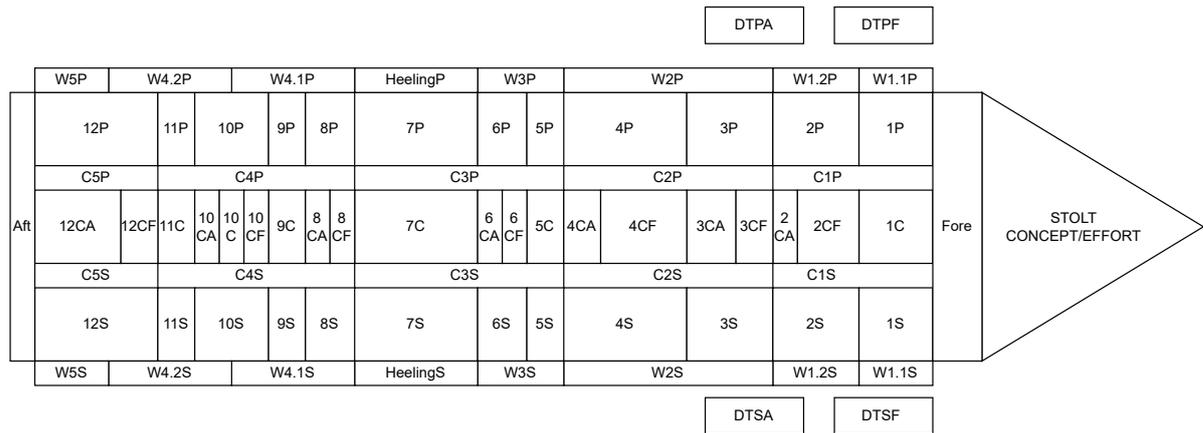


Figure 3.1: Schematic overview Stolt Concept and Stolt Effort

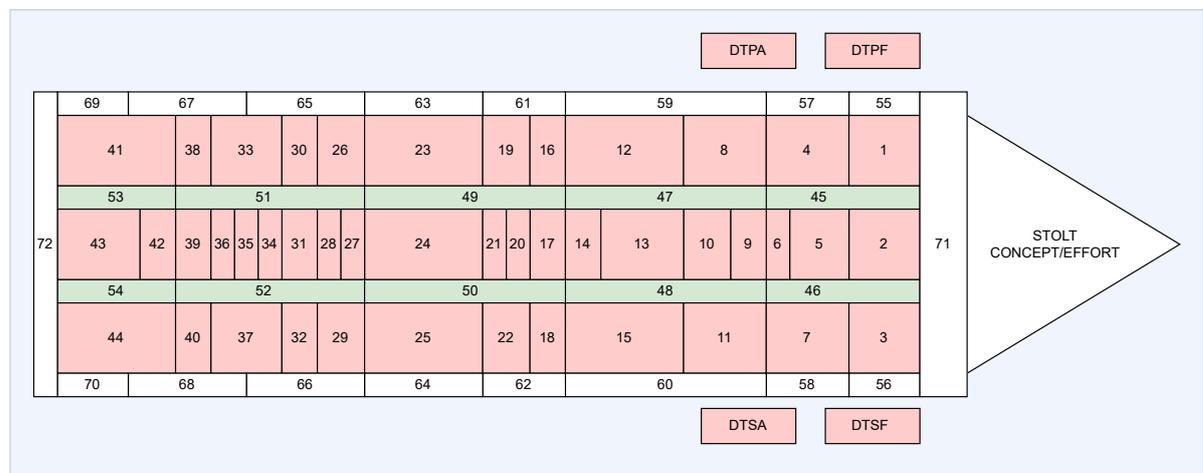


Figure 3.2: Schematic overview Stolt Concept and Stolt Effort, numbered tanks

### 3.1.2. Conversion of fuel consumption

The inflow and outflow per tank is calculated in *joules per time unit*. A paper by Jochemsen-Verstaeten [68] researched the lower heating value of the different fuel types. Heavy fuel oil (HFO) has a lower heating value of 41 *MJ/kg* and marine diesel oil (MDO) has a value 42 *MJ/kg*. This research calculates using the lower heating value of MDO, since vessels are not always allowed to use HFO. The Matlab model determines an energy requirement per tank in *Joule/seconds*. To be able to compare the fuel consumption as determined by Matlab, with the amount of fuel, this must be converted to *ton/hour*.

## 3.2. Features data-driven model

The DDM and the HM both depend on the available data. Subsection 3.2.1 explains the acquisition of the available data. This data is analysed in subsection 3.2.2 and subsection 3.2.3 explains what features must be included in the dataset in order to create the dataset for the models.

### 3.2.1. Dataset acquisition

Within Stolt Tankers there are multiple sources for the data. The 'NR data' is a data source where the vessel performance parameters are noted every day at noon. Another data source is the temperature sensors in cargo tanks. Stolt Tankers has five vessels that are equipped with sensors that provide a measurement every minute. Two of these five vessels are vessels from the same fleet. Each cargo tank is equipped with three sensors, one at the bottom, one at the top and one in the middle between these points. The signal that is registered is the average of these three sensors, and this is sent to shore. Another source of information is the cargo-specific information. This is the information about

what cargo is in each individual tank. The overview can be seen in Table 3.2. The chosen parameters within these datasets can be seen in Table 4.5. This is a selection of all available data fields, which can be found in Appendix A.

Table 3.2: Datatypes and location

Data	Location	Measurement Frequency	Nr of measurements
Noon report data data	PowerBI	Day	232
Sensor data	PowerBI	Minute	9440
Cargo information data	iStow	Trip	70

### 3.2.2. Preliminary Data Exploration

The aim of this research is to predict the fuel consumption for the cargo heating. The literature in subsection 2.1.1 explains that the main influences are the cargo masses, the temperatures within the tank, the air temperature and the sea temperature. With the filtered data, the fuel consumption for cargo heating for 150 consecutive days is plotted with the air and the sea temperature in Figure 3.3. There is no clear linear trend between the fuel consumption for cargo heating and the temperatures. There are peaks in the fuel consumption, while there is not a big change in either the sea or air temperature. There are even periods where the temperatures decrease and the fuel consumption decreases too. There are several possibilities for this. One reason could be that the cargo specifications allow a lower temperature than the current cargo temperature.

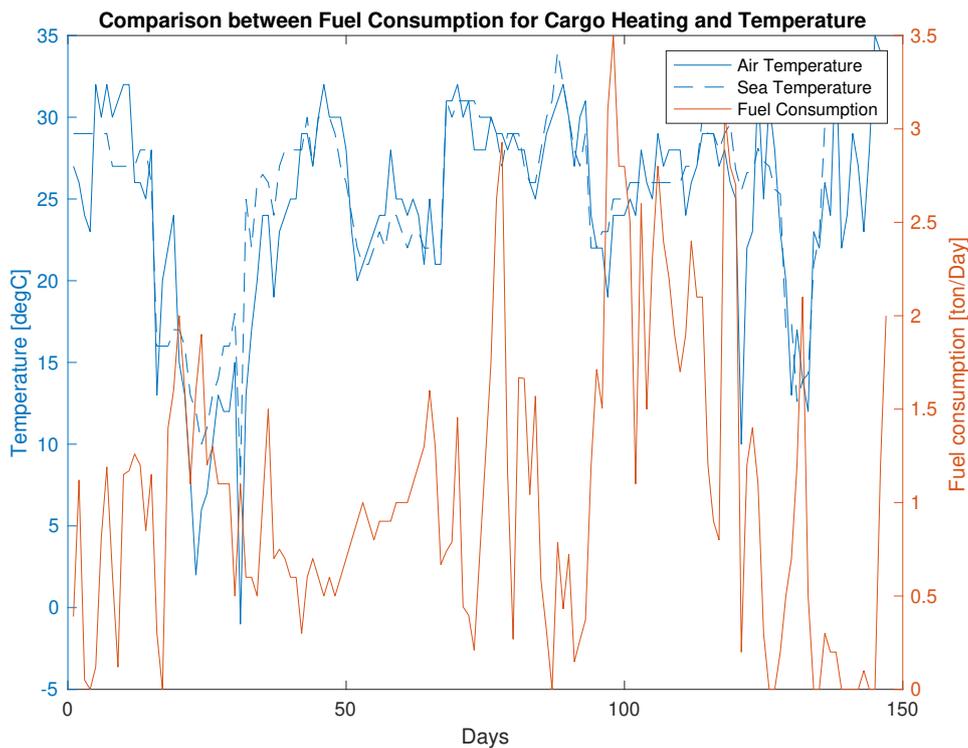


Figure 3.3: Comparison between fuel consumption for cargo heating and air/sea temperature

The second step is to look at the correlation between variables. Pearson’s Correlation Coefficient is suitable for this method. This coefficient has values between -1 and 1. When there is a strong positive linear relation between two parameters, the value of this coefficient approaches 1. When there is a negative relation between parameters, the value is below 0. To calculate these coefficients, Equation 3.1 can be used [66]. The covariance of parameters A and B must be calculated according to Equation 3.2 [66]. The covariance determines the dependency between two variables. The standard

deviation of a parameter can be calculated according to Equation 3.3 and is a measurement for the deviation from the mean of each measurement [66].

$$\text{Corr}(A, B) = \frac{\text{Cov}(A, B)}{\sigma_x \sigma_y} \quad (3.1)$$

$$\text{Cov}(A, B) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (3.2)$$

$$\sigma_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n |x_i - \bar{x}|^2} \quad (3.3)$$

### 3.2.3. Data selection

The aim of the model is that it should be able to predict the fuel consumption based on input parameters. Not all parameters mentioned in Appendix A are relevant to create a model. However, the selected parameters must be able to represent the vessel, environmental conditions and cargo conditions. The five categories where the data is coming from are (1) NR data, (2) temperature data, (3) pressure data, (4) minimum required temperature data and (5) maximum required temperature data. This is shown in Figure 3.5.



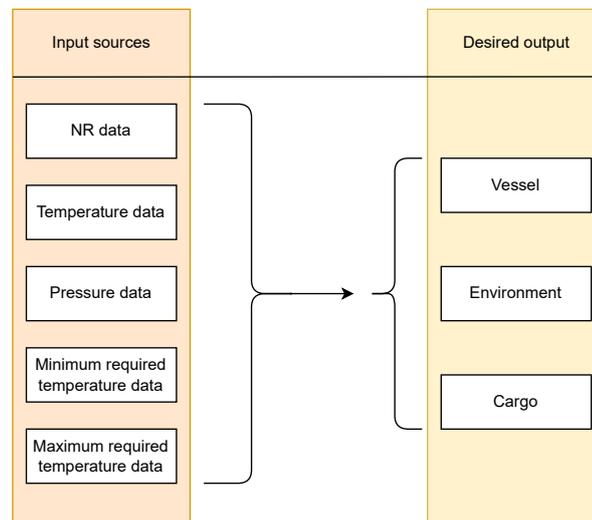
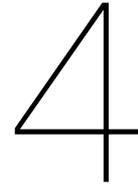


Figure 3.5: From input features to desired output of model



# Simulation setup

This chapter describes the setup for the simulations. First the PBM is discussed in section 4.1. The PBM contains information about the CDM and the CTM, followed by the DDM in section 4.2. The chapter finishes with section 4.3 for the HM.

## 4.1. Physics-based model

First, the setup for the CDM is discussed in subsection 4.1.1. This method is modified in subsection 4.1.3 to be able to simulate a trip for a chemical tanker. The PBM finishes with the experiments for the insulation in subsection 4.1.4.

### 4.1.1. Cooling down model

The CDM consists of 12 blocks and is shown in Figure 4.1b. A top view of this configuration is shown Figure 4.1a. The dimensions and main parameters for the CDM are stated in Table 4.1. The tanks in this geometry have either two or three tanks as neighbours. The length, width and height are three meters, so as to be able to create some volume. The walls of the tanks are 0.1 meters. The reason for these dimensions has to do with the meshing of the model. The mesh size depends heavily on the thickness of the walls. Due to computational power, the ratio between on one side the length, width and height and on the other side, the thickness, should somewhat be in the same range. The material of the walls is chosen to be stainless steel, as this material is stated in the DNV-GL regulations for Chemical Tankers [69]. This model is placed in an environment filled with air. The material properties for air can be found in Table 4.1. The last material involved in the system is the filling of the tanks. For the simulation, it is assumed that all tanks are filled with Vinyl Acetate. Vinyl Acetate is one of the materials that could be transported in a chemical tanker and is often transported.

Table 4.1: Test setup CDM

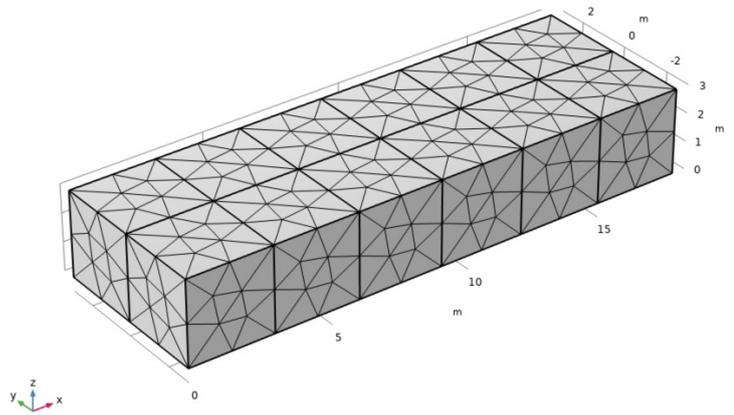
Tank amount	12	Length	3 [m]
Configuration	6*2	Width	3 [m]
Material in tank	Vinyl Acetate	Height	3 [m]
Material wall	Steel	Thickness	0.1 [m]
Material around tanks	Air	Volume	27 [ $m^3$ ]

### 4.1.2. Input for model

The environmental temperature is chosen to be 20 °C, and the cargo initial temperature is set at 35 °C. In one available trip, Vinyl Acetate is transported at 35 °C. The ambient temperature is set at 20 °C to be able to model the cooling down of tanks. A 15 °C temperature difference is a temperature difference that often occurs according to data from Stolt Tankers. The duration of the simulation has to be predefined too. The duration of the simulation is the duration of the time to cool the system down to 20 °C.

SB Aft Tank 1 Fwd PS	SB Aft Tank 2 Fwd PS	SB Aft Tank 3 Fwd PS	SB Aft Tank 4 Fwd PS	SB Aft Tank 5 Fwd PS	SB Aft Tank 6 Fwd PS
SB Aft Tank 7 Fwd PS	SB Aft Tank 8 Fwd PS	SB Aft Tank 9 Fwd PS	SB Aft Tank 10 Fwd PS	SB Aft Tank 11 Fwd PS	SB Aft Tank 12 Fwd PS

(a) Top view CDM



(b) Overview CDM Multiphysics

Figure 4.1: Overview of CDM

Table 4.2: Location of walls

1	Forward
2	Aft
3	Port
4	Starboard
5	Top
6	Bottom

### 4.1.3. Heat Transfer Model for a Chemical Tanker

**Trip conditions** As stated in the research question, a specific voyage and vessel have to be considered to calculate the fuel consumption and temperature behaviour in a cargo tank. The trip that has been considered is from Port Kelang through the Suez Canal towards Rotterdam. The chosen trip is the longest available trip where the vessel is almost completely filled with cargo and has a duration longer than 20 days without a stop in a port. This trip has a duration of 29 days, where no cargo is loaded or discharged. The sea and the air temperature over time are shown in Figure 4.2. During the voyage, there are 40 tanks filled with cargo. The four deck tanks are empty, and are therefore not taken into consideration in the model. The cargo specifications can be found in Appendix C.

### 4.1.4. Insulation

One of the solutions to reduce cargo heating consumption is to insulate the cargo tanks which would reduce heat losses. Currently, the tanks are not insulated. The question arises of how much fuel is saved when the cargo tanks are insulated. One of the benefits of the PBM is that it is able to calculate different physical geometries. The benefit is thus that insulation could be added into the PBM. There are specific places where insulation could be placed due to the existing geometry of the tanks. Some tanks are adjacent and share one wall. It is impossible to place insulation here because the outside of one

Table 4.3: Chemical properties of used materials in CDM

	T [K]	$\rho$ [kg/m <sup>3</sup> ]	$c_p$ J/kg * K	$\mu$ kg/(m * s)	k W/m * K	$\alpha$ m <sup>2</sup> /s
Air [17]	280	1.261	1006	$1.747 * 10^{-5}$	0.02473	$1.879 * 10^{-5}$
	290	1.217	1006	$1.795 * 10^{-5}$	0.02544	$2.078 * 10^{-5}$
	300	1.177	1007	$1.857 * 10^{-5}$	0.02623	$2.213 * 10^{-5}$
	310	1.139	1007	$1.889 * 10^{-5}$	0.02684	$2.340 * 10^{-5}$
Vinyl Acetate [70]		932	1926	$43 * 10^{-5}$	0.15	
Stainless steel UNS S31803 [71]	293.15	7800	500	-	15	$13 * 10^{-6}$
Sea Water [72], salinity = 30 [g/kg]	293.15	1021.1	4060.2	$1.065 * 10^{-3}$	0.602	$1.46 * 10^{-7}$

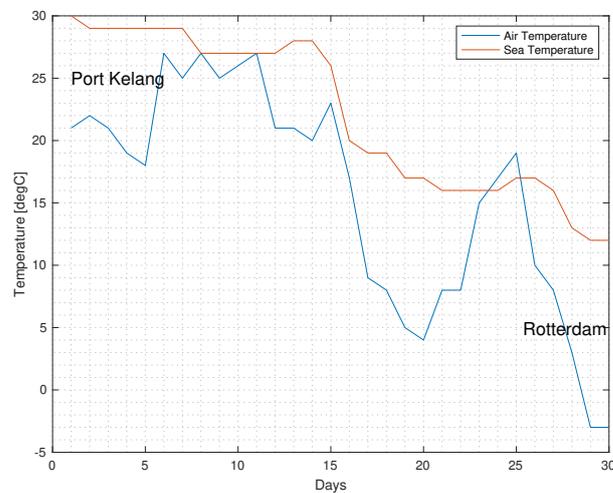


Figure 4.2: Sea and air temperature during the simulated voyage

tank is the inside of another tank. This is shown in Appendix E, Figure E.1. In this figure, the red lines are where insulation is added for the simulation. Furthermore, the bottom and top of the tanks could also be insulated. Within Stolt Tankers, different materials are considered to insulate the cargo tanks. These materials are tested to see the benefits of each individual material. Different materials have a different thermal conductivity. Furthermore, these materials all have a range of different thicknesses at which they could be applied. The maximum and minimum thicknesses are displayed in Table 4.4.

Table 4.4: Different insulation materials

Insulation material	Thermal conductivity [W/mk]	Thickness [mm]
No insulation	-	0
nC ProTherm CUI-HTF [73]	0.04	4 15
ICU Isollat 2 [74]	0.002	1 4
Charcoat CTI [75]	0.001	0.4 4

## 4.2. Data-driven approach

This section explains the DDM. First, the merging of different datasets is explained, followed by the filtering process. After that, the data exploration is explained and it concludes with the data selection for building the model.

### 4.2.1. Merging data

The different data sources, mentioned in Table 3.2, all have a different time format. To merge the three datasets, each dataset has to be converted towards the same time interval. The three different measurement frequencies are: the duration of the trip, once every day and once every minute. The smallest time interval is thus minutes, and the largest is the trip. Only 200 days of data are available from two similar vessels. Since the vessel does not sail every day (some days it is in port), there are only 232 noon reports available since the installation of the sensors. To be able to leverage this out to a dataset that is larger than 1000 datapoints, it was decided to look at the data per hour. This means the NR data should be expanded from a measurement frequency per day towards a measurement frequency per hour. This expansion is possible because most parameters are constant for a certain day or can be estimated based on an hourly basis. The trip information is only measured per trip, but the good part about it is that these parameters are constant over time and it is thus allowed to expand them per hour.

### 4.2.2. Filtering data

In some cases the available data is faulty. A negative pressure or seawater temperature, for example, is not feasible and this should not be taken into consideration. Table 4.5 shows the allowed range for all measurements. There are two options for handling faulty data. The first option is to reconstruct the faulty data with an estimation based on the conditions of other parameters within the dataset. The second option is to remove the row in which this faulty data is. The second option was chosen because the use for the dataset is to predict values based on a dataset. This ensures that the model is trained on reported data and not on constructed data. When the model is trained, these datapoints might be able to be reconstructed, but they should be kept out of the training data. When applying the filters, the dataset reduces from 2861 rows to 2837 rows.

### 4.2.3. Data exploration

Figure 4.3 shows the correlation matrix of the available data, where data is merged together to improve visibility. This figure contains the NR data, an average of the sensor data and the maximum-minimum cargo temperature. The maximum minimum temperature is the highest temperature of all lower boundary temperatures. The last row and column of the correlation matrices are the 'Total Cargo Heating Consumption'. This is the main parameter which needs to be predicted by the DDM. This is the reason why it is the last row and column in the matrix.

Since this parameter does not have a single strong linear relationship with one of the parameters, except boiler heating consumption, the chances are high that a linear model is not sufficient to calculate the cargo heating consumption. The correlation matrix shows that the parameters influencing the cargo heating consumption the most in a positive manner are the average cargo tank temperature and the average temperature of the three tanks with the highest temperature. If the temperatures increase, the cargo heating consumption increases. The figure also shows that when the main engine power increases, the cargo heating consumption decreases. This is expected, since the residual heat of the main engine is used for cargo heating consumption. Furthermore, the parameters about the weather have a higher correlation together. The correlation matrix shows that there is a negative correlation between the 'Main Engine Running Hours [hour]', the 'Main Engine KWHrs [kWh]', the 'Total Propulsion Consumption [tons]' with regard to the 'Total Cargo Heating Consumption'. This is because the residual heat from the main engine is used to heat the cargo tanks. When the main engine produces more energy for propulsion, more residual energy can be used to heat the cargo too. Furthermore, there is no clear correlation shown between the air and sea temperature with regard to the 'Total Cargo Heating Consumption'. This was not expected, since the equations from heat transfer suggest an influence of the temperatures on the 'Total Cargo Heating Consumption'.

Table 4.5: Filter values data

Filter item	Allowed range	Unit	Measurement frequency
Heading Degree	0-360	[Deg]	Day
Aft Draft	5-15	[m]	Day
Fwd Draft	5-15	[m]	Day
Observed Distance	0-480	[nm]	Day
Speed Through Water STHrs	0-17	[kts]	Day
Main Engine Running Hrs	0-25	[hour]	Day
Main Engine KWHrs	0-7000	[kWh]	Day
Total Propulsion Consumption	0-40	[ton]	Day
Boiler RunningHours	0-25	[hour]	Day
Total Boiler Consumption	0-5	[ton]	Day
Total CargoHeating Consumption	0-5	[ton]	Day
Air Temperature	-10-50	[°C]	Day
Sea Temperature	0-40	[°C]	Day
Sea Number	0-7	[-]	Day
Sea Height	0-7	[m]	Day
Sea DirectionDegree	0-360	[deg]	Day
Swell Number	0-7	[-]	Day
Swell Height	0-7	[m]	Day
Swell DirectionDegree	0-360	[deg]	Day
Beaufort	0-7	[-]	Day
Wind Direction Degree	0-360	[deg]	Day
IoT Temperatures	0-100	[°C]	Minute
IoT Pressures	0-50	[bar]	Minute
Minimum Cargo Temperature	0	[°C]	Sailing leg
Maximum Cargo Temperature	100	[°C]	Sailing leg

#### 4.2.4. Data selection for simulation

The information from the vessel is coming from the NR data in the form of features  $x_1 \rightarrow x_7$  from Table 4.6. The environmental conditions are represented by  $x_8 \rightarrow x_9$  and the cargo information is summarised in  $x_{10} \rightarrow x_{12}$ . The vessel data are the drafts, the main engine settings and the distance and speed that is achieved by these main engine settings. It is chosen to not select many weather conditions, because when more weather conditions are taken into account when building a model, more weather conditions must be given as input to predict the total cargo heating consumption. When more weather conditions have to be estimated in order to make a prediction, more uncertainty is entered into the model. Besides this, Bochenski [22] suggested that it is possible to calculate heat transfer for vessels only by taking the sea and air temperature into account as environmental conditions. It is best to create a model with as little as possible features, where not many uncertainties are.

#### 4.2.5. Normalize data

The next step in the simulation is the normalization of the data. Singh [76] states that algorithms work better with normalized data. The benefit of normalized data is that all parameters are in the same order of scale. This is especially the case if the ranges in the data significantly differ. Table 4.5 shows the range of all datafields. The range of the parameters varies from an order of  $10^1$  towards  $10^3$ . When normalized data is used, every parameter makes an equal contribution to the model.

Equation 4.1 shows the equation for normalizing the data, which is required for the two algorithms. This results in values that are between 0-1. The highest value becomes 1 and the lowest value becomes 0. Data can be normalized according to Equation 4.1.

$$X_{normalized} = \frac{X - \min(X)}{\max(X) - \min(X)} \quad (4.1)$$

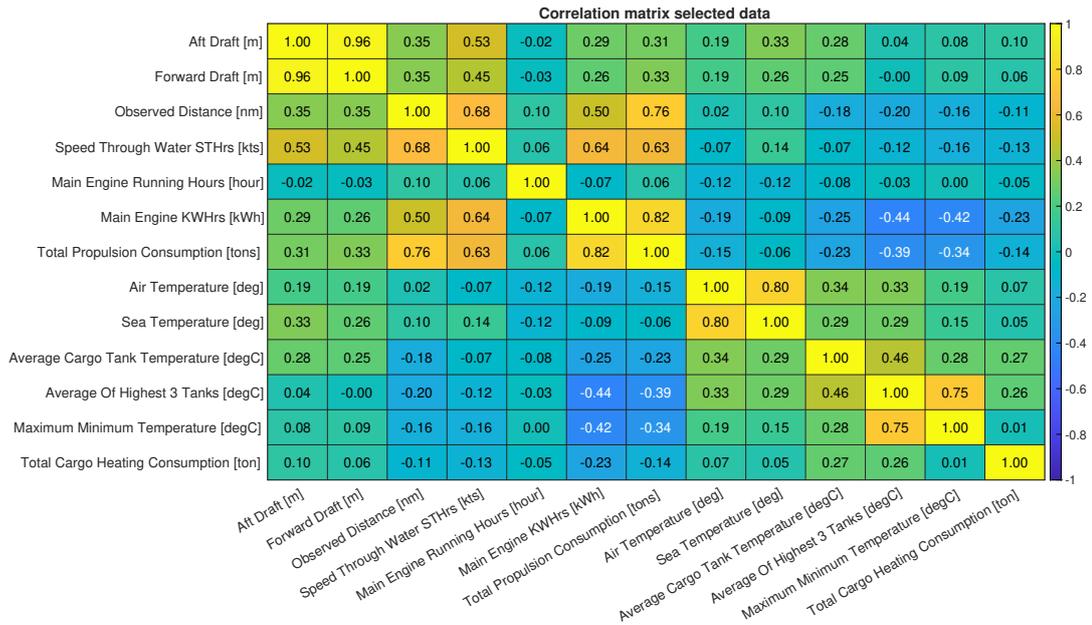


Figure 4.3: Correlation Matrix

Table 4.6: Chosen parameters for data-driven model

$x_1$	Aft Draft [m]
$x_2$	Forward Draft [m]
$x_3$	Observed Distance [nm]
$x_4$	Speed Through Water STHrs [kts]
$x_5$	Main Engine Running Hrs [hrs]
$x_6$	Main Engine Power [kWh]
$x_7$	Total Propulsion Consumption [ton]
$x_8$	Air Temperature [ $^{\circ}C$ ]
$x_9$	Sea Temperature [ $^{\circ}C$ ]
$x_{10}$	Average Cargo Tank Temperature [ $^{\circ}C$ ]
$x_{11}$	Average Cargo Tank Temperature of highest 3 tanks [ $^{\circ}C$ ]
$x_{12}$	Maximum Minimum Cargo Tank Temperature [ $^{\circ}C$ ]
$\hat{y}$	Total Cargo Heating Consumption [ton]

#### 4.2.6. Hyperparameter selection

Both the algorithms used in the DDM require a range for the hyperparameters. For the linear model, the hyperparameter that needs to be selected is called  $\alpha_{linear}$ . The SVR needs two hyperparameters as input. These hyperparameters are called  $C_{nonlinear}$  and  $\epsilon_{nonlinear}$ . The range for the hyperparameters is stated in Table 4.7. The range for the hyperparameters is a tradeoff between the computational time and the accuracy of the model. More hyperparameters in the same range give the model a better option to choose, however, all these hyperparameters have to be tested. The SVR uses two hyperparameters. This means that when the amount of points is increased from 10 points per hyperparameter to 20 points per hyperparameter, the computational time goes up by a factor 4.

Table 4.7: Hyperparameter range

	Start	Stop	Amount	Distribution
$\alpha_{linear}$	0.01	1	10000	Logarithmic
$C_{nonlinear}$	-4	2	10	Logarithmic
$\epsilon_{nonlinear}$	-4	2	10	Logarithmic

### 4.3. Hybrid model

The schematics of the HM are stated in section 2.3. It is shown that the only step that differs from the PBM is step 4 from Figure 2.12. Here, the results from the PBM are merged into the dataset from the DDM. This process can be seen in Figure 4.4. A subset of the dataset of the DDM is merged with the result from the PBM, and in this manner, the new dataset is created to train the HM. This dataset has one parameter more than the DDM. Less data is taken into account. This reduces the amount of data, however, since the PBM has a duration of 29 days. The DDM is built on 2837 rows, while the HM only has 575 rows to train the model.

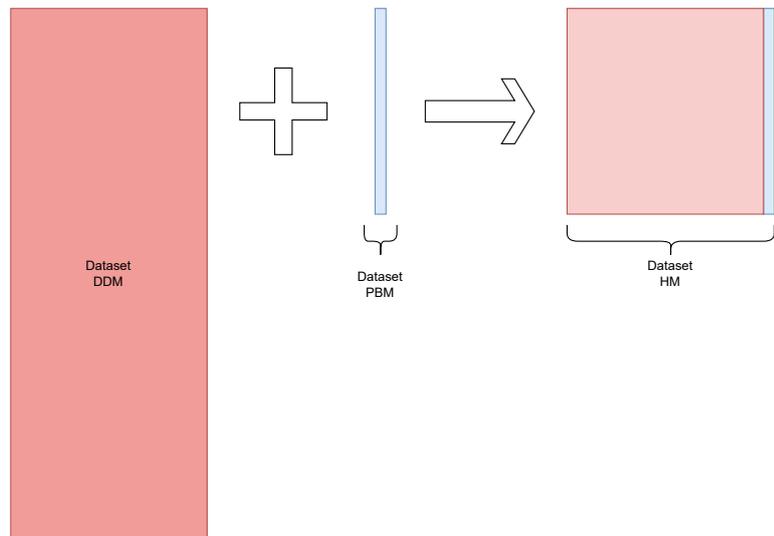


Figure 4.4: HM dataset creation



# 5

## Results

This chapter describes the results from the different models. First, the results from the PBM are described in section 5.1, followed by the results from the DDM in section 5.2. The chapter finishes with the results from the HM in section 5.3.

### 5.1. Physics-based model

This section explains the results from the PBM. First, the CDM is discussed in subsection 5.1.1, where both the Matlab model results are discussed as the results from COMSOL. This is followed by a voyage simulation of a chemical tanker. After this voyage simulation, the results of different type of insulation is shown.

#### 5.1.1. Result cooling down model

Figure 5.1 shows the results of the Matlab calculations. The figure shows that tanks 1, 6, 7 and 12 cool down the fastest, and that tanks 3, 4, 9 and 10 cool down the slowest. The reason for this is that tanks on the outside have four areas connected with the air and thus, cool down faster. It can be seen that the temperature decreases faster at the start of the simulation compared to the end of the simulation. A bigger difference between the tank temperature and the environmental temperature causes a more rapid decrease in temperature.

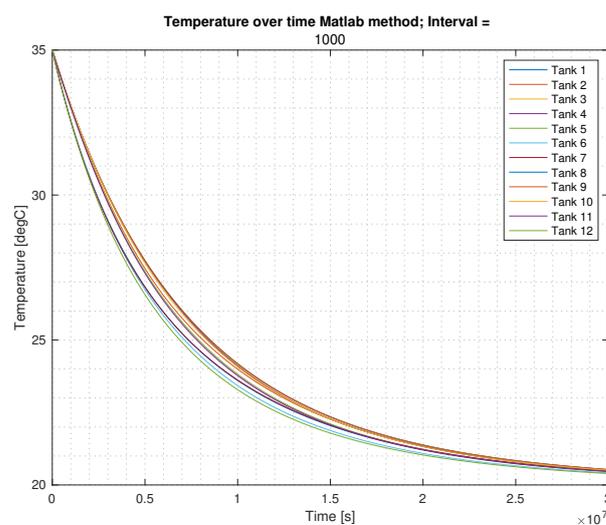


Figure 5.1: Matlab result CDM

### 5.1.2. Mesh refinement and timestep refinement

COMSOL requires the user to select both the mesh size and the timestep for the calculations. The timestep is the time period between two results. The result from COMSOL depends on these two choices. The influence of the mesh size can be seen in Figure 5.3. Furthermore, both the mesh size and the timestep have a high influence on the computational effort of the calculations. A smaller mesh size creates smaller volumes for which the changes are calculated. This means that the computational time increases, since the total amount of volume increases. When the timestep decreases, the computational time increases too, since there are simply more moments to calculate. This can be seen in Table 5.1. On a small scale, the mesh size matters. This is shown in Figure 5.3. In this simulation, all input parameters remain constant, except the mesh size. Meshes *Extra Coarse* and *Extremely Coarse* are the lowest two lines and overlap with each other. The mesh *Coarser* and *Coarse* overlap each other too. The figure show that the finer the mesh becomes, the slower the system cools down. Since the change between meshes *Normal Fine*, *Finer*, *Extra Fine* and *Extremely Fine* is not significant in Figure 5.3, calculations in COMSOL will use the normal mesh. The normal mesh is the fastest of these three meshes and the results are similar.

Table 5.1: COMSOL computational time

Mesh size	Time [s]	Interval	Time [s]
Fine	31951	t = 1000	11156
Normal	11156	t = 5000	2349
Coarse	5701	t = 10.000	1312
Coarser	4033	t = 100.000	195
Extra Coarse	2943		



Figure 5.2: Timestep and interval show

### 5.1.3. Results Matlab versus COMSOL Multiphysics

Figure 5.4 shows the cooling down of the simulation in COMSOL and the simulation in Matlab. It becomes clear both COMSOL and Matlab follow the same cooling down trend when the same boundary conditions are applied. Figure 5.5 shows that the COMSOL result suggests a bigger drop at the initial start. The Matlab model is a bit behind, especially the tanks that are located in the corners. This causes an absolute and relative difference at the start. When looking at Figure 5.4, it can be seen that the absolute difference is caused by a small shift in the horizontal direction (thus in time). The outcome from the Matlab simulation deviates in absolute percentages less than the deviation caused by changing the mesh size of COMSOL. This means it is thus possible to continue with this model.

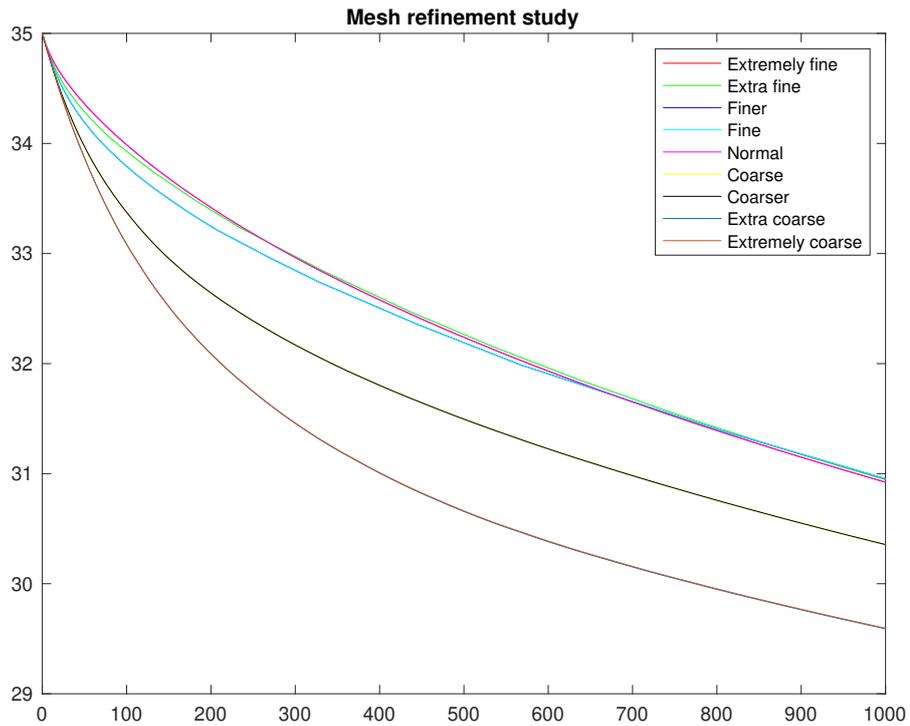


Figure 5.3: Mesh comparison COMSOL

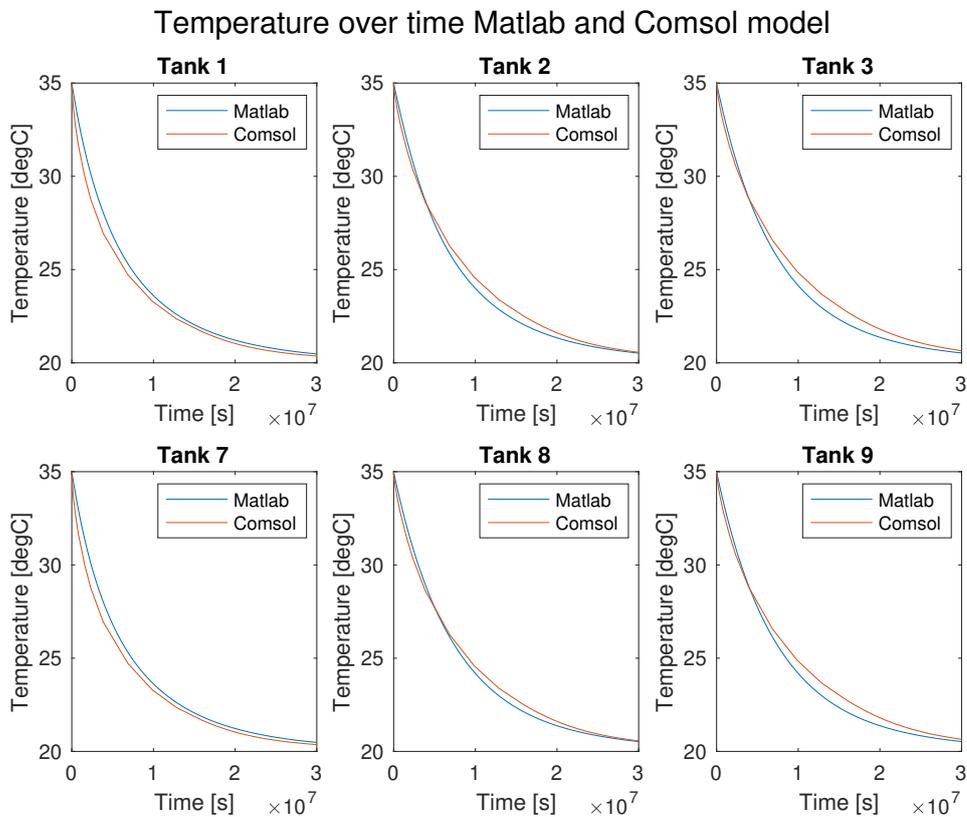


Figure 5.4: Results cooling down model Matlab and COMSOL

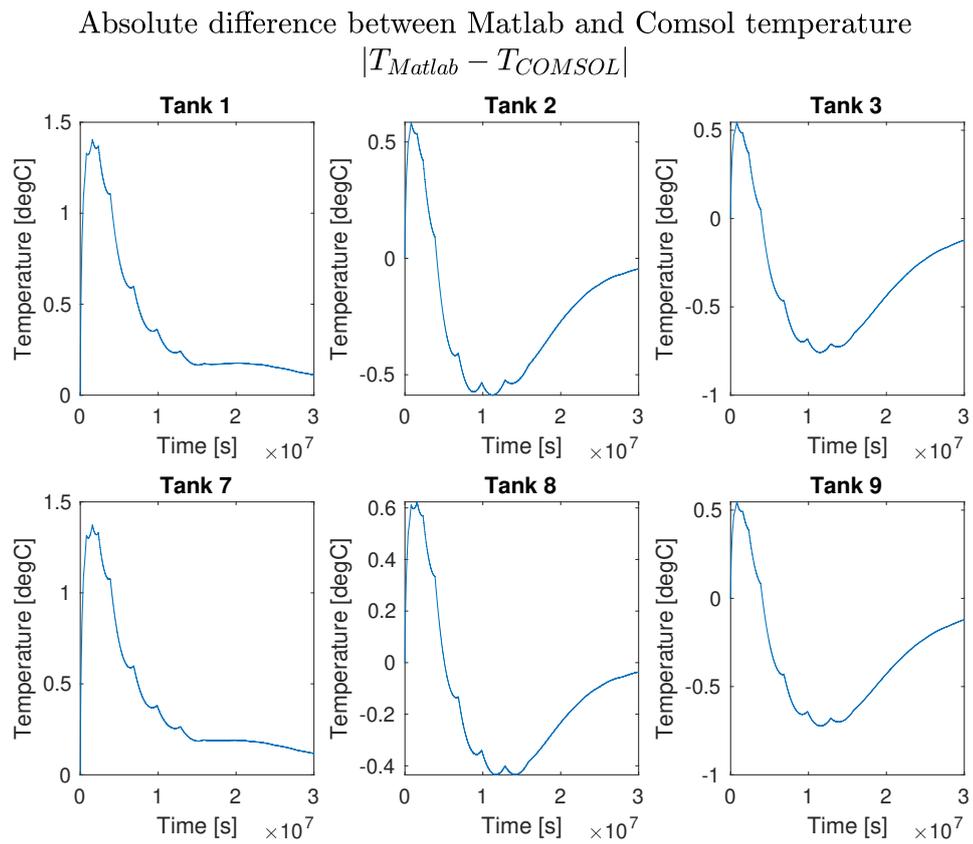


Figure 5.5: Absolute difference between Matlab and COMSOL

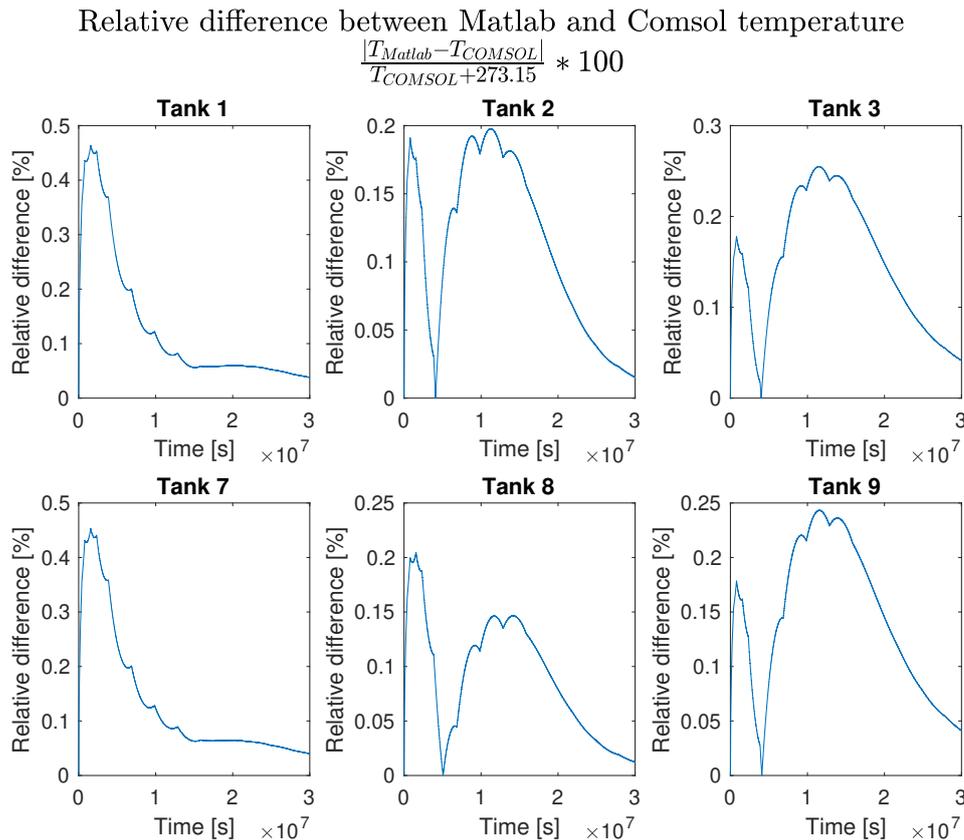


Figure 5.6: Relative difference between Matlab and COMSOL

#### 5.1.4. Chemical tanker model results

Figure 5.7 shows the cargo heating consumption for the simulated trip and the reported cargo heating consumption during this trip. The result from the PBM is that the vessel uses 0.813 tons of fuel per day for cargo heating consumption. In contrast, the NR data reported 0.77 tons per day for cargo heating consumption. This is a deviation of 5.6%. The PBM suggests a more steady consumption, while the reported data shows that fuel consumption varies heavily over time. There are days when the fuel consumption is around two tons per day, whereas the next day there is no consumption. The PBM suggests a more steady decrease in fuel consumption: lower peaks and no shutdowns in fuel consumption. Appendix D shows the temperature behaviour for each individual tank.

The model controls the input for cargo heating consumption, so it is logical that the heating of the cargo follows the same trend. However, the cooling down of the model is not regulated by external controls and gives a better indication of the behaviour of the model. Figure D.4 in Appendix D shows the boxplots of the difference between the predicted temperature and the reported temperature. Tanks 5, 9, 10, 13, 20, 27, 35, 36 and 42 do not have a boxplot. Tanks 5, 9, 13, 20 and 27 do not have measurements to compare against, and tanks 10, 35, 36 and 42 are empty and are thus not considered to be accurate.

The average difference between the considered actual temperature and the predicted temperature is  $0.6860\text{ }^{\circ}\text{C}$ . The boxplot shows that the median of the temperature difference per tank is normally close to 0. At tanks 8, 33, 34, 37 and 38 there is a larger difference in temperature between the reported and simulated temperature, namely close to  $1\text{ }^{\circ}\text{C}$ . These tanks are all above the actual reported temperature. Tank 8 cools down faster than simulated. There is no clear indication of why this occurs. Tank 33, 34 and 37 show that they all cool down according to the reported trend, however, there is a delay at the beginning of the simulation. This delay causes a temperature difference in the tanks, and after the delay the trend is similar. Tank 38 shows significant peaks and drops in the measurements, which is not followed by the model. This could be caused by faulty measurements, or faulty cargo specifications which are fed into the model.

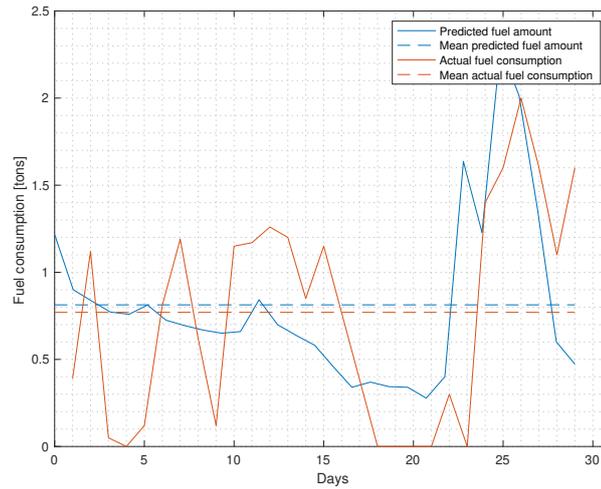


Figure 5.7: Predicted and reported fuel consumption

### 5.1.5. Chemical tanker insulation

As explained in subsection 4.1.4, insulation is beneficial to reduce the fuel consumption for cargo heating. Different types of material could be used for insulation and six types of insulation are simulated. The results of applying insulation can be seen in Table 5.2. All insulation reduces the need for cargo heating, depending on the thickness and the property of the thermal conductivity, different results are achieved. The insulation material that reduces the fuel consumption best is Charcoat CTI when the maximum allowed thickness of insulation is applied. The fuel consumption for cargo heating could theoretically be reduced from 0.813 to 0.2778, a reduction of 64.8%. The fuel consumption demand during the simulated voyage can be seen in Figure 5.8.

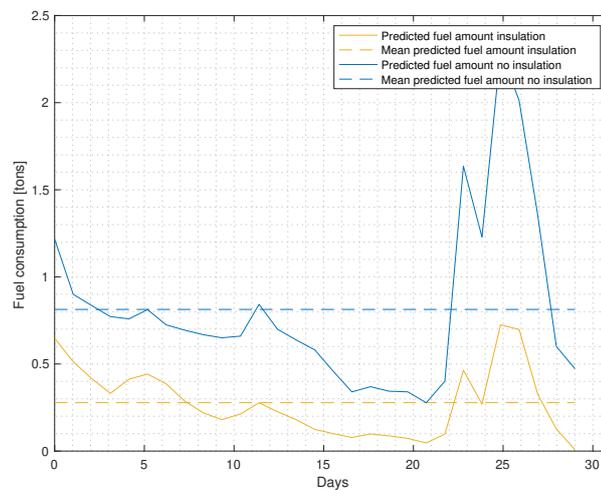


Figure 5.8: Comparison: simulation without insulation and insulation with Charcoat CTI with a thickness of 4 [mm]

Table 5.2: Different insulation materials

Insulation material	Thermal conductivity [ $W/mk$ ]	Thickness [mm]	Fuel consumption [ton/day]
No insulation	-	0	0.813
nC ProTherm CUI-HTF [73]	0.04	4 15	0.5847 0.4312
ICU Isollat 2 [74]	0.002	1 4	0.4045 0.3075
Charcoat CTI [75]	0.001	0.4 4	0.4251 0.2778

## 5.2. Data-driven model

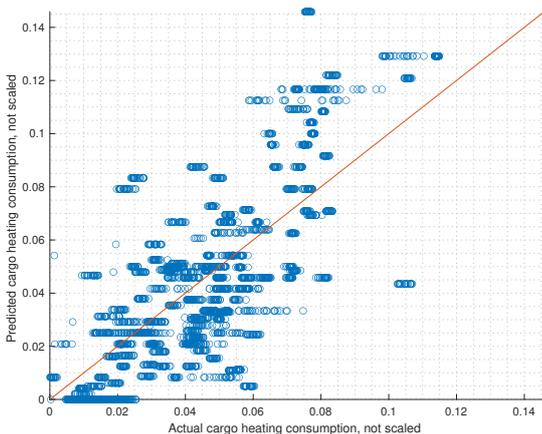
The MAE and MAPE of both models are displayed in Table 5.3. The results are given in the form of mean value  $\pm$  CI. First, the results for the linear model are explained in subsection 5.2.1 and after that, subsection 5.2.2 explains the results for the nonlinear model. After that, this section concludes with the results when extrapolation is tested.

Table 5.3: MAE and MAPE per model

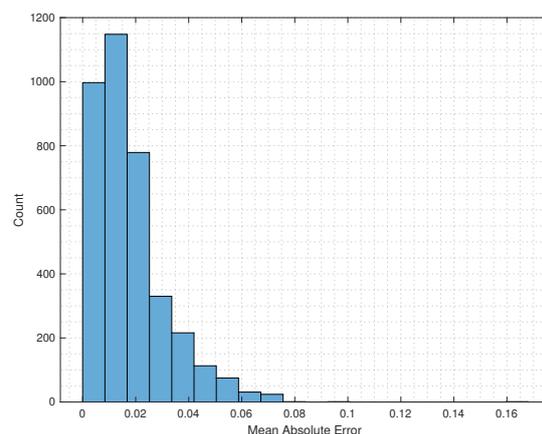
	MAE	MAPE
RidgeCV	$0.017583 \pm 0.00076$	$69.1004 \pm 4.4181$
SVR	$0.0014 \pm 1.7880 * 10^{-04}$	$5.18616 \pm 1.05567$

### 5.2.1. Linear model

The result for the linear regression model is stated in Table 5.3. The MAPE for the RidgeCV method is not deemed sufficient to continue, since the accuracy is not sufficient. The MAE is on average 0.01758 tons of fuel per hour. Figure 5.9a shows the histograms of the MAE. This histogram shows that the occurrence decreases when the error increases. This form of histogram is what would be correct. The larger the error is, the less often this should occur. However, looking at Figure 5.9a, it can be seen that the prediction is not very accurate, as the data points are quite far from the line. This result was expected because the correlation matrix (Figure 4.3) shows no clear linear correlation between the input parameters and the target variable.



(a) Fuel consumption prediction plotted against measured fuel consumption



(b) Histogram of absolute error;  $|Y_{pred} - Y_{actual}|$

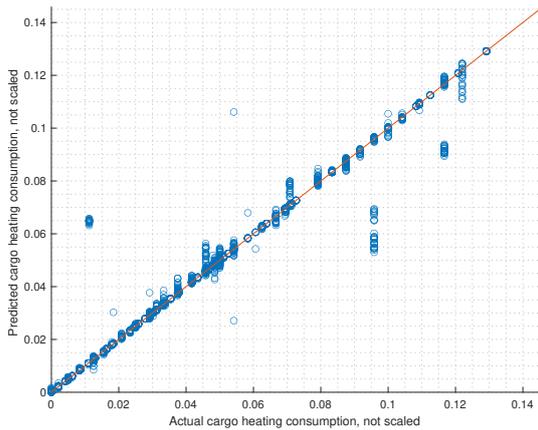
Figure 5.9: Results DDM RidgeCV model

### 5.2.2. Non-linear model

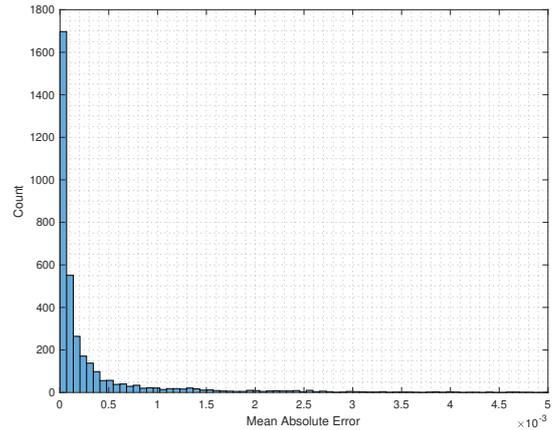
The result for the SVR is stated in Table 5.3. The MAE of the nonlinear model is 0.0014 tons of fuel per hour. When this is expressed as a percentage, in the form of the MAPE, this is 5.18%. In Figure 5.10a

and Figure 5.10b, the results from the SVR method can be seen. Figure 5.10a shows the predicted fuel consumption versus the actual reported fuel consumption. Ideally, each prediction should be the same as the actual reported value. If this were the case, it would mean that the whole graph would be on the diagonal line. Figure 5.10a shows that the model predicts that the vessel uses fuel for cargo heating, however, there are situations where the vessel does not require fuel for cargo heating.

Figure 5.10b shows the frequency of differences between the actual fuel consumption usage and the predicted fuel consumption usage. It can be seen that the MAE decreases when the error increases. In Figure 5.10a there are vertical lines of predictions shown. The actual cargo heating consumption has a variety of about 40 unique values. There are several occasions where the model predicts a different fuel consumption based on the input data, while the reported fuel consumption is the same. The horizontal position in Figure 5.10a is then the same, while the vertical position differs. This causes the vertical lines in the figure.



(a) Fuel consumption prediction plotted against measured fuel consumption



(b) Histogram of absolute error;  $|Y_{pred} - Y_{actual}|$

Figure 5.10: Results DDM SVR model

### 5.2.3. Extrapolation

The model is trained on data within a certain range. In the previous section, the model is tested against data that is in this range. This means that the model is able to interpolate between these values. It could be beneficial to test the performance of the model when it has to calculate extreme values, for example when the average cargo temperature is higher than usual or the sea or air temperature is higher or lower than what has ever been experienced. Figure 2.11 showed how this could be done. The top 10 % for one of the columns is taken out of the dataset to test model performance when calculating extreme values. The model is trained with a subset of the data and then the model is checked against the extrapolation dataset. This has been done for three parameters. The results can be seen in Table 5.4. This table shows the MAPE of the extrapolation. It can be seen that this model does not perform well in the extrapolation of the data and should not be used for this.

Table 5.4: Extrapolation of certain parameters

Parameter	MAE	MAPE
Air Temperature	$0.2787 \pm 0.2700$	$212.789 \pm 4.408$
Sea Temperature	$0.3919 \pm 0.3576$	$190.270 \pm 7.294$
Average Cargo Tank Temperature	$0.2600 \pm 0.2888 \pm$	$227.223 \pm 11.106$

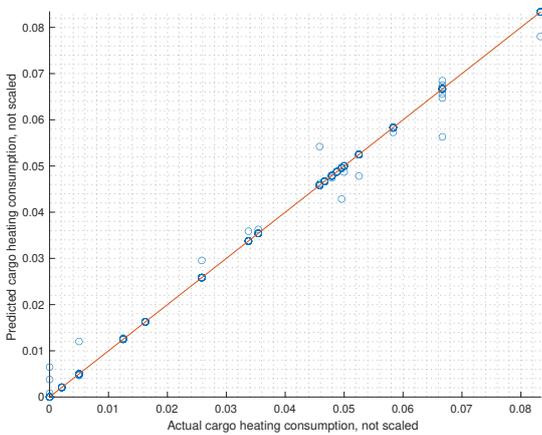
### 5.3. Hybrid model

The results of the HM are displayed in Table 5.5 and in Figure 5.11. The MAE for the HM is 0.0001441 tons of fuel per hour, while the MAPE is 0.8929. The table suggests that the model outperforms the DDM. The histogram of the absolute error shows that the error is close to 0 and the frequency of

occurrence decreases when the error increases. The results displayed are from the nonlinear model, since the DDM showed that a linear model performs worse than the nonlinear SVR model.

Table 5.5: Results SVR hybrid model

MAE	MAPE
$0.0001441 \pm 0.00010$	$0.8929 \pm 1.0955$



(a) Fuel consumption prediction plotted against measured fuel consumption

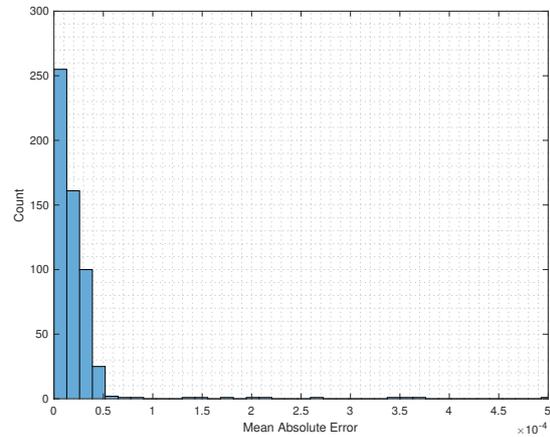
(b) Histogram of absolute error;  $|Y_{pred} - Y_{actual}|$ 

Figure 5.11: Results hybrid model

### 5.3.1. Extrapolation

This model uses the same methodology to extrapolate data as explained in subsection 2.2.3. The results of this extrapolation are displayed in Table 5.6. This model shows that it is not possible to extrapolate the data for the air and the sea temperature. The MAE and MAPE for these two parameters are too high. A MAE of 0.05 is too high, since the maximum fuel consumption is 0.14 ton/hour. The MAE of 0.05 gives thus too much deviation to be able to predict accurately enough. However, this model is able to extrapolate for the average cargo tank temperature. The MAE and MAPE are both low, thus the model is able to predict the fuel consumption accurately enough for the use of a chemical shipping company. This could be an effect of the addition of the PBM into the dataset. The theory behind this suggests a clear relation between the fuel consumption needed for cargo heating and the average cargo tank temperature.

Table 5.6: Extrapolation of certain parameters for the hybrid model

Parameter	MAE	MAPE
Air Temperature	$0.0458 \pm 0.02047$	$98.5806 \pm 0.76453$
Sea Temperature	$0.0458 \pm 0.02048$	$83.1132 \pm 0.57621$
Average Cargo Tank Temperature	$0.00220 \pm 0.002247$	$5.5458 \pm 1.1963$

## 5.4. Combination of results

Every model is able to predict the fuel consumption with an accuracy that is usable for a chemical shipping company. The PBM takes the previous timestep into account, while the DDM and HM are only able to calculate the fuel consumption without taking time into consideration. The reason for this is that there is no time dependency in the dataset that is utilized by the DDM and HM. Furthermore, the dataset used in the DDM and HM did not have data for three days. For this reason, there is no average line plotted for these two models since it does not represent the average over the entire trip. Figure 5.12 shows these predictions. The DDM and HM are displayed as instantaneous predictions

indicated with single points, not connected with a line. The results from the PBM and the reported fuel consumption take time into account and are therefore plotted with a line.

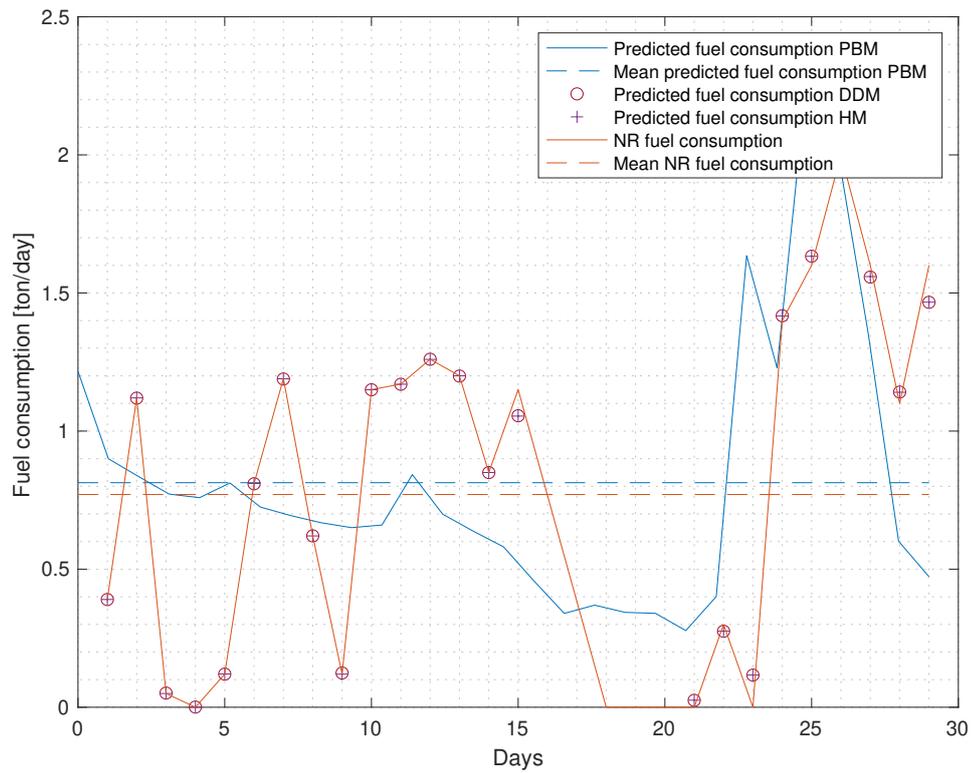


Figure 5.12: Fuel consumption prediction over over time; PBM, DDM and HM

# 6

## Conclusion

This chapter is the conclusion of the thesis. Section 6.1 discusses the conclusions from the physics-based model. This is followed by section 6.2 where the data-driven model is discussed. Section 6.3 explains the conclusions from the hybrid model and section 6.4 answers the research question.

### 6.1. Physics-based model

The introduction to this chapter states the subquestion answered in this chapter. The question is "To what extent are ordinary differential equation models, based on available physical data, solved with numerical Euler methods able to capture the fuel consumption of the heating of cargo per trip?". Figure 5.7 shows that the mean of the predicted and reported consumption for cargo heating is similar. The method is described in chapter 2. The benefit of this method is that it is based on commonly known and accepted laws for heat transfer. The downside, however, is that for each individual vessel the dimensions must be found and manually coupled. The modelled vessel has 48 cargo tanks, all having six walls. This resulted in at least 228 walls where heat transfer could occur. The model depends on the geometric configuration of a vessel. The modelled vessel is a simplification of the real vessel, since not all geometries are fully known. The areas of the walls of the tank are known, but the exact shape was not available. This led to some kind of simplification of the vessel. The Forward Euler method is used to solve the system of ordinary differential equations.

Appendix D shows the results for all individual cargo tanks. It can be seen that the predicted temperature in the tank is equal to the actual tank temperature, or higher. In most tanks, the cooling down trend from the prediction is similar to the reported cooling down of the tanks. In the model, the parameter  $\alpha$  is implemented to mirror the physical conditions. The value of  $\alpha$  is empirically determined by running the model and checking the behaviour of the temperatures. These results seem promising, but the computational time of this method is significant. The processor used in this research is a 2,6 GHz 6-Core Intel Core i7 processor. The computation for the PBM is about two hours to simulate a month of operation. Secondly, all the vessel and tank geometries have to be found and manually coupled into a model. The combination of these two aspects leads to the fact that it might not be realistic to use this model in a day-to-day approach.

This model, however, is useful in explaining of the benefit of insulation. Everyone knows that insulation is beneficial for reducing heat losses. This model quantifies the benefits of different materials and different thicknesses of insulation and this research has shown that this could reduce the fuel consumption by 65%. When this is converted towards total fuel consumption savings in a vessel, this would reduce the fuel consumption of the vessel by 3.25%.

The reported operational data of the chemical tanker vessel fluctuates significantly. Around day 17 to 20, there is no consumption of fuel for cargo heating, followed by a severe increase in the consumption. This fluctuation could be changed when more constant heat is applied.

### 6.2. Data-driven model

To be able to answer the research question stated in the introduction, the following subquestion must be answered: *How can operational data be exploited for the consumption prediction of heating of cargo for*

a *single trip*? The results from this chapter propose a method that is able to predict the fuel consumption for cargo heating. The accuracy of the proposed model is 5%. When the model is built and trained, it could calculate the fuel consumption prediction within seconds. This makes this method suitable for estimations for the fuel consumption. For instance, this model is able to calculate the difference in fuel consumption in case of a weather change.

Furthermore, this model can be adjusted easily for the processing of data generated by other vessels, provided those vessels generate similar data. The downside of the proposed method is that the model only gives accurate results when it has already encountered these circumstances. The model performs badly if it has to predict the fuel consumption when the temperature of the tanks increases or decreases out of the normal operating window. When applying this model, the upper and lower limits for each individual parameter must be defined properly. Since the model is able to calculate the fuel consumption with an MAPE of 5%, this research shows that there are opportunities for a control system for cargo heating. In case the behaviour of the cargo heating is known, this model is able to predict the fuel consumption, albeit further research to optimize the method is a logical follow-up action. This is supported by the fact that the data gathered only covers the past two years. It is obvious that data covering a longer period will improve the model and the output generated by the method applied.

### 6.3. Hybrid model

The hybrid model is a combination of the PBM and the DDM. The results show that the HM is an improvement of the DDM, and that less data is used to train the model. To make this method work, double the work has to be done compared to both the PBM and the DDM since it is a combination of both models. The results improve compared to the DDM. The trained model has a low calculation time to give an answer. Furthermore, the model is able to extrapolate on the average cargo tank temperature. This means that the results for different average cargo tank temperatures can easily be found. The novelty of this approach is that there is no literature available on a HM within the maritime industry for the cargo heating consumption calculations. Since this research is fast to calculate the fuel consumption for a certain trip, further research could focus on a model that is able to forecast the fuel consumption behaviour over time before a journey.

### 6.4. Research question

In the introduction, there is a research question stated here again: *What models can be used to predict the fuel consumption for heating of cargo for a given vessel and specific voyage conditions, taking operational use into account?* This thesis investigates different models. A PBM, DDM and HM are considered. Each model has benefits and disadvantages compared to the other models. An overview is given in Table 6.1 and will be explained in this section.

Table 6.1: Comparison between models

Method	Model preparation	Simulation time	Level of detail	MAPE
PBM	++	2 hours for 30 days	Tank	5.6%
DDM	+	Training model: 45 minutes; using model: instantaneous	Vessel	$5.2 \pm 1.1\%$
HM	+++	Training model: 3 hours; using model: instantaneous	Vessel	$0.9 \pm 1.1\%$

The PBM requires a detailed level of modelling for each individual vessel. When two different vessels are considered, the method to calculate the fuel consumption for a certain voyage is equal. However not the same code could be utilized, since the coupling of all the tanks has to be done dependent on the geometry of the vessel. The DDM requires less preparation in order to use this model. The data has to be downloaded and filtered. However, different types of vessels can use the same code to train a model. This enables the method to be easily scalable to different vessels. The HM requires the most preparations in order to use the HM. This is due to the nature of a HM. A HM is a combination of the two other models and thus, the work for both the PBM and DDM has to be done to create this model.

The time for each model to predict a fuel consumption over a specific timeframe alters too. The DDM and HM both require training of the model. The time to train these models is approximately 45 minutes based on these datasets. When these models are trained, they can predict the fuel consumption instantly. This makes these models very suitable for quick calculations. On the other hand, the PBM has to calculate every timestep during a voyage. The simulation time during the simulated voyage was about 2 hours. There might be options to reduce the computational time of the model, but this model will not become as fast as the DDM and HM.

The benefit of the PBM is that it gives the user more detailed information. The temperatures in each individual tank are modelled and can be seen. This is not possible with the DDM and HM. These models are trained to predict the fuel consumption based on a set of input parameters but do not have detailed information of the individual tanks. Furthermore, the PBM gives the user the freedom to change the physical elements of the vessel. It enables the user to calculate the influence of insulation. These detailed calculations are useful in deciding if it is worth to insulate existing vessels or to insulate vessels that still have to be built.

The accuracy of the PBM and DDM is about equal, namely 5.6% compared to 5.2%. The HM outperforms these two models, being more accurate (a MAPE of 0.9%). This is expected, since the data used in the HM has a value for the fuel consumption due to the PBM. This makes it easier for the HM to predict the fuel consumption.

The PBM is not in place to be run on a daily basis, but could be used to calculate structural changes to a vessel. The DDM could be used as a method to give a quick overview in predicting the fuel consumption, given that the parameters are within the limits of the trained dataset. Depending on resources and desired outcome, different methods and models could be used to predict the fuel consumption for the heating of cargo.



# 7

## Discussion and further recommendations

This chapter discusses some uncertainties in the research. This chapter will first discuss parts from the PBM in section 7.1, followed by the DDM in section 7.2 and continues with the HM in section 7.3. The chapter finishes with further recommendations in section 7.4.

### 7.1. Physics-based model

The first part to discuss is the result of the PBM. This model is only compared against one trip of one vessel. The results during this trip suggest that the trend of the model is similar to the trend that is reported in the vessel. The next step is to simulate different trips to further verify the model. It could, for example, be interesting to simulate a trip where only half the cargo tanks are filled. The empty tanks cause less heat transfer between the cargo tanks and this is not tested in the model.

This research uses a scaling factor to calculate the convection. This scaling factor is called  $\alpha$  and this coefficient is determined based on one trip. This could also be better compared when the geometries are better known. This, however, takes a lot of effort, since all different shapes in the vessel should be modelled separately. When all of this is known, it could be interesting to perform CFD calculations on the vessel to verify the used method.

The PBM calculates the heat transfer per tank. The knowledge of this heat transfer is beneficial for a shipping company because it can improve the safety on board. When the heat transfer is known, it is possible to see if dangerous situations can occur. These situations have happened in the past and could be monitored more closely. The PBM is helpful in this regard.

Furthermore, this research has shown the potential of insulating cargo tanks. The PBM suggests a reduction of 65% in the best scenario. When a chemical shipping company decides to build new ships, it is definitely worth looking into insulation. Insulation will make the building vessel more expensive, but reduces fuel consumption. In the long run, this reduces both the cost and emissions for a shipping company. The 65% reduction is based on one simulated trip. To make this number more general, more trips have to be simulated to gain certainty.

The PBM now considers only one option for insulation. Further research could focus on the location of the insulation and the cost involved in insulating vessels. It might be possible to insulate existing vessels in order to save fuel and is definitely worth looking into.

The last aspect that could lead to more accurate results is to reduce the size of the time step of the Forward Euler method. This increases the computational time, but the truncation error caused by the Forward Euler method reduces with a smaller time step. Another option is to try and make use of the ODE solvers of Matlab. The issue with these solvers is that the heat transfer coefficient depends on the temperature changes, and is modelled explicitly. The heat transfer coefficient has two equations, dependent on the Rayleigh Number. The ODE solver in Matlab did not allow the possibility to use a heat transfer coefficient dependent on the Rayleigh Number. During this thesis, there was no solution found to this problem. However, there might be options to use these solvers.

## 7.2. Data-driven model

The first part to discuss about the DDM is the data gathered by Stolt Tankers. The model uses 12 parameters as input. Nine of these parameters are coming from the NR data. These parameters are scaled to an hourly basis. This means that there are moments in time when the data is not as accurate as possible. When data is gathered for a longer time, it would be beneficial to train the DDM and the HM with new data to have more variety in the data. Another option to increase the volume of data is to increase the measurement frequency of these parameters. Then it is possible to acquire more data in the same amount of time.

The DDM gives the possibility to calculate the fuel consumption for a given set of inputs. The next step could be to investigate a forecast model. A forecast model takes the sequence of past data into account to predict the nearby future. If this works, a forecast model could be used to determine if heating is necessary and beneficial, given future weather conditions. The main benefit of the forecast model is that it accounts for time. The current DDM does not have learned a relation with time. The fuel consumption fed into the model has the unit of [ton/hour], so for a given set of data, the outcome is also in ton per hour. A forecast model knows this time dependency and this results in the possibility to draw graphs as a function over time. With the DDM this is currently not possible.

The DDM uses two types of algorithms: a linear algorithm and a nonlinear algorithm in the form of support vector regression. Further research might focus on testing other algorithms and comparing the results of different algorithms, since other algorithms might produce a more accurate result.

The last point of discussion for the DDM is the selection of parameters used in the model. The parameters that have been selected, since these parameters give an indication of the vessel, cargo and environmental conditions. However, there might be other parameters which could be beneficial in predicting the fuel consumption for the cargo heating. The wind force is an example of one of these parameters, since convection is influenced by the speed of the air along the vessel.

## 7.3. Hybrid model

The HM is based on data from only one trip. The PBM simulated a single trip (29 days), where the DDM used data from 200 days. Hence, a small dataset is used to build the HM. When the PBM simulates more trips, more data can be fed into the HM too. This results in HM that is trained on more data, where the data has more variety. Especially this variety is important. In subsection 5.3.1, the behaviour of the model is tested on performance outside of the training data. It can be seen that this is possible for the parameter 'Average Cargo Tank Temperature', but not for the parameters 'Air Temperature and 'Sea Temperature'. When more data is gathered for the HM, the model can be utilized more frequently.

The HM is trained on a dataset where the cargo inside the tanks stays constant. This results in constant values for some of the parameters in the HM. This is a negative aspect of this model, since it is not tested against other cargo configurations. This should be tested. The model cannot account for a change in the parameter 'Maximum Minimum Cargo Tank Temperature', since it is not trained for this. When multiple voyages are simulated, there will be some variation about cargo-specific information in the data for the HM. This enlarges the possibilities of using this model.

## 7.4. Further recommendations

The first recommendation for Stolt Tankers is to gather all data into one system. For this research, data was pulled from different sources and prepared for use in the model. This results in a reduced scalability of the method. When all these data sources are in the same database, it is easier to combine these and use them for different vessels.

Besides this, closer monitoring of cargo tank temperatures should be considered. The figures in Appendix D show the cargo temperature fluctuations. It is noted that these temperatures vary significantly during a trip. From these graphs, the lower temperature limit cannot be found. For the fuel efficiency, it would be most efficient to reduce the temperatures as much as possible, but making sure that the temperatures are within the limits. When a vessel proves that it is capable of maintaining a steady temperature, it gives Stolt Tankers the confidence to reduce the temperatures closer to the lower limit and thus save fuel.

# Bibliography

- [1] Domenico Gorni and Antonio Visioli. Practical issues in modelling the temperature for the control of smart buildings. *Journal of Physics: Conference Series*, 570:62004, 4 2014. doi:10.1088/1742-6596/570/6/062004.
- [2] Alex J Smola and Bernhard Schölkopf. A tutorial on support vector regression. *Statistics and Computing*, 14(3):199–222, 2004. doi:10.1023/B:STCO.0000035301.49549.88.
- [3] Stolt Tankers. Stolt Tankers' Operating Fleet List, 2022. URL: [https://www.stolt-nielsen.com/media/xkhocjoc/stolt-tankers\\_operating-fleet-list\\_august-2022-v2.pdf](https://www.stolt-nielsen.com/media/xkhocjoc/stolt-tankers_operating-fleet-list_august-2022-v2.pdf).
- [4] R K Rajput. *ENGINEERING THERMODYNAMICS Third Edition SI Units Version*. Laxmi Publications (P) LTD, New Delhi, third edition, 2007.
- [5] IMO. Fourth IMO GHG Study 2020 Executive Summary. Technical report, 2020.
- [6] IMO. INITIAL IMO STRATEGY ON REDUCTION OF GHG EMISSIONS FROM SHIPS Contents. Technical report, 2018. URL: <http://www.imo.org>.
- [7] DNV-GL. EU ETS: Preliminary agreement to include shipping in the EU's Emission Trading System from 2024, 2023. URL: <https://www.dnv.com/news/eu-ets-preliminary-agreement-to-include-shipping-in-the-eu-s-emission-trading-system>
- [8] European Commission Climate Action. Reducing emissions from the shipping sector. URL: [https://climate.ec.europa.eu/eu-action/transport-emissions/reducing-emissions-shipping-sector\\_en](https://climate.ec.europa.eu/eu-action/transport-emissions/reducing-emissions-shipping-sector_en).
- [9] IMO. Rules on ship carbon intensity and rating system enter into force, 11 2022. URL: <https://www.imo.org/en/MediaCentre/PressBriefings/pages/CII-and-EEXI-entry-into-force.aspx>.
- [10] CII – Carbon Intensity Indicator - DNV. URL: <https://www.dnv.com/maritime/insights/topics/CII-carbon-intensity-indicator/answers-to-frequent-questions.html>.
- [11] Chalermkiat Nuchturee, Tie Li, and Hongpu Xia. Energy efficiency of integrated electric propulsion for ships – A review. *Renewable and Sustainable Energy Reviews*, 134:110145, 12 2020. doi:10.1016/J.RSER.2020.110145.
- [12] Congbiao Sui, Douwe Stapersma, Klaas Visser, Peter de Vos, and Yu Ding. Energy effectiveness of ocean-going cargo ship under various operating conditions. *Ocean Engineering*, 190, 10 2019. doi:10.1016/j.oceaneng.2019.106473.
- [13] Yanguo Zhang, Qinghai Li, and Hui Zhou. Emission and Absorption of Thermal Radiation. *Theory and Calculation of Heat Transfer in Furnaces*, pages 45–74, 1 2016. doi:10.1016/B978-0-12-800966-6.00002-8.
- [14] Damiano Zito, Claudio Bruzzese, Ezio Santini, Alberto Tessarolo, D Zito, C Bruzzese, E Santini, A Raimo, and A Tessarolo. Performance and Efficiency Improvement of a Hydraulic Ship Steering Gear by a Permanent Magnet Linear Synchronous Servo-Motor Low speed generators for renewable resources View project GENFAIL View project Performance and Efficiency Improvement of a Hydraulic Ship Steering Gear by a Permanent Magnet Linear Synchronous Servo-Motor. Technical report, 2015. URL: <https://www.researchgate.net/publication/286935799>.

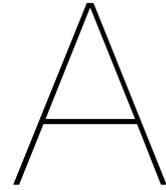
- [15] Foivos Mylonopoulos, Henk Polinder, and Andrea Coraddu. A Comprehensive Review of Modeling and Optimization Methods for Ship Energy Systems. 2023. doi:10.1109/ACCESS.2023.Doi.
- [16] Vassilios Dimoulas. Optimizing Cargo Heating For Tankers. In *ELINT Annual Meeting*. Bureau Veritas, 2017.
- [17] John H Lienhard. *A Heat Transfer Textbook Third Edition*. Cambridge, 2017. URL: <http://web.mit.edu/lienhard/www/ahtt.html>.
- [18] Hari Dass. *The Principles of Thermodynamics*. CRC Press, Boca Raton, 8 2014.
- [19] Frank Incropera, David Dewitt, Theodore Bergman, and Adrienne Lavine. *Fundamentals of Heat and Mass Transfer*. John Wiley & Sons, sixth edition edition, 2007.
- [20] F Baldi, H Johnson, C Gabriellii, and K Andersson. Energy and Exergy Analysis of Ship Energy Systems-The Case study of a Chemical Tanker. *International Journal of Thermodynamics (IJOT)*, 18(2):82–93, 2015. URL: [www.ijotcat.com](http://www.ijotcat.com), doi:10.5541/ijot.70299.
- [21] Jialin Gao, Wenfeng Wu, Jiakuo Zhang, Jiaoyang Tu, Xuxiu Wang, and Shuai Yang. Numerical simulation analysis of oil heating process of oil tanker in Arctic route. In *Proceedings - 2019 2nd World Conference on Mechanical Engineering and Intelligent Manufacturing, WCMEIM 2019*, pages 420–423. Institute of Electrical and Electronics Engineers Inc., 11 2019. doi:10.1109/WCMEIM48965.2019.00089.
- [22] Damian Bocheński and Dominik Kreft. Cargo Ships' Heat Demand - Operational Experiment. *Polish Maritime Research*, 27(4):60–66, 12 2020. doi:10.2478/pomr-2020-0066.
- [23] Alperen Sari, Egemen Sulukan, and Doğanuş Özkan. Analysis and modeling the energy system of a chemical tanker by leap. *Journal of Ship Production and Design*, 37(1):45–53, 2 2021. doi:10.5957/JSPD.07190034.
- [24] Chenglong Shen, Jinshu Lu, and Yanyan Li. Study on characteristics of temperature field in cargo tank during heating of cargo oil. In *Journal of Physics: Conference Series*, volume 2044. IOP Publishing Ltd, 10 2021. doi:10.1088/1742-6596/2044/1/012177.
- [25] Christos Gkerekos, Iraklis Lazakis, and Gerasimos Theotokatos. Machine learning models for predicting ship main engine Fuel Oil Consumption: A comparative study. *Ocean Engineering*, 188, 9 2019. doi:10.1016/j.oceaneng.2019.106282.
- [26] C. Capezza, S. Coleman, A. Lepore, B. Palumbo, and L. Vitiello. Ship fuel consumption monitoring and fault detection via partial least squares and control charts of navigation data. *Transportation Research Part D: Transport and Environment*, 67:375–387, 2 2019. doi:10.1016/j.trd.2018.11.009.
- [27] Khanh Q. Bui and Lokukaluge P. Perera. Advanced data analytics for ship performance monitoring under localized operational conditions. *Ocean Engineering*, 235, 9 2021. doi:10.1016/j.oceaneng.2021.109392.
- [28] Andrea Coraddu, Serena Lim, Luca Oneto, Kayvan Pazouki, Rose Norman, and Alan John Murphy. A novelty detection approach to diagnosing hull and propeller fouling. *Ocean Engineering*, 176:65–73, 3 2019. doi:10.1016/j.oceaneng.2019.01.054.
- [29] Han Cui. Optimization of Preventive Maintenance Cycle of Ship Mechanical and Electrical Equipment Based on MRO System. *Journal of Coastal Research*, 93(sp1):953–959, 9 2019. doi:10.2112/SI93-138.1.
- [30] Clarkson. World Fleet Register, 2023. URL: <https://www-clarksons-net.tudelft.idm.oclc.org/wfr/fleet>.
- [31] UNCTAD. Merchant fleet – UNCTAD Handbook of Statistics 2022, 2022. URL: <https://hbs.unctad.org/merchant-fleet/>.

- [32] IMO. *Rules and Regulations for the Classification of Inland Waterways Ships, Part 5 Main and Auxiliary Machinery - Chapter 13 Piping Systems for Ships Intended for the Carriage of Liquids in Bulk - Section 7 Cargo heating arrangements*. 9.36 edition, 2021. URL: [https://www.imorules.com/IWW\\_PT5\\_CH13\\_7.html](https://www.imorules.com/IWW_PT5_CH13_7.html).
- [33] LTD. KOYO KAIUN CO. Our Chemical Tankers |, 2022. URL: <https://www.koyotky.co.jp/fleet/tanker.html>.
- [34] Andrea Coraddu, Luca Oneto, Francesco Baldi, and Davide Anguita. Vessels fuel consumption forecast and trim optimisation: A data analytics perspective. *Ocean Engineering*, 130:351–370, 1 2017. doi:10.1016/J.OCEANENG.2016.11.058.
- [35] A. Coraddu, M. Kalikatzarakis, L. Oneto, G. J. Meijn, M. Godjevac, and R. D. Geertsmad. Ship diesel engine performance modelling with combined physical and machine learning approach. In *Proceedings of the International Ship Control Systems Symposium*, volume 1. Institute of Marine Engineering Science & Technology, 2018. doi:10.24868/issn.2631-8741.2018.011.
- [36] scikit-learn developers. Support Vector Regression (SVR) using linear and non-linear kernels — scikit-learn 1.2.2 documentation. URL: [https://scikit-learn.org/stable/auto\\_examples/svm/plot\\_svm\\_regression.html](https://scikit-learn.org/stable/auto_examples/svm/plot_svm_regression.html).
- [37] Latif M. (Latif Menashi) Jiji. *Heat convection*. Springer, New York, 2006.
- [38] Terry M Tritt. *Thermal Conductivity Theory, Properties, and Applications*. Kluwer Academic / Plenum Publishers, New York, 2003.
- [39] Stuart W Churchill and Humbert H S Chu. CORRELATING EQUATIONS FOR LAMINAR AND TURBULENT FREE CONVECTION FROM A VERTICAL PLATE. Technical report, 1975.
- [40] Bahman Zohuri. Forced Convection Heat Transfer. In *Thermal-Hydraulic Analysis of Nuclear Reactors*, pages 323–345. Springer International Publishing, 2017. doi:10.1007/978-3-319-53829-7{\\_}9.
- [41] John L. Monteith and Mike H. Unsworth. Heat Transfer. *Principles of Environmental Physics*, pages 151–178, 2013. URL: <https://linkinghub.elsevier.com/retrieve/pii/B978012386910400010X>, doi:10.1016/B978-0-12-386910-4.00010-X.
- [42] Boris Belinskiy, John Graef, and Lingju Kong. Stefan-Boltzmann problem for heat transfer in a fin. *Mathematical Methods in the Applied Sciences*, 44(6):4745–4755, 4 2021. doi:10.1002/mma.7066.
- [43] Davood Domairry Ganji, Yaser Sabzehmeidani, and Amin Sedighiamiri. Radiation Heat Transfer. *Nonlinear Systems in Heat Transfer*, pages 105–151, 1 2018. URL: <https://linkinghub.elsevier.com/retrieve/pii/B9780128120248000035>, doi:10.1016/B978-0-12-812024-8.00003-5.
- [44] Gerard Maugin. THE THERMOMECHANICS OF NONLINEAR IRREVERSIBLE BEHAVIORS An Introduction. Technical report, World Scientific Publishing Co. Pte. Ltd, Singapore, 1999. URL: [www.worldscientific.com](http://www.worldscientific.com).
- [45] Gojko Magazinović. Vertical arrangement of coils for efficient cargo tank heating. *International Journal of Naval Architecture and Ocean Engineering*, 11(2):662–670, 7 2019. doi:10.1016/J.IJNAOE.2019.02.004.
- [46] Yu Wu, Leyang Dai, Ankang Kan, Chao Yang, Fang Wang, and Gangshe Zhang. Theoretical and experimental analysis of the heating operation for cargo oil shipped on an actual voyage. *Ships and Offshore Structures*, 2022. URL: <https://www.tandfonline.com/action/journalInformation?journalCode=tsos20>, doi:10.1080/17445302.2022.2109352.
- [47] Taiwo Ayodele. Types of Machine Learning Algorithms. 2010. doi:10.5772/9385.

- [48] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85):2825–2830, 2011. URL: <http://jmlr.org/papers/v12/pedregosa11a.html>.
- [49] Andrea Coraddu, Luca Oneto, Aessandro Ghio, Stefano Savio, Davide Anguita, and Massimo Figari. Machine learning approaches for improving condition-based maintenance of naval propulsion plants. *Proceedings of the Institution of Mechanical Engineers Part M: Journal of Engineering for the Maritime Environment*, 230(1):136–153, 2 2014. doi:10.1177/1475090214540874.
- [50] Veli Çelik and Erol Arcaklioğlu. Performance maps of a diesel engine. *Applied Energy*, 81(3):247–259, 2005. doi:10.1016/j.apenergy.2004.08.003.
- [51] Ran Yan, Shuaian Wang, and Yuquan Du. Development of a two-stage ship fuel consumption prediction and reduction model for a dry bulk ship. *Transportation Research Part E: Logistics and Transportation Review*, 138, 6 2020. doi:10.1016/j.tre.2020.101930.
- [52] Scikit Learn. Choosing the right estimator, 8 2022.
- [53] Alex J Smola and Bernhard Scholkopf. A tutorial on support vector regression \*. Technical report, Max-Planck-Institut für biologische Kybernetik, Tübingen, 11 2004.
- [54] Harris Drucker, Chris C, Linda Kaufman, Alex Smola, and Vladimir Vapnik. Support Vector Regression Machines. *Advances in Neural Information Processing Systems*, 9, 11 2003.
- [55] Vladimir N Vapnik. The Nature of Statistical Learning Theory. *The Nature of Statistical Learning Theory*, 2000. doi:10.1007/978-1-4757-3264-1.
- [56] Minghua Xie, Decheng Wang, and Lili Xie. A feature-weighted SVR method based on kernel space feature. *Algorithms*, 11(5), 5 2018. doi:10.3390/a11050062.
- [57] Mark Stitson, Alex Gammerman, Vladimir Vapnik, Volodya Vovk, Christopher Watkins, and Jason Weston. Support Vector Regression with ANOVA Decomposition Kernels. 11 1999.
- [58] Saifuldeen Dheyauldeen Alrefaee, Salih Muayad Al Bakal, and Yahya Algamal. Hyperparameters optimization of support vector regression using black hole algorithm. *Int. J. Nonlinear Anal. Appl*, 13:2008–6822, 2022. URL: <http://dx.doi.org/10.22075/ijnaa.2022.6107>, doi:10.22075/ijnaa.2022.6107.
- [59] Daniel Berrar. Cross-validation. In *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, volume 1-3, pages 542–545. Elsevier, 1 2018. doi:10.1016/B978-0-12-809633-8.20349-X.
- [60] Richard Simon, Werner Dubitzky, Martin Granzow, and Daniel P Berrar. *Fundamentals of Data Mining in Genomics and Proteomics*. New York, 2007.
- [61] Annette M Molinaro, Richard Simon, and Ruth M Pfeiffer. Prediction error estimation: a comparison of resampling methods. *BIOINFORMATICS ORIGINAL PAPER*, 21(15):3301–3307, 5 2005. URL: <http://www.bepress.com/ucbbiostat/paper126>, doi:10.1093/bioinformatics/bti499.
- [62] Saeed Banihashemi, Grace Ding, and Jack Wang. Developing a hybrid model of prediction and classification algorithms for building energy consumption. pages 14–16, 2017. URL: [www.sciencedirect.com](http://www.sciencedirect.com), doi:10.1016/j.egypro.2017.03.155.
- [63] Laura von Rueden, Sebastian Mayer, Rafet Sifa, Christian Bauchhage, and Jochen Garcke. Combining Machine Learning and Simulation to a Hybrid Modelling Approach: Current and Future Directions. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12080 LNCS:548–560, 2020. URL: [https://link-springer-com.tudelft.idm.oclc.org/chapter/10.1007/978-3-030-44584-3\\_43](https://link-springer-com.tudelft.idm.oclc.org/chapter/10.1007/978-3-030-44584-3_43), doi:10.1007/978-3-030-44584-3{\\_}43/FIGURES/5.

- [64] Alexei Botchkarev. A New Typology Design of Performance Metrics to Measure Errors in Machine Learning Regression Algorithms. *Interdisciplinary Journal of Information, Knowledge, and Management*, pages 45–79, 2019. doi:10.28945/4184.
- [65] Robert Fildes and Paul Goodwin. Against Your Better Judgment? How Organizations Can Improve Their Use of Management Judgment in Forecasting. *Interfaces*, 37(6):570–576, 2007. URL: <https://about.jstor.org/terms>, doi:10.1287/inte.1070.0309.
- [66] F.M. Dekking, C Kraaikamp, H.P. Lopuhaa, and L.E. Meester. *A Modern Introduction to Probability and Statistics*. Springer, Delft, 2005.
- [67] Ana-Maria Šimundić. Lessons in biostatistics - Confidence interval. *Biochemia Medica*, 18(2):154–161, 2008.
- [68] J Jochemsen-Verstraeten, Drs S E De Vos-Effting, Drs E E Keijzer, S N C Dellaert, Ir A Van Horssen, Ing R N Van Gijlswijk, and Ir J H J Hulskotte. Milieuprofielen van scheepsbrandstoffen ten behoeve van opname in de Nationale Milieudatabase. Technical report, 2016. URL: [www.tno.nl](http://www.tno.nl).
- [69] DNV-GL. RULES FOR CLASSIFICATION Ships, Chemical Tankers. Technical report, DNV-GL, 8 2015. URL: <http://www.dnvgl.com>, .
- [70] Frank Bienewald, Edgar Leibold, Pavel Tužina, and Günter Roscher. Vinyl Esters. In *Ullmann's Encyclopedia of Industrial Chemistry*, pages 1–16. Wiley, 9 2019. doi:10.1002/14356007.a27{\\_}419.pub2.
- [71] Smiths. Stainless Steel. *Technical Datasheet*, 2022.
- [72] Kishor G. Nayar, Mostafa H. Sharqawy, Leonardo D. Banchik, and John H. Lienhard. Thermophysical properties of seawater: A review and new correlations that include pressure dependence. *Desalination*, 390:1–24, 7 2016. doi:10.1016/j.desal.2016.02.024.
- [73] nC Protect. Product Data Sheet | PDS Product Data Sheet | PDS. Technical report, nC Surface Technologies BV, Wormerveer, 2 2023. URL: <https://www.nc-protect.com/Informatiesheets/ProTherm/NL/PDS%20nC%C2%AE%20ProTherm%20CUI-HTF%20NL21.pdf>.
- [74] Isollat-02 Industrial Liquid Thermal Insulation Coating – ICU Marine. URL: <http://icumarine.com/en/isollat-products/isollat-02-industrial-liquid-thermal-insulation-coating/>.
- [75] CharCoat CTI COLOURS. Technical report, CharCoat, Port Moody, Canada, 2 2023. URL: [www.CharCoat.com](http://www.CharCoat.com).
- [76] Dalwinder Singh and Birmohan Singh. Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, 97, 12 2020. doi:10.1016/j.asoc.2019.105524.





# Available data parameters

Table A.1: Cargo information and sensor information

## Cargo information

SNNo  
ProductShortName  
ProductFullName  
IMOName  
SpecificGravity  
CorrectionFactor  
VoyageMinTemp  
VoyageMaxTemp  
DischargeMinTemp  
DischargeMaxTemp  
HeatAdjacent  
VaporPressure  
VaporDensity  
BoilingPoint  
MeltingPoint  
FlashPoint  
Viscosity  
ThermalOilHeatingRequired

## Sensor information

Pressure per tank  
Temperature per tank

Table A.2: Noon report data

**Report information**

Ship class  
 Ship name  
 ReportDateTimeGMT

**Vessel conditions**

AFTDraft  
 MeanDraft  
 FWDDraft  
 EngineOrder  
 AverageRPM  
 SpeedOverGround  
 SpeedThroughWaterMEHrs  
 SpeedThroughWaterSTHrs  
 CorrectedSpeed  
 OrderedSpeed  
 PowerAtEvenKeel  
 Slip  
 EngineDistance  
 ObservedDistance  
 SteamingHours  
 HeadingDegree  
 HeadingDirection  
 Latitude  
 Longitude  
 LoadPerc  
 EnginesCorrect  
 TotalFuelIncinerator

**Fuel**

AuxEngineSFOC  
 AverageFuelPrice  
 CorrectedConsumption  
 FuelRemainingOnBoardIFO  
 FuelRemainingOnBoardLSF  
 FuelRemainingOnBoardMDO  
 FuelRemainingOnBoardMGO  
 ExpectedFuelConsumption

**Main engine**

MainEngineKWHrs  
 MainEngineRunningHrs  
 PropulsionConsumptionIFO  
 PropulsionConsumptionLSF  
 PropulsionConsumptionMDO  
 PropulsionConsumptionMGO  
 PropulsionSFOC  
 TotalPropulsionConsumption

**Generator**

Generator1KWHrs  
 Generator1RunningHours  
 Generator2KWHrs  
 Generator2RunningHours  
 Generator3KWHrs  
 Generator3RunningHours  
 Generator4KWHrs  
 Generator4RunningHours  
 GeneratorConsumptionIFO  
 GeneratorConsumptionLSF  
 GeneratorConsumptionMDO  
 GeneratorConsumptionMGO  
 TotalGeneratorConsumption

**Cargo heating**

CargoHeatingConsumptionIFO  
 CargoHeatingConsumptionLSF  
 CargoHeatingConsumptionMDO  
 CargoHeatingConsumptionMGO  
 TotalCargoHeatingConsumption  
 TotalBoilerConsumption  
 TotalHeatingConsumption  
 BoilerRunningHours

**Tank cleaning**

TankCleaningConsumptionIFO  
 TankCleaningConsumptionLSF  
 TankCleaningConsumptionMDO  
 TankCleaningConsumptionMGO  
 TotalTankCleaningConsumption

**Other consumption**

OtherConsumptionIFO  
 OtherConsumptionLSF  
 OtherConsumptionMDO  
 OtherConsumptionMGO  
 TotalOtherConsumption

**Fresh water**

FreshWaterProduced  
 FreshWaterConsumed  
 FreshWaterROB

**Trip**

BadWeatherDistance  
 BadWeatherHours  
 DistanceToGo  
 ETA  
 ETADeltaPreviousNoonReportHrs  
 ETADeltaSumVoyageLegHrs  
 VoyageLegDurationInDays  
 VoyageLegFromPortName  
 VoyageLegToPortName

**Weather**

AirTemperature  
 Beaufort  
 EffectiveCurrent  
 SeaDirection  
 SeaDirectionDegree  
 SeaHeight  
 SeaNumber  
 SeaTemperature  
 SwellDirection  
 SwellDirectionDegree  
 SwellHeight  
 SwellNumber  
 WindDirectionDegree

**Additional**

ExclusionReason  
 InBunkerConsumptionError  
 IncludeInPoolCalculation  
 IsLastReportVoyageLeg  
 IsVCCbadWeather  
 IsVCCdataEntryError  
 Remark  
 ReportType

# B

## Data visualisation

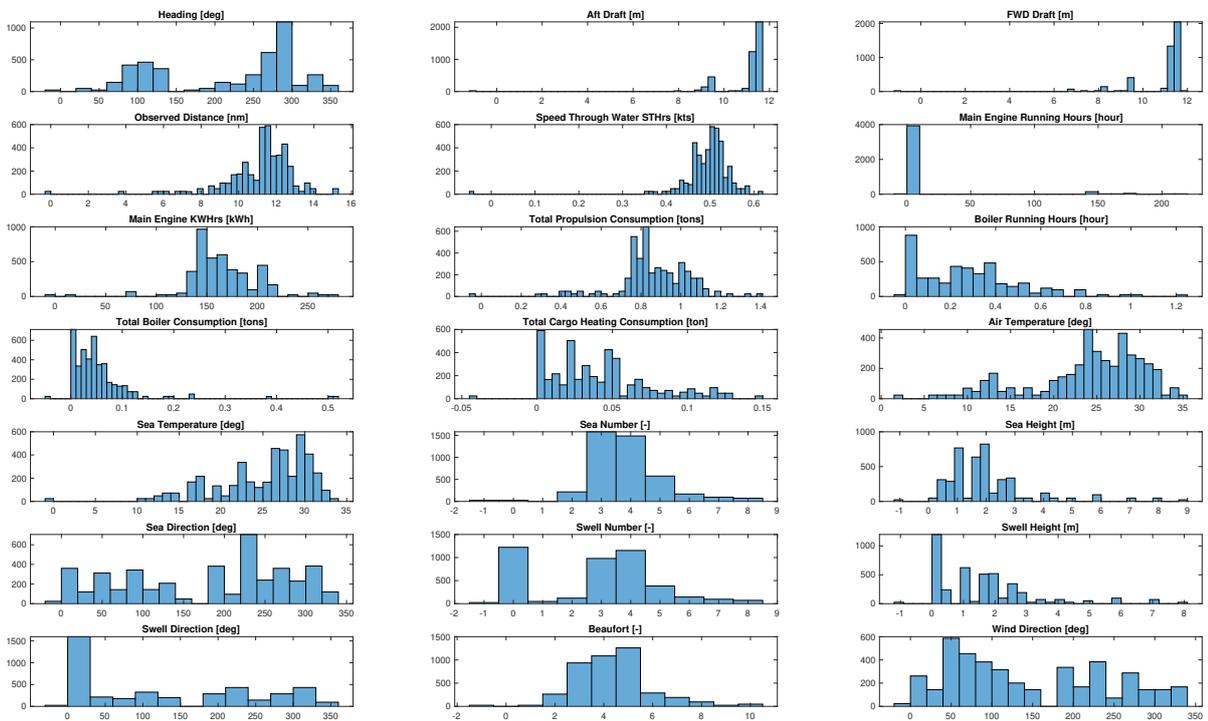


Figure B.1: Histograms of Unfiltered Noon Report Data

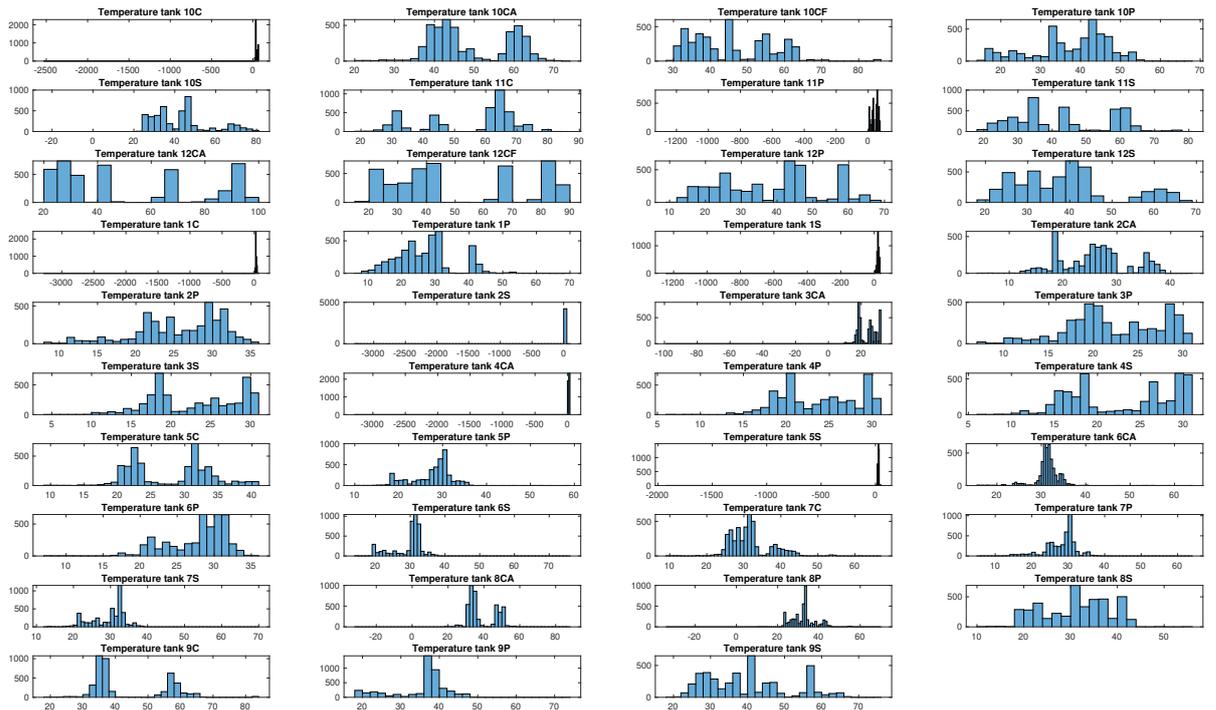


Figure B.2: Histograms of Unfiltered Temperature Data

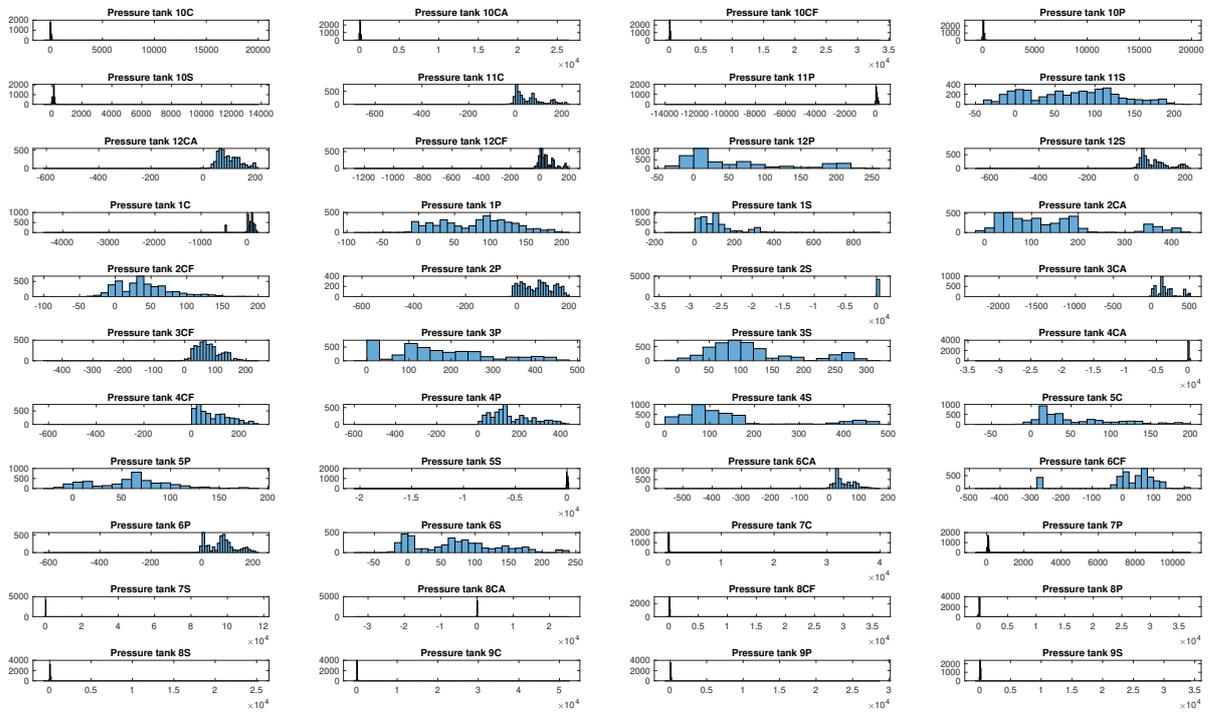


Figure B.3: Histograms of Unfiltered Pressure data

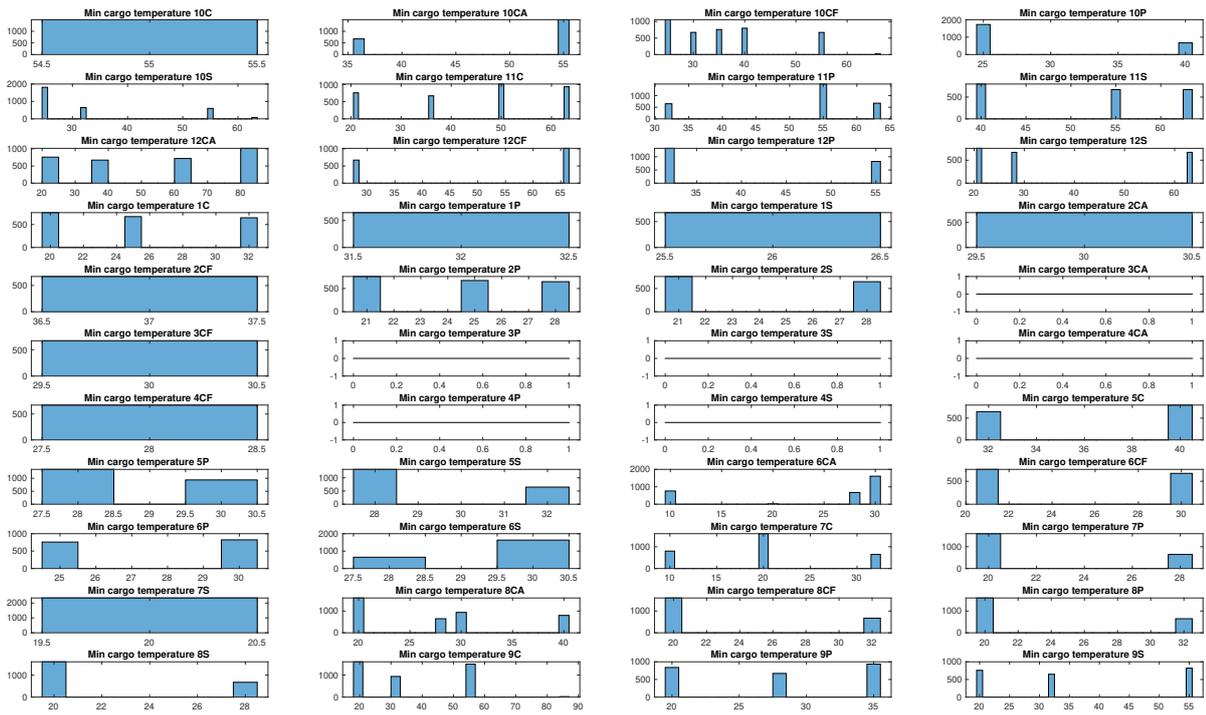


Figure B.4: Histograms of Unfiltered Minimum Cargo Temperature Requirements

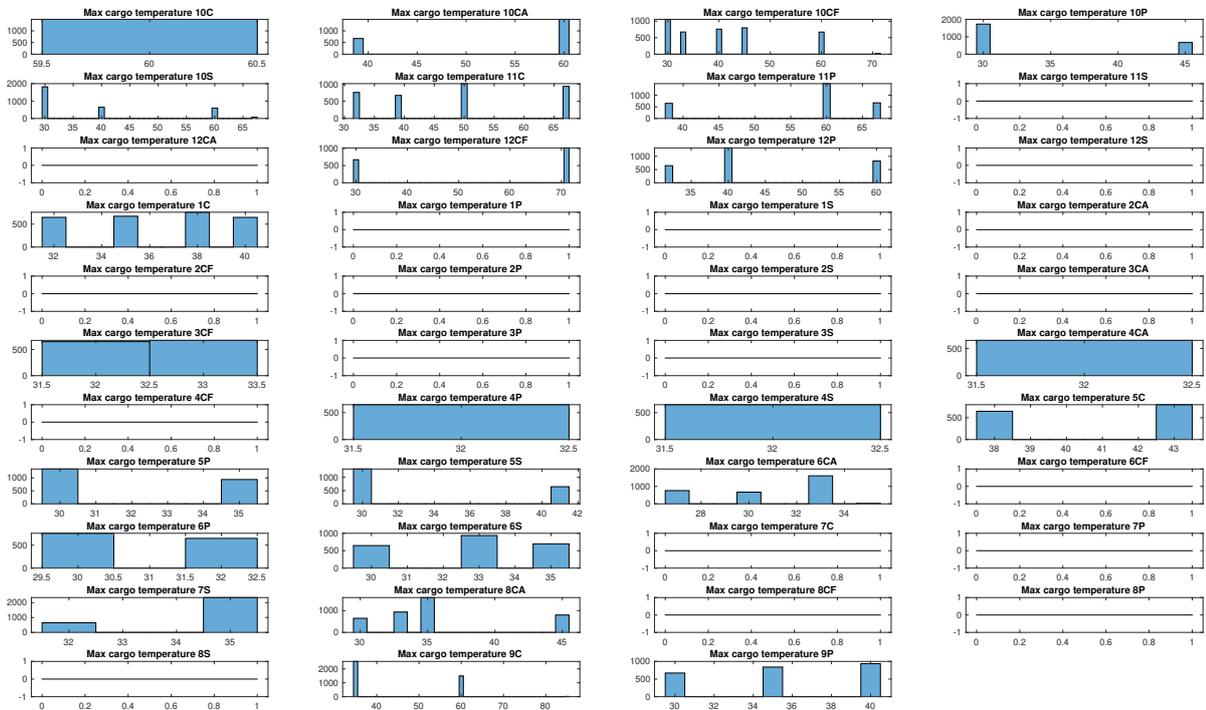
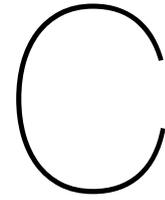


Figure B.5: Histograms of Unfiltered Maximum Cargo Temperature Requirements





# Cargo specifications

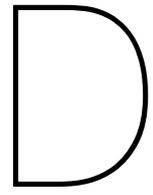
Table C.1: Cargo specifications

Tank Number	Mass [ton]	T_min [degC]	T_max [degC]	cp [J/kg*K]	k [W/m*K]	rho [kg/m <sup>3</sup> ]	mu [Pa*s]	L [m]	T_initial [degC]
Cargo tanks 1-44									
1	479	0	50	1540	0.1350	785	3	8	29
2	845	20	32	2051	0.1426	880	4	8	30
3	549	0	50	1150	0.1500	932	0	8	29
4	990	0	50	1975	0.1510	1000	3	8	29
5	974	0	50	1150	0.1510	932	0	8	29
6	224	0	50	1906	0.1700	950	1	8	29
7	986	0	50	1919	0.0600	1180	1	8	27
8	1010	0	50	1919	0.0600	1180	1	8	28
9	300	20	32	2051	0.1426	880	1	8	29
10	30	0	100	1000	0.0250	1.255	0	8	29
11	984	0	50	1150	0.1500	932	0	8	29
12	1125	20	32	2051	0.1426	880	4	8	30
13	824	0	50	1150	0.1500	932	0	8	28
14	298	20	32	2051	0.1426	880	4	8	28
15	1124	20	32	2051	0.1426	880	4	8	29
16	576	30	35	1804	0.1700	982	49	8	31
17	499	32	38	2051	0.1426	880	6	8	31
18	398	32	41	2051	0.1426	880	4	8	28
19	602	20	32	2051	0.1426	880	4	8	31
20	260	0	50	1540	0.1350	785	3	8	31
21	301	30	33	1861	0.1721	887	48	8	31
22	440	30	33	1861	0.1721	887	48	8	33
23	1079	0	50	1906	0.1700	950	1	8	30
24	1500	0	50	2973	0.1340	970	50	8	30
25	1432	20	32	2051	0.1426	880	4	8	30
26	617	0	50	1923	0.3275	1015	174	8	32
27	305	0	50	1923	0.3275	1015	174	8	33
28	310	30	33	1861	0.1721	887	48	8	33
29	618	0	50	1923	0.3274	1015	174	8	32
30	501	35	40	1861	0.1721	887	65	8	41
31	574	30	35	1804	0.1700	982	49	8	36
32	320	0	50	1906	0.1700	769	2	8	38
33	1604	25	30	2377	0.2850	1261	165	8	43

Continued on next page

Table C.1 – Continued from previous page

Tank Number	Mass [ton]	T_min [degC]	T_max [degC]	cp [J/kg*K]	k [W/m*K]	rho [kg/m <sup>3</sup> ]	mu [Pa*s]	L [m]	T_initial [degC]
34	793	25	30	2377	0.2850	1261	165	8	43
35	30	0	100	1000	0.0250	1.225	0	8	42
36	30	0	100	1000	0.0250	1.225	0	8	52
37	1603	25	30	2377	0.2850	1261	165	8	42
38	499	32	38	2051	0.1426	880	6	8	33
39	500	63	67	1861	0.1721	887	48	8	70
40	485	0	50	1906	0.1700	950	29	8	32
41	1583	20	32	2051	0.1426	880	4	8	32
42	30	0	100	1000	0.0250	1.225	0	8	39
43	1000	0	50	1906	0.1700	950	19	8	39
44	1475	0	50	1906	0.1700	950	29	8	32
Cofferdams 45-54									
45	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
46	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
47	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
48	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
49	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
50	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
51	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
52	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
53	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
54	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
Ballast tanks 55-70									
55	1	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
56	1	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
57	1	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
58	1	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
59	1	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
60	1	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
61	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
62	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
63	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
64	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
65	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
66	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
67	1	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
68	1	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
69	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
70	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
Area forward, aft and bottom 71-73									
71	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
72	0	0	100	1000	0.0250	1.2250	1.8250e-05	8	20
73	5	0	100	1000	0.0250	1.2250	1.8250e-05	1	20



# Thermodynamic Model Results

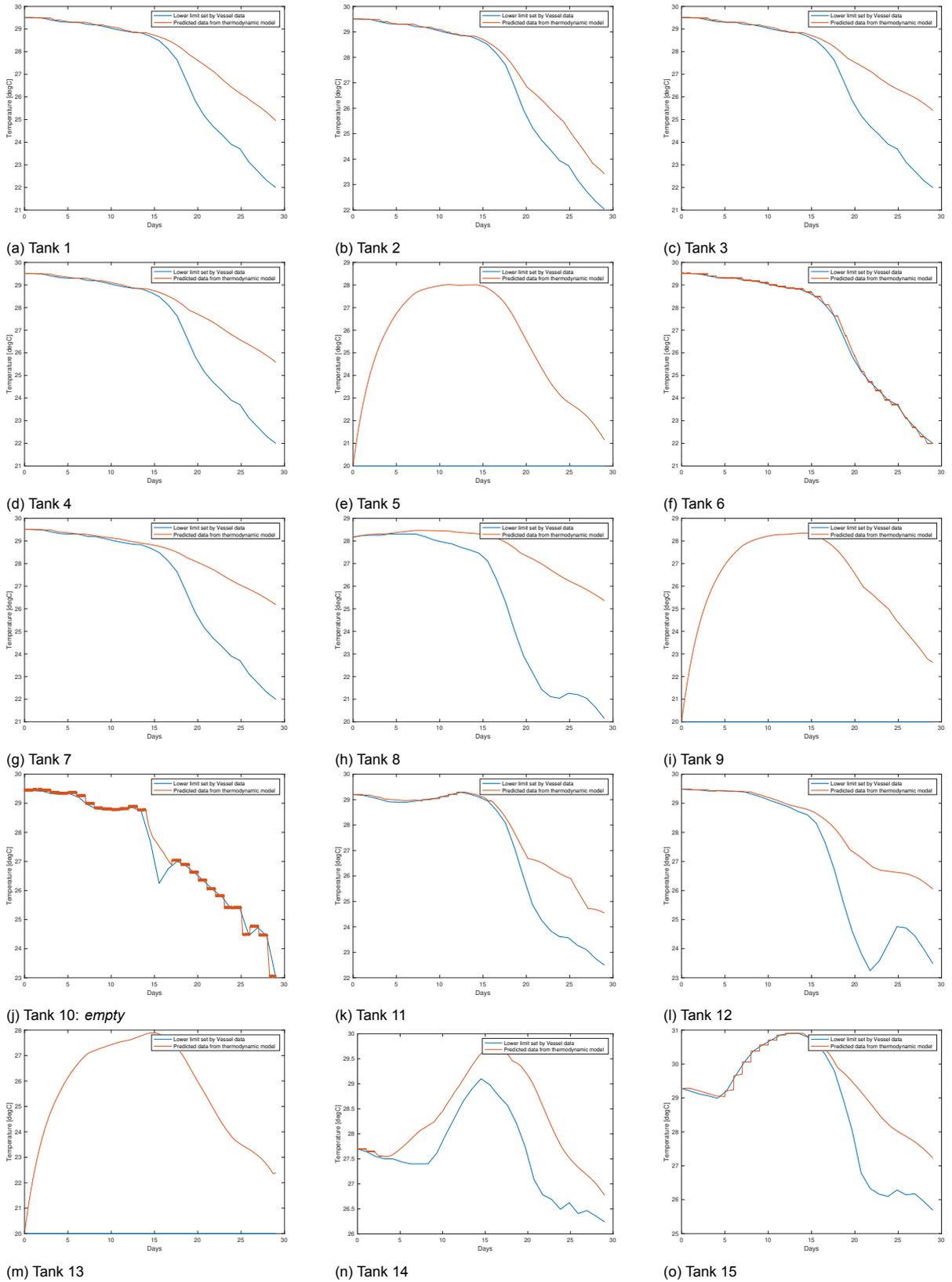


Figure D.1: Temperature profile tanks 1-15

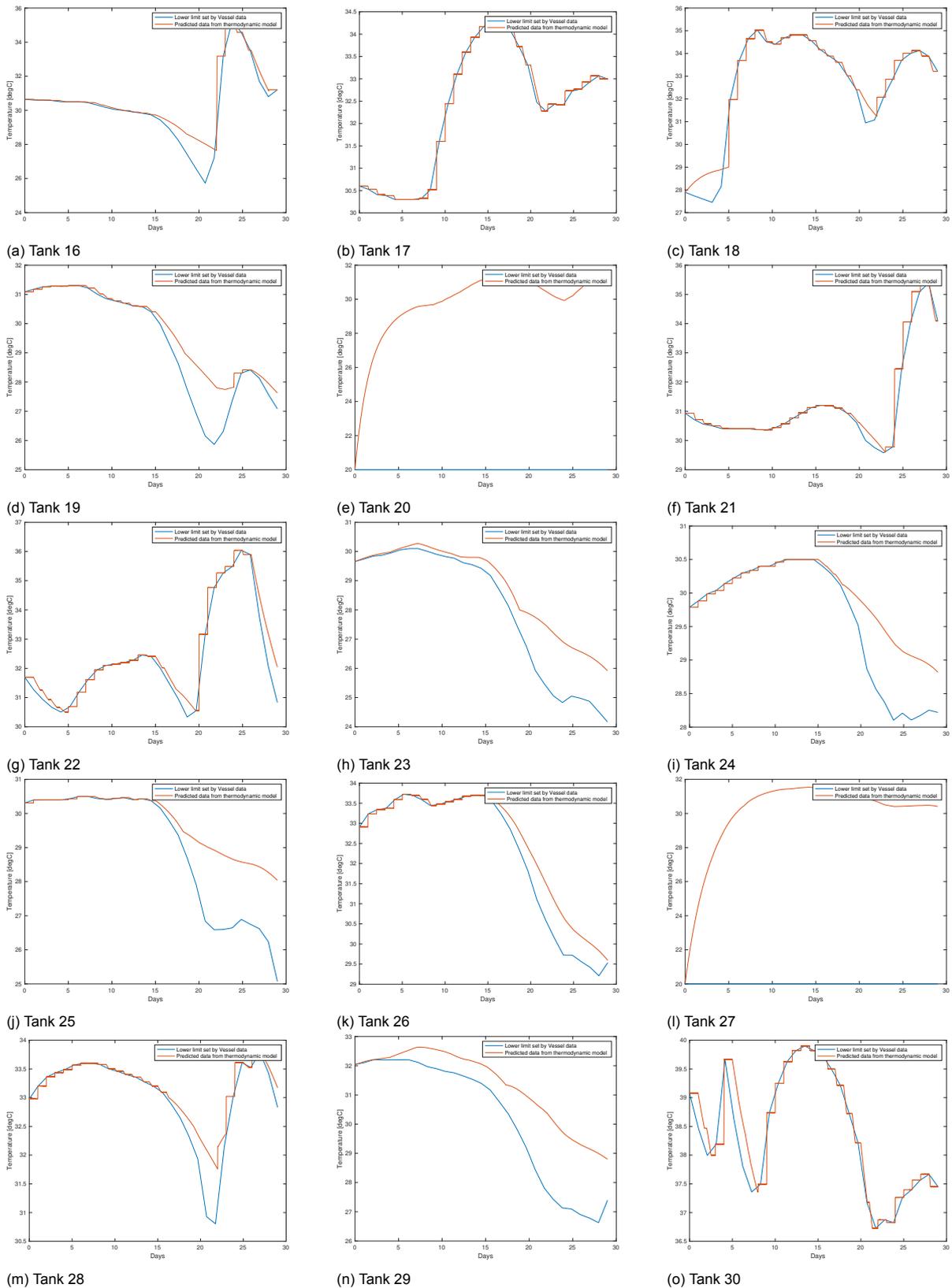


Figure D.2: Temperature profile tanks 16-30

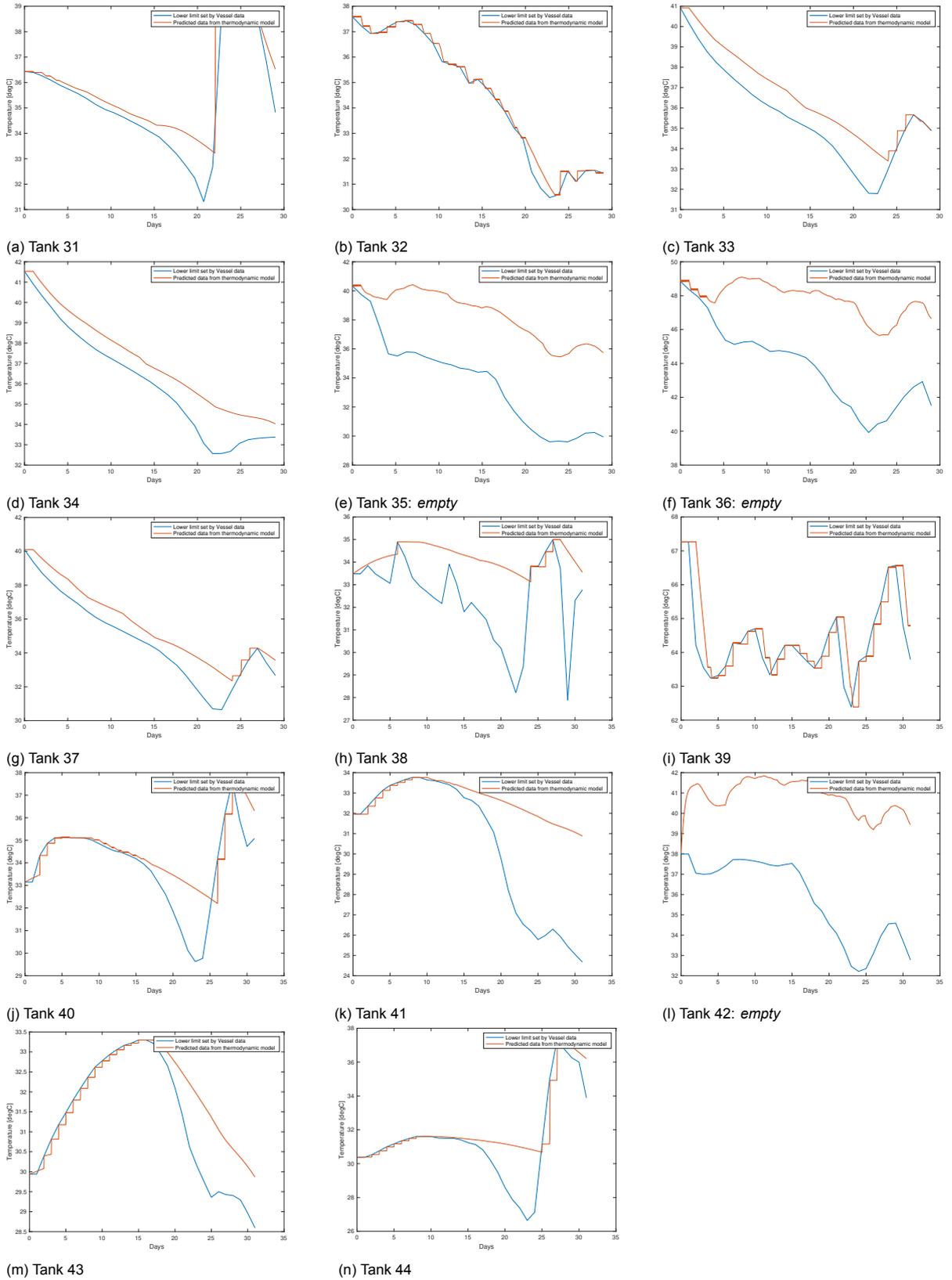


Figure D.3: Temperature profile tanks 31-44

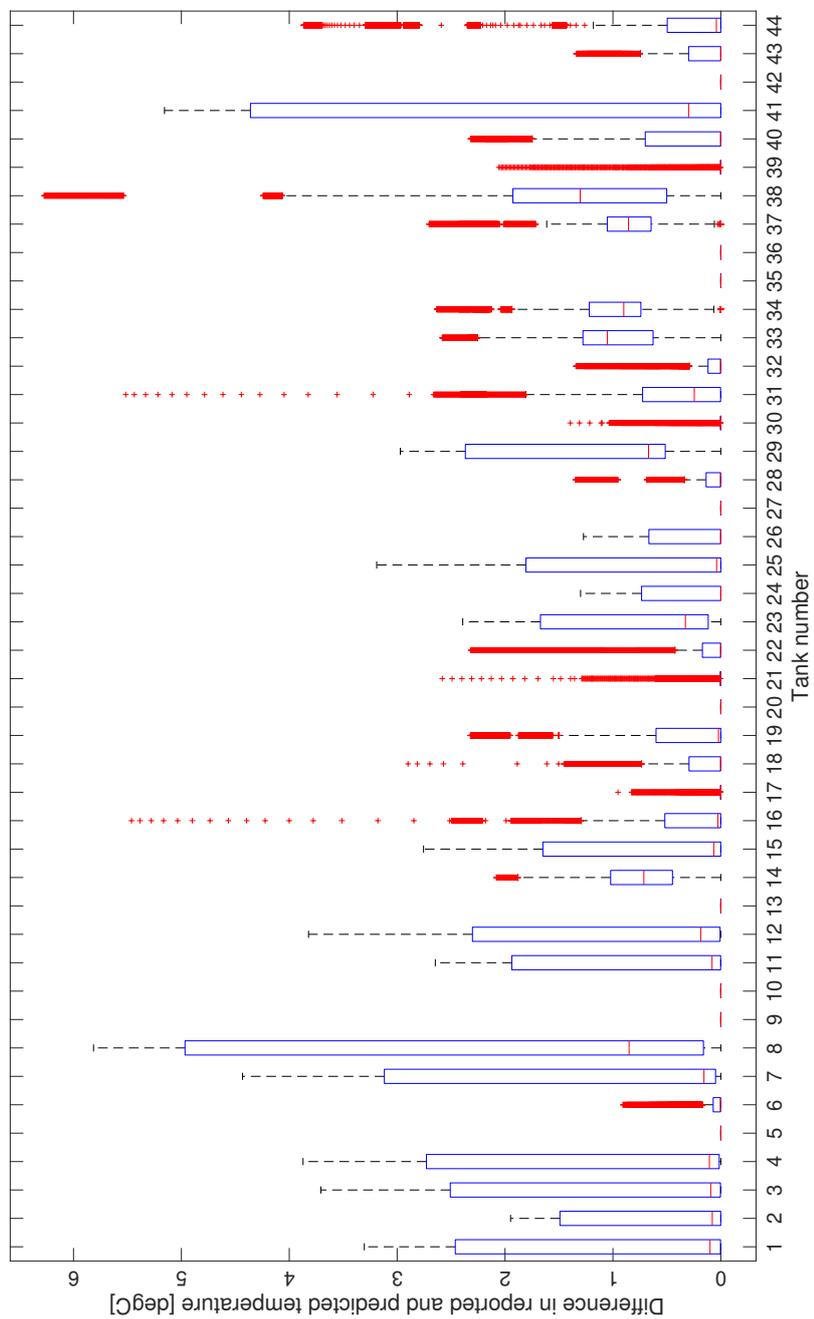
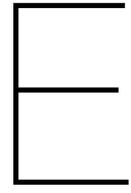


Figure D.4: Difference between reported and predicted temperature per tank [degC]





## Insulation tank layout

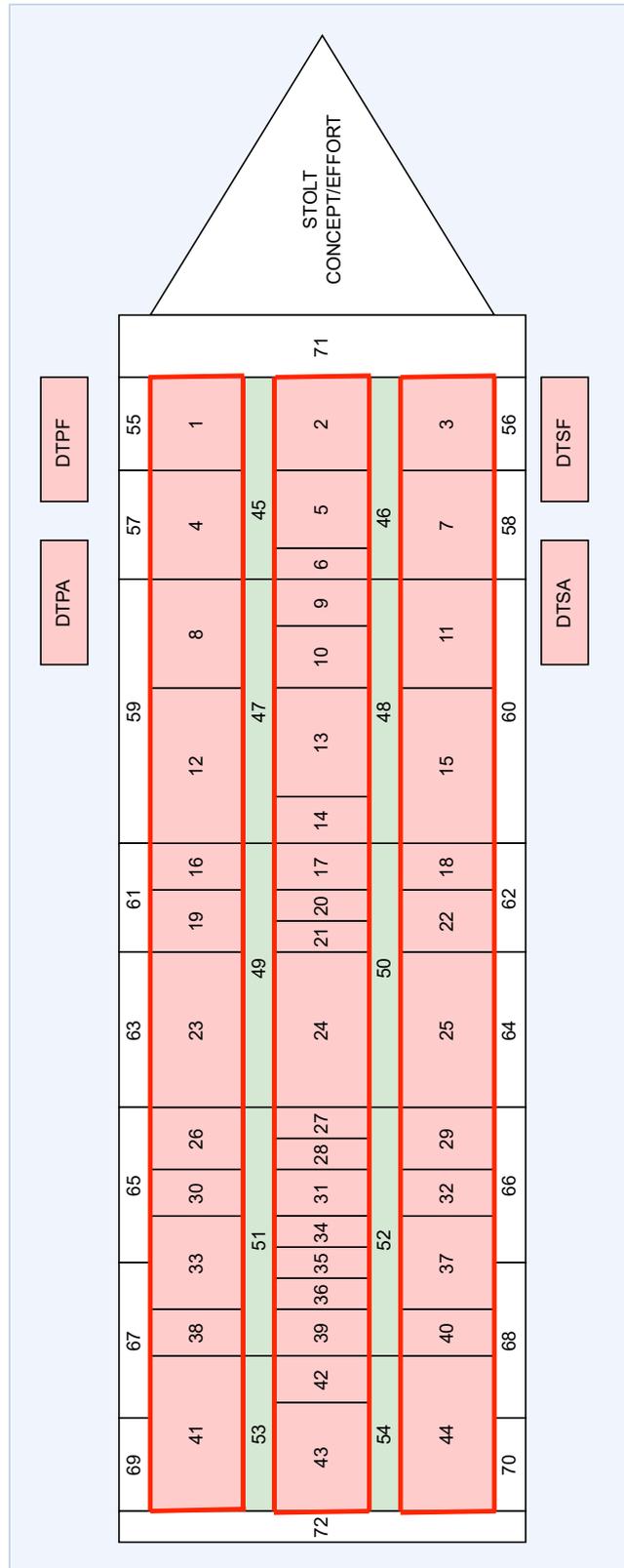


Figure E.1: General overview of the location of insulation of the tanks