MASTER

Robust safety stock levels determination of a tank container operator, dealing with demand uncertainty in a multi-commodity network setting

Groeneveld, R.G.

Award date:
2021

Eindhoven University of Technology

1ML05
Master thesis
Operation Management and Logistics

Robust safety stock levels determination of a tank container operator, dealing with demand uncertainty in a multi-commodity network setting.

October 15, 2021

Version: 3.0

Rafaël Groeneveld
Student number: 0900999

Company supervisor:
ir. Luke van de Bunt

University supervisors:
1st Supervisor: Dr. Ahmadreza Marandi
2nd supervisor: Dr. Shaunak Dabadghao

Eindhoven University of Technology
School of Industrial Engineering
Series Master Thesis Operations Management and Logistics

keywords: *robust optimization, adaptive optimization, static optimization, affine decision rules, forecasting, multi-commodity, repositioning, safety stock, tank containers*

It should be noted that all numbers and costs mentioned in this thesis are fictitious due to confidentiality reasons.

Abstract

Safety stocks levels of a tank container operator depends on the demand uncertainty, the network's structure, the corresponding demand types matched with tank container types, and the timing aspects. This project is executed at Den Hartogh Logistics in search of improvements in its safety stock level determination of tank containers. First, a forecast module is developed to estimate the demand at each hub individually of the different demand types. Secondly, different adaptive robust optimization models are used for the determination of the safety stock levels. Since the large size of the problem, cuts are made in the network based on the corresponding planning horizon for reducing computation time. The different robust optimization models use different approaches to incorporate the adaptive nature and the integrality restrictions of tank containers. These approximation variants are necessary to deal with integrality in the second stage in a computationally convenient manner. Thirdly, a redistribution module is developed to redistribute the tank containers to their initial safety stock level. This redistribution module does not contain uncertainty anymore and can be solved with a mixed linear integer program for the entire network. The results seem promising, and the current estimates of safety stock levels of internal experts are reasonably close in line with the results of the robust model. Although slightly more tank containers in general are preferred for the network of DH. If it is possible to decrease computation time even further for this scale of problems, future research could focus on a multi-period model, even improving the current situation.

Management summary

Introduction
Setting safety stock levels right is vital for tank container operators such as Den Hartogh Logistics. Having too low safety stock levels results in losses of demand and will diminish customer satisfaction. Having too high safety stock levels results in increased costs, which can challenge the competitiveness of Den Hartogh Logistics. Therefore, it is essential to balance these contradictory motives and set safety stock levels just right. Within Den Hartogh Logistics, the current safety stock levels are registered in the SLM file. Den Hartogh Logistics wants a review on these safety stock levels and a scientific method for determination of these safety stock levels. To answer this challenge, the following main research question is formulated:

How should Den Hartogh logistics set their tank-container safety stock levels per tank type and apply its stock level management in the network of Europe?

It is crucial to determine safety stock levels to incorporate the demand uncertainty, the network, the multi-commodity, and the corresponding timing aspect. The pricing component is assumed to be fixed.

Design
The problem consists of three phases. The first phase is revealing the demand structure and construct an adequate forecast for DH. For the first phase a basis ETS model, a SARIMA model, and the highly advanced LOESS model are used for forecasting. Furthermore, the general demand patterns over time are investigated for DH.

The second phase is the translation of the forecast values towards actual safety stock levels for each hub and tank type. It is opted to use adaptive robust optimization for obtaining the safety stock levels. Robust optimization is suitable since no known probability adequately describes the demand data. Adaptability is required since, there exists a timing difference between demand occurrence and its actual loading data, which creates an opportunity to pool inventory between neighboring hubs. The timing difference between demand occurrence and its loading date is called "decision days" in this thesis. Robust optimization is able to meet all the constraints within a specified range which is stated in the uncertainty set for the uncertain parameters. Thus, by incorporating such a structure, Den Hartogh Logistics can guarantee a specific fill rate with their safety stock levels within the range of the uncertainty set. However, due to the large instance of the problem, some approximations methods are necessary. Especially, the adaptability and the integrality of the decision variables make the problem computationally challenging and thus require approximations. Integrality can be approximated by applying continuous relaxation on certain decision variables. For approximating the adaptability affine decision rules are used which is generally computationally fast. Also, a hybrid method using the affine decision rules in an iterative procedure solving both the primal and dual problem obtaining a more accurate solution but general against more computation time. The different approximations and their corresponding performance is evaluated. Additionally, even when using these approximations the problem size remains too large, and it is necessary to divide the network into parts. However, the effect of dividing the network into parts is likely small because only neighboring hubs can pool inventory.

The last optimization phase consists of the redistribution of tank containers towards their initial safety stock levels after demand has occurred in the most cost-efficient manner. Additionally, the corresponding input parameters for the models need to be identified through the transportation structure, the repositioning structure, the planning process, and the related cost structure.

Results
The corresponding planning procedures are mapped. It is found that around 88% percent of the demand arrives before six days of the loading date. To avoid re-planning, most MMP planners start six days before the loading date with tank container assignment. In this thesis, it is assumed that all demand arrives at two decision days. The corresponding transportation time and transportation cost are found by using the available repositioning data. However, these results are somewhat unsatisfactory with the many empty values that occur on the different lanes. These empty values will be assumed as infeasible lanes in the model. The accuracy of the LOESS model in forecasting is the best. Although, finally the ETS model is selected for the forecasting module since it is easier to understand and is only performing slightly worse than the more advanced model.

For the robust models, the model containing the hybrid approximation is outperforming the approximation only using affine decision rules. This is in line with earlier work of Bertsimas & De Ruiter (2016). The performance differs and around 40 % increase in stock and costs, regardless if continuous relaxation is applied. Including integrality restrictions results in an increase of 10% of stock levels and costs on the first and second stage of decision variables respectively. Since the hybrid approximation is more substantial than the integrality restrictions, it is opted to use this approximation. Unfortunately, using integrality restrictions on the hybrid approximation is computationally too extensive. Therefore, the model containing full continuous relaxation with hybrid approximation is opted for usage for DH. To account for the bias of continuous relaxation, the safety stock levels are rounded up.

The results of the robust model are promising and are reasonably close to the values already in the SLM. For example, the SLM uses 2640 tank containers in total, whereas the robust model uses 2880 tank containers without rounding. The robust model and the SLM do thus differ around 9%. If accounted for the continuous relaxation by rounding up, the difference is 18%. However, some exciting deviations are found. In general, the model prefers more 20_Feet_Special tank containers instead of the Swap_Special tank containers. It is possible to use less advanced tank containers to ensure lower costs in the network. Another interesting observation is the need for more safety stock in low-demand hubs. The current SLM values are too low in these hubs. However, after discussing this with MMP planners, this can be partly because low-demand hubs generally have a longer planning horizon. A more detailed analysis for each hub and tank type is present in the results section, in which the robust model is benchmarked against the current values of the SLM.

The results of the redistribution model highlight the substantial fraction of demand that needs to be repositioned back towards hub 1 or hub 39. This model can reposition back towards its initial safety stock levels.

Future improvements data quality
The data quality at Den Hartogh Logistics could be improved. The first improvement is

by starting to register missed demand. By registering the missed demand, it is possible to track better the current performance of the existing stock levels in the SLM. A second data quality improvement is to identify the demand in demand characteristics and not directly translate it into tank types. Tank types are already containing an interpretation of the demand characteristics. Furthermore, it is crucial to correct data errors in the SPO or OT to monitor the demand per demand characteristic correctly. Unfortunately, there are still some errors in this process, and this may guide to misleading input data in the corresponding forecasting models. Also, it would be advisable to mark lanes with an ISO requirement in the SPO or OT templates, making it easier to distinguish the demand of ISO tank containers directly. The last improvement on data quality is better estimating the transportation costs of empty transshipments of indirect order fulfillment. Currently, many missing values may cause unnecessary avoidance of specific lanes in the safety stock determination models.

Future improvements robust optimization
There are also some model improvements possible to obtain better safety stock level determination. Currently, the computation time in all robust models is restrictive for using the entire network. Another possibility could be the includance of the time with integrating the repositioning strategy or usage of different plannings horizons for different types of orders. This could even help Den Hartogh to integrate operational decision-making fully and have floating safety stock levels over time. However, additional research is necessary to reduce the computation time in such a large scale of problems to achieve this.

Business recommendations
The following business recommendations are given to DH based upon this research:

- It becomes apparent in section 6.3.4 that some safety stock levels may be outdated and can be updated such as the hub 7, hub 46, and hub 41. Furthermore, the provided safety stock levels by the robust model provide a basis for critically reviewing the current safety stock levels in the SLM at each hub for different tank types.
- The optimal tank type distribution is fairly close with the current tank type distribution. Although, in general it is favored to have more cheaper non flexible tank containers and slightly more tank containers overall.
- In the sensitivity analysis, it becomes apparent that individual standard deviation at hub level per demand characteristic is the most influential on the cost structure and the number of tank containers in the network. Therefore, it would make sense to invest in advanced forecasting methods since this variance is causing the most substantial costs.
- The approach and models used in this thesis is valuable for DH for safety stock determination in the future. Although, since strong assumptions are necessary, it is still required to review these safety stock levels by MMP planners.
- In this thesis, it becomes apparent that it is difficult to include all problem aspects for safety stock determination. Further automation of the planning process is difficult. Therefore, Den Hartogh should empower MMP planners with as many tools available to enhance human decision-making. Providing them with a more accurate registering of the stock levels, usage of inventory position instead of inventory level, and an accurate safety stock level could help them make better decisions.

Preface

This report represents the final project of my master Operations Management and Logistics at the TU/e. I want to use this moment to thank people who contributed to this project or in general during my studies in Eindhoven.

First of all, I would like to thank Den Hartogh Logistics. The company is very eager to learn, and the employees are passionate about their business. One of the aspects I appreciated the most at Den Hartogh Logistics is the central placement of the student. The emphasis is on your graduation, and you are the manager of the project. Additionally, Den Hartogh Logistics provided me with as much support as possible with data retrieval and understanding of the business processes. In particular, I would like to thank Luke, who supervised and supported me during the whole duration of the project. The conversations we had throughout this thesis were challenging and always in the pursuit of acquiring new knowledge. The fact that you graduated from the TU/e with a master Operations Management and Logistics gave me a reliable partner from both the business and academic perspectives. The amount of professionalism in doing your job is remarkable and cooperating with you was a pleasure.

Secondly, I would like to thank Ahmadreza as supervisor from the TU/e. The conversations I had with you were pleasant. You can motivate and inspire people with academic perspectives that are within reach of the student's capabilities. On top of that, you are very detailed in your feedback and explanation approach, aiming to deliver your thoughts truly. I also would like to thank my second supervisor Shaunak. Although we only met occasionally, your focus is on the goal and just getting things done. Especially in the modelling phase, you guided me through assumptions to keep the flow in the project.

Thirdly, I want to thank my family and girlfriend for their mental support during my studies and thesis. Writing a thesis can be challenging, and they really helped me to keep believing in myself and push through to the end. Lastly, I want to thank my friends of Het Dispuut Pegasus to express gratitude for enriching my student life in Eindhoven.

Thank you to all,

Rafaël Groeneveld, Eindhoven

Contents

List of Abbreviations

Abbreviation	Definition
3PL	Third party logistics.
ACF	Auto correlation function.
ADR	Agreement transport Dangerous good by Road.
B2B	Business to business.
BU	Business unit.
DH	Den Hartogh Logistics.
ETS	Optimizing AIC of Holt Winters model on error, trend, seasonality.
ISO	International Standardization Organization for container stacking.
LOESS	Locally Weighted scatterplot Smoothing.
LL	The business unit Liquid Logistics.
LP	Linear programming.
MAE	Mean absolute error.
MAPE	Mean absolute percentage error.
ME	Mean error.
MPE	Mean percentage error.
MILP	Mixed integer linear programming.
MMP	Multiple days material planning department.
OT	Order template used for demand characteristics in TF2.
PACF	Partial auto correlation function.
RMSE	Root mean square error
RSS	Residual sum of squares, minimization function in regression.
SARIMA	seasonal auto regressive integrated moving average.
SLA	Service level agreement with customer.
SLM	Stock level management tool.
SPO	Standard pre-order used for demand characteristics in TF1.
TEU	Twenty feet equivalent unit, equal to loading size of ISO.
TF	Transfusion the planning system of Den Hartogh.
TCP	Truck Car Planning department.

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Company description

Den Hartogh Logistics
This project is executed at DH, which is a 3PL company in the tank container industry. A 3PL is in charge of the complete logistics from the loading site to the end delivery station of the customer (Hofenk et al., 2011). Therefore, it could integrate intermodal logistics into its process. Intermodal logistics uses multiple transport modalities in one trip (Rondinelli & Berry, 2000). Modalities of logistics are, for example, train, truck, or ferries. DH is thus responsible for the complete logistic trip. Despite the responsibility for the entire trip, DH does not necessarily own all the transportation modes themselves. For example, DH could book trucks, trains, and ships of external companies to transport the tank container. However, DH always owns the tank containers used for transportation. Therefore, container placement is a vital part of DH's business and essential for serving customers effectively.

The industry
The number of companies active in this transportation process of tank containers is limited. In the global tank container industry, DH ranks as the 7^{th} largest tank container operator in 2020 (ITCO, 2020). The top 10 biggest tank container operators contribute 56% of the total market share. The market share of DH results in 8.5 % between the top 10 operators or a 4.8% total market share. Overall, DH is a medium sized tank container operator in a market with not that many significant competitors.

The ITCO (2020) reports a relatively stable annual growth rate between 8-13 % between 2013 and 2019 for the market. The total demand is thus fairly stable. Most of the orders at DH are centred in Europe and thus making DH one of the leading players in the tank container industry in Europe. A leading position at a geographical level is crucial since it can enhance a suitable network structure. Being significant in size allows quicker access to tank containers, firstly because there are more tank containers in the network and secondly because it is easier to access a tank container in a nearby hub. This advantage gives them the ability to offer high-quality service at a competitive price.

The organisational structure
DH consists of business units transporting multiple chemical substances, including dry chemicals, liquefied gasses, chemical liquids, and global. The business unit dry bulk contains dry chemical transportation, which can be chemical products such as sugar. The business unit liquefied gasses transports gasses that have been turned into their liquid state. Gasses can be liquefied by the use of compression or adjusting temperature. The

business unit of chemical liquids transports chemical liquids within Europe. The last business unit is chemical liquids global. This business unit transports the same products as liquid logistics but outside Europe or intercontinental transport between another continent and Europe. The decision to split liquids logistics and global is due to the significant trade inside the network of Europe. The focus of this thesis is only on the BU Liquid logistics and the demand of BU global with one origin or destination in Europe.

1.2 Problem description

This Master project aims to solve the main problem of setting appropriate tank container safety stock levels in Europe. The challenge for DH is to get the right tank container at the right location at the right time while minimizing the costs associated. Demand occurrence has four factors: the type of container is necessary, the location of the demand, the delivery time constraints, and the financial feasibility of the order for DH's network. The first factor of what type of tank container is necessary will be treated in the section 1.2.1. The second factor of demand's location and its timing has to do with the demand uncertainty and will be present in section 1.2.4. Section 1.2.5 will treat the features of the network impacting the profitability of an order and the interaction of the network. Section 1.2.6 deals with the corresponding cost parameters.

1.2.1 Correlated multi-commodity structure

Multi-commodity is the usage of multiple products within the same network (Even et al., 1975). DH uses multiple tank containers in the network, but it is essential to consider that the tank containers themselves are not the product. The product that DH offers is a service of transportation under certain conditions. For example, a customer requests a particular volume of a chemical from A to B with some safety restrictions and some heating equipment. In that case, it is up to DH to decide which tank container meets these demand characteristics. In most cases, multiple tank container types fit the demand characteristics. The most beneficial tank container type is not always trivial. For example, it can be better to load order with no need for a heating system in a tank container with a heating system since the order's destination is low on tank containers with a heating system. While in the other situations, the most basic tank container would be better. The matching of tank types to the requested demand characteristics is the responsibility of an MMP planner. Besides having a multi-commodity structure with multiple demand characteristics, it is also correlated. This kind of puzzling will be referred to as the correlated multi-commodity structure. Therefore, to reveal the true nature of demand, it is essential to consider the demand characteristics and not the number of orders with a specific type of tank container. This subsection will first treat all the demand characteristics, and later on, it will treat the corresponding tank types possibly fitting the demand characteristics.

1.2.2 Demand characteristics

The demand characteristics are on the SPO or the OT. The SPO was used in the first planning system, and the OT is used in the second planning system, but they are more or less the same. If it is not possible to directly meet characteristics, some renegotiating can occur to check other bussiness possibilities. However, in this thesis, it is assumed that the demand characteristics are non-negotiable. Simply because it is impossible to distinguish from the data which order is negotiable or not and if successful in which conditions it would result. Some examples of these negotiations are: downgrading or upgrading the volume, suggest another time window, suggesting another tank container type, or reselling the order to another logistic provider. The following variables are mentioned on the SPO or OT.

Volume
An order contains a specific volume of litres for transportation. The tank containers types do have different sizes to cope with the volume. For instance, an order transporting two small tank containers is changeable into one large tank container order. A small order in a large tank container is done frequently but has some prerequisites. It is essential to fill the tank container outside the 20-80 % of its capacity. Otherwise, it will cause dangerous situations with the liquid's ability to move for specific products. This rule is known as the ADR rule in transportation. However, it is possible to still transport products within the ADR range of 20-80% by placing baffles inside the tank container. Nonetheless, baffles do also have some disadvantages, limiting their use. Firstly, baffles are hard to clean. Therefore, DH limits certain chemicals for transportation in baffled tank containers due to congestion. Secondly, baffles are fixed to the tank container and are not easily be removable. Therefore, it is hard to change the number of baffled tank containers with the existing inventory.

Temperature control
Certain goods require temperature regulation for transportation. Almost all tank containers have temperature control systems, but the level of advancement differs. The standard system on temperature regulation system on a tank container is using steam tubes. However, steam can reach too high temperatures at the tubes' contact point for some products and thus require more advanced temperature regulation. Using glycogen tubes or electrical heating systems instead of the standard steam temperature systems can control temperatures while maintaining low temperatures at the contact points. Therefore, some customers require glycogen or electrical temperature regulation.

Lane requirements
The ISO sets size restrictions for stacking tank containers in transport. Stacking of tank containers is necessary for large vessels and some trains. If a tank container is going to use such a transport modality on a lane, it is necessary to have an ISO tank container for stacking. This requirement is primarily present in the BU Global. Although it may be possible to use only truck transportation within Europe, it can become too costly and infeasible. Nonetheless, covering large distances in Europe with these transport modalities can be cost-efficient and require ISO containers. Within the class of ISO tank containers, there are two size types. The 20 feet ISO tank container and 40 feet ISO tank container. The use of length measurement is sufficient since the containers' width and height have size restrictions. Although the customer does not request ISO, it is necessary to be competitive for DH on specific lanes.

Safety procedures
A tank container can have a handrail to make it easier to access the tank container on the top. Some customers require having such a handrail for safety. Most modern tank containers do have such a handrail, but older models may still not have one. The absence of a handrail occurs in ISO tank containers class. The safety policies in Europe are frequently obligating a tank container with a handrail. The global trade is lagging in this safety policy, and thus ISO tanks with a handrail do have still value in the global network.

Blacklist previous products
Although the cleaning process is extensive, some degree of rest-product of chemicals may still be present. Due to customers' quality regulations, specific chemicals cannot mix with the rest-products. Therefore, certain chemicals have a blacklist based on a tank container's

previous loading's. This problem is rare and only exists with exceptional products. The blacklist requirement thus needs to be checked by verifying the order data and is not a tank container type. The software in plannIT controls this automatically and blocks tank containers containing products on the blacklist.

Conclusion demand characteristics
The demand characteristics considered in this project are the volume, the heating system, ISO requirements and the handrail. These characteristics are selected since they are the most common in orders and the most present in the SPO or OT. Please note that the characteristics easily could be extended, but the number of order demand characteristics combinations will grow excessively, resulting in problems in the computation time of models later on.

The distribution of the selected demand characteristic is in Figure 1.1. Demand characteristic types D1, D5 account for around half of the total demand, and the other demand types consist of smaller percentages. In reality, the number of demand characteristics could be enlarged with more categories. The categories considered consist of 85 % of the total demand present at the BU Liquid logistics and the BU Global considering an origin or destination in Europe.

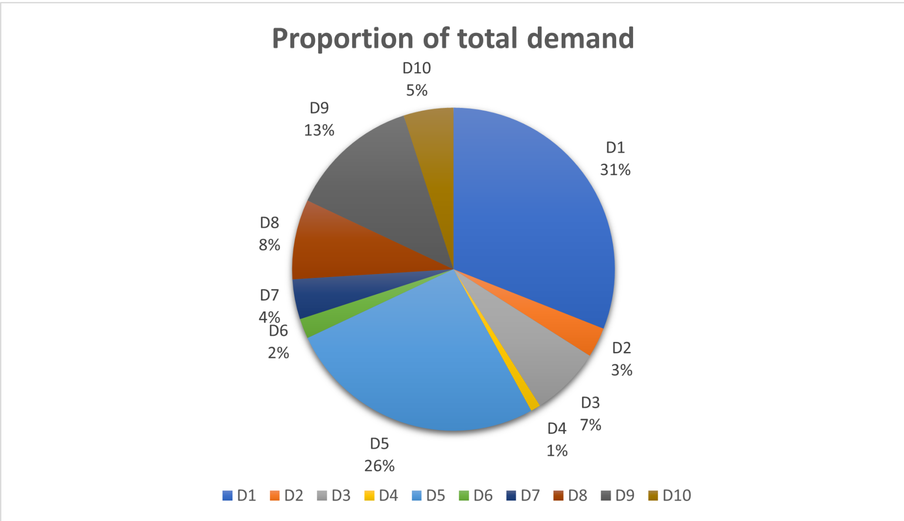

Figure 1.1: Distribution of demand per type

1.2.3 Categorizing tank container types

There are multiple different tank containers options to deal with the different demand characteristics. The different options yield multiple different container types; names of the most important types are shown in Table 1.1. The first choice is the size of the tank resulting in two different types: ISO and swaps tanks. Both tanks do have the same frame measurements. However, the swap tank has more volume by having one or two additional compartments outside the frame. An ISO tank container is capable of all transportation modes, whereas a swap tank is not.

The second option is regarding the heating system. The standard heating system is by using steam tubes, and this is not mentioned in the tank container type's name in Table 1.1. In case a more advanced heating system is installed, it will be called special. Within the ISO category, the special heating system does not exist. However, a separate category of tank containers of that size with no ISO requirement is the 20_Feet_Special. The third option is placing baffles inside the tank container or not. If nothing is mentioned in the name, then there are no baffles present. The fourth option is whether a tank container has a handrail to access the top. However, at the Swap and the 20_Feet_Special tank containers, there is always a handrail to access the top of tank container, and therefore it is not mentioned in the name of Table 1.1. Table 1.1 summarizes the tank container types which apply to this thesis.

Name	Max L	ADR Range	Special Heat	Handrail	ISO stack
20Feet_Special	23750	5000-20000	yes	yes	no
ISO	23750	5000-20000	no	no	yes
ISO_Baffled	23750	N/A	no	no	yes
ISO_HR	23750	5000-20000	no	yes	yes
ISO_HR_Baffled	23750	N/A	no	yes	yes
Swap	29260	6150-24700	no	yes	no
Swap_Baffled	33250	N/A	no	yes	no
Swap_Special	29260	6150-24700	yes	yes	no

Table 1.1: Main different tank container types

Specific container types
Although the main options are present in Table 1.1 which is around 85 % of the fleet operating in Europe, there exist also options on the: ground operated, closed top-loading, and separate compartments within a tank. These could be added to Table 1.1, but as mentioned, it would enlarge the number of types extensively for a relatively small part of the demand.

Conclusion tank container types
From the selected demand characteristics and the selected tank container types, it is possible to deduce a Feasibility matrix stating which tank container fits a demand characteristic with a 1 and a 0 for infeasibility. Due to confidentiality reasons this matrix is omitted.

1.2.4 Demand uncertainty

Contracts
The demand at DH consists of tender contracts and spot contracts. Tender contracts are typically for fixed charges and last for a longer time horizon. In comparison, spot contracts have a shorter time horizon. So the demand has a more stable part with the long contracts of the tenders and a more volatile part of the spot market. However, the exact loading time of an order is mostly not pre-determined in the contract, thus still impacting demand volatility with its timing. The tender contracts typically have an SLA, which has to be met. In case the customer demands a high SLA, it generally has to pay a higher price. At the same time, a low SLA gives DH more flexibility but generally a lower reward. It is thus essential for DH to effectively tackle demand uncertainty to ensure customer satisfaction and still be profitable. It is thus of great importance for DH to have sufficient tank containers at each location to ensure a high fill rate.

Transport uncertainty
In addition to having uncertainty about the number of orders and their timing, there is also some uncertainty in the delivery process itself. During the transport, there can be delays. DH aims to use their slogan "Getting it the first time right" by avoiding risky routes and planning sufficient slack time between switching of transport modalities. Therefore, the duration of the trip will be generally more prolonged, but with a much higher on-time

delivery rate and less rerouting costs. Therefore, transport uncertainty will be out of scope for this thesis.

Forecast global imbalances
Global imbalances of tank containers exist due to differences in the import and export of a continent. Asia, for example, has a surplus in tank containers since more tank containers get imported than exported. The network managers of DH meet once every nine months and try to forecast these global imbalances. Each network manager states the expected imbalances and changes in repositioning policies between continents. Additionally, these meetings contain information surrounding general growth patterns and seasonal influences on the global and continental markets. Adjusting the fleet distribution over the continents may be a possibility regarding the implications of the market. The number of tank containers in the network of Europe is thus not constant and affected by changing global imbalances, seasonal influences and trends. These imbalances need to be adjusted by appropriate pricing or repositioning policies. Overall, this mostly happens in the European ISO category because of its ability to use container ships.

Operational forecasting
A short-term forecast for Europe is not available at DH. However, the following procedure is used in operational stock level management. The MMP planners are responsible for sub-regions in Europe containing multiple hubs. Each MMP planner has regional knowledge and is aware of the contracts and other frequently occurring orders. Therefore, a planner tries to match his expectation of occurring demand within his sub-region with the current existing inventory. The existing inventory is found in SLM, which provides insights into the stock levels at DH over the hubs in Europe. Minimum and maximum stock levels are set in the network by the planners' experience and are incorporated in the SLM tool. The SLM tool visualizes the shortages and surpluses by assigning colours indicating its deviation from its desired stock level as shown in Figure 1.2.

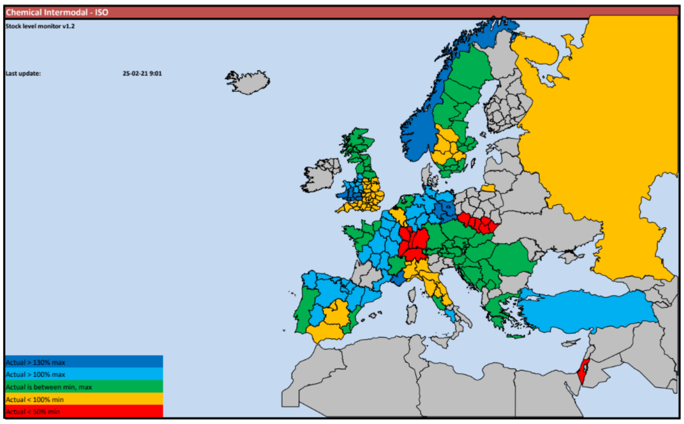

Figure 1.2: Stock levels Europe of certain type tank container

1.2.5 Network

Figure 1.2 shows the network structure present in the SLM. The network structure is important for several reasons.

- A shortage of one hub is solvable by a surplus of nearby other hubs. Transportation times between hubs are crucial for determining which possible hubs can supply other hubs on time.
- The network is not equilibrium due to the inflow and outflow of a hub is not equal to each other. Regardless of the safety stock levels, it is necessary to relocate tank

container towards hubs with a negative balance on import and export after some time.

- Hubs central in the network are ideal for having stock since it increases tank containers' assignment flexibility. The likelihood of assigning to an order increases since these hubs have more links to demand possibilities.
- Having stock at the network's boundaries is not attractive since it has only a few links to demand observing hubs. However, due to the difficulty accessing the network's boundaries, it is necessary to have safety stock at these borders.
- Some hubs prefer stocking due to their low cost for storage or fast transit times to other hubs. An example is Turkey in the network of Europe. The costs for storing are low, and the costs for repositioning decrease by the easy access to the whole Mediterranean area.
- Having a dense network makes it harder for a competitor to enter due to offering unprofitable prices. Therefore, DH must effectively manage to invest in some parts of the network and possibly reject other parts of the network.

1.2.6 Cost efficiency

The quoted costs

Previously, the price-elasticity of the global demand of tank containers at DH was investigated by (Holle, 2019). With pricing policies, it is possible to adjust demand due to its elasticity. Table 1.2 displays the pricing components of an order.

Type cost	**Explanation**
Direct job cost	Consist of cost items that are directly traceable to the order. For example, fees for trains or vessels, hiring drivers and trucks, or operating terminals.
Equipment cost	The cost of having operating equipment of DH in the order.
Overhead cost	This general fee covers general costs in the company such as supportive business functions or employees.
Repositioning saving	The cost that might occur of repositioning, after the order has arrived at its destination. In case the order arrives at a more distant hub, it is not beneficial for the network of DH.Therefore, DH quotes additional costs to cover the possible empty mileage.
Repositioning contribution	A discount by avoiding repositioning costs due to a beneficial origin of the order. The exact opposite of the reposition contribution and determining the possible benefit of the origin of an order.
Expected demurrage	In case the customer keeps the tank container longer than a suggested time frame at the delivery site, it will pay an additional fee which is called demurrage.
Market correction	Based on origin and destination a penalty fee for discouraging certain container flows. Generally, these market corrections are for balancing import and export flows of continents.
Quoted network margin	The profit margin in the order.

Table 1.2: Pricing components of an order

Once the account manager and customer both accepted the order at DH, it is up to DH to operate as cost-efficient as possible. This thesis will assume that pricing is fixed and that creating additional demand by lowering spot rates is not a possibility. This thesis will thus merely focus on the costs that can vary due to the efficient execution of operations in the network.

1.3 Research questions

The main research question central to this project is:

How should Den Hartogh logistics set their tank container safety stock levels per tank type and apply its stock level management in the network of Europe?

In order to answer the main research question, the following sub research questions are formulated.

RQ1: What are the current policies within Den Hartogh surrounding stock level management of tank container distribution in Europe?

Answering this research question should give a complete conceptual understanding of the problem. For instance, the procedures in determining tank types, order handling, forecasting, and fleet management should reveal the corresponding decision structure for safety stock level determination.

RQ2: Would it be advisable for DH to revise some of the stock level management policies?

This research question aims to link scientific literature available on stock level management with the current business process at DH. It is vital to evaluate conceptual design solutions to improve business value for DH. As a company deliverable, it would be valuable to find conceptual solutions for reducing multi-commodity and demand uncertainty or solutions enhancing the network structure of stock level management.

RQ3: Which assumptions and relaxations may be crucial for modelling while still maintaining a representative business environment?

The answers to this sub-question serve as a preparation for the modelling phase. It is a guideline on the musts of inclusion in the model and which details are less relevant. It helps build a structure in the model that sustains as much business value as possible. Additionally, it helps to identify the most influential assumptions for relaxation.

RQ4: How is the demand at DH distributed in Europe over the hubs, and what are the cost and transportation functions between hubs?

This sub-question aims to gather the modelling and optimization stage's input data. The input data gathering consists of three main components: the demand process, the cost structure, and the transportation structure. The demand data will provide insights into the surpluses and shortages on the hub level. Furthermore, the development of a basic forecast will serve as an input for the modelling and optimisation phase. The second component is the cost structure. Some examples of these cost structures are: finding the additional transport costs per hub of serving demand not directly from stock, finding the holding costs at each hub, and finding the penalty costs of not meeting demand. The last component is finding the transportation times between the hubs.

RQ5: How should DH determine its decision parameters for its safety stock levels?

This sub-question aims to translate the problem from a qualitative nature to a quantitative nature and apply scientific models to identify reasonable safety stock solutions. The answering of this sub-question has a cyclic behaviour by trying to improve each time.

RQ6: What is the sensitivity of the design parameters in the found setting, and how should Den Hartogh react?

Investigating the robustness and validity of the found solutions and provide guidance for standardising stock level management is the goal for this sub-question. It gives insights if the found input data or decision parameters slightly deviate. Such studies ensure a stable business environment and help DH react to changing environments. Furthermore, it allows DH to prioritize which parameters are the most important. Lastly, sensitivity analysis helps to evaluate implementation risks and guidance for further improvements.

1.4 Business relevance & company deliverables

Currently, the safety stock levels at DH are determined by expert estimates in the SLM file. DH wants to verify scientifically that these stock levels make sense. The company deliverable of this thesis should thus be accurate estimates of the current safety stock levels. Since these levels are changing due to varying demand patterns, it is crucial to create sustaining value in the future, thus corresponding procedure of finding such safety stock levels is also essential. In case a revised safety stock policy can boost operational utilisation by 1 %, it results in treating 1.56 % more demand and using the mean profit per order indicates that it is a significant amount for research. Since DH occasionally has to reject demand and can generate slightly more demand with lower prices for some orders, it seems realistic that additional capacity will sell. Maybe it will be to a smaller profit margin, but still, it would make a significant impact. Another possibility is that operational utilisation can be improved by having fewer tank containers while still satisfying the same demand levels. Overall, a fleet reduction could save an significant amount in euros for DH yearly. Although both rough calculations may deviate from these numbers, it is clear that only a slight improvement in operational utilization would result in a significant additional amount of profit for DH. Thus researching this topic is of significance for DH from a business perspective. Despite the monetary value in the BU's Liquid logistics and Global, it could have value for other BU's since these BU's face similar characteristics. Please note that, the actual monetary values have been omitted due to confidentiality reasons.

1.5 Thesis outline

This thesis has the following outline. Chapter 2 shows the relevant literature for all the research questions. Chapter 3 discusses the methodology. Chapter 4 report the first results regarding the obtained data. Chapter 5 uses the results in the data found to design an appropriate model. Chapter 6 displays all the results. Finally, Chapter 7 discuss the business implications on the found results of the models and their limitations and implications.

Chapter 2

Literature study

The first section will treat general literature on inventory management and industry characteristic. This literature is necessary to answer the first research question and to benchmark the current performance of DH. The second section will treat literature on forecasting and find the corresponding demand patterns that treat models for research question 4. The third section is the central part of this thesis and describes different modelling approaches to finding the safety stock levels and answering research questions 3 and 5.

2.1 Inventory literature

Inventory control policies
It is essential to define what is considered as safety stock. Safety stock is mostly used to cover uncertainties in demand or its delivery process (Snyder, 1980). It should hedge against a certain degree of uncertainty. DH experiences mainly uncertainty in the demand process and therefore needs additional stock. To have sufficient stock a minimum stock level is present in the SLM. This stock level is containing some expected demand and some reserve stock for covering high demand time periods. DH is thus looking for a certain stock level at each hub which satisfies a certain degree of overall fill rate. The definition of safety stock for DH includes the expected demand and additional part for covering for covering more uncertainty than expected. Bundling the expected demand and its safety stock is done for operational purposes in the SLM for the ease of benchmarking. Another method to reduce demand uncertainty is by the use of strategic incentives. Some of them are mentioned in the book of Nahmias & Olson (2015). An example of tackling uncertainty for DH could be by the usage of supply chain coordination or by incorporating more business flexibility. An example of these strategic incentives at DH could be a discount for customers to determine their time windows early.

The basestock policy in inventory control method in the container industry is done by Buhayenko & Den Hertog (2017). In Figure 2.1 on the left hand side, the corresponding continuous basestock policy (s, S) is shown. If the inventory level drops below s, an order will be placed until the inventory rises towards the level S. However, the additional stock will be delivered after L time periods. On the right-hand side of Figure 2.1, the stock level of a hub in the Network of DH at the start of a day is shown. A few remarks on the difference of the stock level at DH and the continuous basestock level policy. Besides having an outflow of tank containers, there is also a natural inflow of products. Therefore, an increase in the number of tank containers can be the natural inflow or a forced inflow by their repositioning strategy. The outflow consists of the hubs own demand and possible transshipments for fulfilling the demand of its neighbors. It is also a possibility that a shortage of one tank type can be solved with the inventory of another tank container type, which is not deducible from a single inventory level graph. Finally, the data for the stock level at DH is only saved in the database for a specific time point, and Figure 2.1 may suggest fixed review models. However, the inventory strategy of adjustment action of the stock level is monitored throughout the day by MMP planners and is thus continuous. The mathematical framework of the (s, S) is applicable but should be tailored to the industry.

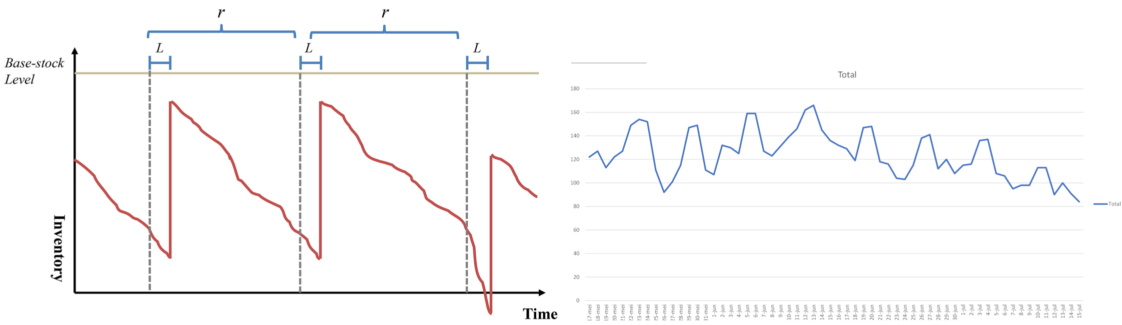

Figure 2.1: Left: Basestock policy (Lee et al., 2018), Right: DH hub stock level

Tank container industry

A general overview of frequently occurring problems in the intermodal 3PL industry is the work of Powell (2003). It describes which resources are present such as the differences in tank containers and the possible service requirements of customers in the industry. Furthermore, the most critical decisions for the 3PL industry are present on operational, planning, financial, and informational level. Knowing these decisions helps to identify the modelling parameters. Lastly, it describes the control functions of planners, information systems, and agents in the tank container industry. This paper contributes to the understanding of the industry and quickly identifying what types of decisions are vital.

The paper of Karimi et al. (2005) starts to sketch the most common problems in this specific industry and proposes an LP model in positioning and hiring tank containers. This paper includes many details such as multi-commodity, time windows, and cost differences in loaded and empty transport. Therefore, it is excellent if the demand is known and a schedule must be found. Nonetheless, since it is not including uncertainty, the value for DH is limited.

A more mathematically orientated analysis of optimizing intermodal container logistics is the Ph.D. dissertation of Sharypova (2014), which discusses several critical problems within intermodal logistics, such as planning horizons, strategic cooperation initiatives, and empty repositioning problems. The emphasis of her work is on the scheduled service design network and creates different models for doing so with the formulation of multiple constraints. A downside is her considerably small problem size, which is, in practice, not realistic. The last chapter of her dissertation contains a novel approach using game theory. By incorporating game theory, she makes it possible to view the complete supply chain of competitors and incorporate the pricing component. Nonetheless, uncertainty is again not used in this project.

A simulation approach for the tank container industry is present in Young Yun et al. (2011). With the use of discrete-event simulation, the performance of several (s, S) policies is tested, and empty reposition transactions and stock levels can be determined in that paper for the tank container industry. By varying the instances of the decision variables, it is possible to identify a good solution for DH. Simulation has the advantage of being able to simulate the company environment very realistically, but the downside of not being optimal and its hardness in significantly large instances. However, as the number of hubs increases, the corresponding possibilities explode and thus its simulation time. A genetic search algorithm was developed to evaluate only valuable solutions. The genetic search algorithm keeps in mind favourable parameter settings and samples solutions with a slight deviation in the previously found favourable setting. The heuristic is called genetic search because it has many similarities to genetic evolution by selecting good genes and dropping unfavourable ones. By using this algorithm, it is possible to evaluate more significant instances in simulation, and this is done for the container industry by (He et al., 2015) and (Jung et al., 2004).

Actual system implementation of finding the appropriate safety stocks and assigning the right amount of tank containers is in the article of Epstein et al. (2012). The procedure for executing such a project is first to determine a forecast and then use this forecast in the corresponding safety stock determination using a network flow model. First, determining a forecast and using this forecast in a network flow model is very appealing. The downside may be the assumption of the normal distribution to determine the appropriate safety stock levels. Additionally, different tank types and demand types are not considered. Lastly, a MILP assigns the right tank containers towards orders. The application structures the process well and gives valuable implementation strategies and managerial insights, which may be helpful for DH when DH opts for implementation. On top of that, the risk of implementation and the financial benefits for that particular company are quantified, which provide support and guidance for implementation.

2.2 Forecasting literature

Before selecting the appropriate forecast model, it is necessary to look at data availability. Currently, the data available at DH is only data on the sales of tank containers. If additional explanatory variables contain additional information on the influence of the demand on tank containers, a multi-variate forecasting model is possible. However, DH is currently not aware of these variables. Since the emphasis of this thesis is on how to deal with variability in the determination of safety stock levels, the search for additional explanatory variables for forecasting is out of the scope of this project. Therefore, it is decided not to treat the category of multi-variate time series forecasting and focus on univariate forecasting methods.

Time series characteristics
The book of Hyndman (2018) describes the procedures of forecasting extensively and treats a wide selection of forecasting methods. Before selecting a model, it is essential to look at the graphics of a time series. The graphics of a time series could contain a trend, seasonality, cyclic behaviour, heteroscedastic error terms, or specific outliers in the demand data. Several data transformation methods are used to remove these patterns or select models that incorporate these features if these features are present. The goal of forecasting is to obtain a forecast as accurately as possible and useability in practice. Selecting a forecast is thus based on its accuracy and practicality.

Holt Winters
One of the most basic forecasting models is the exponential smoothing method. By using an exponentially decreasing weight factor for the demand further in the past, the future demand is estimated (Gardner, 2006). These models can be extended by using additive or multiplicative Holt-Winters to deal with seasonality and other trends (Chatfield, 1978). The multiplicative Holt-winters is more suitable in case the seasonality changes proportionally to the level. In contrast, the more common additive model is used by a more stable

seasonality (Hyndman, 2018). The error trend seasonality models (ETS) are optimizing the parameters of the Holt-Winters model on the Akaike (AIC) or Schwarz (BIC) information criteria. Both information criteria contain a term for the amount of parameters in the model and the log likelihood of the model. These criteria aim to find the best fitting model, but penalizing more parameters in the search of are more parsimonious model. ETS models select forecast models on optimizing this criteria in the search of optimizing a good parsimonious model. ETS models are generally easy to computationally obtain, create a proper fit in most cases, and the corresponding parameters of the model are easy to interpret.

SARIMA models
A more complex method to obtain forecasts is using linear regression on previous demand values and previous forecast error terms to capture auto-correlative relations. These models first remove trends or seasonality by making use of finite differencing. Eventually, the heteroscedastic error terms can be removed by applying the box-cox transformation Nelson & Granger (1979). This class of models are known as the SARIMA models. These models can generate a better fit than the ETS models, but the corresponding parameters may be harder to interpret. For example, if the time series has a trend or seasonality, the time series is transformed. The coefficients are an interpretation of the transformed time series and considerably more difficult to understand.

LOESS regression
A possible disadvantage of a SARIMA model is its adaptation to sudden changes and varying parameters. A SARIMA or ETS model has fixed model coefficients over time. A LOESS regression is a non-parametric technique and uses regression but only uses a subset of the data and applies weighted least squares to find a fit with the data (Atkeson et al., 1997). By doing so, this type of forecasting can alternate trend levels and quickly adapt to varying parameters. The algorithm is quite complex and may be hard to implement and explain. However, in most cases, it will result in a better model.

Other univariate forecasting models
Other highly advanced forecasting methods are also possible such as neural networks with auto-regressive terms (Zhang, 2003), Bayesian forecasting (Phillips-Wren et al., 2010) or spectral analysis (Grzesica & Wicek, 2016). However, these methods are considered beyond the scope of this master thesis. This master thesis emphasises the determination of safety stock levels based on a forecast.

2.3 Safety stock determination literature

Deterministic models
As mentioned, the problem has four main characteristics: demand uncertainty, network structure, multi-commodity, and the time windows. If the demand uncertainty gets omitted, the problem will just be reduced to a significant scheduling problem as already solved by (Karimi et al., 2005). These problems are generally transformed into network flow problems solvable by a MILP if the cost structure is linear. Another excellent example of operating such a model for a container operator is present in Lambert et al. (2009). These types of problems can deal with relatively large instances in a reasonable amount of time. Ignoring the network structure, multi-commodity or the time characteristic will only reduce the size of the problem and making the problem more straightforward. In case the costs are non-linear, the problem becomes much more difficult. Such non-linear problems generally needs to be significantly smaller for solving and can, in most cases, be solved by the use of Lagrangian relaxation, which has been done by Ghadimi et al. (2020). To reduce the size in these non-linear problem ignoring a problem aspect such as

the network, structure, multi-commodity, or the time aspect makes the problem instance much smaller and better solvable. In some cases, it is even possible to solve these problems exact by deriving derivatives and proving convexity as has been done for this industry by Li et al. (2007). However, the demand uncertainty is too high at DH, and therefore these deterministic models are inappropriate.

Models incorporating demand uncertainty
The simplest model with demand uncertainty has no multi-commodity, no time windows, and no network structure. Solving this model result in the news-vendor equations (Petruzzi & Dada, 1999). Adding time windows with these types of models result in the basestock models with an example in Silver & Silver (2017). Incorporating the network structure, multi-commodity, or other problem features combined with the time aspect generally will result in large Markov state-space models (Wensing et al., 2018). On top of that, the exact demand probabilities must be known to solve these models. To tackle these types of problems, three mathematical techniques are considered in this thesis: stochastic optimization, robust optimization, and simulation. These methods require much computation time, and using the right aggregation level is a must. Making your model detailed gives the ability to represent reality better. However, this representation can cause problems in computation time, resulting in computational in-feasibility. In case a higher aggregation level is considered, it may be possible to analyze the full problem. Therefore, choosing the right aggregation level is important for the quality of the solution. Besides changing the aggregation level, there also exist heuristics to decrease the computation time in these types of problems with a review of some methods in He et al. (2015).

2.3.1 Stochastic optimization

Stochastic optimization assume a certain probability distribution for uncertain parameters and tries to optimize these problem incorporating the uncertainty. Whether this method is appropriate depends on the demand at DH's which will be reviewed in Chapter 4. Once the parameters of the demand distribution are found and seems reasonable, a stochastic program can be formulated with uncertain parameters. The book of Powell et al. (1995) contains different stochastic programming models in the logistic industries. This book briefly sums the possibilities in modelling and encountered problems in practice and should help select the appropriate model or DH. A more concrete example of safety stock placement using stochastic programming is Schuster Puga et al. (2019). Typically in these types of problems there exist a time window, the stochastic program contains multiple stages to deal with different timing in decisions. So the first stage contains decision variables made before the realisation of the stochastic parameters, and a second stage contains decision variables made after realising the stochastic parameter. If the stochastic parameters are correlated with each other, it is necessary to fit a multi-variate distribution in stochastic optimization. An example of using such a multi-variate distribution is present in Bertsimas et al. (2010).

2.3.2 Robust optimization

Robust optimization does not assume any probability distribution of the uncertain parameters in the model and is thus more general in comparison to stochastic optimization. Generally, robust optimization can solve an optimization problem with uncertain parameters that are allowed to vary in a specified interval (Ben-Tal & Nemirovski, 1999). Robust optimization searches for a solution in which all the constraints are met under all possible outcomes of the uncertain parameter stated in the uncertainty set. It is essential to understand that the solution of a robust optimization may not be optimal in many demand scenarios, but at least the solution found is feasible in all scenarios. Since safety stock levels are set to cover also a certain degree of uncertainty (Snyder, 1980), robust optimization may become helpful to ensure having enough stock in a specified range of demand. Choosing a large size of the uncertainty set will result in a solution including much inventory, avoiding any risks of stock-outs in the most extreme scenario. Setting a too small size of the uncertainty set, makes it more likely that the true value of an uncertain parameter is outside the range of the uncertainty set. Therefore choosing the proper uncertainty set is an important decision and dependable on the objective of the company and the corresponding data structure (Bertsimas et al., 2018). An excellent review of all the current research on robust optimization is present in the paper of (Gabrel et al., 2014). If the problem does not have any sequential decision-making, it is robust static optimization, and if it has sequential decision-making, it is called adaptive optimization. The paper of (Gabrel et al., 2014) provides a lot of solution methods for both types of robust models. For example, papers that solve robust optimization problems by rewriting the problem with a robust counterpart (Gabrel et al., 2014). A robust counterpart adds additional variables to the original problem and makes it solvable as a regular LP or MILP with these extra variables. However, formulating these robust counterparts can be difficult.

Choosing the uncertainty set
The most straightforward uncertainty set is a box uncertainty set. The box uncertainty set simply restricts the uncertain parameter only to vary a certain amount (Chassein et al., 2018). An extension of that is the budget uncertainty set, which creates an additional restriction on the total amount of variation of all the uncertain parameters (Chassein et al., 2018). In many practical application of robust optimization, the found standard deviation of uncertain parameter is used in determination of the limit of a uncertainty set (Moon & Yao, 2011). Other variations using ellipsoid or even polyhedral uncertainty sets is also an possibility, but generally more complex in modelling (Bertsimas & Brown, 2009). However, when the observed demand displays multi-variate behaviour, advanced construction methods for the uncertainty set are necessary. Advanced machine learning methods on constructing these uncertainty sets are present in the article of Goerigk & Kurtz (2020), and Shang et al. (2017).

Computational challenges
Including all problem aspects within a robust framework may be computationally challenging. These models are advanced and require a lot of computation time (Ben-Tal et al., 2004). The work of Bertsimas & Sim (2003) is important in this area since he was able to reformulate these problems by introducing a robust counterpart. Such reformulation techniques will be required to solve problems of large instances. Other approximation techniques such as the affine decision rules could help (Ouorou, 2011) or the finite scenario approach (Marchetti-Spaccamela & Segal, 2011) to keep the problem tractable.

Robust optimization in tank container industry
The work of Erera et al. (2005) not incorporating demand uncertainty is improved by Erera et al. (2009) by the use of uncertainty sets in robust programming. These two papers are both on the empty repositioning of tank containers. With the corresponding uncertainty set, the problem can be reformulated using its dual problem into a MILP again but incorporating demand uncertainty. The approach and mathematical proof for reformulating such problems are in Bertsimas & Sim (2003). The work of Erera et al. (2009) distinguishes three inventory pooling strategy cases: no inventory pooling between hubs, with inventory pooling between hubs, and have flexible large hubs in the first stage which redistribute over smaller hubs in the second stage. However, a downside of the papers in respect of the problem of DH, is not including tank types and not focusing on safety stock levels, but on the repositioning decisions instead.

Chapter 3

Methodology

3.1 Finding the current inventory policy

The current inventory policy at DH is to be established by conducting interviews with the corresponding MMP planners. The procedures used in tank allocation and redistributing will be sketched in business process diagrams. These allocation procedures are to be examined against the found literature, and if possible, improvement directions for reducing the complexity and demand uncertainty in operations are given.

3.2 The forecast

The demand data of DH needs to be divided into different demand categories. After obtaining these categories, it is possible to forecast the demand per demand category for each hub. Before making the forecast, the corresponding features of the demand such as trends, seasonality, cyclic variation and possibly heteroskedasticity need to be explored. Doing this on beforehand makes it easier to select the right forecasting model. After verifying the model performance and its practicability of the SARIMA, ETS, LOESS forecasting models, it is decided to use the ETS forecast. Since the emphasis of this thesis is not on forecasting, the forecasting model formulation and performances are in Appendix C. Additionally, the forecast errors are to be investigated to see if the demands of hubs are correlated and if possibly a probability distribution can be fitted on the error term. In case a probability distribution can be fitted stochastic optimization is ought to be applied and if not robust optimization.

3.3 Setting the safety stock levels

In case the demand displays multivariate behaviour, a multivariate stochastic program is applicable. However, if the distributions of the forecast error term of the hubs differ, robust programming is the method to use. It will be come apparent that no probability distribution can be fitted and thus robust optimization is the method to use. The different uncertainty sets will be tested to evaluate the model performance. Furthermore, the problem characteristics such as network, multi-commodity, and time windows are gradually introduced in the optimization models. To reveal the sensitivity of the design parameters, input parameters will be varied. For example the fill rate and the forecast error. By doing so, the sensitivity's relation to the total costs becomes more comprehensive. Lastly, the redistribution of tank containers towards their initial safety stock levels which is not an main objective of this thesis is in Appendix F.

Chapter 4

Data analysis

Section 4.1, describes the IT structure at DH and how it is possible to obtain the data used in this project. Section 4.2 explains how the raw data can be transformed and improved, making it ready for usage. The last section describes the first results of the data without interpretation of models.

4.1 Data description

Types of data sets
There exist three types of data sets that can be merged using their unique order ID. The first data set is the order data. The order data contains transportation data, the order arrival date, loading date, delivery date, origin, destination, the tank container assigned, and product related material. The second type is the SPO or OT data. This data contains information on requirements of transportation that the customer and DH account manager agreed upon. The last data set is its financial data; the corresponding quoted prices to the customer, which are mentioned in Table 1.2. Additionally, the actual direct related costs to that order are in that data set for comparison. Merging these three data sets is possible in Microsoft Excel by using sequential bundling functions on their order IDs.

Timing of data sets
The time span of representative order data is relatively small. The data originating before March 2017 is not representative. At March 2017, DH merged with another company and as a result had completely different demand data patterns. The end of the data set is April 2021. In total there is four years of data available. Additionally, it is important to mention that the corona pandemic may have affected demand from March 2020 till April 2021.

Switch of IT system
It is important to recall that this thesis uses data from two different BU's. In the beginning, both systems were running on Transfusion 1 (TF1). However, the BU liquid logistics switched to the more advanced version of Transfusion 2 (TF2) midway 2019. In TF1, the customer requirements were registered in the SPO. Whereas in TF2, the customer requirements were registered in the order template (OT). Since the standard pre-order (SPO) has been updated due to some errors, its improved version is present in TF2 as an OT.

4.2 Data preparation

4.2.1 Merging locations

The amount of hubs present in the operational SLM is 48. However, the demand consisted of much more destinations. Since all cities outside Europe belong to the business unit global, these hubs have been merged in the hub GLOB. After merging these cities, there were still 92 city names remaining. Although a hub consists of a direct city location, DH bundles some cities together. So the corresponding cities of the hubs had to be filtered and renamed to the hub in the SLM. Due to confidentiality reason the map containing all the hubs in Europe, is omitted.

4.2.2 Finding demand characteristics

There are ten different demand characteristics. These demand categories must be mutually exclusive. If they would not be mutually exclusive, an order could arise in multiple demand categories, resulting in multiple orders while it is just one customer order. This section will explain how each demand category is distinguished from the SPO or OT customer data. In the SPO, each variable contained a 0 for no and a 1 for yes. The OT contained for each variable the values yes, no, null. Where null means the customer is indifferent. The OT has also been converted to the 0 or 1 system.

ISO
The variable ISO required in the SPO/OT means that the customer requires an ISO tank container. Nonetheless, the SPO and the OT report a large deviation in percentage of total orders for this variable. In the SPO, this is 8 %, whereas, in the OT, this is 33 %. The MMP planners indicate that this difference does not exist in reality. The problem cause is related to customers checking the boxes of some requirements just to be sure. Furthermore, some of the boxes are automatically checked in the OT, resulting in a much higher requirement rate because customers forget to uncheck them. Therefore this variable is considered useless. However, ISO requirements are primarily obligated by lane usage; often the customer is not aware of this and is not the one to assign the ISO requirement to an order. Therefore, it is possible to estimate ISO requirements only on lane occurrence. The hubs of Russia, GLOB, Piraeus, Gebze, Mediterranean are almost only accessible using ISO tank containers. Therefore, the additional variable ISO gets created, which has a value of 1 in the situation the origin or destination is one of the above mentioned hubs. Please note that, in reality also other lanes may have the ISO requirements. Nevertheless, the number of lanes is 48^2, and on each lane, there are multiple route options, making it extremely time-consuming to verify each lane manually. A fair amount will be estimated using these hubs, but it will probably deviate a bit from the actual situation.

Special
Whether orders belong in the special category is determined through four variables. There must be a special heating requirement, meaning that steam is not allowed, and instead, a special heating system using electricity or glycol is required. Therefore the filter combination of heating required = 1, steaming allowed = 0, glycol allowed = 1 or electrical allowed = 1. Due to data inconsistency in the electrical and glycol allowed criteria, these settings are omitted. The result should represent reality since orders needing even more advanced heating systems such as hot rosin orders are already omitted from the data.

Small/Large
DH characterises all tanks containing a volume above 24.700 litres as large. All containers below will be considered small. It could be possible to divide size differences into more categories. However, this would increase the dimensions again, and in the SLM file itself,

this classification is not used. Using different size categories compared to what is used in the SLM would make it hard to benchmark model performance against SLM stock levels.

Baff
Baffling is necessary when the tank assigned to the order has a volume within the range of 20-80 % of the maximum capacity. Since the tanks differ in size despite being in one size category, the real need for baffling can deviate a bit. Therefore the baffling range is set to a level in which all tank container needs to be baffled. This range also makes sense from a business perspective because if one container needs to be baffled in a large tank container, DH is most likely to assign a smaller tank container, which does not have to be baffled. The range for baffling is set if the requested volume is within the range of 5.200-19.200 litres.

4.2.3 Handling outliers

All the demand records have been included in the analysis since it includes an order. However, sometimes other variables did contain incorrect information. For example, the transportation time between two hubs was, in some cases, negative. These negative values are a data recording error and are therefore omitted in the travel time analysis. In other cases, the transportation time between two hubs was considerably large; in that case, the order was also omitted. For the BU global orders of a duration longer than 225 days and for the BU LL, 75 days was considered too large. Although, these durations were probably correct. Some customers do have demurrage or require having intermediate access to a product. Therefore the tank container stays fully loaded at sites or depot nearby the customer, ready for direct retrieval. This service causes long durations, but the customer directly pays for this long duration. Nonetheless, in this project, the travel time is only relevant; therefore, these kinds of orders can be omitted. Like the cost data, sometimes a series of orders got booked on a single order ID, resulting in non-realistic profit generation. These orders have also been omitted since they were bundling series of orders which were not tractable. Thus, in case an order got a profit contribution of 12.000 euros, which is unrealistic, it got omitted.

4.3 Data analysis

This section will present the data before it gets used in the modelling phase. This section aims to get an understanding of the network and the operating costs. The part of the demand distribution will be present in section 6.2 since the demand requires adjustments for trend and seasonality and thus is linked with forecasting.

4.3.1 Duration

Loaded transport duration
Within DH, there exist two types of transport. The first type is loaded transport, in which the tank container is filled and directly assigned to a customer. The second type is empty repositioning. Empty repositioning is not necessarily coupled to a customer, but only delivering the tank container at a hub in shortage of tank containers. The duration between loaded transport for customers and empty repositioning on the same lane will differ. Firstly, because the customer needs to load and unload at their delivery site which is causing additional time. Secondly, some customers require a tank container to wait fully loaded and pay an additional demurrage fee so that when they require the good it is available immediately. Lastly, to prevent tank containers from arriving late at their

destination, DH aims to plan slack time and only use reliable routes in loaded transport. In contrast, in empty repositioning, this is required to a lesser extent.

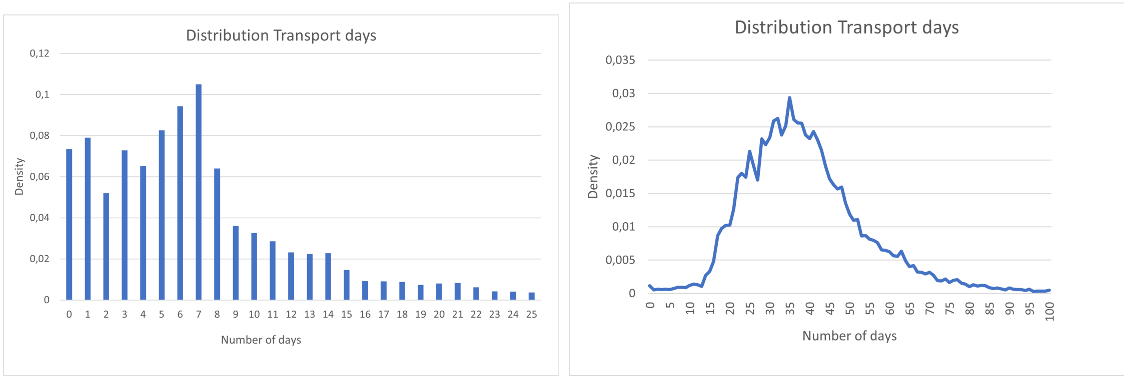

Figure 4.1: Left: Days transport LL, Right: Days transport Global

In Figure 4.1 the corresponding mean loaded duration of a customer order is shown. The duration on a lane is not fixed, since there are multiple routes and the exact time of a route can also depend on weekdays. On average, an order spends nine days in transport for the LL and 33 days for Global. Furthermore, the number of days necessary to transport the tank container to the following order is on average 13 and 55 days for LL and Global.

Empty transport duration

In Transfusion, it is also possible to find the empty repositioning movements and their duration. It is easily obtainable by using a pivot table on the data with the variables HubFrom and HubTo. In case a lane combination does not exist, it will be fixed to a high value. This high value will make the lane very unfavourable or unfeasible. This situation is realistic since this lane may not be financially feasible for DH, or simply no demand occurs on that specific lane.

4.3.2 Transportation cost distribution

Due to the intermodal transportation, the cost of transportation are not accurately predictable by the amount of kilometer or its duration. Many operators do have different freight rates, which differ inside a considerable interval. Using the financial interal KPI JF pivot dataset of DH, linking these costs with the order data is possible. Only the long-range transport modalities such as the trains and ferries are considered. These transportation modes are better for long-range transport because moving by truck is the most expensive mode to cover distances and is mainly used for shunting between terminals of transport modalities. These shunting costs will thus always occur for an order and are therefore less relevant in decision making. The absolute difference in the amount shunting costs are too hard to estimate reasonably and differs too much for each individual lane to incorporate. Lastly, the cost rates between empty transport and loaded transport differ. However, the empty rates are not directly available, and therefore only the full rates are considered in this thesis.

4.3.3 Other costs

Holding cost

The holding costs are found by searching for each hub the depot manually in TF. Holding a tank container at a depot varies highly in price. In most cases, it is a small amount to be paid. However, these depots want to avoid stocking in the long term. Therefore, they increase the price based on its duration. An example would be for the first five days free, the second five days 6 euros per day, and after that, it gets fixed at 15 euros per day. However, using the number of orders a tank container does per year and the duration of a trip, it is deducible that a DH tank container does not frequently stand at a depot for more than ten days. Therefore, the ten-day rate is picked as holding costs at each depot to ensure a linear behaviour of holding costs to ease computations. The assumption of directly applying the ten-day rate may be a bit questionable, but since the holding costs are within the range of 0-12 euro daily rates, which is low compared to the other costs, it does not matter that much. If the depot did not have any empty depot costs on its price list, the mean of all the holding rate of 4.92 euros was picked.

Opportunity cost
The opportunity costs refer to the costs for DH that are associated with owning a tank container. These costs include all costs related to the tank container, such as the investment costs, maintenance costs, and inspection costs. The finance department calculated this already, and Table 4.1 displays the daily cost per tank type.

Tanktype	**Daily Rate**
20_FeetSpecial	€30
ISO_Non_Hr	€18,1
ISO_HR	€18,1
ISO_HR_Baff	€30
Swap	€39,75
Swap_baff	€39,75
Swap_Special	€45,90

Table 4.1: Opportunity cost per tank type

Penalty cost
Penalty costs are the costs of missing an order. These costs are not actually paid, but can be seen as the cost of missing revenue. The BU Global covers much larger distances and thus has a more significant profit margin. Including the profit margin and the contribution on overhead costs of the pricing components in Table 1.2 it is possible to easily obtain the lost cost of a direct order.

It is important to mention that penalty costs can be determined on multiple methods. One could argue that the penalty costs should be the full cost of losing an order, thus the loss of an order. However, it could also be argued that the penalty costs should be discounted for its duration as by Equation 4.1. With the following variables representing: a being the planning horizon, the $\mu_{repostion}$ mean duration of reposition in a trip, and the $\mu_{transport}$ the mean duration of transport. Since the model is capped at single period, the full duration of an order falls potentially not within that time horizon. Correcting it for the corresponding time horizon does take into account that longer time horizon and make the comparison more fair.

$$Penalty = (Loss) * \frac{\alpha}{\mu_{reposition} + \mu_{transport}} \tag{4.1}$$

Chapter 5

Model description

Section 5.1 treats the ETS forecasting model with their parameters. The SARIMA and LOESS forecast models are in Appendix C. Section 5.2 explains robust optimization and how to model it. Section 5.3 shows the robust optimization model tailored for DH. The model approximating the adaptive robust optimization by using affine decision rules is present in section 5.4. The the redistribution model of returning safety stock levels towards their initial state after the occurrence of demand is in Appendix F.

5.1 Forecast models

A forecast model aims to provide a forecast at the start of each planning horizon for each hub and every demand characteristic as accurately as possible.

Holt-Winters additive model
The additive Holt-Winters model is in Equation 5.1 (Hyndman, 2018). The $\hat{y}_{t+h|t}$ term denotes the dependent variable and thus the actual forecast value of the time series value y for the h steps ahead forecast at time moment t. The t represents the current time with its unit in days. The h represents the number of steps in units of t the forecast aims to look ahead.

The additive Holt-winters model consists of three estimation parts. It uses the level estimate ℓ_t, the trend estimate b_t, and the seasonality part s_t. The level estimate is the starting coefficient. The trend estimates the growth factor within a single time unit. The seasonality factor estimates the effect on the demand depending on the season. For some products, it is likely to experience more demand in certain months than other months. The m represents the number of periods in a year, and since seasonality is set monthly, the corresponding value will be 12. The α, β^*, γ are free parameters between 0 and 1.

$$
\begin{aligned}
\hat{y}_{t+h|t} &= \ell_t + hb_t + s_{t+h-m} \\
\ell_t &= \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1}) \\
b_t &= \beta^* t(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1} \\
s_t &= \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}
\end{aligned}
\tag{5.1}
$$

An ETS model will be optimized for these free parameters based on the Aikaike Information Criteria (AIC) value. A low AIC value is preferred, and therefore, the AIC gets minimized. The formula for the AIC value is in Equation 5.2. With the variable k representing the amount of parameters in the model and L the log-likelihood of the model in equation 5.2.

$$AIC = 2k - 2ln(L) \tag{5.2}$$

5.2 Robust optimization modelling

As later will be shown in the results found in 6.2, it is deducible that there is no direct fit with a probability distribution on the demand. Therefore, robust optimization seems to be a suitable method to deal with the demand uncertainty. Before starting with modelling DH's environment, robust optimization in general needs some explanation. First static robust optimization models are treated. Subsequently, the adaptive robust optimization models are explained. Lastly, some approximation methods in solving these models will be shown.

Static robust optimization
First, recall a general optimization problem of the form present in Equation 5.3. With $C \in \mathbb{R}^n$ being a known cost vector vector of size n. The X represents the vector of decision variables with dimension $\mathbb{R}^n$. $A \in \mathbb{R}^{m*n}$ are the coefficients interacting with the decision variables. m is the number of restrictions. Lastly, the vector $B \in \mathbb{R}^m$ consists of the values constraining the problem.

$$\begin{aligned} \text{minimize} \quad & C^T * X \\ \text{subject to} \quad & AX \leq B, \\ & X \geq 0 \end{aligned} \tag{5.3}$$

There can be uncertainty in the parameters of C, A, and B. In the case of uncertainty in A or B, a problem solution that is optimal in the original problem may become infeasible in reality. In the other case of uncertainty in C, the found optimal value can become substantially far from optimality.

To deal with that robust optimization, considers a set in which the C, A, or the B can vary. Recall Equation 5.3, but now uncertainty will be introduced in the A matrix in Equation 5.4. In order to do so with convenient matrix notation, the constraints are stated separately with index i. The formulation is per constraint for the total amount of m constraints with index i. Additionally, the random parameter z influencing the problem is of size k. All in all, resulting in the following dimensions for the parameters in Equation 5.3: $D_i \in \mathbb{R}^n$, $E_i \in \mathbb{R}^{n*k}$, and $z \in \mathbb{R}^k$. The matrix A gets thus decomposed into $D_i + E_i * z$, with the D_i being fixed and no uncertain components, the z being the uncertain parameters, and the E_i being the coefficient of using the uncertain parameters. The expansion $(D + E * z)$ can also be abbreviated with $A(z)$.

$$\begin{aligned} \text{minimize} \quad & C^T * X \\ \text{subject to} \quad & (D_i + E_i * z)^T X \leq B_i \qquad \forall z \in Z, \\ & X \geq 0 \end{aligned} \tag{5.4}$$

Lastly, the z parameters are considered in the uncertainty set, which is shown in Equation 5.5. For the sake of simplicity, the most basic box uncertainty set is used in Equation 5.5 with a lower limit f and an upper limit g.

$$\mathcal{Z} = \left\{ z \in \mathbb{R}^k : f \leq z \geq g \right\} \tag{5.5}$$

Adaptive robust optimization
In static robust optimization, all decisions must be made before the realization of the uncertain parameter z. However, if some decisions must be made before the realization of the uncertain parameter and some decisions after the realization of the uncertain parameters, static robust optimization is conservative. Equation 5.6 shows the formulation of an adaptive robust optimization. The X are the first stage decision variables, and the $Y(z)$ are the second stage decision variables after the realization of uncertain parameter z. The C^T is split into a vector for the first stage $\hat{C}$. The original cost vector C is split into a part containing the first stage decision variables $\hat{C}$ and a part containing the second stage $\tilde{C}$ variables. For the $A(z)$ the same splitting procedure has been applied. Lastly, the definition of the uncertainty set in Equation 5.5 remains valid. By the use of this formulation, it becomes possible to adapt to the uncertain parameter in different timing moments.

$$\begin{aligned} \underset{x,y}{\text{minimize}} \quad & \hat{C}^T * X + \max_{Z} \tilde{C}^T * Y(z) \\ \text{subject to} \quad & \hat{A}(z) * X + \tilde{A}(z) * Y(z) \leq B \qquad \forall z \in Z, \\ & X, Y(z) \geq 0 \end{aligned} \tag{5.6}$$

Affine decision rules
Since adaptive robust optimization can be computationally challenging, approximation procedures are used in practice. One of these approximation methods is the use of affine decision rules. In Equation 5.7 the affine decision rules are stated.

$$Y(z) = Vz + u \tag{5.7}$$

In this case, the uncertain parameter z is of the dimension $\mathbb{R}^k$, the V of the $\mathbb{R}^{n*k}$, and the u of $\mathbb{R}^k$. The adaptive variable $Y(z)$ gets thus approximated by affine decision rules making thus the problem static again, which makes it less computationally challenging.

Solving adaptive robust optimization
A more exact approach of solving adaptive robust optimization is an hybrid approximation method proposed in Bertsimas & De Ruiter (2016). The primal and dual have the same optimal affine policies, but the number of affine constraints may be some magnitude smaller in duality. By using both the primal and dual problem, the algorithm described in Bertsimas & De Ruiter (2016) can provide more accurate and faster lower and upperbounds and solve the adaptive robust optimization with an iterative procedure faster and with only a small optimality gap in respect of the exact solution of robust adaptive optimization. This procedure start with constructing an upperbound with the use of affine decision rules and use the binding scenarios of that particular solution to construct a lowerbound (Bertsimas & De Ruiter, 2016). Since this is thus a mixture of approximation methods, it is reffereed to as Hybrid method in this thesis. It is stated, that it is possible to use this algorithm with the first stage decision variables to be an integer, but not with the second stage decision variables to be an integer (Bertsimas & De Ruiter, 2016). This is due to dualization over the second stage variables making it not possible to use integrality restrictions on the second stage.

5.3 Single period with multi-commodity

First, this section starts with introducing the notation. Secondly, the cost functions which are in the objective are explained. Thirdly, the constraints are formulated. Fourthly, the construction of the uncertainty set is discussed. Lastly, everything is summarized in the main problem formulation.

Problem description
This section aims to provide an understanding of the translation of the problem environment of DH to a modelling environment. An overview of the definitions of all the variables is given in Table 5.1.

First of all, a set containing all the hubs in the network is created with the set $\mathcal{I}$. Since the tank containers flow from one hub towards another hub, it is necessary to have a second set $\mathcal{J}$ containing again all the hubs in the network. The set $\mathcal{C}$ contains all the different combinations of demand characteristics listed in Table **??**. The last set consists of all the different tank types $\mathcal{P}$ considered in this thesis, which are listed in Table 1.1.

There exist four cost coefficient input parameters. The: C_i^h is the price paid at location i to the depot a tank-container for a days in euros. The a represents the number of days DH uses to plan in the rolling time horizon planning. The $C_{i,j}^{TR}$ is the price of transporting an empty tank container from hub i towards hub j. The C_p^O is the price it costs DH to have a tank container of tank type p for a days, regardless it usage of where it is in the network. The last component is the C^P and this is the price it costs DH for losing an order which is shown in equation 4.1.

Regarding the forecast input parameters, there are two main parameters: the $\mu_{i,c}$ which is the expected mean at hub i of demand characteristic c in the corresponding period of the planning horizon a, and the forecast error $\sigma_{i,c}$. The Σ_c is the forecast error for all hubs per demand characteristic c and thus one aggregation level higher. The Ψ is the forecast error of all hubs and all demand characteristics.

There are also input variables which ensure feasibility. $F_{i,j}^{TR}$ ensures feasibility that inventory coming from neighboring hub for demand fulfillment is on time. For example, if hub j wants inventory of hub i, it checks the transportation time $T_{i,j}$ and if it is below the planning horizon a it is feasible. A feasible route gets assigned a one and an unfeasible route gets a 0. The $F_{c,p}$ check feasibility op tank type p for demand characteristic c based on the upcoming results found in Table **??**.

By efficient usage of the found forecast error on different forecast aggregation levels, the size of the uncertainty set will be reduced. The parameters necessary for doing so are: the γ restricts the amount of acceptable standard error for each hub i and each demand characteristic c, the Γ restrict the forecast for the total demand for each demand characteristic c, and the Δ restrict the forecast of all demand.

Lastly, the corresponding decision variables. The $S_{i,p}$ is the first stage decision variable and represents the safety stock level at hub i of tank type p. The $S_{i,p}$ is the main variable and the key variable DH is looking for. The $Q_{i,j,c,p}$ states how much inventory of hub i is used in the demand of hub j for the demand characteristic c and this fulfillment of the actual demand $D_{i,c}$ is done with tank type p.

Notation	Description
Set	
$\mathcal{I}$	the set of hubs, indexed by i.
$\mathcal{J}$	the set of hubs, indexed by j.

Table 5.1 continued from previous page

Notation	Description
$\mathcal{C}$	the set of different demand characteristic, indexed by c.
$\mathcal{P}$	the set of different tank container types, indexed by p.
Input variables	
a	The planning horizon in days at the MMP department.
β	The fill rate which is allowed under the considered size of the uncertainty set.
γ	Amount of individual hub forecast error allowed per demand characteristic.
$F^{TR}_{i,j}$	States if a route is feasible with a 1 and 0 for infeasible for a pair of hub i and j.
$\mu_{i,c}$	Expected demand at hub i of characteristic c.
$\sigma_{i,c}$	Standard error at hub of demand at hub i of characteristic c
Σ_c	Standard error of all the hubs per demand characteristic c
Ψ	Standard error of all hubs of all demand characteristics
Γ_c	Total standard error of demand characteristic c
Δ	Total standard error of demand
$T_{i,j}$	Transportation time between hub i and hub j
$F_{c,p}$	Feasibility matrix of demand characteristic c with tank type p
$C^{TR}_{i,j}$	Cost of transport between hub i and hub j
C^{H}_{i}	Cost holding a tank container at hub i for time period α
C^{O}_{p}	Cost of having a tank container of type p during time period α
C^{P}	Cost of missing an order
1 stage decision variables	
$S_{i,p}$	The amount of stock present at hub i of tank type p
Uncertain parameter	
$D_{i,c}$	The actual demand at hub i of demand characteristic c
2 stage decision variables	
$Q_{i,j,c,p}$	The fulfillment of demand at hub j by hub i of characteristic c with tank type p

Table 5.1: Definition variables in robust optimization

Assumptions

Some assumptions are necessary to model robust optimization. The main assumption in this model is the usage of a single period model. With not having a time index for multiple periods, it is impossible to have intermediate inflows and outflow of inventory in one planning horizon of duration a. Additionally, a tank container can only get assigned once, during the planning horizon of length a. Another assumption is bundling all the demands of BU Global in one hub. However, the impact is low, since stock levels out of Europe are not considered. This fictional, Global hub is necessary to correct for the import and export balance in the network of Europe. The last assumption is having identical penalty costs for each order. In reality, the margins on lanes can differ and thus require more or less conservative stock levels depending on the penalty costs.

Cost functions

As explained in Chapter 4 the related costs occur in the network of DH and thus resulting in an objective function with the following components.

The holding costs are formulated in Equation 5.8. It contains all the stock present at the start of a time period in the first stage decision of $S_{i,p}$ and everything that is used in the second stage is subtracted with the variable $Q_{i,j,c,p}$. The variables $F^{TR}_{i,j}$ and $F_{c,p}$ ensures that the inventory used is feasible for the demand characteristics and is feasible on time. The plus function ensures that the holding costs can not become negative.

$$\text{Holding cost} = \sum_{i\in\mathcal{I}} \left(\sum_{p\in\mathcal{P}} \left(S_{i,p} - \sum_{j\in\mathcal{J}} \sum_{c\in\mathcal{C}} Q_{i,j,c,p}(D_{i,c}) * F_{c,p} * F^{TR}_{i,j} \right)^{+} * C^{H}_{i} \right) \quad 5.8$$

The transshipment costs are calculated in Equation 5.9. All the transshipment of the second stage decision variable $Q_{i,j,c,p}$ are multiplied with the corresponding cost coefficients $C^{T}_{i,j}$ and summed over all indices for obtaining the total transportation costs.

$$\text{Transship cost} = \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \sum_{c\in\mathcal{C}} \sum_{p\in\mathcal{P}} \left(Q_{i,j,c,p}(D_{i,c}) * F_{c,p} * F^{TR}_{i,j} * C^{T}_{i,j} \right) \quad 5.9$$

The penalty costs are calculated in Equation 5.10. The actual demand gets subtracted by the amount of demand DH is able to fulfill in the second stage with variable $Q_{i,j,c,p}$. Again the plus function ensures, that penalty costs can not become non negative.

$$\text{Penalty cost} = \left(\sum_{i\in\mathcal{I}} \sum_{c\in\mathcal{C}} \left(D_{i,c} - \sum^{P} \sum^{J} Q_{j,i,c,p}(D_{i,c}) * F^{TR}_{i,j} * F_{c,p} \right)^{+} \right) * C^{P} \quad 5.10$$

The opportunity costs are calculated in Equation 5.11. The opportunity costs simply consisted of having the tank containers in the system and is thus equal to summing all the first stage decision variables of the safety stock levels and multiplying it with the corresponding cost of that type tank container during the planning horizon.

$$\text{Opportunity cost} = \sum_{p\in\mathcal{P}} \left(\sum_{i\in\mathcal{I}} S_{i,p} \right) * C^{O}_{p} \quad 5.11$$

Explanation constraints

Constraint 5.12 ensures that at least a fill rate of β is achieved in the allowance range of the demand in the uncertainty set Z. It is doing so by comparing the amount of transshipments and the actual demand realized.

$$\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \sum_{c\in\mathcal{C}} \sum_{P\in\mathcal{P}} Q_{i,j,c,p}(D_{i,c}) * F_{c,p} * F^{TR}_{i,j} \geq \beta * \sum_{i\in\mathcal{I}} \sum_{c\in\mathcal{C}} D_{i,c} \quad 5.12$$

Constraint 5.13 makes it impossible to sell more, than there is actual demand present for each hub i of demand characteristic c.

$$\sum_{p\in\mathcal{P}} \sum_{j\in\mathcal{J}} Q_{j,i,c,p}(D_{i,c}) * F_{c,p} * F^{TR}_{i,j} \leq D_{i,c} \quad 5.13$$

Constraint 5.14 ensures using more stock than there is initially allocated at each hub i for each tank type p is impossible.

$$\sum_{c\in\mathcal{C}} \sum_{j\in\mathcal{J}} Q_{i,j,c,p}(D_{i,c}) * F_{c,p} * F^{TR}_{i,j} \leq S_{i,p} \quad 5.14$$

Constraints 5.15 makes the first stage decision variable $S_{i,p}$ integers and positively valued. Furthermore, the second stage decision variables $Q_{i,j,c,p}$ are also integers and positively valued. These values cannot be fractional or negative because partial tank containers or a amount of negative tank containers do not exist.

$$S_{i,p}, Q_{i,j,c,p} \in \mathbb{Z}^{+} \tag{5.15}$$

Defining the uncertainty set

The uncertainty set consists of five restrictions for reducing the size. Restrictions 5.16 and 5.17 are the box uncertainty constraint limiting the demand within γ deviations of its mean in the forecast. Restrictions 5.19 and 5.20 are the budget uncertainty constraints limiting the total deviation of the demand. The last restriction 5.17 is necessary to avoid the situation of negative demand. This is a necessity since DH has some low demand hubs with a high standard deviation, which otherwise would result in scenarios with a negative demand.

$$\begin{aligned}
\mathcal{W} = \Bigg\{ D \in \mathbb{R}^{|\mathcal{I}|*|\mathcal{C}|} : \; & D_{i,c} \leq \mu_{i,c} + \gamma * \sigma_{i,c} \;, && 5.16 \\
& D_{i,c} \geq \mu_{i,c} - \gamma * \sigma_{i,c} \;, && 5.17 \\
& D_{i,c} \geq 0 \;\;, && 5.18 \\
& \sum_{i \in \mathcal{I}} D_{i,c} \leq \Gamma * \Sigma_c \;, && 5.19 \\
& \sum_{i \in \mathcal{I}} \sum_{c \in \mathcal{C}} D_{i,c} \leq \Delta * \Psi, \; i \in \mathcal{I}, \; c \in \mathcal{C} \Bigg\} && 5.20
\end{aligned}$$

The full model formulation

The complete model formulation summarizing all the previous cost functions, constraints, and the uncertainty set of the adaptive robust optimization is in Model 5.21.

Removal plus functions

The holding cost and the penalty cost are dealing with the plus function to prevent having negative costs. Whereas in the objective function, these plus functions are gone. The constraint 5.13 and 5.14 make these plus functions redundant. Since, the $\sum_{p \in \mathcal{P}} \sum_{j \in \mathcal{J}} Q_{i,j,c,p}$ matrix is always smaller or equal than the $D_{i,c}$, the equation of the penalty costs function can never attain negative values. The same reasoning holds for the equation of the holding costs with the other constraint.

Integer demand

The demand, however, is still able to contain some numbers which are not integers. Nonetheless, since $\mu_{i,c}$, γ, and $\sigma_{i,c}$ are all integers, the corner points of the uncertainty set of the $D_{i,c}$ value will likely to be integers. The budget constraint may cause some values to be not integrals at the corner point, but this will likely not consider many points.

$$
\begin{aligned}
\underset{S_{i,p},\,Q_{i,j,c,p}}{\text{minimize}} \quad & \max_{\mathcal{W}} \sum_{i\in\mathcal{I}} \left(\sum_{p\in\mathcal{P}} \left(S_{i,p} - \sum_{j\in\mathcal{J}} \sum_{c\in\mathcal{C}} Q_{i,j,c,p}(D_{i,c}) * F_{c,p} * F^{TR}_{i,j} \right) * C^H_i \right) \\
& + \left(\sum_{i\in\mathcal{I}} \sum_{c\in\mathcal{C}} \left(D_{i,c} - \sum_{p\in\mathcal{P}} \sum^{J} Q_{j,i,c,p}(D_{i,c}) * F^{TR}_{i,j} * F_{c,p} \right) \right) * C^P \\
& + \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \sum_{c\in\mathcal{C}} \sum_{p\in\mathcal{P}} \left(Q_{i,j,c,p}(D_{i,c}) * F_{c,p} * F^{TR}_{i,j} * C^T_{i,j} \right) \\
& + \sum_{p\in\mathcal{P}} \left(\sum_{i\in\mathcal{I}} S_{i,p} \right) * C^O_p \\
\text{subject to} \quad & \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \sum_{c\in\mathcal{C}} \sum_{p\in\mathcal{P}} Q_{i,j,c,p}(D_{i,c}) \quad , \\
& *F_{c,p} * F^{TR}_{i,j} \geq \beta * \sum_{i\in\mathcal{I}} \sum_{c\in\mathcal{C}} D_{i,c} && \forall\, D \in \mathcal{W}, \\
& \sum_{p\in\mathcal{P}} \sum_{j\in\mathcal{J}} Q_{i,j,c,p}(D_{i,c}) * F_{c,p} * F^{TR}_{i,j} \leq D_{i,c} && \forall\, D \in \mathcal{W},\ i \in \mathcal{I},\ c \in \mathcal{C}, \\
& \sum_{c\in\mathcal{C}} \sum_{j\in\mathcal{J}} Q_{j,i,c,p}(D_{i,c}) * F^{TR}_{i,j} \leq S_{i,p} && \forall\, D \in \mathcal{W},\ i \in \mathcal{I},\ p \in \mathcal{P}, \\
& S_{i,p} \in \mathbb{Z}^+ && \forall\, D \in \mathcal{W},\ i \in \mathcal{I},\ p \in \mathcal{P}, \\
& Q_{i,j,c,p}(D_{i,c}) \in \mathbb{Z}^+ && \forall\, D \in \mathcal{W},\ i \in \mathcal{I},\ j \in \mathcal{J}, \\
& && c \in \mathcal{C},\ p \in \mathcal{P}, \\
\mathcal{W} = \Bigg\{ & D \in \mathbb{R}^{|\mathcal{I}|*|\mathcal{C}|: D_{i,c}} \leq \mu_{i,c} + \gamma * \sigma_{i,c}, \\
& D_{i,c} \geq \mu_{i,c} - \gamma * \sigma_{i,c}, \\
& D_{i,c} \geq 0, \\
& \sum_{i\in\mathcal{I}} D_{i,c} \leq \Gamma * \Sigma_c, \\
& \sum_{i\in\mathcal{I}} \sum_{c\in\mathcal{C}} D_{i,c} \leq \Delta * \Psi && \forall\, i \in \mathcal{I},\ c \in \mathcal{C} \Bigg\}
\end{aligned}
\tag{5.21}
$$

5.4 Approximation affine decision rules

Modelling challenges

The model formulated in model 5.7 is now an adaptive robust optimization. These models can be hard to solve directly for large instances. Therefore, these models can be approximated by affine decision rules as stated in equation 5.7. By approximating this with affine decision rules, it becomes a static robust optimization that is computationally easier to solve. Model 5.21 can thus be changed into the model presented in Appendix A.1 to speed up computations.

Chapter 6

Results

In this chapter the results for conceptual research question and the models are shown. Section 6.1 provides the explanation of the current state of planning at DH based on interviews with the responsible planners. Section 6.2 analyzes the demand behavior and its performance in the different forecasting models. Section 6.3 discusses the performance of the robust models.

6.1 Result planning procedures

The general planning process
The planning process of an order starts with the arrival of the order. The arrival of an order differs in the number of days before departing. Typically, orders arrive around three days before loading, and as mentioned, the number of days between an order's loading date and its arrival is called *decision days*. Challenging trips tend to arrive earlier, and there are customers who prefer ordering as late as possible. This causes differences in the amount of decision days for each order. After an account manager has agreed upon a price, the order goes to the MMP department. The MMP department determines which tank container gets assigned to the order and a suggested routing for the tank container. The software system PlannIT proposes an initial schedule for tank container planning, taking all the current orders into account. However, the system is outdated, and the MMP department has to change manually the proposed schedule. After determination of its tank container and route, the order goes to TCP. The TCP department determines the assignment of internal or external trucks to orders. During delivery, transport may become delayed, and management of this process is the responsibility of the MMP and TCP departments.

Time horizon
Each MMP planner has the freedom of its sub-region, and thus each planner does deviate in its planning process to an extent. A general rule for MMP planners to start planning 2 days ahead of its loading date for assigning tank containers. The number of days a planner looks a head, is called the planning horizon. Generally, regions with a high order density usually use a smaller planning horizon because the network changes a lot, and the planners can adapt faster. The opposite is also true, regions with a low order density general use a larger time window .

Verification of planning horizon
Extension of the planning horizon is linked with the corresponding decision days. A MMP planner can not start planning, if the orders of customers are not present. From

analysis. it is deducible by summing the probabilities after 6 decision days, that 89% of the total demand is known when the MMP planners start planning under the current planning horizon policy. The reason for planning six days ahead is to reduce re-planning. Expanding the planning horizon significantly boost the unknown demand and thus its re-planning.

Tank allocation procedure
Although, it is previously stated that a planning horizon of six days is being used by the MMP planners. There are some exceptions for possible difficult to plan orders. Unfortunately, it is not possible to display the table with tank planning procedures due to confidentiality.

Strategic initiatives
For better management of the flow of tank containers strategic initiatives can enhance performance in the supply chain (Nahmias & Olson, 2015). On the problem dimensions demand uncertainty, network, multi-commodity and timing dimension, it is possible to use strategic initiatives. This section discusses how some possible strategic initiatives with cooperation in the supply chain can improve stock level management.

For reducing the amount of multi-commodity, the number of categories must be reduced. If there are fewer demand types, it is possible to do the same number of orders with a smaller inventory. One method is standardization in the industry. For example, standardizing walkways and handrails make it possible to reduce fleet combinations and thus make it easier for tank container operators. Another option is to make tank container types more flexible. For example, portable handrails exist, and such solutions make it possible to use more tank containers of the fleet for required handrail demand. In the future, other innovative methods for other demand characteristics could also be explored by DH to decrease the amount of multi-commodity. The last option is to invest in tank containers capable of as many demand types as possible. Unfortunately, this is also more expensive and whether this is beneficial depends on the pricing parameters.

Methods for tackling the demand uncertainty are related with cooperation of customers. Understanding the production schedule of the most influential customers can help to timely position tank containers for your large customers. Recall, that this is currently already happening for a few customers, but this could be extended. This inclusion helps to reduce the variability in the number of tank containers, which is not known. Furthermore, investing in advanced forecasting methods can help reduce the forecast error and thus the standard deviation in the demand process, resulting in a reduction in safety stock level.

For the time dimension, it is valuable to gain more response time by having more decision days. If there are more decision days, it is possible to use a longer time horizon without additional re-planning. Having a large planning horizon would enlarge the possibilities of on time transshipments and thus reduce safety stock. DH could also introduce some methods on pricing and reward customers communicating their orders timely to shift the planning horizon further ahead.

For strategic initiatives on the network, it can help to identify strong and weak lanes for DH. Identifying vulnerable lanes can support decision-making in rejecting orders in that lane and ensure profitability. If this is not possible, it can be possible to resell the order or search for cooperation with other tank container operators to potentially pool inventory on weak lanes.

SLM issues
During interviews with MMP department employees, it became apparent that the current

version of the SLM needs improvement on several issues:

- There is a difference in the actual stock level and the stock level in the SLM. For example, some tank containers which are present at a hub are not shown and vice versa.
- Tank containers that are soon becoming available after maintenance and re-entering the system are not considered in the SLM.
- Discrepancies in the depicted stock level and the actual stock level may cause decision-making based on wrong information and thus less optimal decision making by MMP planners.
- The SLM is showing inventory levels instead of the inventory position. For example, if many departures of tank containers are already scheduled at a hub, the inventory level can be misleading. The inventory level would indicate that there is more stock than it is available for usage, which may cause less optimal decision-making. The opposite is true that having more arrivals would result in more inventory than the inventory level.
- It is not possible in the SLM to vary the timing towards possible future inventory positions. If this were implemented, future shortages and surpluses could be foreseen, and earlier decision-making in repositioning would increase options.
- Predictions of demand for the upcoming days are not given. Most planners are familiar with a region, observe the patterns in demand, and can thus forecast themselves. However, a forecast could provide extra guidance for planners to verify if their decision-making is indeed valid.
- Moreover, most of the missing parts indicated in the SLM are present in the system plannIT of DH. For example, in PlannIT, it is possible to view each tank container's movement individually and verify inflow, outflow, and hubs. However, it can be time-consuming to manually check this information. Therefore, fully integrating it in the SLM could enhance decision-making and reduce the time for the MMP planners.

6.2 Result forecast

The results have been obtained by programming in R with the use of the package FPP2, which is a R package provided by Hyndman (2018) and effectively tracks forecasting performances.

6.2.1 General time series

First, the monthly total demand in the network of Europe of all the hubs of DH is analyzed and shown in Figure 6.1. For the full time span that is available at DH from 2017 March till 2021 April. The aggregated total demand is visualized since the aim is to discover general patterns in the time series of the demand. The aggregation level of the timing of the time series is monthly to more easily distinguish the effects of trends and seasonality.

Figure 6.1 consist of three parts. The top part is the general time series over time. The left bottom part is the autocorrelation function (ACF). The right bottom part is the partial autocorrelation function (PACF). The ACF represents its correlation of the actual value with previous values (Hyndman, 2018). The lag states the difference in the number of time periods between the actual value and the considered lagged value. However, only looking at the ACF is insufficient. Since the correlation effect of lag 2 is also present at lag 1 and so on. Therefore, the partial autocorrelation function exists, which only measures the correlation of a lag due to that particular lag Hyndman (2018).

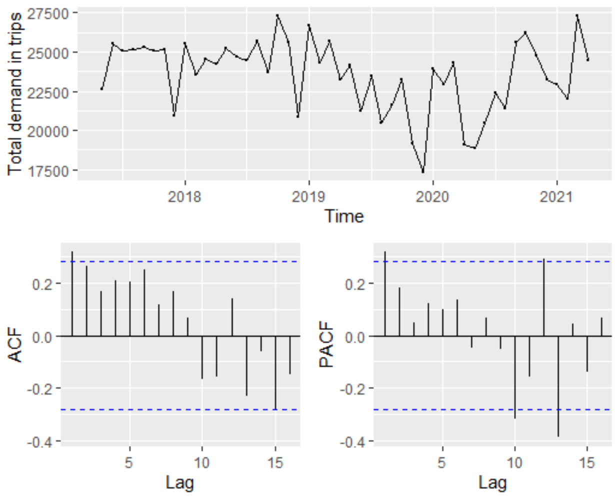

Figure 6.1: Time series of total demand

From the upper part containing the time series, a few observations can be made. Firstly it becomes apparent that the time series does have a sample mean of approximately 23.559 trips a month with a standard deviation of 2.271 trips. Furthermore, the time series seem to behave fairly stationary. There seem to be visually no explicit trend patterns and no explicit seasonality patterns. Regarding possible outliers in the time series, the corona crisis is visible by the drop at the end of 2019. Again this drop seems to be not very substantial and therefore these possible outliers are not deleted.

From the ACF in Figure 6.1 it is apparent that the first lag is above the blue line, which is the 95% significance level. After the first lag, it is slightly decreasing. A decreasing or increasing pattern in the ACF, indicates the presence of a trend (Hyndman, 2018). Thus the ACF may indicate that there is indeed a trend present. For the PACF a significant lag is found for the 1st, 10th, 12th and 13th lag. The first lag significance of the PACF supports again the presence of a trend. The 12th and 13th significance support the presence of seasonality (Hyndman, 2018). The 10th lag significance is not directly explainable.

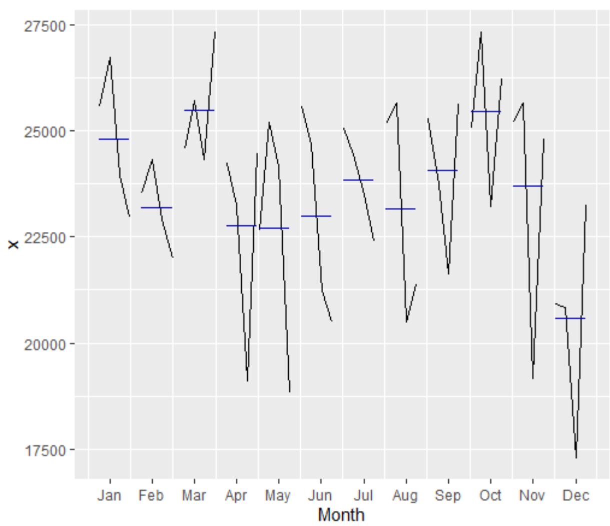

Figure 6.2: Seasonality plot total demand

Because of the found significance in the PACF of seasonality, a seasonality plot is made in Figure 6.2. This Figure plots on the x-axis the months and the value over the years on the y axis with the blue line representing its mean. From the plot, it is visible that, March and October seem to be high volume months, whereas December appears to be a low volume month. However, if looked closely at the y axis, the corresponding jumps are fairly low in respect of its total value. This supports the thoughts of having significant seasonality, but its impact on the demand is fairly low.

Overall, it can be concluded that visually the time series seems to behave fairly stationary. The ACF and PACF hint towards significant effects of a trend and seasonality and thus is present in the time series. These effects are thus significant, but its impact on the demand itself is relatively small.

6.2.2 Flow of demand on hub level

The mean inflow and outflow of hubs
The mean inflow, outflow, and the difference in the flow of the total amount of tank containers is in Appendix B.1 and provides more detailled information regarding the mean and standard deviation of each hub. The three largest hubs, hub 1, hub 12, and hub 39, together have an outflow of 49% and an inflow of 34% of the entire network. The medium to large-sized hubs are hub 22, hub 23, hub 28, hub 11, hub 40, hub 33, hub 24, hub 44, hub 48, which account for 31 % outflow, 32% inflow. So a large part of the demand is covered by only a few hubs. Furthermore, the coefficient of variation in these large hubs is much lower than the smaller ones, implying that the larger hubs have a more stable demand.

The mean difference flow
The inflow and outflow of hubs can differ highly and therefore cause an unbalanced network. From Appendix B.1, it is possible to analyze the mean difference flow and its impact on the network. Unfortunately, it is not possible to discuss the differences and connectability with all the hubs, because of confidentiality reasons.

6.2.3 Testing distribution of forecast error

The results of the complete ETS forecast on Hub level and per demand characteristic are in Appendix D. In this paragraph, the forecast error is used for testing for possible demand distributions. Why are the corresponding error terms used and not the actual demand values? This is done, because the forecast on an individual hub level and per demand characteristic may contain remaining effects of seasonality or trending patterns and thus making the series non-stationary over time Hyndman (2018). For fitting a demand distribution, the corresponding time series must be stationary. Additionally, the forecast errors are checked if the hubs do correlate with each other. There again the forecast is used, to remove patterns of trend and seasonality. In case the demand values would be compared a correlation would be found, what is actually caused by the general trend or seasonality. After the removal of these effects, it is the aim to have an error term series that can be considered as white noise. If it is white noise, the forecast is perfect and no additional information that could potentially improve the forecast is present in the residuals Hyndman (2018).

Normality error term
The Shapiro-Wilk test is used to verify normality Shapiro & Wilk (1963) for the forecast errors. The Shapiro-Wilk test has the H0 hypothesis of the residuals being normally distributed. Of the 48 Hubs, 18 could be immediately rejected under 5% significance of having normal error terms. Note that the other 30 hubs are not automatically normally

distributed, but the test is not powerful enough to reject it, and further analysis would be required to verify normality.

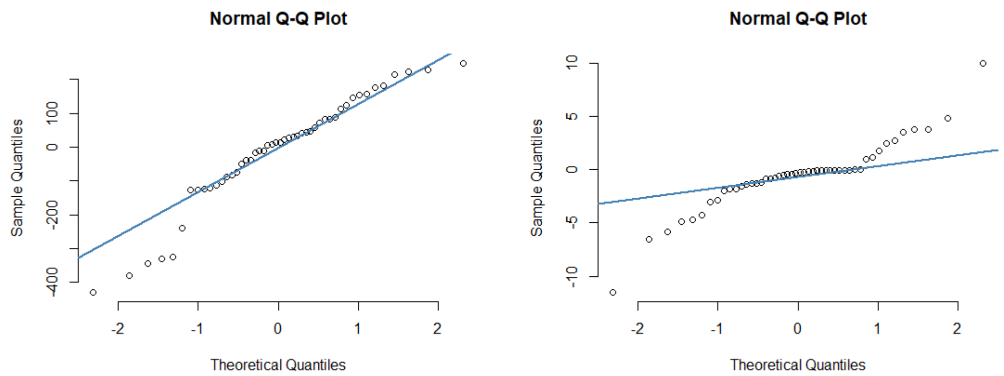

Figure 6.3: Left QQ Plot antwerp, Right QQ plot Mediterranean

The QQ plot, Figure 6.3, demonstrates the difference in the normality of error terms in Hubs. Hub 1 nearly follows the normal distribution, and only its tails may be slightly off. Whereas the hub Mediterranean is not able to fit on the quantile line and is far off. The large hubs generally follow the QQ line for normal error terms more closely, whereas the smaller hubs do not. The inability of the small hubs to follow the normal distribution is due to the lumpy demand behavior of the small hubs, resulting in only a few contracts that are possible to generate demand. The larger hubs have more contracts, and thus, convolutions of these non-identical individual demand distributions will more likely converge to a normal distribution as the number of contracts increases. Nonetheless, it is concluded that assuming normality for all hubs is not representative. Therefore, the use of stochastic optimization becomes obsolete. As mentioned, robust optimization can relax this and only work with the corresponding standard deviations for hedging against uncertainty.

Correlation error term between hubs

The residuals of the forecast are tested for autocorrelation between the hubs. This is tested to verify if hubs correlate in demand behavior and thus if the uncertainty sets should consider this interactive behavior. From Table 6.1 it is deducible that 72 % of the correlation between hubs is within a reasonable range of -0.2 till 0.2. Indicating that most hubs do not correlate that much. Another interesting observation is that most hubs correlate in a positive direction. This correlation is possible because the trend and seasonality patterns are not grasped completely by the forecast, and these cause positive correlation coefficients. For example, suppose the forecast model does not completely grasp the seasonality. In that case, it is more likely that if demand in Paris rises due to seasonality, the demand also rises in Prague. In that particular case, the interaction of demand is not influencing its correlation, but the overall seasonality is. However, the correlation coefficients seem reasonable to not assume multivariate behavior of the forecast error.

Correlation	Percentage
-0,99 till -0,3	1%
-0,3 till 0,2	3%
-0,2 till -0,1	9%
-0,1 till 0	20%
0 till 0,10	24%
0,1 till 0,2	19%
0,2 till 0,3	11%
0,3 till 0,4	5%
0,4 till 0,5	5%
0,5 till 0,99	3%

Table 6.1: Correlation coefficients

Benchmark SLM and forecast
It is possible to construct a safety stock level estimate per hub and per tank type only using the forecast mean and γ times its forecast error. This safety stock level estimate is very conservative since it is on individual level and uses no cooperation in the form sharing inventory. The Robust model will always produce a smaller stock level under the same values for the standard deviation, mean, and planning horizon, since it is able to cooperate between hubs and use multi-commodity. By using the forecast mean and γ times its forecast error as a upper limit, it quantifies the gain of using robust optimization.

6.3 Result robust optimization

The results have been obtained by the use of the Gurobi package in the programming language Python with an additional package RSOME available at Zhi et al. (2019). The package RSOME is specially designed for robust optimization problems.

6.3.1 The different robust optimization models

In Table 6.2 the corresponding model variants of robust optimization are listed. All models use an approximation method to solve the adaptive optimization model. The hybrid method is mentioned in Bertsimas & De Ruiter (2016) and explained in Section 5.2. The other method is by only using the affine decision rules, making the problem static again. Since integrality is costly in computation time and unobtainable for the hybrid method for the second stage variables, it is decided to use continuous relaxation. Continuous relaxation is approximating an integrality constraint by the allowance of continuous variables. By creating the different model variants the effect of hybrid method and integrality in the first and second stage can be observed.

Name	Approximation method	First stage: Continuous relaxation	Second stage Continious relaxation
Model 1	Hybrid method	yes	yes
Model 2	Hybrid method	no	yes
Model 3	Affine decision rules	yes	yes
Model 4	Affine decision rules	no	yes
Model 5	Affine decision rules	no	no

Table 6.2: Different robust optimization models

6.3.2 Computation time

Augmented problem size

Recall the full problem size; there are 48 hubs (for each index i and each index j), ten demand characteristics (with index c), seven tank types (with index p). Resulting in a first stage safety stock level matrix of $S_{i,p} = 48 * 7 = 336$ variables, and a second stage decision transportation of matrix Dimension $Q_{i,j,c,p} = 48*48*10*7 = 161.280$ variables, which have to be calculated. Ideally, there would also be an incorporation of the timing components, but this would logically increase the number of variables again massively.

Set up model

A simulation with pseudo data is used to test which problem size is executable. The computation time is measured on a laptop with a 3.6 GHz processor and 8 GB of ram. This test is run first on all the five model variants. Furthermore, for each optimization run, three separate tests will be used, and the average time of these three will be the computation time for each setting. Since all the values are randomly sampled, each instance will differ in computation time, and it will be more representative by taking the average.

Computation time of models with no multi-commodity

In this paragraph, it is studied to what extent the problem size must be limited to have still computationally tractable optimization problems. Since many instances have to be run, the model gets timed out after 6 minutes to avoid excessive running times. The starting point is the model with one tank type and one demand characteristic and thus without multi-commodity. In Figure 6.4 the computation time is plotted. It is deducible for model 1, that it is possible to evaluate the full network in time without using multi-commodity. With model 2, around half of the network is analyzable. Including integers in the first stage in the hybrid method is thus already much more restrictive. With model 5, it is possible to evaluate the full network and even faster in comparison to model 1. Model 3 and Model 4 are not considered since Model 5, which is more restrictive already is able to reach the full network at ease.

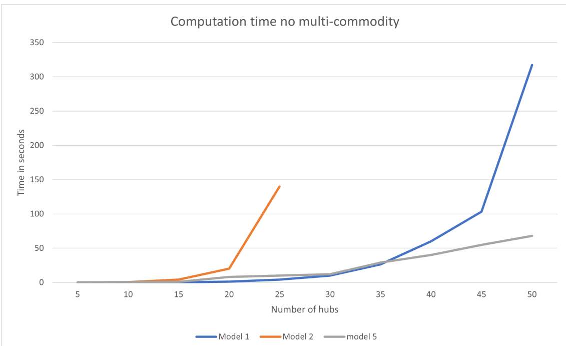

Figure 6.4: Computation time no multi-commodity against amount of hubs

Overall it can be concluded, that the full network is analyzable with ease for all model except model 2. Therefore, it is decided to try to incorporate also the multi-commodity structure.

Computation time models with multi-commodity

By including the multi-commodity which means considering ten of the demand characteristics and seven tank types, the number of hubs where the model can be solved reduces significantly, as shown in Figure 6.5. For better graphical representation, the hybrid approximation models and the affine approximation models are split.

In Figure 6.5 model 1 and model 2 are shown. Model 1 is able to deal with 6 hubs and model 2 is able to deal with 4 hubs within the time limit of 6 minutes. The same pattern is again visible; forcing first-stage variables to be integers is costly in computation time. Unfortunately, both models are only able to do a small part of the full network by including multi-commodity. It is visible that at the end of the graphic a steep jump is present and proceeding with more hubs results in massive increases in computation time.

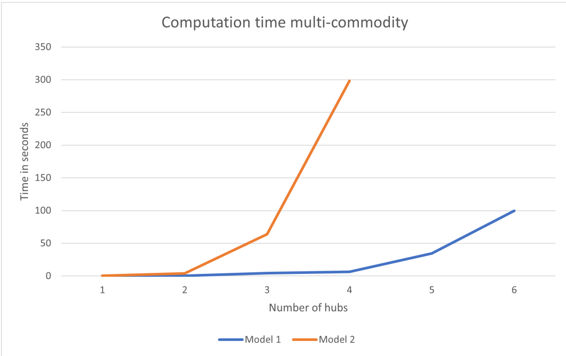

Figure 6.5: Computation time with multi-commodity against amount of hubs

Interestingly, the amount of decision variables for model 1 in no multi-commodity for the full network is $50 * 50 = 2500$. For model 1 with multi-commodity and six hubs, the amount of element is $6 * 6 * 10 * 7 = 2520$. Both situations with more or less the same size of decision variables result in more or less the same amount of computation time. Adding the number of variables is what boosts computation time and not necessarily the number of dimensions.

In Figure 6.6 the problem with multi-commodity is shown for models 3, 4, and 5. Model 3 can handle almost double the number of hubs in respect of model 1.

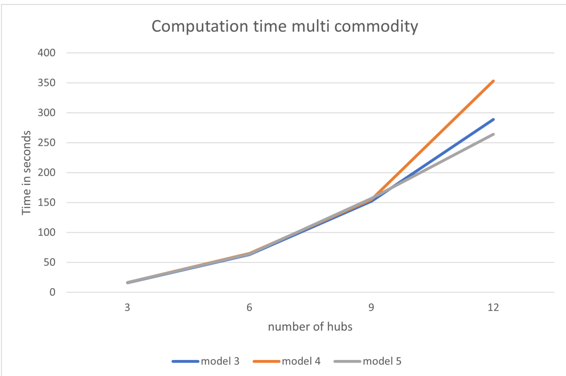

Figure 6.6: Computation time with multi-commodity against amount of hubs

Since both contain full continuous relaxation, the affine decision rule approximation is significantly faster in respect of the hybrid method. Interestingly, including integrality on the first or the second stage does not significantly impact the computation time. This insignificant effect is due since most of the time is devoted to RSOME reformulating the robust counterpart and dealing with the corresponding for loops for the formulation of the problem. This formulation and reformulation is not necessarily taking longer with having integrality restriction. The robust counterpart gets solved by Gurobi, and their integrality does impact the computation time. However, since the time for Gurobi to solve that particular MILP is negligible, this method seems not to differ in computation time for this problem by using static robust optimization.

Conclusions computation time

Model 1 is able to handle the full network and with multi-commodity 6 hubs. Model 2 is able to handle half of the network without multi-commodity and only 4 hubs with multi-commodity. For models 3,4,5 all are able to do the full network without multi-commodity and all models are able to do around 12 hubs for the multi-commodity case. Overall, it can be concluded that computation time is currently limiting analysis of the full network of DH. Since all model show at least polynomial increase in time, it is unlikely solvable since it contains around 161.280 variables whereas models with 2500 variables already seem to have difficulties to be solvable within 6 minutes.

Solutions reducing computation time

Scientific literature provides general reformulation tricks using duality, problem-specific reformulations, or advanced construction methods for reducing the uncertainty set to reduce computation time. Both seem to be out of the scope of this master thesis. Additionally, the gap between augmented problem size and the current problem size is so large, that these techniques are also not able to deal with the full network. Therefore, it is decided to cut the network into parts and solve these parts individually. This will result in sub-optimality.

In Figure 6.7 the division of regions is shown. From Figure 6.5, it was illustrated that under ten demand characteristics, seven tank types it is possible to solve the network for six hubs for model 1 within the limit of 360 seconds. Therefore, each region may at its maximum contain six hubs. Since this division is already containing 11 regions, the total computation of full network analysis is already extensive. Furthermore, the borders of each region are preferred to have geographical borders. Since water is difficult to cross due to more frequent ISO requirements on ferries, slow traveling speed, and generally less frequent operating container lanes, the regional divisions have been done as follows in Figure 6.7.

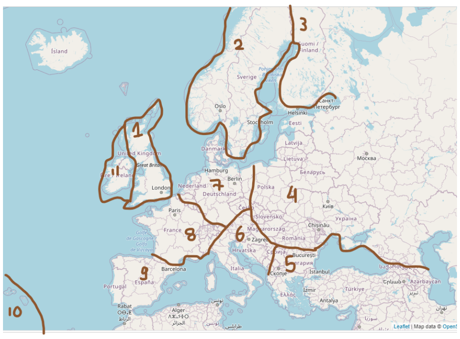

Figure 6.7: Cuts in the network of Europe

Region 10 represents all the global demand. Its exact location is thus not representative in the map; any bundling with other hubs would also make no sense since it has to pool with additional hubs in the network of Global. Region 11 is a leftover of region 1; however, this hub is unreachable within the suggested planning horizon of 2 days for region 1.

For most regions, the cuts in the network will not or slightly impact the solution. By the use of the matrix $F^{TR}_{i,j}$, which states if transportation is feasible between hub i and hub j, it is checked if each hub is reachable on time in the planning horizon a. If considering the full network, many hubs will not be feasible to supply each other on time and therefore should also not be considered. By cutting the network, naturally, a lot of possibilities that are not feasible are already removed from computation. Therefore, cutting the network is justified to some extent. However, there is also a downside, hub 25 could easily reach hub 28 on time, but they are separated from each other by their region definition. So the hubs at the border of a region may, in reality, be able to reposition more than is now possible and thus be able to operate with less safety stock. This problem mainly occurs at the intersection of region 7, region 8, region 6, and region 4. However, it is expected that approximating the safety stock levels by cutting the network in parts will, for most hubs, not have too much impact.

6.3.3 Performance results different robust models

Input parameters
In this section, the full network is analyzed for all model variants. The input parameters are in Table 6.3. The β is set at 0.95 since DH aims to have at least a fill rate of 95 %. The corresponding γ and Γ are set at 2. Since it seems realistic that most demand will lie within a span of two standard deviations by the company. The Chebyshev inequality states that at least 75 % of the demand will be covered (Papoulis & Pillai, 1991). Although, it is believed within DH that it behaves around a percentage much closer to the normal distribution of 95 % percent and thus within two standard deviations.

Model parameter	Parameter values
Group hubs	{13,17,22,23,44,48}
a	6
β	0,95
γ	2
Γ	2
Δ	2

Table 6.3: Region 1: model parameters

Additionally, other parameters are set to: full transport feasibility within the sub-region for $F_{i,j}^{TR}$, penalty costs C^p corrected for day, and a planning horizon a of 6 days. All the corresponding safety stock levels for the different models are in Appendix E.

Regional model performance
In Table 6.4 the performance of the different models is displayed for each region for model 1, model 3, and model 5 and Table 6.5 for model 4 and model 5.

It is immediately visible that using the hybrid method in model 1 in comparison to using only affine decision rules for approximating the adaptive optimization in model 3 has a price. The safety stock levels rise by 37 %, and the cost rise by 42 %. Using the hybrid method, despite its increase in computation time seems a valuable choice and has a significant safety stock reduction.

Forcing towards first stage integrality results in an increase of 5% in costs and 9% in stock for the hybrid method. For the affine decision rule approximation, it results in an increase of 4 % in cost and 9 % in stock. Therefore, it seems that integrality in the first stage is not very restrictive and continuous relaxation is valuable. However, since model 2 and model 4 are both only considering four regions these results are uncertain. Please note that, comparing a model with integrality restrictions against a model without these restriction always results in an higher stock level and higher costs. However, the model with integrality is more representative for reality and therefore preferred. Unfortunately, these integrality restrictions are costly in computation time and the impact of continuous relaxation seems marginal on the first stage.

Model 5 containing second stage integrality and first stage integrality with only affine decision rules in the approximation results in an increase of 21% in objective value and 5% in stock in respect of only having a first stage integrality constraint (model 4). This seems to impact the solution slightly more than the first stage integrality.

	Model 1		**Model 3**		**Model 5**	
	Model OBJ	**Stock**	**Cost OBJ**	**Stock**	**Cost OBJ**	**Stock**
Group 1	8202	443,91	11403	620,97	11901	648
Group 2	3165	185,31	5436	316,53	5766	333
Group 3	2475	122,61	3834	190,38	4044	201
Group 4	3201	194,61	6105	372,15	6330	387
Group 5	1050	81,12	1413	108,48	1488	114
Group 6	3873	230,79	5628	333,78	6036	357
Group 7	18717	1114,68	25035	1358,7	25371	1377
Group 8	5184	313,83	17802	407,55	24345	459
Group 9	2847	182,46	4599	296,4	4932	318
Group 10	2880	241,32	2880	241,32	2979	249
Group 11	267	17,10	732	39,9	768	42
Total	51861	3127,73	84867	4286,16	93960	4485

Table 6.4: Model performance per network group part I

Restrictions in computation time
Unfortunately, the problem size is too large for most of the network groups for model 2 and model 4. The groups consisting of five or more hubs are too large to solve the model within a reasonable amount for both models. Therefore, the effects on the smaller groups is studied and extrapolated towards the larger groups. The fact that Model 2 is only able to solve small groups was in the line of expectations. Figure 6.5 already indicated that after four hubs the computation time exploded for model 2. However, the fact that Model 4 is not able to solve all groups is remarkable, since Figure 6.6 indicates it should be able to do so. A possible explanation could be that the instance of DH is different than the randomized samples and the instance of DH is much harder to solve.

	Model 2		**Model 4**	
	Cost OBJ	**Stock**	**Cost OBJ**	**Stock**
Group 3	2649	129	3978	195
Group 5	1128	114	1488	114
Group 10	2901	243	2901	243
Group 11	768	42	768	42
Total	7446	528	9135	594

Table 6.5: Model performance per network group part II

Conclusions
Overall it is remarkable that model 2 is limited in analyzing and there is a need for more cuts in the network to use model 2 for all groups in the network. However, since this

severely will lead to sub-optimality, it is opted not to do so. Additionally, the gap in optimality by using the hybrid method or only the affine linear decision rule is large. Therefore, if possible, it is preferred to use the hybrid method. However, using the hybrid method makes it not possible to incorporate integrality in the first stage with the corresponding size of the current cuts in the network. Integrality in the second stage is regardless its computation time not possible in the second stage. However, after studying the impact of integrality in the first stage and second stage, the impact seems to be not very substantial for this particular problem. Especially, if we consider the loss that is obtained in respect of only using the affine decision rules. Therefore, model 1 is assumed to be the best performing model for DH despite its inability to treat integrality. Please note, that in reality thus some additional amount of safety stock is necessary for correcting for the continuous relaxation of the decision variables on the first and second stage. Therefore, the safety stock levels will be rounded up to their nearest integer for comparison with the SLM.

6.3.4 Benchmarking robust optimization with SLM total tanks

This section uses the results and elaborates more on the total stock levels present at a hub. For benchmarking, the results in Appendix E are used since the previous section argued its the most suitable for DH. The values of model 1 are rounded and rounded up in Table 6.6 and benchmarked against the SLM which are currently operational at DH.

For the SLM, there do not exist input parameters since it is entirely based on expert opinions and is adjusted over time, based on the experience of performance in the network. The lower limit is just the expected demand and thus the cycle stock at each hub, which is at least expected according to the forecast and concerns no additional stock for covering uncertainty. As elaborated in section 6.2.3 an upper limit can be constructed for the robust model. Although, a hub can have more tanks than the upper limit due to super stocking in that particular hub due to low transport towards other hubs or low holding costs. However, this upper limit indicates how much DH can reduce the stock level instead of only looking at each location and tank type individually. If there is no cooperation between the hubs, the robust model will be the same as an individual forecast model for each hub. Furthermore, it would never make sense to have more total tank containers in the network than the full value of the forecast. The corresponding benchmark of the SLM and the robust optimization model is in Table 6.6. The yellow marked cells indicate a relatively high difference between the SLM and the robust model and if marked a possible explanation for this difference is given in the paragraph of individual results.

Hub	**Mu**	**SLM**	**Model 1 (Round)**	**Model 1 (Round up)**	**Forecast**
1	294	360	363	369	444
2	51	48	42	42	105
3	33	39	15	21	75
4	6	0	0	0	18
5	27	3	42	54	81
6	15	9	30	33	51
7	12	0	18	18	42
8	36	66	105	117	84
9	24	30	39	42	54
10	30	12	39	45	78
11	48	117	90	102	249
12	N#A	N#A	N#A	N#A	N#A

Table 6.6 continued from previous page

Hub	**Mu**	**SLM**	**Model 1 (Round)**	**Model 1 (Round up)**	**Forecast**
13	6	0	15	18	30
14	33	108	57	60	87
15	33	24	21	33	87
16	6	24	12	15	24
17	66	72	87	93	120
18	24	24	39	42	60
19	45	45	3	3	105
20	54	24	78	84	114
21	21	24	33	36	63
22	63	54	93	99	141
23	36	66	51	57	72
24	72	177	0	3	138
25	42	30	165	174	96
26	21	27	30	39	57
27	6	0	15	18	24
28	105	198	153	162	183
29	27	9	42	45	69
30	15	3	3	6	45
31	18	24	30	33	54
32	15	0	27	30	45
33	48	42	42	51	102
34	12	24	27	27	36
35	15	3	9	12	45
36	21	3	33	36	63
37	21	3	30	36	63
38	24	9	45	48	72
39	378	354	453	456	552
40	96	174	132	138	174
41	21	186	27	33	51
42	12	24	24	27	48
43	12	9	21	24	42
44	78	111	102	111	144
45	6	9	63	69	18
46	12	0	24	27	42
47	15	9	24	30	51
48	69	63	96	105	141
	2295	**2640**	**2889**	**3123**	**4539**

Table 6.6: Benchmarking robust model against SLM

Overall results

On average, the results of the SLM and the Robust model are pretty close but the robust models favors in general more tank containers. The Robust model is slightly more conservative with, with over 90 more tanks in total for the rounding model and thus resulting in a 9% difference in the total amount of tanks in the network. For the model with rounding up for each tank type and thus accounting for continuous relaxation used in model 1, a difference of 161 tanks resulting in a 18% difference respectively.

These numbers are corrected for the tanks in the Global hub since this hub is not in the SLM. This also explains why the hub 12 in Table 6.6 has no value. Furthermore, the upper limit of only using a forecast of 4539 tanks, is quite far of the suggested stock level of robust optimization of 2889 and 3123. This indicates that using robust optimization is valuable instead of only looking at each hub and demand characteristic individually. Since it is unknown at DH which percentage of demand under the current operations is lost, it is tough to estimate the fill rate of both safety stock levels and to argue if SLM or the robust optimization model is better. However, we know that under the assumption that the individual demand and the total demand stay between two standard deviations, the robust model can achieve a fill rate of 95 %. Overall, the results imply that for achieving a fill rate of 95 % DH needs to invest in slightly more tank containers. Furthermore, it can be concluded that the results of robust optimization are quite far from its upper limit of treating each hub and tank type individually.

Individual results

From Table 6.6 some interesting deviations in the stock levels of model 1 and the SLM on hub level can be found:

- A frequently observed pattern in the difference between the SLM and the robust model is that the low-demand hubs have a structural shortage in the SLM. An example of these low-demand hubs experiencing these problems are hubs: 5, 7, 27, 36, 37, 38. After verifying with MMP planners why these hubs have such low stock levels, the following three reasons are given. Demand in these hubs generally has a more extended amount of decision days. So the number of decision days in low-demand hubs at the border of the network are usually in the right tail of Figure **??**. This extended time gives planners more time and, therefore, there is no need for additional safety stock over there. Customers in these regions are aware of the low amount of tank containers in that area and plan further ahead to account for this. Another reason is that demand to the borders is frequently in both directions, and thus there is a diminishing need for having stock. Lastly, in most cases, the actual stock level at each hub is relatively high because of the demand flow. Therefore, it can be that the SLM has some low limits, but the actual stock levels can solve these problems, and consequently, it may be hard to detect if these levels are indeed insufficient.
- It is visible that in hub 24, the model indicates 0 tank containers, whereas the SLM shows much more. This difference is because hub 25 has lower holding costs, and the transportation cost between 24 and 25 is set at 0. Therefore, hub 25 will store all these tank containers. However, in reality, the transportation costs are not exactly 0, and consequently, it makes no sense to super stock everything in hub 25. The same thing is happening with hub 3 and hub 45. Therefore, the SLM is probably right there, and if the input data would also contain trucking costs, it would be corrected for these hubs. Additionally, hub 24 is now only linked to the French regions, but in reality it is more connected to the German hubs. On top of that, there are also large customers in hub 24 with high variability, making it beneficial to have some additional stock there.
- It is visible that there must be more tank containers in hub 28, according to the SLM. Apparently, in hub 28, there are customers with extensive blacklists. The blacklist contains previous loadings, which are not allowed to congest with the current product. Due to these comprehensive lists, it is necessary to have more tank containers at these hubs to increase the probability of having a tank container, not on a blacklist of products. On top of that, there is a large customer in hub 28, which

is very volatile and also causing the additional need for safety stock.

- Interestingly, it is observed that the robust model favors more stock in hub 39. After verifying with MMP planners, Rotterdam is indeed in some cases low in stock caused by a relatively high variance.
- In Hub 41, there is a lot of stock in the SLM, which may be too conservative according to the model. An explanation is linked to the construction of hub 41. Hub 41 consists of a bundle of cities quite far from each other, receiving low demand. These distances may be a cause of the significant difference. However, this difference is way too large to be that high. So maybe the adjustment of the safety stock in hub 41 should be better.

6.3.5 Benchmarking tank types of robust optimization with SLM

There are seven tank types considered in this thesis, as shown in Table 1.1. To abbreviate notation, these tank types have been number T1 till T7 in the same order as present in Table 1.1. However, the SLM consist of only four tank types. This project created on purpose with three additional tank types, since further expansion was required for some orders and DH was curious about its impact. By merging T2 T3 and T4 in the ISO category and T5 and T6 in the Swap category, it is possible to compare the tank types of the SLM again. In Figure 6.8 the comparison is shown.

Total distribution of tank types
Overall, the distribution of tank types seems quite close. Impressive since tank containers do have an lifespan of around 10-15 years, and the corresponding demand patterns in demand characteristics can differ significantly over time. However, there are also some minor differences in the distribution of different tank types.

Preferably there would be a smaller percentage of swap_special tank containers and instead of these tank containers a higher percentage of 20_Feet_special tank containers. A swap_special tank container has the same advanced heating system as a 20_Feet_Special, but is larger in size. Having a Swap_Special generally is more flexible, but is also more expensive. The favoring of the 20_Feet_Special is driven by not having enough demand which is only treatable by the Swap_Special and therefore it is better to have a less expensive tank container for these orders.

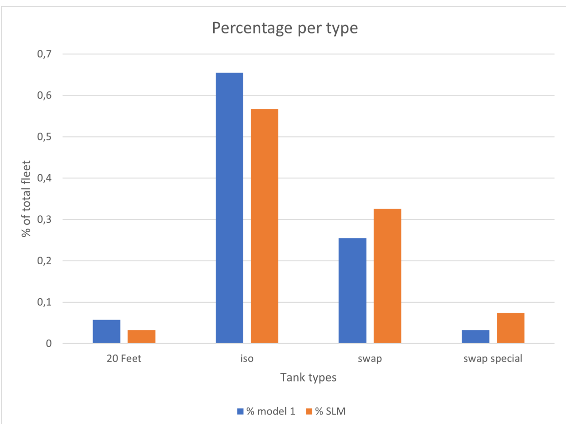

Figure 6.8: Tank type comparison SLM and model 1

Furthermore, a higher rate in ISO tank containers seems preferred which is compensated with a lower percentage in Swap tanks. However, since ISO requirements are frequently wrongly stated in the SPO or OT data, this conclusion must be taken with caution. Again the ISO tank is much cheaper in comparison with the Swap tank, but also their capabilities do differ.

Individual hub differences per tank types
In Table 6.7 the differences per tank types per hub are shown. The NA values represent the Glob hub and this hub is in reality non existent, since it bundles all demand outside Europe in one hub. The SLM has obviously no value for that hub and therefore these values are not present.

The blue color indicates a difference, which is likely due to changing Swap_Special to 20_Feet_Special. Red indicates that the robust model advises significantly fewer tank containers, and green indicates the robust model recommends a higher stock level.

For the 20_Feet category, it can be seen that the SLM means to super stock in hub 1 the 20_Feet_Special. In comparison, the model prefers to have them divided over some smaller hubs, experiencing that type of demand. Super stocking in Antwerp makes sense if planning is going multi-period. For example, if a 20_Feet_Special order is known far in advance, it is possible to load first a normal order in this more advanced tank-container already going towards the loading destination of the 20_Feet_Special order. Therefore, it is explainable that in a multi-period model with relaxation of the arrival of 2 decision days, it can be more beneficial to super stock in hub 1.

For the ISO tanks, it is observed that generally, the small hubs at the border need more stock level, as already was found on the total stock level analysis. Although the explanation of having more decision days in that part of the network may solve this. Interestingly, it is found that hub 39 of the large hubs needs much more ISO tank containers.

There are shortages found in hub 3 and hub 24 for the Swap tanks, which are caused by the holding costs and the low transport costs in in the model, which in reality may differ a bit more than suggested in the model.

Lastly, for the Swap_Special tanks, it is visible that the SLM prefers to stock them all in hub 39 or hub 1, whereas the robust model prefers to divide them more over the hubs experiencing this type of demand. Again the previous argument of going multi-period is a possible explanation for this favoring of super stocking the SLM.

Hub	**Model 20_Feet**	**SLM 20_Feet**	**Model ISO**	**SLM ISO**	**Model Swap**	**SLM Swap**	**Model Swap Spec**	**SLM Swap Spec**
1	12	45	231	180	120	105	6	30
2	0	0	33	24	9	24	0	0
3	0	0	12	9	9	30	0	0
4	0	0	0	0	0	0	0	0
5	12	0	33	0	6	3	3	0
6	6	6	21	0	6	3	0	0
7	0	0	15	0	3	0	0	0
8	0	0	78	30	39	36	0	0
9	0	0	42	30	0	0	0	0
10	6	0	30	6	9	6	0	0
11	6	0	75	81	21	21	0	15
12	N#A	N#A	N#A	N#A	N#A	N#A	N#A	N#A
13	0	0	12	0	6	0	0	0
14	6	0	42	105	6	3	6	0

Table 6.7 continued from previous page

Hub	Model 20_Feet	SLM 20_Feet	Model ISO	SLM ISO	Model Swap	SLM Swap	Model Swap Spec	SLM Swap Special
15	6	0	15	9	12	12	0	3
16	0	0	3	9	6	15	6	0
17	0	3	39	30	54	30	0	9
18	6	0	15	9	15	15	6	0
19	0	0	3	21	0	15	0	9
20	12	0	54	15	12	9	6	0
21	0	0	30	9	6	15	0	0
22	12	3	63	30	18	15	6	6
23	0	3	42	42	15	15	0	6
24	0	3	3	150	0	24	0	0
25	27	0	114	21	15	9	18	0
26	0	0	30	12	9	15	0	0
27	0	0	18	0	0	0	0	0
28	12	3	90	120	54	45	6	30
29	0	0	36	0	9	9	0	0
30	0	0	6	3	0	0	0	0
31	0	3	21	12	6	9	6	0
32	0	0	24	0	6	0	0	0
33	6	0	39	30	6	6	0	6
34	0	0	27	24	0	0	0	0
35	0	0	12	0	0	3	0	0
36	6	0	24	0	6	3	0	0
37	0	0	24	0	6	3	6	0
38	12	0	24	3	6	6	6	0
39	12	9	276	105	162	180	6	60
40	9	0	93	105	30	60	6	9
41	3	0	30	180	0	6	0	0
42	0	0	21	0	6	24	0	0
43	0	0	15	3	3	6	6	0
44	6	6	57	30	48	60	0	15
45	6	0	54	9	9	0	0	0
46	0	0	21	0	6	0	0	0
47	3	0	18	3	9	6	0	0
48	6	3	63	30	30	30	6	0
Total	192	87	2028	1479	798	876	105	198

Table 6.7: Per tank type difference model 1 and SLM

Type differences robust models

In Figure 6.9 the tank types are depicted for the different model solutions. Model 2 and Model 4 are ommited, since these models where not able to analyze the full network. It is visible model 3 and model 5 are relatively close. Furthermore, the adaptive model uses more T4 and T6, which are the more flexible tank containers. The use of only affine decision rule for approximating results in having a fleet consisting of a smaller percentage in flexible tank containers, but generally in more stock. Intuitive this is logical, since being able to better incorporate adaptability makes it more valuable to incorporate flexible tank container in respect of its more expensive price. Again these differences are not that substantial and only seem to be marginal.

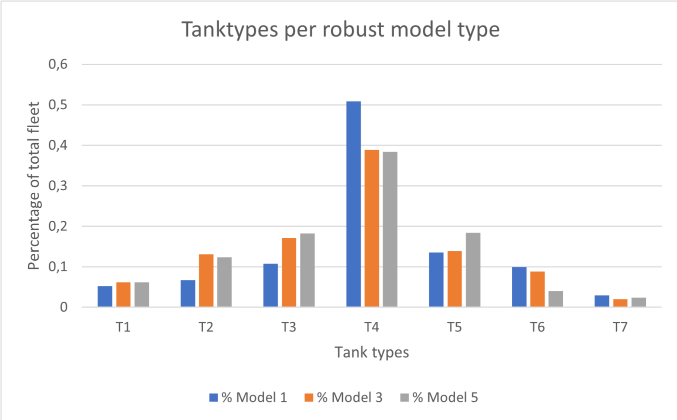

Figure 6.9: Percentage tank types per robust model

6.3.6 Sensitivity to uncertainty set

For the sensitivity analysis only Region 1 of model 1 is used since the sensitivity analysis shows the relationship in the parameters, and alternating all the parameters in the entire network would be too computationally extensive and likely repetitive in analysis. The uncertainty set, shown in Model 5.21 consists of three parameters. The γ controls the individual standard deviation of each hub and demand characteristic. The Γ controls the total standard deviation of the sum of all demand of a demand characteristic and the Δ total demand. Furthermore, the general fill rate constraint β also states how much uncertainty must be covered. There are thus four parameters in which the sensitivity is investigated towards the safety stock levels and the corresponding cost values.

Sensitivity to β

The sensitivity to β is in Figure 6.10. It is visible that only after the value of $0,96$ the fill rate constraint starts to constrain the problem. So setting the constraint lower than 96% fill rate will still result in a fill rate of 96 %. This is important for DH to consider, since they aim to have a fill rate of 95 %. Remarkably optimizing for DH its profit is the around the same value of fill rate and thus achieving this fill rate of 95 % is also from a cost perspective view the most optimal to do. After surpassing the level of a 96 % fill rate, it behaves linearly in respect of the costs and also in respect of its stock level. However, since the demand is restricted in the size of the uncertainty set at γ its individual standard deviation for each hub and each standard deviation, and Γ for the total standard deviation for each demand characteristic and Δ its total standard deviation, the corresponding structure becomes linear again of fulfillment of a certain amount of fill rate.

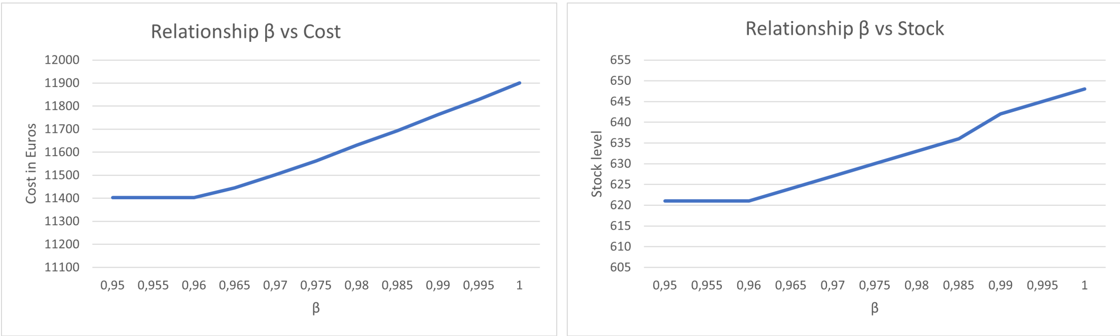

Figure 6.10: Sensitivity β

Sensitivity to γ and $\sigma_{i,c}$

The sensitivity to the parameter γ is in Figure 6.11. It is deducible that the cost and the number of tank containers are relatively linear with a dampening upwards trend. This trend is due to the Γ restriction on the total demand, which hinders the problem from growing further and becoming more significant. Furthermore, the model has been tested with a β constraint of $0,95$ and its relaxation. Interestingly, the models are equal till $\gamma = 5$, after that level the fill rate constraints seem to hinder the problem, and the problem starts to converge to doing no business at all and just pay all the penalty costs. This drop of starting to avoid bussiness due to its large variance and is clearly visible in the righthand side of Figure 6.10. The $\sigma_{i,c}$ is in the same procedure-related and shows the same sensitivity behavior since the γ is a multiple of the $\sigma_{i,c}$. Lastly, since the slope of γ or σ is very high in respect of the stock and costs, it indicates that a good forecast is valuable since a slight decrease of variance results in large reductions. Therefore, for future research it could be opted to use an more advanced forecast. With the knowledge that became apparent during the sensitivity analysis, it would have been a better decision to try to reduce the forecast error even more and use a LOESS model or even a more advanced forecasting model and take the difficulties in understanding for granted.

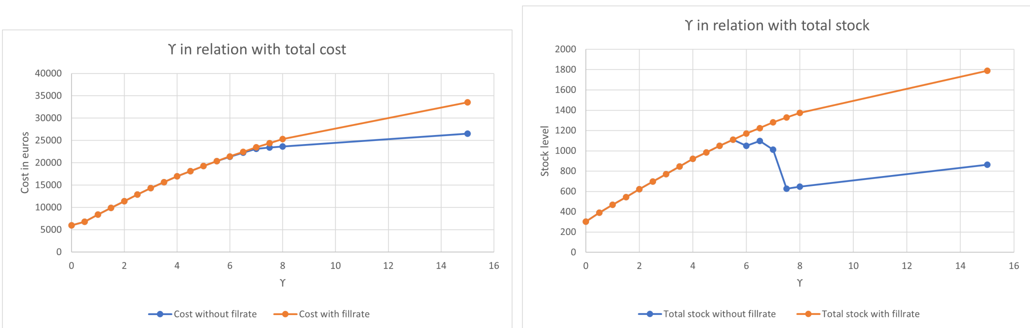

Figure 6.11: Sensitivity γ and $\sigma_{i,c}$

Sensitivity Γ and Δ

The parameter Γ is much less sensitive to an increase in safety stock level and total costs. Especially if it gets compared to the γ variation, it is interesting that the fill rate constraint in this particular example is not constraining the problem. Additionally, after reaching a Γ value of 2, the solution stays the same. The γ is curbing the problem, and therefore, it does not matter how much the Γ increases. After checking the Δ, this variable seem to constrain the problem to an even lesser extent than the Γ and therefore is thus not included in this analysis.

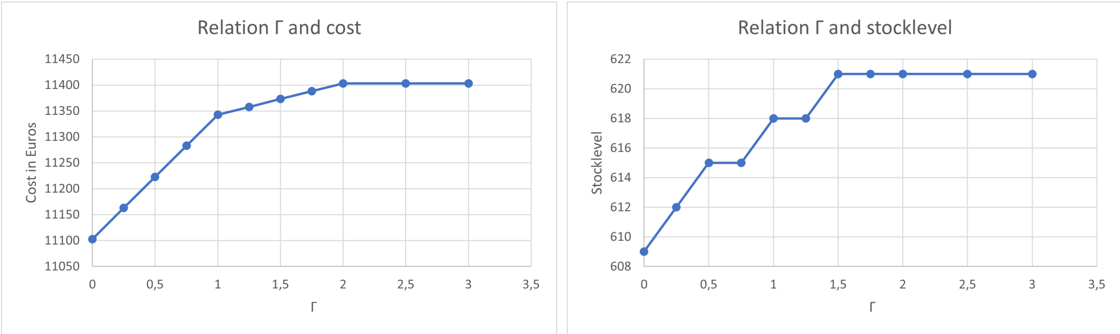

Figure 6.12: Sensitivity Γ

Limit distribution

Constraint violation could be tested if a probability distribution is known. Since the probability distribution is not known, it is not easy to do so. It is possible to assume that the size of the empirical distribution of the forecast error is large enough to represent its probability distribution. Only then, it would be possible in the continuous case to test for constraint violation. However, if the forecast method produces non-stationary error term due to left over effecs of seasonality or trends this method would not be valid. Since the use of testing constraint violation is using thus two large assumptions, it is opted not to do so. The Chebychev inequality remains on estimation the max probability of exceeding a certain amount of standard deviations. Generally, this is way too conservative. However, to give an indication, the Chebychev limit probability is used and plotted against the corresponding cost in Figure 6.13. Although the exact number may not be representative, the figure's corresponding exponential shape is.

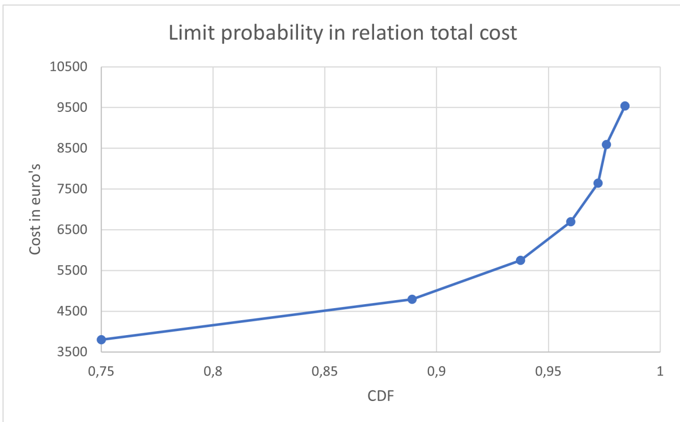

Figure 6.13: Sensitivity limiting CDF

Chapter 7

Conclusion and discussion

This chapter starts in section 7.1 with the recommendations for DH found in this project. The limitations of this project are shown in section 7.2. Section 7.3 describes future academic research directions deducible from this project. Section 7.4 provides a general conclusion and reflects on the main research question.

7.1 Recommendations DH

The following recommendations found in this project are applicable for implementation at DH:

- Update some safety stock levels in the SLM based on the values found in the robust model. Some of the small hubs with no current existing safety stock can be updated and also the very large hubs such as Rotterdam and Antwerpen seem to need some additional stock. A more detailed analysis on hub and tank level is in section 6.3.4. Although it is must be noted, that this must be done with caution and needs review by MMP planners.
- Favor investing in more 20_Feet_Special tank containers over Swap_Special tank containers. The current tank type distribution at DH is in line with what is found with the robust model. A slight deviation is found favoring more the 20_Feet_Special tank containers instead of the Swap_Special tank containers. The large volume orders requiring the Swap_Special tank containers do occur to a lesser extent and the associated costs of being more flexible with these larger tank containers are too high.
- Consider investing in more tank containers overall. It is found that in general there need to be slightly more tank containers, which is around 18%, overall in the network to sustain a fill rate of 95 %. In the sensitivity analysis it is found by relaxation of the fill rate constraint that around a 95 % fill rate, DH is the most cost-efficient. Therefore, investing in slightly more tank containers seems valid. Nonetheless, a critical review is needed since the robust model does incorporate some strong assumptions on the planning horizon.
- In the sensitivity analysis it becomes apparent that the forecast error on hub and tank level is the most influential in the associated cost structure of DH. Therefore, it is valuable to invest in a more advanced forecast than proposed in this project.
- Start registering missed demand. By starting to register missed demand, it becomes known how much demand is missed. This would make it possible to more fairly evaluate the performance of certain safety stock levels.
- Enrich MMP planners with as much support as possible in making planning decisions. Since planners make most decisions and the requirement of many assumptions

in modelling will require to do so in the future. Decision support tools for MMP planners such as an accurate daily forecast, accurate inventory levels in the SLM, working with inventory positions, and the ability to integrate all features in one tool.
- If possible apply strategic initiatives to reduce the presence of the problems aspects such as demand uncertainty, multi-commodity, timing issues, and network issues.

7.2 Limitations

The first limitation is the data quality used in the corresponding robust optimization and forecast models and thus its validity. For example, the demand characteristics are filled in the SPO in TF1 and OT in TF2, and its definitions differ in both systems. Therefore, when comparing the demand characteristics, the demand characteristic distribution starts to deviate after the introduction of TF2, which is in reality not valid. On top of that, some customers check multiple checkboxes on the SPO or OT to be sure. However, planners know what product the customer wants to transport and find a suitable fitting tank container but this is not corrected in the SPO or OT. This results in difficulty distinguishing the actual demand types. Additionally, multiple lanes are not present on the transport data, and not all trucking costs are available. Therefore, it results in many unknown data points which are considered as unfeasible which may actually be feasible. The lack of data quality can dampen the effect of inventory pooling resulting in too conservative safety stock levels.

The second limitation is its inability to test all four problems dimensions at once. First of all, the network is cut in parts, which results in hubs on the border not being able to pool inventory with a nearby hub, which may be possible. Even with cuts in the network, the models dealing with integrality had difficulties in keeping it computationally tractable. Secondly, the timing dimension is fixed. It became apparent that complicated orders or orders at the network border may use earlier planning in the planning process. This aspect is thus not incorporated in the model, causing a bit conservative stock levels at low demand hubs. Thirdly, not going multi-period does not integrate the redistribution of tank container and its safety stock levels. If these had been optimized together, it could have resulted in even better safety stock levels.

7.3 Future research

The nature of the problem at DH is very challenging and making it an excellent motivator for further academic research. Especially since there is much potential business value with substantial cost reductions. Methods on faster computation or approximation of robust optimization would enlarge the practicability of the solution. Some exciting problem extensions are listed below:

- Including integrality on the second stage decision variables for the hybrid approximation method. This method was able to outperform the approximation with only affine decision rules, but with the loss of second stage integrality.
- Extension of the corresponding demand types and tank container types would be possible.
- The ability to analyze the full network without making cuts in the network using multi-commodity with a constant planning horizon.
- The ability to go multi-period and to differ thus decision making on different planning horizons for different types of orders.
- The ability to go multi-period and to incorporate repositioning decision making. Currently, these models run sequentially. A downside in running these models se-

quentially is not incorporating the pooling effect of making it more attractive to reposition tank containers from nearby hubs, which already have to go in that direction due to repositioning.
- The incorporation of the pricing effect could be injected into the model. For example, additional inventory could be sold for lower rates, and incorporating these parameters could result in interesting stock level management solutions.

7.4 Conclusion

This project started with the aim of improving the stock level management process of DH and its safety stock levels. The stock level management process is mapped and strategic initiatives to improve the supply chain are given.

For the determination of safety stock levels, an adaptive robust optimization problem is formulated. By creating cuts in the network on geographical borders, it is possible to analyze the large network of DH and thus use robust optimization in practice. The hybrid approximation of adaptive robust optimization resulted in 40 % lower costs and stock in full continuous relaxation in respect to the use of approximation by only affine decision rules. Unfortunately, the hybrid approximation model is not able to include integrality on the second stage and the first stage is computationally too extensive. After analysis of the effects of integrality by the use of affine decision rules, these effects seem not too restrictive. Therefore, it is decided to use the hybrid approximation with continuous relaxation and round up the number of tank containers to account for the relaxation.

Overall the results are promising and in line with current safety stock levels in the SLM resulting in nice recommendations in section 7.1. However, it must be mentoined that the data quality needs improvement and thus the validity of the solution. Especially, errors in the transportation costs can results in local deviations of safety stock. The provided forecasting models, robust optimization models, and redistribution models can provide support in doing planning operations. DH has thus a valuable scientific method on how to construct safety stock levels in its network. However, due to the challenging environment of DH, it is not possible to incorporate all problem features and thus fully automate the planning process by relying on mathematical models even when the data quality is improved. Therefore, human adjustment of the safety stock levels must remain necessary. Lastly, some exciting model extensions are possible with much business value for DH, which is mentioned in section 7.3, if the literature on adaptive robust optimization is extended.

References

Atkeson, C. G., Moore, A. W., & Schaal, S. (1997). Locally Weighted Learning. *Artificial Intelligence Review*, *11*(1-5), 11–73. doi: 10.1007/978-94-017-2053-3{_}2

Ben-Tal, A., Goryashko, A., Guslitzer, E., & Nemirovski, A. (2004). Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, *99*(2), 351–376. doi: 10.1007/s10107-003-0454-y

Ben-Tal, A., & Nemirovski, A. (1999). Robust solutions of uncertain linear programs. *Operations Research Letters*, *25*(1), 1–13. doi: 10.1016/S0167-6377(99)00016-4

Bertsimas, D., & Brown, D. B. (2009). Constructing uncertainty sets for robust linear optimization. *Operations Research*, *57*(6), 1483–1495. doi: 10.1287/opre.1080.0646

Bertsimas, D., & De Ruiter, F. J. (2016). Duality in two-stage adaptive linear optimization: Faster computation and stronger bounds. *INFORMS Journal on Computing*, *28*(3), 500–511. doi: 10.1287/ijoc.2016.0689

Bertsimas, D., Doan, X. V., Natarajan, K., & Teo, C.-p. (2010). Models for Minimax Stochastic Linear Optimization Problems with Risk Aversion. *Mathematics of Operations Research*, *35*(3), 580–602.

Bertsimas, D., Gupta, V., & Kallus, N. (2018). *Data-driven robust optimization* (Vol. 167) (No. 2). Springer Berlin Heidelberg. doi: 10.1007/s10107-017-1125-8

Bertsimas, D., & Sim, M. (2003). Robust discrete optimization and network flows. *Mathematical Programming*, *98*(1-3), 49–71. doi: 10.1007/s10107-003-0396-4

Buhayenko, V., & Den Hertog, D. (2017). Adjustable robust optimisation approach to optimise discounts for multi-period supply chain coordination under demand uncertainty. *International Journal of Production Research*, *55*(22), 6801–6823. Retrieved from `https://doi.org/10.1080/00207543.2017.1351635` doi: 10.1080/00207543.2017.1351635

Chassein, A., Goerigk, M., Kasperski, A., & Zieliński, P. (2018). On recoverable and two-stage robust selection problems with budgeted uncertainty. *European Journal of Operational Research*, *265*(2), 423–436. doi: 10.1016/j.ejor.2017.08.013

Chatfield, C. . (1978). The Holt-Winters Forecasting Procedure. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *27*(3), 264–279.

Cleveland, R., Cleveland, W., McRae, J., & Terpenning, I. (1990). STL-A-Seasonal-Trend-Decomposition-Procedure-Based-on-Loess. *Journal of Official Statistics*, *6*, 3–73.

Epstein, R., Neely, A., Weintraub, A., Valenzuela, F., Hurtado, S., Gonzalez, G., ... Weintraub, A. (2012). Optimization in a Major Shipping Company. *Interfaces*, *42*(April 2016), 5–16.

Erera, A. L., Morales, J. C., & Savelsbergh, M. (2005). Global intermodal tank container management for the chemical industry. *Transportation Research Part E: Logistics and Transportation Review*, *41*(6 SPEC. ISS.), 551–566. doi: 10.1016/j.tre.2005.06.004

Erera, A. L., Morales, J. C., & Savelsbergh, M. (2009). Robust optimization for empty repositioning problems. *Operations Research*, *57*(2), 468–483. doi: 10.1287/opre.1080.0650

Even, S., Itai, A., & Shamir, A. (1975). On the complexity of timetable and multi-commodity flow problems. *Proceedings - Annual IEEE Symposium on Foundations of Computer Science, FOCS*, *1975-Octob*, 184–193. doi: 10.1109/SFCS.1975.21

Gabrel, V., Murat, C., & Thiele, A. (2014). Recent advances in robust optimization: An overview. *European Journal of Operational Research*, *235*(3), 471–483. Retrieved from `http://dx.doi.org/10.1016/j.ejor.2013.09.036` doi: 10.1016/j.ejor.2013.09.036

Gardner, E. S. (2006). Exponential smoothing: The state of the art-Part II. *International Journal of Forecasting*, *22*(4), 637–666. doi: 10.1016/j.ijforecast.2006.03.005

Ghadimi, F., Aouam, T., & Vanhoucke, M. (2020). Optimizing production capacity and safety stocks in general acyclic supply chains. *Computers and Operations Research*, *120*, 104938. Retrieved from `https://doi.org/10.1016/j.cor.2020.104938` doi: 10.1016/j.cor.2020.104938

Goerigk, M., & Kurtz, J. (2020). Data-driven robust optimization using unsupervised deep learning. *arXiv*.

Grzesica, D., & Wicek, P. (2016). Advanced Forecasting Methods Based on Spectral Analysis. *Procedia Engineering*, *161*, 253–258. doi: 10.1016/j.proeng.2016.08.546

He, J., Huang, Y., & Chang, D. (2015). Simulation-based heuristic method for container supply chain network optimization. *Advanced Engineering Informatics*, *29*(3), 339–354. Retrieved from `http://dx.doi.org/10.1016/j.aei.2014.08.001` doi: 10.1016/j.aei.2014.08.001

Hofenk, D., Schipper, R., Semeijn, J., & Gelderman, C. (2011). The influence of contractual and relational factors on the effectiveness of third party logistics relationships. *Journal of Purchasing and Supply Management*, *17*(3), 167–175. Retrieved from `http://dx.doi.org/10.1016/j.pursup.2011.04.003` doi: 10.1016/j.pursup.2011.04.003

Holle, A. (2019). *Msc. Thesis: Decision support tool for pricing strategies in transportation networks* (Unpublished doctoral dissertation). Technische Universiteit Eindhoven.

Hyndman, R. (2018). *Forecasting: Principles and practices* (2nd editio ed.). Melbourne: OTexts. Retrieved from `https://otexts.com/fpp2/`

ITCO. (2020). *GLOBAL TANK CONTAINER FLEET.* Retrieved from `https://www.international-tank-container.org/storage/uploads/ITCO_2020_Global_Fleet_Report_Final_4_March_2020.pdf`

Jung, J. Y., Blau, G., Pekny, J. F., Reklaitis, G. V., & Eversdyk, D. (2004). A simulation based optimization approach to supply chain management under demand uncertainty. *Computers and Chemical Engineering*, *28*(10), 2087–2106. doi: 10.1016/j.compchemeng.2004.06.006

Karimi, I. A., Sharafali, M., & Mahalingam, H. (2005). Scheduling tank container movements for chemical logistics. *AIChE Journal*, *51*(1), 178–197. doi: 10.1002/aic.10295

Lambert, Y., Saillard, P., & Bathias, C. (2009). Application of the minimum cost flow problem in container shipping. *Fracture Mechanics: Nineteenth Symposium*(June), 318–318. doi: 10.1520/stp33081s

Lee, Y., Song, S. H., & Cheong, T. (2018). The value of supply chain coordination under moral hazard: A case study of the consumer product supply chain. *PLoS ONE*, *13*(3), 1–16. doi: 10.1371/journal.pone.0194043

Li, J. A., Leung, S. C., Wu, Y., & Liu, K. (2007). Allocation of empty containers between multi-ports. *European Journal of Operational Research*, *182*(1), 400–412. doi: 10.1016/j.ejor.2006.09.003

Marchetti-Spaccamela, A., & Segal, M. (2011). A scenario based approach for robust linear optimization. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6595 LNCS*(August). doi: 10.1007/978-3-642-19754-3

Moon, Y., & Yao, T. (2011). A robust mean absolute deviation model for portfolio optimization. *Computers and Operations Research*, *38*(9), 1251–1258. Retrieved from `http://dx.doi.org/10.1016/j.cor.2010.10.020` doi: 10.1016/j.cor.2010.10.020

Nahmias, S., & Olson, T. (2015). *Production and operations analysis* (7th Editio ed.). Long Grove: Wavelend PRess.

Nelson, H. L., & Granger, C. W. (1979). Experience with using the Box-Cox transformation when forecasting economic time series. *Journal of Econometrics*, *10*(1), 57–69. doi: 10.1016/0304-4076(79)90064-2

Ouorou, A. (2011). Affine decision rules for tractable approximations to robust capacity planning in telecommunications. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6701 LNCS*, 277–282. doi: 10.1007/978-3-642-21527-8{_}32

Papoulis, A., & Pillai, U. (1991). *Probability, random variables, and stochastic processes* (4th Editio ed.). Quebec: McGrawHill. doi: 10.1201/9781351105668

Petruzzi, N. C., & Dada, M. (1999). Pricing and the newsvendor problem: a review with extensions. *Operations Research*, *47*(2), 183–194. doi: 10.1287/opre.47.2.183

Phillips-Wren, G., Jain, L. C., Nakamatsu, K., & Howlett, R. J. (2010). Advances in intelligent decision technologies: Proceedings of the Second KES International Symposium IDT 2010. *Smart Innovation, Systems and Technologies*, *4*(September 2014). doi: 10.1007/978-3-642-14616-9

Powell, W. B. (2003). Dynamic Models of Transportation Operations. *Handbooks in Operations Research and Management Science*, *11*(C), 677–756. doi: 10.1016/S0927-0507(03)11013-4

Powell, W. B., Jaillet, P., & Odoni, A. (1995). Stochastic and dynamic networks and routing. *Handbooks in Operations Research and Management Science*, *8*(C), 141–295. doi: 10.1016/S0927-0507(05)80107-0

Rondinelli, D., & Berry, M. (2000). Multimodal transportation, logistics, and the environment: Managing interactions in a global economy. *European Management Journal*, *18*(4), 398–410. doi: 10.1016/S0263-2373(00)00029-3

Schuster Puga, M., Minner, S., & Tancrez, J. S. (2019). Two-stage supply chain design with safety stock placement decisions. *International Journal of Production Economics*, *209*(October 2016), 183–193. doi: 10.1016/j.ijpe.2018.05.018

Shang, C., Huang, X., & You, F. (2017). Data-driven robust optimization based on kernel learning. *Computers and Chemical Engineering*, *106*, 464–479. Retrieved from `http://dx.doi.org/10.1016/j.compchemeng.2017.07.004` doi: 10.1016/j.compchemeng.2017.07.004

Shapiro, S., & Wilk, B. (1963). A Comparative Study of Various Tests for Normality. *Journal of the American Statistical Association*, *63*(324), 1343–1372.

Sharypova, K. (2014). *Optimization of Hinterland Intermodal Container Transportation* (No. 2014). doi: 10.6100/IR766077

Silver, E. A., & Silver, E. A. (2017). Inventory Management : An Overview , Canadian Publications , Practical Applications and Suggestions for Future Research Inventory Management : An Overview , Canadian Publications , Practical Applications and Suggestions for Future Research. , *5986*. doi: 10.3138/infor.46.1.15

Snyder, R. (1980). The Safety Stock Syndrome. *The Journal of the Operational Research Society*, *31*(9), 833–837.

Wensing, T., Sternbeck, M. G., & Kuhn, H. (2018). Optimizing case-pack sizes in the bricks-and-mortar retail trade. *OR Spectrum*, *40*(4), 913–944. Retrieved from `https://doi.org/10.1007/s00291-018-0515-5` doi: 10.1007/s00291-018-0515-5

Young Yun, W., Mi Lee, Y., & Seok Choi, Y. (2011). Optimal inventory control of empty containers in inland transportation system. *International Journal of Production Economics*, *133*(1), 451–457. Retrieved from `http://dx.doi.org/10.1016/j.ijpe.2010.06.016` doi: 10.1016/j.ijpe.2010.06.016

Zhang, P. G. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*, 159–175. doi: 10.1016/S0925-2312(01)00702-0

Zhi, C., Melvyn, S., & Peng, X. (2019). *Robust Stochastic Optimization Made Easy.* Retrieved from `https://www.rsomerso.com/`

Appendix A

Affine decision rule model

The affine decision rule model is in A.1. Please note, the constraints do not have a definition for the set used due to spacing issues. These have However, these definitions should be added to model A.1 and are exactly the same as in model 5.25.

$$
\begin{aligned}
\underset{S_{i,p},Q_{i,j,c,p}}{\text{minimize}} \quad & \max_{\mathcal{W}} \sum_{i\in\mathcal{I}} \left(\sum_{p\in\mathcal{P}} \Big(S_{i,p} - \sum_{j\in\mathcal{J}} \sum_{c\in\mathcal{C}} (V_{i,j,c,p} * D_{i,c} + U_{i,j,c,p}) * F^{TR}_{i,j} * F_{c,p} \Big) * C^H_i \right) \\
& + \left(\sum_{i\in\mathcal{I}} \sum_{c\in\mathcal{C}} \Big(D_{i,c} - \sum_{p\in\mathcal{P}} \sum^{J} (V_{i,j,c,p} * D_{i,c} + U_{i,j,c,p}) * F^{TR}_{i,j} * F_{c,p} \Big) \right) * C^P \\
& + \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \sum_{c\in\mathcal{C}} \sum_{p\in\mathcal{P}} \left(V_{i,j,c,p} * D_{i,c} + U_{i,j,c,p}) * F^{TR}_{i,j} * F_{c,p} * C^T_{i,j} \right) \\
& + \sum_{p\in\mathcal{P}} \Big(\sum_{i\in\mathcal{I}} S_{i,p} \Big) * C^O_p \\
\text{subject to} \quad & \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \sum_{c\in\mathcal{C}} \sum_{p\in\mathcal{P}} (V_{i,j,c,p} * D_{i,c} + U_{i,j,c,p}) \quad , \\
& * F_{c,p} * F^{TR}_{i,j} \geq \beta * \sum_{i\in\mathcal{I}} \sum_{c\in\mathcal{C}} D_{i,c}, \\
& \sum_{p\in\mathcal{P}} \sum_{j\in\mathcal{J}} (V_{i,j,c,p} * D_{i,c} + U_{i,j,c,p}) * F^{TR}_{i,j} * F_{c,p} \leq D_{i,c}, \\
& \sum_{c\in\mathcal{C}} \sum_{j\in\mathcal{J}} (V_{i,j,c,p} * D_{i,c} + U_{i,j,c,p}) * F^{TR}_{i,j} * F_{c,p} \leq S_{i,p}, \\
& S_{i,p} \in \mathbb{Z}^+, \\
& (V_{i,j,c,p} * D_{i,c} + U_{i,j,c,p} \in \mathbb{Z}^+, \\
& \mathcal{W} = \Big\{ D \in \mathbb{R}^{|\mathcal{I}|*|\mathcal{C}|: D_{i,c}} \leq \mu_{i,c} + \gamma * \sigma_{i,c}, \\
& D_{i,c} \geq \mu_{i,c} - \gamma * \sigma_{i,c}, \\
& D_{i,c} \geq 0, \\
& \sum_{i\in\mathcal{I}} D_{i,c} \leq \Gamma * \Sigma_c, \\
& \sum_{i\in\mathcal{I}} \sum_{c\in\mathcal{C}} D_{i,c} \leq \Delta * \Psi \quad \Big\}
\end{aligned}
\tag{A.1}
$$

Appendix B

Mean demand analaysis

Hub	From	From	From	To	To	To	Diff	Diff	Diff
	μ	σ	cv	μ	σ	cv	μ	σ	cv
1	4293	498	0,36	2181	207	0,27	-2112	489	0,69
2	519	90	0,51	705	117	0,48	183	90	1,47
3	174	45	0,75	240	60	0,75	66	72	3,36
4	3	9	9,54	33	15	1,26	30	18	1,74
5	57	18	0,96	348	96	0,84	291	99	1,02
6	27	24	2,61	150	84	1,71	123	72	1,77
7	153	87	1,74	306	69	0,69	153	69	1,35
8	363	78	0,63	327	75	0,69	-36	87	7,08
9	135	63	1,44	543	144	0,81	408	168	1,23
10	126	42	0,96	225	48	0,63	99	48	1,50
11	2424	426	0,54	2508	384	0,45	84	498	18,06
12	687	174	0,75	285	93	0,96	-402	219	1,65
13	30	39	3,90	186	75	1,23	153	60	1,14
14	252	78	0,93	348	81	0,69	93	117	3,78
15	171	60	1,08	258	45	0,54	90	69	2,34
16	0	3	10,59	66	12	0,54	66	12	0,54
17	564	312	1,65	555	144	0,78	-9	276	98,43
18	201	48	0,72	201	93	1,38	3	105	188,91
19	414	87	0,63	297	54	0,54	-117	87	2,25
20	192	114	1,77	174	93	1,62	-18	66	11,19
21	33	33	2,97	216	66	0,90	183	69	1,17
22	651	159	0,72	1620	396	0,75	969	345	1,08
23	558	150	0,81	750	189	0,75	192	213	3,33
24	708	114	0,48	870	339	1,17	162	351	6,48
25	282	66	0,72	276	66	0,72	-6	84	51,06
26	60	24	1,14	162	30	0,57	102	33	0,96
27	9	12	4,44	27	24	2,49	18	27	4,50
28	894	225	0,75	1113	177	0,48	219	156	2,13
29	138	45	0,96	279	63	0,66	141	69	1,47
30	21	30	4,20	18	18	2,88	-6	30	19,86
31	66	24	1,05	192	45	0,72	126	54	1,26
32	6	9	3,99	12	6	1,26	6	9	5,40
33	504	78	0,45	513	87	0,51	6	87	39,78
34	21	21	3,24	72	18	0,75	51	30	1,80

Table B.1 continued from previous page

Hub	**From**	**From**	**From**	**To**	**To**	**To**	**Diff**	**Diff**	**Diff**
35	15	9	1,56	69	21	0,96	54	21	1,17
36	24	12	1,59	108	30	0,84	84	36	1,26
37	45	33	2,34	213	42	0,60	171	60	1,05
38	93	39	1,29	78	27	1,08	-15	45	9,54
39	4764	657	0,42	3339	552	0,51	-1422	573	1,20
40	1332	165	0,36	1212	180	0,45	-120	216	5,43
41	255	75	0,87	648	135	0,63	393	135	1,02
42	108	48	1,32	57	21	1,11	-51	60	3,60
43	54	27	1,38	183	57	0,93	129	54	1,26
44	1071	267	0,75	765	150	0,60	-306	309	3,03
45	3	9	5,97	42	18	1,23	39	18	1,32
46	9	12	4,02	123	42	1,02	114	45	1,23
47	63	21	1,05	102	33	1,02	39	45	3,42
48	987	222	0,66	567	141	0,75	-420	246	1,74
Total	23559	2271	0,30	23559	2271	0,30	0	0	0,00

Table B.1: Mean, STD and CV of flows of demand in hubs

Appendix C

Forecasting Methods

C.1 Model description forecasting

SARIMA model

The SARIMA models use a different approach and utilize auto-regressive terms on previous values to determine the current forecast value. It uses the previous demand value y_t and the forecast error ϵ for the current forecast (Hyndman, 2018). The SARIMA models combine these two regression series, which are depicted in Equation C.1

$$\begin{aligned} \Phi(B)y_i &= (1 - \Phi_1 B^1 - \Phi_2 B^2 - \cdots - \Phi_p B^p)y_t \\ \Theta(B)\varepsilon_i &= (1 + \Theta_1 B^1 + \Theta_2 B^2 + \cdots + \Theta_q B^Q)\varepsilon_t \end{aligned} \tag{C.1}$$

The B is the backshift operator and indicates how many periods a value is lagged. For example, B^2 means that a value is lagging two periods. The ϕ is the regression coefficient for the previous demand values y_t, and the θ represents the regression coefficient for the error terms ϵ_t. For example, the term $\phi_2 * B^2 * y_t$ can be rewritten into $\phi_2 * y_{t-2}$. Please note that Equation C.1 is a series containing multiple regression terms combined with multiple lagged values.

After the model design of both auto-regressive time series, the SARIMA model combines both regression series by the use of Equation C.2. The $\Phi(B^M)$ and $\Theta(B^M)$ denotes the auto-regressive terms for the seasonal part, whereas $\phi(B)$ and $\theta(B)$ represent the regression terms for the other part. $(1 - B)^d$ represents a general lag for removing the trend, with the d representing the amount of finite differencing. $(1 - B^M)^D$ is a lag for the seasonality part. In case there is no trend or seasonality present, it can be decided to drop these terms.

$$\phi(B)\Phi(B^m)(1 - B)^d(1 - B^m)^D y_i \;=\; \theta(B)\Theta(B^m)\varepsilon_i \tag{C.2}$$

For model usage in forecasting, it needs to be decided which lagged values should be included and which can be omitted. This class of models can also be optimized on the AIC; immediately finding the number of parameters and which lags should be included. For determining a specific h steps ahead time interval, fill as time index $t + h$ in for y_t and ϵ_t. Since the backshift operator is used, it will convert to already known previous values, and so a forecast value can be obtained.

Since there exist many possible combinations of SARIMA models, there are abbreviated notations. A SARIMA model is notated in $(p, d, q)(P, D, Q)^m$. The capital letters are the seasonality parameters, and the non-capital letters the other parameters of the forecast. With the p the amount of auto-regressive terms on the y_t, the q the amount of auto-regressive terms on the error term, and the d general lags to correct for a trend or seasonality.

LOESS model
LOESS starts with a weighted linear least squares regression on a moving time window (Hyndman, 2018). A moving time window states it only uses the last n observations for determination of its parameters and the data used for determination of the forecast shift thus over time. By omitting the values further in the past, it is more reactive to change. A good value for the determination of its time window depends on the dataset structure and is an important model design decision. Additionally, a weight function is used to value information even more. Having a weight function and combining it with a moving time window makes it thus even more reactive to sudden change.

The tricube weight function, which is in Equation C.3 is used in LOESS on each observation i considered in the moving time window of size n:

$$w_i(x_i) = (1 - |d_i|^3)^3 \tag{C.3}$$

With the w being the weight factor of the observation and the d being the absolute distance between the fitted value of observation i minus the true demand of observation i in Equation C.3. Please note that this is a moving time windows regression and each time all the weight factors for each of the observations are re-updated and the indices do shift.

After formulating the residual sum of squares and minimizing this function, the minimum is attained having the following estimator in matrix notation in Equation 5.1. The Y is a vector containing the true demand values, the $\hat{X}$ is a vector with at that time point known demand values, and the $W(X)$ is the weight function assigning weights towards each observation.

$$A(x) = YW(x)\hat{X}^T(\hat{X}W(x)\hat{X}^T)^{-1} \tag{C.4}$$

However, LOESS is an advanced decomposition method and uses an algorithmic procedure by using this moving time window weighted OLS for decomposition of the normal time series into a seasonality series, a trend series, and error term series. With an inner and outer algorithmic procedure, LOESS decomposes iteratively the three series which is described in Cleveland et al. (1990). By using these series it is possible to construct a forecast.

C.2 Model performance

Coefficients ETS
The ETS model finds the most optimal model based on minimization of the AIC, the (A,N,N). Resulting in additive error terms, not a trend, and not a seasonality factor. The corresponding coefficients are $\alpha = 0.25$ and $level = 24.444$.

Coefficients SARIMA
The auto SARIMA models find a trend a regression coefficient for moving average $\theta =$

−0.58 on the normal time series, and a $\phi = 0.40$ on the finite difference time series for seasonality.

Coefficients LOESS
The LOESS algorithm decomposes the time series in the components depicted in C.1. It is visible that the LOESS algorithm identifies the corona crisis correctly and can alternate the trend pattern accordingly. The other algorithms did not grasp the trend pattern accurately due to its alternating behavior, whereas LOESS can do that. The found seasonality factor demonstrates similar results as the earlier found seasonality—the high volume months in March and October and a low volume month in December.

C.3 Conclusion forecast

The corresponding performances of the forecasting models are in Table C.1. The ME indicates for all models not too much bias to worry about structural bias. The corresponding RMSE, MAE, MAPE, MPE all favor the LOESS model. Only looking at the accuracy results favors the LOESS model, as second the SARIMA model gets preferred, and lastly, the ETS model gets preferred. However, recall the discussion regarding presenting the seasonality and trending pattern. These effects are significant, but the impact is very low. Using sophisticated forecasting algorithms to reveal additional information on the decomposition of the seasonality and trend patterns will have a relatively small positive impact on reducing the forecast error. DH wants besides a good performing forecasting method, also a forecasting method that is explainable to their MMP planners and management. Since the gains in the SARIMA and LOESS model are small and the comprehensibility of an ETS model is much higher, it is decided to use ETS in forecasting. The ETS model will thus be used to forecast the demand for each hub and demand characteristic individually.

Forecast Method	ME	RMSE	MAE	MPE	MAPE
ETS model	-19,44	2147,7	1636,5	-0,87 %	7,24 %
SARIMA model	113,52	661,1	1983,3	-0,12 %	6,93 %
LOESS model	-43,2	1270,2	1095,3	-0,54 %	4,79 %

Table C.1: Benchmark forecasting methods total demand series

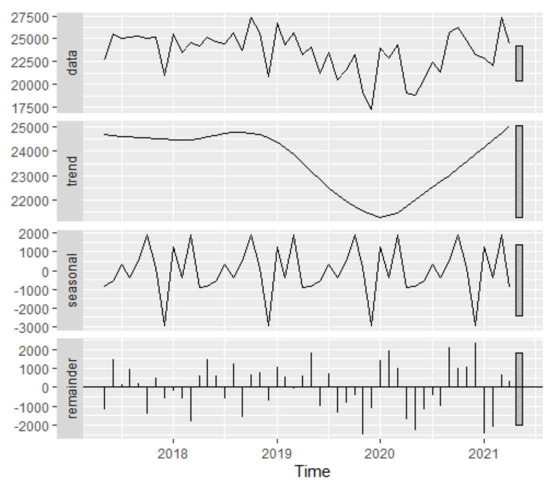

Figure C.1: Decomposition using Loess

Appendix D

Results ETS forecast

D.1 μ values

Hub	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Total
1	111	3	18	3	102	3	6	12	36	9	303
2	18	3	6	0	15	0	3	3	3	3	54
3	12	0	3	0	6	0	3	3	3	3	33
4	3	0	0	0	3	0	0	0	0	0	6
5	3	3	3	3	3	0	3	3	3	3	27
6	3	3	0	0	3	0	0	0	3	3	15
7	6	0	0	0	3	0	0	3	0	0	12
8	12	0	3	0	9	0	3	3	3	3	36
9	0	0	0	0	0	0	3	3	15	3	24
10	6	3	3	0	6	0	3	3	3	3	30
11	0	0	0	0	0	0	6	54	102	12	174
12	12	0	3	0	15	0	3	15	3	3	54
13	0	0	3	0	3	0	0	0	0	0	6
14	9	3	3	0	3	3	3	6	3	0	33
15	9	3	3	3	6	0	0	3	3	3	33
16	0	0	0	0	3	3	0	0	0	0	6
17	18	0	3	0	45	0	0	0	3	3	72
18	6	3	0	0	12	3	0	0	3	0	27
19	15	6	3	3	3	3	3	3	3	3	45
20	18	3	3	3	9	3	3	9	3	3	57
21	3	0	3	0	3	0	3	3	3	3	21
22	27	3	3	3	12	3	3	3	3	3	63
23	15	0	3	0	9	0	3	3	3	0	36
24	18	3	6	3	12	3	6	6	12	3	72
25	12	3	3	0	6	3	3	3	9	3	45
26	3	0	3	0	6	0	0	3	3	3	21
27	0	0	0	0	0	0	0	3	3	0	6
28	27	3	12	3	42	3	3	3	9	3	108
29	6	0	3	0	6	0	3	3	3	3	27
30	3	0	3	0	3	0	0	3	3	0	15
31	3	0	3	0	3	3	0	3	3	0	18
32	3	0	0	0	3	0	0	3	3	3	15
33	15	3	3	0	9	3	3	3	6	3	48

Table D.1 continued from previous page

Hub	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Total
34	3	0	3	0	0	0	0	0	3	3	12
35	3	0	3	3	3	0	0	0	3	0	15
36	3	3	3	0	3	0	0	3	3	3	21
37	6	0	3	0	3	3	3	3	0	0	21
38	3	3	3	3	3	3	0	3	3	0	24
39	183	3	33	3	138	3	3	6	9	6	387
40	51	3	3	0	21	3	3	6	6	3	99
41	0	0	0	0	0	0	3	3	12	3	21
42	3	0	0	0	3	0	0	0	3	3	12
43	3	0	3	0	0	3	0	0	3	0	12
44	30	3	3	0	39	0	3	0	3	0	81
45	3	0	0	0	0	0	0	0	0	3	6
46	3	0	3	0	3	0	0	0	0	3	12
47	3	0	3	0	6	0	0	3	0	0	15
48	27	3	3	0	24	3	3	3	3	3	72

Table D.1: μ per hub and demand type

D.2 σ values

Hub	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
1	12	3	3	3	21	3	6	9	12	3
2	6	3	3	0	6	0	3	3	3	3
3	3	0	3	0	3	0	3	3	3	3
4	3	0	0	0	3	0	0	0	0	0
5	3	3	3	3	3	0	3	3	3	3
6	3	3	3	0	3	0	0	0	3	3
7	3	0	3	0	6	0	0	3	0	0
8	3	0	3	0	6	0	3	3	3	3
9	0	0	0	0	0	0	3	3	6	3
10	3	3	3	0	3	0	3	3	3	3
11	0	0	0	0	0	0	3	15	15	6
12	3	0	3	0	6	0	3	6	3	3
13	3	0	3	0	3	0	0	0	3	0
14	3	3	3	0	3	3	3	3	3	3
15	3	3	3	3	6	0	0	3	3	3
16	3	0	0	0	3	3	0	0	0	0
17	6	0	3	0	15	0	0	0	3	3
18	3	3	0	0	6	3	0	0	3	0
19	3	3	3	3	3	3	3	3	3	3
20	3	3	3	3	3	3	3	3	3	3
21	3	0	3	0	3	0	3	3	3	3
22	9	3	3	3	6	3	3	3	3	3
23	6	0	3	0	3	0	3	3	3	0
24	3	3	3	3	3	3	3	6	3	3
25	3	3	3	0	3	3	3	3	3	3
26	3	0	3	0	3	0	0	3	3	3
27	0	0	0	0	0	0	3	3	3	0

Table D.2 continued from previous page

Hub	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**	**D7**	**D8**	**D9**	**D10**
28	6	3	3	3	9	3	3	3	3	3
29	3	0	3	0	3	0	3	3	3	3
30	3	0	3	0	3	0	0	3	3	0
31	3	0	3	0	3	3	0	3	3	0
32	3	0	0	0	3	0	0	3	3	3
33	3	3	3	0	3	3	3	3	3	3
34	3	0	3	0	0	0	0	0	3	3
35	3	0	3	3	3	0	0	0	3	0
36	3	3	3	0	3	0	0	3	3	3
37	3	0	3	0	3	3	3	3	3	0
38	3	3	3	3	3	3	0	3	3	0
39	21	3	9	3	30	3	3	6	6	3
40	6	3	3	3	9	3	3	6	3	3
41	0	0	0	0	0	0	3	3	6	3
42	3	0	3	0	3	0	0	3	3	3
43	3	0	3	0	3	3	0	0	3	0
44	9	3	3	0	12	0	3	0	3	0
45	3	0	0	0	0	0	0	0	0	3
46	3	0	3	0	3	0	0	0	3	3
47	3	3	3	0	3	0	0	3	3	0
48	9	3	3	0	6	3	3	3	3	3

Table D.2: σ per hub and demand type

Appendix E

Result robust optimization $S_{i,p}$

E.1 Model 1

City Index	T1	T2	T3	T4	T5	T6	T7
1	11,61	0,00	0,00	228,07	0,00	118,63	5,80
2	0,00	0,00	2,40	29,84	9,00	0,00	0,00
3	0,00	5,83	1,31	0,22	0,00	6,85	0,00
4	0,00	0,00	0,00	0,00	0,00	0,00	0,00
5	9,14	0,00	0,00	27,46	0,00	4,57	0,01
6	5,63	0,00	0,00	19,60	0,00	5,63	0,00
7	0,00	14,26	0,00	0,00	0,00	2,85	0,00
8	0,00	13,17	18,51	39,31	33,61	0,26	0,00
9	0,00	0,00	3,00	36,42	0,00	0,00	0,00
10	4,34	0,00	0,00	28,90	0,00	7,19	0,00
11	4,92	10,94	0,31	57,61	15,09	0,00	0,00
12	0,00	1,18	125,46	114,69	0,00	0,00	0,00
13	0,00	2,49	2,49	5,34	0,00	5,34	0,00
14	5,80	0,00	0,00	40,83	0,00	5,80	5,80
15	4,68	7,46	0,05	0,27	9,37	0,00	0,00
16	0,00	2,59	0,00	0,00	0,00	5,45	5,45
17	0,00	0,15	0,00	35,27	49,89	0,89	0,00
18	5,44	0,00	0,00	13,73	0,00	13,73	5,44
19	0,00	2,29	0,00	0,00	0,00	0,00	0,00
20	11,04	25,79	5,52	19,51	3,00	8,37	5,52
21	0,00	0,00	0,00	28,20	0,00	4,69	0,00
22	10,68	0,00	0,00	60,29	0,00	16,39	5,34
23	0,00	0,00	17,30	21,65	1,63	9,68	0,00
24	0,00	0,00	0,05	0,00	0,00	0,00	0,00
25	24,73	9,62	21,91	76,80	7,58	6,00	17,15
26	0,00	0,00	0,00	23,46	0,00	7,55	0,00
27	0,00	0,00	0,00	15,69	0,00	0,00	0,00
28	9,93	40,58	13,93	30,84	23,12	28,97	4,97
29	0,00	0,00	7,85	25,30	0,00	7,71	0,00
30	0,00	4,25	0,00	0,00	0,00	0,00	0,00
31	0,00	0,00	0,00	20,64	0,00	4,81	4,81
32	0,00	0,00	0,00	21,78	0,00	5,45	0,00
33	4,72	0,78	24,40	8,22	0,00	4,72	0,00

Table E.1 continued from previous page

City Index	T1	T2	T3	T4	T5	T6	T7
34	0,00	0,00	0,00	26,00	0,00	0,00	0,00
35	0,00	0,00	0,00	10,33	0,00	0,00	0,00
36	4,57	0,00	0,00	22,86	0,00	4,57	0,00
37	0,00	0,00	0,00	21,70	0,00	4,34	4,33
38	10,89	0,00	0,00	21,78	0,00	5,45	5,45
39	11,25	0,00	0,00	274,93	158,06	2,29	5,63
40	8,42	64,59	8,61	17,04	25,48	0,74	5,68
41	0,00	0,00	0,00	27,41	0,00	0,00	0,00
42	0,00	1,96	0,00	16,38	0,00	4,81	0,00
43	0,00	0,00	0,00	14,43	0,00	1,96	4,81
44	5,34	0,59	35,99	16,03	44,19	0,59	0,00
45	4,66	0,00	14,82	36,12	7,49	0,00	0,00
46	0,00	0,00	2,30	15,52	0,00	5,17	0,00
47	1,49	0,00	1,49	13,03	7,19	0,00	0,00
48	5,34	0,30	30,29	26,71	28,39	0,00	5,34
Totaal	164,65	208,81	338,02	1590,19	423,10	311,44	91,52

Table E.1: Stock levels model 1

E.2 Model 2

Hub	T1	T2	T3	T4	T5	T6	T7	Total
7	0	12	0	15	15	0	0	42
9	0	6	18	18	0	0	0	42
12	0	81	123	39	0	0	0	243
16	0	0	3	0	0	6	6	15
18	6	6	0	6	0	15	6	39
27	0	0	12	6	0	0	0	18
32	0	6	12	6	0	6	0	30
34	0	0	12	15	0	0	0	27
38	12	9	3	9	6	0	6	45

Table E.2: Stock levels model 2

E.3 Model 3

Hub	T1	T2	T3	T4	T5	T6	T7	Total
1	18,00	132,00	57,00	87,00	0,00	141,00	0,00	435,00
2	8,49	8,49	25,65	36,81	19,83	0,00	0,00	99,27
3	0,00	0,00	23,31	31,92	10,35	0,00	0,00	65,58
4	0,00	8,73	0,00	0,00	8,73	0,00	0,00	17,46
5	18,00	0,00	27,00	27,00	0,00	9,00	0,00	81,00
6	8,73	8,73	8,73	14,55	0,00	8,73	0,00	49,47
7	0,00	0,00	11,40	14,25	14,25	0,00	0,00	39,90
8	0,00	25,11	8,22	25,47	19,32	0,00	0,00	78,12
9	0,00	0,00	0,27	49,41	0,00	0,00	0,00	49,68

Table E.3 continued from previous page

Hub	**T1**	**T2**	**T3**	**T4**	**T5**	**T6**	**T7**	**Total**
10	8,46	8,46	11,28	33,87	11,28	0,00	0,00	73,35
11	0,00	0,15	17,10	57,30	22,35	0,00	0,00	96,90
12	0,00	1,17	0,00	240,15	0,00	0,00	0,00	241,32
13	0,00	5,70	5,70	8,55	0,00	8,55	0,00	28,50
14	9,00	0,00	15,00	45,00	9,00	0,00	0,00	78,00
15	17,11	22,82	0,00	25,67	0,00	17,11	0,00	82,72
16	0,00	0,00	0,00	5,70	0,00	8,55	8,55	22,80
17	0,00	0,00	34,38	17,10	63,63	0,00	0,00	115,11
18	8,46	0,00	19,74	0,00	0,00	19,74	8,46	56,40
19	20,61	0,00	38,28	26,49	8,82	0,00	3,72	97,92
20	17,10	20,40	0,00	53,22	0,15	16,56	0,00	107,43
21	0,00	0,00	8,55	42,78	0,00	8,55	0,00	59,88
22	17,10	0,00	60,33	25,68	22,83	0,00	8,55	134,49
23	0,00	0,00	17,67	37,59	0,12	14,46	0,00	69,84
24	0,00	0,00	0,00	11,40	0,00	0,00	0,00	11,40
25	8,58	25,71	23,43	31,71	17,43	0,00	3,60	110,46
26	0,00	0,00	8,55	34,23	0,00	11,40	0,00	54,18
27	0,00	0,00	17,10	5,70	0,00	0,00	0,00	22,80
28	16,44	36,93	8,22	44,37	56,88	0,00	8,22	171,06
29	0,00	0,00	27,42	24,66	0,00	10,95	0,00	63,03
30	0,00	0,00	0,00	34,23	0,00	8,55	0,00	42,78
31	0,00	8,56	17,11	8,56	0,00	8,56	8,56	51,34
32	0,00	0,00	25,68	8,55	8,55	0,00	0,00	42,78
33	8,55	0,00	11,58	54,21	0,00	14,25	3,60	92,19
34	0,00	9,00	9,00	18,00	0,00	0,00	0,00	36,00
35	0,00	0,00	8,49	8,49	0,00	0,00	0,00	16,98
36	9,00	0,00	9,00	36,00	0,00	9,00	0,00	63,00
37	0,00	0,00	0,00	42,33	8,46	0,00	8,46	59,25
38	17,10	0,00	17,10	17,10	0,00	8,55	8,55	68,40
39	17,46	212,31	20,37	87,24	189,03	0,00	0,00	526,41
40	14,67	0,00	77,10	38,16	0,18	32,28	0,00	162,39
41	0,00	8,40	0,00	39,18	0,00	0,00	0,00	47,58
42	0,00	0,00	14,26	22,82	8,56	0,00	0,00	45,64
43	0,00	0,00	0,00	25,67	5,70	0,00	8,56	39,93
44	8,55	0,00	8,55	63,33	57,06	0,60	0,00	138,09
45	0,00	0,00	0,00	16,98	0,00	0,51	0,00	17,49
46	0,00	8,55	0,00	22,83	0,00	8,55	0,00	39,93
47	5,64	8,46	0,00	22,59	0,00	11,28	0,00	47,97
48	8,55	0,00	39,93	43,08	34,23	0,60	8,55	134,94
Total	265,60	559,67	732,51	1666,93	596,74	377,33	87,37	4286,16

Table E.3: Stock levels model 3

E.4 Model 4

Hub	**T1**	**T2**	**T3**	**T4**	**T5**	**T6**	**T7**	**Total**
7	0	21	0	6	0	15	0	42
9	0	9	27	18	0	0	0	54

Table E.4 continued from previous page

Hub	T1	T2	T3	T4	T5	T6	T7	Total
12	0	81	123	39	0	0	0	243
16	0	0	0	6	0	9	9	24
18	9	9	9	0	21	0	9	57
27	0	0	9	15	0	0	0	24
32	0	0	18	18	0	9	0	45
34	0	0	18	18	0	0	0	36
38	18	0	6	27	3	6	9	69

Table E.4: Stock levels model 4

E.5 Model 5

hub	T1	T2	T3	T4	T5	T6	T7	Total
1	18	0	30	246	141	0	0	435
2	9	9	36	30	21	0	0	105
3	0	9	27	27	12	0	0	75
4	0	0	9	0	0	9	0	18
5	18	18	9	27	9	0	0	81
6	9	9	9	15	9	0	0	51
7	0	9	12	6	15	0	0	42
8	0	0	0	63	21	0	0	84
9	0	0	27	27	0	0	0	54
10	9	0	30	27	12	0	0	78
11	0	0	9	69	24	0	0	102
12	0	0	129	120	0	0	0	249
13	0	6	6	9	9	0	0	30
14	9	0	12	48	9	0	0	78
15	18	24	9	18	18	0	0	87
16	0	0	0	6	9	0	9	24
17	0	27	9	18	66	0	0	120
18	9	0	0	21	21	0	9	60
19	21	0	9	57	9	0	9	105
20	18	15	30	27	15	0	0	105
21	0	9	18	27	0	9	0	63
22	18	9	9	72	24	0	9	141
23	0	9	30	18	15	0	0	72
24	0	0	18	12	0	0	0	30
25	9	9	18	57	0	18	9	120
26	0	18	9	18	12	0	0	57
27	0	0	18	6	0	0	0	24
28	18	0	51	45	0	60	9	183
29	0	0	0	57	0	12	0	69
30	0	9	0	27	9	0	0	45
31	0	18	9	9	9	0	9	54
32	0	9	0	27	0	9	0	45
33	9	0	21	48	0	15	9	102
34	0	0	9	27	0	0	0	36
35	0	9	0	9	0	0	0	18

Table E.5 continued from previous page

hub	**T1**	**T2**	**T3**	**T4**	**T5**	**T6**	**T7**	**Total**
36	9	9	9	27	9	0	0	63
37	0	21	6	18	0	9	9	63
38	18	18	0	18	9	0	9	72
39	18	219	39	72	195	0	0	543
40	15	18	0	99	0	33	0	165
41	0	0	33	18	0	0	0	51
42	0	15	9	15	9	0	0	48
43	0	0	18	9	6	0	9	42
44	9	0	57	18	60	0	0	144
45	0	0	9	9	0	0	0	18
46	0	9	0	24	0	9	0	42
47	6	9	15	9	12	0	0	51
48	9	9	9	69	36	0	9	141
Total	276	552	816	1725	825	183	108	4485

Table E.5: Stock levels model 5

Appendix F

Redistribution MILP

F.1 Redistribution tank containers MILP formulation

Problem description
After the occurrence of the demand, the safety stock levels need to be replenished again and tank containers need possibly to be rerouted back again. The model in equation F.1 minimizes the costs of repositioning these tank containers. For doing so, only one new variable needs to be introduced the $IP_{i,p}$. The $IP_{i,p}$ represent the inventory position at hub i of tank type p. Since only one period is considered, the solution does not represent any uncertainty anymore and the problem can be solved with an MILP. The $S_{i,p}$ is already solved in the safety stock determination and the $IP_{i,p}$ account for the previous incoming transshipments. For the cost in the problem, only the transportation costs are important, since it is only determining over which lanes repositioning is necessary.

Mathematical problem formulation
The redistribution assignment of tank containers is in model F.1.

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{c\in\mathcal{C}}\sum_{p\in\mathcal{P}} \left(Q_{i,j,c,p} * FT_{i,j} * C^{TR}_{i,j}\right) \\
\text{subject to} \quad & IP_{i,p} + \sum_{c\in\mathcal{C}} \left(\sum_{i\in\mathcal{I}} (Q_{i,j,c,p} * F^{TR}_{i,j}) - \sum_{j\in\mathcal{J}} (Q_{j,i,c,p} * F^{TR}_{i,j}) \right) \geq S_{i,p} \qquad \forall i \in \mathcal{I},\ p \in P, \\
& \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{c\in\mathcal{C}}\sum_{p\in\mathcal{P}} Q_{i,j,c,p} \leq \frac{1}{2} * \sum_{i\in\mathcal{I}}\sum_{p\in\mathcal{P}} |S_{i,p} - IP_{i,p}|, \\
& Q_{i,j,p} \in \mathbb{Z}^+
\end{aligned}
\tag{F.1}
$$

The first constraint the leftover inventory of hub i of tank type p, plus the corresponding repositioning of tank containers towards i of type p, minus everything that hub i sends away of type p, should be at least equal to its safety stock level $S_{i,p}$.

The second constraint ensures that every tank container only gets repositioned once. This is to avoid, sending tank container from hub A towards B and then from hub B to hub C, instead of sending it directly from hub A to hub C. Unfortunately, this is sometimes more attractive due to inaccuracies in the provided cost input data. However, in reality, this is not the case for DH. Therefore, the last constraint restricts the amount of movements of tank containers to only what is necessary.

F.2 Results tables redistribution MILP

Hub	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	9	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0
5	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	3	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0
7	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0
16	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0
44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table F.1: First part redistribution of safety stock levels

Hub	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	21	0	0	0	27
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0
36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	24	0	0	0	0	0	0	0	0	0
42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table F.2: Second part redistribution of safety stock levels

F.3 Interpretation results redistribution MILP

The redistribution model presented in Equation F.1 solves the instance after the occurrence of demand. This model provides assistance in decision making such that the the corresponding safety stock levels have been replenished to their initial level $S_{i,p}$ while minimizing transportation costs.

The model is augmented to function with the future inventory position, which is modelled with the inventory variable $IP_{i,p}$ and tracks which repositioning actions are on time with the feasibility $FT_{i,j}$ variable.

However, since we do not run a complete simulation there is no known instance after demand occurrence. Since there are many scenarios possible in the uncertainty set, an illustrative example is used. It is decided to use the difference of the mean import and mean export balance of all tank container types for all hubs. This difference between mean import and mean export is called mean difference, which is shown in Appendix B. In that Appendix, the corresponding standard deviation is also shown to indicate how uncertain the difference between import and export of a hub behaves.

The results of the transportation matrix are in Appendix F. Again, it is supported that almost all hubs with a surplus has to reposition tank containers towards hub 1 or hub 39. This pattern was also visible in the mean difference analysis of flow of demand. However the transshipment matrix in Appendix F, also indicates which hubs have to transfer back to hub 1 and hub 39.

For this MILP, it is necessary to have the complete network available since tank containers also need to float between the mentioned sub-regions. But since this network is not uncertain anymore, it reduces the complexity massively, and the whole network is easily analyzed even with integrality.