MASTER

Optimal design of a multimodal distribution network for a chemical company

van den Broek, E.C.M.

Award date:
2020

Department of Industrial Engineering & Innovation Sciences
Master Operations Management and Logistics

Version date: May 1, 2020

Optimal design of a multimodal distribution network for a chemical company

Master Thesis

Eline C.M. van den Broek

First supervisor: dr. ir. R.A.C.M. Broekmeulen
Second supervisor: dr. A.E. Akçay

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

Keywords: Network Optimization, Distribution network (re-)design, Gravity study, Multi-modal Facility Location Problem, Truck-loading problem, Nonlinear programming, Chemical industry.

CONFIDENTIALITY

The research presented in this thesis is performed for a Multinational Chemical company. The research is based on actual costs and volume figures related to the distribution network of this company. However, for confidentially reasons, the name of the company cannot be disclosed, and it is referred to as “the Company”.

In addition to the public version of the Thesis, there are two confidential appendices that will not be published. These appendices will not be referred to in the text and include the following topics:

- Confidential Appendix A: Map of the current warehouse locations;
- Confidential Appendix B: Map of the manufacturing sites in scope.

Abstract

Motivated by the current globalization, it is beneficial for companies to have a proper distribution network that can take advantages of multi-modality. Therefore, this thesis presents a methodology to determine the optimal (re)design for a multimodal facility location problem such that the total distribution network costs are minimized. The methodology focuses on consolidating the logistic activities to one centralized distribution center from where goods are transported to the customer, which can be done directly or through a cross docking facility. In order to find the optimal locations for these facilities, two gravity studies are proposed. At first, a single-facility gravity study is presented to investigate the optimal distribution center location by minimizing the total transportation costs. Secondly, a multi-facility gravity study is presented to determine the optimal cross docking facility locations by minimizing the outbound transportation costs (from facility to customer). By using the solutions of the gravity studies, a set of possible facilities with multi-modal connection is defined. In order to determine the optimal (re)design, a multi-modal facility location problem is developed with the aim to minimize the total distribution network costs by deciding on which facilities to open and which transportation modes to use. The solution of this model depends on a given customer service lead-time.

Executive summary

Because of the current globalization, companies have to compete on regional, national, and international level. In order to compete against global competitors, it is beneficial to have a proper distribution system which can take advantages of multi-modality. Therefore, effective and efficient transportation of goods through the supply chain is crucial in the current global competition. In other words, transportation is a key part of a supply chain. Different companies experienced these globalization problems, which is also the case for a Chemical Company operating in EMEA (Europe, Middle East & Africa). The Company is a global advanced material and specialty additives company and produces a broad range of everyday use items.

The problem for the Company occurs at the logistic department with focus on the packed warehouses. In the current situation, there are several different warehouses, which are operated by different suppliers. It regards eight outsourced packed and re-packing warehouse locations and one packed warehouse operated by the Company. The multiple warehouses are the result of acquisitions with different businesses. This situation results in a complex distribution network accompanied by several inefficiencies. Therefore, the Company initiated a warehouse consolidation project with the aim to reduce costs, improve productivity and quality, reduce complexity, and reduce the CO_2 footprint of the distribution network. These objectives request a redesign of the distribution network and results in the following main research question: *What is the optimal design for the multimodal distribution network of the Company?*
In order to answer this research question, the research is structured according to the following three features: (1) analyzing the current distribution network, (2) developing a methodology to determine the optimal design of the multi-modal facility location problem such that costs are minimized, and (3) analyzing the implementation sensitivity due to different future scenarios.

At first, the current distribution network is analyzed. As discussed before, the distribution network regards nine warehouses (six packed and three re-packing warehouses). In the packed warehouses, goods are stored, whereas in the re-packing warehouses, value-adding services, such as re-packing and re-labelling, are performed. The distribution network consists of three stages: the inbound process, the warehousing process and the outbound process. The inbound process regards the transportation of goods from manufacturing site to the warehouse, the warehousing process regards the handling of goods and storage of inventory, and the outbound process regards the transportation of goods from warehouse to customer. The inbound streams are coming from twenty manufacturing sites that are globally located. For the manufacturing sites outside Europe (57 per cent), goods are shipped to the port of Rotterdam or Antwerp and from there shuttled by truck to a warehouse. It is decided to leave the shipping process out of scope, therefore only the shuttling process from unloading port to the warehouse is considered in the research. For the manufacturing sites within Europe (43 per cent), goods are transported by truck to a warehouse. When a good requires value adding services, it is first served by the re-packing warehouse, where after it is stored in a packed warehouse. Goods are stored in a warehouse until a customer order is placed. When goods are ordered, the outbound process starts. In total there are 1,345 customers served, from which the majority are located within Europe (82 per cent). For customers within Europe, the goods are transported by truck to the

customer. For customers outside Europe (18 per cent), goods are shuttled to the loading port and from there shipped to the customer, therefore only the shuttling process is in scope.
Based on analysis, it is concluded that the total annual throughput is 198,000 metric ton (MT) and the total average daily inventory is 44,000 MT. The demand (and therefore the throughput) is concluded to be stable over time. According to Little's Law, the inventory level is related to the throughput by its turnover rate, which is 4.4 times a year. Due to this relation, it can be concluded that the inventory level is stable as well. The annual transportation costs of the inbound process amount $3.5 million USD, whereas the annual transportation costs of the outbound process amount $14.7 million USD. Furthermore, there are warehousing costs that consist of handling and administration costs (function of throughput) and inventory holding costs (function of inventory level), which amount annually $6.6 million USD. Consequently, the project regards a distribution network with a total cost of $24.8 million USD annually. According to the analysis, the transportation rates are depending on the truckload and the travel distance following an increasing concave down cost structure. Therefore, the distribution network could be improved by optimizing the truckload and/or travel distances.

Secondly, a methodology is developed to determine the optimal design of the multi-modal facility location problem such that the distribution network costs are minimized. In order to minimize the costs, the function of the costs requires research. Based on analysis, it is concluded that the transportation rates are concave related to the truckload and distances. According to these two variables, the function for the transportation rate is structured. (1) The full-truckload (FTL) tariff is determined while considering distance as a concave cost structure following a polynomial distribution (order 2) where the second order coefficient is negative. (2) The less than truckload (LTL) tariff is defined as a function of this FTL tariff and a determined shape parameter for the truckload. This LTL tariff is multiplied with the annual number of movements in order to calculate the annual costs per stream. Based on this equation, the costs are modelled. The ideal situation proposes a consolidation of the warehouses to one distribution center (DC), which should be optimally located relative to the manufacturing sites and customers. Therefore, a gravity study is performed to minimize the total transportation costs by deciding on a set of coordinates for the center of gravity. The problem is defined as a Nonlinear Program due to the increasing concave down cost structures and it regards a single-facility problem against current performances (truckload and number of movements) where the transportation is only performed by truck. Solving the problem results in a gravity center at Rotterdam Port. This gravity center serves as the DC of the network, i.e. the location where the Company concentrated the logistics of goods. From the DC, goods are transported to the customers, which can be done directly or through a cross docking facility (CDF). The CDFs are deployed as a location where products are moved from one transportation mode to another, with minimal warehousing. In order to find the optimal location of the CDF(s), the single-facility gravity problem is transformed into a multiple-facility gravity model, with the following adjustments: deciding on multiple gravity points (facilities) while minimizing the outbound costs from the facilities to the allocated customers. The solution of the multiple-facility gravity model is as follows: a second facility (CDF 1) shows a gravity point in Slovakia or Poland, a third facility (CDF 2) in North Italy, and a fourth facility (CDF 3) in Romania. The advantage of a CDF is the opportunity to combine transportation from DC to CDF and to use different transportation modes that are probably cheaper. Therefore, a multi-modal facility location problem is developed which decides on the facilities to open and transportation modes to use by minimizing the total distribution network costs (i.e. transport and inventory costs) defined as a Mixed Integer Program. Based on the solution of the multiple-facility gravity study, it is investigated which possible locations within a given gravity area (5 per cent error range) are connected to a multi-modal network: road, rail and inland waterways. These defined locations compose a set of possible facilities to open and possible transportation modes to use for which the model select the optimal combination by assuming performances depending on a given customer service lead-time. This customer service lead time is the lead-time between the ordering moment and delivery, which equals the transportation process from DC to a customer. According to various studies, it is known that in general the cheaper the transportation mode, the longer the transportation lead-time. Therefore, the model can only select a cheaper transportation mode when the lead-time allows this.

Hence, different customer service lead-times result in different solutions, which is related to the last feature of the research question. Solving the multi-modal facility location problem for a customer service lead-time of one till ten days (scenarios), shows that the biggest improvement opportunity: annual distribution network cost reduction of $4.5 million USD, can be achieved when the Company delivers against a customer service lead-time of (at least) five days. In this situation, the logistic of goods is consolidated to one DC, located in Rotterdam, and three cross docking facilities are opened, located in Wroclaw in Poland (CDF1), Milan in Italy (CDF2), and Ploiesti in Romania (CDF3). Due to these opened CDFs, orders of different customers can be combined, and the lead-time allows to use the cheapest connected transportation mode (barge between DC and CDF1, and train between DC and CDF2 and CDF3) for the major transportation part. Another scenario is delivering against a customer service lead-time of one or two days, wherefore the logistics are consolidated to one DC and none of the CDFs are opened. This scenario already results in an annual potential cost reduction of $1.2 or $1.8 million USD (for customer service lead-time of one or two days respectively). When extending the customer service lead-time to three days it is possible to open CDF1 (Wroclaw) and transport by train, with an annual potential cost reduction of $2.8 million USD. Further extending of the customer service lead-time to four days, in addition to CDF1, CDF2 (Milan) can be opened, in which transportations to both CDFs are performed by train. This scenario results in a potential cost reduction of $3.5 million USD per year.

Furthermore, a side-aim of the project is to reduce the CO_2 footprint wherefore the distribution network becomes more sustainable. Therefore, the optimal sustainable design is determined by using the multi-modal facility location problem where the costs are adjusted to Greenhouse Gas Emissions (GHG) such that the environmental impact of the distribution network is minimized. While comparing the solutions of the cost optimal and the sustainable optimal design, the main difference is the use of train or barge in which train transport is the most sustainable and barge is the cheapest transportation mode. However, the performance differences are minimal: for the cost optimal distribution network there is an opportunity to reduce the GHG emissions by 25 per cent and the costs by 20 per cent, whereas for the sustainable optimal distribution network there is an opportunity to reduce the GHG emissions by 28 per cent and the costs by 19 per cent. Hence, for both redesigns, cost optimal and sustainable optimal, it is possible to become more efficient and more sustainable.

Preface

This master thesis is the graduation project of the Master Operations Management and Logistics at Eindhoven University of Technology and is performed at a Chemical Company. The master thesis project would not have been possible without the help of everyone that supported me. Therefore, I would like to take some words to thank them.

First of all, I would like to thank my supervisor, Rob Broekmeulen, for his support. Our conversations gave me new insights that always motivated me to keep up the progress. Your critical view and guidance enabled me to improve the quality of my work and to finalize my thesis. In addition, I would like to thank my second supervisor, Alp Akçay, for your useful feedback and for taking the time to evaluate my thesis.

Moreover, I would like to thank my colleagues for helping me in achieving the project goals. Especially, I would like to thank my company supervisor for the good collaboration. Without your help, I would not have been able to collect all the data. Furthermore, I would like to thank the Logistics Director EMEA, for your support during the project and the guidance in making the results relevant for the Company. I am grateful to have been part of the logistics team and to be surrounded by such friendly and inspiring people.

Finally, I would like to thank my family and friends, who always supported me during my study and in the choices I made. You were there to listen to me when needed and to express your confidence in me. In particular when switching from HBO to WO, you gave me the faith that I was able to do that. Thank you for your unconditional love and support for my ambitions during the last seven years. Without your support, it would not have been possible to reach this.

Eline van den Broek, May 2020

Contents

List of Tables

List of Figures

Chapter 1. Introduction

Because of the current globalization, companies have to compete on regional, national, and international level. In order to compete against these global competitors, it is beneficial for companies to have a proper distribution network that can take advantages of multi-modality. Therefore, effective and efficient transportation of goods through the supply chain is crucial in the current global competition. In other words, transportation is a key part of a supply chain.

Different companies experienced these globalization problems in their transportation process, which is also the case for a Multinational Chemical Company operating in EMEA. In the current situation, there are several different packed warehouses due to different acquisitions in the past. It seems likely that the current distribution network is not optimal due to complexity and inefficiencies. In order to reduce costs and become a stronger competitor it is important to improve this distribution network. Therefore, the focus of this thesis lies on a redesign of the multimodal distribution network of this Company such that the costs are minimized. Hence, in this chapter the motivation for requesting this project is described (Section 1) followed by a review of existing literature on distribution network designing (Section 2).

1.1 Motivation

The motivation of this research is mainly driven by the complexity and inefficiency of the current distribution network of the Company. The complexity is expressed in the number of manufacturing sites, the number of suppliers for transportation, number of suppliers for storing, number of customers, number of different transportation flows and the multimodal network. Currently, the logistic network of the Company is served by different suppliers for transporting and storing the goods, which results in a complex situation accompanied with different inefficiencies and high costs for the logistic department. In order to improve this situation, the Company initiated a consolidation project to consolidate the warehouses to one or two remaining hubs that serve as the main gateway for Europe, with the aim to:

1. Reduce warehousing costs, i.e. storage and handling costs by combining volumes;
2. Reduce transport costs due to less shipping points and shuttling;
3. Reduce complexity by reorganizing the network;
4. Improve productivity and quality by making digitalization a hard requirement;
5. Reduce the CO_2 footprint due to improved loading, reduced transportation, locations with multi-modality, and selecting sustainable suppliers.

These objectives request a redesign of the distribution network in which the goal is to find the optimal design for the transportation and warehousing process.

1.2 Literature Study

In this section it is investigated what existing literature is already available on the main subject of the Master Thesis: *distribution network (re)design*. According to (Simchi-Levi, Chen, & Bramel, 2005), a network design is useful when the current set of facilities is deemed to be inappropriate. Therefore, companies want to reorganize or redesign the distribution network in such a way that inventory and transportation costs are minimized, and various service level requirements are satisfied. The distribution network design contains decisions on different levels: decisions about how to improve supply chain performances (section 1.3.1), decisions about where to locate facilities and how to allocate them to manufacturing sites and customers (section 1.3.2). When the optimal design is determined, it is important to consider different tendering strategies in order to select the right supplier (section 1.3.3).

1.2.1 Mixed truckload

The truckload is a crucial performance in reducing transportation costs and environmental impact of transport. (Yuceer & Ozakca, 2010) researched a weighted distribution problem, in which the operational problem is how to load the truck in such a way that the replenishment time of a certain route is maximized. The research developed an efficient heuristic procedure with a main algorithm and two sub algorithms for solving the truck loading problem. Furthermore, (Simchi-Levi, Chen, & Bramel, 2005) researched the shipper problem with the objective to find a production plan, an inventory policy and a routing strategy in order to minimize the total cost and satisfy total demand. In this model, one of the cost components regards truckload, which is assumed to be piece-wise linear. In addition, (Santen, 2017) developed a framework that contributes to a richer understanding of opportunities to increase the loading factor by measuring load factor with a combination of indicators at three levels: packaging, loading and booking level. Beside these methods, another possibility is to exchange transport-capacity between logistic companies. Therefore, the spatial logistic concentration of an area is an important aspect, which effect is researched by (van den Heuvel et al., 2012). The main advantage of a facility location within the concentration is that logistic companies more often exchange their transport-capacity, which results in higher loading factors and lower costs. The researcher developed a mathematical formulation, which made a consideration between the traditional optimal location and the location in concentration area. Based on the reviews, an interesting topic for future research is how to include the truck-load performance in a facility location problem with the objective to minimize the total distribution network costs, i.e. transportation and inventory costs.

1.2.2 Facility location problem

The facility location problem is a branch of operations research with the aim to minimize transportation costs by deciding the optimal placement of facilities. Various studies researched the facility location problem, which results in different variants of this problem. At first, some single-echelon models are reviewed, i.e. models that consider the in- or outbound process. (Ghiani, Laporte, & Musmanno, 2004) presented the Single-Echelon Single-Commodity Location (SESC): the problem of opening a set of facilities and assigning each customer to exactly one open facility while respecting its capacity. Types of single-echelon models that are presented by (Ghiani, Laporte, & Musmanno, 2004) are the p-Centre model, the Location-covering model and the p-Median model. Furthermore, a more complicated single-echelon model is the Single-Source Capacitated Facility Location Problem. In this situation the number of warehouses to locate is not fixed beforehand and it is extended with capacity constraints. Secondly, some multi-echelon models are reviewed, i.e. models that consider the in- and outbound process. (Ghiani, Laporte, & Musmanno, 2004) presents a Two-Echelon Multi-Commodity Location problem (TEMC), in which the model includes inbound and outbound material flow and considered a multi-commodity network. According to (Simchi-Levi, Chen, & Bramel, 2005), the Distribution System Design Problem considers the total network. There are several manufacturing sites that serve the warehouses, and several customers that are served by the warehouse. It is the aim to locate a set of facilities in this distribution network such that the total distribution costs are minimized. In the current literature, this model handles several extensions, such as the location-inventory problem proposed by (Shen, Coullard, & Daskin, 2003), the location problem including disruption presented by (Ahmadi-Javid & Seddighi, 2013), and the location-routing problem developed by (Laporte & Nobert, 1981). More recent studies focus on an extension of the facility location problem by combining it with multimodal transportation networks. Multimodality is the movement of materials from origin to destination by several transportation modes, in which each of these modes have different suppliers (Crainic, 2003). The study of (Fazayeli, Eydi, & Kamalabadi, 2018) integrates multimodal transportation and location-routing problem. The transportation network connecting a supplier to potential DCs supposed to be a multimodal network including three transportation modes: road, rail, and seaways (multimodal), whereas the transportation from DCs to retailers is performed by road (one mode). Furthermore, the study of (Ambrosino & Sciomachen, 2016) faces a capacitated hub location problem on a weighted multimodal network for sea, rail and road. (Tuzkaya, Onut, & Tuzkaya, 2014) presented distribution facilities location problem in a multimodal network. Despite the various developments, the multimodal facility location problem is still under development by researchers since there are suggestions for future research. One of the suggestions is to research the facility location problem considering uncertainties within data.

1.2.3 Tender strategies

Since the distribution network activities of the Company are outsourced, the implementation of the redesign of the distribution network requires a tendering process in order to select the right supplier. In this tendering process, the defined optimal distribution network should be considered as benchmark. Therefore, it is important to evaluate different tender strategies. According to (Ivanov, Tsipoulanidis, & Schonberger, 2017), sourcing strategies can be classified according to three basic features: (1) number of suppliers, (2) geographic, and (3) the shared principles. The number of suppliers depends on the risk that a company faces when relying on one supplier. The geographic sourcing strategy depends on whether to select a supplier on local or global level. In addition, it is important to select a partner on shared principles. One of these principles for the Company is sustainability as stated in their sustainability strategy. According to (Ashnani, Miremadi, Johari, & Danekar, 2015), the transport sector is responsible for 23 per cent of the total global energy-related emissions. Due to this fact, several studies researched sustainable networks. At first, (Govindan, Jafarian, Khodaverdi, & Devika, 2014) proposes a linear complementarity problem in a sustainable network for the vegetable markets. Furthermore, (Najjartabar-Bisheh, Delavari, & Malmir, 2017) analyzes the role of third-party companies in a sustainable supply chain design including location, allocation, and inventory decisions. The study of (Aziziankohan, Jolai, Khalilzadeh, Soltani, & Tavakkoli-Moghaddam, 2017), investigates the effect of Queuing Theory in optimizing energy consumption. Furthermore, (Park, Shin, Chang, & Park, 2010) presented a framework to evaluate and improve a sustainable relationship with suppliers. In the reviewed literature it is noticeable that sourcing decisions are made based on qualitative criteria. Further research should investigate how to change these qualitative selection aspects into quantitative criteria in order to make it possible to include it in the facility location problem.

Chapter 2. Research design

In this chapter the aspects of the research design are described. The chapter starts with an explanation of the problem in more detail (Section 1) followed by the research objective (Section 2). This objective is achieved by answering the defined research questions (Section 3), within a certain project scope (Section 4). In order to answer these questions, a research methodology is implemented that structures the thesis (Section 5).

2.1 Problem Statement

The problem occurs at the logistic department of a Chemical Company operating in EMEA. In the current situation, there are several different packed warehouses, which are operated by different suppliers (third party). It regards eight outsourced packed and re-packing warehouse locations and one packed warehouse operated by the Company. The warehouses are located in the Antwerp-Rotterdam area within a radius of 75km. Due to this current situation, the distribution network becomes complex. The complexity expresses in the number of plants, number of suppliers for transportation, number of suppliers for storing, number of customers, number of different transportation flows and the multimodal network. Based on an analysis of the current situation it seems likely that the complexity results in the following inefficiencies:

- For the inbound process, the same manufacturing sites serve different warehouses, therefore the transportation of goods cannot be combined;
- A part of the containers unloaded at port Antwerp, are shuttled to Rotterdam (port);
- There are different warehouses that store the same products, which might result in higher inventory due to multiple safety stocks and in multiple required movements to one customer for the same product;
- For the outbound process, different warehouses serve the same customers, but the transportation cannot be combined due to the different locations of its origin;
- The outbound truck loading factor is low (comparing to the inbound loading factor), which results in higher transportation costs and emissions per product unit;
- The different contracts with the multiple suppliers result in an unclear overview of the total costs due to different definitions and pricing units;
- The complex distribution network results in the storage of data in different sources, which are contradictory and results in incomplete and incorrect data.

2.2 Research Objective

The inefficiencies due to complexity, as mentioned in the problem statement (Section 2.1), result in high costs for the logistic department. Therefore, the Company wants to improve these activities and initiated a warehouse consolidation project. The goal of this project is to find the optimal design for their distribution network with one or two remaining hubs that serve as the main gateway for Europe. Decisions made in this project have a long-lasting effect on the company. Therefore, it regards decisions on Strategic Level, i.e. decisions regarding the number, location and capacities of warehouses (Simchi-Levi, Chen, & Bramel, 2005). Furthermore, it is important that this network is multimodal due to global activities.

The main research objective is to provide an optimal design of the distribution network of the Company while considering two stages: multimodal network and different scenarios.

Optimal is defined as the situation in which the total costs[1] of the distribution network are minimized while considering different transportation modes and scenarios.

In the first stage, the aim is to propose a mathematical approach to solve the multimodal facility location problem. The aim of the second stage is analyzing the effect of different scenarios (e.g. effect of the customer service level and network sustainability).

2.3 Research Questions

In order to achieve the research objective, some research questions are defined. The main research question of the Master Thesis is as follows:

What is the optimal design for the multimodal distribution network of the Company?

Where, optimal is defined as stated in Section 2.2, multimodal is the movement of materials from origin to destination by several modes of transport (road, rail, water), and the distribution network regards the transportation from manufacturing site to the customer through a warehouse including handling and storage activities.

The main research question is divided into sub-research questions. This splitting is based on the IST-SOLL-GAP method that considered three phases (Welge & Al-Laham, 2007). At first the current situation is defined (IST phase), followed by the ideal situation (SOLL phase). In the GAP phase, it is the aim to define the gaps between the current and ideal situation. Based on a gap-analysis, a solution is proposed that fits the ideal situation best in reality.

In order to analyse the current situation, it is important to define the current distribution network. The aim is to make an overview of the current distribution streams and the network-wide costs. This results in the first sub-research question.

Sub-research question 1. What is the current distribution network design?

The ideal situation will be defined by proposing a mathematical approach that determines the optimal design for the multimodal distribution network. The aim of this approach is to determine the center of gravity of the distribution network, i.e. the optimal location region, with the objective to select a facility that minimize the distribution network costs while considering different transportation modes. This results in the following, second sub-research question.

Sub-research question 2. What methodology is needed that determines the optimal design of a multimodal facility location problem such that costs are minimized?

When the optimal design is proposed, it is important to investigate the differences between the current and ideal situation and how to change to the ideal situation. Therefore, different scenarios regarding the optimal design are considered. This scenario analysis investigates the

[1] Since the distribution network of the Company is served by suppliers, the costs regard the prices that supplier(s) offer.

effect of the customer service level and the sustainability of the network. This results in the third and last sub-research question.

Sub-research question 3. How should the Company implement the distribution network redesign based on the optimal design methodology considering different scenarios?

2.4 Research Scope

In this section the scope is defined, i.e. the range of the project. The project regards nine packed and re-packing warehouses located in the Antwerp-Rotterdam area. The considered process is the distribution of packed goods through the in-scope warehouses. Therefore, the project regards a multi-echelon process: from manufacturing site to the warehouse and from there to the customer. Furthermore, goods that are produced outside Europe (overseas countries), are shipped by seaways to the harbor of Rotterdam or Antwerp. The costs of this process consist of three steps: (1) the shuttling from manufacturing plant to the loading port, (2) the shipment from loading to unloading port, and (3) the shuttling from the unloading port to the warehouse. Since there is not a significant difference in costs (1) and (2), these shipping activities are excluded from the project. Hence, for manufacturing sites outside Europe, the inbound stream included in the project starts from the unloading port and therefore only includes the shuttling costs (3) to the warehouses. The same applies for outbound streams to customers located outside Europe, i.e. only considering shuttling costs to the loading port.

2.5 Methodology and Thesis outline

Different models are used to define the research design of the project. At first, based on the IST-SOLL-GAP method, the research questions are defined. These research questions give the main path for the project design. The research questions are answered by using Network planning to investigate the current and ideal distribution network. While determining the ideal situation a mathematical approach is used. For the developing of the mathematical approach, quantitative research is required. In order to structure this quantitative research, the research methodology of (Mitroff, Raturi, Amoako-Gyampah, & Kaplan, 1974) is implemented. This model gives an overview of quantitative model-based research in operations management in four stages: conceptualization, modelling, model solving, and implementation (see Figure 1).

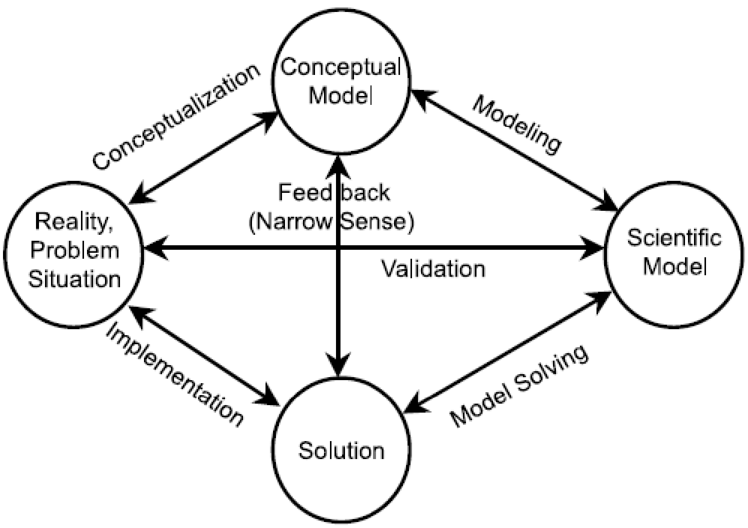

Figure 1: Research model of (Mitroff, Raturi, Amoako-Gyampah, & Kaplan, 1974)

According to pre-research of the problem, it seems likely that one centralized distribution center (DC) is optimal due to the concentrated inbound of goods and the inefficiencies accompanied with several warehouses in close proximity to each other. In addition, a centralized distribution center seems likely to result in bigger improvement opportunities for the outbound process, which is responsible for the main part (81 per cent) of the transportation costs. Furthermore, considering multiple distribution centers decreases the quality of the research and unnecessary increases the complexity since there is no reliable data that makes it possible to link inbound and outbound streams. Therefore, during the definition of the proposal, it is decided to perform the quantitative research by a step-by-step approach consisting of three mathematical models (as shown in Figure 2). At first, a gravity study is presented with the aim to decide on the optimal location for the DC while minimizing the total transportation costs. From this DC goods are transported to the customers, in which cross docking facilities (CDF) might improve the performances of the outbound process. Therefore, a follow up on the single-facility gravity study is the multi-facility gravity study with the aim to find the optimal locations for potential cross docking facilities (CDF) while minimizing the outbound transportation costs. Based on the solutions of the gravity study and the multi-modal connections of locations, a set of potential facility locations is defined. At last, a multi-modal facility location is developed that decides on the optimal combination of potential facilities to open and transportation modes to use subject to a customer service lead-time constraint. This solution serves as a benchmark for the Company, which can be considered and further researched during the tendering process.

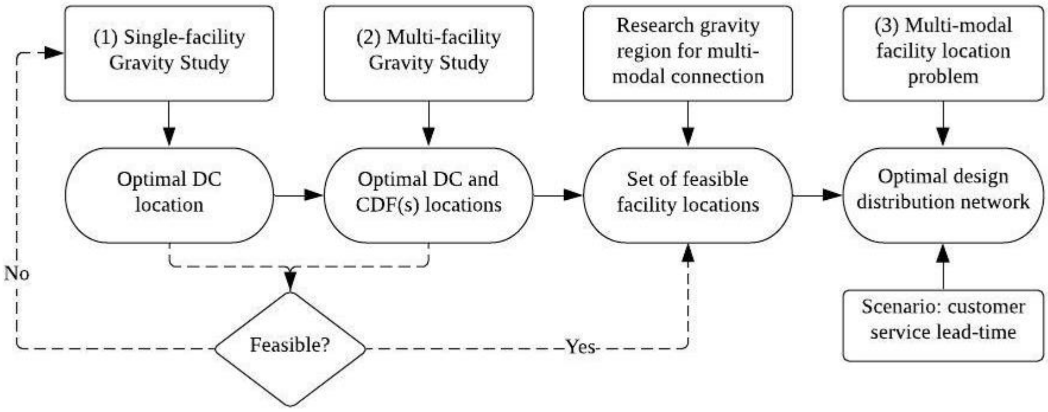

Figure 2. Step-by-step approach of quantitative research

According to the research model, the Master Thesis is structured. In order to improve the current situation, Chapter 3 focuses on defining the performances of the current distribution network. This definition is performed by collecting data and validate the collected data. In Chapter 4, the mathematical model is conceptualized, which is a pre-processing phase in which the model including decisions and assumptions are explained. The first step of the scientific model is to define the cost structure of the transportation activities, which is described in Chapter 5. Based on this cost structure, Chapter 6 presents a single-facility and multi-facility Gravity study. In chapter 7, multi-modality is added as an important aspect resulting in a Multi-modal Facility Location Problem. The aim of Chapter 7 is to propose the optimal distribution network design, in which the solution is depending on the customer service level (scenario). Furthermore, Chapter 8 focuses on the effect of sustainability on the optimal distribution network design. At last, the thesis is concluded in Chapter 9 by summarizing the main results and discussing the research limitations and suggestions for future research.

Chapter 3. Current distribution network

In this chapter the current distribution network is presented, which visualizes the problem. Therefore, it is the aim to answer Sub-research question 1: *What is the current distribution network design?*

The project focuses on the distribution network of packed goods for a Chemical Company operating in EMEA. The core business of the Company is to produces chemical products and therefore the distribution activities are not a core competence in their business plan. Hence, the majority of the transportation and storage of goods is outsourced to different suppliers (third party). These outsourcing activities are performed by six packed warehouses and three re-packing warehouses located in the Antwerp-Rotterdam area within a radius of 75 km. Five out of the six packed-warehouses are outsourced, and one is operated by the Company. The warehouses are part of the distribution network of the Company. A visualization of this distribution network design is shown in Figure 3.

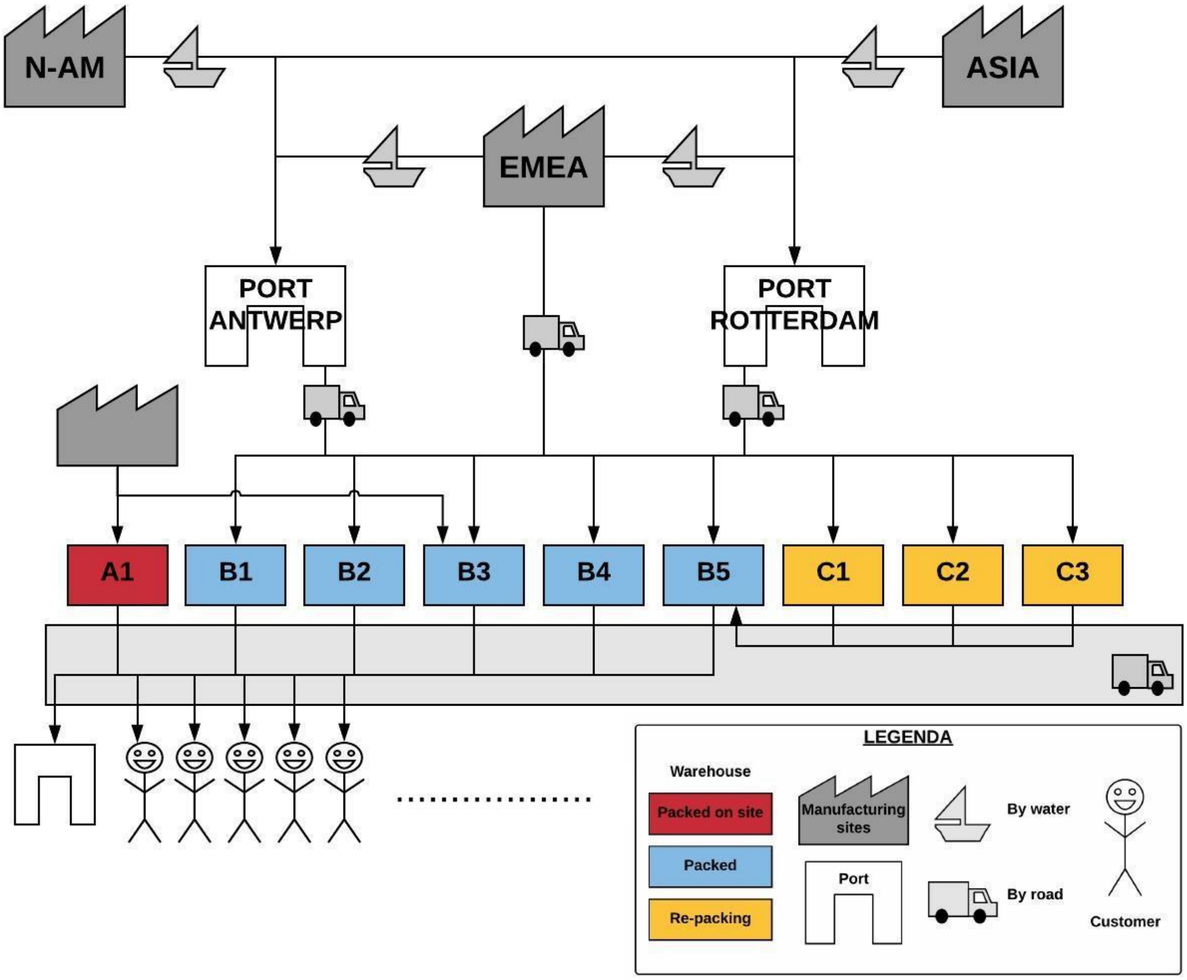

Figure 3: The Current Distribution Network

As can be seen, the network consists of three main stages: the inbound process, the warehousing process and the outbound process. The inbound process regards the transportation of the finished goods, produced in North America (N-AM), Asia and EMEA, to one of the warehouses. Since a significant part of the products originates from overseas countries: N-AM, Asia, Africa and Middle East, the warehouses are located nearby a harbor. The products are shipped per container, unloaded at the port of Rotterdam or Antwerp and from there shuttled by trucks to one of the warehouses. The other part of the products is produced in Europe and transported by truck to the warehouse. Before storing, some products are first served by the re-packing warehouses and after this, moved to a packed warehouse. In the warehouses, a First-In First-Out (FIFO) inventory policy is enforced. When products are ordered, the outbound process: from warehouse to customer, starts. Since the customers are companies, it is regarded as a Business to Business (B2B) process. The majority of the customers are located within Europe, for which the total outbound process is included and accomplished by truck. For customers outside Europe, only the shuttling process from warehouse to the loading port is considered, which is also performed by truck.

Hence, it is concluded that the distribution network consists of mainly two type of activities: transportation and warehousing. For the transportation it regards in- and outbound streams, whereas warehousing activities regard the handling and storing of goods. The structure of this chapter follows these two types of activities: transportation streams (Section 1) and warehouse process (Section 2). Both sections define an overview of the current distribution network; streams, quantities and costs. Beside the transportation and inventory activities, there are order processing activities (Section 3). Furthermore, the analyzed data is validated (Section 4).

3.1 Transportation activities

In order to create an overview of the annual transportation activities, the goods streams during the last year (October 2018 till September 2019) are analyzed by using an internal data source. The transportation activities are divided into inbound and outbound streams. Furthermore, the transportation rates and the distribution of the demand over time are studied.

3.1.1 Inbound streams

The inbound streams are coming from 20 manufacturing sites that are globally located. The majority of the goods (57 per cent) are produced at countries outside Europe, followed by The Netherlands (27 per cent), and Germany (12 per cent). The remaining 4 per cent is coming from other countries within Europe. From these locations, the goods are transported to the nine in scope warehouses (packed and re-packing). These inbound transportation activities have an annual cost of $3.5 million USD for a delivered quantity of 163,000 MT. Furthermore, there is made use of two transportation modes: ocean-road (57 per cent) and road (43 per cent).

The inbound process includes the transportation from manufacturing site to the warehouse. However, for the overseas manufacturing sites, the inbound process regards the shuttling from the port to the warehouse. From the total inbounded of these overseas products, the majority is unloaded at the Rotterdam port, namely 91 per cent, whereas the remaining 9 per cent is unloaded at the Antwerp port. After the unloading process, the containers are shuttled to one of the warehouses. While analyzing the shuttling process, it can be concluded that there are containers unloaded at Antwerp Port, however the warehouse is located in Rotterdam. In these cases, a shuttling between Antwerp and Rotterdam is required. In total approximately 10,000 MT is shipped to Antwerp port, from this number of goods, 1,300 MT is shuttled to Rotterdam.

3.1.2 Outbound streams

From the warehouses, goods are transported to a customer. In total there are 1,345 customers served by the warehouses, which are mainly located in Europe (91 per cent). An overview of the customer locations is visualized in the map in Figure 4.

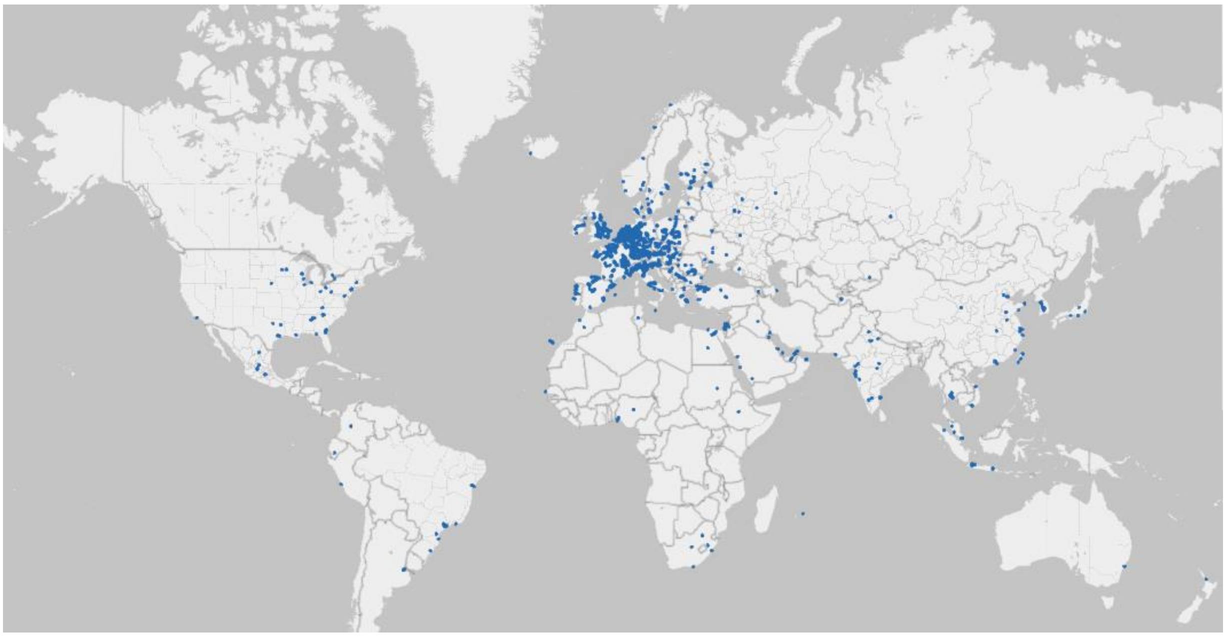

Figure 4. Customer locations (1,345 points)

The customers are served by the six packed-warehouses and one of the repacking-warehouses. These outbound transportation activities have an annual cost of $14.5 million (USD) for a delivered quantity of 198,000 MT. These costs are based on the total planned costs for transporting by land, and the drayage costs[2] for shuttling from warehouses to the loading port. There is made use of two transportation modes: ocean-road (18 per cent) and road (82 per cent). Compared to 163,000 MT inbounded, the outbound values are significant higher, this difference is further analyzed in Section 3.4 (data validation). While comparing the inbound and outbound costs, the outbound costs are roughly four times higher than the inbound costs. This comparison suggests that the warehouse locations are not centrally located relative to the customers.

During analysis it is noticed that goods are transported from different warehouses to the same customers. This is the case for 138 customers, who receive 40 per cent of the total demand (78,500 MT). This situation is probably inefficient, since the transportation of these different warehouses is conducted separate and a combination of the transport can probably result in an improved network, i.e. higher truckload and reduced number of transportation movements.

Besides the inbound and outbound streams there are also transportation streams between warehouses, i.e. intern transportations. In this activity, goods are moved between a packed-warehouse and a re-packing warehouse. This process is necessary because the re-packing activities are not performed at the storing warehouses. In total there is a quantity of 10,000MT transported between warehouses, with a transportation cost of $0.2 million USD.

[2] There are no costs given for the outbound shuttling from warehouse to the port (for overseas customer). Therefore, it is assumed that the shuttling costs of outbound are equal to the shuttling costs of inbound considering drayage costs per kg.

As can be seen in Figure 4, the customers are not uniform distributed, i.e. the customer locations are not equally distributed across the geographical range. When analyzing the geographical distribution of the demand, it is also visible that this is not uniform distributed since the demand is not equally spread over the geographical range, as can be seen in the worldwide gravity of the demand (Figure 5) and a zoomed-in version of Europe (Figure 6).

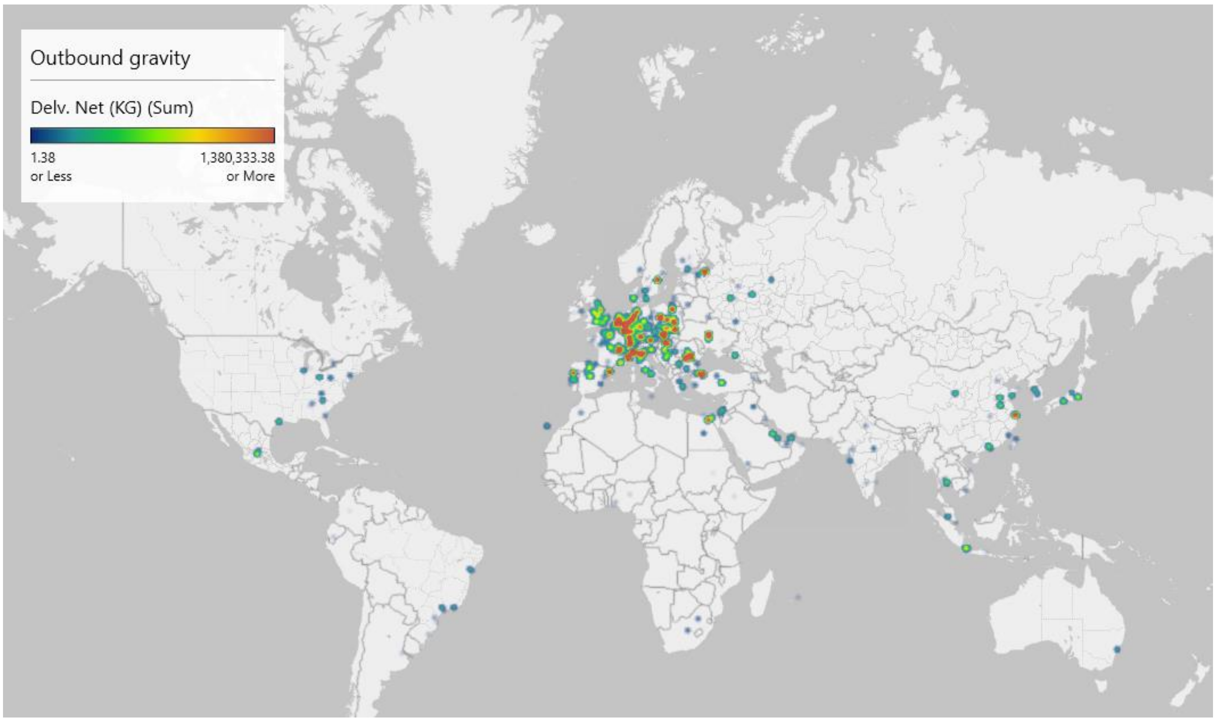

Figure 5. Geographical gravity of demand (global)

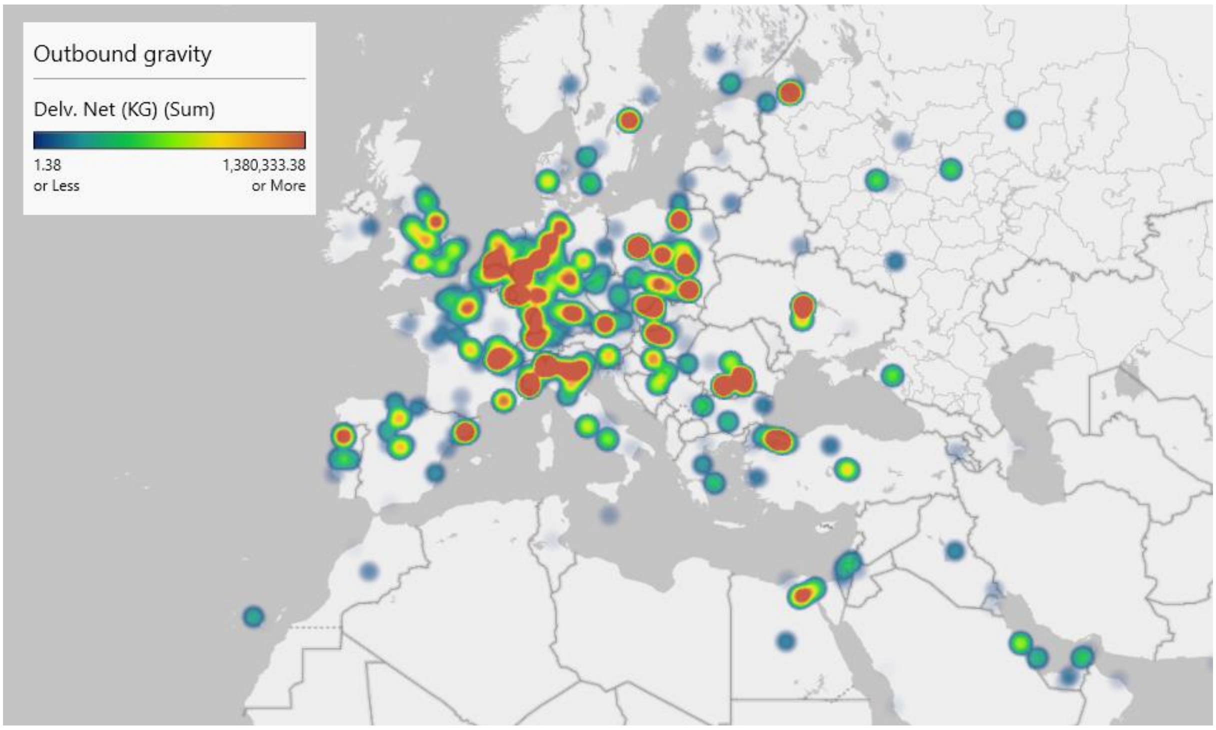

Figure 6. Geographical gravity of demand (Europe)

3.1.3 Transportation rates

As discussed before, the costs of transportation consist of inbound transportation costs and outbound transportation costs. At first, the inbound process is performed by the warehouse supplier, therefore there is a fixed price per container agreed and contracted. For each manufacturing site (or port), to each warehouse, another price per container is specified. Secondly, the outbound process is not performed by the warehouse supplier, but by carriers from other third parties. Therefore, the transportation rate for the outbound process depends on the price that a carrier offers and differs per carrier. When analyzing these transportation rates, it is concluded that the distance to the customer and the truckload of the transportation affects the price that a carrier offers.

In order to show the relationship between the price, the distance, and the truckload, a comparison is made in Table 1. In this table, high is defined as 10 per cent higher than the weighted average, low is defined as 10 per cent lower than the weighted average, and average is others. As can be seen, in most cases a high distance, results in a low price per km, and a high truckload, in a low price per kg and the other way around.

Table 1. Transportation rates overview per carrier

Vendor[3]	% trucks	Price per truck	Distance	Price / km	Truck load	Price / kg
1	19	average	low	high	high	low
2	18	high	high	low	high	average
3	13	low	high	low	low	average
4	9	low	low	high	low	low
5	8	high	high	high	high	high
6	7	low	low	high	low	average
7	6	average	low	high	low	high
8	3	low	low	high	high	low
9	3	low	high	low	low	high
10	3	low	high	low	low	high
11	2	average	high	low	high	low

Since the transportation rate is depending on the two variables, there are two possibilities to lower the costs: optimizing (1) the distance and/or (2) the truck load. In order to improve, it is important to analyse the current situation of both aspects.

At first, the travel distance of the current distribution network is analyzed. The total distance consists of the travelled distance for the inbound and outbound activities, for which there are two options: total distance from manufacturing site or customer to a warehouse (truck distance) or shuttling distance from port to a warehouse (shuttling distance) for overseas locations. An overview of the annual travelled distances is shown in Table 2. As can be seen, the majority of the distribution network distance originates from outbound activities, which is in line with the higher outbound costs compared to the inbound costs.

[3] The eleven shown carriers are responsible for 90 per cent of the outbound transportation, the other carriers are not considered since the truck frequency is low and therefore less reliable.

Table 2. Travel distance current distribution network

Transportation activity	Total distance [km]	Truck distance [km]	Shuttling distance [km]
Inbound	710,000	660,000	50,000
Outbound	13,400,000	13,290,000	110,000
Total	14,110,000	13,950,000	160,000

Secondly, the loading factor is defined as the ratio of the average load to the total truck capacity, which can be calculated by the following equation.

$$Truck\ loading\ factor = \frac{average\ truck\ load\ [MT]}{capacity\ of\ the\ truck\ [MT]}$$

The average load of the truck is obtained by considering the number of trucks and the delivered quantity (kg). In the distribution network there are two capacity types: (1) a container is used for products that are shipped oversea and (2) a truck is used for products that are transported by land. For which the capacity of a container is 25 MT and the capacity of a truck is 24 MT. Since the capacity and the product units are given in MT the data can be used directly in the equation. The average loading factor of the inbound [4] equals 0.87 and for the outbound [5] equals 0.64.

3.1.4 Transportation Distribution

In this section the transportation distribution is studied by analyzing historical demand. Figure 7 shows the distribution of the daily inbound streams during the last 33 months, whereas Figure 8 shows the distribution of the daily outbound streams (demand) during the same time period.

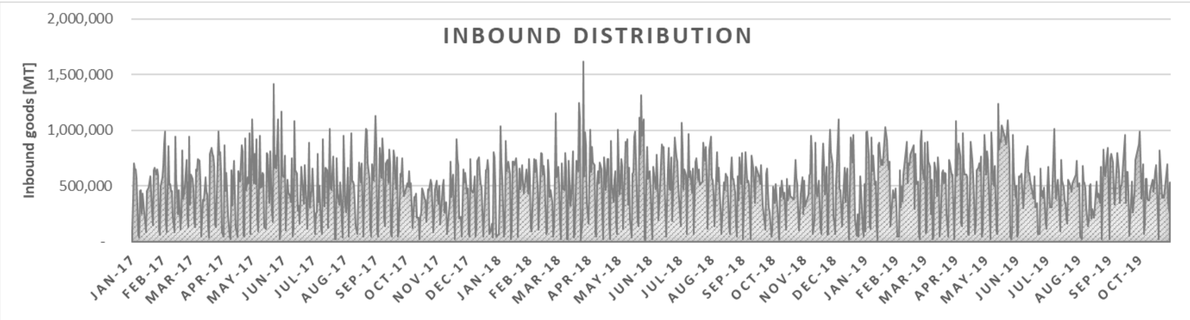

Figure 7. Inbound [kg] distribution (January 2017 - October 2019)

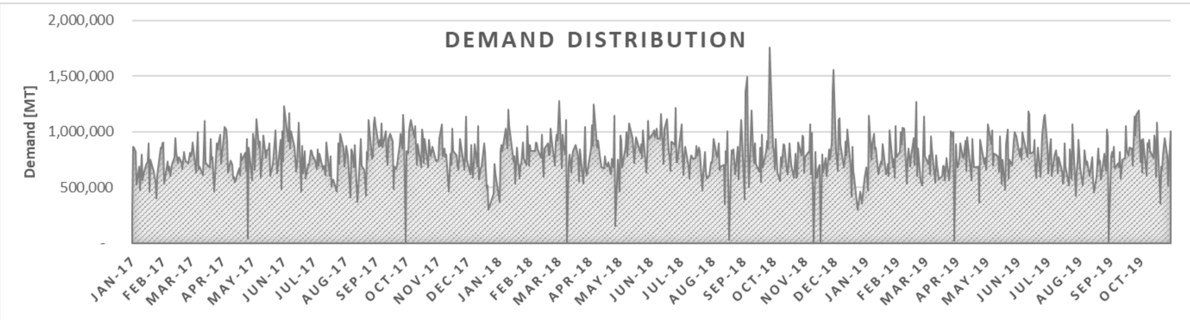

Figure 8. Demand [kg] distribution (January 2017 - October 2019)

[4] Truck loading factor inbound $= \frac{22,193}{24,000} * 0.43 + \frac{20,725}{25,000} * 0.57 = 0.87$

[5] Truck loading factor outbound $= \frac{15,230}{24,000} * 0.82 + \frac{15,957}{25,000} * 0.18 = 0.64$

Based on the graphs there seems no trend visible in the demand, i.e. the average stays the same during the analyzed period. For the outbound streams, there seems a seasonality during December in which the demand is lower than the other months. This seasonality seems to result in a decreased supply of goods during the period of October and November, see Figure 7. Based on the autocorrelation function (ACF) as shown in Figure 9, it is proven that there is no trend in the demand since it is not possible to shift the ACF-graph over time and getting a similar reliable graph. In addition, it is proven that there is some seasonality, since there are waves visible during the year. As can be seen, there are twelve peaks and troughs, which suggests monthly seasonality. However, since the considered future demand is annual based, it is not important to include this seasonality in the forecast. Therefore, it is concluded that the annual demand is stable over time and annual future demand is assumed to be constant.

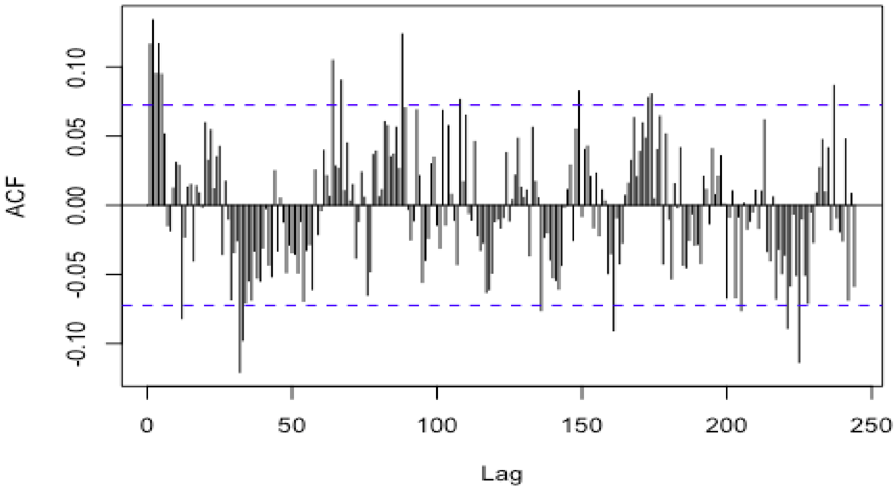

Figure 9. Autocorrelation function of daily demand

3.2 Warehouse activities

In order to create an overview of the annual warehouse activities, the inventory levels during the last year (October 2018 till September 2019) are analyzed by using an internal data source. The warehouse activities regard the handling and storing of goods, for which the data collection results are discussed in this section.

3.2.1 Inventory data

The project regards nine warehouses where inventory can be held. The inventory consists of goods produced at one of the manufacturing sites of the Company. The total daily inventory level is on average 44,000 MT, which consist of 674 SKUs[6]. These SKUs have characteristics that might influence the inventory costs: for goods that are dangerous or require temperature control a surcharge on the holding costs is added. From the total inventory, 9 per cent regards dangerous goods and 16 per cent requires temperature control. The inventory costs consist of three cost types: handling, holding, and administration costs, which regards contracted prices with the suppliers. In total there is an annual average inventory cost of \$6.6 million USD[7] for handling, holding and administrating an average daily inventory of 44,000 MT.

[6] SKUs = Stock Keeping Units, which is defined as GMN (product code) in the systems of the Company
[7] The prices are contracted and given in euros; therefore, an exchange rate of 1.11 is used, i.e. 1 Euro = 1.11 USD

Furthermore, the storage location(s) of each product type is analyzed. It is concluded that there are products stored at multiple warehouses. Therefore, it is counted in how many warehouses each product type is stored. In addition, it is analyzed what the proportional volume of the stored inventory is in a certain number of warehouses. This data can be found in Table 3.

Table 3. Number of storage locations per SKU

	5 warehouses	4 warehouses	3 warehouses	2 warehouses	1 warehouse
Number of SKU	1	14	12	67	580
Inventory volume (% of total)	4	23	7	14	52

Based on Table 3, it can be concluded that the majority of the products are stored only at one warehouse: 580 products, which are responsible for 52 per cent of the inventory in volume. However, 48 per cent of the total inventory is stored in at least two warehouses, from which 14 per cent is stored in two warehouses (67 SKUs), 7 per cent is stored in three warehouses (12 SKUs), 23 per cent is stored in 4 warehouses (14 SKUs), and 4 per cent in five warehouses (1 SKU). The storage of the same goods at different warehouses might result in transportation, inventory and complexity inefficiencies. At first, it seems likely that it requires multiple movements to one customer because a product is stored in multiple warehouses. Secondly, storing one product in different warehouses results in higher inventory since each warehouse stores a safety stock. At last, it gives an unclear overview of the total inventory per product.

3.2.2 Inventory rates

The costs of inventory consist of costs for handling and holding a good for a certain period in a certain warehouse. These inventory costs are contracted with the suppliers and depend on the product specifications of the good, e.g. dangerous goods are more expensive to store than non-dangerous goods. The main inventory costs are handling in and out, holding, and administration. The handling and administration costs are the same for each SKU and are a function of the throughput (demand). The holding costs are based on the average inventory level, for which there is a normal price agreed with surcharges for dangerous goods and temperature control goods[8]. Furthermore, in the current distribution network there is one packed warehouse on site, which is managed by the Company. Therefore, the cost types differ and consist of resource costs such as building, labor, and energy. In contrast to the outsourced inventory, these costs are not fixed per inventory unit but fixed per time unit. In addition, there are activities such as re-packing and (re-)labelling the goods, which are performed by the re-packing warehouses. Goods that require re-packing are handled by at least two warehouses: one for re-packing and one or more for storing. Therefore, these re-packing products requires multiple handling and administration activities, which results in higher costs.

A comparison of the handling, holding and administration costs of the different packed warehouses, results in Table 4. In this table, TC is defined as temperature control, DG is defined as dangerous goods and the surcharge costs is calculated as percentages of the holding costs. In

[8] The temperature control regards goods should be stored below a temperature of 18 or 24 Celsius degree. Since this temperature is not reached in the winter months, the surcharge is only applicable for 6 months during summer and spring.

the table, high is defined as 10 per cent higher than the weighted average, low is defined as 10 per cent lower than the weighted average, and average is others. When a column is empty it means that there is no cost applicable. The row "Handling, holding & administration [per MT]" is the total inventory cost without surcharges. When comparing the warehouses, it is visible that warehouses that offer the service temperature control, have a lower normal price but a high surcharge price. However, the majority of the products stored in these warehouses requires temperature control. Therefore, the differences between the real costs of the warehouses become smaller. The comparison should be more reliable when the surcharge costs are known for all warehouses. However, the additional services are currently not offered by all warehouses and therefore not available.

Table 4. Transportation rates overview packed warehouses

Costs	**B1**	**B2**	**B3**	**B4**	**B5**
Handling, holding & administration [per MT]	high	low	low	average	high
Handling [per MT]	high	low	low	low	high
Holding [per MT/month]	low	low	low	high	high
- Surcharge TC [per MT/month]		+ 196%	+ 196%		
- Surcharge DG [per MT/month]	+ 53%				
Administration [per MT]	high			low	low

When analyzing the costs in the contracts, it is noticeable that the contracts differ per supplier. At first, each supplier uses various types of costs, in total there are 150 different costs mentioned in the contracts. Secondly, each supplier uses other unit prices, such as prices per pallet, per MT, per container, or per order. These different types of costs make it complex to have a clear overview of the total and main costs. In addition, the different unit prices and types of contracts make it hard to compare costs between suppliers. Therefore, the complexity of the contracts makes it unclear what the costs are and how to identify them, which results in inefficiencies.

<u>*3.2.3 Inventory Distribution*</u>

In this section the inventory distribution is studied by analyzing historical data (October 2018 – November 2019). The data is visualized in a distribution graph as shown in Figure 10.

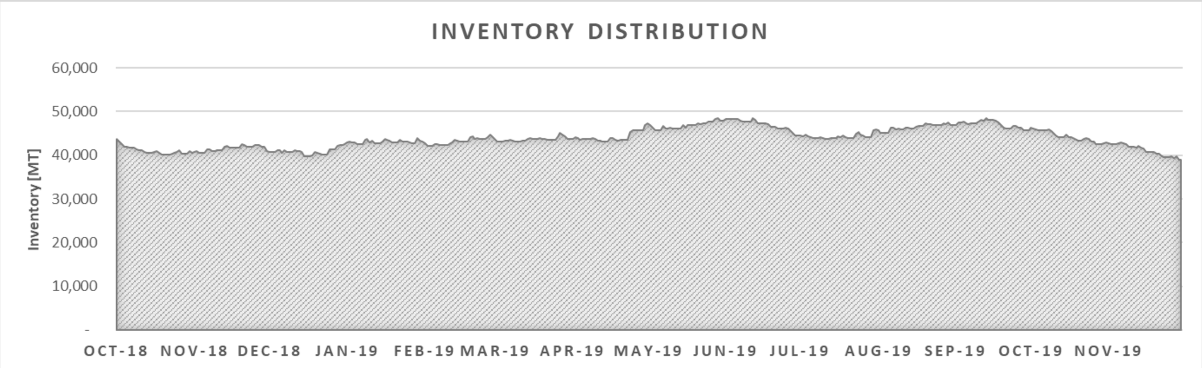

Figure 10. Inventory distribution per product group (October 2018 till November 2019)

As can be seen in the graph, there is some declining seasonality during the end of the year (October, November, December). This inventory seasonality parallels the seasonality in the demand, which makes sense since the inventory level is depending on the transportation of goods, i.e. goods that are coming in (inbound) and going out (outbound) during a period in time. A method to investigate this relationship between the transported quantity of goods and the inventory level is to calculate the cycle time of the goods by using Little's Law (Little & Graves, 2008) with the equation: $CT = \frac{WIP}{TH}$. In this equation, TH is the throughput (i.e. the outbound streams), WIP is the work in progress (i.e. the inventory in the warehouse), and CT the average cycle time that goods spend in the warehouse. Hence, the average cycle time of a good in the current distribution network is 84 days (annual turnover rate of 4.4). Based on this turnover rate and the forecasted demand, the future inventory level will be determined.

3.3 Order processing activities

The order processing activities regard activities that are required to execute the orders in the distribution network. It regards mostly office tasks or activities to maintain and improve the relationship with a supplier. Therefore, the order processing costs consist of time spend on these activities by employees of the Company, which is defined in labor costs. The order processing activities in the distribution network are as follows: handling of invoices, Performance Management, and supplier visits. The size of these activities depends to the number of suppliers in the network; the more suppliers, the more activities required.

3.4 Data validation

In this section, a comparison between the in- and outbound (including intern transportations) per warehouse is made to test the accuracy of the data. The figures are shown in the comparison in Table 5. A negative intern transportation suggests outbound streams from the warehouse, whereas a positive intern transportation suggests incoming streams to the warehouse. In the comparison it is investigated if the quantity of the incoming and outgoing goods (including intern streams) in a certain period are similar and therefore valid.

Table 5. Validation table transportation activities

Warehouse	Inbound [MT]	Outbound [MT]	Intern transport [MT]	Increasing +, decreasing (-) [MT]
A1 & B3 [9]	27,500	56,500	(1,000)	(30,000)
B1	99,500	98,000	(8,000)	(6,500)
B2	1,500	500	(1,000)	-
B4	22,000	20,500	-	1,500
B5	4,000	4,000	-	-
C1	9,000	19,000	10,000	-
C2	*50*	-	-	*50*
C3	*250*	-	-	*250*

[9] In the data, the plant code is similar for warehouses A1 and B3, these warehouses receive from the same local manufacturing site and for these transportation flows there is no distinction made in storage location.

As can be seen in the fifth column of Table 5, the in- and outbound streams for some warehouses differ. It is assumed that a difference is significant when it is greater than the capacity of one shipment a month, which equals the annual capacity of 288MT. Consequently, there are three significant differences: (1) warehouses A1&B3, (2) warehouse B1, and (3) warehouse B4.

The first difference, warehouses A1&B3, can be explained by the fact that there is no transportation registered in the system for goods that are stored in the packed warehouse on site. In this situation, goods are moved from production floor to the warehouse in the same plant, which regards no transportations nor costs. Therefore, the registered inbound is only for warehouse B3, and it is assumed that the difference equals the inbound of warehouse A1. This results in adjusted figures as shown in Table 6.

Table 6. Adjusted in- outbound warehouses A1&B3

Warehouse	Inbound [MT]	Outbound [MT]	Intern transport [MT]	Increasing +, decreasing (-) [MT]
A1	30,000	30,000	-	-
B3	27,500	28,500	(1,000)	-

The other two significant differences: warehouse B1 and B4, suggest a change in inventory for that time period. When the inbound is higher than the outbound there should be an increased inventory and when the outbound is higher than the inbound a decreased inventory. Therefore, these differences will be further discussed in the validation of the inventory.

The inventory is validated by comparing it with the suppliers' invoices and the transactions in SAP. Both comparisons show no (significant) differences, so the collected inventory data is defined as valid. With this inventory data, the validation of the transportation stream can be continued. There are two warehouses (B1 and B4) for which the inbound and outbound streams differ. These differences suggest an increased or decreased inventory level for the measured time period (October 2018 till September 2019). In order to validate, the expected inventory change is compared with the real inventory change as shown in Table 7.

Table 7. Validation inventory change based on in- and outbound

	Expected inv. increasing +, decreasing (-) [MT]	Begin inventory 10-01-2018 [MT]	End inventory 09-30-2019 [MT]	Increasing +, decreasing (-) inv. [MT]	Unexplained difference [MT]
B1	(6,500)	20,500	19,000	(1,500)	(5,000)
B4	1,500	3,000	4,500	1,500	-

Based on the table, it is concluded that the difference in inbound and outbound streams for warehouse B4 is explained by the increased inventory. However, for warehouse B1 there is in reality only a decreased inventory of 1,500 MT instead of the expected 6,500 MT. Therefore, there is still 5,000 MT that cannot be explained by the inventory and requested further research.

Chapter 4. Conceptual Model

In this chapter a representation of the mathematical model is presented. It is a pre-processing phase in which the model including decisions and assumptions are described. It is the aim to qualitatively answer Sub-research question 2: *What methodology is needed that determines the optimal design of a multimodal facility location problem such that costs are minimized?*
The chapter structure follows three aspects of the model: the data input (Section 1), the ideal situation (Section 2), and the different scenarios for implementation in reality (Section 3).

4.1 Data input of the model

Based on the analysis of the current distribution network, the data input of the model is determined. The data input of the model consists of the variables (locations, products and transportation modes) and its values (demand, inventory level, costs). In this section, decisions and assumptions regarding these two aspects are described.

4.1.1 Model variables

The model variables are the aspects that are considered in the model and affects the result of the model. As analyzed in the current distribution network, there are several depending variables. At first, the locations of the manufacturing sites, the warehouse(s), and the customers. Secondly, the products that are transported and stored. At last, the transportation modes that are used to transport products in the distribution network.

Manufacturing sites: Based on the current situation, it can be concluded that the products are imported from 20 manufacturing sites, which are globally located. Goods that are coming from overseas countries (outside Europe) are shipped to the port of Rotterdam or Antwerp and from there shuttled to a warehouse. Since only the shuttling part is in scope of the project, the origin of these goods is no longer the manufacturing site but the unloading port. Therefore, the manufacturing sites outside Europe are aggregated to one point: the unloading port. Since the majority of the goods (91 per cent) is unloaded at the Rotterdam port and even a proportion of the goods unloaded at Antwerp Port is shuttled to Rotterdam, it is decided that the port of Rotterdam is the only considered port in the model.

Warehouse(s): In the current situation, there are nine warehouses (six packed and three re-packing). The packed warehouses perform mainly storing activities, whereas the re-packing warehouses offer only value adding activities such as re-packing and re-labelling. Therefore, the warehouses offer different services. Furthermore, there is one packed warehouse on site. However, the core business of the Company is to manufacture chemical products, so the distribution activities are not a core competence of their business plan. Therefore, the ideal situation is to outsource all distribution activities. The multiple warehouses are the result of the acquisition with different businesses. During this acquisition, the warehouses were taken over, which results in multiple warehouse locations within a radius of 75 km. Compared to the global distribution area, the warehouse locations are located too close to each for transportation advantages. On the contrary, there are several disadvantages accompanied by having multiple warehouses. As can be concluded from the current distribution network, these inefficiencies are as follows. (1) Manufacturing sites serve different warehouses and customers are served by different warehouses. Therefore, goods cannot be transported together from or to the same destination, which results in lower truckload and multiple movements. (2) 13 per cent of the containers unloaded at the Antwerp port, are shuttled to a warehouse in Rotterdam. However, there is also a port in Rotterdam, which makes this shuttling redundant. (3) There are different warehouses that store the same product(s), since each warehouse holds its own safety stock, it seems likely that this leads to higher inventory levels. (4) Furthermore, in this case, the same product for one customer, might come from different warehouses, which requires multiple movements. (5) In addition, as noticed during the analyse, the storing of the same products at different locations, gives an unclear overview of the inventory level of each product, which might result in data errors. (6) Comparing to the inbound truckload, the outbound truckload is low, which results in higher transportation costs per product unit and multiple movements to transport the same quantity. (7) In the current situation, the re-packing and the storing activities are performed by different warehouses. Therefore, there are intern transportations required between warehouses, which leads to higher costs: transportation, handling and administration. (8) The different supplier contracts result in an unclear overview of the total costs due to different mentioned definitions and price units. (9) The order processing activities needs to be performed multiple times by employees of the Company, because of the multiple warehouses. (10) The complex distribution network results in the storage of data in different sources, which are contradictory and results in incomplete and incorrect data. In order to change the current distribution network to an ideal situation, it is the aim to reduce all these inefficient activities. Therefore, it is decided to consolidate the nine warehouses to one distribution center, with probably additional cross docking facilities. The distribution center should be able to perform all the required activities: storing and re-packing. Furthermore, in the current situation the inbound costs and travel distances are lower than the outbound costs and travel distances, so it seems likely that the warehouse location is not optimal located. Therefore, in the ideal situation, the warehouse should be optimally located relative to the manufacturing sites and customers.

Customers: Based on the analysis of the current distribution network, it can be concluded that the customers are on geographical field not uniform distributed. This can be explained by the fact that it regards a business to business process. The customers of the Company are manufacturers with chemicals as raw material. In contrast to retailers, the majority of these companies are not selling directly to the consumer and therefore there is no need to be uniform located. Since the customers are not uniform located, the customer locations need to be considered in the model. However, the data suggests that the number of customers for this problem: 1345 locations, is overwhelming. Therefore, the customers will be aggregated to customer zones[10]. At first, all customers that are located in close proximity to each other are aggregated. According to (Simchi-Levi, Chen, & Bramel, 2005), an effective technique that is commonly used is to aggregate customers according to the zip-code. Therefore, it is decided to aggregate the customers on the first two digits of the zip-code in combination with the country code, i.e. XX-CC, where XX is the country code and CC are the first two digits of the zip-code.

[10] Various researchers report that aggregating data into about 150 to 200 points usually results in no more than a 1 percent error in the estimation of total transportation costs; see (Ballau, 1992) and (House & Jamie, 1981). Furthermore, aggregated data can be useful because it reduced variability, which results in significantly more accurate demand at aggregated level (Simchi-Levi, Chen, & Bramel, 2005)

Secondly, goods that are for customers from overseas countries (outside Europe) are currently shuttled to the port of Rotterdam or Antwerp and from there shipped to a customer. Since only the shuttling part is in scope of the project, the destination of these goods is no longer the customer zone but the port. Therefore, the customer zones with container shipments are aggregated to one point: the loading port. As discussed before, it is decided to only consider Rotterdam port in the model.

Products: Based on the analysis of the current distribution network, there are 674 SKUs. It is decided to aggregate these SKUs to product groups. This aggregation is based on specifications that influence the outcome of the model, i.e. the transportation costs and inventory costs.
At first, for the transportation rate, it is assumed that the costs be similar for each SKU and therefore it is not needed to make a distinction between different product types in transportation. This assumption is based on the low effect on the transportation costs that is accompanied with a splitting in product groups. Furthermore, including product groups in the transportation routes will make the problem highly complex. In addition, there is no reliable data available to make a distinction, i.e. data from different sources need to be combined, which seems likely to make the outcomes less reliable than by ignoring the product types. Secondly, as can be seen in the analysis, there are holding surcharges for dangerous goods and temperature-controlled goods. Therefore, the two specifications: dangerous and temperature control, affect the inventory rate. Hence, the goods will be aggregated on these specifications into three product groups: goods without specifications, dangerous goods, and goods that require temperature control.

Transportation modes: In the current distribution network, the in-scope transportation is conducted by one mode: truck. However, the use of different modes can be an improvement on costs as well as sustainability. In addition to a cost reduction, the Company aims a more sustainable and flexible transportation network. Therefore, multi-modal transportation will be considered in the ideal situation with the following modes: road by truck, rail by train and inland waterway by barge. For the model, it is assumed that only transportation between two facility locations can be performed by rail or inland waterway. Therefore, a facility should have an own connection to rail or inland waterway, such that there is no truck shuttling needed between a (un)loading station or port and the facility. The manufacturing sites and customers are located inland wherefore the water and rail connection with a facility are unknown. Therefore, it is decided to perform all shipments from manufacturing site to a facility and from a facility to a customer zone by truck. Furthermore, each transportation mode has different key transportation features, as shown in Table 8. Based on this information can be concluded that the values such as transportation rates, travel distance, and sustainability are affected by the selected mode.

Table 8. Comparison transportation modes (SEALS, 2008)

Mode	Lead-time	Accessibility	Cost	Capacity	Emissions
Truck	Short	Very high	Moderate	Low	High
Rail	Long	Low	Low	Moderate	Low
Inland waterway	Very long	Low	Very Low	Very High	Very Low

<u>*4.1.2 Model values*</u>

Based on the current distribution network, it is determined which values influence the design and affect the optimal situation. These values are in relation with the variables as described in 4.1.1. Since the main research objective is to provide an optimal design, in which optimal is defined as the situation where the total costs are minimized, the main value is costs. As can be concluded from the analysis there are two type of costs: (1) transportation costs and (2) inventory costs. At first, as concluded in the analysis, the transportation rates depend on the travel distances and the truckload as a function of the number of movements per transportation stream. These movements are based on the total delivered quantity (demand) per stream. Secondly, the inventory rates consist of handling, holding (including surcharges) and administration costs, as a function of the inventory level and throughput. Furthermore, for the model in general, it is decided to use annual values.

Transportation rates: The transportation rate considered in the distribution network is a price paid per movement (per truck). The number of movements is depending on the delivered quantity (inbound and outbound). The quantity that needs to be delivered is equal to the total demand in the system. Therefore, an important value of the model is the future demand. The total costs consist of costs for inbound and outbound, which differ in rates. For inbound from overseas countries, there is a port-shuttling price agreed with the warehouse supplier. Whereas, for the manufacturing sites within Europe, the inbound rates are equally distributed as the outbound rates. Based on the analysis of the outbound rates, it is concluded that the distance to the customer and the truckload affects the price that a carrier offers. In which the relation is as follows: a high distance, results in a low price per km, a high truckload, in a low price per kg and the other way around. Therefore, both variables follow an increasing concave down cost structure relative to the truck price. According to these two variables, the function for the transportation rate is divided into two parts. At first, the full-truck-load (FTL) tariff is determined while considering distance as a concave cost structure following a polynomial distribution (order 2) (Smith, Campbell, & Mundy, 2007). Secondly, based on the defined FTL tariff function, a discount factor (shape parameter) for LTL transportations is obtained to calculate the LTL tariff with the equation according to (van den Vlist & Broekmeulen, 2006). This LTL tariff is multiplied with the number of movements in order to calculate the total costs per stream. Hence, the structure of the transportation rate is as visualized in Figure 11.

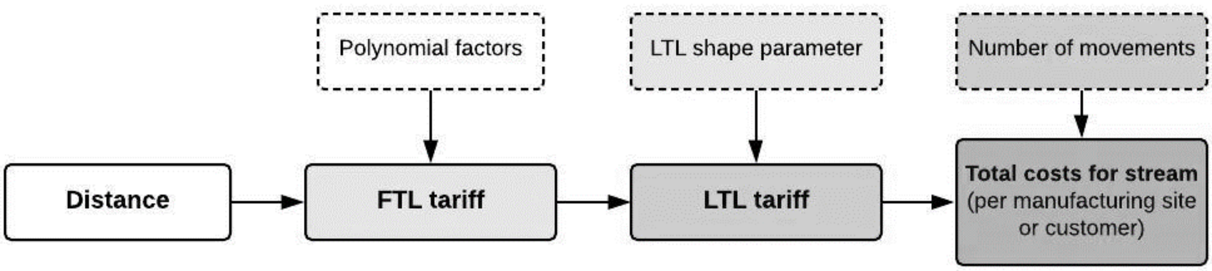

Figure 11. Structure (input and output parameters) of transportation rates

This defined transportation rate function (FTL tariff and LTL tariff) will be used for all transportation movements in the model.

Delivered quantity: The distribution network regards a multi-echelon goods delivery: from manufacturing site to the warehouse (i.e. the inbound quantity) and from warehouse to the customer (i.e. the demand) probably through a cross docking facility. The distribution network is depending on the demand of customers. As can be seen in the analyse, there is no trend in the demand during the last three years (January 2017 – October 2019) so the demand is stable over time. Since the future demand will be forecasted on historical data, it is assumed that this stability will continue in the coming years. As a result, the current annual demand can be used in the model. To serve the customers from the warehouse, there are inbounded goods required. Based on the data validation, the quantity of goods inbounded was 5.000 MT less than the outbounded quantity, which cannot be explained by a change in inventory level. It is most likely that there are missing values in the inbound streams and therefore, it is assumed that the outbound is correct and there is a missing quantity of 5,000 MT on inbound. This missing quantity is divided among the manufacturing sites (proportional) and added to the inbound quantity as input for the model.

Frequency: The frequency is defined as the number of movements per transportation stream, which is related to the demand and truckload, as can be seen in the following equation:

$$Annual\ demand\ [MT] = Annual\ frequency * Average\ truckload\ [MT]$$

Since the demand is stable over time, it is assumed that the frequency is stable against the same performances. In the current situation, it is assumed that the customer frequency is the same as the number of customer orders. However, the consolidation of warehouses and aggregation of customers into customer zones results in less different streams and an opportunity to improve the transportation performances. For these streams, transportation of different orders can be combined subject to a customer service constraint. Therefore, there will be a maximum lead-time determined to deliver goods to a customer. Within this lead-time, orders can be combined in order to maximize the truckload and minimize the number of movements (frequency) per stream. For example, when a certain customer zone demands a daily order, there will be 365 annual deliveries for a service lead-time of 1 day, and 52 delivery moments for a service lead-time of 1 week. Furthermore, different transportation modes are considered in the model. For each transportation mode, the lead-time of transportation seems likely to differ, wherefore the customer service level per mode is different. Therefore, a transportation mode is only possible when the allowed lead-time is higher than the transportation lead-time of a mode.

Truckload: As described above, the truckload is a function of the demand and the frequency in which the demand (number of orders) is stable and the frequency is a function of the customer service level, i.e. the lead-time. Based on the lead-time and the number of orders, it can be determined what the minimum number of deliveries is. However, each transportation mode has a capacity constraint per movement, wherefore the transportation combination of different orders is limited. Therefore, the capacity is an important variable to determine model values. Since the capacity of a transportation mode is defined in the same quantity unit as the products (namely MT), there is no need to transform values. The packed products are very compact packed. However, the packaging and pallets utilize some truck capacity. Therefore, a maximum utilization rate is used to calculate the real capacity for the product units. As described before, in the ideal situation there will be three different transportation modes used: road, rail, and water. Based on the analysis of the current situation, it is known that the capacity of a truck is 24 MT. For transportation by water and rail, the products are packed in a container or railcar and transported by barge or train, which differ in capacity comparing to the truck.

Distances: The distance of a transportation stream depends on the location of its origin and destination. In order to define a location, the longitude and latitude are used. Based on these two variables, the Euclidian distance between two points is determined by using the Haversine formula (Robusto, 1957). In transportation, the actual travel distance is commonly higher than the Euclidian distance. Therefore, the Euclidian distance will be multiplied with a circuity factor to get the actual travel distance.

Inventory rates: Based on the current situation, it is concluded that the inventory rates mainly consist of three cost types: (1) handling, (2) holding, and (3) administration. The handling and administration costs are a function of the throughput, i.e. goods coming in and out. The holding costs depends on the inventory level for each moment in time and is commonly calculated per month. As can be concluded from the analysis, there are surcharge handling costs for dangerous goods and goods that requires temperature control. In the current situation, the costs are agreed in the warehousing contract, so there are different inventory rates per warehouse. An average of these different rates will be considered in the model.

Inventory level: As can be concluded from the analysis of the current distribution network, the inventory is in relation with the demand according to Little’s Law. Therefore, it is assumed that the inventory is stable over time and is equal to the average annual inventory level of the current situation. As discussed before, there are specifications that affect the inventory rates according which the products are divided into product groups. It is required to divide the inventory level among these groups to use as input of the model.

Throughput: The throughput is defined as the quantity of goods passing through the warehouse, i.e. the quantity of goods that is coming in and going out in a certain moment of time. As discussed before, there is a stable inventory level and the quantity of incoming and outgoing goods are equal. Therefore, the annual throughput equals the annual inbound quantity and the annual demand.

4.2 Ideal situation

As can be concluded from Section 1, the ideal situation regards a distribution network in which the current warehouse activities are consolidated to one distribution center (DC). This DC should be optimally located relative to the manufacturing sites and customers. In order to find this location, a gravity study is performed with the objective to minimize the transportation costs. Furthermore, the single-facility gravity study is extended to consider gravity locations of multiple facilities, in order to investigate the location of potential cross docking facilities (CDF). Based on the gravity study, possible facility locations are compared by using a facility location model. In this facility location model, the objective is to minimize the total distribution network costs under different scenarios.

4.2.1 Centre of gravity

Gravity location models are used to determine the center of gravity as a connecter between the manufacturing sites with the customers with the aim to minimizing transportation costs. These transportation costs consist of costs for inbound activities and costs for outbound activities. Both types of cost are a function of the transportation rates per stream and the number of shipments per stream. Therefore, the transportation rates and the frequency are the depending variables in the model. Based on the analysis, it is concluded that the transportation rates are affected by the distance and the truckload. For the inbound activities, the location of a manufacturing site determines the travel distance per stream, whereas this is the customer zone for the outbound streams.

At first, it is determined what the optimal location is for a distribution center by assuming **current performances** (the number of movements and truckload are the same as in the current network) and transportation performed by **truck** only, i.e. single modal. The DC should be optimally located relative to the manufacturing sites and customers. Therefore, the objective is to determine a gravity location for which the total transportation costs: inbound and outbound, are minimized. The input variables of the model are as follows: the manufacturing site locations, the customer zone locations, the inbound quantity or demand per stream, the current frequency per stream (number of orders), the current truck-load per stream, and the polynomial cost coefficients of the distance, and the discount factor for LTL transportations. Consequently, the decision variable of the model is the location of the center of gravity (coordinates), which determines the travel distance per stream. In the model, locations are defined by its coordinates (longitude and latitude) for which the travel distance between two locations is calculated by using the Haversine formula. All possible combinations of coordinates (within selected area) represent a feasible solution and are therefore candidates to become the optimal solution, i.e. the center of gravity. The **single-gravity model** is defined as a Nonlinear Program since the cost function is nonlinear. Therefore, the problem is solved by using a generalized reduced gradient nonlinear heuristic (multiple starts), which results a set of coordinates for the gravity center. This gravity center serves as the Distribution Center of the network, i.e. location where the Company concentrates the logistics of goods. All manufacturing sites deliver to this DC, so all goods are served by the DC. Therefore, it is required that the DC can perform the value adding services, such as re-packing and re-labelling.

From the DC, goods are transported to the customers, which can be done directly or through a cross docking facility (CDF). A CDF is deployed as a location in which products are moved from one transportation mode to another, with minimal warehousing. The advantage of a cross-docking facility is the opportunity to combine transportation from DC to CDF and use different transportation modes that are probably cheaper. In order to find the optimal location of the CDF(s), the gravity problem is transformed into a **multiple-gravity model**, with the following adjustments: deciding on multiple gravity points (facilities) while minimizing the outbound costs from facility to allocated customer. The problem is defined as a Nonlinear Program and solved by using the same generalized reduced gradient nonlinear heuristic (multiple starts). The model is solved multiple times for a different number of gravity points to determine.

For the founded gravity locations, the aim is to test its feasibility and if needed set a gravity region around the gravity point. A location is feasible for the ideal situation when it is connected to three different transportation modes: inland waterway, rail, and road. If a gravity point is not feasible, possible locations in the surrounding are researched by determining a gravity region. This gravity region is based on a five-percentage error from the optimal objective. Within this gravity region, all locations connected to at least two modes and located in a highly concentrated area for freight transport are set as potential facility location.

4.2.2 Facility location model

Based on the potential facility locations, the optimal distribution network is determined by using a facility location model. Facility location models are used to select the right facilities with the aim to minimizing the total distribution network costs, i.e. transportation and inventory costs. For the facility location model, the optimal set of facilities is determined by assuming **performances depending on customer service level** (the number of movements and truck-load depends on customer lead-times) and the transportation between facilities is performed by a **multi-modal** network, i.e. truck (road), train (rail) and/or barge (inland waterways).

All the possible facility locations within the selected gravity regions are candidates to become the optimal solution, i.e. a facility location. Since the multi-modality of the network is considered, it is decided what the possible transportation modes are between the DC location and the other facilities. The structure of the transportation rates in the gravity model, is used for the facility location as well. However, the different transportation modes have different cost coefficients. Therefore, it is expected that the transport rates differ per transportation mode per time unit. The FTL rate per mode is calculated by using a cost relation factor of truck-train and truck-barge per time unit as a function of the mode lead-time. Besides this transportation costs, the facility location model considers the inventory costs, which consist of handling and administration, and holding costs. These costs might differ per warehouse location based on geographical factors and differ per product group, wherefore it is important to consider the product groups in the model. The handling and administration costs are a function of the throughput, whereas the holding costs is a function of the inventory level.

A new considered aspect in the facility location model is the customer service. Based on the lead-time, the truckload is maximized, and the number of movements is minimized (as discussed in section 4.1.1). Therefore, it seems likely that different lead-times result in different optimal solutions with different objective values. These different solutions show opportunities to reduce costs while offering different service levels. Consequently, the decision variables of the model are (1) the set of facilities to open and (2) the modes to use between DC and CDF based on different lead-time scenarios. Furthermore, since all possible facilities locations are known, the distances (from manufacturing site to DC, from DC to CDF, from CDF to customer) are input variables in the model. Based on these distances, customers are allocated to the nearest facility. The allocation of customers to a facility determines the throughput of a certain facility. The **multi-modal facility location** is defined as a Mixed Integer Program, which is solved by using the same solver as the gravity study, i.e. the generalized reduced gradient nonlinear heuristic. The model is depending on the given customer service lead-time, therefore all lead-times from one till ten days are considered. The service lead-time is the lead-time between the ordering moment and delivery, i.e. transport from DC to the customer. Consequently, it can be divided into a lead-time between DC and CDF and a lead-time between CDF and customer. The facility location model is solved multiple times with a trial-and-error for all possible combinations of the partial lead-times within the given service lead-times. As a result, the effect of the service lead-time on the distribution network costs and design is analyzed.

4.3 Implementation in reality

In order to implement the ideal situation in reality, a sensitivity analysis and scenario analysis is performed. At first, since the ideal situation is based on some input assumptions such as the stability of the demand and the parameter setting, it is important to investigate how uncertainty in the data input affects the output of the model. In addition, the scenario analysis proposes the optimal facility location when considering the optimal network in terms of sustainability.

4.3.1 Sensitivity analysis

A sensitivity analysis determines how the uncertainty in the output of a mathematical model can be divided and allocated to different sources of uncertainty in its inputs. During this sensitivity analysis, the outcomes of the model will be recalculated under alternative assumptions, in order to determine the impact of variability. The sensitivity of two aspects will be tested: (1) parameter setting, and (2) demand. At first, in the gravity study the sensitivity of parameters in the model will be evaluated. According to previous studies and analysis, parameters for the model are set. However, based on side effects such as market changes and change in competition, these parameters might change over time. Therefore, parameters that affect the price (such as costs coefficients) and the distances (such as circuity factors) will be changed, wherefore the effect of parameter setting on the optimal situation will be obtained. Secondly, the sensitivity of the demand in the facility location model is evaluated. Based on historical data, it is assumed that the demand is stable in the future. However, there are uncertainties in the future that cannot be seen, which might lead to fluctuations in the demand. Therefore, it will be determined what the effect of a lower and higher demand is and how this might change the optimal distribution network design.

4.3.2 Scenario analysis

A scenario analysis is a process of analyzing future events by considering alternative possible outcomes. During the research, the main objective is to select an optimal distribution network design with the aim to minimize the total distribution network costs. However, as discussed in the motivation of the project, becoming more sustainable (i.e. energy efficiency) is a sub-aim of the project. Therefore, it is researched how alternative outcomes affect the sustainability (classified on carbon footprint[11]) of the distribution network design. Therefore, a comparison of the distribution network costs and environmental impact is made based on three situations: (1) current situation, (2) ideal situation, and (3) an optimal situation in which the objective is to minimize the carbon footprint. The comparison between costs and sustainability of the network will investigate the costs that the company might pay to become more sustainable.

[11] Carbon footprint is defined as the total Greenhouse gas Emissions caused by the distribution network.

Chapter 5. Transportation rates

In this chapter, the function for the transportation rates is defined. In the model, this function determines the transportation costs of the distribution network. Since the optimal distribution network is determined by minimizing these costs, this is an important function. The function is defined by analyzing historical data, which first needs to be removed from outliers (Section 1). After data cleaning, the literature is consulted to decide what models to use in order to define the function structure (Section 2). This function structure is further researched in two models: a full truckload (FTL) function in which the distance is considered (Section 3), and a less than truckload (LTL) function for considering the cost structure of the truckload (Section 4).

5.1 Data cleaning

In this chapter, a transportation rate function is defined that is used in the gravity model to calculate the transportation costs for a certain movement. In order to define the transportation rate function, historical data of different transportations is researched. In total there are 11,444 measurements considered. These measurements are raw data; therefore, it requires data cleaning. The variables in the data are truck price, truckload, and travel distance. Based on a box plot per variable, it can be concluded what the outliers are for each considered variable. As can be seen in Figure 12, the range of each considered variable is as follows.

- Truck price ($) = {1, 2709}
- Truck load (kg) = {7, 24000}
- Distance (km) = {5, 2102}

After data cleaning, there are 11,254 measurements remaining (190 outliers removed).

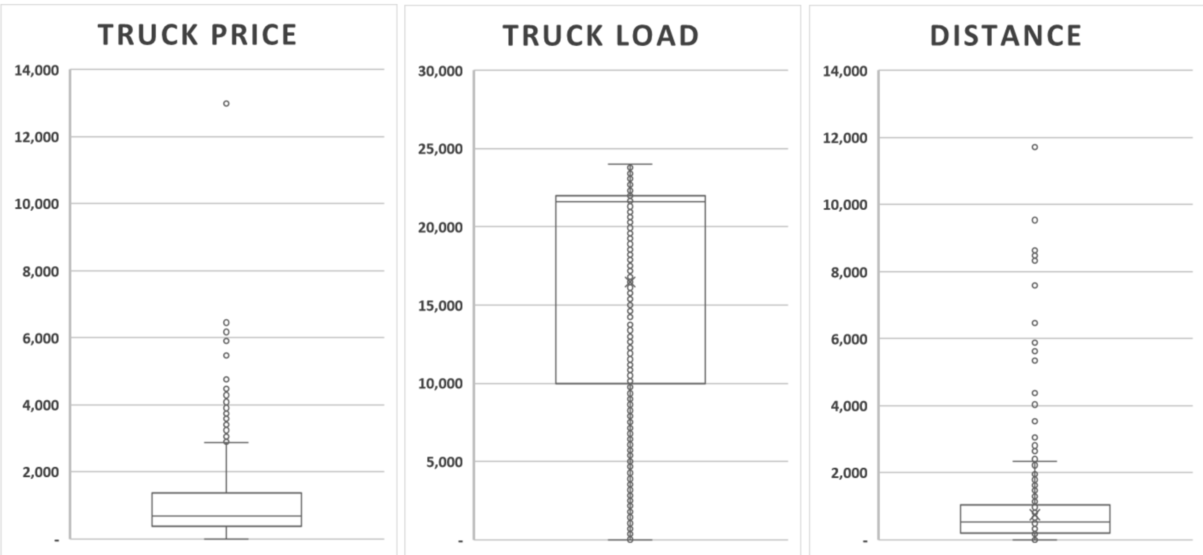

Figure 12. Box plots of model variables

5.2 Function structure

Based on the analysis of the current distribution network, the transportation rates are depending on two variables: (1) truckload, and (2) distance. It was concluded that both variables follow an increasing concave down cost structure relative to the truck price, i.e. a high distance, results in a low price per km, and a high truckload, in a low price per kg and the other way around. Various studies researched the concave cost function of LTL transportation rates. According to (Swenseth & Godfrey, 1996), the freight rates increase as shipping weights decrease. The effect of LTL on the rate per distance is evaluated by using three different freight rate functions: proportional (linear) function, exponential (nonlinear) function, and adjusted inverse (nonlinear) function. In addition, (Muriel & Simchi-Levi, 2003) presented a piece-wise linear concave cost function. According to (van den Vlist & Broekmeulen, 2006), the less than truckload (LTL) tariff can be approximate by using a shape parameter $0 \leq r \leq 1$ expresses the routing efficiency. More recent articles presented a model for LTL rates as a function of the truckload and the distance. At first, (Kay & Warsing, 2009) provided a model, where the LTL rates are inversely proportional to three shipment characteristics: density, weight, and distance of shipment. Secondly, (Smith, Campbell, & Mundy, 2007) proposed that transportation revenues are related to the fundamental cost drivers: number of shipments, the weight of shipments, and the distance of shipment. The presented model is a second-order polynomial approximation of the rates in relation to the frequency, truckload and distance of shipment.

While analyzing the data, it becomes clear that there is no (piecewise) linear trend for one of the variables, since the cost coefficient changes continue. Therefore, most of the cited models are inapplicable. In order to define a function for the transportation rate, the concave relation of both variables, distance and truckload, is considered separately by using two models. One of the models is used to define a function for LTL tariff in which the truckload is considered. This LTL tariff is a function of the FTL tariff and a discount factor for LTL (van den Vlist & Broekmeulen, 2006). In this function, the used FTL tariff is defined by using another model: considering distance as a polynomial cost structure (second order) for FTL (Smith, Campbell, & Mundy, 2007). These two models and its values are detailed explained in the next sections.

5.3 FTL tariff considering distance

With the cleaned data set the concave relation of the price relative to the distance is analyzed. Since the transportation rate is also related to the truckload, it is required that the effect of the truckload is removed from the data. Therefore, the concave function of the distance is investigated while considering only FTL tariffs. The truckload from which the FTL tariff counts is determined by plotting the truckload against the truck tariff. It is assumed that the FTL tariff counts for truckloads which have less than 1 per cent error from the maximum tariff (at maximum capacity of 24,000 kg). Based on the trendline function as figured in Figure 13, this is the case for truck-loads equal or greater than 22,000 kg.

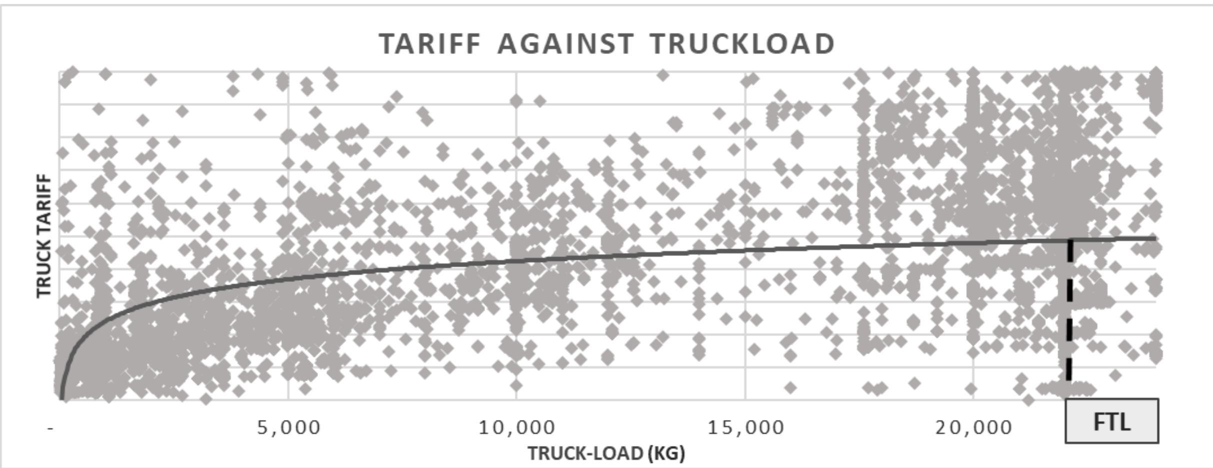

Figure 13. Considering FTL tariff for truckload

From the total number of measurements, there are 5,625 measurements with a truck-load equal or greater than 22,000 kg. These measurements are used to define the FTL tariff function while considering travel distance. As discussed in Section 5.2, it is assumed that the distance is polynomial distributed (order 2) relative to the FTL tariff, resulting in the following function.

$$FTL\ tariff = \beta_0 + \beta_1 * distance + \beta_2 * distance^2$$

In order to find the right coefficients, a regression analysis is conducted with depended variable: truck price, and independent variables: distance and squared distance. There are two possibilities for this regression analysis: (1) regression with intercept (0,0), (2) regression with intercept not (0,0). Comparing the results of these two regression analyses, the one with intercept (0,0) gives better statistical evidence. The coefficients to determine the FTL tariff and the statistical evidence data are as shown in Table 9.

Table 9. Regression results polynomial function FTL tariff

Variable	Coefficient	Std. error	P-value	Lower 95%	Upper 95%
$\boldsymbol{\beta_0}$	0	-	-	-	-
$\boldsymbol{\beta_1}$	2.3352	0.0188	.000	2.2984	2.3720
$\boldsymbol{\beta_2}$	-0.00069	0.00001	.000	-0.00072	-0.00066

The regression analysis results in a R^2 of 0.91, which means that 91 per cent of the variation is explained by the model. Furthermore, both p-values are equal to zero, which investigates a significant model (p-value < .05). This results in the following equation for the FTL tariff.

$$FTL\ tariff = 0 + 2.3352 * distance - 0.00069 * distance^2$$

The objective of the gravity model is to minimize the costs by changing the distance. Since the function has a negative second coefficient, the FTL tariff will decline and even become negative from a certain distance point. Therefore, it is expected that it results in a warehouse location with high distances (resulting negative tariffs). Furthermore, for a minimization problem, the cost structure should be smooth, otherwise it has multiple local minima (solutions). Therefore, it is assumed that the costs will be constant from the point that the function reach its maximum FTL tariff, which is for distances equal or greater than 1690 km. The distribution of the FTL tariff is shown in Figure 14, in which the dotted line represents the original function and the solid line represent the function which is constant from its maximum.

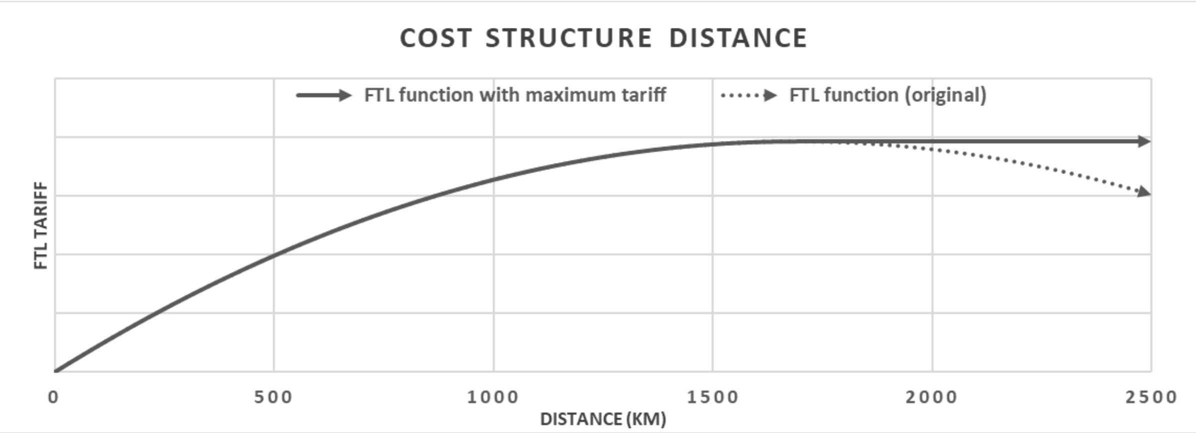

Figure 14. Polynomial cost structure distance (FTL tariff)

5.4 LTL tariff considering truckload

With the function of the FTL tariff, the relation between the truck price and the distance is defined. The other depended variable of the truck price is the truckload, for which the relation relative to the truck price is defined in this section. Therefore, the LTL tariff equation is defined, which is a function of the FTL tariff and a discount rate for LTL shipments. As discussed in Section 5.2, the used equation to determine the LTL tariff is as follows.

$$LTL\ tariff = FTL\ tariff\left(\left(\frac{LTL\ load}{FTL\ load}\right)^{(1-r)}\right)$$

In order to determine the discount factor (r) by a regression analysis, the LTL tariff function requires transformation such that the power function (1-r) becomes a coefficient. This results in the following equation.

$$log\ \left(\frac{LTL\ tariff}{FTL\ tariff}\right) = (1-r) * log\left(\frac{LTL\ load}{FTL\ load}\right)$$

By using $log\left(\frac{LTL\ tariff}{FTL\ tariff}\right)$ as dependent variable, and $log\left(\frac{LTL\ load}{FTL\ load}\right)$ as independent variable, a regression analysis is conducted for all 11,254 measurements (cleaned data) for which the FTL tariff is calculated by using the equation as defined in Section 5.3. The regression analysis with intercept (0,0) results the coefficient (and discount factor) as shown in Table 10.

Table 10. Regression results discount factor LTL tariff

Variable	Coefficient	Std. error	P-value	Lower 95%	Upper 95%
$\mathbf{1-r}$	0.3521	0.0044	.000	0.3435	0.3607
$\boldsymbol{r}$	0.6479	0.0044	.000	0.6565	0.6393

The regression analysis results in a R^2 of 0.36 per cent, wherefore the model does only explain 36 per cent of variation. However, the p-value is equal to zero, which means that the model is significant (p-value < .05). Based on these figures, the following LTL tariff equation is considered in the model with a cost structure as shown in Figure 15.

$$LTL\ tariff = FTL\ tariff\left(\left(\frac{LTL\ load}{FTL\ load}\right)^{(1-0.65)}\right)$$

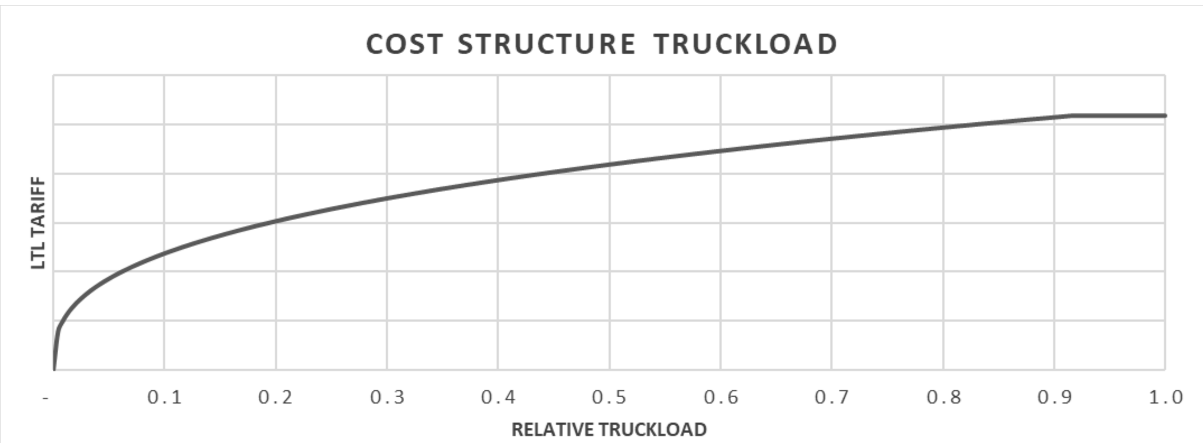

Figure 15. Cost structure truckload (LTL tariff)

Chapter 6. Gravity Study

In this chapter a gravity model with the objective to determine the center of gravity of the distribution network is presented. First, the general model is defined, where after the model is used for the case study of the Company. The gravity model partly answers Sub-research question 2: *What methodology is needed that determines the optimal design of a multimodal facility location problem such that costs are minimized?*

The chapter structures as follows: modelling (Section 1), setting of parameters for case study (Section 2), the results of the case study (Section 3), an adjusted model to find multiple gravity points (Section 4) and a sensitivity analysis to investigate the effect of uncertainty (Section 5).

6.1 Modelling

In the modelling phase, the general mathematical gravity model is defined. First, the required (input) variables for the model are defined. By using these variables, a gravity model is created that determines the center of gravity for a multi-echelon distribution network, i.e. considering inbound and outbound transportation streams.

In this situation, the facility serves as a connecter between the manufacturing sites with the customers with the aim to minimizing transportation costs. Therefore, a set of manufacturing sites and customer zones are geographically dispersed in a region. Therefore, the location data variables in the gravity model are the following:

- M = number of manufacturing sites (inbound origins), also let m = {1, 2, …, M}
- C = number of customer (zones), also let c = {1, 2, …, C}
- x_m = coordinate x (radius) of manufacturing site *m*
- y_m = coordinate y (radius) of manufacturing site *m*
- x_c = coordinate x (radius) of customer zone *c*
- y_c = coordinate y (radius) of customer zone *c*

The objective of the gravity study is to minimize the annual transportation costs. As discussed in Chapter 5, the transportation costs regard an increasing concave down cost structure relative to the travel distance and the truckload. These two relationships are expressed into an FTL tariff function and LTL tariff function as follows:

$$LTL\ tariff = FTL\ tariff\ \left(\frac{Load\ LTL}{Load\ FTL}\right)^{(1-r)}$$

$$FTL\ tariff = \beta_0 + \beta_1 * Distance + \beta_2 * Distance^2$$

Where, r is the discount factor for LTL shipments, β_0 is the intercept of the FTL tariff, β_1 is a positive coefficient per distance unit (order 1), and β_2 is a negative coefficient per squared distance unit (order 2). The Load LTL divided by Load FTL equals the relative truckload, where the Load FTL is the minimum truckload for which an FTL tariff is applicable, i.e. the load from where the truck tariff is constant. Trucks with a load equal or greater than this minimum truckload are considered as FTL. Furthermore, there is a maximum distance considered in the model from where the FTL tariff is constant, in order to negotiate the negative second order coefficient from that point and make the function smooth.

These cost functions results in the following additional data variables for the gravity model.

- Q_m = average truckload between manufacturing site *m* and the facility
- Q_c = average truckload between the facility and customer zone *c*
- r = discount factor for less than truckload
- β_0 = intercept of FTL tariff
- β_1 = positive coefficient of travel distance (order 1)
- β_2 = negative coefficient of squared travel distance (order 2)
- T = minimum truckload for FTL tariff
- S = maximum distance considered in the model

Since the FTL tariff is used for truckloads equal or greater than *T*, there is made use of a proportional function for the relative load (P) with the following interval variables:

$$P_m \begin{cases} \frac{Q_m}{T}, & if\ Q_m < T, \\ 1, & otherwise. \end{cases} \quad \forall\, m \in M, i \in \mathrm{I} \qquad P_c \begin{cases} \frac{Q_c}{T}, & if\ Q_c < T, \\ 1, & otherwise. \end{cases} \quad \forall\, c \in C, i \in$$

The transportation costs for each movement are multiplied with the annual number of movements for a certain lane to get the annual costs. Therefore, the frequency is a model input:

- F_m = annual number of movements between manufacturing site *m* and the facility
- F_c = annual number of movements between the facility and customer zone *c*

The problem of the gravity model is to determine where to locate a facility with the aim to minimize the annual transportation costs. Therefore, the **decision variables** in the model are the gravity location coordinates:

- x_w = coordinate x (radius) of center of gravity *w*
- y_w = coordinate y (radius) of center of gravity *w*

For both variables, a maximum and minimum range will be defined:

- (minx, maxx) = x coordinate (radius) range of area in which facility can be located
- (miny, maxy) = y coordinate (radius) range of area in which facility can be located

To decide on the center of gravity it is required to translate two sets of coordinates to a distance (in km). This is defined as a Haversine function (Robusto, 1957) in the gravity model, where,

- D_m = travel distance between manufacturing site *m* and the facility
- D_c = travel distance between the facility and customer zone *c*
- e = earth radius
- rf = road factor (circuity factor)

Based on this maximum distance, a function for the considered distance is defined by:

- K_m = considered travel distance between manufacturing site *m* and the facility
- K_c = considered travel distance between the facility and customer zone *c*

The considered distance is a function of the real-distance and the following interval variables.

$$a_m \begin{cases} 1, & if\ D_m \in (0, S] \\ 0, & otherwise. \end{cases} \quad \forall\, m \in M, i \in \mathrm{I} \qquad a_c \begin{cases} 1, & if\ D_c \in (0, S] \\ 0, & otherwise. \end{cases} \quad \forall\, c \in C, i \in \mathrm{I}$$

$$b_m \begin{cases} 1, & if\ D_m\ \in\ (S,\infty] \\ 0, & otherwise. \end{cases} \quad \forall\, m \in M, i \in \mathrm{I} \qquad b_c \begin{cases} 1, & if\ D_c\ \in\ (S,\infty] \\ 0, & otherwise. \end{cases} \quad \forall\, c \in C, i \in \mathrm{I}$$

Based on the concave cost structures, the Gravity Model, denoted as Problem 1, can be formulated as the following Nonlinear program.

$$Min \sum_{m=1}^{M} \left(FTL_m \left({P_m}^{(1-r)}\right)\right) F_m + \sum_{c=1}^{C} \left(FTL_c \left({P_c}^{(1-r)}\right)\right) F_c$$

Subject to:

$$D_m = \operatorname{acos}(\sin(x_m)\sin(x_w) + \cos(x_m)\cos(x_w)\cos(y_m - y_w))\, e * rf \qquad \forall\, m \in M \quad (1.1)$$

$$D_c = \operatorname{acos}(\sin(x_c)\sin(x_w) + \cos(x_c)\cos(x_w)\cos(y_c - y_w))\, e * rf \qquad \forall\, c \in C \quad (1.2)$$

$$K_m = a_m D_m + b_m S \qquad \forall\, m \in M \quad (1.3)$$

$$K_c = a_c D_c + b_c S \qquad \forall\, c \in C \quad (1.4)$$

$$a_m, b_m, a_c, b_c\ \in \{0,1\} \qquad \forall\, m \in M, c \in \mathrm{C} \quad (1.5)$$

$$a_m + b_m = 1 \qquad \forall\, m \in M \quad (1.6)$$

$$a_c + b_c = 1 \qquad \forall\, c \in C \quad (1.7)$$

$$FTL_m = \beta_0 + \beta_1 K_m + \beta_2 (K_m)^2 \qquad \forall\, m \in M \quad (1.8)$$

$$FTL_c = \beta_0 + \beta_1 K_c + \beta_2 (K_c)^2 \qquad \forall\, c \in C \quad (1.9)$$

$$K_m, K_c \le S \qquad \forall\, m \in M, c \in C \quad (1.10)$$

$$P_m, P_c\ \in (0,1) \qquad \forall\, m \in M, c \in C \quad (1.11)$$

$$D_m, D_c \ge\ 0 \qquad \forall\, m \in M, c \in \mathrm{C} \quad (1.12)$$

$$x_w\ \in (minx, maxx) \quad (1.13)$$

$$y_w\ \in (miny, maxy) \quad (1.14)$$

The objective of the gravity location model is presented by the minimization equation, where the first part equals the inbound transportation costs, and the second part equals the outbound transportation costs. Constraints (1) and (2) are used to determine the travel distance of each stream, wherefore the considered distance (based on the maximum distance) is determined by constraint (3) and (4) with two integer variables. These integer variables are binary (5) and just one of both is selected for a certain stream, this is defined by constraint (6) and (7). Based on the considered distances and the distance coefficients for the polynomial concave cost structure, the FTL tariff per stream is determined by equation (8) and (9). Furthermore, the restriction of the considered distance is the maximum distance as denotes in constraint (10). The proportional function can lay between interval (0,1) as defined in equation (11). Furthermore, it is known that the distances (decision variables) are non-negatives see constraint (12). At last, in order to decide on the location (x and y coordinates) a minimum and maximum interval is considered based on the range of the area, as defined in constraint (13) and (14).

6.2 Parameter setting

In order to implement the Gravity Model for the distribution network of the Company, the parameters out of the model are defined in this section. At first, the locations in the distribution network (manufacturing sites and customers) are aggregated. Secondly, based on the defined transportation rate function, parameters for the costs are set. At last, assumptions are made for the remaining parameters.

6.2.1 Aggregation of locations

There are two type of locations considered in the Gravity model: (1) Manufacturing sites, and (2) customers. As discussed in the conceptual model, the manufacturing sites are aggregated on unloading port and the customers are aggregated based on the Zip-Code (2 digits) and if applicable on loading port. In total there are 20 manufacturing sites, however, the majority is located outside Europe and these origins are changed to one point: the unloading port (Rotterdam), which results in eight different points. Though, in the model the port points will still be considered separately, because of the difference in truckload and its effect on the LTL prices. Therefore, there are twenty manufacturing sites considered in the model.

- M = 20, also let m = {1, 20}

In the current distribution network, there are 1345 customers, which requires aggregation to customer zones. As discussed in the conceptual model, it is decided to aggregate the customers on the first two digits of the zip-code in combination with the country code. This results in 666 customer zones. Furthermore, goods that are for customers from overseas countries (outside Europe) are changed to one destination point: the loading port (Rotterdam). This results in 525 different customer points. However, in the model the port points will still be considered separately, because of the difference in truckload and its effect on the LTL prices. As a result, there will be 666 customer zones considered in the model.

- C = 666, also let c = {1, 666}

6.2.2 Transportation rates

The parameters regarding the transportation rates are discussed in Chapter 5. Based on this chapter, the cost parameters are defined in this section. At first, the FTL tariff is applicable for truckloads equal or greater than 22,000 kg. Secondly, it is known that the FTL-tariff follows the following polynomial concave cost structure.

$$FTL\ tariff = 0 + 2.3352 * Distance - 0.00069 * Distance^2$$

Furthermore, it is shown that the maximum distance (the point at which the function reaches its maximum FTL tariff) is 1690 km. At last, a regression analysis shows that the discount factor for LTL tariff is 0.65. Therefore, the parameter values are as follows:

- T = 22,000 kg
- β_0 = 0
- β_1 = 2.3352
- β_2 = -0.00069
- S = 1690 km
- r = 0.65

6.2.3 Remaining parameter values

As discussed in the conceptual model, the gravity model is used to determine the optimal location against the same performances as in the current model. Therefore, the truckload per movement and the number of movements per lane (frequency) are equal to the analyzed data.

- F_m and F_c are equal to analyzed numbers $\forall\, m \in M, c \in C$
- Q_m and Q_c are equal to analyzed numbers $\forall\, m \in M, c \in C$

Furthermore, for the coordinates of the manufacturing sites and the customer zones the latitude and longitude is known, where x represents the longitude, and y represents the latitude. However, these values are in degrees, and the model calculates the distance with radius, therefore the coordinates are translated to radius by using the following equations.

$$coordinate\ (radius) = \frac{coordinate(degrees)}{180°}\ \pi$$

In addition, the area in which a facility can be located is globally[12]. Therefore, the decision area is worldwide, which results in the following feasible coordinate ranges.

- minx = -3.15, and maxx = 3.15
- miny = -1.58, and maxy = 1.58

In order to calculate the Euclidian distances in km, the earth radius is required, which is 6,378 km. Furthermore, a road factor is multiplied by this Euclidian distance to determine the actual travel distance. According to (Ballou, Rahardja, & Sakai, 2002), the Europe road factor is used.

- e = 6,378 km
- rf = 1.46

6.3 Solution Gravity Study

In this section, the results of the Gravity Model (Section 6.1), by using the case study parameters (Section 6.2), are shown and discussed. The Gravity Model is solved by using the solver function in Excel with the Generalized Reduced Gradient (GRG) Nonlinear solving method. This method looks at the slope of the objective function as input values, change and determines that it reached an optimum solution when the partial derivates equal zero. A disadvantage of this method is that the solution is highly depending on the initial conditions and therefore may results a local solution. In order to obtain the global minimum, the multi-start function is used with population size 1,000. Therefore, the algorithm creates a randomly distributed population of initial values that are each evaluated using the traditional GRG Nonlinear algorithm. In other words, it starts multiple times from different initial conditions within the range of the decision variable, which increases the probability that the solution found is the global minimum.

Based on this heuristic, solving Problem 1 results in the following solution:

- y_w of 0. 90553 (radius), which is 51.8832 (degrees)
- x_w of 0. 07428 (radius), which is 4.2560 (degrees)

[12] The valid range of latitude in degrees is -90 and +90 for southern and northern hemisphere respectively. Longitude is in the range -180 and +180 specifying coordinates west and east of the Prime Meridian, respectively. These figures in degrees are translated to radius (and rounded on two decimals) to use in the model.

By using these coordinates, it is found that the optimal location for a consolidated warehouse (considering inbound and outbound, equal performance as in current situation, and only truck-transportation) is Rotterdam (port). This solution makes sense, since the majority of the inbound streams is starting from this point and it is the end point for a part of the outbound streams. For these streams the modelled transportation costs amount $0 USD. The objective function results the minimum transportation costs for the distribution network, which is $15.3 million USD. From this total amount, $1.6 million USD regards the inbound transportation costs, and the remaining $13.7 million USD regards the outbound costs. Comparing to the current transportation costs of $18 million USD, a consolidation of nine warehouses to one warehouse in Rotterdam port seems to reduce the annual transportation costs with $2.6 million USD. However, in the optimal situation it is modelled that there is no transportation cost for products inbounded or outbounded from Rotterdam port, which differ from reality, since there is always a shuttling cost. These shuttling costs[13] needs to be subtracted from the reduction of $2.6 million according to the model. Furthermore, the current performances are considered in the gravity model and therefore not all potential improvements are implemented yet. These cost reduction opportunities are further researched in the Facility Location Model (Chapter 7). In addition, direct cost reductions of consolidating are the intern transportation costs ($0.2 million USD) and costs made for order processing activities, which are not counted in the model.

The requirement for a feasible location is that the location has multi-modal transportation network, which means connected to road, rail and inland waterway. It can be concluded that Rotterdam (port) is connected to all of the different modes. Therefore, it is concluded that the founded gravity location is feasible for the model.

6.4 Gravity model for multiple facilities

The location at the gravity point, serves as distribution center (DC) for EMEA, i.e. location where the Company concentrated the logistics of goods. All manufacturing sites deliver to this distribution center. Therefore, all the goods are served by the distribution center. It is required that the distribution center can perform the value adding services, such as re-packing and re-labelling. From the DC the goods are transported to the customers, which can be done directly or through a cross docking facility (CDF). A CDF is deployed as a location in which products are moved from one transportation mode to another, with minimal warehousing. The advantage of these facilities is the opportunity to combine transportation from a DC to a CDF, for which there is a high probability that the truckload will increase and becomes more profitable to make use of different transportation modes. Therefore, the use of higher truckloads and different modes (rail and inland waterways instead of road) might reduce the transportation costs. Furthermore, this reduction in movements and use of transportation modes that produce less emissions leads to a more sustainable distribution network, which correspond the side aim of the project. In this section, it is researched if these CDFs can be beneficial in the distribution network of the Company. In order to find the optimal CDF locations(s), the single-facility Gravity Study (Problem 1) from Section 6.2 is followed up by a multi-facility Gravity study.

[13] The shuttling costs per movement can be multiplied with 4,604 annual inbound movements, and 2,134 annual outbound movements to/from the Rotterdam port. For example, when the shuttling costs amount $150 USD per movement (average of current shuttling costs Rotterdam port), the annual shuttling costs results in an amount of $1.0 million USD.

6.4.1 Multi-facility Gravity Model

The multi-facility Gravity Model is a follow up on the single-facility gravity model. Therefore, the single-facility gravity model is adjusted in order to determine the optimal location of CDF(s) by considering the outbound transportation streams and the gravity location in Rotterdam (port) as DC location. The DC and CDF(s) together form the number of facilities denoted by W.

- W = number of facilities, also let w = {1, 2, …, W}
  Where, w = 1 denotes the DC and w ∈{2, W} denotes the CDFs

For the inbound process it is known that everything is transported to the DC; therefore, the inbound is negotiated in this adjusted gravity model. The considered costs are the outbound costs: from the DC to a customer or from determined CDF(s) to a customer. Each customer is allocated to one facility, this allocation is done by using the following binary decision variable.

$$X_{wc} \begin{cases} 1, & if\ customer\ is\ served\ from\ facility\ w \\ 0, & otherwise. \end{cases} \qquad \forall\, c \in C,\ w \in W$$

Based on the allocation, the transportation costs per outbound stream are calculated. Therefore, the transportation cost function as discussed in Chapter 5 is used (FTL tariff and LTL tariff). The sum of these outbound transportation costs amounts the objective value, i.e. the annual cost for serving all customers from a DC or CDF. The additional transportation of goods from DC to CDF is not considered in this model, since these costs will change while combining goods (i.e. improved performances) and using different transportation modes. These costs are added later in the project, namely in the Facility Location Model (Chapter 7).

Based on the adjustment, the Gravity Model for multiple facilities, denoted as Problem 2, can be formulated as the following Mixed Integer Nonlinear program.

$$Min \sum_{c=1}^{C} \left(FTL_c \left(P_c^{\,(1-r)}\right)\right) F_c$$

Subject to:

$$D_{wc} = \operatorname{acos}(\sin(x_c)\sin(x_w) + \cos(x_c)\cos(x_w)\cos(y_c - y_w))\ e * rf \qquad \forall\, c \in C,\ w \in W \qquad (2.1)$$

$$\textstyle\sum_{w=1}^{W} X_{wc} = 1 \qquad \forall\, c \in C \qquad (2.2)$$

$$D_c = \textstyle\sum_{w=1}^{W} D_{wc} X_{wc} \qquad \forall\, c \in C \qquad (2.3)$$

$$K_c = a_c D_c + b_c S \qquad \forall\, c \in C \qquad (2.4)$$

$$a_c, b_c \in \{0,1\} \qquad \forall\, c \in C \qquad (2.5)$$

$$a_c + b_c = 1 \qquad \forall\, c \in C \qquad (2.6)$$

$$FTL_c = \beta_0 + \beta_1 K_c + \beta_2 (K_c)^2 \qquad \forall\, c \in C \qquad (2.7)$$

$$K_c \le S \qquad \forall\, c \in C \qquad (2.8)$$

$$P_c \in (0,1) \qquad \forall\, c \in C \qquad (2.9)$$

$$X_{wc} \in \{0,1\} \qquad \forall\, c \in C,\ w \in W \qquad (2.10)$$

$$D_c \ge 0 \qquad \forall\, c \in C \qquad (2.11)$$

$$x_w = gravityx \qquad \forall\, w = 1 \qquad (2.12)$$
$$y_w = gravityy \qquad \forall\, w = 1 \qquad (2.13)$$
$$x_w \in (minx, maxx) \qquad \forall\, w \in \{2, W\} \qquad (2.14)$$
$$y_w \in (miny, maxy) \qquad \forall\, w \in \{2, W\} \qquad (2.15)$$

The objective of the gravity location model is presented by the minimization equation, which equals the outbound transportation costs for delivery from DC or CDF to the customer. Constraints (1) is used to determine the travel distance between a facility and a customer. Based on constraint (2) a customer is allocated to one facility, resulting in one selected distance per customer, as determined by constraint (3). The maximum distance is still applicable in this model, wherefore the considered distance is determined by constraint (4) with two integer variables. These integer variables are binary (5) and just one of both is selected for a certain stream, this is defined by constraint (6). Based on the considered distances and the distance coefficients for the polynomial concave cost structure, the FTL tariff per stream is determined by equation (7). Furthermore, the restriction of the considered distance is the maximum distance as denoted in constraint (8). The proportional function can lay between interval (0,1) as defined in equation (9). In addition, the allocation variable is binary (constraint 10) and it is known that the distances are non-negatives (constraint 11). Based on Problem 1, it is already decided that the DC is located at the gravity point, with coordinates (*gravityx, gravityy*) as can be seen in constraint (12) and (13). In order to decide on the location (x and y coordinates) of the CDF(s) a minimum and maximum interval is considered, as defined in constraint (14) and (15).

6.4.2 Parameter setting

The majority of the parameters is equal to the used values in Problem 1, for these parameters is referred to *Section 6.2 Parameter Setting*. The parameters that are changed or added to Problem 2 are as discussed in this section.

First, an added parameter is the number of facilities. In order to find different solutions, the W is adjusted for each run. For the first run, the number starts at two, wherefore only the DC and one CDF are considered in the model. In order to find the multiple gravity points, the number is adjusted per solving (+1) until it gives repeating or non-improving results. Secondly, the coordinates of the DC are added in this model. Based on the solution of Problem 1, the DC coordinates are as follows:

- gravityx = 0.90553 (radius)
- gravityy = 0.07428 (radius)

6.4.3 Solution

The Gravity Model for multiple facilities is solved by using the GRG-Nonlinear solving method in Excel including the multi-start function with population size 1,000. As discussed above, the model is solved multiple times by adjusting the number of facilities (input variable W). This results in the solutions as shown in Table 11. As can be seen, for four facilities, the model results in similar locations as for three facilities, and just 10 per cent of the demand is served by one facility. Furthermore, it is visible that the cost reduction of adding a facility becomes smaller. Therefore, it is decided to not solve the model for more than four facilities.

Table 11. Results of Gravity Model with multiple facilities

Objective value (USD)	1 facility	2 facilities	3 facilities	4 facilities
$13.7 million	Rotterdam port (NL) *100% demand*			
$11.1 million	Rotterdam port (NL) *62% demand*	Puchov (SK) *38% demand*		
$9.1 million	Rotterdam port (NL) *44% demand*	Cairate (IT) *25% demand*	Bielsko-Biala (PL) *32% demand*	
$8.2 million	Rotterdam port (NL) *44% demand*	Cairate (IT) *25% demand*	Gliwice (PL) *22% demand*	Campina (RO) *10% demand*

In all the situation, the DC is located in Rotterdam. When adding one CDF, this facility is optimally located in Puchov (Slovakia). For which 62 per cent of the demand is served by the DC and 38 per cent by the CDF. This results in a difference of $2.6 million USD of outbound transportation costs. For a distribution network with 2 CDFs, the facilities are located in Cairate (Italy) and Bielsko-Biala (Poland), which results in a difference of $4.6 million USD comparing to one facility (one DC and no CDF). For a distribution network with 3 CDFs, the locations are in Cairate (Italy), Gliwice (Poland), and Campina (Romania), resulting in a difference of $5.5 million USD compared to one facility. These cost differences should be compared with the changed costs for transporting goods from DC to CDF(s) to determine the optimal situation.

6.4.4 Feasibility of gravity points

A multi-modal transportation network of the facilities might have transportation and sustainable benefits. Therefore, the requirement for a feasible location is that the location has a multi-modal network, i.e. connected to road, rail and inland waterway. Except the gravity point at Rotterdam Port, none of the locations are connected to the three modes. Therefore, the surrounding of the gravity points is researched on alternative locations with an extended network to multi-modality. Based on a maximum error of 5 per cent from the objective value (transport costs per facility), a radius is selected as shown in Figure 16, which determines the gravity region for each gravity point. Therefore, the model gives that all possible facility locations within the following (Euclidian) radius are in scope: for Puchov (SK) 216 km, for Cairate (IT) 131 km, for Bielsko-Biala (PL) 204 km, for Gliwice (PL) 153 km, and for Campina (RO) 40 km.

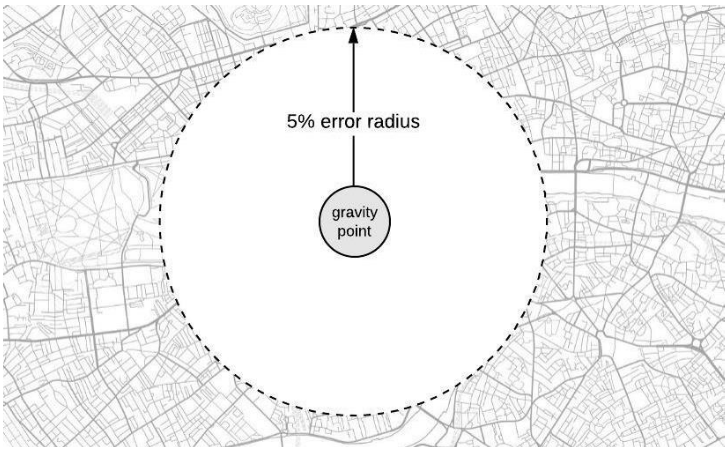

Figure 16. Example gravity region

6.5 Sensitivity analysis

A sensitivity analysis determines how the uncertainty in the output of a mathematical model can be divided and allocated to different sources of uncertainty in its inputs. During this sensitivity analysis, the outcomes of the model are recalculated under alternative assumptions in order to determine the impact of variability. In this section, the parameter setting sensitivity of the following parameters: the road factor, the discount factor, and the (squared)distance cost coefficients, are considered.

6.5.1 Road factor

The road factor used in the gravity study, is the average road factor for Europe. However, the road availability differs per country and region. Therefore, it is investigated what the effect is when lowering or raising this factor. According to (Ballou, Rahardja, & Sakai, 2002), the country within Europe with the lowest road factor is Italy (1.18) and the highest is France (1.65). These two values are set as lower and upper bound parameters.

- Original: 1.46 , for Europe
- Lower bound: 1.18 , for Italy
- Upper bound: 1.65 , for France

Based on the analysis, a lower road factor results in lower travel distances and lower tariffs per movement. For a higher road factor, it is the other way around, i.e. results in higher tariffs per movement. Therefore, the objective value raises for the upper bound and lowering for the lower bound road factor. When solving the Gravity Models (Problem 1 and 2) with the new parameters, it is founded that the decided locations are similar to original. The DC results in each situation exactly the same location, and the CDF(s) are similar to original with a maximum deviation of 210 km for four facilities, 23 km for three facilities, and <1 km for two facilities.

6.5.2 Discount factor

The discount factor considers the effect of the truckload on the price per movement. This factor is determined by using a regression analysis of the LTL-tariff (observations) against the FTL-tariff and discount factor. Based on this regression, the discount factor is obtained to be 0.65. In this sensitivity analysis it is considered how a lower or higher discount factor affects the price and the solution of the gravity study. Therefore, the performances of a lower bound (r = 0.3) and upper bound (r = 0.8) are investigated. The effect of these changing parameters on the LTL tariff is as shown in Table 11.

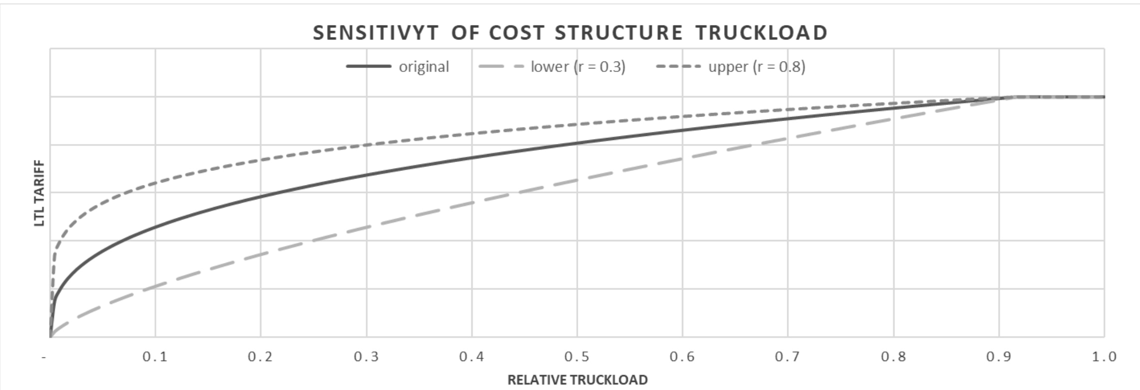

Figure 17. Sensitivity of discount factor for LTL tariff

As can be seen, the lower bound results in lower costs for LTL shipments. Therefore, the objective value is also lower than original. Furthermore, it can be seen that the upper bound results in higher costs for LTL shipments. As a result, the objective value is higher than original. Therefore, it can be concluded that the discount factor affects the objective value. However, while solving the Gravity Models the decision variables (x and y coordinates) are for both situations, upper and lower bound, equal to the result with original value (deviation < 1km) wherefore can be concluded that the decision variables are not affected by the discount factor.

6.5.3 Distance cost coefficients

For the FTL tariff, the effect of the distance cost structure is investigated. During analysis, it is found that the function regards a polynomial concave cost structure with intercept (0,0), i.e. the original values. The FTL tariff coefficients are determined by a regression analysis, which also determines the 95 per cent lower and upper bound. Therefore, the effect of this lower and upper bound on the gravity model are investigated. Furthermore, it is possible to use a polynomial concave cost structure with intercept not fixed at (0,0), for which the function is determined by using a regression analyse. In addition, various literature assume that the distance cost function is linear. Therefore, it is investigated what the effect is of a linear distance function. This results in four alternative distance coefficients for FTL tariff as shown in Table 12.

Table 12. Sensitivity of distance coefficients for FTL tariff

Distribution function:	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$
Polynomial concave (original)	-	2.3352	(0.00069)
Lower bound (95%)	-	2.2984	(0.00072)
Upper bound (95%)	-	2.3720	(0.00066)
Linear function	-	1.5061	-
Polynomial concave with intercept	203	1.7880	(0.00041)

Based on the changed parameters, the lower bound results a lower FTL tariff, wherefore the objective value decreases, and the upper bound results a higher FTL tariff, wherefore the objective value increases. Solving Problem 1 shows that there is no effect of the distance-price function on the gravity center, which results in each situation the same decision variables as in the original situation. However, solving Problem 2 (determining the CDF(s)) shows a variation in the decision variables. For the lower and upper bound (95 per cent) there is no to minimal effect (deviates < 3km). When analyzing a polynomial concave cost structure with intercept not at (0,0) it is found that there is a minimal effect, since it deviates < 20km from the original CDF(s) locations. Furthermore, while considering a linear function, the deviation is largest, namely maximum of 60 km for three and four facilities, and 17 km for two facilities. In addition, the objective value for both functions, linear and polynomial with intercept not at (0,0) is fluctuating (higher and lower) comparing to the original values.

6.5.4 Conclusion sensitivity analysis

Based on the sensitivity analysis, it is concluded that the parameters affect the objective value. In most cases, the lower bounds of a variable results in lower costs than original, and for the upper bound the other way around. Figure 18 shows the objective values (annual outbound costs) while changing parameters in four scenarios: 1, 2, 3, or 4 facilities to be opened.

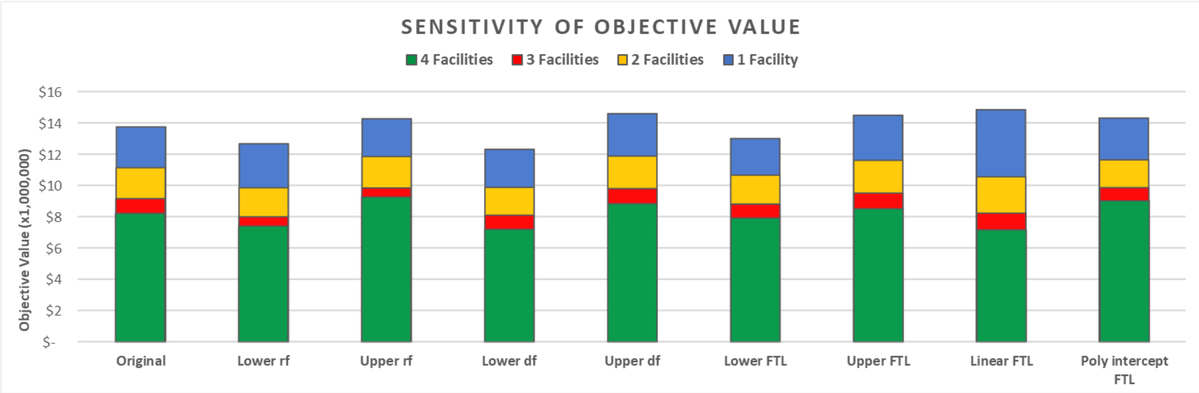

Figure 18. Sensitivity of objective value

As can be seen in the figure, the objective value is affected by the changed parameters. The parameter setting with the smallest effect are the lower and upper bound of the FTL tariff. Therefore, it can be concluded that with 95 per cent reliability the costs for the FTL tariff (with the used polynomial concave cost structure) are within these bound objective values. The parameter changing with the highest effect are for a lower objective value: lowering discount factor, and for a higher objective value: using linear distance function for FTL tariff.

The additional aim is to investigate the sensitivity of the solution, i.e. the decision variables. Therefore, Figure 19 shows the different determined facility locations by solving the gravity models against changed parameters. Based on this figure, it can be concluded that the solutions are robust. However, the higher the number of facilities, the higher the deviation.

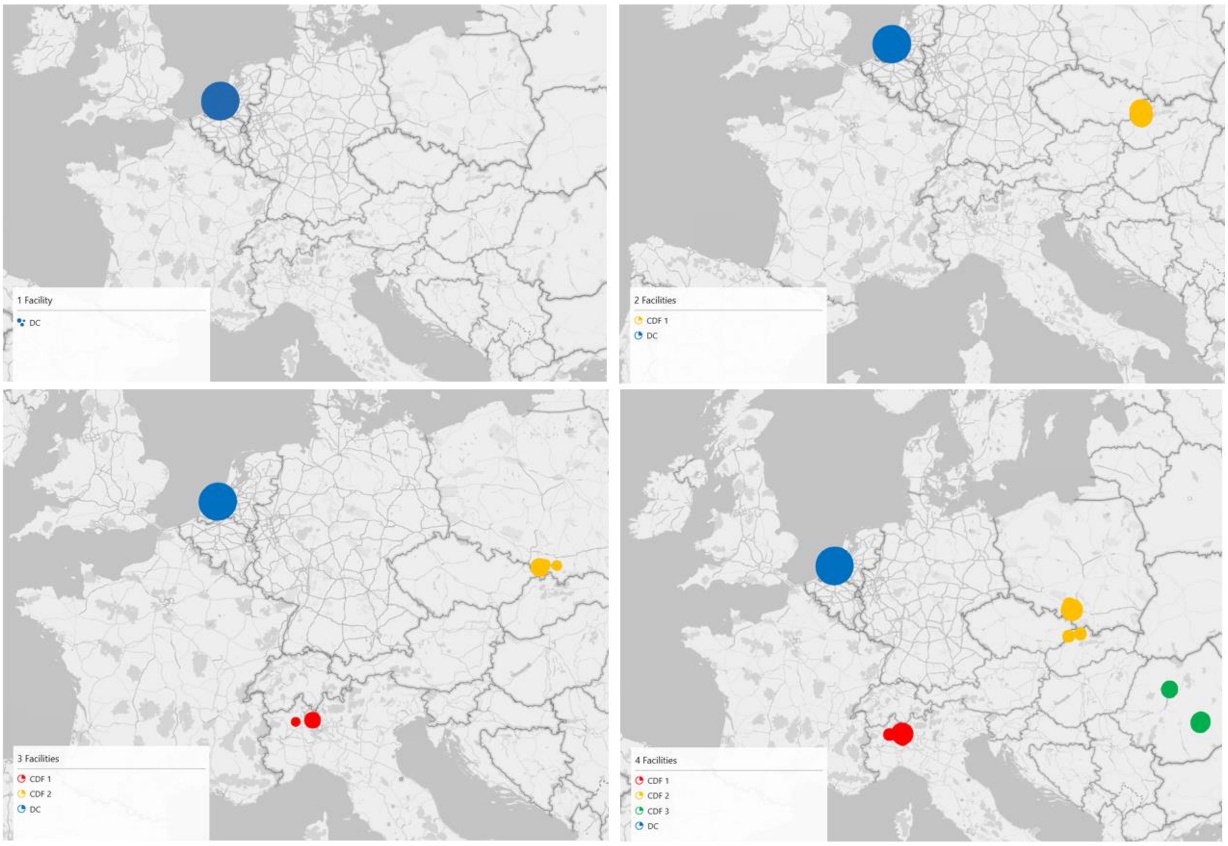

Figure 19. Sensitivity of gravity points

Chapter 7. Multimodal Facility Location

In this chapter a facility location model is presented with the objective to determine the optimal set of facilities to open and to decide the multi-modality of the network. With this Multimodal Facility Location Model, the final answer is given on Sub-research question 2: *What methodology is needed that determines the optimal design of a multimodal facility location problem such that costs are minimized?* Furthermore, Sub-research question 3: *How should the Company implement the distribution network redesign based on the optimal design methodology considering different scenarios?* is partly answered by considering different customer service level scenarios. The chapter structures as follows: assumptions for the model (Section 1), the modelling phase (Section 2), setting of parameters for the case study (Section 3), the results of the case study (Section 4), and a sensitivity analysis to investigate the effect of uncertainty in demand (Section 5).

7.1 Model Assumptions

In this section, some important input variable assumptions of the model are explained. First, it is decided what the possible facility locations are considering the multi-modal opportunities of the selected gravity regions. Furthermore, customer service is added as an important aspect for this model, since the performances of the network (frequency and truckload) are depending on it. In addition, the facility location model considers inventory costs.

7.1.1 Facility locations

Based on the solution of the gravity study, it is decided what the possible locations are for a facility. First, it is investigated what the feasibility of the founded gravity points is, i.e. what are the multi-modal possibilities for the locations. When a gravity point is not multi-modal, there is searched for locations within the gravity region that has an extended network to multi-modality, i.e. connected to at least two out of the three possible modes for freight transportation: road, rail according to (Rail Net Europe, 2018) and Inland Waterways according to (UENCE, 2018). Furthermore, it is decided to only select locations within a logistic concentration area.

- *Facility 1 (DC)*: The gravity study results the port of Rotterdam as the optimal location for the DC. Investigating the modality possibilities of this location, it is found that the gravity point is connected to all three modalities: road, rail and water. Therefore, the first facility is fixed at the Rotterdam port.
- *Facility 2 (CDF1)*: When analyzing a distribution network with two facilities, it is founded that the second facility is optimally located at Puchov in Slovakia, which is only connected to road for freight transport. Further research found that there are two close locations in Slovakia with freight rail connection: Zilina and Leopoldov. Furthermore, in the gravity region (radius of 216 km) is searched for locations connected to all modes, which is proposed to be Bratislava (Slovakia), Vienna (Austria), and Budapest (Hungary). When analyzing a distribution network with three facilities, it is founded that the second location is optimally located at Bielsko-Biala in Poland, this location has a connection to road and rail. However, it is a small city wherefore the rail connection is not applicable for freight transportation. When analyzing 4 facilities, the second facility is optimally located at Gliwice in Poland. This location is three-modal connected and within the gravity region of Bielsko-Biala. However, there is a small

harbor and therefore limited. Therefore, the gravity region for both, 205 km radius for Bielsko-Biala and 153 km for Gliwice, is researched for better alternatives. The closest location for a three-modal connection is Wroclaw in Poland, which is 204 km away from Bielsko-Biala and 145 km from Gliwice and therefore feasible.

- *Facility 3 (CDF2)*: The third location is optimally located at Cairate in Italy, which is only connected to road. The locations within the gravity region (131 km radius) that are connected to road and rail for freight transportation are: Milan and Novara in Italy. Because of the mountain landscape in this region, the closest location with a three-modal connection is Basel in Switzerland which is located 232 km from the gravity point and therefore not within the gravity region. Hence, there is no three-modal connection feasible for the third facility (CDF2).
- *Facility 4 (CDF3)*: The fourth location is Campina in Romania, which is only connected to road for freight transportation. A road and rail connected location within the gravity region (radius of 40 km) is Ploiesti in Romania. The closest alternative location that is connected to all possible modes is Ruse in Bulgaria. However, this location is not within the gravity region and therefore there is no three-modal connection feasible for the fourth facility (CDF3).

The defined locations compose a set of possible facilities to open as input variable for the model for each alternative distribution network, i.e. one, two, three or four facilities. The overview of the location and multi-modal options are as shown in Table 13.

Table 13. Multi-modality of possible facility locations

Distribution network	Facility	w	Location options	Road	Rail	Water
1, 2, 3 and 4 Facility	DC	1	Rotterdam (NL)	■	■	■
2 Facilities	CDF1	2	Puchov (SK)	■		
		3	Zilina (SK)	■	■	
		4	Leopoldov (SK)	■	■	
		5	Bratislava (SK)	■	■	■
		6	Vienna (AT)	■	■	■
		7	Budapest (HU)	■	■	■
3 and 4 Facilities	CDF1	8	Bielsko-Biala (PL)	■		
		9	Gliwice (PL)	■	■	
		10	Wroclaw (PL)	■	■	■
3 and 4 Facilities	CDF2	11	Cairate (IT)	■		
		12	Milan (IT)	■	■	
		13	Novara (IR)	■	■	
4 Facilities	CDF3	14	Campina (RO)	■		
		15	Ploiesti (RO)	■	■	

Since the network regards outsourced activities, it is assumed that the facilities do not have a capacity constrain. Therefore, it is not needed to consider capacity in the allocation of customers to a facility. In order to prevent planning complexity, it is decided to allocate all demand of a customer to one facility, which is the closest open facility (DC or CDF). Therefore, it is still possible to direct serve customers from the DC, when this is the closest open facility.

7.1.2 Network

As discussed in the conceptual model, different transportation modes can be used between the DC and CDF locations and the transportation of goods from manufacturing site and to the customer are still performed by truck (as in current situation). A visualization of this network is shown in Figure 20.

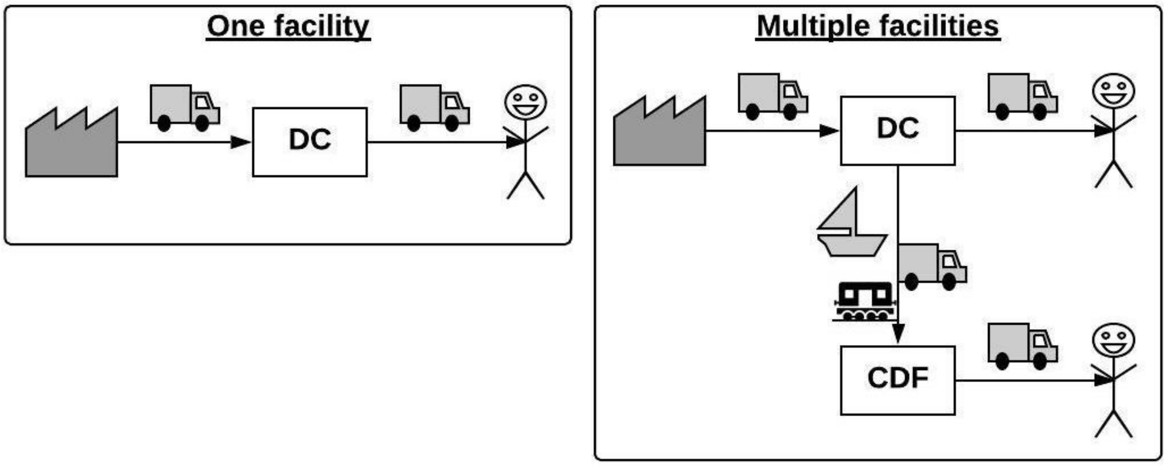

Figure 20. Multi-modality of alternative distribution network

Since the transportation of goods by train or barge seems likely to be cheaper than by road, the extension with multi-modality gives improvement opportunities. Furthermore, it becomes possible to combine transportations of different customers (allocated to the same CDF) for a major part of the transport process, which results in less movements and higher loading factors. In order to combine transportation of different customers between DC and CDF, it is decided to divide the outbound streams into two streams. This results in a Multi Echelon network with total three streams: (1) inbound of goods from manufacturing site to the DC, (2) outbound of goods from DC to a CDF, and (3) outbound of goods from CDF to customer. The goods that needs to be transported between the DC and a CDF equals the total demand during a certain time period of all customers allocated to this CDF.

7.1.3 Customer service

Based on the analysis of the current situation, it is concluded that the distance and the truckload affect the transportation costs. Therefore, the gravity study focused on reducing the total distribution network distance. In order to further optimize the distribution network, it is important to investigate the improvement opportunities related to an increased truckload and therefore a reduced number of movements.

According to the literature study, several researches presented methods to solve the truck-loading problem. The main way to improve the truckload, is to change the transportations from pull to push based on requirements of the system, which is in this situation the lead-time (daily delivery, weekly delivery etc.). In the studies, the aim is to maximize the transportation lead-time, subject to a capacity constraint of the selected transportation mode. While maximizing the lead time, the number of movements can be reduced, and the truckload can be increased.

Various studies researched the truck-loading problem as a separate problem. Based on the literature study, it is investigated that an interesting topic for future research is how to include the truckload as an aspect in a facility location problem, i.e. including transportation lead-time as a variable. Therefore, the truck-loading problem is included into the facility location model. This extension is possible, since the problem minimizes the transportation costs, which are a function of the transportation rates (depending on the truckload) and the number of movements and therefore a maximized truckload results in a reduction of the costs.

Furthermore, it seems likely that the transportation of goods by rail or barge are cheaper than truck. However, the transportation lead-time is higher. Therefore, it is only possible to use a cheaper transportation mode when the service lead-time allows this. When a cheaper option is possible in the given service lead-time, this transportation mode is chosen for all transportations between that DC and CDF.

For the model, the service lead-time will be an input variable, in which the truckload and number of movements are optimized based on this given lead-time. Therefore, different service lead-times might result in different solutions that are analyzed as different scenarios. As a result, the model investigates the effect of the service lead-time on the distribution network costs (objective value) and on the facilities to open and transportation modes to select (decision variables). Based on the model results, the Company can select the scenario (service lead-time) that fits best in reality.

7.1.4 Inventory costs

The distribution network costs consist of transportation costs and costs for handling and holding the goods at facilities. At the DC, goods are not allocated to a customer yet and therefore the goods that are stored is inventory. However, the goods stored at the CDF are already sold to a customer and only stored for a short time-period in the CDF to change from transportation mode. Therefore, these storing of goods serves as buffer in between two transportation streams, but it is assumed that the handling and holding costs are equal to the inventory costs.

The inventory level differs per storage location (for multiple facilities) and is a function of the throughput and cycle time according to Little’s Law:

$$WIP = TH * CT$$

Where, WIP is the work in progress (i.e. the inventory at a facility), TH is the throughput (i.e. the outbound streams for a facility), and CT the average cycle time that goods spend in the warehouse. Based on the cycle time, the number of cycles a year can be calculated, i.e. the annual inventory turnover rate. By using Little’s Law, the inventory (buffer) level per facility is determined. This inventory level is multiplied with the annual inventory holding costs, which differs per product group and might differ between regions. Therefore, it is important to consider different product groups and different costs per facility, which are depending on the average labor costs of that region.

7.2 Modelling

In the modelling phase, the general mathematical multimodal facility location model is defined. First, the required (input) variables for the model are defined. By using these variables, a problem is created that decides what facilities to open and what transportation mode(s) to use between DC and CDF (if multiple facilities opened). The objective of the facility location model is to minimize the annual distribution network costs, which consist of four different cost types: (1) inbound costs from manufacturing site to DC, (2) the outbound costs from DC to CDF, (3) the outbound costs from DC or CDF to customer, and (4) the inventory costs at all facilities.

In this distribution network, there is a set of manufacturing sites, one distribution center location, a set of possible facility locations, and a set of customer zones included. Each actor in these sets is pointed at a location, wherefore the location data variables in the Facility Location Model are the following:

- M = number of manufacturing sites (inbound origins), also let m = {1, 2, …, M}
- n = the distribution center location
- W = number of possible facilities (DC and CDFs), also let w = {1, 2, …, W}
  Where, w = 1 is the distribution center location
- C = number of customer zones, also let c = {1, 2, …, C}

As discussed in the model assumptions, each customer is allocated to one facility, which is the closest open facility, as formulated in the following binary variable and equation.

$$X_{wc}\begin{cases} 1, & if\ facility\ w\ is\ closest\ open\ facility\ for\ customer\ c, \\ 0, & otherwise. \end{cases} \qquad \forall\ w \in W, c \in C$$

$$\sum_{w=1}^{W} X_{wc} = 1 \qquad \forall\ c \in C$$

7.2.1 Multi-modality

An important extension of the facility location model is the multi-modal transportation network between DC and CDF. Therefore, the considered modes require a notation, which is as follows.

- K = number of considered transportation modes, also let k = {1, 2, …, K}

Each facility is connected to a set of transportation modes, as discussed in Table 13. The possible transportation modes per facility are given by the following integer variable.

$$Z_{nw}^{k}\begin{cases} 1, & if\ mode\ k\ is\ possible\ between\ DC\ n\ and\ facility\ w \\ 0, & otherwise. \end{cases} \qquad \forall\ w \in W, k \in K$$

7.2.2 Transportation costs

The annual transportation costs are a function of: (1) transportation rates per movement LTL, and (2) annual number of movements F. This transportation cost function is defined as α per network stream: between a manufacturing site m and the DC n (α_{mn}), between the DC n and a facility w with mode k (α_{nw}^{k}), and between a facility w and a customer zone c (α_{wc}).

$$\alpha_{mn} = LTL_{mn}F_{mn} \qquad \alpha_{nw}^{k} = LTL_{nw}^{k}F_{nw}^{k} \qquad \alpha_{wc} = LTL_{wc}F_{wc}$$

7.2.2.1 Transportation rates (LTL)

Since the locations of the facilities are known, the FTL tariff per relevant lane and for the possible modes can be calculated on beforehand. Therefore, it is possible to use these calculated tariffs as an input variable for the model. This FTL tariff is further used to calculate the LTL tariff based on the relative truckload and a discount factor for LTL (r), as following:

$$LTL\ tariff = FTL\ tariff * P^{(1-r)}$$

Where, P is the relative truckload of a transportation mode between two locations. The relative truckload equals one for all loads (Q) greater or equal to the minimum FTL tariff load (T). Therefore, the following range variables for the relative load is set.

$$P\begin{cases} \frac{Q}{T}, & if\ Q\ \in\ (0,T] \\ 1, & otherwise. \end{cases}$$

The discount factor (r) might differ per transportation mode. The lower the discount factor (r), the higher discount given for LTL shipment. When it is only possible to buy FTL shipments, the r equals 1, wherefore the relative truckload always results 1 and there is no discount given for LTL shipments.

7.2.2.2 Number of movements (F)

The number of movements are depending on the average loading factor of the shipments. As can be seen in the analysis of the current situation, the average loading factor for the inbound is 0.87 and for the outbound is 0.64. In the proposed distribution network, warehouses are consolidated, wherefore transportation of goods can be combined under a capacity constraint, i.e. increasing the average loading factor. Since it regards packed products, it is not possible to utilize the whole capacity of a transportation mode. Therefore, a maximum feasible loading factor is assumed, and the load of a truck cannot exceed this feasible loading capacity (U).

$$Q \leq U$$

Based on this capacity constraint and the volume of goods that needs to be transported between two points (E), a minimum number of movements is determined under the following equation.

$$F \geq \frac{E}{U}$$

Furthermore, for the outbound streams a service lead time is given. It is important that a customer can be served within the service lead time when it places an order. During the service lead-time, different orders can be combined for transportation. Therefore, the number of movements can be determined based on the lead-time (L) and the annual number of time units (A) as denoted in the following equation.

$$F \geq \frac{A}{L}$$

However, it is possible that the time between orders for a customer is greater than the requested lead-time. Therefore, orders cannot be combined, and it is not needed to deliver each lead-time moment. Hence, in this situation the time between orders (λ), which is normally distributed, determines the minimum number of movements as shown in the following equation.

$$F \geq \frac{A}{\lambda}$$

These steps can be combined and be used to calculate the average truckload (Q). Furthermore, the number of movements can only be an integer value; therefore, it is roundup to 0 decimals.

$$F = \frac{E}{Q} = \max\left\{\frac{E}{U}, \, min\left\{\frac{A}{S}, \frac{A}{\lambda}\right\}\right\}$$

7.2.3 Inventory rates

Besides the transportation costs, the inventory costs are considered in the model. The inventory costs consist of handling, administration and holding costs. The holding costs are a function of the average inventory level, whereas the handling and administration costs are a function of the throughput. The holding costs differ per product group, which is denoted as follows.

- G = number of different product groups, also let g = {1, 2, …, G}

In order to calculate the annual holding costs (β), the average inventory level (I) is multiplied with the corresponding annual holding costs per inventory unit (HO) for each facility w and product group g, as formulated in the following equation.

$$\beta_w^g = I_w^g HO_w^g \qquad \forall\, w \in W, g \in G$$

The average inventory level per product group at a CDF is calculated by using Little's Law. At first, the annual demand (E) is multiplied by the product group proportion (PR) and divided by the annual number of time units to get the average demand per time unit. Secondly, this average demand per time unit is multiplied with its service lead-time (S) to get the average inventory level. Furthermore, the annual facility demand equals all customer demand allocated to the facility (X_{wc}). At last, the inventory level at the DC is expected to equal the current inventory level (CIL) minus the average inventory level(s) at the CDF(s).

$$I_w^g = \frac{E_w PR^g}{A} * S \qquad \forall\, w \in \{2, W\}, g \in G$$

$$E_w = \sum_{c=1}^{C} E_c X_{wc} \qquad \forall\, w \in W$$

$$I_w^g = CIL * PR^g - \sum_{w=2}^{W} I_w^g \qquad \forall\, w = 1, g \in G$$

In order to calculate the annual handling and administration costs (γ), the annual throughput (equal to annual demand) is multiplied by the annual handling and administration costs per throughput unit (HA), which are region depending and therefore specified per facility. The handling and administration costs are considered for each facility that a product is going through. Therefore, it regards (1) handling and administration at the DC and (2) at a CDF[14].

$$\gamma_n = \sum_{c=1}^{C} E_c \, HA_n$$

$$\gamma_w = E_w HA_w \qquad \forall\, w \in \mathrm{W}$$

[14] When customers are allocated to the DC (w = 1), there is no double handling and administration costs and therefore the handling and administration costs at this facility equals zero.

7.2.4 Customer service lead-time

For a truck-loading problem, the aim is to maximize the truckload of the selected transportation mode, subject to a lead-time constraint. By maximizing the loading factor, the number of movements can be reduced such that the costs reduce as well. The lead-time depends on the customer service level that the Company offers their customers, i.e. what is the maximum lead-time from ordering moment till delivery. This service lead-time regards the transportation from DC to customer probably through a CDF. Since the outbound of goods is divided into two parts, the service lead-time is also divided into two partial service lead-times: (1) from DC to CDF, and (2) from DC or CDF to customer. The transportation lead-time should not exceed this service lead-time. Based on the transportation mode, travel distance, and travel time, a transportation lead-time per possible lane is determined. This lead-time determines if a mode is possible for a certain customer service level. Therefore, by changing the customer service lead-time, the solution might change due to different mode options. This explanation is visualized in Figure 21, where the variables are denoted as follow:

- S = maximum customer service lead time between DC and customer
- L_{nw}^{k} = transportation lead time from DC *n* to facility *w* by mode *k*
- L_{wc} = transportation lead time from facility *w* to customer zone *c*
- X_{nw}^{k} = binary decision variable for selecting mode k

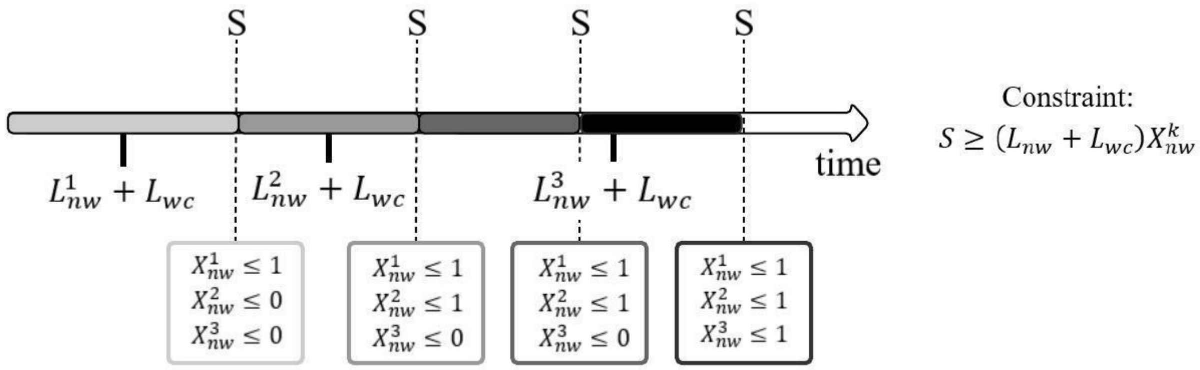

Figure 21. Explanation real, transportation and service lead-times

7.2.5 Decision variables

There are several decision variables in the model. At first, the model decides on the number of facilities and what facilities to open from the set of potential facility locations. For these open facilities, it is decided what transportation mode to use between the DC n and facility w. These two decisions are denoted by the following two binary decision variables.

$$Y_w \begin{cases} 1, & if\ facility\ w\ is\ open \\ 0, & otherwise. \end{cases} \qquad \forall\, w \in W$$

$$X_{nw}^{k} \begin{cases} 1, & if\ mode\ k\ is\ used\ between\ DC\ n\ and\ facility\ w \\ 0, & otherwise. \end{cases} \qquad \forall\, w \in W, k \in K$$

At last, the model decides on the optimal combination of the two partial allowed lead-times as function of the given service-lead time (S) for the total outbound stream: from DC to customer.

- $S1$ = allowed service lead-time between a DC *n* and facility *w*
- $S2$ = allowed service lead-time between a facility *w* and customer *c*

7.2.6 The Multi-Modal Facility Location Problem

The objective of this Facility Location Problem is to minimize the total distribution network costs, which consist of three cost types: transportation costs (α), inventory holding costs (β), and inventory handling and administration costs (γ). The transportation costs regard inbound costs from manufacturing site to DC (*mn*), outbound costs from DC to facility (*nw*), and outbound costs from facility to customer (*wc*). Based on a given service lead-time, it is decided what facilities to open and transportation mode (*k*) to use. The Multi-modal Facility Location Problem, denoted as Problem 3, is formulated as the following Mixed Integer program.

$$Min \sum_{m=1}^{M} \alpha_{mn} + \sum_{w=1}^{W}\sum_{k=1}^{K} \alpha_{nw}^{k} X_{nw}^{k} + \sum_{w=1}^{W}\sum_{c=1}^{C} \alpha_{wc} X_{wc} + \left(\sum_{w=1}^{W}\sum_{g=1}^{G} \beta_{w}^{g} Y_w + \gamma_n + \sum_{w=1}^{W} \gamma_w Y_w \right)$$

Subject to:

$S = S1 + S2$		(3.1)
$S1 \geq L_{nw}^{k} X_{nw}^{k}$	$\forall\ w \in W, k \in K$	(3.2)
$\sum_{k=1}^{K} X_{nw}^{k} = Y_w$	$\forall\ w \in W$	(3.3)
$X_{nw}^{k} \leq Z_{nw}^{k}$	$\forall\ w \in W, k \in K$	(3.4)
$S2 \geq L_{wc} X_{wc}$	$\forall\ w \in W, c \in C$	(3.5)
$Y_w \geq X_{wc}$	$\forall\ w \in W, c \in C$	(3.6)
$\sum_{w=1}^{W} Y_w \geq 1$		(3.7)
$Y_w, X_{nw} \in \{0,1\}$	$\forall\ w \in W$	(3.8)
$S1, S2 \in \mathbb{Z}$		(3.9)

The given service lead time (S) regards the maximum lead-time to serve a customer from the ordering moment, i.e. the lead-time to transport goods from DC to the customer. This transportation is performed into two steps, i.e. from DC to CDF, and from (DC or) CDF to customer. Therefore, the service lead-time is divided according these two steps: S1 and S2, see constraint (1). The model will consider all different combinations of S1 and S2 given a certain S and some selection constraints. At first, the first partial lead-time, S1, cannot exceed the transportation lead-time of a selected mode, i.e. when the transportation lead-time of a mode is higher than the allowed service lead-time, it is not possible to select that mode, see constraint (2). Furthermore, it is given that there should be one mode selected between the DC and each open CDF, see constraint (3). Each facility is connected to a set of transportation modes and it is only possible to select a mode when it is connected to that mode, see constraint (4). Secondly, it is given that the second partial lead-time, S2, cannot exceed the transportation lead-time between the facility (DC or CDF) and its allocated customers, see constraint (5). Furthermore, customers can only be allocated to an open facility, see constraint (6). Therefore, at least one facility should be opened, see constraint (7). At last, the decision variables are binary (constraint 8) and the partial lead-times should be any combination of integer values (constraint 9).

7.3 Parameter setting

In order to implement the Multi-Modal Facility Location Problem for the distribution network of the Company, the case study parameters are set in this section. At first, the variables for the locations, the modes, and the product groups are defined. Secondly, the transportation parameters are described. At last, the inventory parameters are set.

7.3.1 Variable parameters

Some of the input variables of the model are the locations of the manufacturing sites, the possible locations for DC and the facilities, and the customer zones. Based on the analysis of the current situation and the gravity model the following parameters are defined.

- M = 20, also let m = {1, 20}
- n = 1 , which is the DC location at Rotterdam Port.
- W = 15, also let w = {1, 15} , where the facilities are as mentioned in Table 13.
- C = 666, also let c = {1, 666}

Secondly, the facility location problem focusses on a multi-modal network. As discussed before, there are three modes considered between the DC and a CDF, namely (1) road transportation, (2) rail transportation, and (3) inland waterway transportation. Therefore:

- K = 3, also let k = {1,3}

At last, as discussed in the analysis, there are two aspects where the inventory holding costs depends on, which are dangerous goods and temperature-controlled goods. These two aspects, results in the dividing of the inventory into three product groups: (1) goods without specification, (2) dangerous goods (class 1 till 9), and (3) goods that require temperature control. Therefore, the following parameter for the product group variable is set.

- G = 3, also let g = {1,3}

7.3.2 Transportation parameters

For the transportation parameters it is required to set parameters for the capacity per transportation mode. From analysis, it is already known that the capacity of a truck equals 24 MT. In addition, according to (Rodrigue, 2017), it is assumed that goods are transported in containers for rail and barge, which have a capacity of (multiples of) 25 MT.

Since the network regards packed goods, there is a maximum utilization capacity that is lower than the maximum capacity of the transportation mode. According to the analysis, it is concluded that the average loading factor of the outbound is low (0.64) which requires improvement. The inbound loading factor is already high in the current situation, namely 87 per cent. According to the inbound loading factor in the current situation, it is assumed that it is feasible to improve both loading factors to an average of 0.90. Therefore, the maximum utilization capacity (U) is the mode capacity multiplied with this maximum loading factor.

Furthermore, as can be seen in the analysis of the transportation rates, the FTL tariff is applicable for a loading factor of 0.92. By using this FTL loading factor, the minimum load for FTL tariff per mode (T) can be calculated, which is used to calculate the relative truckload and the discount for LTL.

The transportation costs per mode variate highly. According to (Mathisen, Hanssen, & Jorgensen, 2012), estimates show that compared to transport by truck, the average freight cost per MT per time unit is 12 times as low for transport by rail and 35 times as low for transport by water. However, this low-cost water transport is ranked as the slowest and therefore need a comparison in transportation lead-time, which is the same for the moderate speed of rail transportation. Therefore, the FTL tariff for rail and inland waterway is a function of the mentioned cost relation (1/12 or 1/35), the mode capacity (25 MT for train and barge compared to 24 MT for truck), and the transportation lead-time (L) per mode and is calculated as follows.

$$For\ train\ (k=2): \qquad FTL_{nw}^{2} = \frac{FTL_{nw}^{1}}{12} * \frac{24\ MT}{25\ MT} * \frac{L_{nw}^{2}}{L_{nw}^{1}} \qquad \forall\ w \in W$$

$$For\ barge\ (k=3): \qquad FTL_{nw}^{3} = \frac{FTL_{nw}^{1}}{35} * \frac{24\ MT}{25\ MT} * \frac{L_{nw}^{3}}{L_{nw}^{1}} \qquad \forall\ w \in W$$

For all lanes (transport between two locations within Europe), the transportation lead-time by truck is assumed to equals one day. The distances for road transportations are calculated in the same way as during the gravity study, i.e. by using the coordinates in the Haversine formula and multiply this with a road factor of 1.46. According to (SeaRates, 2020), the transportation lead-times for rail are determined, whereas the inland waterway transportation lead-times are determined according to (Searoutes, 2020). In order to get these lead-times, average speed of a train is set on 25 km/h and barge on 13 km/h (according to (Liguori, 2018), (European Court of Auditors, 2016), and (Karttunen, Vaatainen, Antti , & Ranta, 2012)).

An overview of the connected transportation modes (Z), FTL tariffs, and transportation lead-times for each lane between the DC (n) and facility (w) are shown in Table 14.

Table 14. Parameters (connection, distance and lead-time) between DC and facilities

Facility	**Mode connection**			**FTL tariff ($-USD)**			**Lead time (days)**		
w	Z_{nw}^{1}	Z_{nw}^{2}	Z_{nw}^{3}	FTL_{nw}^{1}	FTL_{nw}^{2}	FTL_{nw}^{3}	L_{nw}^{1}	L_{nw}^{2}	L_{nw}^{3}
1	1	1	1	-	-	-	0	0	0
2	1	0	0	1,951	-	-	1	-	-
3	1	1	0	1,959	470	-	1	3	-
4	1	1	0	1,948	468	-	1	3	-
5	1	1	1	1,939	465	319	1	3	6
6	1	1	1	1,914	459	263	1	3	5
7	1	1	1	1,970	473	324	1	3	6
8	1	0	0	1,748	-	-	1	-	-
9	1	1	0	1,782	428	-	1	3	-
10	1	1	1	1,767	283	194	1	2	4
11	1	0	0	1,408	-	-	1	-	-
12	1	1	0	1,957	470	-	1	3	-
13	1	1	0	1,942	466	-	1	3	-
14	1	0	0	1,970	-	-	1	-	-
15	1	1	0	1,970	630	-	1	4	-

In contrast to truck transportation, it is assumed that it is not possible to order LTL shipments for railcar or barge-container, i.e. only possible to buy against FTL container tariff. Therefore, the LTL-discount rate (r) equals 1 for barge and rail. The discount factor for truck, is equal to current, namely 0.65.

Furthermore, the majority of the inbound streams is starting from the DC point and this location is the end point for a part of the outbound streams. Based on the transportation rate function, there is no transportation costs. However, in reality there is always a shuttling cost required. Currently, the shuttling costs for one movement from Rotterdam warehouse to Rotterdam port is contracted at $150 USD. Therefore, this shuttling costs is considered as LTL tariff for each movement from or to the port, i.e. aggregated manufacturing site or customer zone is port.

In this facility location problem, there are several alternative facilities considered with the same purpose: serving as CDF1, CDF2 or CDF3. However, there can only be one facility opened for the same purpose. Therefore, an additional requirement (constraint 3.10) is to open maximum one facility out of the alternative facility locations for CDF 1 (facilities 2 till 10), for CDF 2 (facilities 11 till 13), and CDF 3 (facilities 14 and 15).

$$\sum_{w=a}^{b} Y_w \leq 1 \qquad \forall\ (a,b) \in (2,10), (11,13), and\ (14,15) \qquad (3.10)$$

7.3.3 Inventory parameters

For the inventory parameters there are two types of costs: handling administration and holding. The handling administration costs (HA) are depending on the throughput, whereas the holding costs (HO) are specified per product group and depending on the inventory level at a facility. Based on the current analysis and average inventory rates, the parameters are as follows.

- PR^1 = 0.78
- PR^2 = 0.06
- PR^3 = 0.16
- HO^1 = $73 per MT per year
- HO^2 = $105 per MT per year
- HO^3 = $145 per MT per year
- HA = $14 per MT

Since the operation costs differ geographical, it is researched what the inflation rate is of all considered facilities compared to the current facilities. According to (Eurostat, 2019), the following rates are set, which are multiplied with current costs to get the specific facility costs.

- Inflation rate w ∈ {1,1} (NL) = 1.00
- Inflation rate w ∈ {2,5} (SK) = 0.32
- Inflation rate w ∈ {6,6} (AT) = 0.95
- Inflation rate w ∈ {7,7} (HU) = 0.26
- Inflation rate w ∈ {8,10} (IT) = 0.79
- Inflation rate w ∈ {11,13} (PL) = 0.28
- Inflation rate w ∈ {14,15} (RO) = 0.19

7.4 Solution

In this section, the results of the Multi-modal facility location problem (Section 7.2) are shown by using the model assumptions (Section 7.1) and the case study parameters (Section 7.3). The solution of the model is depending on a given service lead-time and there are different lead-times considered in the model. According to the current allowed lead-times, it is decided that the service lead-time is maximum ten days. Therefore, the considered scenarios in the model are a service lead-time of one day till ten days. The given service lead-time is the total lead-time from order moment till delivery, i.e. the transportation from DC to customer. This process is divided into two parts: (1) from DC to facility, and (2) from facility to customer. The lead-time of these two parts should be equal to the given service lead-times, wherefore there are different combinations possible. For all possible combinations of a given service lead-time, the model is solved by using the solver function in Excel with the Generalized Reduced Gradient (GRG) Nonlinear solving method. The combinations are considered by programming a trial and error VBA in Excel. Hence, for each service lead-time, one day till 10 days, the optimal solution (combination of partial lead-times, open facilities and transportation modes) is the situation with the lowest total network costs. Since the combination of partial lead-times is done by trial and error, the solver only has to decide on binary values: the facilities to open and the mode to use. Therefore, there are no local minima and the GRG Nonlinear algorithm results the optimal solution, i.e. there is no multi-start function required.

The solution of the Multi-modal Facility Location problem (Problem 3) is shown in Figure 22, where the decision variables (number of facilities to open and the transportation mode to use) and the corresponding objective value (total network costs) are given per scenario (service lead-time). As can be seen, consolidating to one warehouse directly reduce the total network costs. Furthermore, opening a CDF is profitable from the moment that the lead-times allows the model to use different transportation modes.

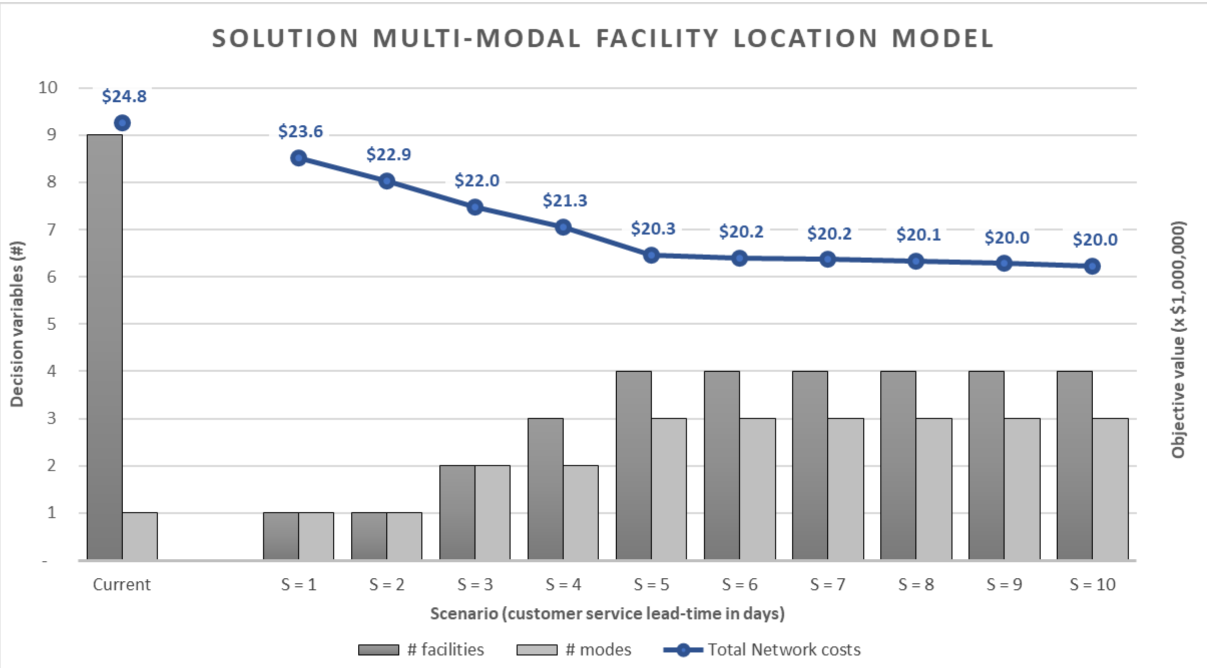

Figure 22. Solution Multi-modal Facility Location problem

The current distribution network costs amount annually $24.8 million USD, which consist of $18.2 million USD transportation costs and $6.6 million USD inventory costs. The network regards nine warehouses for which transportations are only performed by truck (single modal). The aim of the model is to minimize the total network costs (shown in Table 15) by deciding on the optimal combination of decision variables (shown in Table 16). As a result, performances of the distribution network improve, i.e. higher loading factor (shown in Table 17).

At first, it can be seen that consolidating to one distribution center (located in Rotterdam) results in a potential reduction of $1.2 million USD of the transportation costs (S=1). This situation occurs because transportations from a manufacturing site or to a customer can be consolidated as well. Secondly, when the service lead-time is extended with one day (S=2), it becomes possible to combine customer orders within these two days, therefore the annual transportation costs further reduce with $0.7 million USD. At third, from a service lead-time of three days (S=3), it becomes possible to transport goods by train between the DC and CDF1 (located in Wroclaw, Poland). For these transportations, it is possible to combine demand of different customers in the transportation between DC and CDF, resulting in higher performances (shown in Table 17) for the major part of the outbound stream. Furthermore, train transportations are cheaper than truck transportations. Therefore, the transportation costs reduce when opening one CDF and using a multi-modal network. However, the annual inventory costs increase with $0.2 million USD because of handling and administration costs at two facilities for 38 per cent of the goods (see Table 18). At fourth, when extending the service lead-time to four days (S=4), it is also possible to transport by train between the DC and CDF2 (located at Milan, Italy), wherefore the annual transportation costs further reduce with $1.1 million USD. However, the annual inventory costs increase with $0.5 million USD, since there is double handling and administration required for 57 per cent of the goods (see Table 18). At fifth, extending the service lead-time to five days (S=5) allows to open CDF 3 (located at Ploiesti, Romania) and performing the transport by train. Furthermore, this lead-time allows the use of barge transport between DC and CDF1 against lower costs than train, resulting in an additional annual potential reduction of $1.0 million USD. From a service lead-time of five days, the cheapest options for transportation modes and facilities to open are selected. Furthermore, the maximum loading factor for transportation between DC and CDF(s) is reached, namely 90 per cent. Therefore, the effect of extending the service lead-time to six, seven, eight, nine or ten days is minimal. The small reduction in these scenarios is the result of combined transportation between CDF and customer due to a higher lead-time between these two points and lower holding costs due to an increasing of the inventory level at a cheaper CDFs comparing to the more expensive DC.

Table 15. Objective value (solution multi-modal facility location problem)

Situation	**S=1**	**S=2**	**S=3**	**S=4**	**S=5**	**S=6**	**S=7**	**S=8**	**S=9**	**S=10**
Network costs (x 1,000,000)	$23.6	$22.9	$22.0	$21.3	$20.3	$20.2	$20.2	$20.1	$20.0	$20.0
Transport costs (x 1,000,000)	$17.0	$16.3	$15.1	$14.0	$13.1	$13.0	$12.9	$12.9	$12.8	$12.7
Inventory costs (x 1,000,000)	$6.6	$6.6	$6.8	$7.3	$7.3	$7.2	$7.2	$7.2	$7.2	$7.2

Table 16. Decision variables (solution multi-modal facility location problem)

Situation	**S=1**	**S=2**	**S=3**	**S=4**	**S=5**	**S=6**	**S=7**	**S=8**	**S=9**	**S=10**
S1 (nw)	-	-	2	3	4	4	4	4	4	4
S2 (wc)	1	2	1	1	1	2	3	4	5	6
# facilities	1	1	2	3	4	4	4	4	4	4
CDF 1	-	-	w10 Train	w10 Train	w10 Barge	w10 Barge	w10 Barge	w10 Barge	w10 Barge	w10 Barge
CDF 2	-	-	-	w12 Train	w12 Train	w12 Train	w12 Train	w12 Train	w12 Train	w12 Train
CDF 3	-	-	-	-	w15 Train	w15 Train	w15 Train	w15 Train	w15 Train	w15 Train

Table 17. Performances (solution multi-modal facility location model)

Situation	**S=1**	**S=2**	**S=3**	**S=4**	**S=5**	**S=6**	**S=7**	**S=8**	**S=9**	**S=10**
Load-factor mn	.89	.89	.89	.89	.89	.89	.89	.89	.89	.89
Load-factor nw	-	-	.88	.90	.90	.90	.90	.90	.90	.90
Load-factor wc	.64	.64	.64	.64	.64	.64	.65	.66	.66	.67

Table 18. Proportional demand per facility (solution multi-modal facility location problem)

Situation	**S=1**	**S=2**	**S=3**	**S=4**	**S=5**	**S=6**	**S=7**	**S=8**	**S=9**	**S=10**
DC (w=1)	1.00	1.00	.62	.43	.43	.43	.43	.43	.43	.43
CDF 1 (w=10)	-	-	.38	.33	.22	.22	.22	.22	.22	.22
CDF 2 (w=12)	-	-	-	.24	.24	.24	.24	.24	.24	.24
CDF 3 (w=15)	-	-	-	-	.10	.10	.10	.10	.10	.10

Overall, the biggest improvement opportunity: annual reduction of $4.5 million USD, is to deliver against a customer service lead-time of 5 days or greater, while consolidating to one DC (located at Rotterdam) and opening three CDFs (located at Wroclaw, Milan, and Ploiesti). These CDFs make it possible to combine orders of different customer and to use the cheapest connected transportation mode (train and barge) for the major part of the transportation process. This solution serves as a benchmark for the Company, which can be considered and further researched during the tendering process.

7.5 Sensitivity analysis

A sensitivity analysis determines how the uncertainty in the output of a mathematical model can be divided and allocated to different sources of uncertainty in its inputs. Therefore, the outcomes of the model are recalculated under alternative assumptions in order to determine the impact of variability. During the gravity study, the sensitivity of several parameters (road factor and rates parameters) is tested and concluded to be low sensitive. However, uncertainties in demand are not considered, since a uniform lower or higher demand will not affect the gravity points because the proportional demand per customer location stays the same and there were equal performances (truckload and number of movements) assumed. In contrast to the gravity study, the multi-modal facility location considers the improvement opportunities for the performances. Therefore, it is possible to check the sensitivity of uncertainties in demand. Hence, in this sensitivity analysis it is investigated how uncertainty in demand affect the solution. Therefore, the problem is solved for two situations: (1) 10 per cent lower demand, and (2) 10 per cent higher demand. Where variation in demand is assumed to be uniform distributed over the customers. Furthermore, it is assumed that the number of orders per customer is equal, however the demand per order differ. Since the inventory is in relation with the demand (according to Little's Law), the average inventory level changes respectively to the demand.

An overview of the solution for the lower demand is given in Figure 23 and for the higher demand in Figure 24. As can be seen, the same decisions (opening of facilities and use of transportation modes) are made as in the original situation. The costs aspects are lower or higher because there is less or more demand handled. Furthermore, the height of reductions differs due to lower or higher improvement opportunities, e.g. the possibility to combine transport of goods is lower when the demand decreases. However, the uncertainty in the demand does not affect the decisions made by the model, wherefore it is concluded to be (demand) insensitive.

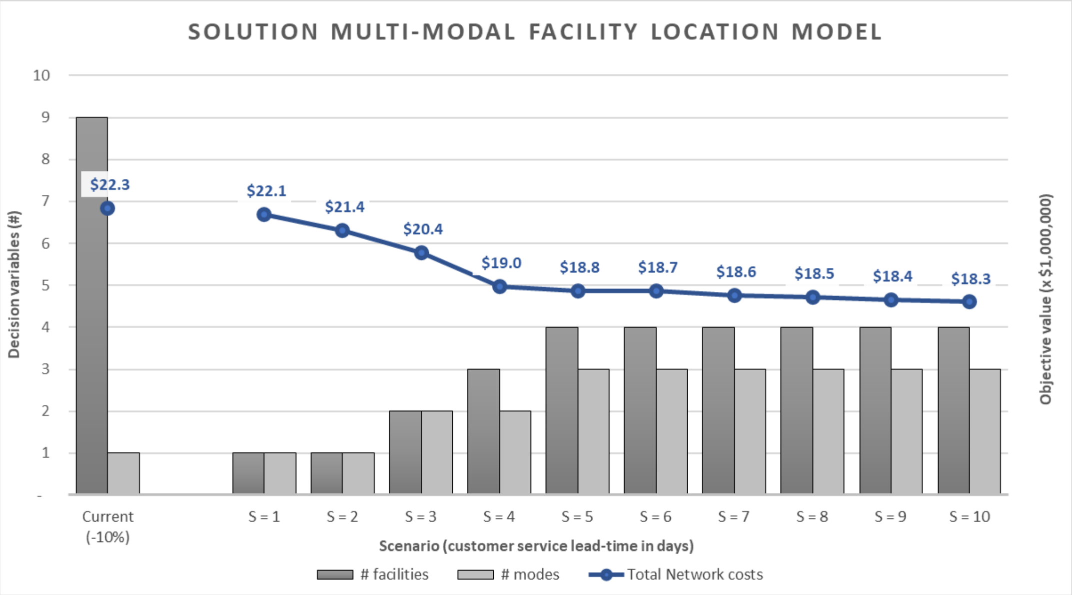

Figure 23. Sensitivity analysis 10% lower demand (solution multi-modal facility location problem)

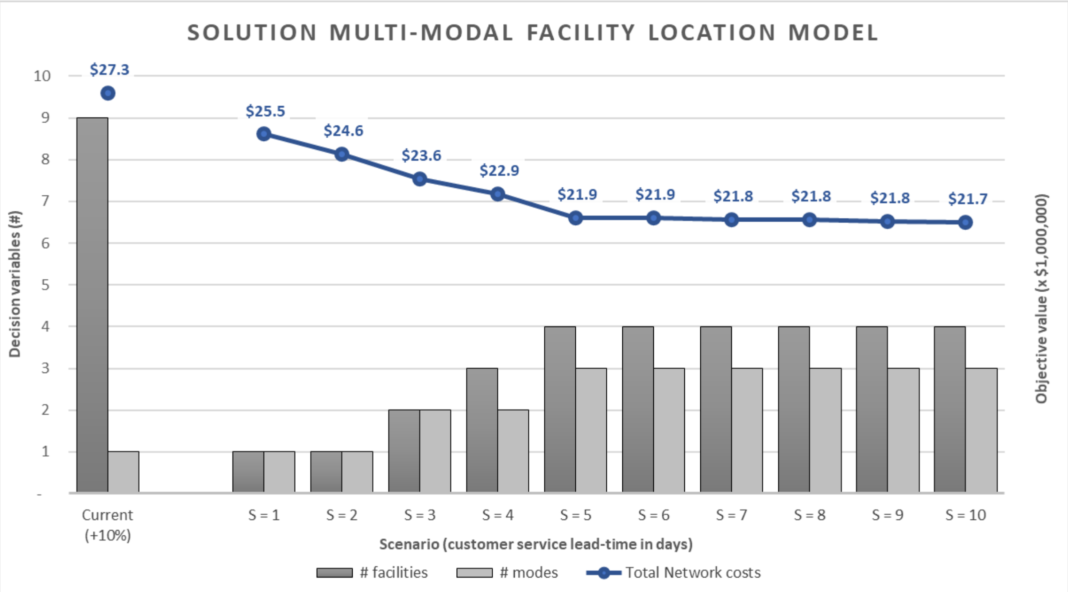

Figure 24. Sensitivity analysis 10% higher demand (solution multi-modal facility location problem)

Chapter 8. Sustainable Distribution Network

In this chapter the optimal distribution network in terms of sustainability is investigated. Where, optimal is defined as the situation with the lowest environmental impact. Therefore, the aim is to minimize the total Greenhouse Gas (GHG) Emissions of the distribution network. Considering optimal in terms of sustainability is a different scenario than considering optimal in terms of costs, wherefore in this chapter a final answer is given on Sub-research question 3. *How should the Company implement the distribution network redesign based on the optimal design methodology considering different scenarios?*

In order to determine the optimal sustainable design, the aim is to minimize the total GHG Emissions. The structure of this problem is similar to the multi-modal facility location problem (Problem 3), i.e. deciding which facilities to open and what transportation mode(s) to use. However, for this model it is the aim to minimize the environmental impact instead of the costs, wherefore the model input variables differ. Therefore, Problem 3 is used wherefore the transportation cost aspects are adjusted to sustainable aspects and the inventory is ignored.

According to (McKinnon & Piecyk, 2020), the GHG emissions per tonne-km for the European Chemical Transport sector are as follow: for road transport (truck) 62 gCO_2/tonne-km, for rail transport (train) 22 gCO_2/tonne-km, and for inland waterway (barge) 31 gCO_2/tonne-km. For each lane the environmental impact is calculated by multiplying these values with the distance and corresponding demand. The result of this calculation replace the LTL tariff values in the Multi-modal facility location problem of Chapter 7. The adjusted model is solved for the different service lead-time scenarios, which results in the comparison between the objective values in (1) costs optional and (2) sustainable optimal situation, as shown in Figure 25.

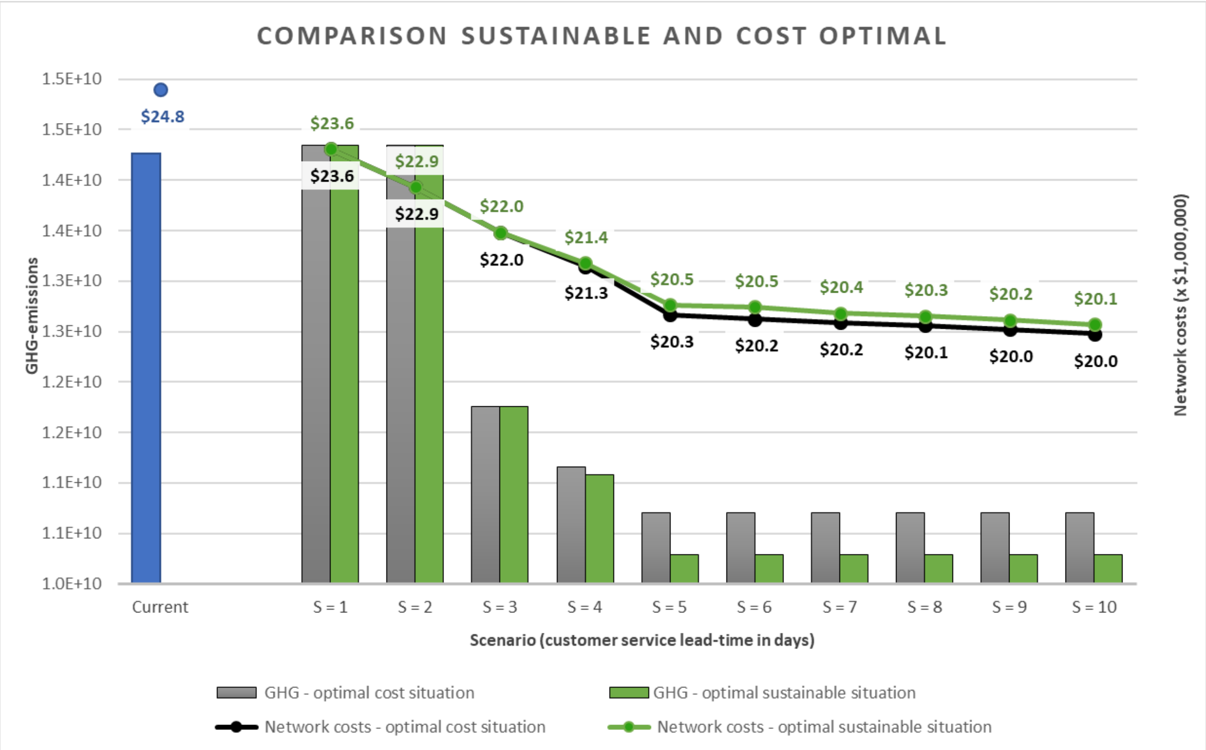

Figure 25. Comparison Sustainable and Cost optimal situation

As can be seen, in both situations: cost optimal and sustainable optimal, there is a reduction of the GHG emissions as well as the costs of the distribution network. Furthermore, when analyzing the decision variables for the sustainable optimal situation as shown in Table 19, the decisions are similar to the solution of the cost optimal (shown in Table 16). However, the comparison visualizes also some differences. At first, it is noticeable that the GHG-emissions of the current situation are somewhat lower than for the consolidated warehouse (S=1 or S=2), whereas the reduction in costs is high. This can be explained by the fact that the truckload is not considered in the emissions (i.e. average GHG-emission per MT are considered), whereas the truckload is very depending for the costs. Furthermore, there is calculated with exact distances in the current situation, whereas for the modelled situations there is made use of an equation based on coordinates and a road factor. Secondly, for a service lead time of four days (S=4) it becomes possible to transport by train to Gliwice, wherefore it is more sustainable to open CDF1 in Gliwice instead of Wroclaw but also more expensive. However, this is only the case when there are three facilities opened, when opening four facilities it is again more sustainable and cheaper to open the facility in Wroclaw. This can be explained by the fact that a part of the customers is now allocated to CDF3 in Romania and therefore the total travel distances changes. At last, the main difference between the cost optimal situation and the sustainable optimal situation is the use of train or barge, in which train transport is the most sustainable and barge is the cheapest transportation mode. Therefore, if the lead-time and connection allow, train is selected for the sustainable optimal situation and barge is selected for the cost optimal situation. However, the differences in performance are minimal (considering service lead-time of 5 days till 10 days):

- For the cost optimal distribution network there is the opportunity to reduce the GHG emissions by 25 per cent and the costs by 18-20 per cent;
- For the sustainable optimal distribution network there is the opportunity to reduce the GHG emissions by 28 per cent and the costs by 17-19 per cent.

Table 19. Solution of optimal sustainable situation (decision variables)

Situation	**S=1**	**S=2**	**S=3**	**S=4**	**S=5**	**S=6**	**S=7**	**S=8**	**S=9**	**S=10**
S1 (nw)	-	-	2	3	4	4	4	4	4	4
S2 (wc)	1	2	1	1	1	2	3	4	5	6
# facilities	1	1	2	3	4	4	4	4	4	4
CDF 1	-	-	w10 Train	w9 Train	w10 Train	w10 Train	w10 Train	w10 Train	w10 Train	w10 Train
CDF 2	-	-	-	w12 Train	w12 Train	w12 Train	w12 Train	w12 Train	w12 Train	w12 Train
CDF 3	-	-	-	-	w15 Train	w15 Train	w15 Train	w15 Train	w15 Train	w15 Train

Chapter 9. Conclusion and discussion

In this chapter, a conclusion and discussion are given about the results and used models of this thesis. At first, the main results are summarized (Section 1), which answers the main research questions: *What is the optimal design for the multimodal distribution network of the Company?* Secondly, limitations of the model and the case study are reflected (Section 2) followed by suggestions for future research (Section 3).

9.1 Main Results

In order to determine the optimal distribution network design, a methodology is developed with the objective to minimize the distribution network costs. Based on analysis, it is concluded that these distribution network costs depend on the truckload and the distances and follow an increasing concave down cost structure. Therefore, a nonlinear function for the price per movement is defined in which the truckload and distance is considered. The ideal situation proposes a consolidation of the warehouses to one distribution center (DC), which should be optimally located relative to the manufacturing sites and customers. Therefore, a single-facility gravity study (defined as a Nonlinear Program) is performed with the objective to minimize the total transportation costs by deciding on the center of gravity. Solving the problem results in a gravity center at Rotterdam Port. At this distribution center the logistic of goods is concentrated, from where goods are transported to customers directly or through a cross docking facility (CDF). In order to find the optimal CDF location(s), a multi-facility gravity model is developed (defined as a Nonlinear Program) with the objective to minimize the outbound transportation costs by deciding on multiple gravity points. The solution shows that a second facility (CDF 1) should be optimally located in Slovakia or Poland, a third facility (CDF 2) in North Italy, and a fourth facility (CDF 3) in Romania. The advantage of a CDF is the opportunity to combine transportation of different customers for the major part of the transportation process (between DC and CDF) and to use cheaper transportation modes. Therefore, a multi-modal facility location problem is developed (defined as a Mixed Integer Program) with the objective to minimize the distribution network costs by deciding on the facilities to open and the transportation modes to use subject to a given customer service lead-time. According to this customer service lead-time, the performances of the distribution network change. At first, a longer lead-time gives the opportunity to combine the transportation of different orders, wherefore the truckload increases, the number of movement decreases, and therefore, the costs decrease. It is possible to optimize the truckload, since the cost aspects in the model are related to this truckload: i.e. higher loading factors results in lower costs per product unit. Secondly, based on literature, it is concluded that the cheaper a transportation mode, the longer the travel lead-time, therefore a higher customer service lead-time, might allow a cheaper transportation option. The solution shows that the biggest improvement opportunity: annual cost reduction of $4.5 million USD, can be achieved when delivering against a customer service lead-time of (at least) five days. In this situation, the logistic of goods is consolidated to one distribution center, located in Rotterdam port, and three cross docking facilities are opened: CDF1 located in Wroclaw (Poland), CDF2 in Milan (Italy), and CDF3 in Ploiesti (Romania). Due to these opened CDFs, orders of different customers can be combined, and the given lead-time allows to use the cheapest connected transportation mode for the major transportation part. By extending the multi-modal facility location problem with the truck-loading problem and the customer service lead-time, a gap in the literature is filled: how to include the truck-loading problem and uncertainty (different scenarios) in the facility location problem. Therefore, the thesis gives new insights and it is regarded as academic relevant.

Furthermore, a side-aim of the project is to reduce the CO_2 footprint wherefore the distribution network becomes more sustainable. Therefore, the optimal sustainable design is determined by using the multi-modal facility location problem where the costs are adjusted to Greenhouse Gas Emissions (GHG) and therefore the environmental impact of the distribution network is minimized. While comparing the solutions of the cost optimal and the sustainable optimal design, the main difference is the use of train or barge, in which train is the most sustainable and barge is the cheapest transportation mode. However, the performances differ minimal: for the cost optimal distribution network there is an opportunity to reduce the GHG emissions by 25 per cent and the costs by 20 per cent, whereas for the sustainable optimal distribution network there is an opportunity to reduce the GHG emissions by 28 per cent and the costs by 19 per cent. Hence, for both redesigns: cost optimal and sustainable optimal, it is possible to become more efficient and more sustainable.

9.2 Reflection on Limitations

In this section, the limitations of the thesis are reflected. In total there are five main limitations which consist of assumptions made in the model or case study that might affect the solutions.
At first, the defined cost function in the model is based on historical data. These historical data regard more than eleven thousand measurements of prices that carriers requested during the last years to perform the transportation of goods. Comparing these measurements, shows a significant relation of the cost structure and therefore it is assumed to use this concave cost structure to model transportation rates in the research. However, due to uncertainties, such as differences in oil prices, competition or capacity, transportation rates might change in the future. Furthermore, there are factors that might impact the relative cost of train and barge comparing to truck. For the gravity points it is already researched that model-decisions are not affected by changed costs (e.g. from concave to linear). Though, when the offered prices of suppliers in the tendering process differ in cost structure or relative prices between modes, it is recommended to resolve the multi-modal facility location problem with the new costs. Secondly, it is assumed that the same customer service lead-time is agreed with all customers. In reality, it might be that different customers requested different lead-times. However, the requested lead-time might depend on the performances of the distribution network, i.e. a customer might consider higher lead-times when the process becomes more efficient or sustainable. Therefore, it is complex to investigate on beforehand what the requested lead-time of each customer for each possible scenario is and there is assumed on a general customer service lead-time. Though, when there cannot be agreed with all customers on a general lead-time this might have two possible disadvantages: (1) when a high general lead-time is implemented, customers that request a lower lead-time might no longer order at the Company, whereas (2) when a low general lead-time is implemented the improvement potential for customers that are willing to agree on a higher lead-time cannot be achieved. Therefore, when it is not possible to agree with all customers on the same lead-time, a segmentation of customers based on customer service level might results a solution that is better applicable in practice. At

third, the possibilities of multi-modality are only considered for the outbound streams. This decision is made because it seems likely that the effect of multi-modality in the inbound streams is minimal. This is based on the fact that it is only possible for the minority of manufacturing sites (the sites within Europe) to consider the extension of a multimodal network. Furthermore, relative to the costs of the outbound process, the costs of the inbound process are low, therefore potential improvement opportunities are low as well. However, multi-modality for the inbound process probably results in even further reduction of the distribution network costs. At fourth, the model determines the required inventory level based on the relation to the demand by using the current turnover rate. Therefore, there is no optimal inventory policy included in the redesign. Since the demand is stable over the time, the inventory level results to be stable as well. However, consolidating from nine warehouses to one DC might results a different optimal inventory level due to the possibility to centralize the safety stock. Therefore, including an optimal inventory policy seems likely to lower the inventory level, resulting in even further reduction of the distribution network costs. At last, it is assumed to not consider any facility costs in the model. Due to the outsourcing of the transport and storing of goods, the facilities are not owned by the Company, therefore the facility costs are included in the defined costs (supplier prices). Additional facility costs for the Company consist of Ordering Processing costs, which regard activities to maintain and improve the relationship with suppliers. It is complex to specify these costs per facility and even more complex to determine the relative costs for a CDF comparing to a warehouse. Furthermore, these additional facility costs are negligible comparing to the reduced transportation costs when opening an additional facility. In addition, a consolidation of nine warehouses to one till four facilities (and probably one supplier) seems likely to result in lower facility costs.

9.3 Future research

The thesis initiates four main directions for future research before it is possible to implement the optimal design. At first, by consolidating nine warehouses to one distribution center the logistic complexity reduces. However, the use of cross docking facilities and different transportation modes might increase the planning complexity. Therefore, it should be future researched what the impact of the redesign is for the planning tasks. Secondly, in the current situation it is known that there is a lack of capacity of truck-chauffeurs. Therefore, the proposed redesign might help in solving this problem due to the transportation by other modes. However, it is not researched what the trend is in capacity of these modes, therefore, it is recommended to future research the capacity of required transportation in the coming years. At third, due to the current Corona crisis, it is required to segment customers into customer groups, which are served against different customer service levels. Therefore, future research should investigate what effect this segmentation has on the optimal distribution network and how to include the use of different customer service lead times in the multi-modal facility location model. At last, the Company should decide on the scenario (customer service lead-time and the sustainability of the network) that fits best in reality. Based on the selected scenario, future research is required to find a supplier who can perform the optimal distribution network redesign against the given objective value, i.e. a supplier that have facilities at the right locations with a connection to multi-modality and offers acceptable prices. Hence, the solution of this research can serve as benchmark during the tendering process.

Bibliography

Ahmadi-Javid, A., & Seddighi, A. (2013). A location-routing problem with disruption risk. *Transp Res Part E Logist Transp Rev, 53*, 63-82.

Ambrosino, D., & Sciomachen, A. (2016). A capacitated hub location problem in freight logistics multimodal networks. *Optim Lett, 10*, 875-901.

Ashnani, M., Miremadi, T., Johari, A., & Danekar, A. (2015). Environmental impact of alternative fuels and vehicle technologies: A Life Cycle Assessment perspective. *Procedia Environmental Sciences, 30*, 205-210.

Aziziankohan, A., Jolai, F., Khalilzadeh, M., Soltani, R., & Tavakkoli-Moghaddam, R. (2017). Green supply chain management using the queuing theory to handle congestion and reduce energy consumption and emissions from supply chain transportation fleet. *Journal of Industrial Engineering and Management, 10*(2), 213-236.

Ballou, R., Rahardja, H., & Sakai, N. (2002). Selected country circuity factors for road travel distance estimation. *Transportation Research Part A, 36*, 843-848.

Crainic, T. (2003). Long-haul freight transportation. *Hall RW (ed), 56*, 451-516.

European Court of Auditors. (2016). *Rail freight transport in the EU: still not on the right track.* European Union.

Eurostat. (2019). *Lonen en Loonkosten*. Retrieved 03 10, 2020, from Eurostat Staitsics Explained: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Wages_and_labour_costs/nl

Fazayeli, S., Eydi, A., & Kamalabadi, I. (2018). A model for distribution centers location-routing problem on a multimodal transportation network with a meta-heuristic solving approach. *J Ind Eng Int, 14*, 327-342.

Ghiani, G., Laporte, G., & Musmanno, R. (2004). *Introduction to Logistics Systems Planning and Control.* Chichester: John Wiley & Sons Ltd.

Govindan, K., Jafarian, A., Khodaverdi, R., & Devika, K. (2014). Two echelon multiple vehicle location-routing problem with time windows for optimization of sustainable supply chain network of perishable food. *Int J Prod Econ, 152*, 9-28.

Ivanov, D., Tsipoulanidis, A., & Schonberger, J. (2017). Chapter 5. Sourcing strategies. In *Global Supply Chain and Operations Management: A Decision-oriented Introduction to the Creation of Value* (pp. 97-118). Switzerland: Springer International Publishing.

Karttunen, K., Vaatainen, K., Antti , A., & Ranta, T. (2012). The Operational Efficiency of Waterway Transport of Forest Chips on Finland's Lake Saimaa. *Silva Fennica*, 395-413.

Kay, M., & Warsing, D. (2009). Estimating LTL rates using publicy available emperical data. *International Journal of Logistics: Research and Applications, 12*(3), 165-193.

Laporte, G., & Nobert, Y. (1981). An exact algorithm for minimizing routing and operating costs in depot location. *Eur J Oper Res, 6*, 224-226.

Liguori, P. (2018, 02 26). *Europe's rail freight can better connect ports to the hinterlands*. Retrieved 03 23, 2020, from JOC: https://www.joc.com/regulation-policy/transportation-policy/international-transportation-policy/europe-rail-truckers-work-us-not-against-us_20180628.html

Little, J., & Graves, S. (2008). Chapter 5: Little's Law. In D. Chhajed, & T. Lowe, *Insights From Basic Operations Management Models and Principles* (pp. 81-100). LLC: Springer Science + Business Media.

Mathisen, T., Hanssen, T., & Jorgensen, F. (2012). Generalized transport costs in intermodal freight transport. *Procedia - Social and Behavioral Sciences, 54*, 189-200.

McKinnon, A., & Piecyk, M. (2020). *Measuring and Managing CO2 Emissions of European Chemical Transport.* Edinburgh, UK: CEFIC The European Chemical Industry Council.

Mitroff, I., Raturi, A., Amoako-Gyampah, K., & Kaplan, B. (1974). On managing science in the systems age: two schemas for the study of science as a whole systems phenomenon. *Interfaces, 4*(3), 46-58.

Muriel, A., & Simchi-Levi, D. (2003). Supply Chain Design and Planning - Applications of Optimization Techniques for Strategic and Tactical Models. In A. Kok, & S. Graves, *Handbooks in OR & MS* (p. Chapter 2). Elsevier B.V.

Najjartabar-Bisheh, M., Delavari, M., & Malmir, B. (2017). Role of third-party companies in sustainable supply chain design. *Int J Logist Syst Manag*, (in press).

Park, J., Shin, K., Chang, T., & Park, J. (2010). An integrative framework for supplier relationship management. *Industrial Management & Data, 110*(4), 495-515.

Rail Net Europe. (2018). *Rail Freight Corricors (RFCS) General Information*. Retrieved 03 10, 2020, from RNE: http://rne.eu/rail-freight-corridors/rail-freight-corridors-general-information/

Robusto, C. (1957). The Cosine-Haversine Formula. *The American Mathematical Monthly, 64*(1), 38-40.

Rodrigue, J.-P. (2017). *The Geography of Transport Systems.* New York: Hofstra University.

Santen, V. (2017). Towards more efficient logistics: increasing load factor in a shipper's road transport. *The International Journal of Logistics Management, 28*(2), 228-250.

SEALS. (2008). *Statistical Coverage and Economic Analysis of the Logistics Sector in the EU.* SEALS Consoritum.

SeaRates. (2020). *Sea rates distances-time*. Retrieved 03 10, 2020, from SeaRates: https://www.searates.com/services/distances-time/

Searoutes. (2020). *Searoutes routing*. Retrieved 03 10, 2020, from Searoutes: https://www.searoutes.com

Shen, Z., Coullard, C., & Daskin, M. (2003). A joint location-inventory model. *Transp Sci, 1*(37), 40-55.

Simchi-Levi, D., Chen, X., & Bramel, J. (2005). *The Logic of Logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management.* New York: Springer Science+Business Media.

Smith, L., Campbell, J., & Mundy, R. (2007). Modeling net rates for expedited freight services. *Transportation Research Part E, 43*, 192-207.

Swenseth, S., & Godfrey, M. (1996). Estimating Freight Rates for Logistics Decisions. *Journal of Business Logistics, 17*(1), 213-231.

Tuzkaya, U., Onut, S., & Tuzkaya, G. (2014). A Strategic planning methodology for the multimodal tarnsportation systems: a case study from Turkey. *J Appl Math*, ID 931456:23.

UENCE. (2018). *Map of the European Inland Waterway Network*. Retrieved 03 10, 2020, from United Nations Economic Commission for Europe: https://www.unece.org/fileadmin/DAM/trans/main/sc3/AGN_map_2018.pdf

van den Heuvel, F., de Langen, P., van Donselaar, K., & Fransoo, J. (2012). *Spatial concentration and location dynamics in logistics: the case of a Dutch province.* Eindhoven: University of Technology Eindhoven.

van den Vlist, P., & Broekmeulen, R. (2006). Retail Consolidation in Synchronized Supply Chains. *ZfB, 76*, 165-176.

Welge, M., & Al-Laham, A. (2007). *Strategic Management.* Wiesbaden: Gabler.

Yuceer, U., & Ozakca, A. (2010). A truck loading problem. *Computers & Industrial Engineering, 58*, 766-773.