MASTER

Relocating a subset of road transport depots of a 3PL company

Bouwmans, F.J.

Award date:
2021

Department of Industrial Engineering and Innovation Sciences
Operations Planning Accounting & Control Group

Relocating a subset of road transport depots of a 3PL company

Master Thesis

Frans Bouwmans

In partial fulfilment of the requirements for the degree of **Master of Science** in **Operations Management and Logistics**

University supervisors:
Dr. L. Martin, *TU/e*, OPAC
Dr. J. Kinable, *TU/e*, OPAC
Dr. A. Schrotenboer, *TU/e*, OPAC

Company supervisor:
Christian Fricke, *Rhenus Logistics*, Operations Manager

Eindhoven,
26th July 2021

TU/e, School of Industrial Engineering
Master Theses Operations Management and Logistics

Keywords: Third party logistics, large road freight forwarder, relocating a subset of road freight forwarding depots, capacitated multi-allocation hub location problem, multi-commodity network flow problem

Abstract

Rhenus logistics is a large third party logistics (3PL) company. Amongst other solutions, they offer freight forwarding from initial senders to final consignees by road. Rhenus Road NL controls road freight forwarding services from, to, over, and within the Netherlands. The depots that Rhenus Road NL operates are from a theoretical point of view different than any other relevant road transport depot in the road freight forwarding network. Rhenus Road NL is looking to relocate the depot that they currently operate in Venlo, such that total costs comprised of transportation costs and fixed depot establishment costs are minimized. To Rhenus Road NL it is not about whether they should relocate the depot, but where to. Therefore, we develop a methodology for optimally relocating a subset of depots of Rhenus Road NL. Our problem represents a new variant of the Median Capacitated Multi-allocation Hub Location Problem (CMHLP), and we define it as the p-Median Subset CMHLP (SCMHLP). We introduce a model formulation for the general case of relocating a subset of depots of Rhenus Road NL. A special case of the problem occurs when we relocate only one depot of Rhenus Road NL. For this special case we propose a Multi-commodity Network Flow Problem (MNFP) approach in which we solve a MNFP per candidate depot location. In the reference results, the lowest total cost result by relocating the depot in Venlo to the candidate location in Venlo. The appeal of candidate locations is relatively insensitive to realistic changes in parameters and circumstances. The (percentage) differences in total costs however imply that the MNFP approach has difficulties differentiating between (arguably some) candidate depot locations. Therefore, qualitative characteristics that the company considers (e.g. retaining personnel) are decisive. We conclude that the candidate site in Venlo is the best option, at least when relocating only the depot in Venlo. We also conduct numerical experiments on artificial scenarios of relocating a single other depot or more than one depot that Rhenus Road NL operates to obtain managerial insights and test the methodology's performance. We need to be careful drawing conclusions on artificial scenarios. Despite the limitations of this research, we recommend the company to relocate the depot that is currently located in Venlo to the candidate site in Venlo.

Executive summary

Rhenus logistics is a large 3PL company. Amongst other solutions, they offer freight forwarding from initial senders to final consignees by road. Rhenus Road NL controls road freight forwarding services from, to, over, and within the Netherlands. The depots that Rhenus Road NL operates are from a theoretical point of view different than any other relevant road transport depot in the road freight forwarding network. Rhenus Road NL is looking to relocate the depot that they currently operate in Venlo, such that total costs comprised of transportation costs and fixed depot establishment costs are minimized. To Rhenus Road NL it is not about whether they should relocate the depot, but where to. They consider eight candidate depot locations. Table 1 presents the reference results sorted by annual total costs. Δ represents the yearly difference with the lowest total costs.

Table 1: Reference results

Candidate location	**Total costs**	**% Δ**
Venlo	€ 38,332,108	- %
Eindhoven	€ 38,503,258	0.45 %
Oss	€ 38,525,741	0.51 %
Maastricht	€ 38,575,296	0.63 %
Tilburg	€ 38,763,784	1.13 %
Arnhem	€ 38,775,083	1.16 %
Amersfoort	€ 39,244,697	2.38 %
Dordrecht	€ 39,483,445	3.00 %

The lowest annual total cost result by relocating the depot to the candidate location in Venlo. The (percentage) differences with the lowest annual total costs are relatively low for candidate locations relatively close to the candidate location in Venlo. The appeal of candidate locations is relatively insensitive to realistic changes in parameters and circumstances. Given that we make assumptions, and that there is inevitably going to be inaccuracy in the data, the (percentage) differences in annual total costs imply that we cannot decide on the optimal location solely based on numerical results. Qualitative characteristics and requirements that the company considers are decisive. We conclude that the candidate site in Venlo is the best option, at least when relocating only the depot in Venlo. Although Rhenus Road NL does not consider relocating another depot in the foreseeable future, this research also investigates two types of artificial scenarios that may be interesting for e.g. distant future relocations. Firstly, we analyze relocating a single other depot that Rhenus Road NL operates instead of the depot in Venlo. We find that Rhenus Road NL can gain noticeable transportation costs savings by relocating the depot in Hillegom to an artificial candidate site in Nieuwegein. Secondly, we examine relocating more than one depot that Rhenus Road NL operates. We show that no cost savings can be gained by relocating depots simultaneously instead of consecutively, nor by relocating more than one depot instead of only one. Hence, we see no reason to delay single depot relocations. We do however need to be careful drawing conclusions based on artificial scenarios. Despite the limitations of this research we recommend the company to relocate the depot that is currently located in Venlo to the candidate site in Venlo.

Preface

This thesis is the result of research that I have conducted at Rhenus Logistics, in order to obtain the master's degree in Operations Management and Logistics at Eindhoven University of Technology. Many people have supported me throughout this research, and a few stand out.

Dr. Martin, you have been an amazing first supervisor. I highly appreciate the time that you spent on meetings and providing feedback. You challenged me to get the most out of myself, and thereby you definitely improved the quality of this thesis. I would also like to thank Dr. Kinable, who has been my mentor for the largest part of my master studies, and later became my second supervisor. Your feedback was very helpful and it was great that you were very much on the same page with Dr. Martin.

In addition, I would like to thank Mr. Fricke. You have shown great involvement and I appreciate that from the beginning onwards I have been given trust and autonomy. Our meetings made me feel confident about the course of the research and have kept me highly motivated. I would also like to thank all the colleagues at Rhenus Logistics that provided me with the information needed.

Finally, I would like to express special thanks to my parents, Thea and Michael Bouwmans, and to my girlfriend, Sukunya Hathayang, for supporting me throughout my life as a student at Eindhoven University of Technology. It has been an amazing experience.

I have written this thesis with passion and enjoyment, and I wish you excitement reading it.

Frans Bouwmans

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Outsourcing logistics functions to 3PL companies is a source of competitive advantage for many companies (Ghaffari-Nasab et al., 2016). Companies cite greater operational efficiency, improved customer service levels, and a better focus on their core businesses (König and Spinler, 2016). 3PL companies offer a wide range of logistics services and have an increasingly important role in supply chain management (Motaghedi-Larijani et al., 2012). Road transport operations are among the most common outsourced logistics activities across all continents, and play a key role in reducing emissions (Perotti et al., 2012; Stojanović, 2017). To reduce the carbon footprint and be competitive, 3PL companies should have an efficient and highly responsive road freight forwarding network design that is able to meet customer demands.

1.1 Problem

Rhenus Road NL controls road freight forwarding services from, to, over, and within the Netherlands. They do so from their depots in Venlo, Oldenzaal and Hillegom. The manager of Rhenus Road NL Venlo faces two problems. Firstly, the depot in Venlo is split over two sites that are in close proximity to one another, causing operational inefficiencies. The manager of Rhenus Road NL Venlo aims to control all transfers from one building (with possibilities to expand) and for that purpose the current sites are not large enough. Secondly, the decision of placing a depot in Venlo was made thirty years ago based on gut feeling and meanwhile the customer base has developed. To counter the problems, Rhenus Road NL aims to investigate where they should relocate the depot in Venlo to, such that they minimize costs comprised of transportation costs and fixed depot establishment costs. To Rhenus Road NL it is not about whether they should relocate the depot, but where to. The company also considers qualitative characteristics and requirements in decision-making, on which we elaborate separately from quantitative modelling. Relocating depots is a strategic decision for Rhenus Road NL: they make such decisions less than once every decade. The depots in Oldenzaal and Hillegom are not split over two sites and replacing or expanding more than one facility is too large of a financial investment. It is also unlikely that Rhenus Road NL will add more depots in the foreseeable future.

1.2 Company description and contribution

Rhenus logistics is a large 3PL company that offers solutions globally for a wide range of industries along the entire supply chain. One of the most prominent solutions they offer for their customers is freight forwarding by road, denoted by Rhenus Road. The core business of Rhenus Road is as follows. From a depot of Rhenus Road, a truck picks up commodities at an initial sender. Thereafter, goods usually flow over the depot of departure, where they are consolidated based on common destination. From there, the goods can either be delivered to the final consignee, or they can be sent to another road transport depot that is more efficiently able to forward the goods. To achieve economies of scale, road transport depots arrange roundtrips with one another. A roundtrip is a trip to a place and back usually over the same route. The initial sender and final consignee together constitute an origin-destination (o-d) pair. In reality, orders of multiple o-d pairs are combined in forwarding freight. In this project we however abstract from the former because required data (e.g. time windows) is not documented and itineraries are planned manually.

Rhenus Road observes transfers to Rhenus Logistics' departments other than those to their road transport departments as transfer to any other client. E.g., transportation orders to move commodities by road from a facility of Rhenus Contract Logistics to a facility of Rhenus Air, are treated as any other order. The former is also reflected in the data. E.g., in the data of road transport depot A, the initial sender is always the location at which a truck of depot A picked up the goods, except when the goods came from another road transport depot. In addition, Rhenus Logistics controls their processes on a national, departmental level. E.g., in the Netherlands, the road freight forwarding operations are controlled from three locations (subdepartments) and together they form a road transport department: Rhenus Road NL. The subdepartments of a countries' road transport department collaborate relatively closely in allocating orders to meet their shared customer demand. Therefore, the depots that Rhenus Road NL operates are different from a theoretical point of view than any other road transport depot in the road freight forwarding network. Figure 1.1 depicts the relevant road transport depots from the point of view of Rhenus Road NL.

Figure 1.1: Relevant road transport depots

By relocating the depot that Rhenus Road NL currently operates in Venlo, the orders may be allocated differently over the depots that Rhenus Road NL operates. However, because we do not have access to the data of the depots that Rhenus Road NL does not operate, we assume that an order will be forwarded by the same relevant road transport depots that Rhenus Road NL does not operate as before the relocation. This is valid to assume according to company experts. Consequently, in order to properly reallocate orders, our interest is in the real connected transfer locations from the point of view of Rhenus Road NL. Figure 1.2 shows what this implies. For the relocation decision we are not interested in final consignee A, but in depot A. I.e., the original o-d pair is initial sender A - final consignee A, but the o-d pair that we consider in the relocation decision is initial sender A - depot A. The resulting logistics structure that we consider in this project is the one that Figure 1.3 depicts. Rhenus Road NL considers multi-allocation of clients to their depots, meaning that clients can be allocated to any depot of Rhenus Road NL. We assume that paths between o-d pairs visit at least one and at most two depots that Rhenus Road NL operates. Depot-depot connections (roundtrips) are at economies of scale. The depots that Rhenus Road NL operates have capacities and are fully interconnected. In contradiction, each depot that Rhenus Road NL operates only has connections with a subset of relevant road transport depots that Rhenus Road NL does not operate. The depot-depot connections do not change with a relocation of a depot of Rhenus Road NL. The new depot should have the same depot-depot connections as the current depot in Venlo. I.e., some links have zero capacity and hence, where the allocation of regular clients is unrestricted, the allocation decisions related to relevant road transport depots that Rhenus Road NL does not operate are limited by the fixed depot-depot connections. Compared to regular transfers, there is a relatively strict time frame in which roundtrips must take place. Therefore, some depot-depot connections have a reachability limit (maximum distance between depots).

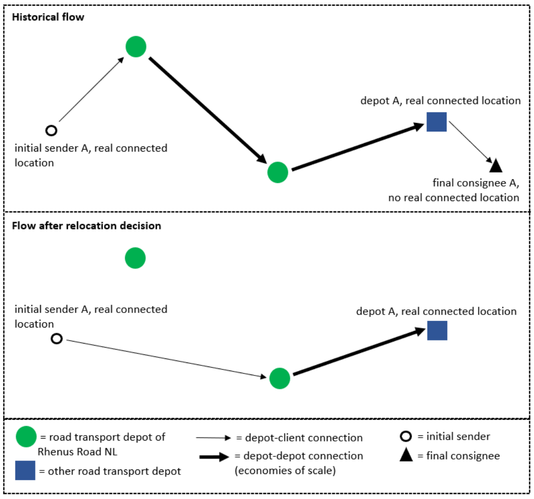

Figure 1.2: Example real connected transfer locations

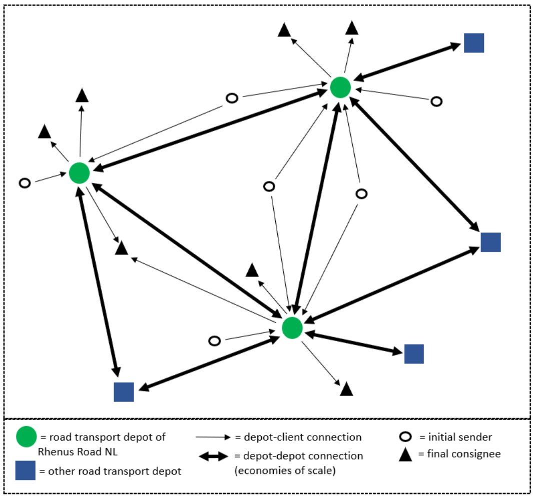

Figure 1.3: Logistics structure that we consider in this project

1.3 Research questions

To counter the problem of Rhenus Road NL, we develop a methodology for optimally relocating a subset of road transport depots for companies that have a logistics structure like Rhenus Road NL. We optimize the network in terms of minimizing costs comprised of transportation costs and fixed depot establishment costs. We apply the methodology on the problem of Rhenus Road NL in order to provide recommendations on relocating the depot that is currently located in Venlo. We also conduct numerical experiments on artificial scenarios of relocating a single other depot or more than one depot that Rhenus Road NL operates to obtain managerial insights and test the methodology's performance. The goal of this thesis is to answer the following main research question:

How can the network design of a company with a logistics structure like Rhenus Road NL be optimized by relocating a subset of their depots while respecting restricted node allocations and reachability limits for roundtrips with relevant road transport depots?

We formulate a set of sub-research questions that contribute in answering the main research question. First, the characteristics and requirements of the logistics structure of Rhenus Road NL need to be clear. Therefore, the following sub-research question is formulated:

1. *What are the characteristics and requirements of the logistics structure of Rhenus Road NL?*

The first sub-research question is answered in this chapter. This answer serves as input for conducting a literature review that contributes in finding how the strategic problem of this thesis can best be solved. This thesis contains a summary of the literature review that we wrote in a separate document. We formulate the second sub-research question as follows:

2. How can the problem of relocating a subset of road transport depots for companies with a logistics structure like Rhenus Road NL best be solved according to literature?

The answers to the previous sub-research questions serve as input for developing our own methodology. As such, we formulate the third sub-research question as follows:

3. How should the problem of relocating a subset of road transport depots for companies with a logistics structure like Rhenus Road NL be modelled mathematically?

We perform a data analysis and compute numerical results on the problem of Rhenus Road NL with the developed methodology, to answer the fourth sub-research question:

4. Which candidate depot locations are appealing in terms of costs to relocate the depot that is currently located in Venlo to?

We elaborate on qualitative characteristics and requirements of candidate locations, in order to decide upon which candidate location is the best option when relocating only the depot in Venlo. We also conduct numerical experiments on artificial scenarios of relocating a single other depot or more than one depot that Rhenus Road NL operates, to answer the fifth sub-research question:

5. How does our developed methodology perform when relocating other depots that Rhenus Road NL operates, and what managerial insights do we obtain?

Finally, we discuss the limitations of this research, and provide Rhenus Road NL with recommendations.

1.4 Outline of the report

Chapter 2 provides a summary of the conducted literature review that contributes in finding how the strategic problem of this thesis can best be solved. Hence, Chapter 2 answers sub-research questions 2. Chapter 3 presents the problem definition and methodology in order to answer sub-research question 3. In Chapter 4 we analyze the data. Chapters 5 and 6 elaborate on numerical results of the methodology, and answer sub-research questions 4 and 5, respectively. Chapter 7 provides the conclusion and discussion, and answers the main research question.

Chapter 2

Background and literature

This chapter provides a summary of the conducted literature review that aimed to determine what is known on the topic of this thesis, and how well this knowledge is established. The literature review contributes in finding how the strategic problem of this thesis can best be solved. Note that we aim to develop a methodology for relocating a subset of depots (and not just one). Location Science is a branch of optimization science that concerns itself with investigating where to physically locate a set of facilities so as to minimize or maximize an objective while satisfying customer demand and being subject to a set of constraints (Hakimi, 1964). The Location Problem is at the core of the Location Science principle. There is a large body of literature (dating back to Huff (1963)) in which customers self-select facilities, however, in our Location Problem we assign clients to facilities. We classify our problem as a Hub Location Problem (HLP). HLPs incorporate the spatial interaction between o-d pairs as well as between homogeneous facilities (Campbell and O'Kelly, 2012; Crainic, 2003).

2.1 HLP

O'Kelly (1986a), O'Kelly (1986b) and O'Kelly (1987) provide a key impetus for the growth of the HLP as a distinct Location Science research area. The HLP is an extension of the Facility Location Problem (FLP) and simultaneously includes elements of the Quadratic Assignment Problem (QAP) (Campbell and O'Kelly, 2012). Because both FLPs and QAPs are NP-hard, their combination tends to make HLPs at least as difficult (O'Kelly, 1987). In HLPs, facilities (hubs) act as consolidation, transshipment, and switching points for flow distribution between origins and destinations (spokes) (Alumur and Kara, 2008). Campbell and O'Kelly (2012) distinguish between five key features of HLPs. Firstly, there is flow between o-d pairs. Secondly, flows can or must go through hub facilities. Thirdly, hubs are facilities to be located. Fourthly, there is a benefit of routing flows via hubs (Alumur and Kara, 2008). Lastly, the objective is to minimize the sum of fixed hub establishment costs and transportation costs. Campbell (1994) elaborates on some other features that remain commonly assumed to investigate many distinct variants of HLPs (Campbell and O'Kelly, 2012). These are that most HLPs do not allow for direct flows between o-d pairs, paths between o-d pairs visit at most two hubs, and hubs are fully interconnected.

2.1.1 Traditional network definition

Static HLPs always consider a pure phase-in problem (and literature on multi-period HLPs only consider expansions). I.e. in case of static HLPs, prior to the location decision, there are no hubs present in the network, and the total set of nodes usually represents (artificial) demand centers that forward freight. With the location decision, some of the demand centers are converted into hubs, and the other nodes, the spokes, remain demand centers (O'Kelly and Miller, 1994; Yaman et al., 2012). The demand centers are regional facilities and in case they are artificial, they are created by aggregating the data to e.g. the post code district level. As a result, HLPs consider a complete graph of nodes. HLP literature also commonly considers only one set of nodes, in which each node corresponds to an origin, destination, and potential hub location simultaneously. Therefore, for most variants of HLPs, networks containing 200 nodes are regarded as being computationally very difficult. In this thesis, we refer to the hubs as depots that Rhenus Road NL operates and to the spokes as real connected transfer locations. So, unlike in traditional HLPs, in our problem we do not aggregate to a higher node level for the spokes, and we only combine spokes according to the actual orders.

2.1.2 Related business applications and problems

Research on Location Problems of large long-haul (3PL) road freight forwarders is relatively scarce. Such problems have mostly been modelled as a HLP (Abbasi et al., 2021; Alumur and Kara, 2009; Campbell, 2005; Cheong et al., 2007; Cunha and Silva, 2007; Ghaffari-Nasab et al., 2016; Hu et al., 2018; Kara and Tansel, 2002; Yaman et al., 2012). These papers however all consider (artificial) demand centers and a complete graph. The paper of Üster and Agrahari (2011) is the only paper that approaches a Location Problem for large road freight forwarders different. They consider a strategic level Multi-commodity Network Flow Problem (MNFP) with paired origins and destinations, in which they define each o-d pair as a commodity. MNFPs can be interpreted from the HLP perspective (Campbell and O'Kelly, 2012). In fact, if the location of the hubs are known a priori, the HLP reduces to a MNFP (Ebery et al., 2000). The decision variables in the formulation that Üster and Agrahari (2011) proposes are however not binary but integer about base capacities. MNFPs can be solved in polynomial time (Ouorou et al., 2000). When the number of potential hub configurations is low, it may be interesting to iteratively solve a MNFP. For brevity of this literature review, with the commonalities in the HLP and MNFP, and since we aim to develop a general methodology for relocating a subset of hubs and not just one, we leave MNFP literature out.

2.2 Median CMHLP

Because we consider a Median problem, capacitated hub nodes and multi-allocation of spokes to hubs, we look into literature on the Median Capacitated Multi-allocation Hub Location Problem (CMHLP) (Campbell, 1994). In addition to the key distinguishing and commonly assumed features of general HLPs that we mentioned earlier, the majority of Median CMHLP literature considers the following features: non-modular exogenous hub capacities imposed on all incoming flows, uncapacitated links, splittable commodities, endogenous number of hubs to place, single objective, discrete and static problem, the triangle inequality holds. The former is the case for all papers that we discuss in the remainder of this literature review, unless mentioned otherwise.

Campbell (1994) proposed the first Mixed Integer Programming (MIP) formulation for the (Median) CMHLP. Ebery et al. (2000) were the first to devote a paper to just the (Median) CMHLP, and also the first to provide a heuristic algorithm and computational study on the (Median) CMHLP. Ebery et al. (2000) introduce a Mixed Integer Linear Programming (MILP) formulation that is considerably more efficient than the one of Campbell (1994). In their formulation they only impose hub capacities on flows coming from nonhub nodes. Their heuristic algorithm is based on shortest paths, and they incorporate the upper bound that they obtain from this heuristic in a Linear Programming (LP) based branch-and-bound solution procedure. Computational experiments for up to 50 nodes show that the shortest path approach yields tight upper bounds. Boland et al. (2004) outline properties of the optimal solutions of Ebery et al. (2000), and develop preprocessing and cutting procedures that lead to a reduction in the computation time while maintaining similar gaps. For the same problem, Marín (2005) introduces a formulation that relaxes from the assumption that the triangle inequality holds. Contreras et al. (2012) present models for both splittable and nonsplittable commodity variants of the Median CMHLP. They consider modular hub capacities and only impose hub capacities on flows coming from nonhub nodes. They propose an exact algorithm based on Benders decomposition that is able to solve instances with up to 300 nodes and 5 distinct capacity levels to optimality in almost half a day. Rodriguez-Martin and Salazar-Gonzalez (2008) study a MILP formulation for the Median CMHLP in which they consider capacitated hubs and links, an incomplete backbone network, and more than two hub visits. They also do not assume that the triangle inequality holds. Sender and Clausen (2013) present a formulation for a Median CMHLP that considers modular hub capacities and an incomplete backbone network. Sasaki and Fukushima (2003) provide a model for the one-stop Median CMHLP, that incorporates both capacity constraints on hubs and links. Demır et al. (2019) introduce a model formulation for the multi-objective Median CMHLP and routing problem. They do not assume a complete backbone network, and consider both hub and link capacities. Monemi et al. (2021) present a model formulation for a bi-objective load balancing Median CMHLP.

Static and deterministic models do not capture many of the characteristics of real-world problems. Alumur et al. (2016) propose a MILP formulation for the multi-period modular Median CMHLP. They consider an incomplete backbone network, fixed costs for operating hub links, variable operational costs for the flow in the hubs, and more than two hub visits. Boukani et al. (2016) provide a robust optimization model that deals with uncertainty in the fixed setup cost and hub capacities. Correia et al. (2018) propose a formulation for the stochastic multi-period Median CMHLP with modular hub capacities, of which the uncertainty is in the demand.

2.3 Position of this research in literature

The papers of Campbell (1994), Ebery et al. (2000), Boland et al. (2004), and Contreras et al. (2012) are closest related to this research. These papers all consider a discrete and static Median CMHLP with deterministic parameter types, a single-criteria objective function, fixed hub establishment costs, a discounted inter-hub flow, a minimum of one and maximum of two hub stops, a complete backbone network, exogenous hub capacities, splittable commodities, and that the triangle inequality holds. Unlike in our study, these papers all consider an endogenous amount of hubs to place. Campbell (1994) is the only paper that incorporates hub capacities on all incoming flows. Contreras et al. (2012) is the only paper that considers non-modular hub capacities. The spoke types, and as such the level to which the data is aggregated, are demand centers in Campbell (1994) and Contreras et al. (2012), and post code districts in Ebery et al. (2000) and Boland et al. (2004).

Our problem represents a new variant of the Median CMHLP. On a high level, it is different from literature on Median CMHLPs in two ways. Firstly, literature on static Median CMHLPs always deals with a pure phase-in problem. We however study a static Median CMHLP that does not build the entire network from scratch but instead relocates a subset of facilities. Secondly, Median CMHLPs traditionally consider a complete graph of nodes where all nodes before the location decision are referred to as (artificial) demand centers (and the data is aggregated to this level). The optimal hub locations are then chosen from a subset of these (artificial) demand centers. We however use all real connected transfer locations as spokes, and we only combine spokes according to the actual orders. The optimal hubs are chosen from a set of candidate hub locations that does not intersect with the set of spokes. Our study is also the only paper on the Median CMHLP that considers reachability limits and capacity limits for some links.

Chapter 3

Problem formulation

In this chapter we present the methodology for relocating a subset of road transport depots for companies that have a logistics structure like Rhenus Road NL. Our problem represents a new variant of the Median CMHLP. Since we consider relocating a subset of hubs and a p-Median problem, we define our problem as the p-Median Subset CMHLP (SCMHLP).

3.1 Problem statement

The p-Median SCMHLP relocates a subset of p hubs to p optimal locations, through minimizing costs comprised of fixed hub establishment costs and transportation costs, by assigning spokes to hubs. These location-allocation decisions are solved at the strategic level. We refer to the depots that Rhenus Road NL operates as hubs, to the real connected transfer locations (see Chapter 1) as spokes, and to the road freight forwarding facilities that Rhenus Road NL does not operate as relevant road depots. We develop a static and discrete model with initially deterministic parameter types. The backbone network is complete and hubs have exogenous and non-modular capacities that are imposed on all incoming flows. Each hub only has connections with a subset of relevant road depots. I.e., some links between hubs and relevant road depots have zero capacity. Every to be relocated hub is replaced by exactly one other hub, and the relevant road depot connections remain the same as before the relocation. The hub that replaces another hub has the same relevant road depot connections as the hub that it replaces. All hubs have maximum allowed distances between one another, whereas only a subset of relevant road depots has reachability limits with hubs. Which relevant road depots have a reachability limit deviates per hub, but the maximum allowed distance is the same for all reachability limited connections. Freight is forwarded more effectively on roundtrips between hubs and with relevant road depots than on regular trips. We assume a minimum of one and maximum of two hub stops and we abstract from routing to combine orders of multiple o-d pairs. Furthermore, we assume continuous allocation of an o-d pair's flow: when the number of hubs is small compared to the number of o-d pairs that need to be routed, splittable variants provide a tight relaxation for nonsplittable variants of CMHLPs (Contreras et al., 2012). We also assume that each hub can be relocated to every candidate hub location, and that the capacity and fixed costs associated with a potential hub location depend on the hub placed.

We use the formulation of Ebery et al. (2000) as a starting point, which is found to be efficient in CMHLP literature. The defined problem is equivalent to a p-Median SCMHLP on graph $G\langle V, A\rangle$. V represents the set of vertices and A the set of arcs. The vertice set contains hubs R that the 3PL company can relocate to candidate locations P, and static hubs S (hubs that may not be relocated). P and S represent the set of hubs H that can be actively present in the network. The vertice set also contains customer origins O and destinations D. C represents the set of o-d pairs: unlike most HLPs, we consider a MNFP setting where not each origin sends commodities to each destination. Like the hubs, relevant road depots T permit for transshipments at economies of scale. Allowed relevant road depot connections may differ per hub so we introduce T^h. We let T_r represent the set of reachability limited relevant road depots for $r \in R$. For the zero capacity links with relevant road depots, we introduce O^h and C^h. O^h is the set of origins that are in T but not in T^h, for $h \in (R \cup S)$. C^h is the set of o-d pairs that have a destination in T but not in T^h, for $h \in (R \cup S)$. Regarding the parameters, there is a fixed cost F_p^r associated with opening hub $r \in R$ at candidate location $p \in P$. Q^s denotes the capacity for hub $s \in S$, and Q_p^r the capacity for a hub at $p \in P$ replacing hub $r \in R$. d_{ij} is the distance between i and j for origin-hub, hub-hub and hub-destination movements. w_{ij} is the flow between o-d pair $\langle i, j \rangle \in C$. B is the maximum allowed distance between some locations organizing roundtrips. B is independent of the locations. f is the transportation cost per distance for regular trips, and g for roundtrips ($g \leq f$). We consider three decision variables. h_k^r denotes the binary variable about whether hub $r \in R$ is relocated to candidate location $k \in P$. y_{kl}^i represents the continuous variable for flow from node $i \in O$ via hubs $k, l \in H$. x_{lj}^i denotes the continuous variable for flow between $\langle i, j \rangle \in C$ via hub $l \in H$. In y_{kl}^i and x_{lj}^i, k refers to the first hub, and l to the second. Table 3.1 provides an overview of the sets, parameters and decision variables.

Table 3.1: Definition of p-Median SCMHLP sets, parameters and decision variables

Sets	
V	Set of all nodes in the network
A	Set of arcs
H	Set of hubs that can be actively present in the network
S	Set of static hubs (hubs that may not be relocated)
R	Set of hubs that are eligible for relocation
P	Set of potential hub locations
C	Set of o-d pairs
O	Set of origin nodes
D	Set of destination nodes
T	Set of relevant road depots
T^h	Set of relevant road depots for $h \in (R \cup S)$
T_r	Set of relevant road depots for $r \in R$ with reachability limit
O^h	Set of origins that are in T but not in T^h, for $h \in (R \cup S)$
C^h	Set of o-d pairs that have a destination in T but not in T^h, for $h \in (R \cup S)$
Parameters	
F_p^r	Fixed costs for opening a hub $p \in P$ replacing hub $r \in R$
Q^s	Flow capacity for a hub at $s \in S$
Q_p^r	Flow capacity for a hub at $p \in P$ replacing hub $r \in R$
d_{ij}	Distance between i and j
w_{ij}	Flow between $\langle i, j \rangle \in C$
B	Maximum allowed distance between some locations
f	Transportation cost per distance for regular transfers
g	Transportation cost per distance for roundtrips
Decision variables	
h_k^r	Binary variable about whether hub $r \in R$ is relocated to candidate location $k \in P$
y_{kl}^i	Continuous variable for flow from node $i \in O$ via hubs $k, l \in H$
x_{lj}^i	Continuous variable for flow between $\langle i, j \rangle \in C$ via hub $l \in H$

Figure 3.1 depicts an example explaining the core parts of the formulation. We consider an origin i that is paired with destination (relevant road depot) j. Regarding y_{kl}^i, goods flow from origin i over static hub k and potential hub l (replacing one of the to be relocated hubs). For x_{lj}^i, goods flow from i to potential hub l, and thereafter to destination j (the i,l connection is at zero cost).

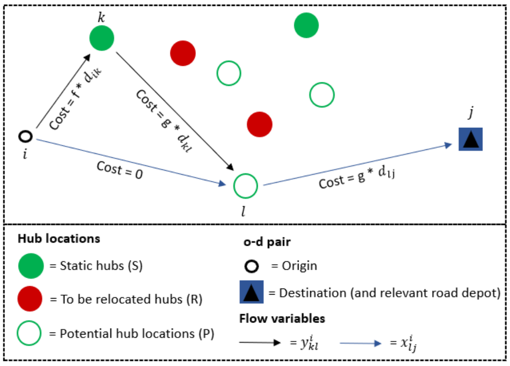

Figure 3.1: Example core formulation

3.2 Mathematical model formulation

We state the quadratically-constrained MILP for the p-Median SCMHLP below. We let c_{ij} be the transportation costs per unit flow for transfers between i and j. $c_{ij} = gd_{ij}$ for hub-hub movements and for trips between hubs and relevant road depots. $c_{ij} = fd_{ij}$ for any other trip.

$$\min \quad \sum_{k\in P}\sum_{r\in R} F_k^r h_k^r + \sum_{i\in O}\sum_{k\in H}\sum_{l\in H} c_{ik} y_{kl}^i + \sum_{i\in O}\sum_{k\in H}\sum_{l\in H} c_{kl} y_{kl}^i + \sum_{\langle i,j\rangle\in C}\sum_{l\in H} c_{lj} x_{lj}^i \tag{3.1}$$

$$\text{s.t.} \quad \sum_{k\in P} h_k^r = 1 \qquad \forall r \in R \tag{3.2}$$

$$\sum_{r\in R} h_k^r \leq 1 \qquad \forall k \in P \tag{3.3}$$

$$\sum_{l\in H} x_{lj}^i = w_{ij} \qquad \forall \langle i,j\rangle \in C \tag{3.4}$$

$$\sum_{\langle i,j\rangle\in C} x_{pj}^i \leq \sum_{r\in R} Q_p^r h_p^r \qquad \forall p \in P \tag{3.5}$$

$$\sum_{\langle i,j\rangle\in C} x_{sj}^i \leq Q^s \qquad \forall s \in S \tag{3.6}$$

$$\sum_{i\in O}\sum_{l\in H} y_{kl}^i \leq \sum_{r\in R}\sum_{\langle i,j\rangle\in C} w_{ij} h_k^r \qquad \forall k \in P \tag{3.7}$$

$$\sum_{k\in H} y_{kl}^i = x_{lj}^i \qquad \forall \langle i,j\rangle \in C, l \in H \tag{3.8}$$

$$h_k^r y_{kl}^i = 0 \qquad \forall i \in O^r, k \in P, l \in H, r \in R \tag{3.9}$$

$$y_{sl}^i = 0 \qquad \forall i \in O^s, l \in H, s \in S \tag{3.10}$$

$$h_k^r x_{kj}^i = 0 \qquad \forall \langle i,j\rangle \in C^r, k \in P, r \in R \tag{3.11}$$

$$x_{sj}^i = 0 \qquad \forall \langle i,j\rangle \in C^s, s \in S \tag{3.12}$$

$$h_k^r d_{ik} \leq B \qquad \forall i \in (T_r \cup S), k \in P, r \in R \tag{3.13}$$

$$h_k^r h_l^s d_{kl} \leq B \qquad \forall k,l \in P, r,s \in R \tag{3.14}$$

$$x_{lj}^i \geq 0 \qquad \forall \langle i,j\rangle \in C, l \in H \tag{3.15}$$

$$y_{kl}^i \geq 0 \qquad \forall i \in O, k,l \in H \tag{3.16}$$

$$h_k^r \in \{0,1\} \qquad \forall k \in P, r \in R \tag{3.17}$$

The objective function (3.1) minimizes costs comprised of hub investment costs and transportation costs. The transportation costs comprise costs for origin-hub, hub-hub, and hub-destination movements. Constraints (3.2) and (3.3) ensure that every to be relocated hub is replaced by exactly one hub and that at most one hub can be located at a candidate hub location, respectively. Constraints (3.4) establish that goods of each o-d pair must flow over hubs. Constraints (3.5) guarantee that incoming flows imposed on a potential hub do not exceed the capacity of the hub that it replaces. Constraints (3.6) assure that the hub capacities of static hubs are not exceeded. Constraints (3.7) ensure that there are no flows over first hubs that are not open. Constraints (3.8) establish that for each o-d pair the flow arrives in a second hub via the first hubs (flow conservation). The first and second hub can also be the same hub. Equalities (3.9)-(3.12) assure that there is no flow over some links between hubs and relevant road depots, whereas inequalities (3.13) and (3.14) take care of the maximum allowed distance between some locations organizing roundtrips. Constraints (3.15) and (3.16) set the domain of the flow variables to be continuous. Constraints (3.17) set the domain of hub placement variables to be binary. We linearize constraints (3.9), (3.11) and (3.14) as follows.

Lemma 3.2.1. *Let $u_{kl}^{ir} = h_k^r y_{kl}^i$, and M be sufficiently large. Then, constraints (3.9) are equivalent to (3.18) and (3.19):*

$$u_{kl}^{ir} = 0 \qquad \forall i \in O^r, k \in P, l \in H, r \in R \tag{3.18}$$

$$u_{kl}^{ir} \geq y_{kl}^i - M\left(1 - h_k^r\right) \qquad \forall i \in O^r, k \in P, l \in H, r \in R \tag{3.19}$$

Proof. h_k^r is a binary variable. If h_k^r takes value zero, then inequality (3.18) ensures that u_{kl}^{ir} will be zero as well, and inequality (3.19) ensures that u_{kl}^{ir} must be greater than or equal to a negative number. If h_k^r is one, inequality (3.19) establishes that u_{kl}^{ir} must be greater than or equal to y_{kl}^i, and inequality (3.18) guarantees that y_{kl}^i must then be zero. Hence, $h_k^r y_{kl}^i = u_{kl}^{ir}$. ■

Lemma 3.2.2. *Let $v_{kj}^{ir} = h_k^r x_{kj}^i$, and M be sufficiently large. Then, constraints (3.11) are equivalent to (3.20)-(3.21):*

$$v_{kj}^{ir} = 0 \qquad \forall \langle i, j\rangle \in C^r, k \in P, r \in R \tag{3.20}$$

$$v_{kj}^{ir} \geq x_{kj}^i - M\left(1 - h_k^r\right) \qquad \forall \langle i, j\rangle \in C^r, k \in P, r \in R \tag{3.21}$$

Proof. h_k^r is a binary variable. If h_k^r takes value zero, then inequality (3.20) ensures that v_{kj}^{ir} will be zero as well, and inequality (3.21) ensures that v_{kj}^{ir} must be greater than or equal to a negative number. If h_k^r is one, inequality (3.21) establishes that v_{kj}^{ir} must be greater than or equal to x_{kj}^i, and inequality (3.20) guarantees that x_{kj}^i must then be zero. Hence, $h_k^r x_{kj}^i = v_{kj}^{ir}$. ■

Lemma 3.2.3. *Let the binary variable $z_{kl}^{rs} = h_k^r h_l^s$. Then, constraints (3.14) equal (3.22)-(3.26):*

$$z_{kl}^{rs} d_{kl} \leq B \qquad \forall k, l \in P, r, s \in R \tag{3.22}$$

$$z_{kl}^{rs} \leq h_k^r \qquad \forall k, l \in P, r, s \in R \tag{3.23}$$

$$z_{kl}^{rs} \leq h_l^s \qquad \forall k, l \in P, r, s \in R \tag{3.24}$$

$$z_{kl}^{rs} \geq h_k^r + h_l^s - 1 \qquad \forall k, l \in P, r, s \in R \tag{3.25}$$

$$z_{kl}^{rs} \in \{0, 1\} \qquad \forall k, l \in P, r, s \in R \tag{3.26}$$

Proof. Inequalities (3.23) and (3.24) ensure that z_{kl}^{rs} will be zero if either or both the binary variables h_k^r or h_l^s are zero. Inequality (3.25) makes sure that z_{kl}^{rs} will take value one if both h_k^r and h_l^s are set to one. Hence, $h_k^r h_l^s = z_{kl}^{rs}$. ■

3.3 Special case: relocating only one hub

A special case of the p-Median SCMHLP occurs when we relocate only one hub. We focus on this special case because this is the problem of Rhenus Road NL. When relocating one hub, we consider two options. The first option is to slightly change the notation for the p-Median SCMHLP. The second option is related to the fact that if the location of the hubs are known a priori, the HLP reduces to a MNFP (Ebery et al., 2000). By defining each o-d pair as a commodity (Üster and Agrahari, 2011), we can solve a MNFP per candidate hub location. We prefer the MNFP approach due to its simplicity in formulation and solving. Given the computational simplicity of the MNFP approach and that customer demand for Rhenus Road NL has a stochastic nature, we incorporate stochastic demand in the MNFP approach and capture it by means of finite sets of scenarios (Correia et al., 2018). In order to integrate correlations in demand, each scenario contains demand of one day. Initially, we compose a scenario set per weekday (e.g. the scenario set of Monday contains all Mondays). Depending on statistical differences in weekday demand (see Chapter 4), scenario sets may be clustered. We assume that each scenario has the same probability of occurrence. Algorithm 1 provides an overview of the MNFP approach with stochastic demand.

Algorithm 1: MNFP approach with stochastic demand

```
Input: Prepared data
Output: Optimal hub location for the to be relocated hub
for every candidate hub location do
    for every scenario set do
        Compute the input data for the candidate hub and scenario set iteration;
        Solve the LP formulation for the MNFP with stochastic demand (see below) to
          obtain the minimum transportation costs;
    Compute the minimum total costs by summing the minimum transportation costs
      over all scenario sets, and adding the fixed costs for establishing the hub at the hub
      location under consideration;
Determine which candidate hub location results in the minimum total costs;
```

The problem with respect to one candidate hub iteration and one scenario set L is equivalent to a MNFP with stochastic demand on graph $G\langle V, A\rangle$. V represents the set of vertices and A the set of arcs. The vertice set contains static hubs S, the to be relocated hub R, and the candidate hub location P considered in the candidate hub iteration. P and S represent the hubs H_p that are present in the network in a candidate hub iteration. The vertice set also contains customer origins O and destinations D. A customer location is not present in a scenario if it has zero outgoing or incoming flow in the scenario: O^l and D^l are the customer origins and destinations present in scenario $l \in L$, respectively. Not each origin sends commodities to each destination, and C^l represents the set of o-d pairs (commodities) in scenario $l \in L$. A^l is the set of arcs in scenario $l \in L$ in a candidate hub iteration. Commodities flow from origin to hub, hub to hub, and hub to destination. $\delta_l^+(i)$ represents the set of outgoing arcs for node i in scenario $l \in L$ in a candidate hub iteration, and $\delta_l^-(j)$ the set of incoming arcs for node j in scenario $l \in L$ in a candidate hub iteration. Relevant road depots T permit for transshipments at economies of scale, and T^l represents the ones present in scenario $l \in L$. Allowed relevant road depot connections may differ per hub so we introduce T_h^l. We let T_r represent the set of reachability limited relevant road depots for $r \in R$. With respect to the parameters, each hub has a capacity that represents an upper bound on its incoming flow. Q_h denotes the daily flow capacity for hub $h \in H$ for scenario set L. d_{ij} is the distance between $\langle i, j\rangle \in A^l$ in scenario $l \in L$. w_k^l is the flow associated with $k \in C$ in scenario $l \in L$, and B is the maximum allowed distance between some locations organizing roundtrips. f is the transportation cost per distance for regular transfers, and g for roundtrips. x_{ijk}^l is the continuous variable for the arc flow over $\langle i, j\rangle \in A^l$ of o-d pair $k \in C$ in scenario $l \in L$. Table 3.2 provides an overview of the sets, parameters and decision variables.

Table 3.2: Definition of MNFP sets, parameters and decision variables

Sets	
V	Set of all nodes in the network
A	Set of arcs
L	Set of demand scenarios
H_p	Set of hubs present in the network in a candidate hub iteration
S	Set of static hubs
R	Set for the hub that is eligible for relocation
P	Set for the candidate hub location considered in the candidate hub iteration
O	Set of origin nodes in the scenario set
D	Set of destination nodes in the scenario set
C^l	Set of o-d pairs (commodities) in scenario $l \in L$
O^l	Set of origin nodes in scenario $l \in L$
D^l	Set of destination nodes in scenario $l \in L$
A^l	Set of arcs in a candidate hub iteration in scenario $l \in L$
$\delta_l^+(i)$	Set of outgoing arcs for node i in a hub iteration in scenario $l \in L$
$\delta_l^-(j)$	Set of incoming arcs for node j in a hub iteration in scenario $l \in L$
T	Set of relevant road depots
T^l	Set of relevant road depots present in scenario $l \in L$
T_h^l	Set of relevant road depots for $h \in H_p$ in scenario $l \in L$
T_r	Set of relevant road depots for $r \in R$ with reachability limit
Parameters	
Q_h	Daily flow capacity for hub $h \in H$ for scenario set L
d_{ij}^l	Distance between $\langle i,j\rangle \in A^l$ in scenario $l \in L$
w_k^l	Flow associated with $k \in C^l$ in scenario $l \in L$
B	Maximum allowed distance between some locations
f	Transportation cost per distance for regular transfers
g	Transportation cost per distance for roundtrips
Decision variables	
x_{ijk}^l	Continuous variable for the arc flow over $\langle i,j\rangle \in A^l$ of o-d pair $k \in C$ in scenario $l \in L$

We state the LP formulation for the MNFP with respect to one candidate hub iteration and one scenario set below. We let c_{ij}^l be the transportation costs for transfers over arc $\langle i,j\rangle \in A^l$ in scenario $l \in L$. $c_{ij}^l = g d_{ij}^l$ for hub-hub movements and for trips between hubs and relevant road depots. $c_{ij}^l = f d_{ij}^l$ for any other trip.

$$\min \quad \sum_{l\in L} \sum_{\langle i,j\rangle\in A^l} \sum_{k\in C^l} c_{ij}^l x_{ijk}^l \tag{3.27}$$

$$\text{s.t.} \quad \sum_{j\in\delta_l^+(i)} x_{ijk}^l = w_k^l \qquad \forall k \in C^l, i = O^l(k), l \in L \tag{3.28}$$

$$\sum_{i\in\delta_l^-(j)} x_{ijk}^l = w_k^l \qquad \forall k \in C^l, j = D^l(k), l \in L \tag{3.29}$$

$$\sum_{i\in\delta_l^-(j)} x_{ijk}^l - \sum_{i\in\delta_l^+(j)} x_{jik}^l = 0 \qquad \forall j \in H_p, k \in C^l, l \in L \tag{3.30}$$

$$\sum_{j \in \delta_l^+(i)} \sum_{k \in C^l} x_{ijk}^l \leq Q_i \qquad \forall i \in H_p, l \in L \tag{3.31}$$

$$x_{ijk}^l = 0 \qquad \forall i \in (T^l \backslash T_j^l), j \in H_p, k \in C^l, l \in L \tag{3.32}$$

$$x_{ijk}^l = 0 \qquad \forall j \in (T^l \backslash T_i^l), i \in H_p, k \in C^l, l \in L \tag{3.33}$$

$$d_{pj} \leq B \qquad \forall p \in P, j \in (S \cup T_r) \tag{3.34}$$

$$x_{ijk}^l \geq 0 \qquad \forall \langle i, j \rangle \in A^l, k \in C^l, l \in L \tag{3.35}$$

The objective function (3.27) minimizes costs comprised of investment costs for new hubs and transportation costs. Constraints (3.28) and (3.29) guarantee that for each commodity in each scenario all demand leaves the origin node, and arrives in the destination node, respectively. Constraints (3.30) ensure that no units will be left behind in the hub nodes in any of the scenarios. Constraints (3.31) guarantee that the total incoming flows handled by a hub on a day do not exceed the daily hub capacities. Links between relevant road depots that are not allowed in a scenario are determined by equalities (3.32) and (3.33), whereas inequalities (3.34) take care of the maximum allowed distance between some locations organizing roundtrips. Constraints (3.35) control that the decision variable is non-negative.

3.4 Concluding remarks solution approach

For the relocation of one hub, we apply the MNFP approach with stochastic demand. With this approach we can solve the required instances in a reasonable amount of time. It is insufficient to use deterministic o-d pair demand: if we use one scenario containing the average daily volume transferred for each o-d pair, the appeal of candidate locations is significantly different. When relocating more than one hub, a similar MNFP approach as the one described in this chapter is applicable. However, as the number of hubs to replace as well as the number of candidate locations grows, applying the p-Median SCMHLP instead of the MNFP approach makes increasingly more sense: the number of possible hub configurations is equivalent to $|P|^{|R|}$, and the MNFP approach is very likely going to be impractical in terms of computation time when relocating more than one hub. Moreover, because the company is not interested in relocating more than one hub in the foreseeable future, we rather apply the p-Median SCMHLP on artificial scenarios of relocating more than one hub to test the model performance as academic contribution. The p-Median SCMHLP is a computationally difficult problem and for this thesis it is sufficient to solve smaller instances under the assumption of deterministic demand.

Chapter 4

Data analysis

To be able to make a decision about where to relocate the hub that is currently located in Venlo to, we need clean data in the right format, and an understanding of patterns in the data.

4.1 Data preparation

In this section we provide a high level overview of the data preparation steps required to obtain the reference data. The reference data serves as input for computing the reference results on the problem of Rhenus Road NL, for which we apply the MNFP approach with stochastic demand. Appendix A provides the detailed data preparation description. We are provided raw data of orders between 2018 and 2020 of the hubs. In collaboration with company experts, we prepare the data for the (reachability limited) relevant road transport depots (per hub), and we use this data to compute the (coordinates of the) real connected transfer locations. Rhenus Road NL handles many different types of commodities. Therefore, the company uses different metrics to express order size. We decide in collaboration with company experts on which data to use for the flow parameter: the volume (m^3) of shipments. The volume is filled out for 85.9% of the orders, and we extrapolate missing values. We prepare the flow related data for each of the scenario sets using historical data. The data insights that we discuss in the next section contributes in preparing the flow data in terms of e.g. dealing with outliers. We split up the flow data in a training (years 2018 and 2019) and test (year 2020) data set. We define the current and eight potential hub locations by their coordinates. The hub capacities of the static hubs remain the same. In the reference data, the potential hub locations have the same capacities as the current hub in Venlo. We are provided the hub establishment costs for the potential hub locations. Assuming euclidean distances, we prepare the transportation cost per distance parameters: if we input the current hub configuration, the computed transportation cost should be equivalent to the actual transportation costs, with a similar percentage of orders using a hub-hub link.

4.2 Data insights

The information in this section is kept confidential and is only available to the university and company supervisors. The thesis is readable without this confidential information.

4.3 Concluding remarks data analysis

The most important points brought up in this chapter are as follows. We described how we prepare the reference data, which serves as input for computing the reference results on the problem of Rhenus Road NL. In computing results, we should take into account that the volume handled increases insignificantly over time, and that there is a slight western shift in demand over time.

Chapter 5

Results problem Rhenus Road NL

In this chapter we present results of the numerical experiments that we perform on the problem of Rhenus Road NL. For this problem we apply the MNFP approach with stochastic demand. We employ the Python programming language to implement the MNFP approach and use Gurobi to solve the MNFP model. We carry out the experiments on an Intel(R) Core(TM) i7-8750H @ 2.20 GHz and 16.0 GB RAM. Appendix C elaborates on the computation time of the MNFP approach.

5.1 Reference results

The reference data described in Chapter 4 serves as input for computing the reference results. The reference data is split up in training (years 2018 and 2019) and test (year 2020) data. Table 5.1 presents the training reference results, sorted by annual total costs. Δ represents the yearly difference with the lowest total costs. The total costs are a sum of optimal transportation costs and fixed investment costs. Costs for personnel, fuel, depreciation, maintenance, and overhead constitute the transportation costs. The fixed costs mainly comprise costs for buying land and construction. Table 5.1 does not include the current situation because we cannot properly benchmark against it in terms of total costs: the current hub in Venlo is split over two rented sites (causing operational inefficiencies) which have no possibilities to expand. This research is not about whether the hub should be relocated, but where to. The actual annual transportation costs are € 37,191,254 in the training data. The lowest annual total costs result by relocating the hub to the candidate location in Venlo. The transportation costs mainly cause the total cost differences.

Table 5.1: Results training reference data

Candidate location	**Transportation costs**	**Fixed costs**	**Total costs**	Δ	% Δ
Venlo	€ 37,185,463	€ 416,667	€ 37,602,129	€ -	- %
Eindhoven	€ 37,317,615	€ 473,333	€ 37,790,948	€ 188,819	0.50 %
Oss	€ 37,391,987	€ 413,333	€ 37,805,320	€ 203,191	0.54 %
Maastricht	€ 37,459,331	€ 410,000	€ 37,869,331	€ 267,202	0.71 %
Arnhem	€ 37,621,087	€ 426,667	€ 38,047,754	€ 445,625	1.19 %
Tilburg	€ 37,611,482	€ 443,333	€ 38,054,815	€ 452,686	1.20 %
Amersfoort	€ 38,048,688	€ 470,000	€ 38,518,688	€ 916,559	2.44 %
Dordrecht	€ 38,297,223	€ 453,333	€ 38,750,557	€ 1,148,428	3.05 %

Table 5.2 presents the test reference results, sorted by annual total costs. The actual annual transportation costs are higher (€ 39,362,996) in the test data, because demand increases at relatively constant distances between o-d pairs over time. Again, the annual total costs are lowest for the candidate location in Venlo, and the candidate site rankings are similar over the data sets.

Table 5.2: Results test reference data

Candidate location	**Transportation costs**	**Fixed costs**	**Total costs**	**Δ**	**% Δ**
Venlo	€ 39,375,400	€ 416,667	€ 39,792,067	€ -	- %
Eindhoven	€ 39,454,545	€ 473,333	€ 39,927,878	€ 135,812	0.34 %
Oss	€ 39,553,250	€ 413,333	€ 39,966,583	€ 174,517	0.44 %
Maastricht	€ 39,577,225	€ 410,000	€ 39,987,225	€ 195,158	0.49 %
Tilburg	€ 39,738,388	€ 443,333	€ 40,181,722	€ 389,655	0.98 %
Arnhem	€ 39,803,075	€ 426,667	€ 40,229,741	€ 437,674	1.10 %
Amersfoort	€ 40,226,715	€ 470,000	€ 40,696,715	€ 904,648	2.27 %
Dordrecht	€ 40,495,887	€ 453,333	€ 40,949,220	€ 1,157,154	2.91 %

The (percentage) differences with the lowest annual total costs however imply that the MNFP approach has difficulties differentiating between (arguably some) candidate locations. Because the percentage differences with the lowest annual total costs are even smaller for all locations in the test data set, differentiating between candidate locations may be even more difficult when taking into account future trends. Given that we make assumptions, and that there is inevitably going to be inaccuracy in the data, we cannot decide on the optimal location solely based on numerical results. Zooming in on e.g. domestic level data to resolve this issue is not an option to the company, because this will bias the outcomes. Qualitative characteristics and requirements that the company considers may be decisive, and we will discuss them in Chapter 7.

Figures 5.1 and 5.2 depict the geographical patterns of appealing candidate hub locations, for the training and test data, respectively. Because the new hub must be operated by Rhenus Road NL, it must be located in the Netherlands. The markers in the figures represent the current hub locations (in blue) and the candidate locations. The numbers in the markers refer to the position in the total cost ranking. The annual total costs are relatively low for candidate locations relatively close to the current hub in Venlo. The decrease in percentage differences with the lowest annual total costs over the data sets can be explained by the slight western geographical shift in demand over time, causing candidate sites located more west to be more appealing with more recent data. The candidate locations in Amersfoort and Dordrecht violate the reachability limit constraints. By relaxing these constraints we examine whether it may be interesting to redesign the network more radically in terms of roundtrip schedules. The results do however not encourage this.

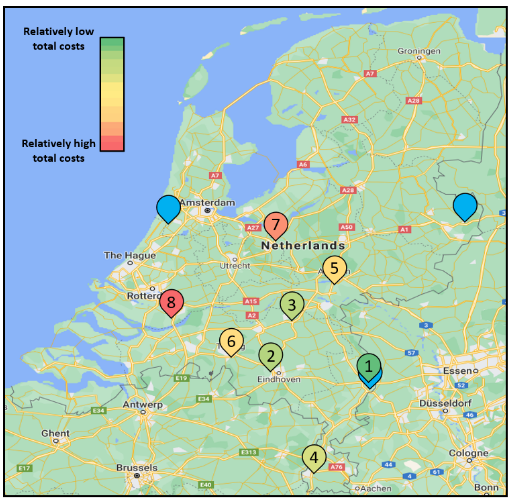

Figure 5.1: Geo-patterns appeal training

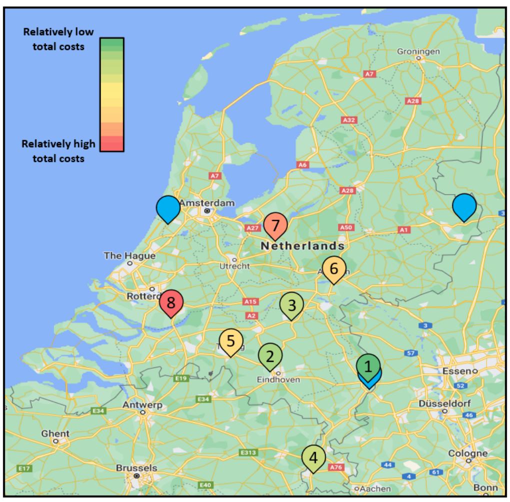

Figure 5.2: Geo-patterns appeal test

With the current hub configuration, 10.5% of the orders are assigned to different hubs. This shows that although we make assumptions, the MNFP approach does a decent job realistically assigning o-d pairs to hubs. Table 5.3 presents the percentage of orders that are assigned to different hubs for each of the candidate locations. Naturally, these percentages are lower for candidate locations relatively close to the current hub in Venlo. Candidate locations are generally more appealing when more orders are assigned to the *same* hubs as before the relocation. There is no pattern in the orders that are assigned differently. Table 5.3 also provides insights on the the hub link usage percentage (the percentage of orders using a hub-hub link). This percentage is higher when the distance between a candidate location and the static hubs is higher, which makes sense because then there is a larger distance over which the benefits of hub links can be exploited. Although candidate locations with a relatively high hub link usage percentage are in general relatively appealing, it is not necessarily advantageous to locate the hub as far as possible from the static hubs. The appeal of a candidate hub location also depends on the o-d pair (al)locations in relation to the entire candidate hub configuration, and the impact of sub-optimal routings that are a result of hub capacities.

Table 5.3: Reference results other insights

Candidate location	**% orders assigned to different hubs**	**Hub link usage %**
Venlo	10.7 %	12.2 %
Eindhoven	13.0 %	11.4 %
Oss	12.9 %	10.8 %
Maastricht	11.5 %	13.6 %
Arnhem	17.1 %	10.5 %
Tilburg	20.1 %	11.2 %
Amersfoort	22.4 %	9.5 %
Dordrecht	25.8 %	10.1 %

5.2 Outline remainder of this chapter

In the remainder of this chapter, we put the reference results in perspective. First, we validate the reference results, by examining the patterns over the working days, and the effects of randomly assigning training and test data. Thereafter, we perform a sensitivity analysis in which we investigate how changes in parameters influence the appeal of hub configurations. We change the flow data (by using self-created stochastic demand), the transportation cost per distance parameters, and the hub capacities. Finally, we perform a scenario analysis that takes the form of controlled experiments, to examine the effects of increasing demand, the western shift in demand, and Brexit.

5.3 Validation

5.3.1 Working day analysis

Table 5.4 presents results in terms of (sorted) annual total costs per working day. The location rankings are similar over the working days. For each of the working days, the lowest annual total costs result by relocating the hub to the candidate location in Venlo. The candidate locations in Oss, Eindhoven, and Maastricht tend to switch places in the ranking, and so do the candidate locations in Arnhem and Tilburg. On Mondays, the center of demand is slightly more northern. On Tuesdays, Wednesdays and Thursdays, the center of demand is slightly more western. On Fridays, the center of demand is slightly more south-eastern. The relatively low distances between o-d pairs on Mondays cause that the annual total costs are relatively low on Mondays. The annual total costs are relatively high on Fridays because the volume handled on Fridays is relatively high.

Table 5.4: Results per working day

Monday		Tuesday		Wednesday	
Location	**Total costs**	**Location**	**Total costs**	**Location**	**Total costs**
Venlo	€ 5,847,205	Venlo	€ 7,023,247	Venlo	€ 6,922,388
Oss	€ 5,865,284	Eindhoven	€ 7,034,253	Eindhoven	€ 6,960,189
Eindhoven	€ 5,869,912	Oss	€ 7,046,162	Oss	€ 6,966,066
Maastricht	€ 5,900,700	Maastricht	€ 7,057,731	Maastricht	€ 6,984,168
Arnhem	€ 5,905,498	Tilburg	€ 7,073,979	Tilburg	€ 7,012,757
Tilburg	€ 5,912,959	Arnhem	€ 7,109,186	Arnhem	€ 7,018,896
Amersfoort	€ 5,982,888	Amersfoort	€ 7,186,130	Amersfoort	€ 7,107,131
Dordrecht	€ 6,035,406	Dordrecht	€ 7,204,264	Dordrecht	€ 7,160,327
Thursday		**Friday**			
Location	**Total costs**	**Location**	**Total costs**		
Venlo	€ 7,613,640	Venlo	€ 10,925,628		
Eindhoven	€ 7,647,052	Maastricht	€ 10,964,919		
Oss	€ 7,651,369	Eindhoven	€ 10,974,846		
Maastricht	€ 7,661,281	Oss	€ 10,984,640		
Tilburg	€ 7,695,586	Tilburg	€ 11,042,874		
Arnhem	€ 7,698,840	Arnhem	€ 11,048,945		
Amersfoort	€ 7,788,040	Amersfoort	€ 11,176,666		
Dordrecht	€ 7,832,362	Dordrecht	€ 11,237,096		

5.3.2 Randomly assigning training and test data

In this subsection we examine the influence of seasonal factors and trends, by randomly assigning orders to the training and test data sets (instead of based on years like in the reference data). Table 5.5 presents the results in terms of (sorted) annual total costs for the randomly assigned data. The average daily demand is relatively high for the training data, and relatively low for the test data. Therefore, the annual total costs are quite different over the two data sets. The candidate location rankings and percentage differences with the lowest annual total costs remain similar. Hence, we can disregard the influence of seasonal factors on the appeal of candidate hub locations. The random assignment of data causes the future trend related pattern to disappear.

Table 5.5: Results randomly assigned training and test data

Training data			Test data		
Location	**Total costs**	**% Δ**	**Location**	**Total costs**	**% Δ**
Venlo	€ 38,925,202	- %	Venlo	€ 37,145,922	- %
Eindhoven	€ 39,076,115	0.39 %	Eindhoven	€ 37,307,545	0.44 %
Oss	€ 39,119,442	0.50 %	Oss	€ 37,338,340	0.52 %
Maastricht	€ 39,174,323	0.64 %	Maastricht	€ 37,377,242	0.62 %
Arnhem	€ 39,362,567	1.12 %	Tilburg	€ 37,542,225	1.17 %
Tilburg	€ 39,374,564	1.15 %	Arnhem	€ 37,600,115	1.22 %
Amersfoort	€ 39,841,359	2.35 %	Amersfoort	€ 38,051,371	2.45 %
Dordrecht	€ 40,103,948	3.03 %	Dordrecht	€ 38,242,437	3.05 %

5.4 Sensitivity analysis

5.4.1 Flow data

In this subsection we analyze the sensitivity of the results to the flow data by using self-created stochastic demand as input, instead of historical demand. This will allow us to draw stronger conclusions from our results as we mitigate the influence of variance in the input data. We generate stochastic demand by taking samples from demand distributions that we determine from the historical data. Appendix B provides a more detailed description on the demand generation procedure. Table 5.6 presents the results in terms of (sorted) annual total costs for the self-created stochastic demand. The candidate locations in Venlo and Maastricht are more appealing, because the center of demand is slightly more south-eastern. Since we do not take into account trends in self-creating demand (because the infrequencies of o-d pairs are such that we cannot estimate demand e.g. per year), we do not observe a future trend related pattern.

Table 5.6: Results self-created stochastic demand

Training data			Test data		
Location	**Total costs**	**% Δ**	**Location**	**Total costs**	**% Δ**
Venlo	€ 38,210,650	- %	Venlo	€ 38,381,876	- %
Maastricht	€ 38,308,272	0.25 %	Maastricht	€ 38,450,198	0.18 %
Eindhoven	€ 38,549,052	0.89 %	Eindhoven	€ 38,681,383	0.78 %
Oss	€ 38,559,470	0.93 %	Oss	€ 38,799,870	1.09 %
Arnhem	€ 38,742,505	1.39 %	Arnhem	€ 38,955,726	1.50 %
Tilburg	€ 38,814,674	1.58 %	Tilburg	€ 39,018,329	1.66 %
Dordrecht	€ 39,252,548	2.73 %	Dordrecht	€ 39,432,713	2.74 %
Amersfoort	€ 39,488,265	3.34 %	Amersfoort	€ 39,563,269	3.08 %

5.4.2 Transportation cost per distance parameters

In this subsection we investigate the sensitivity of the results to the transportation cost per distance parameters. We distinguish between two transportation cost per distance parameters (f and g) because trucks on roundtrips forward freight more effectively. For brevity of this analysis, we do not distinguish between training and test data.

Change f and g simultaneously with the same factor

Any transportation cost component influencing both transportation cost per distance parameters equally can cause this types of change. Example cost components are higher or lower costs for personnel, and a higher or lower diesel price. Realistic changes are within the -25% and 25% interval, and we change f and g in steps of 5%. Figure 5.3 shows that the annual transportation costs divided by the change factor changes linearly for all candidate locations. This implies that the assignment of orders to hubs is independent of the change factor. Figure 5.4 shows that the annual total costs divided by the change factor decreases nonlinearly for all candidate locations. This nonlinear decrease originates from the fixed costs being held constant, while being part of the annual total costs that is divided by the change factor. Overall, the appeal of candidate locations is insensitive to changing f and g simultaneously with the same factor: the largest change in the percentage differences with the lowest annual total costs is 0.04%, and the candidate location rankings are the same for all changes. For candidate locations with lower fixed costs than the candidate location in Venlo, the percentage differences with the lowest annual total costs are slightly higher when increasing f and g. For candidate locations with a higher fixed costs than the candidate location in Venlo, the percentage differences with the lowest annual total costs are slightly lower when increasing f and g. This makes sense because the impact of the fixed costs is lower when the transportation costs increase. The findings are reversed when decreasing f and g.

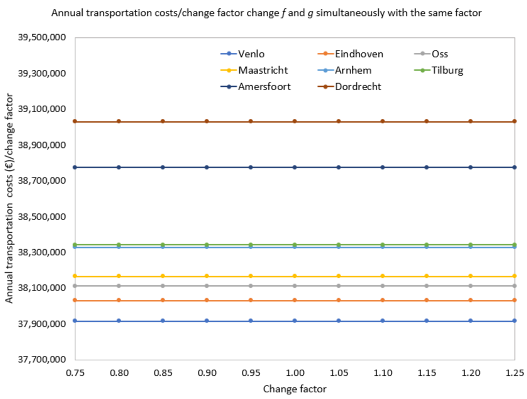

Figure 5.3: Change f and g simultaneously: annual transportation costs/change factor

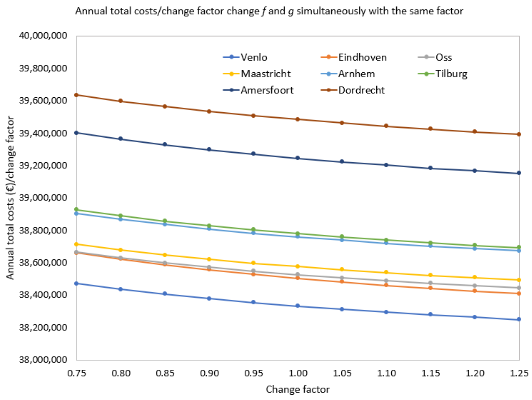

Figure 5.4: Change f and g simultaneously: annual total costs/change factor

Change only f

Different costs per unit flow that only influence regular trips can cause this type of change (e.g. different truck capacity utilization for regular trips). Realistic changes are within the -20% and 20% interval (then, $g \leq f$), and we change f in steps of 5%. Figure 5.5 shows that the annual transportation costs divided by the change factor changes nonlinearly for all candidate locations. The nonlinearity implies that as expected, the assignment of orders to hubs depends on the f to g ratio. The decrease can be explained by the fact that g is held constant. Figure 5.6 shows that the annual total costs divided by the change factor decreases nonlinearly for all candidate locations. The nonlinear decrease is caused by both the fixed costs and transportation costs. The appeal of candidate locations is relatively insensitive to changing f: the largest change in the percentage difference with the lowest annual total costs is 0.38%, and the candidate location rankings are the same for almost all changes. From an increase in f of approximately 10% onwards, the candidate location in Maastricht is more appealing than the candidate locations in Oss and Eindhoven. When increasing f, the percentage differences with the lowest annual total costs increase for all candidate locations except Maastricht. This makes sense because with higher f it is more important that a candidate location is better able to exploit the advantages of using hub links. Therefore, candidate locations with a relatively high hub link usage percentage are more appealing. The findings are reversed when decreasing f.

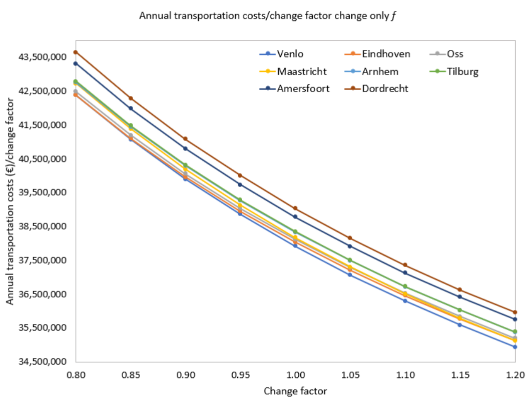

Figure 5.5: Change only f: annual transportation costs/change factor

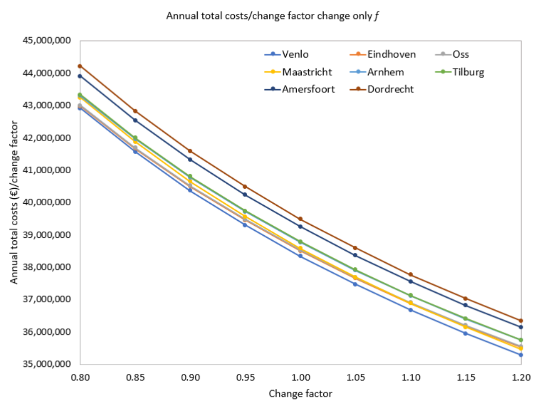

Figure 5.6: Change only f: annual total costs/change factor

Change only g

Different costs per unit flow that only influence roundtrips can cause this type of change. Realistic changes are within the -20% and 20% interval (then, $g \leq f$), and we change g in steps of 5%. Figure 5.7 shows the annual transportation costs divided by the change factor for all candidate locations. Figure 5.8 shows the annual total costs divided by the change factor for all candidate locations. The appeal of candidate locations is relatively insensitive to changing g: the largest change in the percentage differences with the lowest annual total costs is 0.37%, and the candidate location rankings are the same for almost all changes. From a decrease in g of approximately 10% onwards, the candidate location in Maastricht is more appealing than the candidate location in Oss. From a decrease in g of approximately 15% onwards, the candidate location in Maastricht is more appealing than the candidate location in Eindhoven. When decreasing g, the percentage differences with the lowest annual total costs increase for all candidate locations except Maastricht. This makes sense because with lower g it is more important that a candidate location is better able to exploit the advantages of using hub links. Therefore, candidate locations with a relatively high hub link usage percentage are more appealing. The findings are reversed when increasing g.

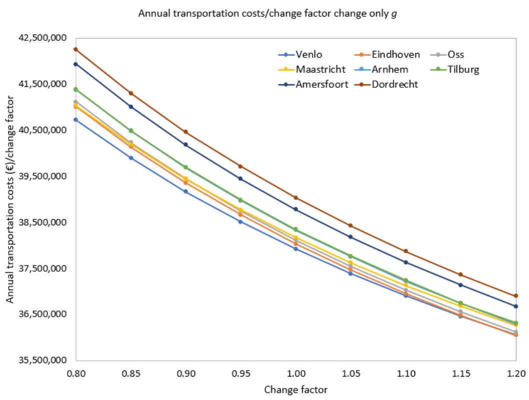

Figure 5.7: Change only g: annual transportation costs/change factor

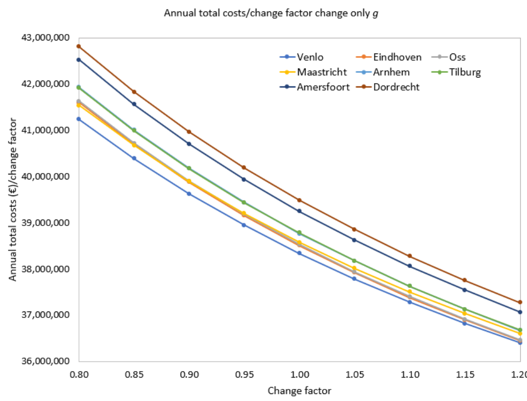

Figure 5.8: Change only g: annual total costs/change factor

5.4.3 Hub capacities

The hub capacities are estimated (see Appendix A how) and therefore we analyze the sensitivity of the results to changing hub capacities in this subsection. For brevity of this analysis, we do not distinguish between training and test data. Realistic changes (ensuring feasibility of the model) are within the -10% and 10% interval, and we change the hub capacities in steps of 2.5%. Figure 5.9 depicts the annual total costs for each of the candidate locations. The hub capacities have a significant impact on the annual total costs. Lower hub capacities increase the annual total costs (until the model is infeasible) because more o-d pair flows are routed sub-optimally. Higher hub capacities decrease the annual total costs (until the capacities result in no sub-optimal routings) because more o-d pair flows are routed optimally. The descending slope decreases as the hub capacities increase, which makes sense because sub-optimal routings that are a result of hub capacities get worse as the hub capacities decrease. The appeal of candidate locations is relatively insensitive to changing the hub capacities: the largest change in the percentage differences with the lowest annual total costs is 0.14%, and the candidate location rankings are the same for all changes. When decreasing the hub capacities, the percentage differences with the lowest annual total costs decrease for all candidate locations except Maastricht. This makes sense, because with lower hub capacities there are more sub-optimal routings, and sub-optimal routings generally have a larger negative impact on candidate locations for which the distance to the static hubs is larger. This negative impact seems however not to be very large, which can be explained by the fact that there is a larger increase in the number of sub-optimal routings for candidate locations for which the distance to the static hubs is lower. The findings are reversed when increasing the hub capacities.

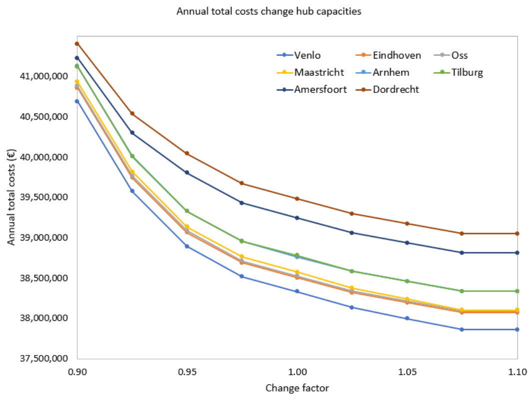

Figure 5.9: Change hub capacities: annual total costs

5.5 Scenario analysis

5.5.1 Increasing demand

The volume that Rhenus Road NL handled per week increases insignificantly over time. In this subsection, we consider increasing demand with constant geographical spread. Rhenus Road NL aims to absorb most of the increasing demand by installing more capacity at the hub that replaces the current hub in Venlo. The candidate locations have similar possibilities in terms of future expansion: approximately 50% more capacity in comparison with the capacity of the current hub in Venlo. By increasing demand with 24%, we consider a similar demand/capacity ratio as in the current situation. Table 5.7 presents the results in terms of (sorted) annual total costs for an increase in demand of 24%. The annual total costs increase on average 30%. Naturally, this is higher than the 24% increase in demand, partially because the fixed costs are higher when installing more capacity, but mainly because one hub absorbs most increasing demand. The percentage differences with the lowest annual total costs increase for all candidate locations. I.e., the location in Venlo is more appealing when increasing demand, which partially makes sense because demand increases with constant geographical spread. This however also shows that the impact of sub-optimal routings is not very large: more commodities are routed sub-optimally because the capacities of the static hubs are held constant, and sub-optimal routings have a larger negative impact on candidate locations for which the distance to the static hubs is larger. The increase in percentage differences with the lowest annual total costs is indeed relatively large for the candidate location in Maastricht.

Table 5.7: Results increase demand

Training data			**Test data**		
Location	**Total costs**	**% Δ**	**Location**	**Total costs**	**% Δ**
Venlo	€ 48,602,767	- %	Venlo	€ 52,229,873	- %
Eindhoven	€ 48,947,846	0.71 %	Eindhoven	€ 52,464,907	0.45 %
Oss	€ 49,030,471	0.88 %	Oss	€ 52,517,137	0.55 %
Maastricht	€ 49,186,000	1.20 %	Maastricht	€ 52,616,374	0.74 %
Arnhem	€ 49,317,227	1.47 %	Tilburg	€ 52,825,293	1.14 %
Tilburg	€ 49,365,830	1.57 %	Arnhem	€ 52,877,523	1.24 %
Amersfoort	€ 49,827,556	2.52 %	Amersfoort	€ 53,446,829	2.33 %
Dordrecht	€ 50,153,195	3.19 %	Dordrecht	€ 53,781,100	2.97 %

5.5.2 Brexit

Data shows that demand in the United Kingdom (UK) was in particular relatively high in 2020, which explains for the most part the western geographical shift in demand and therefore candidate location appeal. Brexit, the withdrawal of the UK from the European Union (EU), may however cause a significant decrease in demand in the UK. Therefore, instead of analyzing a future western shift in demand, we rather investigate the influence of Brexit. As of January 1st 2021, new rules apply to trade between the EU and UK (European Commision, 2021). This has caused road freight to be down by approximately a third between the EU and UK in early 2021 (The Guardian, 2021). We do not have flow data of Rhenus Road NL of 2021, but company experts also experience a significant decrease of demand in the UK. Table 5.8 presents the results in terms of (sorted) annual total costs for a decrease of demand in the UK of 33%. The candidate sites located more (south-)east are more appealing, while the percentage differences with the lowest annual total costs increase for all candidate locations. The percentage differences with the lowest annual total costs is higher for all candidate locations except Maastricht in the test data set. Hence, we rather expect a future (south-)eastern geographical shift in the appeal of candidate locations.

Table 5.8: Results Brexit

Training data			**Test data**		
Location	**Total costs**	**% Δ**	**Location**	**Total costs**	**% Δ**
Venlo	€ 34,933,737	- %	Venlo	€ 37,124,347	- %
Maastricht	€ 35,279,948	0.99 %	Maastricht	€ 37,481,313	0.96 %
Oss	€ 35,341,406	1.17 %	Oss	€ 37,563,451	1.18 %
Eindhoven	€ 35,422,720	1.40 %	Eindhoven	€ 37,716,541	1.60 %
Arnhem	€ 35,541,370	1.74 %	Arnhem	€ 37,797,008	1.81 %
Tilburg	€ 35,815,620	2.52 %	Tilburg	€ 38,214,507	2.94 %
Amersfoort	€ 36,327,242	3.99 %	Amersfoort	€ 38,620,050	4.03 %
Dordrecht	€ 36,771,950	5.26 %	Dordrecht	€ 39,094,379	5.31 %

Chapter 6

Results artificial scenarios

Rhenus Road NL does not consider relocating another hub than the current hub in Venlo. The current hub in Venlo is split over two rented sites (causing operational inefficiencies) which have no possibilities to expand. In addition, there is not enough budget to relocate multiple hubs. Still, it is interesting to investigate several artificial scenarios for e.g. distant future relocations and methodology performance. First, we elaborate on managerial insights of relocating a single other hub. Thereafter, we elaborate on the computation time and managerial insights of relocating more than one hub. For brevity of the analyses we do not distinguish between training and test data. When relocating the hub in Venlo, and inputting the reference data without splitting it, the lowest annual total costs are € 38,332,108. To be consistent with the problem of Rhenus Road NL, we assume that the hubs under consideration are relocated anyways, to a buyable site. Because the current sites are rented, we do not benchmark against the current situation in terms of total costs. The actual annual transportation cost are € 37,915,168.

6.1 Relocating a single other hub

In this section we apply the MNFP approach with stochastic demand to examine whether Rhenus Road NL can gain significant cost savings by relocating a single other hub. We locate artificial candidate sites at a distance of 0, 50 and 100 kilometer in every cardinal and ordinal direction from the hub that we relocate, while ensuring that the candidate sites are located in the Netherlands. We let the fixed hub establishment costs depend on the municipality in which a candidate site is located, as well as on the current site size of the hub that we relocate to the candidate location. The fixed costs mainly comprise costs for buying land and construction, and we use municipal reports on real estate to estimate these costs.

Hillegom
Table 6.1 presents the results in terms of (sorted) annual total costs when we relocate the hub in Hillegom. The lowest annual total costs result by relocating the hub in Hillegom to the artificial candidate location in Nieuwegein. When relocating the hub in Hillegom instead of the hub in Venlo, the lowest annual total costs are € 250,904 lower per year. Most of this difference originates from lower transportation costs. Figure 6.1 shows that candidate sites located more south(-west) are more appealing when relocating the hub in Hillegom.

Table 6.1: Results relocating the hub in Hillegom

Candidate location	**Transportation costs**	**Fixed costs**	**Total costs**	**∆**	**% ∆**
Nieuwegein	€ 37,791,204	€ 290,000	€ 38,081,204	€ -	- %
Rotterdam	€ 37,859,277	€ 300,000	€ 38,159,277	€ 78,072	0.21 %

Hillegom	€ 37,915,168	€ 266,667	€ 38,181,834	€ 100,630	0.27 %
Roosendaal	€ 37,970,805	€ 253,333	€ 38,224,139	€ 142,934	0.38 %
Almere	€ 37,997,886	€ 260,000	€ 38,257,886	€ 176,681	0.46 %
Uden	€ 38,059,622	€ 268,000	€ 38,327,622	€ 246,418	0.65 %
Enkhuizen	€ 38,277,018	€ 260,000	€ 38,537,018	€ 455,813	1.20 %
Schagen	€ 38,297,125	€ 265,000	€ 38,562,125	€ 480,921	1.26 %
Apeldoorn	€ 38,397,503	€ 258,333	€ 38,655,837	€ 574,632	1.51 %
Heerenveen	€ 38,648,320	€ 250,000	€ 38,898,320	€ 817,116	2.15 %

Oldenzaal

Table 6.2 presents the results in terms of (sorted) annual total costs when we relocate the hub in Oldenzaal. The lowest annual total costs result by relocating the hub in Oldenzaal to the artificial candidate location in Oldenzaal. When relocating the hub in Oldenzaal instead of the hub in Venlo, the lowest annual total costs are € 200,273 lower per year. This difference fully originates from lower fixed costs: the hub in Oldenzaal is smaller than the hub in Venlo, and the candidate sites are different. When relocating the hub in Oldenzaal instead of the hub in Hillegom, the lowest annual total costs are € 50,631 higher per year. Figure 6.2 shows that candidate sites located more east are more appealing when relocating the hub in Oldenzaal.

Table 6.2: Results relocating the hub in Oldenzaal

Candidate location	**Transportation costs**	**Fixed costs**	**Total costs**	**Δ**	**% Δ**
Oldenzaal	€ 37,915,168	€ 216,667	€ 38,131,835	€ -	- %
Emmen	€ 38,054,054	€ 216,667	€ 38,270,721	€ 138,886	0.36 %
Doetinchem	€ 38,146,809	€ 218,333	€ 38,365,142	€ 233,307	0.61 %
Deventer	€ 38,171,715	€ 216,667	€ 38,388,382	€ 256,547	0.67 %
Veendam	€ 38,197,307	€ 211,667	€ 38,408,973	€ 277,139	0.73 %
Meppel	€ 38,238,141	€ 206,667	€ 38,444,808	€ 312,973	0.82 %
Nijmegen	€ 38,364,841	€ 226,667	€ 38,591,507	€ 459,673	1.21 %
Ermelo	€ 38,425,781	€ 205,000	€ 38,630,781	€ 498,946	1.31 %
Heerenveen	€ 38,471,419	€ 208,333	€ 38,679,752	€ 547,917	1.44 %

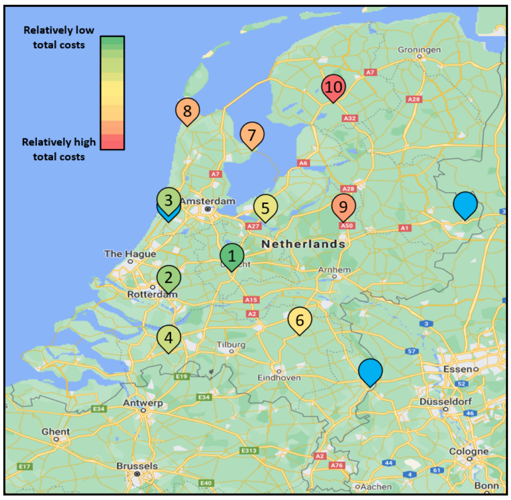

Figure 6.1: Geo-patterns appeal Hillegom

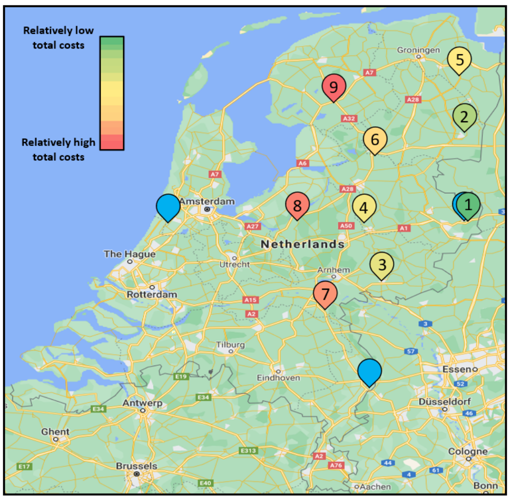

Figure 6.2: Geo-patterns appeal Oldenzaal

6.2 Relocating more than one hub

In this section we examine whether significant cost savings can be gained by relocating more than one hub instead of only one, and by relocating hubs simultaneously instead of consecutively. When relocating hubs simultaneously, all potential hub configurations are evaluated, and the optimal hub configuration is guaranteed. When relocating hubs consecutively, not all potential hub configurations are evaluated, and the optimal solution is not guaranteed. If relocating hubs simultaneously instead of consecutively is beneficial, there is reason to delay single hub relocations until there is e.g. more budget to relocate multiple hubs. If not, there can still be a benefit of relocating more than one hub instead of only one, but there is no reason to delay the single hub relocation. When relocating more than one hub simultaneously, we apply the p-Median SCMHLP (see Chapter 3 for an explanation why). The p-Median SCMHLP is computationally difficult to solve, and we can only solve relatively small instances under the assumption of deterministic demand. In the results of relocating only one hub that we discussed earlier, we consider stochastic demand and relatively large instances. Therefore, we should not benchmark these results against the results of relocating more than one hub. To properly benchmark against relocating only one hub and relocating multiple hubs consecutively, we apply the MNFP approach and input the same data as in the p-Median SCMHLP. For a single hub relocation using all deterministic instead of stochastic demand, the relocations that result in the lowest annual total costs are as follows. Relocate Venlo to Maastricht. Relocate Hillegom to Roosendaal. Relocate Oldenzaal to Oldenzaal. The differences in appeal between using stochastic and deterministic demand are caused by how hub configurations account for sub-optimal routings that are a result of hub capacities.

Given the number of hubs in the network of Rhenus Road NL, we can only analyze relocating two or three hubs. There is a trade-off between the number of o-d pairs and the number of candidate hub locations. We aim to find a good balance between the two, such that we can use it in computing managerial insights. We employ the Python programming language to implement the p-Median SCMHLP and use Gurobi to solve it. We carry out the experiments on an Intel(R) Core(TM) i7-8750H @ 2.20 GHz and 16.0 GB RAM. Appendix D provides insights on the computation time of several instances. Based on these experiments, we consider the following instances in computing managerial insights. When relocating two hubs, we consider 5,000 o-d pairs, and 15 candidate locations. When relocating three hubs, we consider 500 o-d pairs and 15 candidate locations.

6.2.1 Relocating two hubs

When relocating two hubs, we choose 15 candidate locations from the ones used in earlier analyses. The candidate locations considered depend on the hubs that we relocate, while ensuring a sufficient geographical spread. We randomly select 5,000 o-d pairs and scale the transportation cost per distance parameters to obtain realistic transportation costs.

Venlo and Hillegom

When relocating the hubs in Venlo and Hillegom simultaneously, the lowest annual total costs are € 38,107,489. The corresponding solution is to relocate the hub in Venlo to the candidate site in Venlo and the hub in Hillegom to the candidate site in Roosendaal. The solution is the same when relocating these hubs consecutively: first relocate the hub in Hillegom and thereafter the hub in Venlo. I.e., there is no benefit of relocating these hubs simultaneously instead of consecutively. Figure 6.3 depicts the solution. The markers in the figures represent the current locations (in blue), the candidate locations to which no hub is relocated (in white), and the candidate locations to which a hub is relocated (in green). Using the same input data, the lowest annual total costs are € 37,684,822 by relocating only one out of two hubs (Hillegom to Roosendaal). There is no benefit of moving both the hubs in Venlo and Hillegom instead of only the hub in Hillegom. Note that the costs for relocating one hub can be lower because we assume that Rhenus Road NL wants to relocate the hubs under consideration anyways. I.e., the total costs for relocating two hubs include fixed costs for relocating two hubs instead of one.

Venlo and Oldenzaal

When relocating the hubs in Venlo and Oldenzaal simultaneously, the lowest annual total costs are € 38,436,897. The corresponding solution is to relocate the hub in Venlo to the candidate site in Maastricht and the hub in Oldenzaal to the candidate site in Oldenzaal. The solution is the same when relocating these hubs consecutively: first relocate the hub in Oldenzaal and thereafter the hub in Venlo. I.e., there is no benefit of relocating these hubs simultaneously instead of consecutively. Figure 6.4 depicts the solution. Using the same input data, the lowest annual total costs are € 38,131,835 by relocating only one out of two hubs (Oldenzaal to Oldenzaal). There is no benefit of moving both the hubs in Venlo and Oldenzaal instead of only the hub in Oldenzaal.

Hillegom and Oldenzaal

When relocating the hubs in Hillegom and Oldenzaal simultaneously, the lowest annual total costs are € 37,901,489. The corresponding solution is to relocate the hub in Hillegom to the candidate site in Roosendaal and the hub in Oldenzaal to the candidate site in Oldenzaal. The solution is the same when relocating these hubs consecutively: first relocate the hub in Hillegom and thereafter the hub in Oldenzaal. I.e., there is no benefit of relocating these hubs simultaneously instead of consecutively. Figure 6.5 depicts the solution. Using the same input data, the lowest annual total costs are € 37,684,822 by relocating only one out of two hubs (Hillegom to Roosendaal). There is no benefit of moving both the hubs in Hillegom and Oldenzaal instead of only the hub in Hillegom.

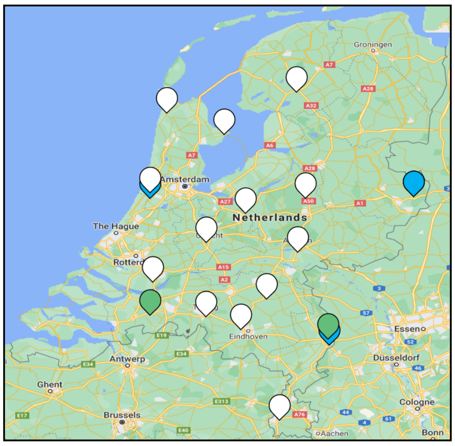

Figure 6.3: Relocate Venlo & Hillegom

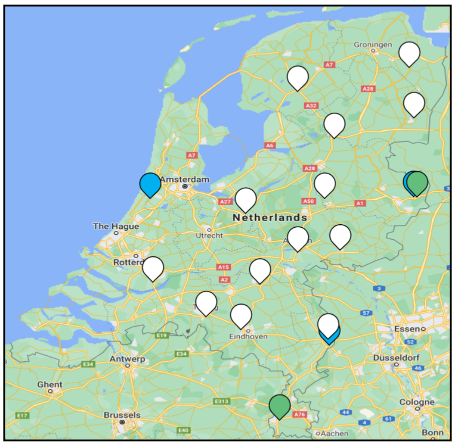

Figure 6.4: Relocate Venlo & Oldenzaal

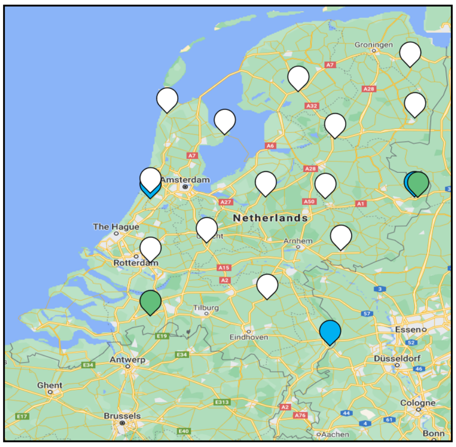

Figure 6.5: Relocate Hillegom & Olden.

6.2.2 Relocating three hubs

When relocating three hubs, we choose 15 candidate locations from the ones used in earlier analyses, while ensuring a sufficient geographical spread. We randomly select 500 o-d pairs and scale the transportation cost per distance parameters to obtain realistic transportation costs. The lowest annual total costs are € 38,302,715 for relocating three hubs simultaneously. The corresponding solution is to relocate the hub in Hillegom to the candidate site in Roosendaal, the hub in Oldenzaal to the candidate site in Oldenzaal, and the hub in Venlo to the candidate site in Venlo. The solution is the same when relocating the hubs consecutively; first relocate the hub in Hillegom, then the hub in Oldenzaal and thereafter the hub in Venlo. Figure 6.6 depicts the solution. Using the same input data, the lowest annual total costs are € 37,667,378 by relocating only one out of three hubs (Hillegom to Roosendaal), and € 37,884,045 by relocating only two out of three hubs (Hillegom to Roosendaal and Oldenzaal to Oldenzaal). There is no benefit of moving all three hubs instead of only the hub(s) in Hillegom (and Oldenzaal).

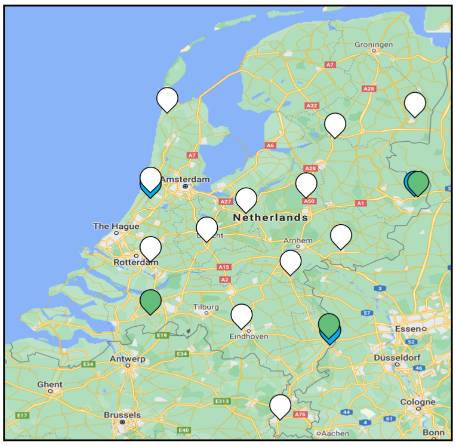

Figure 6.6: Relocate all three hubs

6.3 Concluding remarks artificial scenarios

Based on the findings of relocating only the hub in Venlo or Oldenzaal, we suspected that the data may be shaped towards the current hub locations. The results of the relocation of the hub in Hillegom however contradict this suspicion: by relocating the hub in Hillegom to the artificial candidate site in Nieuwegein, Rhenus Road NL can gain noticeable transportation costs savings. In addition, we showed that no cost savings can be gained by relocating hubs simultaneously instead of consecutively, nor by relocating more than one hub instead of only one. Therefore, we see no reason to delay single hub relocations. We do however need to be careful drawing conclusions based on the artificial scenarios due to several reasons. Firstly, to be consistent with the problem of Rhenus Road NL, we assume that the hubs are relocated anyways, to a buyable site. Rhenus Road NL may want to consider hiring (and benchmarking against the current situation) for the other hub relocations. Secondly, we disregard drivers other than transportation costs and fixed hub establishment costs that may also influence whether to relocate hubs or not (e.g. split in sites causing operational inefficiencies). Thirdly, some of the candidate locations are artificial. Fourthly, due to the computational difficulty of the p-Median SCMHLP, we can only solve instances with a relatively small amount of candidate locations and (randomly selected) o-d pairs. Finally, we assume deterministic demand in analyzing relocating more than one hub.

Chapter 7

Conclusion and discussion

Rhenus logistics is a large 3PL company. Amongst other solutions, they offer freight forwarding from initial senders to final consignees by road. Rhenus Road NL controls road freight forwarding services from, to, over, and within the Netherlands. The depots that Rhenus Road NL operates are from a theoretical point of view different than any other relevant road transport depot in the road freight forwarding network. Rhenus Road NL is looking to relocate the depot that they currently operate in Venlo, such that total costs comprised of transportation costs and fixed depot establishment costs are minimized. To Rhenus Road NL it is not about whether they should relocate the depot in Venlo, but where to. Therefore, we develop a methodology for optimally relocating a subset of road transport depots for companies that have a logistics structure like Rhenus Road NL. In optimizing the network, we consider that each depot that Rhenus Road NL operates only has connections with a subset of relevant road transport depots that Rhenus Road NL does not operate, and that some of these connections have a reachability limit. The goal of this research is to answer the following main research question:

How can the network design of a company with a logistics structure like Rhenus Road NL be optimized by relocating a subset of their depots while respecting restricted node allocations and reachability limits for roundtrips with relevant road transport depots?

For the general case of relocating a subset of depots of Rhenus Road NL, we introduce a model formulation for the p-Median SCMHLP. We refer to the depots that Rhenus Road NL operates as hubs. A special case of the p-Median SCMHLP occurs when we relocate only one hub. For this special case we develop a MNFP approach with stochastic demand. The MNFP approach does a decent job realistically assigning o-d pairs to hubs. In the reference results, the lowest total cost result by relocating the hub in Venlo to the candidate location in Venlo. The (percentage) differences in total costs are relatively low for candidate locations relatively close to the candidate location in Venlo. The appeal of candidate locations is similar over the working days, and when randomly assigning training and test data. A sensitivity analysis shows that with our self-created stochastic demand the candidate locations in Venlo and Maastricht are more appealing, and that the results are insensitive to changes in the transportation cost per distance parameters and hub capacities. In the data analysis we find that demand increases insignificantly over time, and a scenario analysis shows that increasing demand with constant geographical spread makes the candidate location in Venlo more appealing. Lower demand in the UK as a result of Brexit makes candidate sites located more (south)-east more appealing. Given that we make assumptions, and that there is inevitably going to be inaccuracy in the data, the (percentage) differences in total costs imply that we cannot decide on the optimal location solely based on numerical results. Qualitative characteristics and requirements that the company considers may be decisive. The company considers four qualitative properties that affect the desirability of their logistics location: proximity to highway access points, possibilities for future growth, retaining personnel, and road congestion. The candidate locations are very similar in terms of the first two properties. Regarding retaining personnel, the closer to the current hub in Venlo, the better. With respect to road congestion, TomTom (2020) measures road congestion in terms of congestion levels. Congestion levels are relatively high in Amersfoort, Arnhem, Eindhoven and Tilburg. We conclude that the candidate site in Venlo is the best option, at least when relocating only the hub in Venlo.

We also conduct numerical experiments on artificial scenarios of relocating a single other hub or more than one hub to obtain managerial insights and test the methodology's performance. By relocating the hub in Hillegom to the artificial candidate site in Nieuwegein, the company can gain noticeable transportation costs savings. We also show that the company cannot gain cost savings by relocating hubs simultaneously instead of consecutively, nor by relocating more than one hub instead of only one. Therefore, we see no reason to delay single hub relocations. We do however need to be careful drawing conclusions based on artificial scenarios.

7.1 Limitations

The quality of the findings depend on the quality and availability of data, as well as on the quality of modelling. Due to the variety in products that Rhenus Road NL ships, order size is stored in multiple ways. It depends on the properties of an order's commodities which demand indicators are filled out. In collaboration with company experts we decide that it is the best option to use the volume of shipments as flow indicator in modelling. The volume is filled out for most orders, and we extrapolate missing values. This negatively influences accurately imposing on hub capacities and scaling the transportation costs. Furthermore, in reality, orders of multiple o-d pairs are combined in forwarding freight. In this project we however abstract from the former because required data (e.g. time windows) is not documented and itineraries are planned manually. Therefore, transportation costs are approximated inaccurately. Moreover, since we only acquired order data between 2018 and 2020, we cannot investigate long-term trends nor e.g. the actual influence of Brexit. In addition, we assume that euclidean distances are sufficient.

Because we do not have access to the data of the depots that Rhenus Road NL does not operate, we assume that an order will be forwarded by the same relevant road transport depots that Rhenus Road NL does not operate as before the relocation. Although this is valid to assume according to company experts, it is not completely realistic. We considered including non-zero daily link capacities between road freight forwarding depots (determined based on historical data) in the models, but we chose not to implement this because this is too restrictive. We also considered completely unrestricted allocation decisions. However, a realistic road freight forwarding network requires structured roundtrip links and schedules. Redesigning the network in terms of roundtrip links and schedules requires data that is currently unavailable (e.g. interdependencies in roundtrip links and schedules across depots, types of goods a depot can handle, and fixed costs for opening roundtrip links). A component of inaccuracy of the modelling approach that we apply in this thesis originates from the fact that the company does not store raw data on the real connected transfer locations, nor on their coordinates. We decided in collaboration with company experts on how to compute the real connected transfer locations, and we employed geocoding software to obtain their coordinates. We expect there to be slight inaccuracies in the preparation of this data.

We could have improved the robustness of the findings for relocating more than one hub by applying a MNFP approach with stochastic demand that is similar to the one described in this thesis, instead of the p-Median SCMHLP. However, we do not have enough time to compute results with such an MNFP approach: we need to assure a sufficient geographical spread in candidate locations, and given the computational power of our machine, the computation time is too high. It is still an option to input fewer days of flow data, but this deteriorates the reliability of the results.

Despite the limitations of this research, we are confident that the results are reliable. We recommend Rhenus Road NL to relocate the hub that is currently located in Venlo to the candidate site in Venlo.

7.2 Future research

Location Problems generally deal with pure phase-in problems. We however develop a methodology for a Location Problem that does not build the entire network from scratch but instead relocates a subset of facilities. Therefore, this research can serve as a source of inspiration for future research on any Location Problem that considers relocating a subset of facilities. The p-Median SCMHLP is a computationally difficult problem, which inspires to develop a heuristic approach in future research. Future research may also focus on model performance of the p-Median SCMHLP with more to be relocated hubs and static hubs. Suggestions for extensions of the p-Median SCMHLP and MNFP approach are: modular hub capacities, routing to combine orders of multiple unique o-d pairs, redesign the network in terms of roundtrips links and schedules (with an endogenous number of hubs to place), multi-period decision-making, and stochasticity in parameters (other than demand).

Rhenus Road NL does not aim to expand, replace, or add other depots to their freight forwarding network in the foreseeable future. In the distant future, Rhenus Road NL could however perform other strategic level quantitative research to explore these options. It may be of particular interest to further investigate relocating the depot in Hillegom. To improve the robustness of the findings of relocating more than one depot, Rhenus Road NL could apply a similar MNFP approach with stochastic demand as the one describe in this thesis, and carry out computational experiments on a more powerful machine. In addition, Rhenus Road NL could investigate adding a central hub to their road freight forwarding network. By adding a central hub and thereby managing processes more centralized, the company can consolidate better and transfer commodities more efficiently. On the other hand, horizontal collaboration reduces complexity, lead time uncertainty and variability, as well as supply chain costs (Ghaderi et al., 2016). Diminishing the number of stages in the transport chain also leads to higher customer service levels. Rhenus Road NL could also perform tactical and operational level quantitative research. E.g., redesign the network in terms of roundtrip links and schedules, and include routing to combine orders of multiple unique o-d pairs. If this were to be analyzed by Rhenus Road NL, it will require a considerable amount of structured data that is currently unavailable, as well as tactical and operational level domain knowledge. Finally, although Rhenus logistics offers solutions along the entire supply chain, Rhenus Road NL does not explicitly offer multimodal solutions. Rhenus Logistics could explore the benefits of synchronizing and optimizing a multimodal distribution network design.

References

Abbasi, S., Saboury, A., and Jabalameli, M. S. (2021). Reliable supply chain network design for 3PL providers using consolidation hubs under disruption risks considering product perishability: An application to a pharmaceutical distribution network. *Computers & Industrial Engineering*, 152:107019.

Alumur, S. and Kara, B. Y. (2008). Network hub location problems: The state of the art. *European Journal of Operational Research*, 190(1):1–21.

Alumur, S. and Kara, B. Y. (2009). A hub covering network design problem for cargo applications in Turkey. *Journal of the Operational Research Society*, 60(10):1349–1359.

Alumur, S. A., Nickel, S., Saldanha-da Gama, F., and Seçerdin, Y. (2016). Multi-period hub network design problems with modular capacities. *Annals of Operations Research*, 246(1):289–312.

Boland, N., Krishnamoorthy, M., Ernst, A. T., and Ebery, J. (2004). Preprocessing and cutting for multiple allocation hub location problems. *European Journal of Operational Research*, 155(3):638–653.

Boukani, F. H., Moghaddam, B. F., and Pishvaee, M. S. (2016). Robust optimization approach to capacitated single and multiple allocation hub location problems. *Computational and Applied Mathematics*, 35(1):45–60.

Campbell, J. F. (1994). Integer programming formulations of discrete hub location problems. *European Journal of Operational Research*, 72(2):387–405.

Campbell, J. F. (2005). Strategic network design for motor carriers. *Logistics systems: Design and optimization*, pages 245–278.

Campbell, J. F. and O'Kelly, M. E. (2012). Twenty-five years of hub location research. *Transportation Science*, 46(2):153–169.

Cheong, M. L., Bhatnagar, R., and Graves, S. C. (2007). Logistics network design with supplier consolidation hubs and multiple shipment options. *Journal of Industrial and Management Optimization*, 3(1):51.

Contreras, I., Cordeau, J.-F., and Laporte, G. (2012). Exact solution of large-scale hub location problems with multiple capacity levels. *Transportation Science*, 46(4):439–459.

Correia, I., Nickel, S., and Saldanha-da Gama, F. (2018). A stochastic multi-period capacitated multiple allocation hub location problem: Formulation and inequalities. *Omega*, 74:122–134.

Crainic, T. G. (2003). Long-haul freight transportation. In *Handbook of Transportation Science*, pages 451–516. Springer.

Cunha, C. B. and Silva, M. R. (2007). A genetic algorithm for the problem of configuring a hub-and-spoke network for a ltl trucking company in brazil. *European Journal of Operational Research*, 179(3):747–758.

Demır, İ., Ergın, F. C., and Kıraz, B. (2019). A new model for the multi-objective multiple allocation hub network design and routing problem. *IEEE Access*, 7:90678–90689.

Ebery, J., Krishnamoorthy, M., Ernst, A., and Boland, N. (2000). The capacitated multiple allocation hub location problem: Formulations and algorithms. *European Journal of Operational Research*, 120(3):614–631.

European Commision (2021). The EU-UK Trade and Cooperation Agreement. Last accessed 5 July 2021.

Ghaderi, H., Dullaert, W., and Amstel, W. P. V. (2016). Reducing lead-times and lead-time variance in cooperative distribution networks. *International Journal of Shipping and Transport Logistics*, 8(1):51–65.

Ghaffari-Nasab, N., Ghazanfari, M., and Teimoury, E. (2016). Hub-and-spoke logistics network design for third party logistics service providers. *International Journal of Management Science and Engineering Management*, 11(1):49–61.

Hakimi, S. L. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research*, 12(3):450–459.

Hu, L., Zhu, J. X., Wang, Y., and Lee, L. H. (2018). Joint design of fleet size, hub locations, and hub capacities for third-party logistics networks with road congestion constraints. *Transportation Research Part E: Logistics and Transportation Review*, 118:568–588.

Huff, D. L. (1963). A probabilistic analysis of shopping center trade areas. *Land Economics*, 39(1):81–90.

Kara, B. Y. and Tansel, B. Ç. (2002). The latest arrival hub location problem for cargo delivery systems with stopovers. *Transportation Research Part B: Methodological*, 41(8):906–919.

König, A. and Spinler, S. (2016). The effect of logistics outsourcing on the supply chain vulnerability of shippers. *The International Journal of Logistics Management*, 27(1):122–141.

Marín, A. (2005). Formulating and solving splittable capacitated multiple allocation hub location problems. *Computers & Operations Research*, 32(12):3093–3109.

Monemi, R. N., Gelareh, S., Nagih, A., and Jones, D. (2021). Bi-objective load balancing multiple allocation hub location: A compromise programming approach. *Annals of Operations Research*, 296(1):363–406.

Motaghedi-Larijani, A., Jabalameli, M.-S., and Tavakkoli-Moghaddam, R. (2012). A network design model considering inventory cost from a third party logistics perspective. *International Journal of Management Science and Engineering Management*, 7(1):29–35.

O'Kelly, M. E. (1986a). Activity levels at hub facilities in interacting networks. *Geographical Analysis*, 18(4):343–356.

O'Kelly, M. E. (1986b). The location of interacting hub facilities. *Transportation Science*, 20(2):92–106.

O'Kelly, M. E. (1987). A quadratic integer program for the location of interacting hub facilities. *European Journal of Operational Research*, 32(3):393–404.

O'Kelly, M. E. and Miller, H. J. (1994). The hub network design problem: A review and synthesis. *Journal of Transport Geography*, 2(1):31–40.

Ouorou, A., Mahey, P., and Vial, J.-P. (2000). A survey of algorithms for convex multicommodity flow problems. *Management Science*, 46(1):126–147.

Perotti, S., Zorzini, M., Cagno, E., and Micheli, G. J. (2012). Green supply chain practices and company performance: the case of 3pls in italy. *International Journal of Physical Distribution & Logistics Management*.

Rodriguez-Martin, I. and Salazar-Gonzalez, J. J. (2008). Solving a capacitated hub location problem. *European Journal of Operational Research*, 184(2):468–479.

Sasaki, M. and Fukushima, M. (2003). On the hub-and-spoke model with arc capacity constraints. *Journal of the Operations Research Society of Japan*, 46(4):409–428.

Sender, J. and Clausen, U. (2013). Heuristics for solving a capacitated multiple allocation hub location problem with application in german wagonload traffic. *Electronic Notes in Discrete Mathematics*, 41:13–20.

Stojanović, Đ. (2017). Road freight transport outsourcing trend in Europe–what do we really know about it. *Transportation Research Procedia*, 25:772–793.

The Guardian (2021). Road freight between Britain and EU is down by a third, data shows. Last accessed 15 June 2021.

TomTom (2020). TomTom Traffic Index. Last accessed 18 June 2021.

Üster, H. and Agrahari, H. (2011). A Benders decomposition approach for a distribution network design problem with consolidation and capacity considerations. *Operations Research Letters*, 39(2):138–143.

Yaman, H., Karasan, O. E., and Kara, B. Y. (2012). Release time scheduling and hub location for next-day delivery. *Operations Research*, 60(4):906–917.

Appendix A

Data preparation

A.1 (Reachability limited) relevant road transport depots (per hub)

There is no raw data available on which road transport depots are relevant to Rhenus Road NL. Table A.1 presents the raw data required in order to compute the former.

Table A.1: Raw data for relevant road transport depots

Principal	**Agent**
...	...

Each row represents an order handled by a specific hub and contains information on the Principal and Agent. When from the point of view of Rhenus Road NL the freight was transferred from another relevant road transport depot, the Principal cell is filled out with a label corresponding to the relevant road transport depot. In contradiction, when from the point of view of Rhenus Road NL the freight was transferred to another relevant road transport depot, the Agent cell is filled out with a label corresponding to the relevant road transport depot. Due to operational reasons that are not relevant to this project, the Principal and Agent columns are also utilized to store different information. As a result, extracting which road transport depots are relevant to Rhenus Road NL based on these columns is not completely accurate, but according to experts in the field it provides a good approximation. Algorithm 2 describes how we obtain the relevant road transport depots.

Algorithm 2: Computing relevant road transport depots

Input: 9 original files of raw data: one per hub per year
Output: Dataframe T with all relevant road transport depots

1 Concatenate 9 original files into 1 dataframe;
2 Compute the frequency of each unique label of concatenated Principal and Agent columns;
3 Check manually for each unique label in PA that has a frequency of at least 50 whether the label corresponds to a road freight forwarder that is not a hub. If so, put the label in T[RD];
4 Let expert in the field confirm information in T and adapt according to feedback;

Similarly, we obtain sets for the relevant road transport depots per hub. The differences are that we concatenate the files per hub, and use that the frequency of a unique label must be more than 20 while excluding depots that are not in T. For T_r we first compute the distance between the current hub in Venlo and every location in T and S. Then, we remove the locations that are not within a radius of $500km$ of the current hub in Venlo. The remaining depots should be within a radius of $600km$ (B) after the relocation.

A.2 Real connected transfer locations and coordinates

There is no raw data available on the real connected transfer locations. Besides, the raw data does not include coordinates for locations. Table A.2 presents the raw data required in terms of columns to compute the former. Each row contains information on the initial sender and final consignee of an order. The available information of the initial sender is as follows: Sender Name (SN), Address (SA), Post Code (SPC), City (SCI), and Country (SCO). Similarly, the information available on the final consignee is: Consignee Name (CN), Address (CA), Post Code (CPC), City (CCI), and Country (CCO). Furthermore, each row contains information on the Principal (Pr), Agent (Ag), True Type (TT) and Type (Ty). TT is either of category Export, Import, Domestic, or Cross-border, whereas Ty is either of category Export, or Import. The True Type is self-explanatory. The Type provides relevant information in case of Cross-border transfers; it facilitates determining in which direction a hub is connected to a relevant road transport depot that they do not operate.

Table A.2: Raw data for real connected transfer locations and coordinates

SN	**SA**	**SPC**	**SCI**	**SCO**	**CN**	**CA**	**CPC**	**CCI**	**CCO**	**TT**	**Ty**	**Pr**	**Ag**
...	...	...	...	...	...	...	...	...	...	...	...	...	...

The initial sender and final consignee always correspond to the real connected transfer locations, except when the freight was forwarded by another relevant road transport depot. Every time the freight corresponding to a specific order is handled by another road transport depot, a new row of data is generated in the data of the depot that is handling the order. Both the initial sender and final consignee information remains the same, but the Principal and Agent change. In theory, the Principal and Agent columns facilitate computing the real connected transfer locations very easily. However, the opposite is true. In many cases where the Principal is a relevant road transport depot, the same relevant road transport depot is filled out as Agent. Again, this is due to operational reasons that are not relevant to this project. With the high frequency the former occurs, it is also not possible to correct for these difficulties manually. As a result, we need to come up with an alternative to compute the real connected transfer locations and their coordinates: Algorithm 3. Note that in case two hubs handle the same order, this order is twice in the overall data sets, and we need to account for this.

Algorithm 3: Computing real connected transfer locations and coordinates of every order

Input: 9 original files of raw data and dataframe T

Output: Dataframe df with location labels and coordinates of the real connected transfer locations of every order

1 Concatenate 9 original files into 3 dataframes: one per location of Rhenus Road NL;
2 Remove all shipments in the Oldenzaal dataframe that were forwarded from and to Venlo and Hillegom according to the Principal and Agent columns;
3 Remove all shipments in the Venlo dataframe that were forwarded from and to Hillegom according to the Principal and Agent columns;
4 Concatenate 3 dataframes without double shipments into 1 dataframe df;
5 ...;

```
for i in range(len(df)) do
    Aggregate sender information in df[Se][i] (use , to separate column information);
    Aggregate consignee information in df[Co][i] (use , to separate column information);
    if df[TT][i] == "Export" then
        for j in range(len(T)) do
            if df[Ag][i] == T[RD][j] then
                df[Co][i] = df[Ag][i];
            else if df[Ag][i] ≠ df[Pr][j] then
                if df[Pr][i] == T[RD][j] then
                    df[Se][i] = df[Pr][i];
    if df[TT][i] == "Import" then
        for j in range(len(T)) do
            if df[Pr][i] == T[RD][j] then
                df[Se][i] = df[Pr][i];
            else if df[Ag][i] ≠ df[Pr][j] then
                if df[Ag][i] == T[RD][j] then
                    df[Co][i] = df[Ag][i];
    if df[TT][i] == "Domestic" then
        if df[Ag][i] ≠ df[Pr][i] then
            for j in range(len(T)) do
                if df[Pr][i] == T[RD][j] then
                    if df[CA][i] ≠ Talhoutweg 2 (exception) then
                        df[Se][i] = df[Pr][i];
                else if df[Ag][i] == T[RD][j] then
                    df[Co][i] = df[Ag][i];
    if df[TT][i] == "Cross" then
        if df[Ag][i] ≠ df[Pr][i] then
            for j in range(len(T)) do
                if df[Ag][i] == T[RD][j] then
                    df[Co][i] = df[Ag][i];
                else if df[Pr][i] == T[RD][j] then
                    df[Se][i] = df[Pr][i];
        else if df[Ag][i] == df[Pr][i] then
            if df[Ty] == "Export" then
                df[Co][i] = df[Ag][i];
            else if df[Ty] == "Import" then
                df[Se][i] = df[Pr][i];
Concatenate df[Se] and df[Co] and compute its set to get all unique real connected
    transfer locations;
Geocode the real connected transfer locations, assess geocoding quality and account for
    deficiencies (geocoding software is very accurate and indicates confidence levels of having
    computed the coordinates correctly);
Manually compute coordinates of labels in T[RD] and let expert in the field confirm
    retrieved information (labels in T[RD] have no geographical meaning);
Match coordinates to original orders' real connected transfer locations.
```

A.3 Flow related data

Table A.3 presents the demand metrics that Rhenus Road NL uses to express order size in the data. Rhenus Road NL handles many different types of commodities. It depends on the properties of an order's commodities which of the fields of the first three columns in Table A.3 are filled out. The loading meter column is filled out for 20.7% of the orders, the volume column for 85.9% of the orders, and the weight column for 99.9% of the orders. The payweight, also referred to as the chargeable weight, is a mandatory field to be filled out and for the company the reference for charging customers. The payweight is computed by taking the maximum over the values in the first three columns of the order, after having them multiplied by a specific fixed factor that is unique for the column. Given the variety in commodities shipped, according to company experts, the payweight and weight are inappropriate to scale transportation costs with, nor to impose hub capacities on. Instead, we should use the volume of shipments as the flow indicator. In case the volume is filled out, the correlation between the volume and payweight data is 0.877. As such, we extrapolate the missing values in the volume column using the payweight data. Because the volume data may not be filled out for a reason (e.g. inappropriate for the type of commodities), we check for the mean and standard deviation in the volume before and after the extrapolation. These are relatively similar.

Table A.3: Raw demand data format

Loading meters (m)	**Volume (m^3)**	**Weight (kg)**	**Payweight**
...	-	-	...

So, we use the volume data to create the flow related data (L, w_k^l, C^l). For the reference data, the scenario sets contain the prepared historical data (according to the preparation described in this Appendix, but also in the data insights section of Chapter 4). Initially, we compute for each working day a distinct scenario set. E.g. in the scenario set of Monday each scenario contains all flow of one unique Monday. Thereafter, we cluster scenario sets if the difference in the volume handled between two working days is statistically insignificant. Finally, we split up each scenario set in a training (years 2018 and 2019) and test (year 2020) data set.

A.4 Remaining hub related data

Table A.4 shows the data format of a potential hub location (P). We are provided a set of candidate locations, and for each candidate location the fixed costs for establishing a hub. We compute the annual fixed costs by assuming that the company replaces a specific hub every thirty years. Table A.5 presents the data format of the current hubs (hubs in R and S).

Table A.4: Candidate hub locations

Label	**Latitude**	**Longitude**	**Fixed costs**
...	...	...	...

Table A.5: Static hub locations

Label	**Latitude**	**Longitude**
...	...	...

With respect to Q_h, for the reference data, the new hub should have the same capacity as the hub that is currently located in Venlo. The hub capacities of the static hubs remain the same. According to company experts, the current hub in Venlo was at its maximum capacity during the last three months of 2020. Therefore, we let the daily hub capacity per scenario set of the hub in Venlo be equal to the average daily volume handled in the corresponding scenario set during this period of time. We then use the m^2 ratios of the current hubs to calculate the daily volume that can be handled by the two static hubs per scenario set. We let the hub capacities differ per scenario set in order to take into account storage requirements over the days.

A.5 Transportation costs related parameters

Assuming euclidean distances, we prepare the transportation cost per distance parameters (f and g): if we input the current hub configuration, the computed transportation cost should be equivalent to the actual transportation costs, with a similar percentage of orders using a hub-hub. This allows us to adequately compare transportation costs with fixed costs, and differences in transportation costs between hub configurations. The transportation costs comprise costs for personnel, fuel, depreciation, maintenance, and overhead. The fuel cost component causes the difference in f and g. Table A.6 presents the resulting transportation costs per distance (€$/km$) for the two types of trips. These are relatively low due to the fact that we abstract from routing to combine orders of multiple unique o-d pairs.

Table A.6: Transportation cost per distance vehicle types

Parameter	**Value**
$\boldsymbol{f}$	0.0332
$\boldsymbol{g}$	0.0265

Appendix B

Self-creating stochastic demand

In self-creating stochastic demand, we distinguish between three types of frequencies.

Type 1: o-d pairs with a relatively high frequency of flows on a working day.
To estimate o-d pair volume, we need the distribution and parameters per o-d pair per working day. Therefore, we need at least a few observations: we use a cut-off value of 5. I.e. in this type 1 section we explain how we deal with o-d pairs that have a flow frequency of at least 5 on a working day. We analyze the patterns of o-d pair 1, which represents most type 1 o-d pair working days relatively well. Figure B.1 depicts the number of orders handled per working day for o-d pair 1. On Thursday, the o-d pair frequency is only 3. Thus, we fit distributions for the o-d pair volume per working day for all working days, except for Thursday. For simplicity, we only fit normal and Poisson distributions. For o-d pair 1, for all working days, the estimation error is lowest with the Poisson distribution. So, for this type we estimate the volume of each o-d pair on each working day using the Poisson distribution. The mean of the distribution depends on the average volume of an o-d pair on the working day.

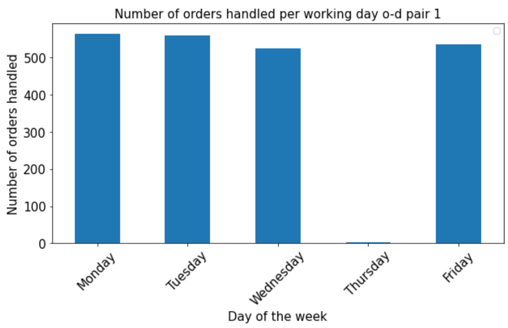

Figure B.1: # working day flows o-d pair 1

Type 2: o-d pairs with a relatively low frequency of flows on a working day, of which the origin has a relatively high frequency of flows on a working day.
This type is inspired by the fact that there are significantly fewer unique origins than destinations. To estimate origin volume, we need the distribution and parameters per origin per working day. Therefore, we need at least a few observations: again, we use a cut-off value of 5. I.e. in this type 2 section, we explain how we deal with origins that have a flow frequency of at least 5 on a working day, excluding type 1 o-d pair working days. We analyze the patterns of origin 2, which represents most type 2 origin working days relatively well. Figure B.2 depicts the number of orders handled per working day for origin 2. On Wednesday, the origin frequency is only 4. Thus, we fit distributions for the origin volume per working day for all working days, except for Wednesday.

For simplicity, we only fit normal and Poisson distributions. For origin 2, for all working days, the estimation error is lowest with the Poisson distribution. So, for this type we first estimate the volume of each origin on each working day using the Poisson distribution. The mean of the distribution depends on the average volume of an origin on the working day. We then distribute the volume to destinations using the probability of the destination given the origin. Finally, we exclude type 1 o-d pair working days.

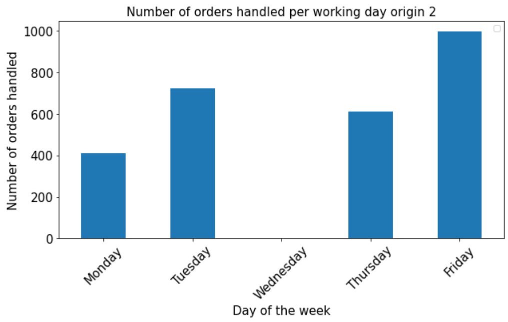

Figure B.2: # working day flows origin 2

Type 3: remaining o-d pairs.
We now have covered estimating demand for 91.7% of the orders. The remaining 8.3% of the orders contain o-d pair working days for which we cannot properly estimate demand. Since this is a relatively low percentage, we copy their volumes from the reference data, and randomly assign them to a scenario.

Appendix C

Computation time MNFP approach with stochastic demand

Table C.1 presents insights on the runtime of the MNFP approach per hub iteration, for the reference data without splitting in training and test data. We distinguish between runtime to solve the model and computation time to create the sets, parameters and variables. We elaborate on the latter because the computation time is not so much in solving the MNFP model itself iteratively, but due to the relatively high number of unique o-d pairs mostly in creating the input data that is unique per candidate hub iteration. The computation time for creating the input data is relatively high for the first hub iteration. For the remaining hub iterations we can more efficiently create the input data that is unique to the candidate hub iteration by using the data of the first hub iteration.

Table C.1: Computation time reference data

Hub iteration	**Model**	**Preparation**
1	00:01:20	03:41:27
2	00:01:18	00:09:50
3	00:01:16	00:09:48
...	...	...

Appendix D

Computation time p-Median SCMHLP

For the relocation of two hubs, we consider increments of 5,000 in the number of randomly selected o-d pairs. We increase the number of candidate locations in steps of 5. Given that the new hubs must be located in the Netherlands, to assure a sufficient geographical spread, we consider a minimum of 10 candidate locations. Table D.1 presents the computation time for relocating two hubs while considering 10 and 15 candidate locations. For instances with 20 candidate locations or more, we already run out of memory at an amount of 5,000 o-d pairs.

Table D.1: Computation time (hh:mm:ss) $|R| = 2$

$\|C\|$	$\|P\| = 10$	$\|P\| = 15$
5,000	01:43:04	05:31:27
10,000	22:18:37	Out of memory
15,000	Out of memory	Out of memory

The instances that we can solve are relatively small. The number of o-d pairs that we can consider is higher when considering 10 candidate locations. On the other hand, 15 candidate locations allows for a better geographical spread in candidate locations. Because we analyze the effect of relocating hubs simultaneously instead of consecutively, with fewer candidate locations, there is a larger probability that the optimal locations are the same for both cases. Therefore, we compute managerial insights with 15 candidate locations (and 5,000 o-d pairs) instead of 10.

For the relocation of three hubs, we consider increments of 500 in the number of randomly selected o-d pairs. We increase the number of candidate locations in steps of 5. Table D.2 presents the computation time for relocating three hubs while considering 10 and 15 candidate locations. For instances with 20 candidate locations or more, we already run out of memory at an amount of 500 o-d pairs.

Table D.2: Computation time (hh:mm:ss) $|R| = 3$

$\|C\|$	$\|P\| = 10$	$\|P\| = 15$
500	01:35:57	09:12:43
1,000	21:53:19	Out of memory
1,500	Out of memory	Out of memory

Due to the same reasons as with relocating two hubs, we rather compute managerial insights with 15 candidate locations (and 500 o-d pairs) instead of 10.