Master of Science – Logistics and Transport Management

Location Planning Considering Delivery Time and Service Level Constraints

A Heuristic Solution Approach to a Linear Optimization Problem from an Automotive Spare Parts Network

Author
Alexander Zienau

Supervisors
Jonas Flodén
University of Gothenburg

Felix Zesch
4flow AG

May 27, 2018

UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

4flow research.

Abstract

This thesis addresses the challenge to model a warehouse location problem under consideration of delivery time and service level constraints. For identification of appropriate modelling techniques, literature research is conducted to identify relevant models using similar approaches. A European spare parts network is chosen for model application, which is why qualitative research combined with an expert interview supports model development from a content perspective. The developed model requires customer classification into different delivery time categories, for which a desired service level is set as input data. A sensitivity analysis of the model shows the impact of an increasing service level on the objective function value, and is thus allowing to quantify costs of measures towards a more decentralized network structure. In order to guarantee applicability for a large amount of data, a heuristic combining Lagrangian relaxation with a knapsack problem approach has been developed and solved within 8 hours of computation time.

Key words: network design, location problem, delivery time, outliers, linear optimization, heuristic, Lagrangian relaxation, knapsack problem, spare parts network, automotive industry

Acknowledgement

I would like to use this opportunity to express my deep gratitude to all the people who supported me along the journey of writing this thesis. Firstly, I would like to thank my supervisor Jonas Flodén from University of Gothenburg for his guidance and his valuable feedback throughout the whole thesis process despite the geographical distance. Moreover, I would like to thank my second supervisor Felix Zesch and all the other colleagues involved at 4flow in Berlin for their daily practical support and for providing the framework for such an intensive and interesting thesis project. I would also like to acknowledge my research colleagues Frieder Smolny and Karl Däubel from TU Berlin who were always there to give advice on the mathematical aspects of this thesis. Last but not least, I would like to thank my friends and family for their mental support and continuous encouragement.

Table of Content

List of Abbreviations

chap.	chapter
3PL	Third-Party Logistics Provider
FTL	Full Truckload
LTL	Less Than Truckload

List of Figures

List of Tables

1 Introduction

In this chapter, the meaning of location problems in supply chain networks is introduced together with the main tradeoffs faced in network planning. Further, the chosen topic, methodological direction and reference example are motivated, before the chapter ends with developing the research question and giving a logical sequence of this research.

1.1 Background

In general, supply chain network design decides on the number of manufacturing plants and warehouses to be opened, their location and their capacity and face the challenge to provide the optimal framework from a strategic, tactical and operational perspective on supply chain activities. due to large costs involved in opening facilities, network design decisions are mostly long-term decisions, but at the same time have a deep impact on tactical and operational supply chain planning decisions like the choice of transportation modes, capacities and vehicle routing (Melo *et al.*, 2009; Simchi-Levi *et al.*, 2011; Crainic, 2000). Since many networks are becoming more and more time-sensitive due to decreasing lead times, distances between network locations play an important role. When it comes to network design, companies are faced with a tradeoff: They either choose decentralized warehousing in order to fulfill delivery time constraints or they opt for centralization to minimize the network costs. However, decentralization impedes use of risk pooling, which would lead to a reduced stock level by creating the flexibility to supply several customers from a warehouse (Du and Evans, 2008; Simchi-Levi *et al.*, 2011).

Warehouse location planning is thus one part of network design and is in its main features a common and widely discussed problem with very different areas of interest (see for example Melo *et al.*, 2009). Among others, Simchi-Levi *et al.* (2011) point out the difficulty to provide a certain service level to remote customers as a core challenge, as maximum distances between warehouses and customers are usually set in order to deliver within a specific distance or time frame. Therefore, it can become very expensive to provide a service level that includes all customers. In the case of a retail store chain, the example provided by Charikar *et al.* (2001) states that a network reaching 88% of the US population would be economically satisfying. In order to include so-called outliers, the number of warehouses and corresponding costs would not justify the increased service level at all.

A service level in terms of delivery time can be differently motivated. The following examples present various reasons: it can be imposed by the market through competitors' supply chain performance (Papathanasiou and Manos, 2007), by products, which can be perishable (Pulido *et al.*, 2015) or by customers demanding quick delivery, like in spare part networks (Yaobao *et al.*, 2013) or humanitarian logistics (Balcik and Beamon, 2008). Consequently, a variety of delivery time categories can appear within a network.

For clarification of terminology, the following definitions hold throughout the thesis:

delivery time: the time it takes to transport a product from a warehouse to a customer

service level: the share of customers supplied in time in a network.

1.2 Problem Discussion

As discussed, it is important for a company to identify the most economic service level within the network planning stage, when long-term decisions on facility locations come up. For that reason, the ability to vary the service level is suggested in network design (Simchi-Levi *et al.*, 2011), so that the tradeoff between investment and service level can be evaluated on different scenarios. Consequently, the need to investigate the use of quantitative models arises in order to illustrate the monetary effect.

In general, location problems are extensively discussed in literature as mathematical optimization problems (Melo *et al.*, 2009). This originates from (supply chain) network characteristics. Those are often complex and it is not possible to include all the existing data. Models can however include simplified assumptions that focus on a specific question within the complex system (Flodén *et al.*, 2017). Further, such models provide the opportunity to quantify the effect of different scenarios in network design (Melo *et al.*, 2009). The attention among location problem models reaches from varieties in different network structures and characteristics to different solution techniques and find application in strategic supply chain network decisions. Thus, this thesis discusses a location problem within the framework of mathematical optimization, considering service level and delivery time constraints.

In supply chain networks, a large number of network members (suppliers, warehouses, customers) and the combination of strategic and tactical planning levels can cause the abovementioned complexity (Simchi-Levi *et al.*, 2011). This requires large amounts of data and computation resources. It is therefore crucial for a model to be scalable – to be able to handle larger input data sets. Moreover, as this thesis results from a collaboration with a research project that investigates how algorithms can optimize logistics networks, while considering large amounts of data, scalability is an important factor of this thesis' problem. Thus, the research's focus is firstly, to apply techniques, which make the optimization problem solvable for large amounts of data and secondly, to provide a scalable framework, which can be further extended by the use of more sophisticated techniques.

Spare parts meet the characteristics of the initial problem in being time-sensitive and therefore require short transport distances so that warehouse locations are crucial. Furthermore, there is a demand in research on strategic spare part logistics planning (Wagner *et al.*, 2012), to which warehouse location planning is a sub topic. Moreover, in practice, spare parts or aftermarket sales have become a profit opportunity, especially for automotive Original Equipment Manufacturers (OEMs) and transport providers. The most important reasons, among others, are opportunities to balance declining economies or product groups through a relatively constant demand or to reach greater profit margins compared to conventional transport services (Barkawi *et al.*, 2006; Li, 2015). On the downside, longer product life cycles and short product innovation cycles increase complexity to meet customers' service demands (Wagner *et al.*, 2012).

1.3 Research Question and Thesis Disposition

This thesis aims to develop a warehouse location planning model whith the main characteristics to (1) consider customers with different delivery time categories in order to achieve defined service levels and (2) the ability to vary the share of customers supplied in every delivery time category. As a result, the effect of different service levels on network design is to be demonstrated. Due to the access to data from that industry, an aftermarket spare parts network from the automotive industry is chosen as a reference example for the initial problem throughout the thesis. In research, the question for an economically reasonable tradeoff in terms of promised delivery time and number of customers supplied is brought up by Schittekat and Sörensen (2009) in their practical research on Toyota's European spare parts network, with comparably short delivery times throughout Europe.

Therefore, the following research question is to be answered throughout this thesis:

How can the impact of different service levels for different delivery time categories on warehouse location problems in automotive spare parts networks be modeled in an appropriate way?

To answer this question, the *methodology* chapter discusses ways to develop mathematical models as well as procedures for model validation and verification and ends in the research methods guideline.

Next, the chapter *theoretical background* shows solution approaches for location problems considering service level and delivery time variation and discusses their applicability for the present problem. In addition, automotive spare parts networks characteristics are presented in detail, so that case-specific model requirements are identified. The chapter ends with the identified research gap.

The chapter *model development* defines characteristics of the developed model and continues with the mathematical formulation of the warehouse location problem. Then, input data is presented, followed by model and data validation as well as model verification with a small instance and clear expected results.

Within *solution approach*, a heuristic solution is developed to make the initial problem solvable for large amounts of data.

Lastly, the chapter *analysis* discusses the model output with regards to model characteristics as well as practical implications for a spare parts network.

2 Methodology

This chapter presents drivers to be considered when developing a supply chain network model. In addition, need as well as techniques for model validation and verification is introduced and determined. Subsequently, this results in the methods framework for this research.

2.1 Research Paradigm

Following Collis and Hussey (2014), model development falls into positivism rather than interpretivism. While interpretivism refers to research among others based on subjective impressions of reality, positivism is based on focusing on ‘logical or mathematical proof’ for claims in the research. In addition, ‘casual relationships’ and ‘casual laws’ are assumed to investigate relationships between variables. Thus, positivism is related to quantitative measurement of data. At the same time, the authors question the applicability of a pure separation between a quantitative view on social phenomena and qualitative social contexts. Moreover, the researchers’ objectivity would be hard to ensure, since they would indispensably slip in subjective interests and assumptions. Similar critique is brought up by Flodén *et al.* (2017) who claim that a model would always be individualistic, and that it would be unlikely that two researchers would independently build the same model.

2.2 Model Development

The question whether or not warehouses should be opened in a network challenges a company to cope with a complex decision-making problem. In such problems, one or more case-specific criteria, on which the decision is based, are to be determined. Here, the most common criterion is profit maximization: locations are to be planned so that the (positive) difference between profit and costs is maximized. Thus, a mathematical model would allow to quantify different scenarios and be a logical choice to approach location problems, due to the previously mentioned network complexity (Maßmann, 2006). Although some aspects in strategic logistics are hard to quantify, such as political stability around a possible location or growth rate of a developing market, location problem models are popular tools in literature (Schmidt and Wilhelm, 2000). When actually developing a model, some crucial aspects need to be taken into account: According to Melo *et al.* (2009), model development would always be a tradeoff between scope, realism, complexity and solvability.

In terms of realism, it should first be stated that models are a reduction of reality focusing only on aspects which influence the goal of the decision (Maßmann, 2006; Flodén *et al.*, 2017). Moreover, only that part of reality can be modeled for which input data is available (Flodén *et al.*, 2017). Especially for strategic decisions, data is often not likely to be available or of poor accuracy (Simchi-Levi *et al.*, 2011). However, not every model aims to map reality at all costs, as the classifications of ReVelle *et al.* (2008) into four location planning model types (see Table 1) shows. Regarding that classification, analytic models simplify reality and aim at showing relationships between variables. Similar characteristics account for continuous models, in which straight-line distances are used and facilities can be located at any location (also called "greenfield" planning). Contrary to that, discrete location models would be most suitable for practical use, as they consider discrete sets of demands and candidate locations. This is extensively discussed in literature (ReVelle *et al.*, 2008). The use of discrete and continuous models is also discussed in Maßmann (2006), where a set of possible locations is determined qualitatively, followed by discrete modelling on a more detailed planning level. To sum up, models aim to be as realistic as possible within the simplified observed section of reality.

	Model Type	**Characteristics**
(1)	Analytic Models	• simplifying assumptions, like fixed costs for all facilities and distance-dependent linear transport costs • provide insight into relationships between variables, but lack of practical relevance
(2)	Continuous Models	• facilities can be located anywhere while demand is located at fixed locations • few real-life applications like video cameras
(3)	Network Models	• include network structure, where demand arises at links and nodes • possible application for emergency highway services
(4)	Discrete Location Models	• assume discrete demand and sets of candidate locations • practically applicable

Table 1: Methodological Model Classification in Location Planning (from ReVelle *et al.*, 2008)

Due to company collaboration, data availability and in order for the model to be as realistic as possible, a discrete location model is chosen in this thesis. In addition, an expert interview with a consultant experienced in automotive spare parts networks is conducted to obtain up-to-date practical information. A semi-structured interview type is considered to be appropriate, since it allows the interviewer to follow-up on model-relevant topics from spare parts networks and at the same time gives the interviewee room to elaborate on topics that appear most relevant. This in turn should give the interviewer the opportunity to get a broader understanding of the supply chain network (Collis and Hussey, 2014). The interview has been conducted face-to-face for about thirty minutes. The interview guide is presented in Appendix 9.1, and the results are used for model development in chapter 0. Moreover, knowledge about spare parts networks is gained from literature research. Both knowledge sources are to support model development as well as analysis with respect to spare parts networks characteristics.

In most cases, models are applied for complex systems, such as production systems or production facilities. Although simplifications might lead to removal from reality, a problem as such can still be complex. Depending on the problem type, this could be the case for large numbers of possibilities to choose from, as an increase in solution elements (like warehouses) in turn makes computational effort increase exponentially (Klein and Scholl, 2012). For that reason, simplifications are necessary to create a solvable model. This can be done by aggregation of data, for example regarding demand destinations, average transport costs or average speeds (Flodén *et al.*, 2017)

2.3 Heuristic Solutions

If models are still not solvable with commercial mathematical programming software, heuristic techniques are used to simplify the model's computational effort. Large numbers of decision variables, numerous constraints or a large amount of input data usually increases model complexity. Therefore, rather than always reaching exact solutions, heuristic ones are commonly presented, as they provide a solution close to optimality. For a heuristic solution, a range of accepted quality can be previously defined (Melo *et al.*, 2009). Heuristics further require less computational costs, which also makes them more applicable to real-life decisions (Litvinchev and Espinosa, 2012). Cordeau *et al.* (2006) state that optimal solutions of real-life problems are often not meaningful, as the error margin of input data would already exceed 1%. Consequently, optimal solutions would only be optimal in terms of computation, but already contain an error resulting from the input data and simplifications being made. Thus, solutions up to 1% optimality based on heuristic techniques would be a realistic tradeoff between realism and computational costs.

There are certain standard techniques and classifications for heuristic solutions (Litvinchev and Espinosa, 2012), but they are mostly specific to individual problems. For that reason, heuristic solution techniques cannot be discussed in detail before the initial model has been developed. Hence, this topic will be theoretically introduced and applied in chapter 5.

To conclude, models need to be well elaborated for practical problem solving to actually support decision-making (Melo *et al.*, 2009).

2.4 Verification and Validation

The previous chapter illustrates the individualistic nature of models caused by the specific problem as well as the chosen simplifications and the complexity of the observed systems. Since models further search for unknown results, it is crucial for a model to be tested regarding functionality and reliability. Although verification and validation are core elements of model development, there is no standard procedure. In general, model verification relates to making sure that the model gives the right result and works correctly, while validation investigates whether the model represents reality or the topic being investigated (Collis and Hussey, 2014; Janová, 2012). Per definition, a model is a simplification of reality, which is why it can only be 'good enough' in terms of validation rather than 'right' (Kleijnen, 1995). Consequently, Collis and Hussey (2014) state that it would be usually easier to design a verified research than reaching high validity.

First, a model can be verified regarding the calculation to check the correctness of the mathematical formulation. In addition, the programming code can be verified in the model implementation stage (Thacker *et al.*, 2004). For that, the authors suggest to use a small data set. In optimization problems, Janová (2012) further claims to check whether the model's constraints are fulfilled. Depending on the problem, it might however be time-consuming to gather the data for that. In this thesis, a small and partly fictive data set is therefore used to verify functionality of the model including its constraints.

For validation, a quantitative approach is suggested by Thacker *et al.* (2004) and Janová (2012), which compares the model output with actual results. For that, the necessary data needs to be accessible, which is not the case for this thesis. The qualitative technique is generally applicable to investigate whether the initial question has been answered by the model, regardless the simplifications being made (Janová, 2012). The previously mentioned data dependence as well as deviation from the model optimum caused by the input data further calls for validation of input data. Model validation in this thesis therefore discusses how realistic the developed model of an automotive spare parts network is by contrasting identified network characteristics with chosen model characteristics. For data validation, input data is presented and – if possible – compared to research findings. Since this thesis' model development is built up from a mix of general modeling techniques and specific characteristics of spare parts networks, both sides are validated.

The previously presented drivers and challenges in model development result in the methodical procedure for model development as shown in Figure 1. Here, verification is done twice, for the initial model as well as for the heuristic solution technique.

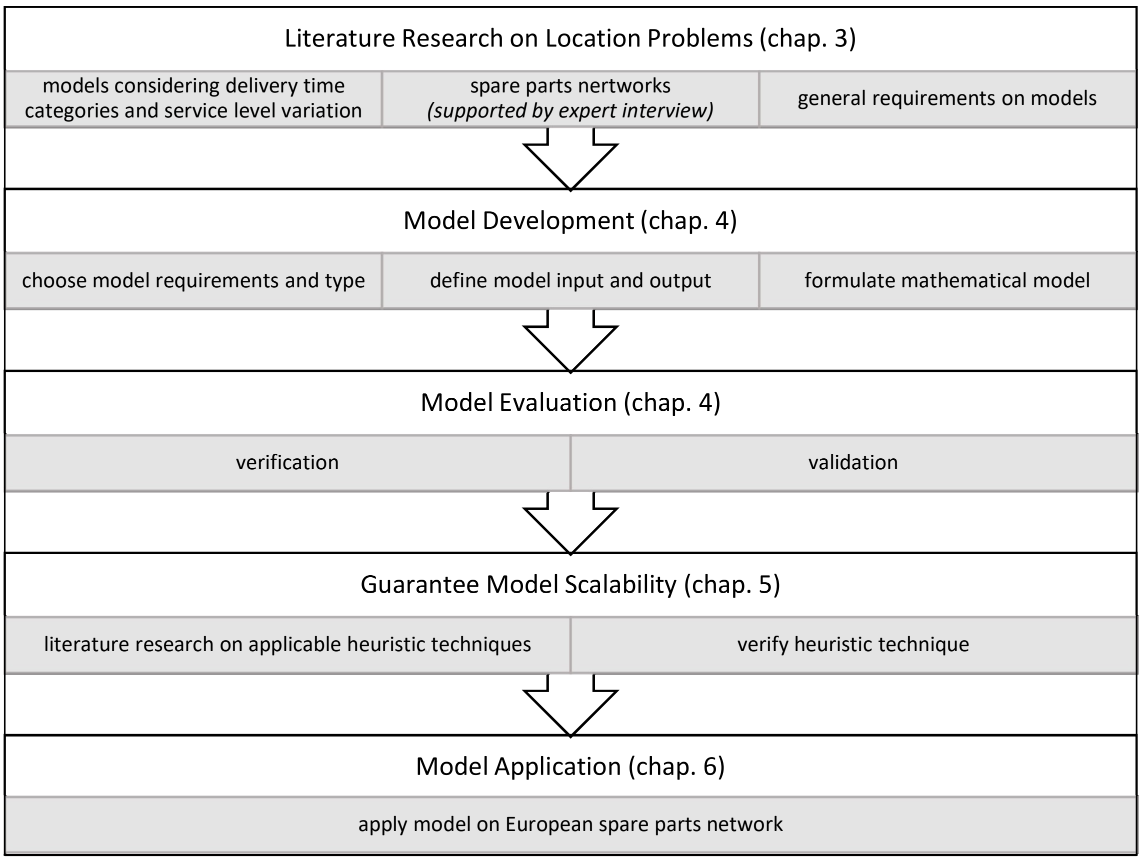

Figure 1: Methodical procedure for model development

3 Theoretical Background

This chapter provides the theoretical background behind the desired model. For that, relevant model classifications are listed, followed by related problems from different areas considering delivery time categories and service level variation. In addition, spare parts network characteristics are introduced, and finally, an illustration of the research gap concludes this chapter.

3.1 Location Problem Classification in Literature

The most prominent basis for discussion of warehouse location problems is the 'Weber Problem', which minimizes the sum of weighted distances to deliver the optimal location. It is commonly quoted in literature and formulated as (Drezner *et al.*, 2010):

$$\min W(x,y) = \sum_{i=1}^{n} w_i \, d_i(x,y) \tag{3.1}$$

Here, w_i is defined as weight (i.e. transport costs), d_i as distance between points x and y, and W as total costs. Seeing this as a starting point, location problems are built individually according to its purpose. Further reaching examples for that can be minimization of average time to market, maximization of distances from the public or the setting of railway stations so that unpredictability of delivery schedules is minimized (Zanjirani Farahani and Hekmatfar, 2009). The model would thus adjust the objective function and add case-specific decision indices and constraints to be fulfilled.

Next to the classifications introduced in chapter 2.2, another fundamental classification is made by Zanjirani Farahani and Hekmatfar (2009) into location problems, allocation and location-allocation problems. Further detailed classifications of facility location problems, focus on different network types and large numbers of sub problems like consideration of warehouse installation costs, number of products, time periods or single sourcing constraints (Bagherpoor *et al.*, 2009; Melo *et al.*, 2009; ReVelle *et al.*, 2008). These detailed classifications of location problems show the extensive attention they have received in research. However, the different categories share the common ground of minimizing the total cost of opening facilities at locations and allocating customers to them (Zanjirani Farahani and Hekmatfar, 2009). The focus of this thesis is on model classifications linked to the two main characteristics of the model to be developed: service level variation and delivery time categories.

For that reason, the field of location routing is neglected in this thesis. These problems add the aspect of tours starting from a facility visiting multiple customers and thus makes them become more complex problems, since demand nodes are not only allocated to facilities, but also to routes. For those, the optimal order of customers to be visited needs to be set (Hassanzadeh *et al.*, 2009). These tours could then be constrained by total delivery time or distance (Rath and Gutjahr, 2014) and would therefore allow a high degree of detail in terms of time-constrained transports. However, the computational effort would increase significantly.

Concerning service level variation, set covering problems are to be named as appropriate model type, since they are characterized by maximizing the number of connected demand nodes. They are further divided into total covering problems and partial covering problems by Fallah *et al.* (2009). In partial covering problems, i.e. time, distance or budget constraints do not allow all demand nodes to be connected, which would often be the case in real life. In such case, for example, the maximum number of warehouses can be given (exogenous) or those demand points, which are not connected, create penalty costs. The solution would consequently return the optimal number of customers supplied so that network costs are minimized. However, the model requires to identify these penalties first (Zanjirani Farahani and Hekmatfar, 2009).

Furthermore, covering problems aim to connect every demand node with a hub node regardless of demand weights. In real life, this methodology– next to product distribution and warehouse location – could be applied for mail delivery or emergency service facilities (Fallah *et al.*, 2009; Charikar *et al.*, 2001). The latter case is a good example for the need to reach every demand node within a certain predefined distance or time constraint where no demand preferences are used. Rather than choosing the location according to the possibility of reaching a majority of buildings in a very short time, it would be chosen in a way, that also suburbs could be reached in a reasonable time, although the 'demand' might be lower in those areas, due to a lower population density. Finally, following Fallah *et al.* (2009), one of the covering problems' main characteristics is that facility capacities are not considered.

Within discrete location problems, ReVelle *et al.* (2008) classify 'median and plant location problems' as the opposite of 'center and covering problems'. The former would aim to minimize weighted distances between nodes, which is why in most cases, the nearest nodes are selected. The popular p-median problem, in which the number of warehouses to be located is constrained by p, also belongs to this category. The opposite of this, when the number of warehouse or links are not constrained, is called the uncapacitated facility location problem (Melo *et al.*, 2009).

Delivery time constraints are not explicitly classified, but it is observed that distances are often used instead of time units (Toregas *et al.*, 1971; Balcik and Beamon, 2008). In these cases, the authors restrict every single link not to exceed a certain distance.

3.2 Related Research on Location Problems

3.2.1 Delivery Time Categories

In general, time is a crucial factor, when it comes to location planning, but can affect the problem in different ways. Either time relates to direct transports or, like in location-routing problems, total tour duration is limited, as the following examples show.

The first reference research by Papathanasiou and Manos (2007) investigates a network problem regarding distribution of perishable products. The goal is to locate the production facility for every product with a different maximum distribution time so that products arrive at the customers' sites in time. The maximum transport time originating from the customer is assumed to be known, so that a binary variable indicates whether a warehouse location is within range or not. Consequently, one constraint ensures that total supply volume at the manufacturing plant for one customer lies within its predetermined range by only allowing to open facility locations in range.

Balcik and Beamon (2008) chose a similar approach in their covering problem and also determined in advance, whether or not a facility would be in range to fulfil a delivery time constraint. Here, the authors plan warehouse locations for humanitarian logistics where different delivery time categories originate from criticality of relief items. The authors de-

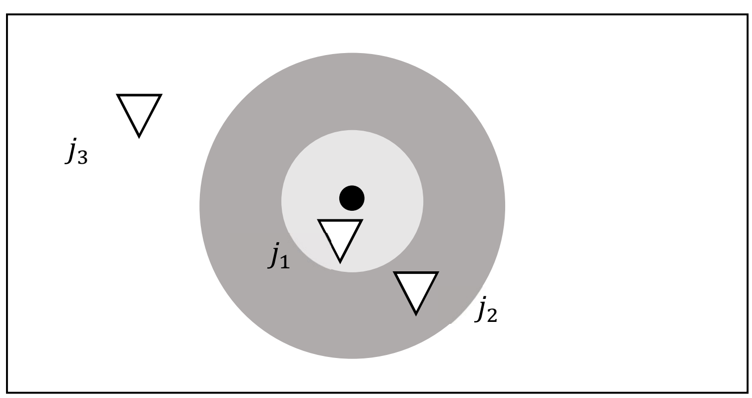

Figure 2: Subsets of potential warehouse in range (adjusted from Balcik and Beamon, 2008)

termine possible disaster locations and characterize them with perimeters standing for different coverage levels. These coverage levels implicate the arrival time at the possible disaster location. Figure 2 shows exemplarily how potential locations are classified: warehouse location $j_1$can supply the disaster location within the first time window, location j_2 within the second time window and j_3 would not reach the disaster location in time to have any impact. Hence, potential warehouse locations are assigned to subsets of the set of potential warehouse locations N. For every disaster location, there is a subset of N called $N_s(l_k)$, containing all potential warehouse locations that can provide the coverage level l for item k for demand point s. In the objective function, only subsets are considered as potential locations, which maximize total demand covered of the network:

$$\max \sum_{s} \sum_{k} \sum_{l_k} \sum_{j \in N_s(l_k)} p_s \, d_{sk} \, w_k \, \alpha_k^{l_k} f_{sjk} \qquad 3.2$$

with $p_s$being the probability of scenario s, d_{sk} the expected demand for item k in scenario s, w_k the weight of item k, $\alpha_k^{l_k}$ the coverage level weight and f_{sjk} the proportion of item type k demand satisfied by distribution center j in scenario s. This approach requires to determine all possible connections in range from a facility location or demanding location in advance. Subsequently, this causes a more detailed preparation of the model's input data depending on the number of nodes and accuracy.

The same approach has been chosen by Altiparmak *et al.* (2006). A detailed description however, is not provided here, as it can already be concluded that classification into subsets appears to be an appropriate way to model the supply within several delivery time categories for direct transports from warehouses to customers.

3.2.2 Service Level Variation

In order to model different service levels or, in other words, varying or choosing the number of demand nodes to be connected, three techniques are found that are relatable to the initial problem.

Charikar *et al.* (2001) built an analytical model of an uncapacitated facility location problem (covering problem) adding penalty costs to the objective function value for every customer j, in case it is not connected to any facility.

$$\min \sum_i f_i * y_i + \sum_{ij} c_{ij} * x_{ij} + \sum_j p_j * r_j \quad 3.3$$

subject to

$$x_{ij} \le y_i, \forall ij \quad 3.4$$

$$\sum_i x_{ij} + r_j \ge 1, \forall j \quad 3.5$$

$$x_{ij}, y_i, r_j \ge 0 \quad 3.6$$

Here, f_i are fixed costs to open facility i, c_{ij} the connection costs between facility i and customer j, and x_{ij}, y_i and r_j are binary indicators for whether or not i and j are connected (x_{ij}), whether or not i is opened (y_i) and whether or not customer j is connected to a facility (r_j). By adding individual penalty costs, customers can be prioritized according to their importance for the network and how excluding them would be weighted. However, the necessity to have information in advance to be able to determine needs is designated as the model´s main weakness by the authors. For a spare part network with a large number of customers, it can be quite time-consuming to reasonably weight those. Additionally, effects on (i.e. a multi-level) network structure are not always trivial.

Another model suggested by Charikar *et al.* (2001) assumes that there is a share p of outliers, which can be excluded. Thus, with n defined as the number of total customers, the authors define the maximum number of customers to be excluded as $l = n - p$. This leads to the following model formulation:

$$\min \sum_i f_i * y_i + \sum_{ij} c_{ij} * x_{ij} \quad 3.7$$

subject to

$$x_{ij} \le y_i, \forall ij \quad 3.8$$

$$\sum_i x_{ij} + r_j \ge 1, \forall j \quad 3.9$$

$$\sum_j r_j \le l \quad 3.10$$

$$x_{ij}, y_i, r_j \ge 0 \quad 3.11$$

Thus, the authors allow to vary the total amount of customers supplied in the network. Contrary to the previous approach, no extra customer information is necessary, but the model excludes l customers with the deepest impact on total network costs. Depending on the cost structure, selection of outliers is based on transport costs, which can – in other problems - be heavily influenced by either distance or transport volume. Therefore, not only remote, but also important customers might be excluded.

Contrary to the models of Charikar *et al.* (2001), Balcik and Beamon (2008) limit their maximal covering problem by total budget provided in case of an disaster. In terms of service level, the goal is to minimize the number of deaths, meaning to maximize the (weighted) demand supplied. For that, the technique of building subsets (see Figure 2) of facility locations in range of a demand (disaster) location is used. Due to the time constraints, service can either be provided or not. Thereby, for every connection between facility and demand node, it is known in advance whether the service can be provided. Unlike the previous models, service level is in this case not restricted by an imposed percentage or additional costs but limited through time and monetary resources provided. According to Fallah *et al.* (2009), this would be close to real-life network design decisions, which are mostly limited by monetary constraints.

To sum up, the described approaches slightly differ in the reasons for customers to be excluded and therefore show different weaknesses. The first model from Charikar *et al.* (2001) requires a well-elaborated sophisticated monetary value for every customer, who is not being supplied, while the second approach from Charikar *et al.* (2001) might tend to exclude important customers in case of volume-dependent transport costs. The covering-problem specific technique of Balcik and Beamon (2008) requires to limit total network costs, which impedes to identify the actual costs to offer a higher service level.

3.2.3 General Model Requirements

The review on location problems by Melo *et al.* (2009) gives insight into model characteristics and research gaps for more practical oriented modelling. Firstly, the authors claim modelling of different transport modes. This clearly involves non-road transport, but also goes along with the research on transport costs by Lapierre *et al.* (2004), distinguishing between full truckload (FTL), less-than truckload (LTL) and parcel (see Figure 3). Different transport tariffs also settle the claim by Melo *et al.* (2009) for non-linear transport costs, since those would not be the case in reality and lead to inaccurate results.

In addition, the authors demand the use of road distances instead of less accurate distance measurement technique like the Euclidean distance (Melo *et al.*, 2009). Such accuracy would provide the basis for networks in which not always the closest facility is chosen, but the economic one due to the cost structures in Figure 3 (ReVelle *et al.*, 2008).

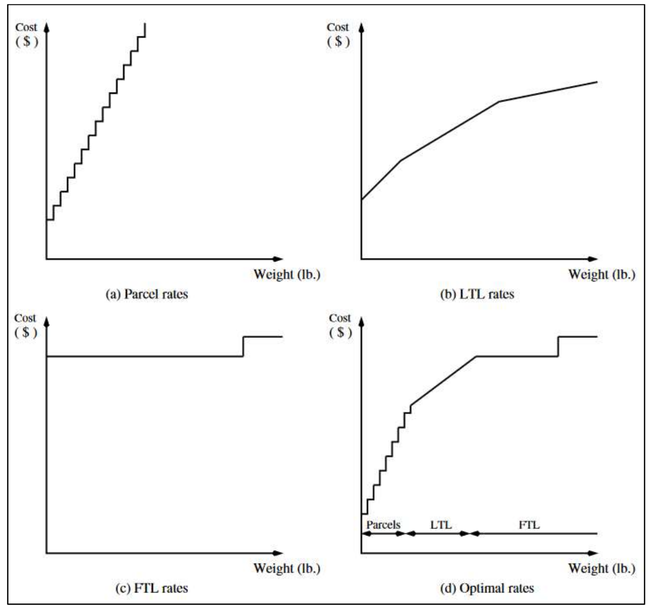

Figure 3: Road Transport Cost Functions (Lapierre et al., 2004)

Melo *et al.* (2009) go one step further with not only aiming for more-than-two layer problems, but pointing out the necessity to not solely supply customers from the lowest layer. Such network structure would mean to for example not only supply from regional warehouses, but also directly from a central warehouse. In reality, this might be convenient for large deliveries or great urgency.

Moreover, Melo *et al.* (2009) identify a research gap in including multiple commodities into location problems, as their different characteristics would be a relevant issue in practice. Other relevant areas are intra layer flows, which describe the balancing of inventory between warehouses of the same layer, and capacities and their expansion over time.

3.3 Location Problems in Spare Parts Networks

3.3.1 Spare Parts Network Characteristics

Large numbers of suppliers and repair shops as well as even larger numbers of end-customers, which in case of the automotive industry would be vehicle owners or operators, characterize spare part networks. A company like Volvo Group (combining “trucks, buses, construction equipment, marine and industrial engines and aero”, Huiskonen, 2001) deals with availability of around 100,000 different parts. This should satisfy customers’ high service expectations regarding delivery speed and long-term repair possibility. The core challenge is to provide sporadic parts supply, which can be very difficult to forecast, to customers of current and previous product families (Holmqvist and Pessi, 2006; Wagner *et al.*, 2012; Huiskonen, 2001). Service requirements can however vary according to customer priority or location (i.e. due to size of city) or due to the service level offered by the OEM. To determine these service classifications are other core issues in spare parts management (Schittekat and Sörensen, 2009; Wagner *et al.*, 2012). According to Huiskonen (2001), ABC analysis would be a common tool to classify parts into price and demand categories, but with increasing complexity, the need for multi-criteria classifications in inventory management arose. While ABC analysis relates to annual turnover, it can be expanded by XYZ analysis indicating usage regularity. Following both approaches, a classification matrix arises where AX parts would be of high value and constant demand, while CZ values would be of low value and sporadic demand (Scholz-Reiter *et al.*, 2012).

For distribution of aftermarket spare parts in Europe, the car manufacturer Toyota uses two central warehouses - in the Czech Republic and in Germany - and the service of third-party logistics providers, which use their network consisting of warehouse and transportation capacities (see Figure 4). Calls for spare parts come in during the day and can vary in urgency: Toyota promises ability to deliver the next day, but customers might ask for no-urgent order to be supplied later. In such orders, urgency can either be based on customer (repair shop) preference or part urgency (Chen *et al.*, 2006; Schittekat and Sörensen, 2009). Following Schittekat

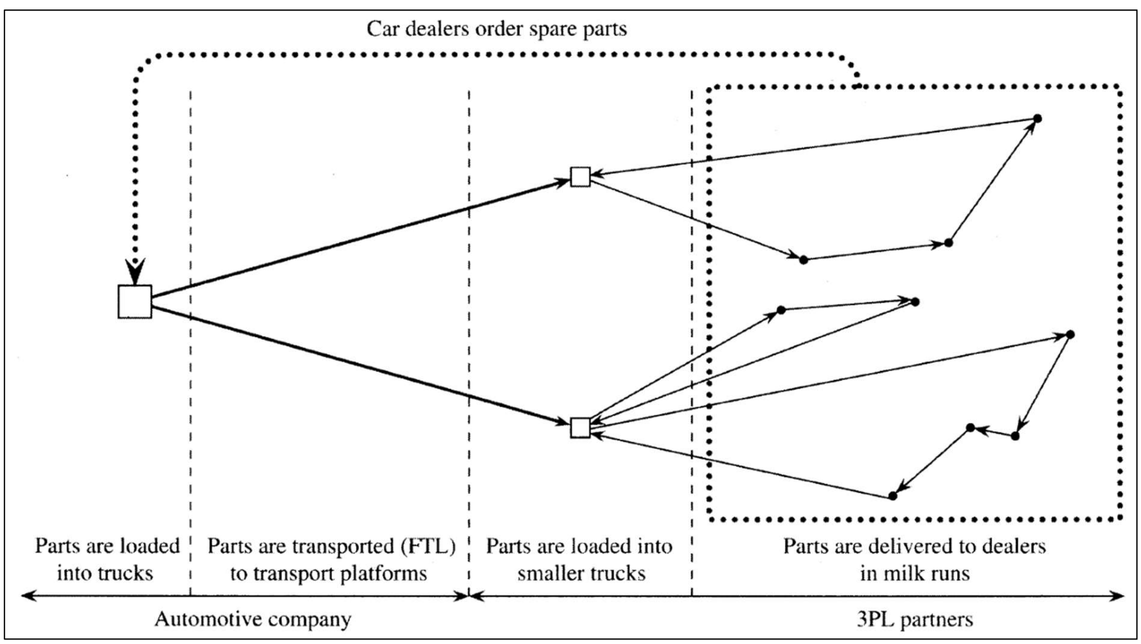

Figure 4: Spare Parts Network Structure at Toyota (from Schittekat and Sörensen, 2009)

and Sörensen (2009), spare parts are usually shipped as FTL to warehouses offering handling and crossdocking. There they are loaded onto smaller trucks and distributed via delivery tours to repair shops. In Toyota's case, European operations are taken over by different 3PLs, each covering one or more small regions. This network separation leads to a constantly changing situation for Toyota in terms of service level, costs, inventory management but also in terms of communication and negotiation. However, the authors point out relatively high costs for Toyota's setup.

The approach of Toyota running few central warehouses in combination with regional warehouses seems to be the predominating network structure in the automotive industry (Li, 2015, 2015; Wagner *et al.*, 2012). In other industries, different spare part network designs have been found by Wagner *et al.* (2012). In their research, some OEMs distribute from one centralized warehouse only, where focus companies benefit from lower overhead costs compared to decentralized approaches but increased outbound transportation distances and costs.

Others use a mix of local and central warehouses, where the OEM chose the locations close to its main customers. In the paper machines industry, one OEM even keeps spare parts in local warehouses, which allows it store in proximity to its customers. This can be beneficial, if for instance customs processes are part of the transport process and might delay delivery. On the other hand, such network structure only works for standard parts, since they allow to replenish in larger batches and by that to reduce probability to run out of stock (Huiskonen, 2001; Wagner and Lindemann, 2008). Moreover, this network design coincides with Yaobao *et al.* (2013) simplified two-stage model of an automotive network in which spare parts are supplied from factories to warehouses, and then to the repair shops.

To sum up, those companies aiming for cost efficiency in their spare part supply choose a centralized warehouse, while those trying to create long-term partnerships try to set up a local and responsive network (Wagner *et al.*, 2012). Obviously, strategies are not always that easy to classify, and this is also not a tradeoff limited to spare part network design, but faced in general supply chain issues.

3.3.2 Reference Location Problems

After gaining insight into the classical warehouse location problem and maximal covering problems in the previous chapter, reference spare parts network problems are introduced in the following.

The two identified location choice frameworks in spare parts networks aim to find the lowest combination of transport and warehouse costs. Yaobao *et al.* (2013) mathematically formulate the tradeoff as a linear optimization problem (see Figure 5). The model of Yaobao *et al.* (2013) delivers the optimal number and location out of a set of potential warehouse locations but for that requires the corresponding input data to be available. In this case, these are total demand volumes, warehouse fixed costs, distance- and volume-dependent transport costs, distances between all supplier – warehouse and warehouse – customer connections. Further, the authors add a budget constraint and a maximum number of opened warehouses and propose a heuristic to solve the problem (Yaobao *et al.*, 2013).

Schittekat and Sörensen (2009) investigated the network of Toyota described in chapter 3.3.1 and viewed the problem from a slightly different angle. Instead of assuming direct transport connections with a linear cost structure, different transport tariffs are included in their model to minimize total network costs. In this case, Toyota ships parts as FTL from central warehouses to crossdocking warehouses and then a 3PL distributes in milk runs. In order to consider the different cost structure in milk runs, the authors viewed the problem as a location-routing problem, making the warehouse locations dependent on route length, volume and stops. Their overall goal was to improve 3PL selection at Toyota, since their transport networks offered to Toyota differ through warehouse and vehicle assets. However, taking into consideration the quick-changing transport providers, this problem is more likely to belong to the tactical planning level rather than the strategic one.

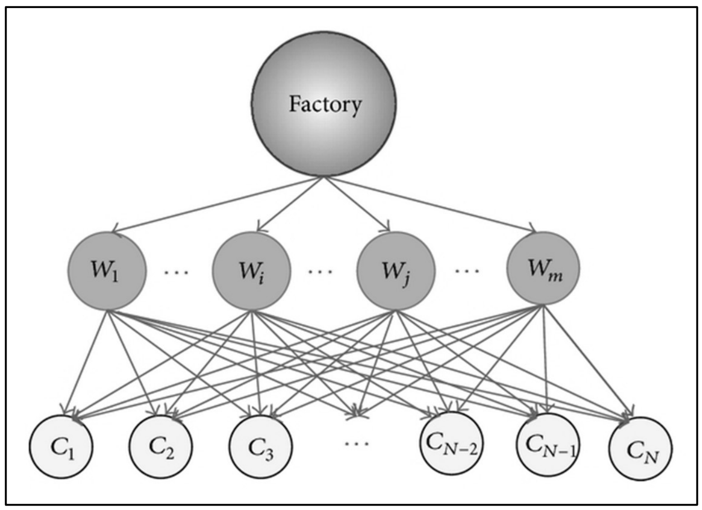

Figure 5: Network structure of warehouse location problem in a spare parts network (Yaobao *et al.*, 2013)

Combined with the access to real data, this is the main reason for the Toyota case study to show a significantly higher detail degree compared to Yaobao *et al.* (2013). In Schittekat and Sörensen (2009), for example delivery time windows, different vehicle speeds and capacities and a driving time constraint by legislation are part of the model which allow them to solve a real-life problem in their research. Contrary to that, Yaobao *et al.* (2013) optimize a facilitated network structure, which is rather supported by assumptions on core issues like transport costs.

With regard to the spare parts characteristic of uncertainty in demand, stochastic location models address this issue. Uncertainty is however not limited to demand, but also concerns travel time, facility costs, distances, availability of a facility and number of facilities (Current *et al.*, 2010). Discrete location problems also provide the opportunity to include subjective weights and to run different scenarios. If however parameters like demand change predictably over time, a multi-period approach with changing parameters is suggested by Melo *et al.* (2009).

3.4 Research Gap

With regard to the research question, both presented location problems from automotive spare parts networks neglect the opportunity to exclude outliers. This is not relevant for the Toyota case, since the company's guarantee to deliver – even the next day - to all European customers does not allow the option to exclude the ones not making good economic sense. Relating to Wagner *et al.* (2012), Toyota's strategy is in accordance with the overall goal of the spare parts business to improve a company's image through a high service level and short delivery times, rather than being primarily cost-oriented. Although not explicitly discussed, (geographical) outliers have a deep impact on tour planning. Including an outlier in this scenario means reducing the number of customers supplied in that particular tour and thereby causing an additional tour with an additional vehicle and the related transport costs. To conclude, it is likely to be a cost-sensitive issue for Toyota.

By setting one constant delivery time category of one day, Toyota also needs to limit daily delivery tours with time constraints (Schittekat and Sörensen, 2009). Since all customers need to be supplied the next day, a customer exceeding a route time capacity would also result in creating an additional route causing the related transport costs. In order to compute the cost-optimal delivery time to be offered, some significant changes in the model would probably be necessary. It is concluded though, that different delivery time categories within a network have not been considered mathematically in spare parts networks yet. In addition, research on strategic spare parts management is justified by an observable and expected growth in profit and complexity in the aftermarket business (Wagner *et al.*, 2012). The developed managerial framework by Wagner *et al.* (2012) how to align a company's spare parts strategy aims to provide a holistic concept. Network design is one sub problem in that framework, which the Toyota case expands with rather mixed strategic and tactical questions on service level and delivery time categories arising.

Within the literature research on related location problem methodology, the author has found no approach combining service level constraints and different delivery time categories (see). Instead, many covering problems are constrained by for example total network costs and location-routing problems by route length or duration.

Author(s), Year	**Mathematical location problems including…**				
	...number of facilities	*... location of facilities*	*...time constraints*	*...delivery time classifications*	*...outliers*
Papathanasiou and Manos, 2007	✗	✓	✓		
Badi et al., 2017	✓	✓	✓	✓	
Balcik and Beamon, 2008	✓	✓	✓	✓	
Rath and Gutjahr, 2014	✓	✓	✓	✓	
Tavakkoli-Moghaddam et al., 2010	✓	✓	✓	✓	
Rottkemper et al., 2012			✓	✓	
Charikar et al., 2001	✓	✓			✓
Xu and Xu, 2009	✓	✓			✓

Table 2: Identified location problems on problem-relevant aspects

As a result, including strategic spare parts networks characteristics into quantitative location planning fills a research gap from content-related perspective as well as in generally applicable model development requirements.

4 Model Development

This chapter presents the research gap, chosen model characteristics and the model itself. After that, input data and necessary pre-processing steps are described. Lastly, the model is validated and verified.

4.1 Model Characteristics

Again, the focus of the model to be developed lies on service level variation and different delivery time categories, which is why complexity is to be reduced in other areas rather than in these core ones. For that reason, the modeled network structure is based on the more simplified Yaobao *et al.* (2013) consisting of suppliers, warehouses and customers (see Figure 5). The overall question the model is supposed to answer, is which warehouses to open and how to allocate customers to warehouses so that network costs are minimized, while considering different service levels for delivery time categories. Therefore, the model falls under the classification of location-allocation problems (Zanjirani Farahani and Hekmatfar, 2009). Moreover, a set of potential warehouse location is assumed to be given, which makes the problem become a discrete optimization problem. Also, locations of customers and suppliers are known. However, the model is not constrained by maximum number of warehouses to be opened or by maximum network costs and hence does not belong to covering problems. Instead, it aims to quantify the effect of an increasing service level, for which no upper limit is set and can hence be classified as an uncapacitated warehouse location problem.

Warehouse Locations

Following Melo *et al.* (2009), warehouses are capacitated. As Schittekat and Sörensen (2009) and the expert interview revealed, an automotive spare parts network contains several warehouse levels (i.e. central, regional and cross-docking warehouse). In order to reduce complexity, the different levels with their strongly increasing number of possible connections are not adopted, but at every warehouse location, there is a set of different warehouse types to choose from. These warehouse types differ in fixed costs, variable costs and capacity. By that, it is possible to model the difference between a smaller warehouse type of lower capacity, lower fixed costs and higher variable costs, and a larger warehouse type of higher capacity, higher fixed costs and lower variable costs (Maßmann, 2006). Warehouse costs are further country-dependent, and thus allow the model to distinguish between locations in high and low-wage countries and to take into account drivers behind geographical shifts to i.e. Eastern Europe (Schmidt and Wilhelm, 2000).

According to the expert interview, it is common in the allocation of customers to warehouses, to always supply customers from the same warehouse. This would ease operational processes on transport and receiving side and moreover facilitate control of delivery time requirements (expert interview). Thus, every customer is supplied from one warehouse.

Products

Further, the spare-parts typical variety of products is to be considered. For that, customers' demand for every product is assumed to be known and total demand volume equals total supply volume. Here, the previously discussed single sourcing is modeled by connecting every customer to exactly one warehouse. Moreover, it is assumed that every product is only supplied from one supplier, but one supplier can manufacture several products. This simplification is motivated by reduction of computation time as the number of possible flows between suppliers and warehouses is significantly reduced. Consequently, customer demand and warehouse capacity are measured in kg. Here, kg is chosen as single unit to facilitate the interdependency between capacity, variable warehouse costs and demand volume.

Delivery Time Category and Transport Tariffs

It is modeled that customers can be supplied within different delivery time windows. So, every customer is assigned to a delivery time category t, which restricts maximum transport time

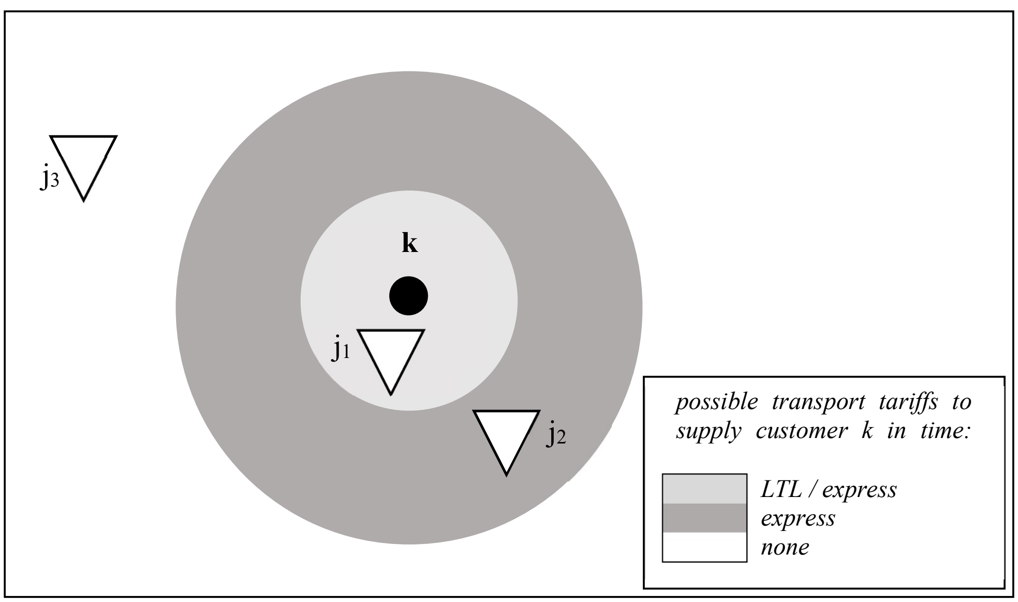

Figure 6: Possible Transport Tariffs Depending on Distance

from warehouse to customer to i.e. $t_1 = 4h$. For that, transport time is converted into road distance in km. Furthermore, transport time is dependent on the chosen transport tariff. It is possible to supply a customer with an LTL transport in its assigned delivery time from warehouses within a certain perimeter. Moreover, a second perimeter exists from which the customer can also be supplied in time with a faster transport at higher costs (see Figure 6). With this, the model fulfills the demand for different modes of transports. Figure 7 gives an overview about the modeled network structure and the possible transport tariffs. Consequently, the model addresses the question whether it would be economic to accept higher transport costs instead of opening an additional warehouse for remote customers.

Spare parts' sporadic demand calls for the use of ABC/XYZ classifications and statistical distributions on product level, but the expert interview shows that due to the high variety of products, a classification from a transport perspective is done on customer level. While forecasting of product demand would be relevant for inventory management, for route planning, customers would be categorized by number of supplies per day. Following the XYZ logic, an 'X customer' would thus for example be supplied three times per day. Therefore, the model considers different delivery frequencies for the different delivery time categories (t_1 has the highest frequency and shortest maximum distance between warehouse and customer). This further strengthens the effect of customer classification, as transports to t_1-customers not only require warehouse proximity, but also impede consolidation.

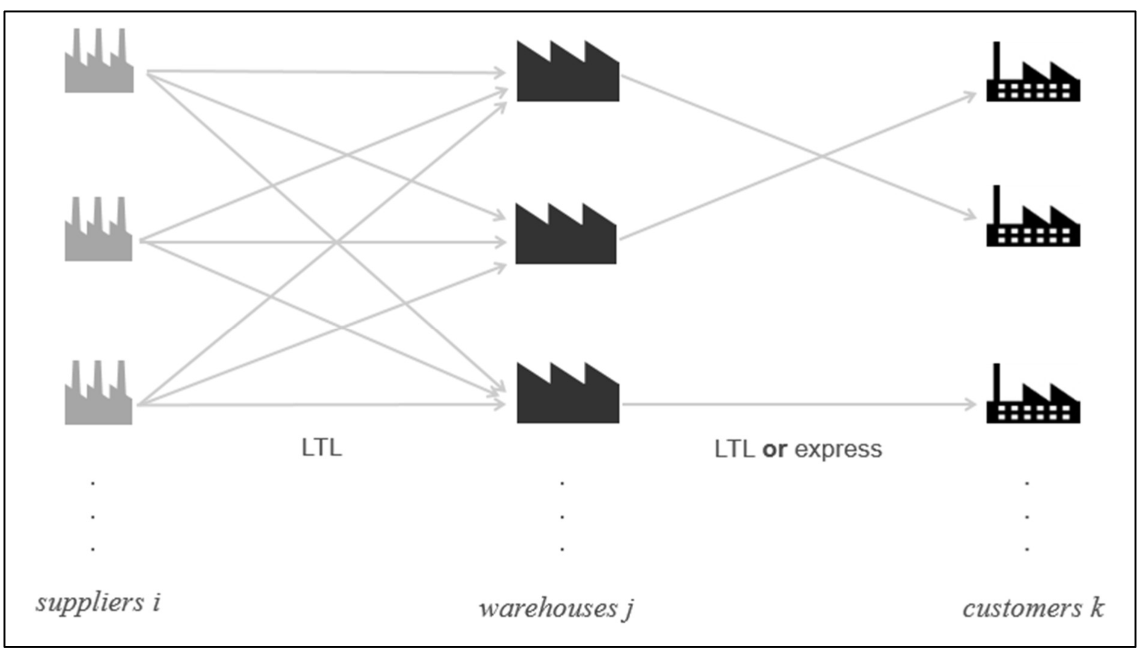

Figure 7: Simplified Model Structure

Service level

Finally, the monetary effect of an increasing service level is modeled by running different scenarios. One service level is determined for every delivery time category so that i.e. 90% of all customers of delivery time category t_1 are supplied in time. However, also if not supplied in time, every customer is allocated to a warehouse and thus supplied later than required. Hence, the service level refers to whether or not a customer is supplied in time, and it is not possible to exclude customers. This goes along with Toyota's strategy to supply all customers directly (Schittekat and Sörensen, (2009).

4.2 Mathematical Formulation

Resulting from the chosen model requirements, the following sets, parameters and variables are defined.

sets

J	set of potential warehouse locations $J, j \in J = \{j_1, j_2, \dots, j_q\}$
K	set of customers $K, k \in K = \{k_1, k_2, \dots, k_r\}$
$J(k) \subseteq J$	set of potential warehouse locations *j* that can supply customer k within its assigned delivery time category t
$J'(k) \subseteq J(k)$	can supply customer k within delivery time with standard tariff
$J''(k) \subseteq J(k)$	can supply customer k within delivery time only with express tariff, $J'(k) \cap J''(k) = \emptyset$
$K_t \subseteq K$	set of customers assigned to delivery time category $t, k \in K_t = \{k_1, k_2, \dots, k_m\}$
T	set of delivery time categories $T, t \in T = \{t_1, t_2, \dots, t_n\}$
B	set of transport tariffs $B, b \in B = \{b_1, b_2, \dots, b_m\}$
W	set of warehouse types $W, w \in W = \{w_1, w_2, \dots, w_v\}$

parameters

h_{jw}	capacity for warehouse type w $[kg]$
f_{jw}	fixed costs for warehouse type w at location j [€]
g_{jw}	variable costs of warehouse type w at location j $[€/kg]$
l_{ka}	demand volume of every product of customer k during the observed time frame $[kg]$
l_k	total demand volume of customer *k* during the observed time frame $[kg]$

c_{jkb}	transport costs from warehouse j to customer k with tariff b including transport of a from i to j and considering l_k and distance
s_t	share of customers supplied in time per delivery time category t [%]

decision variables

Y_{jw} $\begin{cases} 1 \text{ if warehouse type } w \text{ at location } j \text{ is chosen} \\ 0 \text{ otherwise} \end{cases}$

Z_{jkwb} $\begin{cases} 1 \text{ location } j \text{ with type } w \text{ is assigned to customer } k \; with \; tariff \; b \\ 0 \text{ otherwise} \end{cases}$

According to Melo *et al.* (2009), the use of binary decision variables would be typical for strategic network decisions, while continuous variables would rather be associated with tactical and operational decisions.

Moreover, transport costs c_{jkb} should be clarified at this point. Since a decision variable indicates whether or not transport costs between warehouse and customer occur, this allows a more sophisticated pre-processing of parameters. As for every customer the demand for every product is known, as is the corresponding supplier, total transport costs for a customer's demand volume along the supply chain is only dependent on the warehouse chosen. Thus, every warehouse-customer connection can be added to the transport costs for the products from the corresponding suppliers to the particular warehouse (see Figure 8). By that, all the decision-relevant transport costs can be summarized in one parameter, which then requires a binary decision variable to choose this warehouse-customer connection. Considering this, the following linear minimization problem is defined.

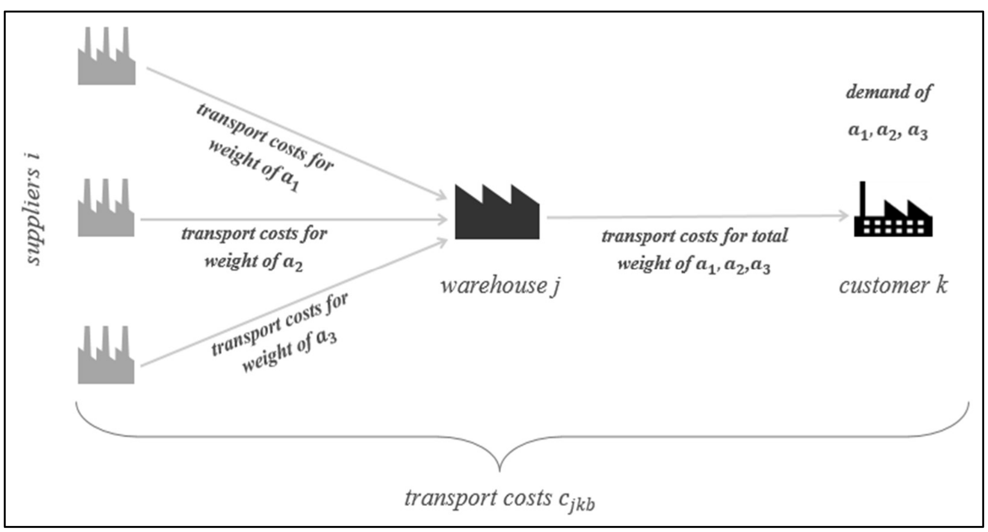

Figure 8: Visualization of transport costs calculation

$$\min Z = \sum_{j=1}^{q} \sum_{w=1}^{v} (Y_{jw} f_{jw} + g_{jw} \sum_{k=1}^{r} \sum_{b=1}^{m} Z_{jkwb} l_k) + \sum_{j=1}^{q} \sum_{k=1}^{r} \sum_{w=1}^{v} \sum_{b=1}^{m} Z_{jkwb} c_{jkb} \tag{4.1}$$

subject to

$$\sum_{j=1}^{q} \sum_{w=1}^{v} \sum_{b=1}^{m} Z_{jkwb} = 1, \forall k \in K \tag{4.2}$$

$$\sum_{w=1}^{v} Y_{jw} \leq 1, \forall j \in J \tag{4.3}$$

$$\sum_{k=1}^{r} \sum_{w=1}^{v} \sum_{b=1}^{m} Z_{jkwb} l_k \leq \sum_{w=1}^{v} Y_{jw} h_{jw}, \forall j \in J \tag{4.4}$$

$$\sum_{k \in K_t} \left(\sum_{j \in J'(k)} \sum_{w=1}^{v} \sum_{b=1}^{m} Z_{jkwb} + \sum_{j \in J''(k)} \sum_{w=1}^{v} Z_{jkwb_{expr}} \right) \geq s_t |K_t|, \forall t \in T \tag{4.5}$$

$$Y_{jw}, Z_{jkwb} \in \{0,1\} \tag{4.6}$$

The objective function consists of warehouse costs (fixed costs and demand-dependent variable costs) and transport costs. Constraint 4.2 makes sure that every customer is supplied from exactly one warehouse location, one warehouse type and with one transport tariff. Constraint 4.3 limits the number of warehouse types chosen at a location to maximum one. Capacity constraint 4.4 guarantees that only connected warehouses are opened and that when allocating customers to a warehouse, their total demand does not exceed the warehouse's capacity. The problem-specific service level constraint 4.5 determines the minimum number of customers to be supplied within their delivery time category. This number needs to be exceeded by the sum of customers supplied from warehouses in range for LTL transport and the ones in range for express transport. Finally, constraint 4.6 defines Y_{jw} and Z_{jkwb} as binary variables. It is necessary to consider warehouse type w in both decision variables, since the connection between opened warehouse type and customer allocated is relevant for variable warehouse costs being dependent on warehouse type and customer demand. Output of the model is the objective function value giving the total network costs of the most cost-efficient network design. Further, for the best network design, the decision variables show which warehouses and types to open and how to allocate customers respectively.

4.3 Validation

It should be shown that the modeled aspects are the relevant ones to be considered when modeling a spare parts network. For that, realism is evaluated by validating the model requirements for both general model development and spare parts networks.

Spare parts networks

Regarding the expert interview, sporadic demand would be one of the core challenges in spare parts logistics and thus calls for the use of statistical techniques on the tactical planning level. However, for strategic issues, Melo *et al.* (2009) see discrete models as appropriate.

A far-reaching simplification is the assumption of a three-level network (supplier, warehouse, and customer). Although such procedure is used in literature to model spare parts networks, the predominant structure consisting of up to three levels of warehouses finds support from the expert interview. Typical warehouse types would be central warehouses, regional warehouses and cross-docking warehouses. Consequently, Z parts would be suitable for a central warehouse, while X parts for regional warehouses. Cross-docking warehouses play an important role in over-night delivery. In Toyota's network, cross-docking warehouses are supplied over night in order to start delivery tours to customers early in the morning.

Furthermore, the model does not consider FTL transports from suppliers to warehouses. Although this is supported by literature findings (Schittekat and Sörensen, 2009), it has been neglected in this model, since it would have further increased complexity. The corresponding constraint 4.7 would have converted the problem into a network flow problem, where N_{ij} is the number of FTL transports from supplier i to warehouse j.

$$cap_{FTL} * N_{ij} \geq \sum_{a \in A(i)} \sum_{k=1}^{r} \sum_{b=1}^{t} Z_{jkwb}\, l_{ka}, \forall\, i \in I, j \in J, w \in W \qquad 4.7$$

However, FTL transports are approximated in the model, as transports from suppliers to warehouses are computed with LTL tariffs, but consolidated for cost calculation.

Moreover, the assumption of single sourcing on both supplier and warehouse level needs to be carefully evaluated. Customers being supplied from only one warehouse is supported from the expert interview with regards to service level fulfillment: this would not only make scheduling and communication easier, but also help to maintain a certain delivery time guarantee. On the suppliers side, purchasing a part from only one source can mean a low risk aversion, as the

OEM takes the risk of supplier default. In spare parts, the fact that only one supplier is available, might be the case (sole sourcing). However, single sourcing does not apply for all parts of a car as such sourcing strategy would always consist of both single and multiple sourcing (Costantino and Pellegrino (2010), expert interview).

Lastly, the simplification of demand and capacity into kg causes a loss of detail. This is done to reduce the number of dimensions per article. Otherwise, capacity, variable warehouse costs and transport costs for demand or transport volumes would have to be measured in both weight and volume.

General model development

With regards to general aspects in model development, demand is not fix (Maßmann, 2006). For this model, a typical assumption is made, where supply volume equals demand volume. Moreover, the model considers multiple products, which is demanded by Melo *et al.* (2009)

The use of road distances provides a high degree of detail as a basis of delivery time measurement. However, this is weakened through the assumption of direct transports from warehouses to customers instead of milk runs. Otherwise, the model would include location-routing and by that significantly increase the number of possible connections. For the same reason, the model does not consider inventory balancing between warehouses, although this would be common in practice and is suggested by Melo *et al.* (2009).

As demanded by Lapierre *et al.* (2004), no linear transport costs are used, but instead transport costs origin from a tariff database, considering some or all of the following input: start and destination country, weight and distance. By that, realistic transport rates can be applied for the model, which strengthen the model characteristic of two different transport modes.

To conclude, all simplifications discussed within validation find support from literature and/ or the expert interview. Thus, it can be stated that simplifications are reasonable and do therefore not impede model validity to an inappropriate high degree. However, as presented in detail in chapter 2.4, it is hard to draw more concrete conclusions related to model validity.

4.4 Input Data

The input data describes a European spare parts network of one OEM from the automotive industry. This means, that all customers (car dealers), warehouse locations and suppliers are located within Europe. The data originates from an anonymized automotive spare parts network provided by 4flow. It consists of 10,000 customers, 77 potential warehouse locations and 1400 suppliers, all of them located in Europe. Further, every article is assigned to one supplier. Historical demand for every customer is known for more than 40,000 different products. Consequently, the large product variety and the fact that not every part is ordered by every customer go along with research findings. Moreover, road distances are available for every supplier-warehouse and warehouse-customer connection.

Customer demand accounts for the timeframe of one year and ranges from about 5,000 to 7,500 kg per customer. Demand is characterized by a large product variety in terms of weight per product. According to the number of deliveries and depending on the customers' delivery time category, transports are calculated for every delivery, so that transport costs c_{jkb} are the sum of costs for these transports.

At every potential warehouse location, there are three different types of warehouses to choose from. In terms of overall capacity, there is always a small (warehouse type 1), a medium (2) and large warehouse type (3). Across different countries, capacities are the same over the observed timeframe for every warehouse type: 800,000 kg (1), 10,000,000 kg (2) and 12,000,000 kg (3). Warehouse fixed and variable costs are however country specific: thus, it is for example more expensive to open a large warehouse in Sweden than one in Bulgaria, due to higher labor costs. The cost structure within a warehouse location is always analogue. The small warehouse type has the lowest fixed costs and highest variable costs, and the large warehouse type the highest fixed costs and lowest variable costs.

Moreover, every customer is assigned to a delivery time category. For that classification, no data is available for which allocation is made arbitrarily. The number of delivery time categories is set to three, where t_1 has the shortest delivery time and t_3 the longest. Corresponding distances are listed in Table 3. Those maximum distances are set so that feasible solutions with a high service level are possible given the set of potential warehouse locations and customers. Short maximum distances would reduce the number of potential customers being supplied in time. For that reason, maximum distances are longer than one would expect for e.g. multiple deliveries per day. Allocation of customers to delivery time categories is made according to size of the city where the customer is located: repair shops in cities with more than 500,000 inhabitants belong to t_1, while t_2 contains customers in cities with a population between

100,000 and 500,000 and t_3 contains customers in cities with less than 100,000 inhabitants. This allocation gives 973 customers in t_1, 2,151 customers in t_2 and 6,874 customers in t_3.

	max. delivery time [h]	**max distance normal [km]**	**max distance express [km]**
t_1	5	400	500
t_2	7	600	700
t_3	9	800	900

Table 3: Input data for delivery time categories

If the model proposes not to supply a customer in time, this is done based on total 'connection' costs (depending on distance, transport volume and tariff). In order not to choose important but remote customers to be delivered later, it can be assumed that important customers are assigned delivery time category one (fastest delivery) and for those customers, service level s_{t_1} can be set quite high.

As previously discussed, before running the model, pre-processing is necessary to prepare the data for the model input. Firstly, subsets are to be defined in order to determine whether a warehouse can supply a customer within its delivery time requirement. Therefore, every customer is assigned one subset $J'(k)$ with warehouses in range to supply with LTL tariff, and another one with warehouses within the "express" perimeter $J''(k)$ around the customer. These subsets are based on the customers' delivery time category.

Secondly, it is necessary to merge transport costs from warehouse to customer and from supplier to warehouse, as described in chapter 4.2.

4.5 Model Verification

In order to verify the mathematical formulation of the problem, the mathematical solver *Gurobi* (version 7.5.2) is installed in Python (version 3.6). The algorithm starts with accessing data from an Excel file before actually computing the optimization. Thus, it is necessary to verify both calculation and import code. For that, a small test instance is created, consisting of ten customers, two delivery time categories and three potential warehouse locations (see Appendix 9.2 and 9.3). Suppliers are neglected here, as they only account for transport costs. These are constant for every warehouse-customer connection (LTL: 5.00 €, "express": 10.00 €). In addition, other parameters are chosen, so that complexity is low and results are foreseeable.

As expected, the pre-processing algorithm gives the warehouses in range to supply in time for every customer. For the current setup it is shown, that there are only two possible express connections in delivery time category t_1 and one in t_2. Subsequently, it is not possible to fulfil a service level of 100% and to supply all customers in delivery time category t_2 within the desired time.

So, if service level input is set to 100% for t_1 and 80% for t_2, model output is to open warehouses in Gothenburg and Bremen and to supply Stockholm, Berlin and Paris with express transport to reach the desired service level (see Figure 9). In that figure, customers supplied with LTL tariff are marked green, the ones supplied with the express tariff are marked orange and the customer in Madrid, not contributing to service level fulfillment, is marked red. Moreover, total network costs give the expected value of 265.10 €. The parameters capacity, transport costs and fixed costs are also tested under random scenarios.

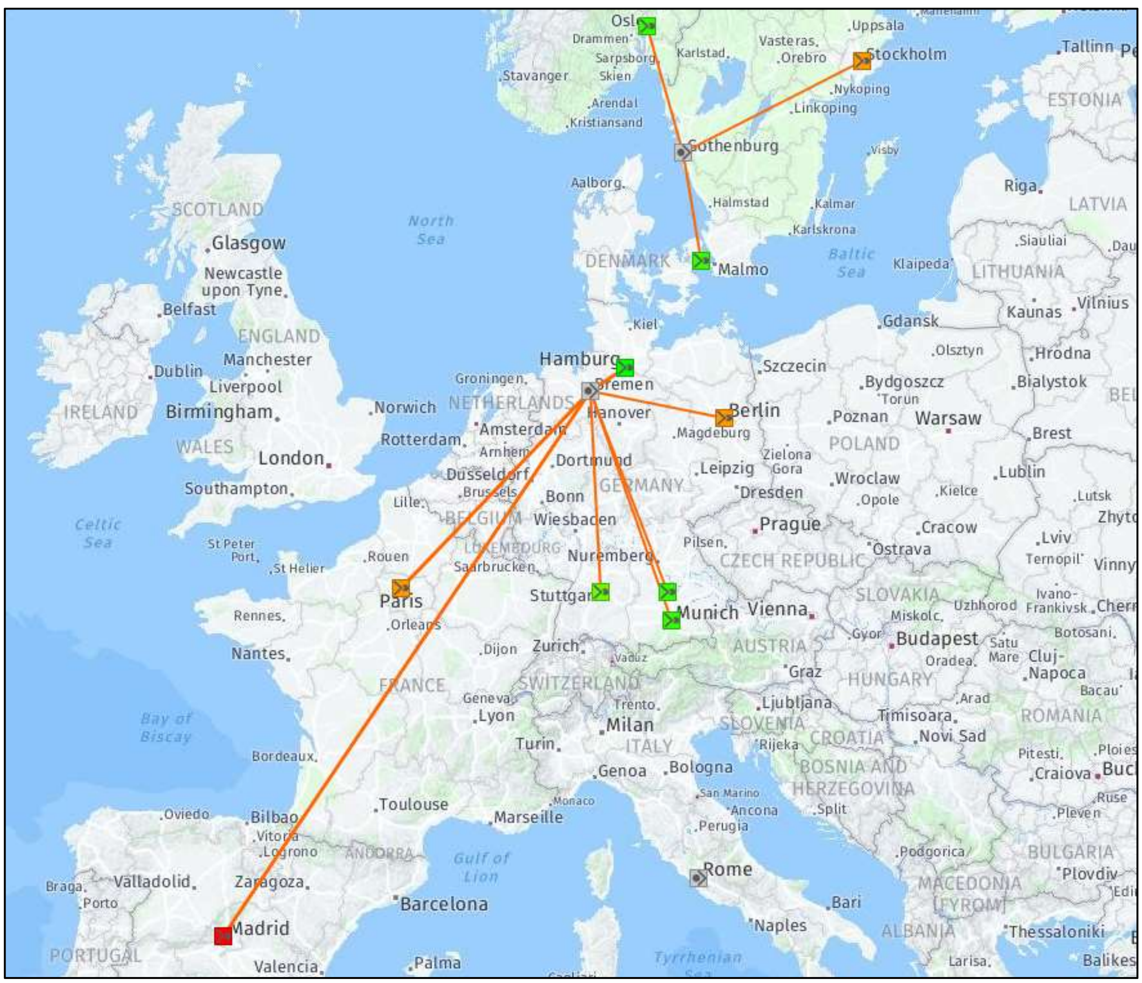

Figure 9: Verification visualization (from 4flow vista, transport planning software)

Consequently, this model is the partial answer to the research question. The model gives the desired output including number and warehouses to be opened, customers and tariffs allocation and total network costs. Further, it allows varying the service level per delivery time category, which refers to whether or not a customer is supplied within a certain period, while still supplying all of the customers. However, it does not allow the input of the actual network data as the problem size exceeds the maximum computable size. With regards to the research question, the model is lacking the appropriateness to be applied for the actual size of a European spare parts network. A heuristic solution technique is therefore developed to provide appropriateness in terms of solvability.

5 Solution Approach

This chapter identifies appropriate heuristic techniques and develops a solution approach for large instances. This solution is verified and converted into a feasible solution.

5.1 Heuristic Techniques

5.1.1 Lagrangian Relaxation

There is a large variety of heuristic techniques in literature. For that reason, several papers with heuristic solutions to location problems have been investigated. Focus has been on linear programs with a similar network structure, including binary decision variables or single sourcing constraints. Within this process, the minimization problem by Litvinchev and Espinosa (2012) has been identified as a reference solution, as it is also characterized by binary variables and similar constraints. To solve their two-stage mixed integer linear program, the authors used the Lagrangian relaxation to generate a solution close to the optimum with lower computation time. Further, it cannot be guaranteed that the relaxed constraints of the initial problem are fulfilled. For that reason, the authors heuristically converted the good solution into a feasible solution.

The Lagrangian heuristic's main characteristic is the elimination of constraints. Mostly, one to two constraints are 'relaxed' (Litvinchev and Espinosa, 2012). This means that constraints are attached to the objective function and multiplied with the so-called Lagrangian multipliers, so that not fulfilling the constraint would be penalized by increasing (in case of a minimization problem) the objective function value (Korte and Vygen, 2012). The following simplified example problem from Korte and Vygen (2012) shows the application of Lagrangian multipliers for a linear objective function.

$$\max c * x \tag{5.1}$$

subject to

$$Ax < b \tag{5.2}$$

$$Dx < e \tag{5.3}$$

Assuming that constraint 5.2 makes the problem much harder to solve, one would relax that constraint by use of the Lagrangian multiplier λ:

$$\max c * x + \lambda(b - Ax) \tag{5.4}$$

$$Dx < e \tag{5.5}$$

Subsequently, the following step is to identify the best value for the Lagrangian multiplier that penalizes not fulfilling the constraint in just the right way.

For the initial problem, two constraints are chosen to be relaxed (see Equation 5.6). In accordance with Litvinchev and Espinosa (2012), constraints are chosen in a way, that the original problem can be simplified into sub-problems, depending on as few variables as possible. In this case, constraints 4.2 and 4.5 are chosen, as the remaining constraints depend on j. Thus, λ_k is introduced as Lagrangian multiplier for customer k and μ_t as multiplier for delivery time category t. Therefore, Equation 5.6 needs to be simplified so that sub problems depending on j can be extracted. Therefore, brackets are eliminated in Equation 5.7. In the result, terms from the relaxed service level constraint are still dependent on customer k, which is why subsets of warehouses in range are converted to depend on j (Equations 5.9 and 5.10). It is further possible to define a sub problem for every $j - w$ combination. The resulting Equation 5.8 shows the remaining part of the Lagrangian objective function to be minimized. In other words, one optimization problem would be solved for every warehouse location-warehouse type combination.

$$\begin{aligned} \min L = {} & \sum_{j=1}^{q}\sum_{w=1}^{v} Y_{jw} f_{jw} + \sum_{j=1}^{q}\sum_{w=1}^{v} g_{jw} \sum_{k=1}^{r}\sum_{b=1}^{m} Z_{jkwb} l_k \\ & + \sum_{j=1}^{q}\sum_{k=1}^{r}\sum_{w=1}^{v}\sum_{b=1}^{m} Z_{jkwb} c_{jkb} + \sum_{k=1}^{r} \lambda_k \left(\left(\sum_{j=1}^{q}\sum_{w=1}^{v}\sum_{b=1}^{m} Z_{jkwb} \right) - 1 \right) \\ & + \sum_{t=1}^{n} \mu_t \left(s_t |K_t| - \sum_{k \in K_t} \sum_{j \in J'(k)} \sum_{w=1}^{v}\sum_{b=1}^{m} Z_{jkwb} - \sum_{k \in K_t} \sum_{j \in J''(k)} \sum_{w=1}^{v} Z_{jkwb_{expr}} \right) \end{aligned} \tag{5.6}$$

$$\begin{aligned} \min L = {} & \sum_{j=1}^{q}\sum_{w=1}^{v} Y_{jw} f_{jw} + \sum_{j=1}^{q}\sum_{w=1}^{v} g_{jw} \sum_{k=1}^{r}\sum_{b=1}^{m} Z_{jkwb} l_k + \sum_{j=1}^{q}\sum_{k=1}^{r}\sum_{w=1}^{v}\sum_{b=1}^{m} Z_{jkwb} c_{jkb} \\ & + \sum_{k=1}^{r}\sum_{j=1}^{q}\sum_{w=1}^{v}\sum_{b=1}^{m} \lambda_k Z_{jkwb} - \sum_{k=1}^{r} \lambda_k + \sum_{t=1}^{n} \mu_t s_t |K_t| - \sum_{t=1}^{n} \sum_{k \in K_t} \sum_{j \in J'(k)} \sum_{w=1}^{v}\sum_{b=1}^{m} \mu_t Z_{jkwb} \\ & - \sum_{t=1}^{n} \sum_{k \in K_t} \sum_{j \in J''(k)} \sum_{w=1}^{v} \mu_t Z_{jkwb_{expr}} \end{aligned} \tag{5.7}$$

$$\min L = \sum_{t=1}^{n} \mu_t s_t |K_t| - \sum_{k=1}^{r} \lambda_k + min\left\{\sum_{j=1}^{q}\sum_{w=1}^{v}\left(Y_{jw} f_{jw} + g_{jw}\sum_{k=1}^{r}\sum_{b=1}^{m} Z_{jkwb} l_k\right.\right.$$
$$\left.\left. + \sum_{k=1}^{r}\sum_{b=1}^{m} Z_{jkwb}\left(c_{jkb} + \lambda_k\right) - \sum_{t=1}^{n}\sum_{k\in J_t'^{-1}(j)}\sum_{b=1}^{m}\mu_t Z_{jkwb} - \sum_{t=1}^{n}\sum_{k\in J_t''^{-1}(j)}\mu_t Z_{jkwb_{expr}}\right)\right\} \quad 5.8$$

with

$$J_t'^{-1}(j) = \{k \in K_t : j \in J'(k)\} \quad 5.9$$

$$J_t''^{-1}(j) = \{k \in K_t : j \in J''(k)\} \quad 5.10$$

Simplification of the Lagrangian objective function gives one constant part and sub problems to be minimized for every *j-w* combination. This also follows Litvinchev and Espinosa (2012), who divided their Lagrangian objective function into several sub-problems of lower computational effort.

To determine the best value for a Lagrangian multiplier, the sub gradient method[1] is chosen in accordance with Litvinchev and Espinosa (2012). The sub gradient method is a technique to iteratively optimize the value of the Lagrangian multiplier. For that, derivatives with respect to the multipliers are determined:

$$\frac{\partial L(Y_j, Z_{jk}, \lambda_k, \mu_t)}{\partial \lambda_k} = \left(\sum_{j=1}^{q}\sum_{k=1}^{r}\sum_{w=1}^{v}\sum_{b=1}^{m} Z_{jkwb}\right) - 1 \quad 5.11$$

$$\frac{\partial L(Y_j, Z_{jk}, \lambda_k, \mu_t)}{\partial \mu_t} = s_t|K_t| - \sum_{k\in K_t}\sum_{j\in J'(k)}\sum_{w=1}^{v}\sum_{b=1}^{m} Z_{jkwb} - \sum_{k\in K_t}\sum_{j\in J''(k)}\sum_{w=1}^{v} Z_{jkwb_{expr}} \quad 5.12$$

Resulting from that, with step length r^i, the sub gradient method updates multipliers with every iteration step i:

$$\lambda_k^{i+1} = \lambda_k^i + r^i\left(\left(\sum_{j=1}^{m}\sum_{w=1}^{v}\sum_{b=1}^{m} Z_{jkwb}\right) - 1\right), \forall k \in K \quad 5.13$$

[1] For further techniques to determine the value of the Lagrangian multiplier, the reader is referred to Litvinchev and Espinosa (2012) and Maßmann (2006).

$$\mu_t^{i+1} = \mu_t^i + r^i\left(s_t|K_t| - \sum_{k\in K_t}\sum_{j\in J'(k)}\sum_{w=1}^{v}\sum_{b=1}^{m} Z_{jkwb} + \sum_{k\in K_t}\sum_{j\in J''(k)}\sum_{w=1}^{v} Z_{jkwb_{expr}}\right), \quad \forall t \in T \qquad 5.14$$

For the sub gradient method, it would be crucial to choose the right step length r (Wolsey, 1999). Wolsey (1999) presents three iterative methods to compute the best step length, which cannot be described in detail at this point for space reasons. For this thesis, the one that is described as best suited for practical use is applied:

$$\mu_k = \mu_0 p^k \qquad 5.15$$

Here, μ_k is the Lagrangian multiplier at iteration k, μ_0 is a large enough constant and p a parameter $p < 1$. Although μ_0 and p can be chosen individually, it is stated that the larger both values are, the more likely it would be to find a solution close to the optimum, since the step length remains relatively high and is less likely to converge at a local optimum. Thus, Equations 5.16 and 5.17 give the adjusted sub gradient method to the Lagrangian relaxation.

$$\lambda_k^{i+1} = \lambda_k^i + \mu_0 p^i\left(\left(\sum_{j=1}^{m}\sum_{w=1}^{v}\sum_{b=1}^{m} Z_{jkwb}\right) - 1\right), \forall k \in K \qquad 5.16$$

$$\mu_t^{i+1} = \mu_t^i + \mu_0 p^i\left(s_t|K_t| - \sum_{k\in K_t}\sum_{j\in J'(k)}\sum_{w=1}^{v}\sum_{b=1}^{m} Z_{jkwb} + \sum_{k\in K_t}\sum_{j\in J''(k)}\sum_{w=1}^{v} Z_{jkwb_{expr}}\right), \quad \forall t \in T \qquad 5.17$$

5.1.2 Knapsack Problem

In Litvinchev and Espinosa (2012), the use of Lagrangian multipliers and simplification of the Lagrangian objective function leads to two sub problems. Each of the problems then contains only one decision variable. Next, the authors fix one decision variable and instead of a location and allocation problem, only a network flow problem with continuous variables needs to be solved. In logistic terms, a warehouse is set as ‘opened’ and only allocation of customers is computed. Similar to Litvinchev and Espinosa (2012), here, the use of Lagrangian multipliers allows to simplify the problem into a knapsack problem, a standard allocation problem of low computational costs, for every warehouse location – warehouse type combination.

A knapsack problem can be illustrated by filling a bag with several items, where every item is assigned weight and value to the hiker. The problem is to fill the bag, so that the total value is maximized while fulfilling the bag's total weight constraint. Problems being solvable with this technique are therefore characterized by binary variables, like the present location-planning problem, that indicates whether or not allocation should be made rather than solving problems with continuous variables (Domschke and Drexl, 2007).

To apply the standard knapsack problem, decision variable Y_{jw} is set to one for every $j - w$ combination (see Equation 5.18), which leaves allocation of customers and transport tariffs.

$$\min L_{jw} = g\sum_{k=1}^{r}\sum_{b=1}^{m} Z_{kb} l_k + \sum_{k=1}^{r}\sum_{b=1}^{m} Z_{kb}\,(c_{kb} + \lambda_k) - \sum_{t=1}^{n}\sum_{k\in J_t^{\prime -1}(j)}\sum_{b=1}^{m} \mu_t Z_{kb} - \sum_{t=1}^{n}\sum_{k\in J_t^{\prime\prime -1}(j)} \mu_t Z_{kb_{expr}}\,, \forall j \in J, w \in W \tag{5.18}$$

subject to

$$\sum_{k=1}^{r}\sum_{b=1}^{m} Z_{kb} l_k \le h \tag{5.19}$$

$$Z_{kb} \in \{0,1\} \tag{5.20}$$

In this problem, the elimination of constraints allows extracting a knapsack problem, in which warehouse location j and warehouse type w correspond to the capacity constraint and demand of customer k corresponds to an item's weight e. Every customer k can be assigned to a warehouse with two possible transport tariffs. Thus, every customer-tariff combination is one possible item to allocate to the warehouse. Every item's weight e is subsequently defined by the customer's demand:

$$e(k,b) = l_k. \tag{5.21}$$

The value c however is more complex, as it considers all the costs that occur if a customer-tariff combination is allocated to a warehouse. This clearly includes demand-dependent variable warehouse costs and transport costs. Moreover, connection costs are also influenced by the distance between warehouse and customer. So, the Knapsack problem formulation (Equation 5.18) gives the connection costs for every $k - b$ pair (Equation 5.22). In addition, the Lagrangian multipliers are to be considered as they make sure that allocation is made in accordance with the initial problem's constraints: every customer needs to be allocated and the desired service level per delivery time category needs to be fulfilled.

$$-c(k,b) := \begin{cases} g * l_k + c_{kb} + \lambda_k - \mu_t & if\ k \in J_t^{-1'}(j), \forall b \\ g * l_k + c_{kb} + \lambda_k - \mu_t & if\ k \in J_t^{-1''}(j), b =\ express \\ g * l_k + c_{kb} + \lambda_k & else \end{cases} \quad 5.22$$

As the standard Knapsack problem aims to maximize the value, an adjustment needs to be made for the initial minimization problem by converting c into a negative value. The detailed procedure to identify the best solution from the Knapsack problem is described in Table 4.

	Task
1	set $\mu = 0$ and $\lambda = 0$
2	for 1,000 iterations:
3	for every warehouse location j
4	for every warehouse type w
5	compute $connection\ costs$ for every combination of customer k and transport tariff b with every $j - w$ pair
6	compute knapsack problem for $k - b$ pairs with $connection\ costs \geq 0$:
7	$capacity\ =\ warehouse\ capacity$ $weight\ =\ demand$ $value\ =\ connection\ costs$
8	save those $k - b$ pairs, that are assigned to $j - w$ pair
9	compute sum S of connection costs (value c) of all $k - b$ pairs assigned to jw
10	compute $A = S + f_{jw}$
11	find min A and corresponding w, $w^* = w$
12	if $A \leq 0$:
13	$Y_{jw^*} = 1$
14	for assigned $k - b$ pairs
15	$Z_{jkw^*b} = 1$
16	else: do not connect $j - w - k - b$
17	$Y_{jw} = 0$ $Z_{jkwb} = 0, \forall\ w, k - b$ pairs
18	update Lagrangian multipliers
19	go to 3

Table 4: Pseudo code for heuristic solution

To conclude, the knapsack problem, solved for every $j - w$ combination, provides a heuristic solution on how to allocate customers and their transport tariffs to warehouses, supported by Lagrangian multipliers, which make sure, that allocation is done in accordance with the relaxed constraints. The heuristic then chooses the warehouse locations and warehouse types to open based on the Knapsack allocation. However, the only constraint being fulfilled for sure is the capacity constraint. The others, due to the solution's heuristic nature, are expected to be mostly fulfilled, but no guarantee is given.

5.2 Verification of Heuristic Solution

This heuristic is implemented in Python and runs on the same test instance as in chapter 4.6. To find the best step length, several scenarios with different values for μ_0 and p are run. The best heuristic solution compared to the optimal solution is found with $\mu_0 = 10{,}000$ and $p = 0.7$ for both λ_k and μ_t. As expected, this heuristic gives a solution close to the optimal solution, as provided in chapter 4.6. The resulting warehouses to be opened and the allocation of customers are mostly identical. However, the customer in Madrid is not allocated at all. This is due to the special circumstance that that customer cannot be supplied in time by any warehouse. Taking a closer look at the Lagrangian multipliers, it can be seen that λ_{Madrid} has the lowest value of all λ_k. While all λ_k are positive, λ_{Madrid} is closer to zero. Only a negative λ_k leads to an increase of the objective function value and the heuristic would then allocate Madrid to a warehouse, as connection costs would be lower than the penalty costs caused by the Lagrangian multiplier. The special case of Madrid exists because allocation of Madrid to a warehouse does not affect μ_t. Having this in mind, one can choose input parameters for the update of the Lagrangian multiplier (Equation 5.16) accordingly.

Moreover, the Lagrangian objective function gives a negative value being influenced by the large values of the multipliers. Hence, the objective function value cannot be used as an indicator for total network costs.

5.3 Feasibility

Lastly, the good solution from the heuristic needs to be converted into a feasible solution. As constraints 4.2 and 4.5 might not be fulfilled, the correct allocation of customers and tariffs needs to be done. In the heuristic solutions, customers can therefore either not be allocated at all (like the fictive customer in Madrid in the test instance), or can be allocated multiple times or the service level can be not reached.

For that reason, if there is a customer not being allocated, it is checked whether the delivery time category's service level of the customer is already fulfilled or not. If so, the customer is assigned to the warehouse with remaining capacity and the lowest connection costs or simply any warehouse that provides the service to the customer. If the service level is not yet reached, the customer is allocated so that its allocation contributes to the service level fulfilment and if necessary, it is also possible to open a warehouse that has not yet been opened. An equivalent procedure is used to eliminate customers being allocated twice or more times.

This allocation of remaining customers is however not optimal nor is a holistic logic behind it. The implementation partly iterates over customers and chooses the connection with the first match. Consequently, it is assumed that the better the heuristic solution, the better the resulting feasible solution since the room for arbitrariness is lower. The verification model is too small to test this hypothesis, since never several warehouses can supply a customer in time.

Finally, the feasible solution gives the total network costs by transferring it into the objective function of the initial problem (see Equation4.1). The feasible solution is generated for every iteration, which is why the best solution obtained from the heuristic is the one with the lowest objective function value.

6 Analysis

In the following, results of model application on the European spare parts network are presented and discussed. Again, this is done including aspects both related to model development and to practical applicability for actual spare parts networks.

6.1 Model Performance

6.1.1 Results

Returning to the initial challenge described in the introduction, different scenarios are compared to identify the costs of an increasing service level in order to determine an appropriate service level from the OEM's point of view. Such sensitivity analysis would be a common method to test a model's robustness against changes (Werners, 2006; Domschke and Drexl, 2007).

The number of scenarios is set in a way that the effect of an increase in service level in every single delivery time category can be illustrated (Table 5). For that, a random scenario (scenario 1) serves as a reference scenario, from which several changes are measured. Scenarios 2-4 increase the service level of one delivery time category by 4% respectively, while scenario 5 increases all three service levels by 4%. The reference scenario is designed under the assumption that important clients (i.e. in large cities) belong to delivery time category t_1, which is why its service level is comparably higher. Throughout the different scenarios, it is found that the exact predefined service level is not always met. When converting the knapsack solution into a feasible solution, allocation can be done to a warehouse with lowest connection costs. This can contribute to service level fulfilment even if it is already reached. Therefore, the service level can be higher than required, but not lower.

delivery time category	**scenario** service level				
	1	**2**	**3**	**4**	**5**
t_1	95%	99%	95%	95%	99%
t_2	90%	90%	94%	90%	94%
t_3	85%	85%	85%	89%	89%

Table 5: Chosen scenarios for sensitivity analysis

Due to confidentiality, results can neither be visualized nor presented in detail, but the solution network's core characteristics are shown in the following. For scenario 1, the model suggests to open 37 warehouses, all of them being of the large type. This is due to the high capacity and the significant share of warehouse variable costs. The latter are the lowest for the large warehouse type. Total network costs are 1,467,217,983 € and the share of express transports is 0.0021%.

For scenarios 2-5, the model gives an increase in total network costs of 0.2%, 1.5%, 2.8% and 4% respectively (see Table 6). It is found that the service level increase in scenario 2 can be handled with additional express transports instead of opening an additional warehouse. Interestingly, for scenario 3 a solution is found with one warehouse less than in the reference solution. However, higher network costs arise through more express transports in that solution. In scenario 4, where the service level for customers in delivery time category t_3, to which belong repair shops in small cities, is raised, the number of warehouses increases significantly. Assuming that the potential warehouse locations are chosen based on their accessibility to highly populated areas of Europe, this explains that it would be much more expensive to offer a higher service to customers in small cities, as this means a higher weighting of remote areas. Finally, scenario 5 shows the highest increase in number of warehouses and total costs. With regards to warehouse types, the number of small warehouses increases with the higher service level for customers in t_3: in scenario 3, nine of the warehouses are small while it is three in scenario 5.

criteria	**scenario**				
	1	**2**	**3**	**4**	**5**
increase in network costs (compared to scenario 1)	-	0.2%	1.5%	2.8%	4%
number of opened warehouses	37	37	36	46	52
share of express transports	0.0021%	0.0031%	0.0054%	0.0048%	0.004%

Table 6: Comparison of results from all scenarios

To sum up, model application on the network goes mostly along with general literature findings on warehouse centralization: the higher the service level or warehouse decentralization, the higher the total network costs. However, the monetary effect in this network is not as high as literature indicates which is caused by the network's cost structure of relatively low warehouse fixed costs (a high increase in number of warehouses leads to only 4% increase in total network costs). In this model, both faster transports and additional warehouse are measures to provide a higher service level. However, the input data clearly shapes the results by problem-specific distances chosen for delivery time categories and the impact of warehouse fixed costs.

6.1.1 Computation time

Throughout model development and testing, computation time has always been a core issue. For the different scenarios, computation time has been around 8 hours for 300 iterations, which has been the cancellation criterion. On the test data, the number of iterations could be set up to 10,000, but the large data set does not allow this computational effort. The model has been run on a computer with 8GB RAM and an Intel(R) Core(TM) i7 processor with 3.4 GHz.

Especially the first iterations are of high computational costs, since the Lagrangian multipliers not yet influence allocation of customers towards the optimum. Therefore, i.e. the knapsack problem is solved for all customer-tariff pairs (see Table 4). Since all multipliers start at zero in the first iteration, the model neither opens any warehouses nor allocates any customers. The conversion into a feasible solution is therefore time-consuming as it involves all the customers in the network. Thus, with the Lagrangian multipliers becoming more and more meaningful, computation time per iteration decreases. Nevertheless, computational resources are crucial for applying the model.

6.1.2 Optimality

Despite findings of expected trends, this does not allow any statement with regards to optimality of the solution as no reference value is available. In other words, it cannot be said how far the heuristically generated feasible solution is away from the optimal solution, or if there is any network structure at all leading to lower total network costs. Testing of the heuristic solution has been conducted in chapter 5.2, but the range of possible network setups is too limited to draw any conclusion on quality of the solution. Instead, for every customer, there is in most cases only one warehouse which is able to provide delivery in time. Consequently, a part of the European network is chosen, which is just large enough to allow different network setups while still being solvable with the commercial solver. After some attempts, it is found that the actual network with the number of customers limited to 100 can be solved with the commercial solver. To increase reliability of comparing the heuristic solution with the optimal solution, this is done three times with different sets of customers (but each consisting of 100 customers). Running these comparisons on scenario 1 gives total network costs being 14%,15% and 17% higher than the respective optimal objective function value. It can therefore be assumed that the solution provided by the heuristic approach lies within such range from the optimal solution.

When applying the heuristic on the large data set, long computation time reduces the number of possible runs to investigate appropriate values for the start value μ_0 and step length p^i of the Lagrangian multipliers. Compared to the optimality test, the heuristic solution consisting of Lagrangian relaxation and the knapsack problem approach performs relatively poor, which can be seen in the number of customers allocated by the heuristic. For that reason, the results found in Table 6 are to a larger extend depending on the conversion of the heuristic solution into a feasible solution. It can thus be expected that additional runs varying μ_0 and p^i same as an increase in the number of iterations would lead to better results.

6.2 A Spare Parts Network Perspective

As previously mentioned, the solution network is primary driven by the maximum distances set for warehouse-customer connections, since those determine whether the allocation can contribute to service level fulfillment. Consequently, referring to the Toyota network, the model would reflect regional warehouses (from where delivery tours to customers start) rather than the central warehouses. For the same reason, the high degree of detail in transport costs and country-dependent warehouse fixed costs does contribute to model validity and trustworthiness of costs, but is not explicitly reflected in the network structure (e.g. no concentration of warehouses in Eastern Europe due to lower labor costs). Furthermore, chosen maximum distances together with number and location of potential warehouses shape the overall network structure to a large extend by determining coverage (as presented by Figure 6). Here, the modeled network also aims to combine warehouse characteristics that are divided among the different warehouse levels in the Toyota network: low maximum distances to customers (cross-docking warehouse) and high coverage throughout Europe (regional/ central warehouses).

Concerning the chosen warehouse types, the approach to open many 'large' warehouses can be questioned for practical use. In the Toyota network, proximity to customers is reached by cross-docking warehouses with low storage capacities. However, the model considers a reduced number of warehouse levels, which is why the warehouses need to meet requirements on customer proximity, capacity and consolidation. Consequently, the chosen warehouse types can serve as indicators for planning of several warehouse levels.

To sum up, the model quantifies the effect of varying service levels for different delivery time categories. Category t_3 has been identified as the subset of customers with the deepest impact on network structure and costs, as it contains a high share of remote customers in small cities. Hence, one could draw the cost-based conclusion to either lower the service level for that delivery time category or to increase the promised delivery time.

7 Conclusion

7.1 Findings

A model has been implemented in Python that gives number, location and type of warehouses to be opened and customer allocation in a spare parts network while applying different service levels for different delivery time categories. Moreover, the model considers two transport tariffs differing in delivery speed and costs so that the trade-off between higher transport costs and an additional warehouse is modeled. Other main assumptions are customers being supplied from one warehouse and articles only being supplied from one supplier. Thus, the number of possible flows in the network is reduced to improve the applicability for large amounts of data, i.e. spare parts typical product variety is aggregated. A heuristic solution approach has been developed to apply the model for a European automotive spare parts network which is characterized by a large number of customers (repair shops). The heuristic solution combines Lagrangian relaxation with a knapsack problem approach.

Application of this model has quantified costs of an increasing service level for different delivery time categories and proved general applicability for such warehouse centralization problems. With regards to appropriateness, simplifications with the deepest impact are most likely the assumption of one warehouse level and direct transports from warehouses to customers. Furthermore, application is obviously not limited to spare parts networks, but also allows modelling of similar networks characterized by, among others, single sourcing constraints, time sensitivity and several transport tariffs or modes.

Computation time for the given network has been about 8 hours and the knapsack solution has not delivered the expected result within this timeframe. Hence, the use of more sophisticated aggregation techniques could lower computation time and thus allow more iterations to obtain a solution closer to the optimum.

This thesis fills a research gap in including service level variation for several subsets of nodes in location planning. In terms of general modelling aspects, the use of road distances and real transport tariffs as well as the possibility to model large networks are strengths of the developed model.

7.2 Outlook

The computation time of several hours calls for the investigation of more sophisticated optimization techniques. In this model, products are aggregated as well as transports on supplier – warehouse and warehouse - supplier side. As next steps, it could be investigated how i.e. customers or transport tariffs can be aggregated in an efficient way. This would be especially beneficial since, – in this network – customers are quite similar in demand size and significantly contribute to complexity with their large number.

Based on model development, it is possible to follow-up in different directions. With regards to the model as such, some points have been neglected in this thesis but might be worth a closer look. Firstly, it might be interesting to consider overnight delivery more detailed. As in the case of Toyota, milk runs to repair shops begin early in the morning and thus transports to cross-docking warehouses occur overnight. Thus, considering Toyota's network structure, distances could be significantly influenced by this constraint. Mainly total distance - if arrival is scheduled for the next morning -, but also route-specific constraints (caused by noise) for nighttime transports might influence route length.

Moreover, the expert interview showed that car manufacturers would distinguish between high- and low-runners when it comes to part classification. While low-runners would be typically stored centrally, high-runners are mostly stored in regional warehouses. The consideration of inventory balancing, as suggested by Melo et al. (2009), goes into a similar direction and is an appropriate tool to cope with sporadic demand. These aspects might be possible to model without an underlying network structure of several warehouse levels.

Looking into future trends that might affect spare parts networks, 3D printing is named as potential influencing factor (VTT, 2017). With increasing applicability, it could be investigated how this affects transport of spare parts, for example what type of parts would be predominantly transported. It is possible that mostly small parts are printed at site, so that the focus of spare parts transport shifts to rather bulky parts. This would then be reflected in the relation between transport and warehouse costs, and subsequently new conditions would emerge when answering the question of warehouse centralization.

8 References

Altiparmak, F., Gen, M., Lin, L. and Paksoy, T. (2006), "A Genetic Algorithm Approach for Multi-Objective Optimization of Supply Chain Networks", *Computers & Industrial Engineering*, Vol. 51, pp. 196–215.

Badi, I., Stevic, Z. and Novarlic, B. (2017), "Emergency Medical Service Location Problem: A Case Study in Misurata, Libya", *Transport & Logistics - The International Journal*, Vol. 17 No. 43, pp. 30–35.

Bagherpoor, Z., Parhizi, S., Hoseininia, N. and Yaghin, R. (2009), "Warehouse Location Problem", in Zanjirani Farahani, R. and Hekmatfar, M. (Eds.), *Facility Location: Concepts, Models, Algorithms and Case Studies*, *Contributions to management science*, Physica, Heidelberg, pp. 295–314.

Balcik, B. and Beamon, B. (2008), "Facility Location in Humanitarian Relief", *International Journal of Logistics*, Vol. 11 No. 2, pp. 101–121.

Barkawi, K., Baader, A. and Montanus, S. (2006), *Erfolgreich mit After Sales Services: Geschaftsstrategien Fur Servicemanagement Und Ersatzteillogistik*, Springer, Dordrecht.

Charikar, M., Khuller, S., Mount, D. and Narasimhan, G. (Eds.) (2001), *Algorithms for Facility Location Problems with Outliers*.

Chen, M.-C., Hsu, C.-M. and Chen, S.-W. (2006), "Optimizing Joint Maintenance and Stock Provisioning Policy for a Multi-Echelon Spare Part Logistics Network", *Journal of the Chinese Institue of Industrial Engineers*, Vol. 23 No. 4, pp. 289–302.

Collis, J. and Hussey, R. (2014), *Business Research: A Practical Guide for Undergraduate & Postgraduate Students,* 4th ed., Palgrave Macmillan, Basingstoke, Hampshire.

Cordeau, J.-F., Pasin, F. and Solomon, M.M. (2006), "An integrated model for logistics network design", *Annals of Operations Research*, Vol. 144 No. 1, pp. 59–82.

Costantino, N. and Pellegrino, R. (2010), "Choosing between single and multiple sourcing based on supplier default risk. A real options approach", *Journal of Purchasing and Supply Management*, Vol. 16 No. 1, pp. 27–40.

Crainic, T. (2000), "Service Network Design in Freight Transportation", *European Journal of Operational Research*, Vol. 122 No. 2, pp. 272–288.

Current, J., Daskin, M. and Schilling, D. (2010), "Discrete Network Location Models", in Drezner, Z. and Hamacher, H. (Eds.), *Facility Location: Applications and Theory*, 1st ed., Springer, Berlin, pp. 81–118.

Domschke, W. and Drexl, A. (Eds.) (2007), *Einführung in Operations Research,* 7th ed., Springer, Berlin.

Drezner, Z., Klamroth, K., Schöbel, A. and Wesolowsky, G. (2010), "The Weber Problem", in Drezner, Z. and Hamacher, H. (Eds.), *Facility Location: Applications and Theory*, 1st ed., Springer, Berlin, pp. 1–36.

Du, F. and Evans, G. (2008), "A Bi-Objective Reverse Logistics Network Analysis for Post-Sale Service", *Computers & Industrial Engineering*, Vol. 35, pp. 2617–2634.

Fallah, H., NaimiSadigh, A. and Aslanzadeh, M. (2009), "Covering Problem", in Zanjirani Farahani, R. and Hekmatfar, M. (Eds.), *Facility Location: Concepts, Models, Algorithms and Case Studies*, *Contributions to management science*, Physica, Heidelberg, pp. 145–176.

Flodén, J., Meers, D. and Macharis, C. (2017), "Modelling of Intermodal Systems", in Monios, J. and Bergqvist, R. (Eds.), *Intermodal freight transport & logistics*, CRC Press, Boca Raton, FL, pp. 213–238.

Hassanzadeh, A., Mohseninezhad, L., Tirdad, A., Dadgostari, F. and Zolfagharinia, H. (2009), "Location-Routing Problem", in Zanjirani Farahani, R. and Hekmatfar, M. (Eds.), *Facility Location: Concepts, Models, Algorithms and Case Studies*, *Contributions to management science*, Physica, Heidelberg, pp. 395–417.

Holmqvist, M. and Pessi, K. (2006), "Agility Through Scenario Development and Continuous Implementation. A Global Aftermarket Logistics Case", *European Journal of Information Systems*, Vol. 15 No. 2, pp. 146–158.

Huiskonen, J. (2001), "Maintenance Spare Parts Logistics. Special Characteristics and Strategic Choices", *International Journal of Production Economics*, Vol. 71 No. 1-3, pp. 125–133.

Janová, J. (2012), "Crop planning optimization model. The validation and verification processes", *Central European Journal of Operations Research*, Vol. 20 No. 3, pp. 451–462.

Kleijnen, J.P.C. (1995), "Verification and validation of simulation models", *European Journal of Operational Research*, Vol. 82 No. 1, pp. 145–162.

Klein, R. and Scholl, A. (2012), *Planung und Entscheidung: Konzepte, Modelle und Methoden einer Modernen Betriebswirtschaftlichen Entscheidungsanalyse,* 2nd ed., Franz Vahlen, München.

Korte, B. and Vygen, J. (2012), *Kombinatorische Optimierung: Theorie und Algorithmen,* 2nd ed., Springer, Berlin.

Lapierre, S., Ruiz, A. and Soriano, P. (2004), “Designing Distribution Networks. Formulations and Solution Heuristic”, *Transportation Science*, Vol. 38 No. 2, pp. 174–187.

Li, F. (2015), “Study on the Spare Part Logistics Management Model Facing the Automotive Aftermarket”, available at: https://www.atlantis-press.com/proceedings/meici-15/25286 (accessed 21 February 2018).

Litvinchev, I. and Espinosa, E. (Eds.) (2012), *Solving the Two-Stage Capacitated Facility Location Problem by the Lagrangian Heuristic: Proceedings*.

Maßmann, M. (2006), *Kapazitierte Stochastisch-Dynamische Facility-Location-Planung: Modellierung und Lösung eines strategischen Standortentscheidungsproblems bei unsicherer Nachfrage,* 1st ed., Deutscher Universitäts-Verlag, Wiesbaden.

Melo, M., Nickel, S. and Saldanha-da-Gama, F. (2009), “Facility Location and Supply Chain Management. A Review”, *European Journal of Operational Research*, Vol. 196, pp. 401–412.

Papathanasiou, J. and Manos, B. (2007), “An Approximation Algorithm for the Location of Dairy Enterprises under Time Constraints”, *European Journal of Operational Research*, Vol. 182, pp. 1479–1487.

Pulido, R., Muñoz, J. and Gazmuri, P. (2015), “A Continuous Approximation Model for Locating Warehouses and Designing Physical and Timely Dsitribution Strategies for Home Delivery”, *EURO Journal on Transportation and Logistics*, Vol. 4 No. 4, pp. 399–419.

Rath, S. and Gutjahr, W. (2014), “A Math-Heuristic for the Warehouse Location–Routing Problem in Disaster Relief”, *Computers & Operations Research*, Vol. 42, pp. 25–39.

ReVelle, C.S., Eiselt, H.A. and Daskin, M.S. (2008), “A Bibliography for Some Fundamental Problem Categories in Discrete Location Science”, *European Journal of Operational Research*, Vol. 184 No. 3, pp. 817–848.

Rottkemper, B., Fischer, K. and Blecken, A. (2012), “A transshipment model for distribution and inventory relocation under uncertainty in humanitarian operations”, *Socio-Economic Planning Sciences*, Vol. 46 No. 1, pp. 98–109.

Schittekat, P. and Sörensen, K. (2009), “Supporting 3PL Decisions in the Automotive Industry by Generating Diverse Solutions to a Large-Scale Location-Routing Problem”, *Operations Research*, Vol. 57 No. 5, pp. 1058–1067.

Schmidt, G. and Wilhelm, W.E. (2000), “Strategic, tactical and operational decisions in multi-national logistics networks. A review and discussion of modelling issues”, *International Journal of Production Research*, Vol. 38 No. 7, pp. 1501–1523.

Scholz-Reiter, B., Heger, J., Meinecke, C. and Bergmann, J. (2012), "Integration of demand forecasts in ABC-XYZ analysis. Practical investigation at an industrial company", *International Journal of Productivity and Performance Management*, Vol. 61 No. 4, pp. 445–451.

Simchi-Levi, D., Kaminsky, P. and Simchi-Levi, E. (2011), *Designing and Managing the Supply Chain: Concepts, Strategies, and Case Studies,* 3rd ed., McGraw-Hill, Boston.

Tavakkoli-Moghaddam, R., Makui, A. and Mazloomi, Z. (2010), "A new integrated mathematical model for a bi-objective multi-depot location-routing problem solved by a multi-objective scatter search algorithm", *Journal of Manufacturing Systems*, Vol. 29 No. 2, pp. 111–119.

Thacker, B.H., Doebling, S.W., Hemez, F.M., Anderson, M.C., Pepin, J.E. and Rodriguez, E.A. (2004), *Concepts of Model Verification and Validation*.

Toregas, C., Swain, R., ReVelle, C. and Bergman, L. (1971), "The Location of Emergency Service Facilities", *Operations Research*, Vol. 19 No. 6, pp. 1363–1373.

VTT (2017), "Spare parts go digital. A boost for the industrial spare parts business", Technical Research Centre of Finland Ltd, available at: https://www.vttresearch.com/media/news/spare-parts-go-digital (accessed 20 May 2018).

Wagner, S., Jönke, R. and Eisingerich, A. (2012), "A Strategic Framework for Spare Parts Logistics", *California Management Review*, Vol. 54 No. 4, pp. 69–92.

Wagner, S. and Lindemann, E. (2008), "A-Case Study-Based Analysis of Spare Parts Management in the Engineering Industry", *Production Planning & Control*, Vol. 19 No. 4, pp. 397–407.

Werners, B. (2006), *Grundlagen des Operations Research: Mit Aufgaben und Lösungen*, Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg.

Wolsey, L.A. (1999), *Integer programming, Wiley-Interscience series in discrete mathematics and optimization*, J. Wiley, New York.

Xu, G. and Xu, J. (2009), "An Improved Approximation Algorithm for Uncapacitated Facility Location Problem with Penalties", *Journal of Combinatorial Optimization*, Vol. 17 No. 4, pp. 424–436.

Yaobao, Z., Ping, H. and Shu, Y. (2013), "An Improved Particle Swarm Optimization for the Automobile Spare Part Warehouse Location Problem", *Mathematical Problems in Engineering*.

Zanjirani Farahani, R. and Hekmatfar, M. (Eds.) (2009), *Facility Location: Concepts, Models, Algorithms and Case Studies, Contributions to management science*, Physica, Heidelberg.

9 Appendix

9.1 Interview Guide

- Which automotive spare part networks did you get to know? How did their network design differ (network members and their number, stages, etc.)?

- Which do you think are operational reasons for customers only being supplied from one warehouse?

- Can you make a statement about purchasing a spare part from only one supplier (single sourcing)? How common is it and if so, under which circumstances?

- Which classifications are made in automotive spare parts networks (regarding articles, customers, etc.)?

- Can you relate ABC / XYZ classifications or others to network structure?

- What do you think are relevant future trends in spare parts networks?

9.2 Distance Matrix [km]

Customer Location	**Warehouse Locations**		
	Gothenburg	Bremen	Rome
Oslo	293	1046	2495
Hamburg	639	127	1659
Madrid	2793	2052	1957
Munich	1288	773	915
Stockholm	471	1231	2545
Copenhagen	317	587	1901
Ingolstadt	1214	697	998
Stuttgart	1286	641	1071
Berlin	744	406	1502
Paris	1527	801	1422

9.3 Test Data

Customer Location	Demand [kg/year]	Delivery Time Category
Oslo	1	t_1
Hamburg	1	t_1
Madrid	1	t_2
Munich	1	t_2
Stockholm	1	t_1
Copenhagen	1	t_1
Ingolstadt	1	t_2
Stuttgart	1	t_2
Berlin	1	t_1
Paris	1	t_2

Delivery Time Category	max. distance for LTL transport [km]	max. distance for express transport [km]
$t_1 = 5\ h$	400	500
$t_2 = 10\ h$	800	900

Warehouse Locations	Warehouse Types	Warehouse Characteristics	
Gothenburg	w_1	capacity	10 [kg/year]
		fixed costs	100,00[€]
		variable costs	0.01[€/kg]
	w_2	capacity	10 [kg/year]
		fixed costs	100,00 [€]
		variable costs	0.01 [€/kg]
Bremen	w_1	capacity	10 [kg/year]
		fixed costs	100,00 [€]
		variable costs	0.01 [€/kg]
	w_2	capacity	10 [kg/year]
		fixed costs	100,00 [€]
		variable costs	0.01 [€/kg]
Rome	w_1	capacity	10 [kg/year]
		fixed costs	100,00 [€]
		variable costs	0.01 [€/kg]