MASTER

Network design for next day deliveries of spare parts at IBM

Koppes, M.M.

Award date:
2008

Eindhoven, April 2008

Network Design for Next Day Deliveries of Spare Parts at IBM

by
Myrthe M. Koppes

BSc Industrial Engineering and Management — Eindhoven University of Technology 2005
Student identity number 0531191

in partial fulfilment of the requirements for the degree of

Master of Science
in Operations Management and Logistics

Supervisors:
Prof. dr. ir. G.J. van Houtum TU/e, OPAC
Dr. T. Tan, TU/e, OPAC
Ir. I.C. Reijnen, TU/e, OPAC
Drs. J.J Felten MBA, IBM SPO
J. Smit MBA, IBM SPO
Ir. M.A. Kuipers MBA, IBM SPO

TUE. Department of Technology Management
Series Master Theses Operations Management and Logistics

ARW 2008 OML

Subject headings: Spare parts, Inventory control, Multi-locations, Lateral transshipments, Location/allocation, Design network, “Service logistiek”

I. Abstract

This master thesis describes an integrated design approach for spare parts network design of next day orders at IBM. The integrated design approach for network design is new and unique because it integrates inventory decisions and specifically takes lateral transshipments into account. This network design therefore not only contain the determination of the optimal number and locations of hubs but also the determination of inventory levels and the method to dispatch the required parts from facilities to the customers in need. Using an integrated design approach without lateral transshipments and a decoupled design approach as benchmark, the integrated approach with lateral transshipments significantly influences the solution and costs of the network design. Therefore, it is recommended to further use and extent research about this integrated design approach including lateral transshipments for network design of spare parts.

II. Management summary

This report is the result of a Master thesis project at IBM. IBM offers next to advanced information systems, high quality after-sales service. This high quality after-sales service is required to satisfy the **Service Level Agreement** (SLA) that defines the level of service the customer requires. A differentiation in the level of service is made by IBM into same, next, next business and second business day orders. As the names already imply, the available time span differs per service level.
This study has been conducted at the organizational department Service Parts Organizations (SPO) that manages the spare parts. A logistics network capable of serving customers in a time-responsive manner is crucial for the delivery of high quality after-sales service. Due to changes in the business environment, a review of the current network design is desirable. Hereby the network design is defined as the number and location of stocking points to fulfil customer demand. The assignment has therefore been formulated as:

"Develop a tool that provides insight into the optimal network design for next day orders"

A network is defined as optimal if costs are minimal and the predefined service constraints are met.

The network design has been focused on next day orders because, due to the design of fault-tolerant systems, the importance of this type of orders increases. Furthermore, the original design of the current network focused on same and not on next day orders. Recall and repair processes and replenishment of same day locations have been considered out of scope. The geographical scope is Europe (for an exact overview see ***Appendix B***). Moreover, the network has been designed by a **green field approach.** So, current locations or regulations are not constraining the network design.

The structure of the current network of IBM has been characterized as a multi-echelon network. An **echelon** is a layer in the network that consists of stocking points with the same function. Next day orders are fulfilled by either the Central Buffer (CB) or the Regional or Local Stocking Hub (RSH/LSH). This depends on the parts availability at the RSH and LSHs and the coverage of the CB. The difference between a LSH and RSH is that a LSH covers one country and a RSH covers more than one country. In general suppliers replenish the CB and in turn the CB replenishes the RSH/LSHs. The locations used to fulfil next day demand are defined as 'hubs'.

In literature 2 network design approaches were identified: the integrated design approach and the decoupled design approach. The **integrated design approach** simultaneously determines the traditionally strategic decisions and the inventory decision. Traditionally strategic decisions are the number and locations of hubs and the assignment of customers to the hubs. The **decoupled design approach** first determines these traditionally decisions and uses the output as a fixed input for traditionally tactical decisions like inventory.
Literature review showed that an integrated design approach compared to a decoupled design approach in a spare parts environment leads to a cost reduction and a different solution in terms of number and locations of hubs. This can be explained by the strong interaction between the inventory and the traditionally defined strategic questions and the relative high inventory costs in the spare parts business.
Furthermore, literature on spare parts showed that the use of lateral transshipments leads to significant inventory reduction. **Lateral transshipments** are defined as the provisioning of a part by a stocking point to a customer of another stocking point that is out of stock. So, it is a back up option to fulfil demand in response to stock outs. Related to network design, we expected that the inclusion of lateral transshipments into the design approach leads to a different solution in the network design.

Based on literature, we expected that the inclusion of both the inventory decision and lateral transshipments into a network design approach leads, compared to current network design approaches, to another solution in network design. Therefore the network has been designed based on an integrated design approach with the inclusion of lateral transshipments. It should be noted that this approach is not found in literature and can therefore be defined as new and unique. Furthermore, two other design approaches have been addressed to determine the value of the new developed design approach:

I. Decoupled design approach
II. Integrated design approach without lateral transshipments
III. Integrated design approach with lateral transshipments

Based on the design approach III, 4 design questions have been identified:

1. What is the optimal number of hubs?
2. What is the optimal location of hubs?
3. What method should be used to dispatch the part from hub to the customer?
4. What are the optimal inventory levels of parts at the selected hubs?

To answer the design questions, a mathematical model have been presented that models a two echelon, single-item, multiple and discrete location problem with service constraints. The locations of the hubs are discrete because the geographical area is divided into 400 customer groups. The centre of each customer group represents a potential location of a hub. The service level is defined as the percentage of orders that is fulfilled by the network. As the service contract guarantees service for every failure, a single item model is suitable. The aim of the model is to minimize the total network costs. The total network costs include transportation and warehouse costs but exclude opening costs of a hub.

With a **scenario analysis** the first two design questions have been addressed. A scenario analysis is defined as the process of analyzing the best suitable network by considering alternative possible outcomes of scenarios. By picking a sensible set of scenarios, a near-to-optimal solution can be found. Each scenario consists of the number and location of hubs. For design approach II and III, the number of hubs has been varied from 5-20 hubs and the location of the hubs has been determined by common sense. For design approach I the output of the Warehouse Site Planner (WSP) have been used as input for scenarios. The WSP is an IBM toolbox that minimizes the amount of kilometres that have to be crossed from the hubs to fulfil demand of customers. As it does not take inventory into account, this toolbox is based on the decoupled design approach.

In design approach III, a part is dispatched from a hub to a customer by a regular, lateral or an emergency (trans)shipment. The order and hub that is selected to fulfil demand is based on the transportation costs and constrained by the available time. In case the first choice hub does have inventory, a so called **regular shipment** is used. Otherwise, another hub can fulfil the demand which is defined as a **lateral transshipment**. If none of the hubs in the array has stock to fulfil demand, the demand is fulfilled by a central source, defined as an **emergency shipment**. In design approach II lateral transshipments are excluded from the demand fulfilment options.

The output of the three previous described design questions have been used as input in the sub model for the heuristic optimization of base stock levels of the hubs, the fourth design question, with a **greedy algorithm**.

Based on a business case consisting of nearly 30,000 SKUs and around 400 potential locations answers to the design questions have been retrieved that include the comparison between the different design approaches.

1. Number of hubs

The total network costs of design approach III are negatively related to the number of hubs. Moreover, the total costs function is decreasing over the whole domain. This is caused by the exclusion of opening costs of hubs and the inclusion of lateral transshipments. This implies that the network design is more robust for expansion of customer area and less sensitive for the design of LSPs. Furthermore, the solution (in number of hubs) of design approaches I and II differ from the solution of design approach III.

2. Locations of hubs

The total costs in design approach III are not significantly influenced by the locations of the hubs. Furthermore, using the WSP based on a decoupled approach to determine the locations, does not lead to minimal costs. Therefore, design approach III is preferable over design approach I.

3. Method to dispatch parts from hubs to customers

Results from design approach III show that the number of lateral and emergency (trans)shipments is respectively positively and negatively correlated to the number of hubs. To evaluate the value of the inclusion of lateral transshipments into the design approach, a comparison is made between design approach II and III. Analysis showed that both the solution and costs differ significantly: The solution of design approach III showed significantly lower (inventory) costs while reaching a higher service level. So, lateral transshipments significantly reduces costs while significantly increasing service level

4. Inventory of hubs

Results of the analysis showed that the inventory costs strongly influence the total network costs. Furthermore, the value of the inclusion of lateral transshipments into the design approach is supported by the significant lower inventory investment of design approach III compared to design approach II.

Implementing the proposed network design, the difference in structure between the current and proposed new network design should be acknowledged (illustrated in ***Figure 1***). The new network structure requires intensive collaboration between hubs.

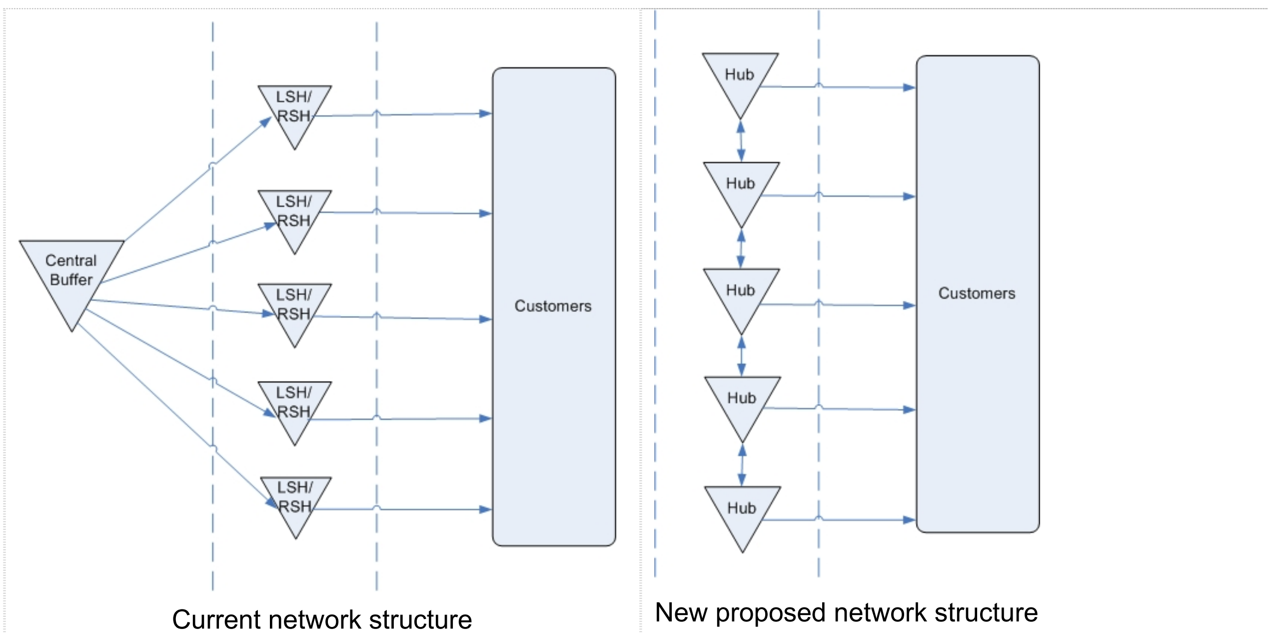


Figure 1: Difference between current and new proposed network structure

Based on the above results, we come to the following main recommendations:

- We advise IBM to use the integrated design approach with lateral transshipments for their network designs
- We recommend to investigate the integration of the network designs of same and next day orders
- We advise to include lateral transshipments in inventory planning

III. Preface

This master thesis is the final piece of my study Industrial Engineering and Management. In this master thesis I focus on service logistics. Service logistics is a rapid growing area of interest for both companies and research. In the Service Logistics Forum (SLF) knowledge and skills about service logistics are shared. Furthermore, it tightens the relation between research and practice. IBM is one of the companies that participate in the SLF and proposed a very challenging project because neither practice nor research did give clear guidelines how to handle the research question. The opportunity to contribute both to IBM spare business as the literature did complicate but also motivated me to let the project succeed.

I would like to thank Geert-Jan van Houtum for his supervision during my master. His enthusiasm and knowledge made sure that our meetings were pleasant and really helped me moving on during the project. Due to the innovative character of the project for both business and literature, I worked closely with Ingrid Reijnen, AIO at Eindhoven University of Technology. I really appreciate the time and effort she spent on adjusting the greedy algorithm and supervising me. Next, I would like to thank Koos Huibers for helping me programming the model into Delphi. Finally, I would also like to thank Tarkan Tan, my second supervisor for his feedback on my reports.

At IBM, I would like to thank Jorn Felten for being my daily supervisor during the project. Furthermore, I would like to thank Joost Smit and Michiel Kuipers for proposing the challenging project in the SLF research program. Next, I would like to thank Dennis Faber for providing feedback and helping my gathering the data. Moreover, I would like to thank all colleagues of IBM SPO. I really appreciate the time and effort you spend on helping me!

Finally, I would like to thank my family and friends for their support during my study. You made sure that I enjoyed this time!

Myrthe Koppes
Eindhoven, April 2008

IV. Table of contents

V. Introduction

Due to the competitive value of after-sales service for their advanced information systems, IBM defined this service as one of their focal areas. After-sales service is defined as providing necessary service engineers and replacement of parts to existing, geographically dispersed customers when they experience any problems with a purchased product (Candas and Kutanoglu, 2006). In this thesis we will concentrate on the management of spare parts. The spare parts business is in the literature recognized as a special research area caused by the special combination of high service requirements, high value of parts and low demand rates. Therefore, streamlining the logistic system of spare parts is complicated, expensive and strategically important (Huiskonen, 2001).

Due to the worldwide, geographically dispersed customers and the quick service they demand, it is necessary that IBM has several stocking points in the customer area.
In this thesis a **network design study** is conducted that is defined as designing the network with stocking points to fulfil customer demand. Based on the used integrated approach this contains determining the number and locations of stocking locations but also deciding inventory levels and determining the method to dispatch the required parts from facilities to the customers in need.

In preparation of this master thesis project two preliminary studies have been conducted. First of all, a literature study to identify existing models for network design. As the amount of available literature on network design for spare parts is limited, this review gives an overview of all types of models for network design. Hereby the value of the different types of models for a spare parts environment is assessed. The second preliminary study is the research proposal that contains an outline of the current spare parts business, the formulation of the assignment and the scope of the assignment.

The two preliminary studies are input for the first chapter. This chapter will give an introduction of the context in which the study is conducted and formulates the research assignment. Furthermore, the approach to handle the network design study is defined based on the output of the literature study.
The remaining part of the report can be split into two parts. In the first part the model and method that are used to fulfil the research assignment are described. The application of the first part to the business environment of IBM is discussed in the second part. This part furthermore includes extensions to the basic model, an implementation plan, conclusions and recommendations. An extensive outline of these two parts of the report is given at the end of ***Chapter 1***.
It should be noted that all used concepts are bolded in the text and can be look up at the concept list at the end of the report. Furthermore, a list of variables and abbreviations are included.

1. Research assignment

This chapter gives an overview of the research assignment. First, an introduction about the company at which the assignment is conducted, is presented. The second section zooms in on the business environment of the assignment, service logistics. Based on trends in this business environment, the assignment is formulated. The fourth section describes the scope of the assignment. Next, the current network design is presented. To determine the approach to handle the assignment, a literature framework is presented in section six. Based on this literature framework, four design questions are defined that have to be answered to design the new network. Finally, the position of this study within the company and an outline of the remainder of the report is given.

1.1 Company description

Nearly 100 years ago the International Business Machines corporation (IBM) was founded through a merger of three separate corporations. Although the electronic computer (the product IBM is currently associated with) was not developed yet, all products of IBM were and still are designed and developed for the purpose of information sharing. Products contain technologies to record, process, communicate, store and retrieve information from its first scales, tabulators and clock to today's powerful computers and vast global networks.

The mission of IBM is *`strive to lead in the creation, development and manufacturing of the industry's most advanced information technologies, including computer systems, software, networking systems, storage devices and microelectronic and to translate these advanced technologies into value for our customers through our professional solutions and services businesses worldwide'*. Translating this mission into the current business strategy generates 91.4 billion dollars a year which ranked IBM on the 10th place in the Fortune 500 in 2006.

IBM is an international organization that has more than 355,000 employees worldwide and serves customers in 200 different countries (Figures 2006, website). IBM's customers include many different kinds of enterprises: from sole proprietorships to the world's largest organizations.

1.2 Service logistics

IBM offers, next to advanced information systems, after sales service for these systems. The product of the purchase of this service is a **Service Level Agreement (SLA)** which states the level of service IBM has affirmed to its customer. IBM offers a wide range of service levels ranging from 2 hours to 3 days. From a marketing viewpoint, a categorization can be made into:

1. Same day orders
2. Next day orders
3. Second business day orders
4. Second next business day orders

As the names already suggest, the categories indicate the time span in which the service needs to be delivered. The outcome of the service should be an up and running system.
These SLAs can be arranged for all types of information systems (for an overview of all product types see Koppes, 2007b). However, in general the more advanced systems are linked to higher service levels. This is caused by the fact that customers who require high advanced systems, rely on these systems for their core processes. Therefore, uptime is crucial for their business processes. This causes the request for extremely quick service.

The service process consists of three main stages. The first stage encompasses the detection of a failure. The detection of the failure is reported by a customer engineer or by the customer itself which results in an

order. In case the failure is caused by a failed part, the part has to be replaced by a new part. Otherwise, the system can be fixed by the customer engineer and the service process stops at this stage.
The second stage consists of the delivery of the spare part to the customer. Finally, in the third stage the machine is repaired by replacing the old failed part by the new delivered part.

As illustrated by the outline of the service process, in case the failure is caused by a failed part, the availability of spare parts is crucial. To make sure that the spare parts are available at the right time and location at minimal costs, IBM has a dedicated department to manage the spare parts: the Service Parts Organization (SPO). The SPO department located in the Netherlands is responsible for the region Europe, Middle-East and Africa (EMEA). This department manages a product portfolio of more than 5000 types of machines. These machines together have 200,000 different part numbers of which 40,000 are defined as active parts. Parts are defined 'active' if they have had demand within the past year. The demand rates of these active parts are relatively low and vary strongly: 50 percent of these spare parts have a demand rate of a maximum of 3 demands per year while only 1 percent has a demand rate of more than 1000 demands/year.

1.3 Assignment

In this section the assignment is formulated. This assignment is based on the developments in the business environment of IBM and the current network design. The developments in the business environment consist of the approaching end of the contract with Logistic Service Provider(s) (LSP) and environmental changes. The main environmental changes are (for details see ***Appendix A***):

1) Shift from same day to next day orders due to fault-tolerant design of systems
2) Emerging markets like Russia and Eastern Europe
3) After-sales service is becoming a strategic asset in companies' offerings (Saccani et al., 2007)

During the '90s the original network was designed upon the outcome of an IBM developed toolbox called the Warehouse Site Planner (WSP). In the last years only pieces of the network design have been adjusted. As the business environment changed during these years, it is interesting to investigate the network design again to see if and how the network can be improved. Furthermore, the drawbacks of the available toolbox of IBM (for details see research proposal (Koppes, 2007b) raise the request for another method to improve the current network.

Improvements in the network design are particularly expected for next day orders because the WSP (used for the original design) only focused on same day orders. Furthermore, the importance of next day orders increases because of the design of fault tolerant systems (one of the environmental changes, see ***Appendix A***).

The direct cause of the initialization of the project is the approaching end of the contract with the current Logistic Service Provider. The fact that IBM can review offers of all Logistic Service Provider provides more freedom for the design of the network because current restrictions and constraints of the Logistic Service Provider do not have to be taken into account.

The assignment can be formulated as follows:

"Develop a tool that provides insight in the optimal network design for next day orders."

A network design is optimal for IBM if supply chain activities are executed at minimal costs and service levels stated in the service contracts are achieved.

1.4 Scope

This section defines the scope of the study. It is based on several aspects which are each described in a separate sub section. The last sub section captures the scope by formulating what is defined as being "out of scope".

1.4.1 Geographical

IBM serves customers over the whole world but this assignment is executed for the department IBM SPO located in the Netherlands that is responsible for EMEA. This contains Europe, Middle East (including Russia) and Africa. However, to avoid infeasible solutions due to legal restrictions, this assignment will focus on Europe. Furthermore, the geographical scope is made consistent with other network design related projects of IBM. An illustration of the geographical scope is shown in ***Appendix B***).

1.4.2 Functional

The functional aspect in the scope definition refers to the product types that are considered in the study. As formulated in the assignment, the study focuses on next day orders to anticipate to the shift from same day to next day orders due to the design of fault-tolerant systems. Next day orders can be placed for both low and high end products in the product portfolio. Therefore, all product types are considered in the study.

1.4.3 Green field approach

The network design will be developed from a **green field perspective**. This means that current locations are not constraining the network design. The philosophy of decoupling the analysis of the current design is that it will generate more insight and innovative solutions. Furthermore, this approach is suitable because within a short time period the contracts with the current LSPs will expire so the design is not constrained by the conditions of the current LSPs.

1.4.4 Processes

Several processes like order fulfilment, repair, recall and replenishment can be identified within the network. In this study only next day order fulfilment and replenishment from suppliers to hubs are considered. Replenishment of same day stores, repair processes and recall processes are left out of scope because these processes can be decoupled from the considered next day order fulfilment process.

1.4.5 Out of scope

To clearly summarize the main constraints of the assignment, an overview of the issues that are left out of scope is presented:

- Same day contracts, orders or stores
- Current locations
- Replenishment of same day stores
- Inbound logistics
- Reverse logistics

1.5 Current network

Before we describe the design method to create the new network structure, we will discuss the current network design. This section illustrates the structure of the current network design and the next day order fulfilment and replenishment process. Furthermore, the control mechanisms of the network are presented.

1.5.1 Structure of network

Figure 2 illustrates the structure of the current network design. The structure of the network can be characterized by identifying echelons, illustrated by the vertical lines in ***Figure 2***. An **echelon** is a layer in the distribution network that consists of stocking points with the same function. Due to the focus on next day orders, we will neglect the fourth echelon that consists of locations for same day orders. Details about same day orders in the current network design can be found in Koppes (2007b).

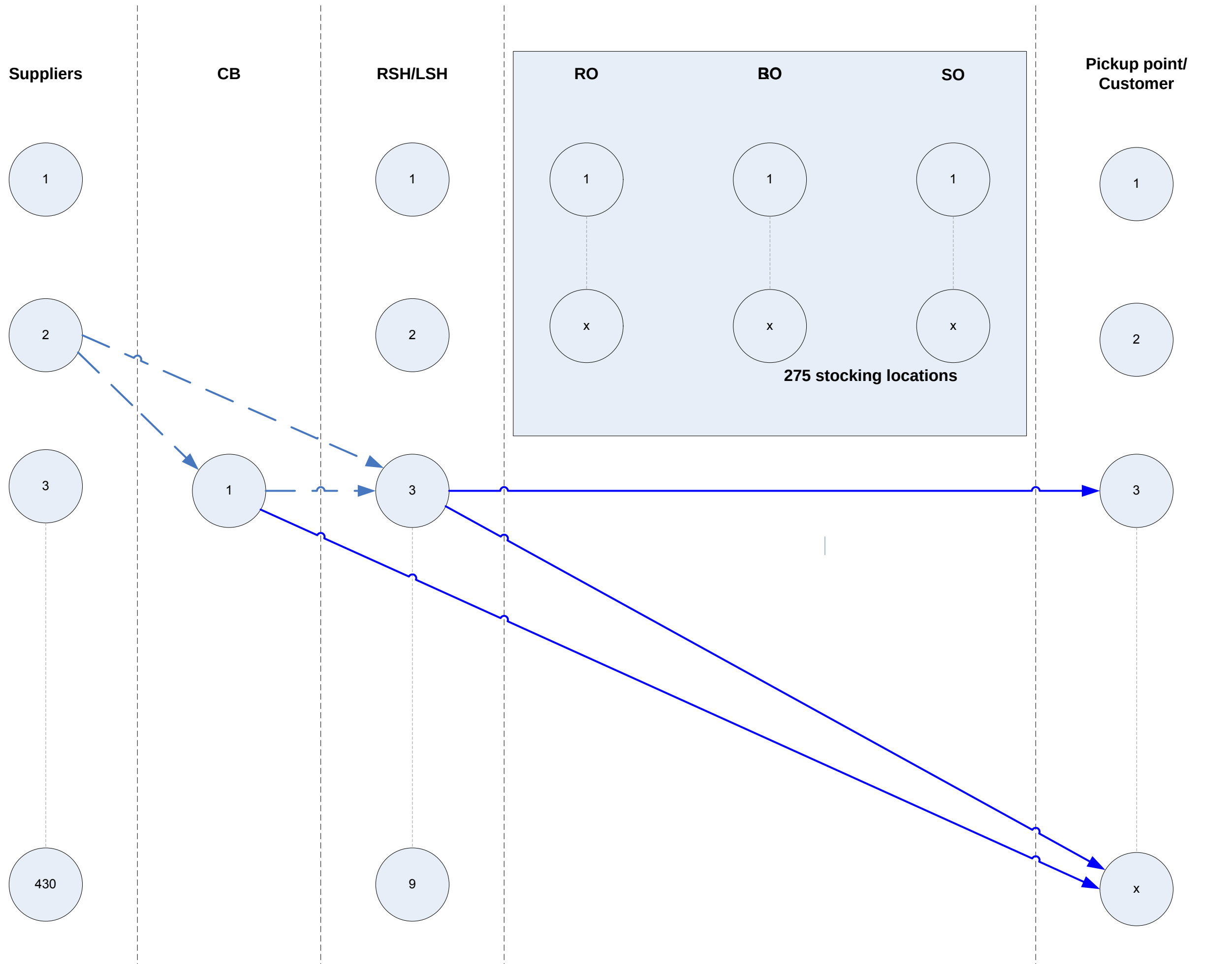


Figure 2: Next day order fulfilment (solid lines) and replenishment (dashed lines) process in current network design

1.5.2 Next day order fulfilment

Illustrated with the solid lines, ***Figure 2*** shows that next day orders can be fulfilled by either a shipment from the Central Buffer (CB), or the third echelon that consists of Regional Stocking Hub (RSH) and Local Stocking Hubs (LSH). The difference between a RSH and a LSH is related to the number of countries they cover: a RSH covers several countries and a LSH is dedicated to one country. Most countries in Europe have their own central stocking point and therefore most countries have an LSH.

Whether a next day order is fulfilled from the Central Buffer or from a RSH/LSH depends on the stock availability and the coverage of the Central Buffer. The Central Buffer is currently located in the Netherlands and covers the demand in the Netherlands and the nearby countries Germany, Belgium and Luxembourg. Customers in Portugal can also be reached within time because the current Logistic Service

Provider uses, due to a cheap airline connection, an airplane to reach Portugal. Furthermore, the Central Buffer does cover parts of Great-Britain, Austria and France. France is not geographical partly covered like Great-Britain and Austria but it is partly covered in terms of type of orders: next day sunrise orders are fulfilled from the LSH but all other orders with a less tied time constraint are supplied from the Central Buffer. In case an LSH or RSH should fulfil the demand of an order but is out of stock, the demand is also fulfilled by the Central Buffer.
As the Central Buffer, Regional and Local Stocking Hubs have the same function in the order fulfilment process of next day orders namely fulfilment of the orders, these stocking points are defined as ‘hub’ in this thesis.

To guarantee the service level of the same day customers, next day orders are never fulfilled by the locations that are dedicated to the fulfilment of same day orders (echelon 4). Furthermore, shipments from these stocking locations are more expensive, because in most cases a general transport line is not available, so everything has to be transported with an expensive transport mode, the taxi.

We can conclude that in the next day order fulfilment process not only time constraints but also costs influence the decision from which stocking point the order is fulfilled. Therefore, next day orders are fulfilled from one of the hubs depending on the coverage of the Central Buffer and the availability of the parts at the Local and Regional Stocking Hubs.

1.5.3 Replenishment of hubs

Figure 2 shows that the Central Buffer acts in the replenishment process both as receiver and supplier. It receives parts from almost all suppliers (99%) and supplies the parts to the Regional Stocking Hub or Local Stocking Hub (RSH/LSH). The remaining 1% of the suppliers replenishes the Regional or Local Stocking Hubs (RSH/LSH) directly, because in this case the part is only required at one specific Regional or Location Stocking Hubs (RSH/LSH) due to the local orientation of the product.

1.5.4 Control mechanisms

The current network is controlled by three mechanisms: 1) inventory policy; 2) performance measure; and 3) Control Tower Concept.

Inventory

IBM SPO controls the inventory of the hubs by reviewing the inventory daily and ordering parts if the inventory is below a so called ´critical stock value´ with a fixed order quantity. This inventory control policy can be defined as an (r, s, Q) policy with *r* representing the continuous review (every day), *s* the critical stock value and *Q* equal to the fixed order quantity. Because the inventory is not checked at every moment in time, it could also be argued that this control policy uses a discrete instead of continuous review period. However, as the inventories are checked each time the system updates these values in combination with the relatively low demand rate of parts, it can be seen as continuous review.

Performance measure

To evaluate the performance of the system and to track customer satisfaction, IBM uses two service measures. First, the Parts Availability Level (PAL) is used. **PAL** indicates in how many cases a part request on a stock location is fulfilled (expressed in %). The PAL is measured on stock location, country and Central Buffer level. As this measure does not reflect the customer satisfaction correctly (for details see Koppes, 2007b), another service measure is used in all countries since the start of 2008. This service measure is the Parts Delivery Time (PDT). **PDT** indicates the percentage of customer deliveries that is delivered on time. The start of the time measurement is the call of the customer and the time stops if the requested part is delivered at the customers´ requested place.

Control Tower Concept
The **Control Tower Concept** consists of outsourced business processes due to the fact that these business processes are not seen as a core competence by management. IBM defined three business processes in the Control Tower Concept:

1. The physical processes distribution and warehousing
2. Repair and warranty redemption
3. Operational planning and order execution

1.6 Theoretical framework

To get insight into the different approaches that can be used for network design, we will take a look at a review of the literature that is available in the field of network design for spare parts. An extensive review on network design for spare parts can be found in preliminary work (Koppes, 2007a). This section starts with a description of the difference between a decoupled and an integrated approach. As the integrated approach that includes inventory decisions shows benefits, the second subsection gives an overview of literature on inventory planning in a spare parts environment.

1.6.1 Integrated vs. decoupled design approach

In the literature spare parts network design management is a new and growing research area. Therefore, only a few studies are executed that investigate network design for spare parts. A distinction in network design studies can be made into studies with a decoupled and studies with a integrated design approach. The difference between these approaches is related to the number and traditionally defined levels of decisions they include. A distinction in levels of decision is made between strategic, tactical and operational decisions. Traditionally, strategic defined questions that encompasses network design are:

1. How many stocking point are needed?
2. Where to locate stocking points?
3. How to allocate customers to stocking points?

The last question is called the allocation decision and determines which stocking point supplies which customer. Traditionally, tactical defined decisions are for example inventory levels and routing of shipments (Candas and Kutanoglu, 2006).
The **decoupled design approach** is defined as first making the strategic decisions of the network design and using this outcome as a fixed input for optimizing the tactical issues. The **integrated design approach** is defined as simultaneously making the traditionally strategic decisions of network design and the tactical inventory decision. Therefore, the integrated network design encompasses:

1. How many stocking point are needed?
2. Where to locate stocking points?
3. How to allocate customers to stocking points?
4. What are the optimal inventory levels of stocking points?

The reason to include the inventory decision into the network design is a strong interaction between the four decisions. Given the time constraint in the spare parts business, it is essential to meet the required service level. This requires not only that customers have to be located close enough to a stocking location but also that the stocking locations do have a part available if demand occurs. The availability of parts is defined as the product fill rate capability and is influenced by the inventory levels (Mason et al., 2004). Therefore, not only the location but also the inventory levels of stocking points determine the performance of the network design.

Available literature shows that an integrated design approach leads to cost reductions (Candas, 2005; Candas and Kutanoglu, 2006; Jeet and Kutanoglu, 2005; Mason et al., 2004; Shen and Daskin, 2005). Jeet and Kutanoglu (2005) showed that this difference is caused by the existence of a service time constraint. Candas and Kutanoglu (2006) showed that the cost reduction using an integrated instead of a decoupled approach increases when the holding costs, lead times, level of service or variability of demand increases. As ±50% of the total costs of the spare parts business of IBM are warehousing costs and 33% of the total costs direct holding costs, it is expected that the integrated approach will benefit over the decoupled approach. The direct holding costs are expected to increase even more due to the trends component miniaturization and increasing costs of diagnosing product failure (Cohen et al, 1996).

Next to the cost reduction, Candas and Kutanoglu (2006) showed that the integrated design approach compared to the decoupled approach led to a substantial difference in the solution in terms of the number and location of stocking points. The decoupled approach selected less stocking points in their optimal solution than the integrated approach.

The results show that using an integrated approach creates value to the network design. As the integrated network design includes the inventory decision, an overview of available spare parts literature on inventory is presented in the next subsection.

1.6.2 Inventory planning in spare parts environment

Determining the inventory levels is complicated in a spare parts environment. This is caused by the characteristics of spare parts. Huiskonen (2001) reflects these characteristics in four identified control characteristics of spare parts: criticality, specificity, demand pattern and value of part. The criticality of spare part items is high which leads to high service requirements. As penalty costs are often high and clearly stated in the service contract, great pressure is put on attaining high service levels. Besides, the demand for spare parts may be extremely sporadic and therefore difficult to forecast and the prices of individual parts may be very high.

Next to the characteristics, there are also several trends that increase the complexity of the inventory planning of spare parts. The first trend is the decrease in product life cycles and in turn an increasing introduction speed of new product. The second trend is the increasing variability of products which causes a reduction in installed base. Both a shorter life cycle and lower installed bases decrease the availability of customer information. This in turn makes it hard to forecast demand and therefore complicates the determination of the optimal inventory levels.

To cope with this difficulty in determining inventory levels for spare parts, specific research has been done in this area. Several studies (Alfredsson and Verrijdt (1999); Grahovac and (2001); Kranenburg; 2006; Wong, et al., 2005b; Wong et al., 2006) show that the use of lateral transshipments leads to significant reductions in inventory costs. **Lateral transshipments** are defined as the provisioning of a part by a stocking point to a customer of another stocking point that is out of stock. So, it is a back up option to fulfil demand in response to stock outs. Lateral transshipments can be seen as a form of pooling. Physically, there are multiple stock points, but they have access to each other’s inventory when needed. This means that the inventory at multiple stock points should be optimized jointly (Kranenburg, 2006).

Lateral transshipments lead to a reduction in inventory costs due to the fact that in general the costs of lateral transshipments are much lower than the downtime costs (Wong, et al., 2005). The use of lateral transshipments is especially beneficial for spare parts because short product life cycles and part obsolescence make holding stock expensive and risky (Cohen, 1997). Furthermore, the high value of spare parts creates high inventory costs and therefore sharing of inventory can reduce the inventory costs significantly.

Next to the inventory reduction and therefore effect on the inventory and in turn total network costs, we expect that lateral transshipments also influence the solution of the network design. First of all, because the extent to which lateral transshipments can be used, depend on the number and location of hubs. Furthermore, the expectation is that an extra hub does not have to lead to more inventory because the inventory in case of lateral transshipments is calculated for all hubs simultaneously. This causes that the safety stock per hub can be lower and therefore, the amount of stock in the added hub can also be lower.

As in network design the trade off is made between transportation costs and warehouse costs, the solution can differ if the inventory levels are calculated simultaneously instead of separately.

As we expect that, based on literature review, the use of an integrated design approach and inclusion of lateral transshipments leads to other solutions in network design, these two issues will be used in this network design study. This is caused by the strong interaction between the strategic network design decisions and the inventory decision and by the fact that inventory costs are large part of the total network costs. It should be noted that this approach is not found in literature and can therefore be defined as unique.

1.7 Design questions

In this section we will see how the network design study will be approached. The **network design study** can be defined as designing the network with stocking points to fulfil customer demand. Based on the integrated approach, this requires determining the number and locations of stocking locations but also deciding inventory levels and determining the method to dispatch the required parts from hubs to the customers in need (Candas and Kutanoglu, 2006). The method to dispatch the required parts from hubs to the customers is affected by the use of lateral transshipments. The use of lateral transshipments makes it possible that the demand of a customer is not always fulfilled by the same stocking point. In case the first choice stocking point does not have inventory, another stocking point can fulfil the demand which is defined as a lateral transshipment. Therefore, the amount of lateral transshipments depends on the inventory levels of the hubs. The order and possibility of hubs to fulfil customer demand by a lateral transshipment is a design issue that has to be determined. This question will be captured in the decision which method will be used to dispatch the required parts from hubs to the customers.
The four design questions that encompass the integrated network design study can be defined as:

- What is the optimal number of hubs?
- What is the optimal location of hubs?
- What method should be used to dispatch the part from hub to the customer?
- What are the optimal inventory levels of parts at the selected hubs?

As these design questions are strategic and consider a long time period, this study will design the network on a high abstraction level. This means that details defined as coordination decisions by Silver et al. (1998) such as vehicle routing, choice of transportation capacity and warehouse design are not taken into account. Therefore, answers to follow-up and more detailed questions should be derived from and positioned within the framework of answers given to the above research questions. The answers of the abovementioned design questions thus function as constraints of the solution space of the lower level coordination questions.

1.8 Position of study

This study is part of the project 'network design' that is initiated by a department of SPO EMEA. The project 'network design' also contains another project that uses the Warehouse Site Planner tool to investigate the network design. This project uses a decoupled approach to determine the best suited network design for IBM. Furthermore, the Warehouse Site Planner tool takes into account more detailed issues from practice like a route planner to determine the exact time and distance between two points. So it addresses the above mentioned coordination decisions.
This study is also related to the IBM project Horizon 2010 that investigates the selection of Logistics Service Providers and handles the contracting phase. This project requires insight into optimal network design and therefore this study can be considered as input to this project.

1.9 Outline of report

This section describes the structure of the remaining chapters in this report. These chapters can be categorized into 2 parts. The first part describes the model and method of the integrated network design approach. In the second part this will be applied to the business environment of IBM. Below, an outline is presented of the chapters that fall respectively in the first and second part.

Part 1
Based on the approach defined in this chapter, ***Chapter 2*** presents a basic model that includes all four design questions. Due to the fact that it is from a mathematically view too complicated to solve all four design question simultaneously, this basic model is split into a main and sub model with each two design questions.

Chapter 3 shows the methodologies that are used to solve the main and sub model and answer the related design questions. A solution of the main model that consists of the design questions the number and location of hubs is found with a scenario analysis. A scenario consists of the number and locations of hubs. The input for the number of hubs is retrieved randomly or by managerial insight. The locations of the hubs can be determined by logical sense or with the Warehouse Site Planner tool. Furthermore, this chapter presents greedy algorithm that determines the inventory of the hubs. Based on these inventory levels, the design question how the parts are dispatched from the hubs to the customers can be answered.

Part 2
Chapter 4 presents the results of the integrated approach in combination with the inclusion of lateral transshipments in the planning of inventory. It provides insights into the best suited network design based on the investigated influence of the number and location of hubs. Furthermore, it shows a comparison in total network costs and solution between an integrated approach including and excluding lateral transshipments.

Chapter 5 describes suggestions to extend the basic model to reflect the IBM case better. The proposed extensions of the model are 1) differentiation of service target levels and 2) the inclusion of an air hub into the basic model.

Chapter 6 gives a short overview of implementation issues of the recommended network design. It first defines the main differences between the current and recommended network design. Next, stakeholders and their resistance to changes are described. Based on this analysis, an intervention strategy is presented. Moreover, some practical issues like the time frame of the implementation, are addressed.

Finally, ***Chapter 7*** discusses the conclusions and recommendations of this network design study. A distinction in recommendations is made between recommendations for IBM and recommendations for further scientific research.

It should be noted that all used concepts are bolded in the text and can be looked up in the concept list at the end of the report. Furthermore, a list of variables and abbreviations is included.

2. Model

This chapter describes the model that will be used to answer the four design questions formulated in the previous chapter. The basic model that is presented in the first section integrates all four design questions. The main assumptions are given in section 2. Mathematically seen, it is too complex to handle the four design questions simultaneously. Therefore, section 3 describes a main and a sub model and the interaction that exists between these models. Finally, section 4 contains a conclusion that gives an overview of the models that are presented in this chapter.

2.1 Basic model

Before the design questions can be handled in the main and sub model, the basic model presents components that are used in both models. In the network, customers generate failure based demand that hubs have to fulfil. The model aggregates the demand of customers to customer groups and uses these customer groups to define the possible locations of the hubs. The creation of customer groups is described in subsection 1 and the possible location of hubs is described in subsection 2. The third subsection presents the method that is used to calculate the distance between two points, for example a customer group and a hub. Next, the modelling of demand and the demand fulfilment process is explained. Subsection 6 and 7 describe respectively the inventory control policy and the network costs that are considered in the model. Finally, the service level constraint is added to the model and an overview of the model formulation including notation is given.

2.1.1 Customer groups

IBM has customers with installed bases in the field who generate failure based demand. As the number of IBM customers is high and analysis of the order data shows a wide spread of customers over the whole region, demand is aggregated to the customer group level (for used customer data see ***Appendix E***). The criterion for belonging to a specific customer group is the geographical location of the customer. Two methods, depending on the availability of data, are used to create these customer groups.

The general method to create customer groups is called the 'square' method. The starting point of the general 'square' method is the conversion of the postal codes of the customers that are available in the order data, into coordinates. Based on the maximum and minimum of the latitude and longitude coordinates (illustrated in ***Figure 3***), each country can be approached with a square. The rectangulars are created per country because the postal code system differs per country. Next, the country square is split into equal sized rectangulars with a maximum size of 150 km in latitude and longitude. The numbers in these rectangulars shown in ***Figure 3*** correspond to the customer group number. Based on the postal code information of customers, customers are assigned to a customer group. Details about the creation of customer groups are presented in ***Appendix B***.

In case postal code information was not available or could not be converted into coordinates, the second method is used to create customer groups. The demand of these countries is aggregated to country level and the customers are assigned to the country customer group. Analysis showed that only 2% of the total orders placed did not have postal code information. Furthermore, analysis showed that there is a strong relation between the number of orders per country placed and the availability of the postal code information: the relatively small countries (with low demand) have less or no zip code information. Therefore the aggregation to country level does not harm the analysis significantly.

The outcome of applying these two methods to all countries creates almost 1000 customer groups. 600 of these 1000 customer groups are located in sea due to islands and the shape of the countries (the rectangular of Italy for example does contain a lot of sea). Of the other 400 customer groups that are located on land, 18 are aggregated country customer groups. The remaining customer groups are based on the general 'square' method. Illustration of all customer groups is shown in ***Appendix D***. The set of customer groups is defined as N and indexed by n.

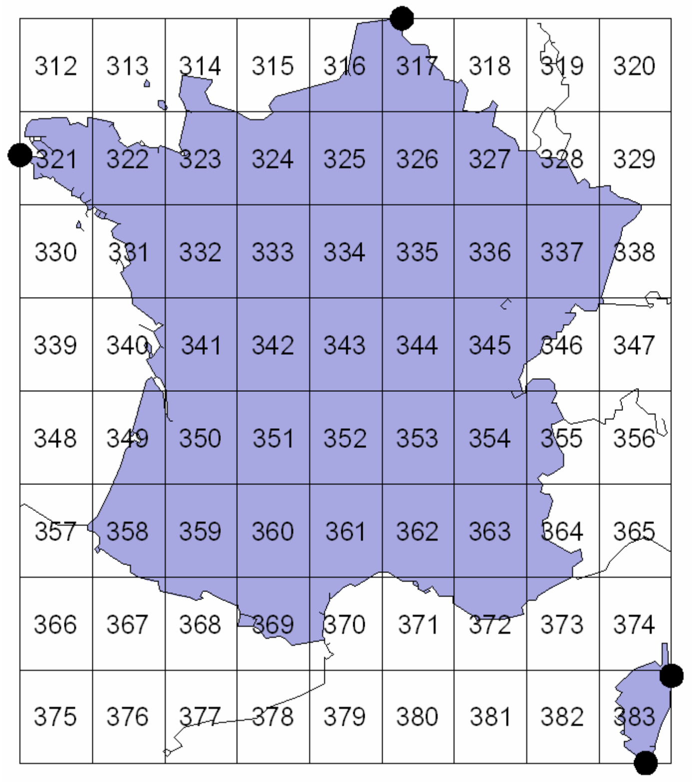


Figure 3: Illustration of the ′square′ method used to define customer groups

2.1.2 Locations of hubs

To fulfil the demand of the customer groups, IBM has several stocking points in the network. In the model these stocking points will be defined as hubs and the set of hubs is defined as J. The possible location for a hub is the centre of each customer group which is indexed by n and illustrated in ***Figure 4***. Details of the calculation of the centre of the customer groups is presented in ***Appendix E***. Indicating the decision variable, referring to the location of the hubs, the binary variable Y_n is introduced in the model. This binary variable is equal to 1 if location n is selected as hub indexed by j and equal to 0 otherwise.

2.1.3 Distance

To determine the distance between two points that are characterized by a coordinate, for example between a hub j and a customer group n, the Haversine formula is used. This mathematical formula is suited for relatively long distances because it takes into account the curve of the sphere. The input of this formula is the coordinates of the two points consisting of a longitude and latitude value (in radians). As a difference in altitude also has to be included, a roads closeness factor is added to the formula, which is set equal to 1,38, a typical value for the Netherlands (Broekmeulen, 2004). Although the value of this roads closeness factor differs between countries, one fixed value is used for this parameter. Otherwise, the mathematical complexity would increase extremely which would be imbalance with the result of the inclusion. The result of the inclusion would be more accurate measure of distance. However, the influence of distance in the model is limited and therefore the accuracy does not have to be extremely high.

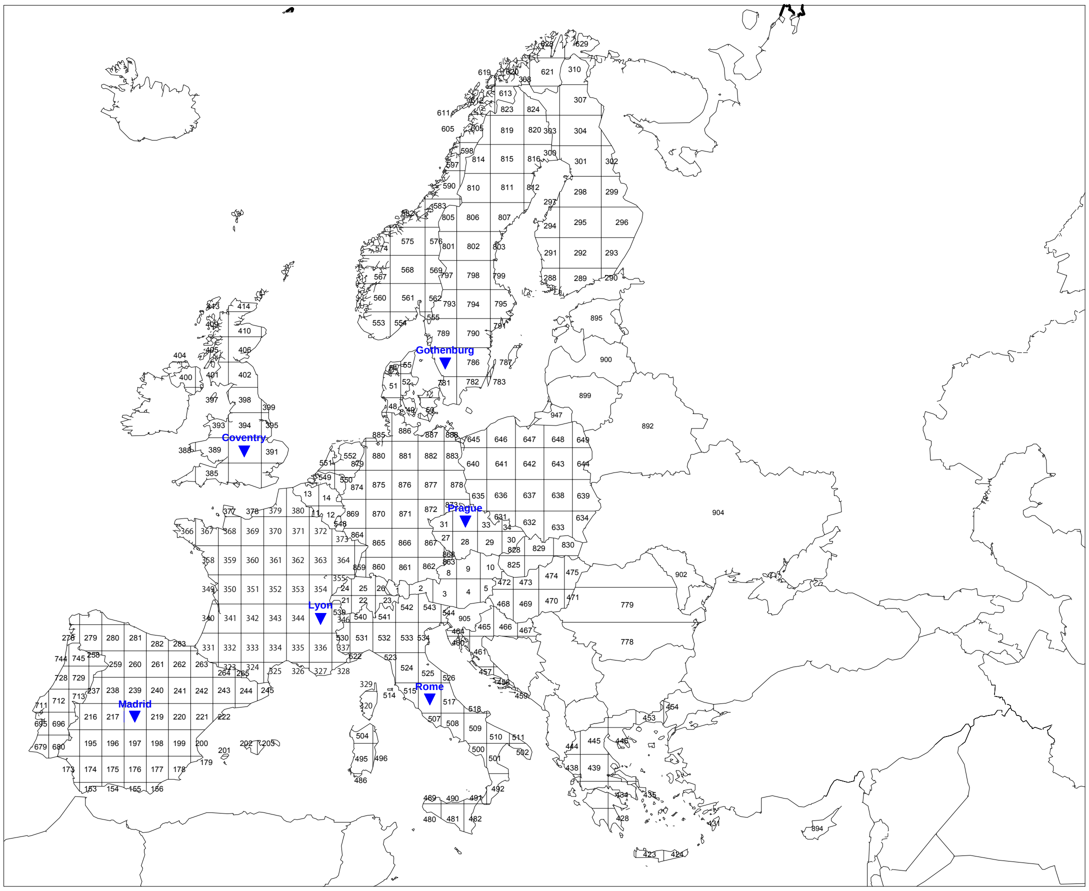


Figure 4: Illustration possible locations of hubs

$$D_{j,n} = l \cdot 2R \cdot \sin^{-1} \sqrt{\sin\left(\frac{lat_j - lat_n}{2}\right)^2 + \cos\left(lat_j\right) \cdot \cos\left(lat_n\right) \cdot \sin\left(\frac{lon_j - lon_n}{2}\right)^2}$$

Where:

l = Road closeness factor (= 1,38)
R = Radius of the earth (= 6367 km)
lat_j = Latitude value of coordinate of hub *j*
lat_n = Latitude value of coordinate of group *n*
lon_j = Longitude value of coordinate of hub *j*
lon_n = Longitude value of coordinate of group *n*

Another option to determine the distances is using the AND data that IBM has available. The AND data takes into account the kind and location of roads and therefore works like a route planner. However, as the available AND data does not cover all countries in the project scope and does not recognize all postal codes, it was chosen to use the mathematical formula. Moreover, a test sample showed that the difference between the distance based on the mathematical formula and the AND data tool of IBM was small: less than 5% of the data showed a large difference (±50%). This large difference between calculated distances by the formula and the AND data tool incurred in case the sea had to be crossed, for example a shipment between Great Britain and Finland. As most shipments have to cross particularly land instead of sea, the probability of a large difference between the two methods will be small. Furthermore, in 90% of the cases the difference between the distances of the AND data and the mathematical formula was less than 20%. These percentages are considered acceptable and moreover supported by the fact that the reliability of the AND data is not proved either.

2.1.4 Demand

As IBM offers a wide range of systems, the installed base of IBM customers varies per type of systems. All these systems consist of several parts of which one unit is defined as one SKU. This is indexed by k in the model. In total, around 30,000 active parts can be identified with a unique SKU identification number. A part is defined 'active' in case it has had demand in the last year.

These SKUs can differ in terms of price, size, vitality level and weight but also in failure pattern. It is assumed that a Poisson process can describe the failure pattern of SKUs. This failure pattern is characterized by a failure rate. The Poisson failure pattern is supported by the fact that most parts are electronic devices. The life span of an electronic device is in general exponential so the failure rate is constant. Due to the constant failure rate, the failure pattern can be defined memory less, one of the characteristics of the Poisson distribution. Furthermore, several case studies showed (Corbijn van Willemswaard, 2006; Van Sommeren, 2007) that Poisson rates are reflective for demand of spare parts. As each customer has a specific set of SKUs that each generate failure driven demand with a SKU specific rate, the demand rate is indicated with rate μ_{ik}, the Poisson rate of customer i for SKU k. Next to the memory less property of the Poisson distribution, the Poisson distribution has another important mathematical property. This property is that the sum of two Poisson distributions is also Poisson distributed with a rate that is equal to the sum of the two Poisson rates. As the demand in the model is aggregated to customer group level, this last property is useful in determining the demand of the customer groups. The demand of SKU k of customer group n is equal to:

$$\mu_{n,k} = \sum_{r \in R_n} \mu_{r,k}$$

Where:

R_n	Set of customers in customer group n, indexed by r
$\mu_{r,k}$	Demand rate of customer r for SKU k

The centre of each customer group represents the aggregated demand of all customers in that customer group. This location is similar to the location of the hubs. Details about the retrieved customer demand data are described in ***Appendix E***.

2.1.5 Demand fulfilment

In case there are several hubs in the network, the model has to determine which hub fulfils the demand for SKU k of customer group n. This decision is constrained by the service time window of the hub and influenced by the transportation costs of the shipment. These two factors will therefore be described in the next two sections.

Service time window

IBM customers have Service Level Agreements which define the maximum allowable time span between the moment a machine breaks down and the moment it is up and running again. To monitor the performance of the service and track whether these agreements are met, the service measure Part Delivery Time (PDT) is used. The PDT indicates the time from the report of a failure by the customer until the delivery of a spare part at the requested location. Hereby it is assumed that the Customer Engineer (CE) can repair the machine with the requested part and that the time to repair can be neglected. In the model, the service time constraint is reflected in the **service time window** indexed by φ that indicates the maximum time it can take to get the spare part from the hub to the customer.

The business of IBM shows that besides the transport time, the service time also consists of a so called ′cut off time′ referring to the time it takes to prepare the spare part for transport, like picking the part from the warehouse. As this 'cut off time' is arranged with the Logistic Service Providers, the assumption that the ′cut off time′ is equal to a fixed time unit (of one hour) is justified. The remaining and largest part of the service time constraint, the transport time, is a function of the distance between the hub and the customer group and a speed profile. The speed profile is defined as the average speed of the shipment in km/hour. Based on data from a Logistic Service Provider, this variable is set at 50 km/hour.

$$\tau_{j,n} = \frac{D_{j,n}}{s}$$

Where:

$\tau_{j,n}$	Transport time of shipment between hub *j* and customer group *n*
$D_{j,n}$	Distance from hub *j* to customer group *n* (km)
s	Speed profile (km/hour)

Due to differentiation of service contracts (into for example sunrise or pre 12) all orders have a different available time span. The facts that a so called 'cut off time' is used and that failures can occur randomly in time, makes that there are more options of available time span (continuously ranging from 6 to 36 hours) than just these marketing service levels. For example, the available time for a customer order with a next day order pre 12 is two hours more in case the failure is reported at 12 PM instead of 14 PM. As distinguishing the available time for each failure does complicate the model extremely, it is decided to use a fixed average service time window. This value is set equal to a weighted average of 18 hours.

To include the service time window into the model as a constraint, the binary parameter $\partial_{j,n}$ is introduced. The value of this parameter is equal to 1 if customer group *n* can be reached within the service time window ($\tau_{\tilde{j},j} \leq \varphi$) and equal to 0 otherwise. This parameter can also be illustrated by circles drawn from the hub locations (see ***Figure 5***): the value of the parameter is equal to 1 for all customer groups *n* that fall within the circle of hub *j* with size *φ*. Therefore, this parameter can also be defined as the geographical coverage.

This geographical coverage can also be determined the other way around: in case a circle with size *φ* is drawn from the central point of customer group *n*, the hubs *j* that are covered by this circle will have parameter $\partial_{j,n}$ equal to 1 (see ***Figure 6***). This binary parameter is illustrated by the green and grey coloured triangles which reflect hubs within and outside the geographical coverage of the customer group respectively.

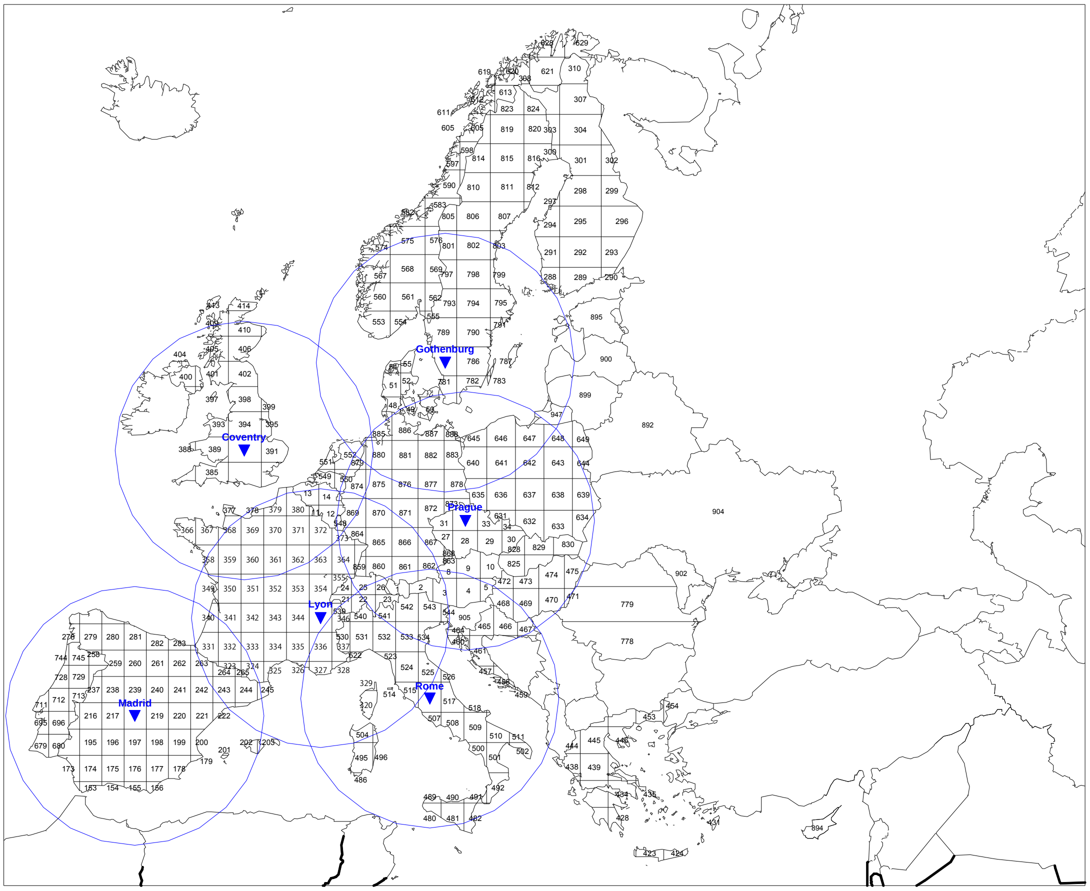


Figure 5: Service time constraint of hubs

Transportation costs

The decision which hub *j* in the service window constraint (so within the circle) fulfils the demand of customer group *n* for SKU *k*, is driven by the transportation costs. The transportation costs are a function of distance between hub *j* and customer group *n* and the weight of the shipment. Details about the exact calculation of these costs will be presented later in this chapter (section 7) that describes all network costs. As the weight of the shipment is SKU dependent, the transportation costs for SKU *k* only depend on the distance between hub *j* and customer group *n*. Therefore, the hub that is located nearest to the customer group logically generates the lowest transportation costs.

The preference order of hubs for demand fulfilment of customer group *n* is also based on the transportation costs. This order is defined by the **lateral transshipment array** $\sigma_n(i)$ for customer group *n* with *i* indicating the place of hub *j* in the array. Therefore, the hub with the lowest transportation costs (so located nearest to customer group *n*) will have the first position in this array indicated by *i* equal to 1. The array only includes the hubs that satisfy the service time window constraint and thus can be characterized with $\partial_{j,n} = 1$. The variable p_n is introduced in the model to indicate the number of hubs that can fulfil demand of customer group *n*. As the number and location of hubs differs per scenario, this number of hubs in the lateral transshipment array is scenario specific. This scenario specific lateral transshipment array can be retrieved from the general lateral transshipment array that contains all possible locations of hubs (so all customer groups within the service time window). The scenario specific lateral transshipment array presents the hubs that are selected from this array and deletes all other possible hub locations.

Allocation fractions

Figure 6 illustrates the several modes through which demand of customer group *n* for SKU *k* can be fulfilled. Indicated by the arrows in ***Figure 6***, the demand of customer group *n* can be fulfilled by one of the hubs within the service time window (indicated by the circle) or by the central source. The order of the hubs that fulfil demand, illustrated by the numbers at the arrows in ***Figure 6*** and determined by the transportation costs, is reflected by the lateral transshipments array. The order of demand fulfilment is now explained in detail.

The first option to fulfil demand is the hub that has the first position in the lateral transshipment array and is defined as a **regular shipment**. In case this hub does not have inventory for SKU *k*, the demand can be fulfilled by the hub that has the successor position in the lateral transshipment array. This is called a **lateral transshipment**. In case this hub also does not have inventory for SKU *k*, the process repeats until one of the hubs in the lateral transshipment array has inventory of SKU *k*. So, all options to provide a lateral transshipment are checked. In case none of the hubs present in the lateral transshipment array of customer group *n* has SKU *k* on stock, it is assumed that a central source will fulfil the demand. This type of demand fulfilment is defined as an **emergency shipment**.

The fraction of demand of customer group *n* for SKU *k* that is fulfilled by hub *j* is indicated by $\alpha_{n,j,k}(S)$. It is dependent on the amount of inventory because this determines the probability of being out-of-stock. Given the base stock levels of the hubs and the demand rates of the customer groups, this fraction can be calculated. The fraction of demand of customer group *n* for SKU *k* that is satisfied by an emergency shipment is included in the model by $\theta_{n,k}(S)$. All fractions are unique for each customer group *n* and SKU *k* combination. Assuming the central source has infinite inventory, every demand can fulfilled be fulfilled. This can be mathematically expressed as:

$$\sum_{j\in\sigma_n(i)} \alpha_{n,j,k}(S) + \theta_{n,k}(S) = 1$$

Where:

$\alpha_{n,j,k}(S)$	Fraction of demand of customer group *n* that is fulfilled by location *j* at base stock level *S*
$\theta_{n,k}(S)$	Fraction of demand of customer group *n* that is fulfilled through an emergency shipment by the central warehouse at base stock level array *S*

2.1.6 Inventory control

It is assumed that the inventory of each hub is controlled by a base stock policy. The **base stock policy** is also known as the one-for-one replenishment policy, because once a SKU *k* is used to fulfil a customer order, immediately a new SKU *k* is ordered to replenish the warehouse. Assuming that the pipeline stock is included in the inventory position, the inventory position of a base stock policy is always equal to $S_{n,k}$, the base stock level of SKU *k* of hub *j*.

Literature (Kranenburg, 2006) supports the use of a base stock policy in a spare parts environment because the relatively low demand in combination with the relatively high value of the SKUs makes the policy well-suited. Although IBM operates in the service logistics, not all SKUs can be characterized by a low demand and high value. Therefore, there could be argued that this inventory policy can not be applied to the IBM service environment. However, the difference between the current (s,Q) inventory policy and the base stock policies lies in replenishment control. The (s,Q) policy trades off the order and replenishment costs to determine the optimal replenishment quantity. The base stock policy does not calculate this quantity because this is always equal to one. The similarity between the (s,Q) and base stock policy is that both policies determine the safety stock. As replenishment and order costs are left out of scope there can be concluded that the use of a base stock policy is justified for this network design study.

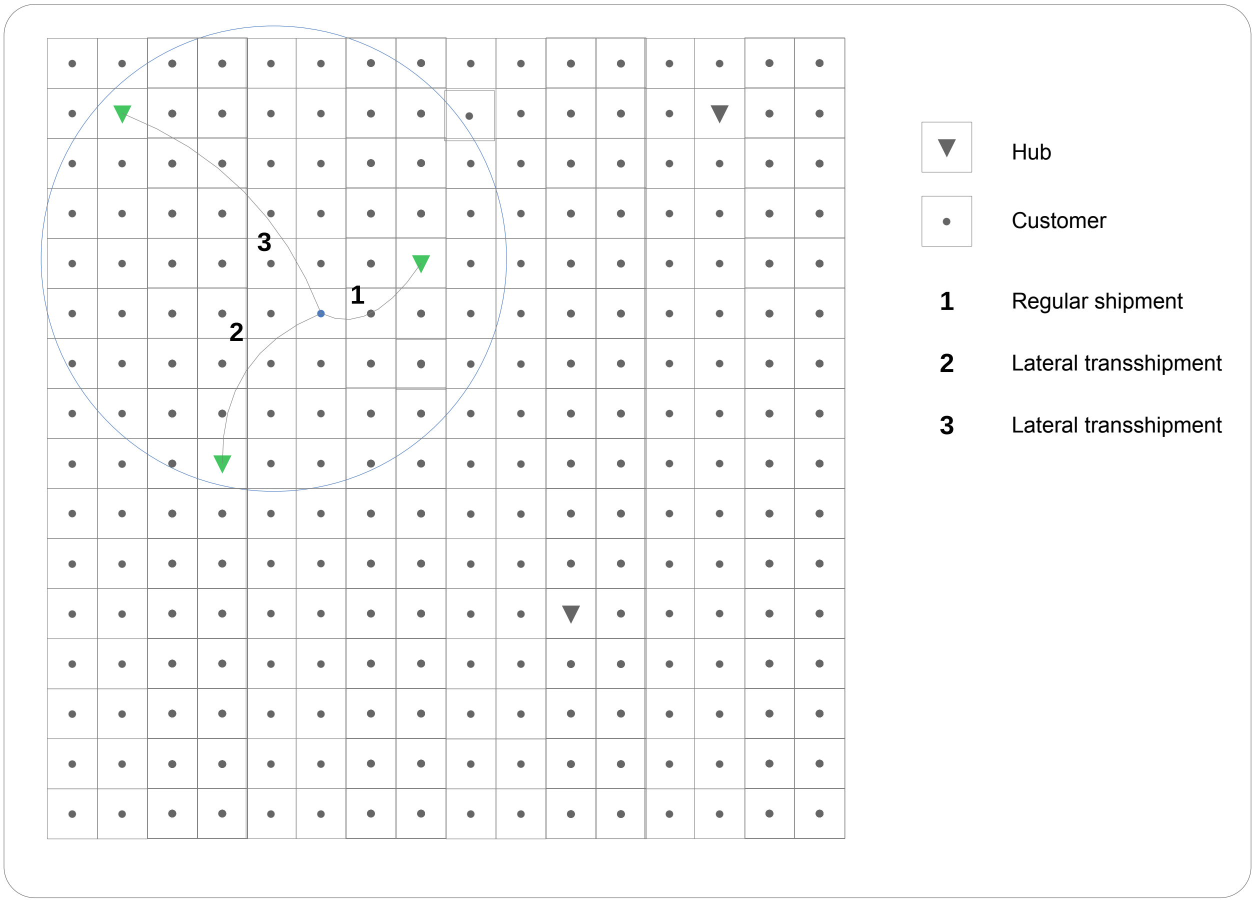


Figure 6: Demand fulfilment of customer group n

2.1.7 Costs

The network costs in the model consist of transport and warehouse costs (see ***Figure 7***). The selection of these costs is based on both literature and practical experiences of IBM professionals. As this measure has to be considered for each order of SKU *k*, there can be concluded that the model is independent for each SKU *k*. Furthermore, the model can be defined as a single item model because it is essential that all demand is delivered as quickly as possible. The subsections will elaborate the different network costs. Details about the data that is used to determine the value of the cost parameters are presented in ***Appendix F***.

Transportation costs

Transportation costs are defined as the costs that have to be paid to Logistic Service Providers for the shipment of parts. This includes for example the wage of the driver and fuel charges. As already indicated, the transportation costs are dependent on the distance and weight of the shipment. Balakrishan et al. (2000) showed that it is sufficient to identify a fee for a couple of distance ranges to calculate the distance. As the maximum distance that can be reached within the service time window is equal to 800 km, three distance ranges (0-200, 200-400 and >400 km) are identified. The distance of 800 km is based on a service window of 18 hours and a speed profile of 50 km/hour. The available time span, reflected in the service time window is subtracted by 2 hours to correct for the distance that has to be crossed to transport the part to the customer instead of the centre of the customer group.

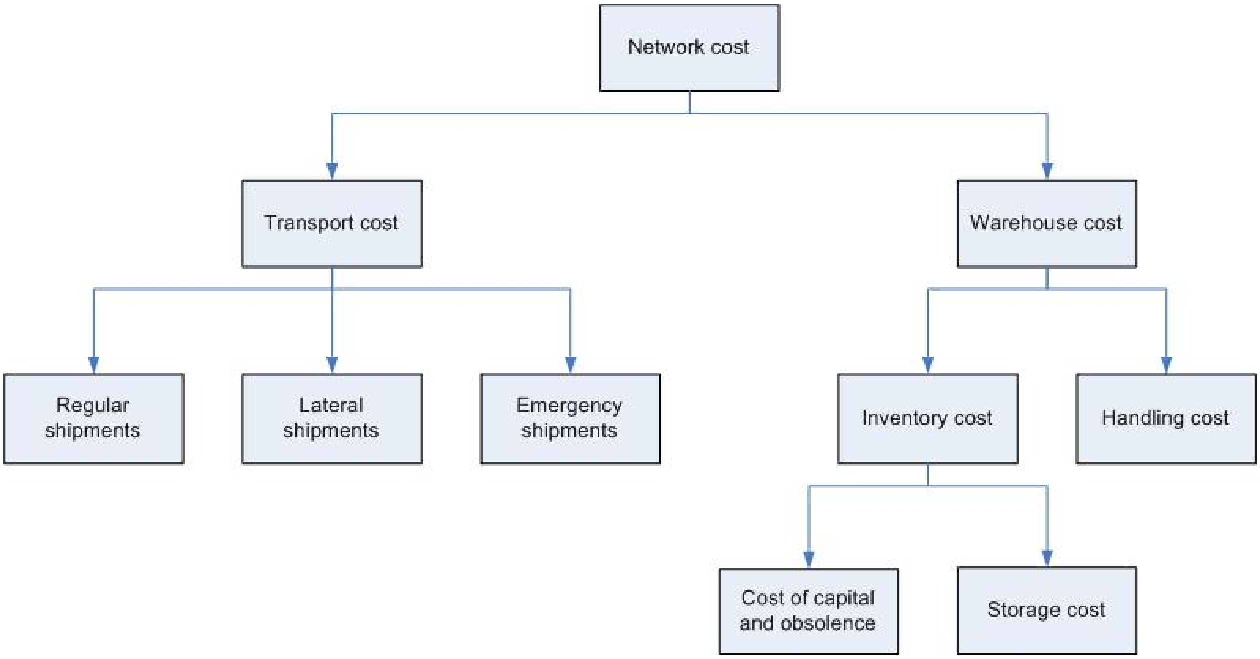


Figure 7: Overview of network costs

The tariffs of these distance ranges have the unit euro per kg. Hereby it should be noted that the minimal weight for the price of a shipment is equal to 2 kg, so even if the shipment is smaller than 2 kg, still 2 times the rates (in euro/kg) has to be paid for the shipment. As the transportation costs can differ per country, the parameter c_j is included into the transport cost calculation.
The transportation costs from hub *j* to customer group *n* for SKU *k* is calculated using:

$$T_{j,n,k} = c_j \cdot f\left(D_{j,n}\right) \cdot \max\left(2, g_k\right)$$

Where:

$T_{j,n,k}$	Transportation costs from hub *j* to customer group *n* for SKU *k*
c_j	Correction factor for shipment from hub *j*
$f\left(D_{j,n}\right)$	Fee value for a shipment which is a function of the distance between hub *j* and customer group *n*
g_k	Weight of SKU *k*

These transportation costs have to be paid for a regular shipment from hub *j* to customer groups *n* for SKU *k*. However, demand can also be fulfilled by lateral transshipments or emergency shipments. As these shipments can be considered a backup option, these transportation costs are higher than the transportation costs for the regular shipment. This is indicated with uplift λ, which value will be equal or greater than 1. For lateral transshipments it is assumed that a value of 1,2 is reflective in case lateral transshipments are taken into account in negotiations with the Logistic Service Providers. Moreover, it takes into account the probability that a cross border shipment is used.
Emergency shipment costs have un uplift factor of 1,5 to reflect the quick, expensive transport mode that is used. This value is based on the current difference in tariffs between air- and road transport modes. The virtual character of the central source has as a consequence that the emergency shipment costs are independent of the distance between hub *j* and customer groups *n*. To make the transport cost function general for all type of shipments, the uplift factor is added.

$$T_{j,n,k}^{reg} = c_j \cdot f\left(D_{j,n}\right) \cdot \max\left(2, g_k\right) \cdot \lambda_{reg} \qquad \text{with } \lambda_{reg} = 1$$
$$T_{j,n,k}^{lat} = c_j \cdot f\left(D_{j,n}\right) \cdot \max\left(2, g_k\right) \cdot \lambda_{lat} \qquad \text{with } \lambda_{lat} = 1{,}2$$
$$T_{j,n,k}^{em} = f\left(3\right) \cdot \max\left(2, g_k\right) \cdot \lambda_{em} \qquad \text{with } \lambda_{em} = 1{,}5$$

As the probability of the use of a specific type of transport option is reflected by the allocation fractions, the multiplication of these fractions with the demand of customer group *n* for SKU *k*, reflects the amount of regular, lateral and emergency shipments.

The total yearly transportation costs for SKU *k* of customer group *n* are then given by:

$$C_{n,k}^{t} = \mu_{n,k}\left(\alpha_{n,\sigma_n(1),k}\left(S\right) \cdot T_{\sigma_n(1),n,k}^{reg} + \sum_{j=\sigma_n(2)}^{\sigma_n(p_n)} \alpha_{n,j,k}\left(S\right) \cdot T_{j,n,k}^{lat} + \theta_{n,k}\left(S\right) \cdot T_{j,n,k}^{em}\right)$$

Considering all customer groups, the network transportation costs for SKU *k* are:

$$C_k^t = \sum_{n \in N} C_{n,k}^t$$
$$C_k^t = \sum_{n \in N}\left(\mu_{n,k}\left(\alpha_{n,\sigma_n(1),k}\left(S\right) \cdot T_{\sigma_n(1),n,k}^{reg} + \sum_{j=\sigma_n(2)}^{\sigma_n(p_n)} \alpha_{n,j,k}\left(S\right) \cdot T_{j,n,k}^{lat} + \theta_{n,k}\left(S\right) \cdot T_{j,n,k}^{em}\right)\right)$$

Where:

$\mu_{n,k}$	Aggregated demand rate of customer group *n* for SKU *k*
$\alpha_{n,j,k}\left(S\right)$	Fraction of demand of customer group *n* that is fulfilled by location *j* at base stock level *S*
$T_{j,n,k}^{reg}$	The transportation costs for a regular shipment from hub *j* to customer group *n* for SKU *k*
$T_{j,n,k}^{lat}$	The transportation costs for a lateral transshipment from hub *j* to customer group *n* for SKU *k*
$T_{j,n,k}^{em}$	The transportation costs for an emergency shipment from hub *j* to customer group *n* for SKU *k*
$\sigma_n\left(i\right)$	Lateral transshipment array for customer group *n* with the position in the array indexed by *i*
p_n	Number of locations that can reach customer group *n* within the service time window φ
$\theta_{n,k}\left(S\right)$	Fraction of demand of customer group *n* that is fulfilled through an emergency shipment by the central warehouse at base stock level array *S*

Warehouse costs

Warehouse costs consist of inventory and handling costs (see ***Figure 7***). As the warehousing costs have to be paid for each hub, the total network costs are summed over all hubs. However, as the number and location of hubs are decision variables of the model, the binary parameter Y_n is included into the warehousing costs function. In that way, the warehouse costs are summed over all customer groups *N* instead of the set of hubs *J*. In case the centre of a customer group is not selected as a location of a hub, no warehouse costs have to be paid. In that case, the value the binary parameter Y_n will be equal to zero and therefore the warehouse costs will be multiplied by 0. Therefore, the warehouse costs are also equal to zero if the possible location is not selected as the location of a hub. Next, the inventory and handling costs will be discussed in detail.

A. Inventory costs

Inventory costs can be split into costs of capital and obsolescence and storage space costs (see ***Figure 7***). Costs of capital and obsolescence cover the opportunity costs of the money invested. These opportunity costs are influenced by the probability of deterioration of stock, damage, theft and obsolescence. The costs of capital and obsolescence are equal to a fixed (interest) percentage of the value of an SKU. The base stock level reflects the average inventory of the whole year at hub *j* as the assumption is made that both physical and pipeline stock is charged for interest. The cost of capital and obsolescence for SKU *k* at hub *j* is therefore calculated with:

$$C_k^h = \sum_{n \in N} h \cdot v_k \cdot S_{n,k} \cdot Y_n$$

Where:

h	Opportunity percentage representing cost of capital and obsolescence for one unit of value/year
v_k	WAC value of SKU *k* (in euros)
$S_{n,k}$	Base stock level at potential location, the centre of customer group *n*, of SKU *k*
Y_n	Binary decision variable that is equal to 1 if the centre of customer group *n* is selected as location of a hub indexed by *j*, and equal to 0 otherwise

The storage space costs have to be paid to Logistic Service Providers for storing 1 m^2 per year. Therefore, these costs depend on the amount of required space which is a function of the size of SKU *k* and the average number of SKU *k* in stock. This average inventory is, as already supported above, equal to the base stock level of SKU *k* at hub *j*. Furthermore, the storage space costs depend on the location of the hub. In general the value of land influences the storage space costs and will be lower in areas with fewer buildings. Taken into account these variables, the network storage space costs for SKU *k* are given by:

$$C_k^u = \sum_{n \in N} u_n \cdot o_k \cdot S_{n,k} \cdot Y_n$$

Where:

u_n	Storage costs for storing 1 m^2 of storage space/year at location *j* (in euros/ m^2 /year)
o_k	Required storage space for SKU *k* (in euros/ m^2)
$S_{n,k}$	Base stock level at potential location, the centre of customer group *n*, of SKU *k*
Y_n	Binary decision variable that is equal to 1 if the centre of customer group *n* is selected as location of a hub indexed by *j*, and equal to 0 otherwise

B. Handling costs

Handling costs have to be paid for the inbound and outbound activities that require resources. Similar to the storage space costs, the handling costs are dependent on the location of the hub. In for example Eastern Europe, the handling rates will be lower due to lower labour wages. Although in practice the inbound and outbound costs are determined separately, due to the influence of consolidation in especially inbound costs, one fixed cost factor is used to determine the handling costs per unit. The effect of consolidation is not taken into account because due to the assumption of a base stock policy, one-for-one replenishment takes place and therefore inbound costs can be determined per SKU *k*. Furthermore, the outbound activity in the model is demand driven and in combination with the assumption that the demand is equal to one SKU *k*, the outbound costs can also be calculated per SKU *k* without taking consolidation into account. Due to the demand driven character of the handling costs function, the demand rate per customer group for each SKU *k* is included. Due to the possibility of lateral transshipments, the actual demand of customer group *n* for SKU *k* at hub *j* is reflected by multiplication of the allocation fractions by the demand rates. The handling costs for SKU *k* are given by:

$$C_k^g = \sum_{n \in N} w_n \cdot \mu_{n,k} \cdot \alpha_{n,j,k}\left(S\right) \cdot Y_n$$

Where:

w_n	Handling costs for handling one SKU at potential location, the centre of customer group *n*
$\mu_{n,k}$	Aggregated demand rate of customer group *n* for SKU *k*
$\alpha_{n,j,k}\left(S\right)$	Fraction of demand of customer group *n* that is fulfilled by location *j* at base stock level *S*
Y_n	Binary decision variable that is equal to 1 if the centre of customer group *n* is selected as location of a hub indexed by *j*, and equal to 0 otherwise

The total yearly warehouse costs of the network for SKU *k* are equal to:

$$C_k^w = C_k^h + C_k^u + C_k^g$$

$$C_k^w = \sum_{n \in N} Y_n \left(S_{n,k}\left(h \cdot v_k + u_n \cdot o_k\right) + w_n \cdot \mu_{n,k} \cdot \alpha_{n,j,k}\left(S\right)\right)$$

As these costs cover the main warehouse activities, it is assumed that the fixed costs like energy costs are included in the rates of these activities.

Excluded costs

- As we assume a green field approach, migration or opening costs of hubs are not taken into account
- Customs and foreign trade (import and export) costs are not considered in the model because these depend on both the location of the hub and the location of the customers. This will complicate the model too much
- As the recall and return process is kept out of scope, scrap, recall and rebalancing costs are excluded from the model

To obtain the total costs, the costs are summated over all SKUs because the model is a single item model. The total costs function (per year) including the transport and warehousing costs is therefore given by:

$$C = \sum_{k \in K} C_k^t + C_k^w$$

2.1.8 Service constraint

The fraction of orders that can be fulfilled by a regular or lateral (trans)shipment is reflected in the service measure. Hereby it is assumed that in case the part can be delivered within the service time window, the order is also repaired within the time that is available and stated on the order request. This implies that the time to repair is assumed to be negligible.

To assure that the required service level is met by the network, a service level constraint is included in the model. This constraint reflects that the service constraint of the network has to be equal or higher than the required service level. Otherwise, the solution is not feasible. As it is a single item model, the total network costs per SKU k are minimized given the service level constraint. The service level and service level constraint can be mathematically expressed as:

$$P_k \geq P_{req}$$

$$P_k = \frac{\sum_{n\in N}\left(1-\theta_{n,k}\left(S\right)\right)\mu_{n,k}}{\sum_{n\in N}\mu_{n,k}} = \frac{\sum_{n\in N}\sum_{j\in J|j\in\sigma_n}\alpha_{n,j,k}\left(S\right)\mu_{n,k}}{\sum_{n\in N}\mu_{n,k}}$$

P_{req}	Required service level for each SKU
P_k	Service level of SKU k
$\theta_{n,k}\left(S\right)$	Fraction of demand of customer group n that is fulfilled through an emergency shipment by the central warehouse at base stock level array S
$\mu_{n,k}$	Aggregated demand rate of customer group n for SKU k
$\alpha_{n,j,k}\left(S\right)$	Fraction of demand of customer group n that is fulfilled by location j at base stock level S

The data used to determine the value of the target service measure is described in ***Appendix E***.

2.1.9 Main model formulation

This section will give an overview of all formulae used in the model. At the end of the report a definition list of all variables is given.

Min $\quad C = \sum_{k\in K} C_k^t + C_k^w \qquad \forall n \in N \qquad$ j,k $\in$ N

subject to

$$P_k \geq P_{req}$$

$$P_k = \frac{\sum_{n\in N}\left(1-\theta_{n,k}\left(S\right)\right)\mu_{n,k}}{\sum_{n\in N}\mu_{n,k}} = \frac{\sum_{n\in N}\sum_{j\in J|j\in\sigma_n}\alpha_{n,j,k}\left(S\right)\mu_{n,k}}{\sum_{n\in N}\mu_{n,k}}$$

$$C_k^t = \sum_{n\in N}\left(\mu_{n,k}\left(\alpha_{n,\sigma_n(1),k}\left(S\right)\cdot T_{\sigma_n(1),n,k}^{reg} + \sum_{j=\sigma_n(2)}^{\sigma_n(p_n)}\alpha_{n,j,k}\left(S\right)\cdot T_{j,n,k}^{lat} + \theta_{n,k}\left(S\right)\cdot T_{j,n,k}^{em}\right)\right)$$

$$C_k^w = \sum_{n\in N} Y_n\left(S_{n,k}\left(h\cdot v_k + u_n\cdot o_k\right) + w_n\cdot\mu_{n,k}\cdot\alpha_{n,j,k}\left(S\right)\right)$$

$$T_{j,n,k}^{reg} = c_j\cdot f\left(D_{j,n}\right)\cdot\max\left(2,g_k\right)\cdot\lambda_{reg} \qquad \text{with } \lambda_{reg}=1$$

$$T_{j,n,k}^{lat} = c_j\cdot f\left(D_{j,n}\right)\cdot\max\left(2,g_k\right)\cdot\lambda_{lat} \qquad \text{with } \lambda_{lat}=1,2$$

$$T_{j,n,k}^{em} = c_j\cdot f\left(D_{j,n}\right)\cdot\max\left(2,g_k\right)\cdot\lambda_{em} \qquad \text{with } \lambda_{em}=1,5$$

$$\mu_{n,k} = \sum_{r\in R_n}\mu_{r,k}$$

2.2 Assumptions

- Demand of all customer groups per SKU k is Poisson distributed with mean $\mu_{j,k}$
- There is one virtual central warehouse with infinite inventory
- Base stock inventory control policy is used for all hubs j
- Inventory costs are charged for both physic and pipeline stock
- Emergency costs are independent of distance
- Road transport is used as only transport mode. This is supported by the fact that this type of transport has relative low costs and is quick enough to fulfil next day orders
- When parts are delivered within PDT, no penalty costs are encountered: the Customer Engineer (CE) is always able to fix the machine when a part is delivered within the PDT
- The replenishment lead time for every hub and SKU is fixed

2.3 Division of model

Defined in the 4 design questions, the model aims to determine: the number and location of hubs, the method to dispatch parts from hubs to customers and the base stock levels of these hubs. The overall aim is to minimize the total network costs which consist of transport and warehousing costs.
The model formulation shows that the design questions influence each other: the decision variable Y_n, that determines the locations of the hubs, is a required input for the other two design questions. The base stock levels $S_{n,k}$ determine the allocation fractions and in turn affect the cost function, which has to be minimized to determine the optimal number and location of hubs.
Finding the optimal values of all four design questions simultaneously would make the model mathematically too complex unless a lot of simplifications would be made to the model. These simplifications would harm the applicability of the model too much. Therefore, it is chosen to split the design questions into 2 parts. The first 2 design questions concerning the network structure, the number and locations of hubs, will form the first main part. The second part will contain the other 2 design questions and is input for the main part (see ***Figure 8***). Both parts are explained in the next sections.

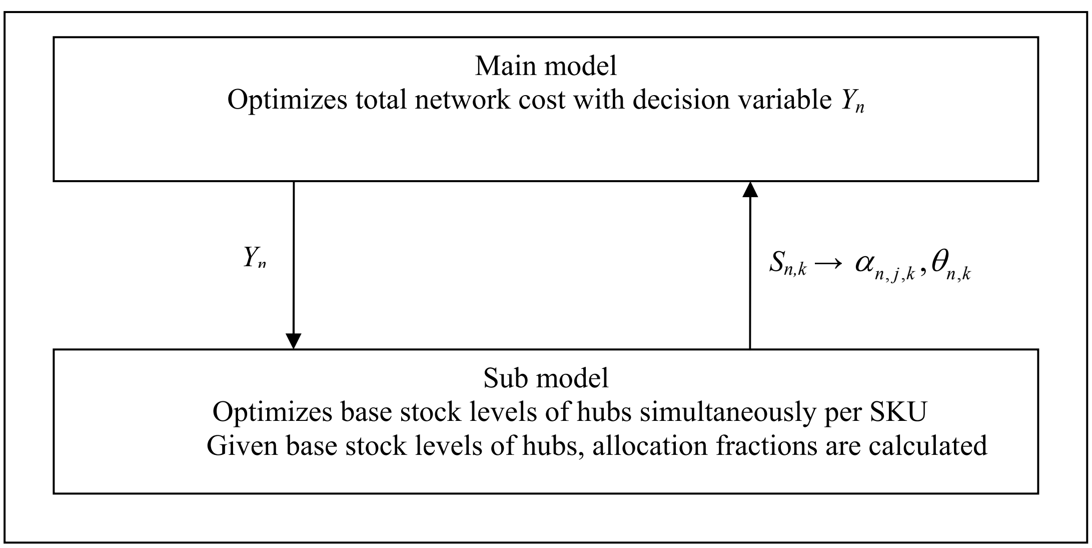


Figure 8: Relation between main and sub model

2.3.1 Main model

As the objective of the project is to provide insight in the optimal network structure, the determination of the optimal number and location of hubs can be considered the main model. The aim of this model is minimization of the total network costs. The base stock levels of the hubs per SKU and the allocation fractions that follow from these base stock levels are used as an input to the model.

Min $\quad C = \sum_{k \in K} C_k^t + C_k^w \qquad \forall n \in N \qquad$ j,k $\in$ N

Subject to

$$P_k \geq P_{req}$$

$$P_k = \frac{\sum_{n \in N} \left(1 - \theta_{n,k}(S)\right) \mu_{n,k}}{\sum_{n \in N} \mu_{n,k}} = \frac{\sum_{n \in N} \sum_{j \in J | j \in \sigma_n} \alpha_{n,j,k}(S) \mu_{n,k}}{\sum_{n \in N} \mu_{n,k}}$$

$$C_k^t = \sum_{n \in N} \left(\mu_{n,k} \left(\alpha_{n,\sigma_n(1),k}(S) \cdot T_{\sigma_n(1),n,k}^{reg} + \sum_{j=\sigma_n(2)}^{\sigma_n(p_n)} \alpha_{n,j,k}(S) \cdot T_{j,n,k}^{lat} + \theta_{n,k}(S) \cdot T_{j,n,k}^{em} \right) \right)$$

$$C_k^w = \sum_{n \in N} Y_n \left(S_{n,k} \left(h \cdot v_k + u_n \cdot o_k \right) + w_n \cdot \mu_{n,k} \cdot \alpha_{n,j,k}(S) \right)$$

$$T_{j,n,k}^{reg} = c_j \cdot f\left(D_{j,n}\right) \cdot \max\left(2, g_k\right) \cdot \lambda_{reg} \qquad \text{with } \lambda_{reg} = 1$$

$$T_{j,n,k}^{lat} = c_j \cdot f\left(D_{j,n}\right) \cdot \max\left(2, g_k\right) \cdot \lambda_{lat} \qquad \text{with } \lambda_{lat} = 1,2$$

$$T_{j,n,k}^{em} = c_j \cdot f\left(D_{j,n}\right) \cdot \max\left(2, g_k\right) \cdot \lambda_{em} \qquad \text{with } \lambda_{em} = 1,5$$

$$\mu_{n,k} = \sum_{r \in R_n} \mu_{r,k}$$

Decision variable

Y_n	Binary decision variable that is equal to 1 if the centre of customer group *n* is selected as location of a hub indexed by *j*, and equal to 0 otherwise

2.3.2 Sub model

As indicated above, the 2 research questions concerning 1) the method to dispatch the parts from hubs to customers and 2) the base stock level will be answered in the sub model. As illustrated in ***Figure 8***, the output of this sub model is used as an input for the 2 design questions concerning the network structure. As the binary parameter Y_n is essential for the answer on these questions, this parameter in turn is input of this sub model and directly determines the lateral transshipment array $\sigma_n(i)$. Therefore, interaction between the two models is essential (see ***Figure 8***).

Indicated by the lateral transshipments array, the method to dispatch a part from a hub to a customer is based on the transportation costs. The number of hubs in the lateral transshipments array is constrained by the service time window. The allocation fractions that reflect the probability that the part is dispatched from hub *j* to customer group *n*, depends as explained before on the fill rates of the hubs. Therefore, the design question concerning the inventory of the hubs does influence the method to dispatch a part from a hub to a customer. It does not influence the general dispatch method but the allocation fractions that indicate the amount of regular, lateral and emergency (trans)shipments.

Given base stock levels of hubs, demand rates of customers and replenishment lead time, the fill rate of the hubs can be calculated. In case of Poisson distributed demand, the fill rate can be calculated with the Erlang loss probability function with base stock levels representing the number of servers and a rate equal to the multiplication of the replenishment lead time and total demand of hub *j* (Reijnen, 2008):

$$\beta_{j,k}(S) = 1 - L\left(S_{j,k}, t_j^{reg} \hat{M}_j\right)$$

Where:

$S_{j,k}$	Base stock level at hub *j* for SKU *k*
t_j^{reg}	Replenishment lead time for hub *j*
$\hat{M}_j$	Total demand at hub *j*

In this formula the parameters are indexed by *j* instead of *n* in the main model because the decision variable Y_n is input of the model. Therefore the set of hubs *J* can be considered fixed and the summation over all potential locations, equal to customer groups, is not necessary anymore.

To determine the optimal value of the base stock levels, a trade off between the transport and warehouse costs has to be made. These costs are influenced by the demand rate $\mu_{n,k}$ and value of SKU *k*. Next to the demand from the customer groups, the hubs can also receive demand from hubs. This is caused by the possibility of lateral transshipments. Therefore, the optimal value of the base stock level is dependent on the base stock level of the other hubs and the lateral transshipment array of the customer groups. Due to the service constraint that has to be satisfied for each demand, the model has the objective to minimize the total network costs for each SKU *k*:

$$\text{Min} \quad C_k = C_k^t + C_k^w$$

$$P_k \geq P_{req}$$

$$P_k = \frac{\sum_{n \in N} \left(1 - \theta_{n,k}(S)\right) \mu_{n,k}}{\sum_{n \in N} \mu_{n,k}} = \frac{\sum_{n \in N} \sum_{j \in J | j \in \sigma_n} \alpha_{n,j,k}(S) \mu_{n,k}}{\sum_{n \in N} \mu_{n,k}}$$

$$C_k^t = \sum_{n \in N} \left(\mu_{n,k} \left(\alpha_{n,\sigma_n(1),k}(S) \cdot T_{\sigma_n(1),n,k}^{reg} + \sum_{j=\sigma_n(2)}^{\sigma_n(p_n)} \alpha_{n,j,k}(S) \cdot T_{j,n,k}^{lat} + \theta_{n,k}(S) \cdot T_{j,n,k}^{em} \right) \right)$$

$$C_k^w = S_{j,k}\left(h \cdot v_k + u_j \cdot o_k\right) + w_j \cdot \mu_{n,k} \cdot \alpha_{n,j,k}(S)$$

<u>Decision variables</u>

$S_{n,k}$	Base stock level at potential location, the centre of customer group *n*, for SKU *k*

2.4 Conclusion

This chapter presented the basic model that contains all variables to answer the four design questions. As it is, from a mathematically viewpoint, too complex to answer the four design questions simultaneously, we identified a main and a sub model. Therefore, we need to determine for both the main and sub model how these models are solved. In the next chapter two different methods are identified to answer the relevant design questions in the main and sub model.

3. Method

In ***Chapter 2*** a main and sub model are introduced that both address 2 design questions. In this chapter we present the method to answer the relevant design questions for both the main and sub model. In the first section the method, a greedy algorithm, to solve the sub model is described. Next, the method of the main model, a scenario analysis, is discussed. In section 3 the implementation of both the main and sub model into a software tool is described. Finally, a conclusion is drawn.

3.1 Greedy algorithm

The greedy algorithm will be applied to the sub model that addresses the design questions concerning the method to dispatch parts and the inventory level of the hubs. As indicated in ***Chapter 2***, the values of the options that are defined in the dispatch method are determined in the sub model, and not the method to dispatch parts. These values of options are reflected in the allocation fractions and depend on the inventory levels of the hubs, the fourth design question. Therefore the decision variables of the sub model and thus of the greedy algorithm, are equal to the base stock levels of the hubs.

A **greedy algorithm** is an approximation technique and known as the ´biggest-bang-for-the-buck´ method: it iteratively chooses the alternative that provides the ´biggest-bang-for the-buck´ until a certain stopping criterion is reached. In this case, the stopping criterion is minimal costs given a service target level. The alternatives are an increase of the base stock levels of the hubs. So the base stock level of the hub that provides most benefits by increasing the base stock level by one SKU will be increased by one.

As the greedy algorithm is an approximation technique, the values of the decisions variables are not necessarily optimized. However, validation of the greedy algorithm shows close-to-optimal solutions (Kranenburg, 2006; Reijnen et al., 2008). Due to the large size of the problem with 400 customer groups, 400 possible locations and nearly 30,000 SKUs, it is mathematical too complex to optimize the problem without simplifications. Furthermore, a close-to-optimal solution can be seen as sufficient in a rapid changing environment taking into account that the used data and parameters are also estimates. Therefore, the greedy algorithm is considered a suitable method for the sub model.

The algorithm consists of an optimization and an evaluation algorithm which are used in turn. The evaluation part calculates the total costs (C_k) given the locations of the hubs (Y_j) and the inventory levels of the hubs ($S_{j,k}$). The optimization part aims, given a set of locations (Y_j), to minimize costs while reaching the target service level with the inventory levels for each SKU k as decision variable. Therefore interaction between the parts is needed iteratively: the evaluation part needs the base stock levels of the hubs and the optimization part requires the costs given the base stock levels (see ***Figure 9***). Details about both algorithms are described below.

3.1.1 Evaluation algorithm

As illustrated in ***Figure 9***, the evaluation algorithm determines, given base stock levels of the hubs $S_{j,k}$ for SKUs k and the total demand at each hub ($\hat{M}_{j,k}$), the allocation fractions $\beta_{j,k}, \alpha_{n,j,k}, \theta_{n,k}$ for each customer group j and SKU k combination. The allocation fractions are used to determine the costs as shown in the model formulation section presented in ***Chapter 2***.

The complexity of calculating the base stock levels is caused by the possibility of lateral transshipments. The consequence of the possibility of lateral transshipments is that the hub not only experiences demand from customers, but also from other hubs. Therefore, the base stock levels of the hubs each depend on the base stock levels of the other hubs, the demand pattern of customers in the network and the lateral transshipment array.

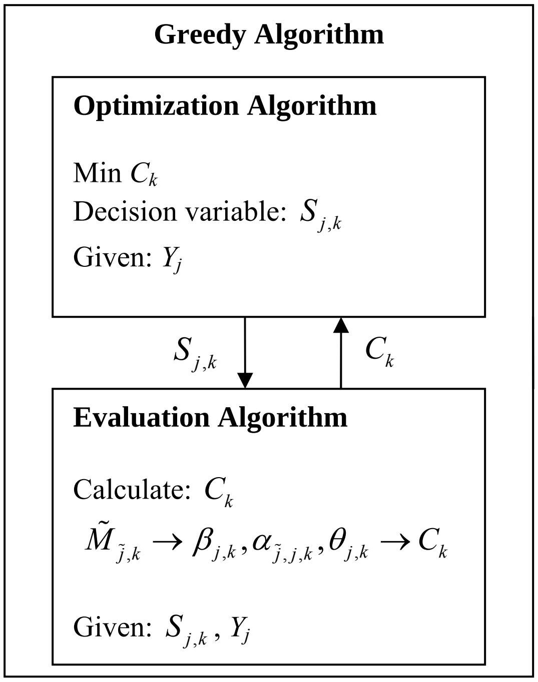


Figure 9: Greedy algorithm

Reijnen et al. (2008) present an evaluation algorithm that takes into account these factors. This evaluation algorithm is a variant of the evaluation algorithm of Kranenburg (2006) that also takes into account lateral transshipments. The differences between the models on which the evaluation algorithm of Reijnen et al. (2008) and Kranenburg (2006) are based, are described in ***Appendix G***.

The evaluation algorithm (Reijnen, 2008) is mathematically formulated by:

Step 1: For all hubs $j \in J$, $k \in K$, $\beta_{j,k}(S) := 1 - L\left(S_{j,k}, t_j^{reg} \sum_{n \in N | \sigma_n(1) = j} \mu_{n,k}\right)$

Step 2: For all customer groups $n \in N$, $k \in K$, $\hat{M}_{n,\sigma_n(1),k} := \mu_{n,k}$

Step 3: For all customer groups $n \in N$, $k \in K$:

Step 3-a: For $2 \le i \le p_n$:

$$\hat{M}_{n,\sigma_n(i),k} := \left(1 - \beta_{\sigma_n(i-1),k}\right) \hat{M}_{n,\sigma_n(i-1),k}$$

Step 4: For all hubs $j \in J$, $k \in K$:

Step 4-a: $\hat{M}_{j,k} := \sum_{n \in N | j \in \sigma_n} \hat{M}_{n,j,k}$

Step 4-b: $\beta_{j,k}(S) := 1 - L\left(S_{j,k}, t_j^{reg} \hat{M}_{j,k}\right)$

Step 5: Repeat step 3 and 4 until $\hat{M}_{j,k}$ does not change more than ε for each $j \in J$

Step 6: For all customer groups $n \in N$ and hubs $j \in J | j \in \sigma_n$, $\alpha_{n,j,k}(S) := \dfrac{\beta_{j,k}(S) \hat{M}_{n,j,k}}{\mu_{n,k}}$

Step 7: For all customer groups $n \in N$, $k \in K$ $\theta_{n,k}(S) = 1 - \sum_{j \in J | j \in \sigma_n} \alpha_{n,j,k}(S)$

Notation

N	Set of customer groups which centre is an option for locating a hub indexed by n
K	Set of Stock Keeping Units (SKUs), indexed by k
J	Set of locations that are selected from set N, defined 'hub' and indexed by j
R_n	Set of customers in customer group n, indexed by r
$\mu_{n,k}$	Demand rate for customer group n
σ_n	Lateral transhipment array for customer group n
$\sigma_n(1)$	The first-choice hub of customer group n
p_n	The number of hubs that can reach group n within the service time window
S	Base stock level array
t_j^{reg}	Replenishment lead time from hub j
$\beta_{j,k}(S)$	Fill rate of hub j for SKU k
$\alpha_{n,j,k}(S)$	Fraction of demand of customer group n that is fulfilled by hub j at base stock level S
$\theta_{n,k}(S)$	Fraction of demand of customer group n that is fulfilled through an emergency shipment by the central source at base stock level array S
$\hat{M}_{n,j,k}$	Demand from customer group n to hub j for SKU k
$\hat{M}_{j,k}$	Total demand at hub j for SKU k

To determine the total demand at hub j consisting of demand from customers as well as from other hubs, steps 1 to 5 have to be fulfilled.

First, the total demand of hub j is set equal to the summation of demand of all customer groups n that have hub j on the first position in the lateral transhipment array σ_n. The fill rate of hub j, is, given the inventory levels of the hubs, calculated with an Erlang distribution with rate equal to the multiplication of this total demand with the replenishment lead time. However, this demand rate does not reflect the actual demand due to the possibility of lateral transshipments. These transshipments are used to cope with the fact that the fill rates of the hubs do not have to be equal to 100%. In case the fill rate of hub j is lower than 100%, the fraction of demand that can not be fulfilled is forwarded to hub j that has a subsequent position in the lateral transhipment array. As this array depends on customer group n, the demand for hub j from customer group n is equal to the fraction of the total demand at predecessor hub j that can not be fulfilled multiplied by the demand from customer group n at predecessor hub j. Taking into account all demand that originates from lateral transshipments, the total demand that arrives at hub j changes. As a consequence, the fill rate of hub j also changes which in turn affects the fractions of demand that are forwarded to the backup hubs. The iterative process of a change in fill rates that affects the amount of demand that arrives at hub j from predecessor hub j in array σ_n and therefore causes a change in the total demand and in turn again causes a change in the fill rate continues until the total demand at hub j only changes slightly. Reijnen et al. (2008) showed that this iterative process converges (quickly).

Finally, based on the fill rates of the hubs and the final demand rates at the hubs, the allocation fractions for each customer group n and SKU k combination can be determined. As it is assumed that all demand is fulfilled using emergency shipments in case none of the hubs is able to fulfil demand, the fraction of emergency shipments can be determined by subtracting the sum of all allocation fractions from 1.

3.1.2 Optimization algorithm

The aim of the optimization algorithm is to determine a feasible solution with minimal costs. A solution is defined feasible if the service target constraint is met. The optimization algorithm stepwise builds up inventory at the hubs for each SKU *k*. The algorithm, similar to Kranenburg (2006), consists of three steps. First the optimization algorithm is mathematically formulated. Next, an explanation of the steps of the optimization algorithm is presented.

<u>Step 1: Initialization</u>
Set all inventory levels of the hubs $S_{j,k}$ equal to 0
Set $S_{j,k} := 0, j \in J, k \in K$

<u>Step 2: Base stock levels</u>
For each SKU $k \in K$:
Step 2a: Calculate $\Delta C(k, j),\ j \in J$
Step 2b: While min $\{\Delta C(k, j)\} \leq 0$

1. Determine $\hat{j}$ such that $\Delta C(k, \hat{j}) \leq \Delta C(k, j),\ j \in J$
2. Set $S_{k,\hat{j}} = S_{k,\hat{j}} + 1$
3. Calculate $\Delta C(k, j),\ j \in J$

<u>Step 3: Service level constraint</u>
For each SKU $k \in K$:
Step 3-a Calculate $\Delta R_{j,k}(S_{j,k}),\ j \in J,\ k \in K$
Step 3-b While $P_k < P_{k,required}$ and $\Delta P_k > \Delta P_{req}$

1. Determine $\hat{j}$ such that $\Delta R_{\hat{j},k}(S_{\hat{j},k}) \geq \Delta R_{j,k}(S_{j,k})$
2. Set $S_{k,\hat{j}} = S_{k,\hat{j}} + 1$
3. Calculate $\Delta R_{j,k}(S_{j,k})$ for all $j \in J$ and $k \in K$

Where:

$$\Delta R_{j,k}(S_{j,k}) = \frac{\Delta P_k}{\Delta C_k} = \frac{P_k(S_{j,k}+1) - P_k(S_{j,k})}{abs\left[C_k(S_{j,k}+1) - C_k(S_{j,k})\right]}$$

$$C_k = C_k^t + C_k^w$$

P_k	Service level of SKU *k*
$\theta_{n,k}(S)$	Fraction of demand of customer group *n* that is fulfilled through an emergency shipment by the central source at base stock level array *S*
$\mu_{n,k}$	Aggregated demand rate of customer group *n* for SKU *k*
$\alpha_{n,j,k}(S)$	Fraction of demand of customer group *n* that is fulfilled by hub *j* at base stock level *S*
$\Delta R_{j,k}(S_{j,k})$	Relative increase in service level per cost unit for hub *j* with base stock level $S_{j,k}$ for SKU *k*
ΔP_k	Difference in service level for SKU *k* due to change of base stock level
ΔC_k	Difference in costs for SKU *k* due to change of base stock level
$P_k(S_{j,k})$	Service level with base stock level $S_{j,k}$ for hub *j* and SKU *k*

C_k	Total network costs, consisting of transport and warehouse costs, for SKU *k*
C_k^t	Total transportation costs, consisting of regular, lateral and emergency shipments, for SKU *k* (in euros)
C_k^w	Total warehousing costs, consisting of inventory and handling costs, for SKU *k* (in euros)

The first step is the initialization phase. In this phase the base stock levels are set equal to a zero.

The second step is evaluating the increase of base stock levels with one SKU *k* for every hub $\tilde{j} \in J$. This step continues until the stopping criterion, minimal costs, is reached. As we assume a convex cost function, this implies that this step continues as long as for a SKU *k* an increase of a base stock level S_{jk} would lead to a cost decrease. The base stock level of the hub that gives the largest cost decrease by one will be selected to increase the base stock level by one. At the start, raising the inventory will raise inventory costs, but will be cheaper than the expensive other modes like emergency shipments. Therefore, the increase in inventory costs will be smaller than the decrease of emergency shipment costs. At a certain moment the set of base stock levels of the hubs will generate minimal costs. At that moment addition of stock is more expensive than the decrease in emergency shipment costs. Then the algorithm stops with this step and starts with the third step of the algorithm.

Step 3 starts with a check whether the service constraint is met. In case the service constraint is met, the optimization algorithm generates the solution that is output of step 2. In case the service constraint is not met, the base stock levels will be raised with one. This is based on the expectancy that an increase of inventory level will always lead to a higher fill rate (Kranenburg, 2006). The base stock level of the hub that has the highest relative increase in service level per cost unit, is increase by one first.
The third step of the optimization algorithm ends if the required service level (given as input) is reached or if the difference between the previous and new service level is less than the required (given as input) difference in service levels. In the last case, the service level will not reach the required service level due to the coverage of the hubs. For example if a scenario contains three hubs, it is impossible to attain a required service level of 95% because large part of the demand lies outside the geographical coverage of the hubs. The demand outside the geographical coverage will still need emergency shipments which cause an upper limit for the service level.

3.2 Scenario analysis

To find a solution for the main model, a scenario analysis is used. A **scenario analysis** is defined as the process of analyzing the best suited network by considering alternative possible outcomes of scenarios. By using a logical scenario set, a near-to-optimal network design can be found. The aim of the scenario analysis is to improve decision making by allowing more complete consideration of outcomes and their implications.
The decision variable Y_n in the model represents the scenario: the number and locations of hubs. As the decision variable Y_n has a binary value, the number of options seems to be small. However, the high number of customer groups (400) cause that the number of possible combinations of locations is extremely high.
Next to the extremely high number of possible combinations of locations, there is also another reason to use a scenario analysis instead of a mathematical optimization method: This method is also useful to investigate the implications of constraints in the network concerning location choice or investigate the effect of movement of a location. For example as several activities are outsourced, the network is constrained by the network of Logistic Service Providers.

The creation of the scenario requires determining the number and location of the hubs, the decisions that are in line with the design questions that are addressed in the main model.
The number of hubs is varied from 5-20 hubs. This determination of the value of the number of hubs can be randomly or systematically, for example based on comparison of several scenarios.
The locations of the hubs can be determined by logical insight or the Warehouse Site Planner (WSP). The WSP was already mentioned in the first chapter because this tool was used to design the current network. The input for the locations of the hubs uses a version of the WSP that minimizes the number of kilometres that has to be crossed to transport parts to customers.

3.3 Implementation

Based on the above mentioned methods, the main and sub model are implemented in a software tool. In the tool a stochastic and dynamic environment with customer demand generation, demand fulfillment and hub replenishment is simulated.
To make sure that the proposed solution is correct, it has to be checked whether the model is build right (**verification**) and whether it is the right model for the intended application environment (**validation**). An overview of the complete verification and validation process is presented in ***Appendix H***. The process shows the robustness and practicality of the model.
The program can be used as a standalone executable file. Due to the relatively small size, the program can run on a standard computer. The running period strongly depends on the number of hubs, number of SKUs and the number of scenarios. Details about the implementation of the model into software can be found in ***Appendix K***.

3.4 Conclusion

Due to the mathematical complexity, both methods used to handle the main and sub model are approximation techniques. Therefore, the solution of the model will not be optimal, but the decision values are expected to approach near-to-optimal values if a logical set of scenarios is used. In the next chapter, part 2 of the report starts that applies the developed model and chosen methods to the business environment of IBM.

4. Results

This chapter presents the results of the case study. First, the design of the experiment is discussed. Next, findings are described per design question for the developed design approach; the integrated design approach with lateral transshipments. The third and fourth section show a comparison of this developed approach with two other existing design approaches: respectively the integrated design approach without lateral transshipments and the decoupled design approach. Finally, a conclusion is drawn.

4.1 Design of experiment

This section first describes the design approaches that are included in the design of the experiment. Next, the types of scenario sets that are used to address the design questions of the main model are presented.

4.1.1 Design approaches

To have a benchmark for the performance (in terms of total network costs) and the solution of the developed design approach III in this thesis, a comparison is made with the performance and solution of two other existing design approaches. The design of the experiment includes the design approaches:

I. Decoupled design approach
II. Integrated design approach without lateral transshipments
III. Integrated design approach with lateral transshipments

The main difference between design approach I and III is that the design approach I does not take inventory into account determining the optimal network design. SPO EMEA does have a customized tool, the Warehouse Site Planner (WSP), which determines the optimal location of hubs given the number of hubs. The tool determines the optimal location of hubs by minimizing the amount of kilometres that have to be crossed to fulfil customer demand. This is based on the supposition that transportation costs are dominant in total network costs and that the transportation costs are linear related the number of kilometres. As the inventory decision is excluded from the determination of the network design, this tool is based on a decoupled design approach. To make a comparison between approach I and III, the output of this tool is therefore compared to the solution of design approach III. The comparison is based on the total network costs that can be determined by the computerized model. Hereby the method for design question 3 and 4, the method to dispatch parts and the determination of inventory, is kept similar so equal to the method of approach III presented in ***Chapter 3***. Furthermore, the difference between the integrated and decoupled approach is analyzed by investigating the influence of the inventory costs on the solution and on total network costs.

The difference between design approach II and III is related to the method that is used to dispatch parts, so design question 3. In design approach II the part is always dispatched from the hub that has minimal transportation costs, so the ‘first choice hub’. The method to dispatch parts is reflected in the lateral transshipments array and therefore this array only contains the ‘first-choice’ hub for design approach II. The lateral transshipments array for design approach III contains all hubs that fall within the service time window. The exclusion of lateral transshipments also affects the method used to answer design question 4. Design question 4, the determination of the inventory, has to be adjusted (compared to design approach III) for design approach II because demand in approach II only consists of demand from customers. The demand in design approach III does, due to lateral transshipments consist of demand from customers as well as from other hubs. Therefore, the evaluation part of the greedy algorithm can jump over steps 3 to 5 for design approach II. The method for addressing the remaining 2 design questions, the number and location of hubs, does not differ for design approaches II and III.

4.1.2 Scenario sets

To answer the design questions that are part of the main model solved by a scenario analysis, two types of scenario sets are used as input. First the types of scenario sets are described. Next, it is indicated for which design questions the types of scenario sets are used.

1. "Stepwise scenario set" is defined as a set of scenarios in which each successor scenario has one additional hub. This is the only difference between scenario (A) and scenario ($A+1$). So, A hubs are identical under both scenarios and in scenario $A+1$ there is an extra hub $A+1$.
2. "Optimal set of scenarios" is a set of scenarios that are independent of each other. So, the locations per scenario differ. Furthermore, the locations of hubs for each scenario are chosen in a geographical optimal way.

To investigate the first design question, the number of hubs, the first type of scenario, "stepwise scenario set" is used as input. This scenario set ensures that no other effects than an increase in number of hubs can influence the performance of the network. The second type of scenario set, "optimal set of scenarios", is used if the second design question, the location of hubs, is addressed.

4.2 Design approach III

This section describes the findings from the results of approach III. The following sub sections present the findings per design question.

4.2.1 Number of hubs

The first research question addresses the number of hubs that provides minimal costs at a given service constraint. In ***Figure 10*** the service constraint is illustrated by the hatched transparent rectangular. It indicates that all values behind this rectangular, so from 1 to 7 hubs, are not feasible because they do not satisfy the service constraint.

Figure 10 shows that the total costs are negatively related to the number of hubs so if the number of hubs increases, the total costs decreases. This can be explained by the fact that by adding a hub, the amount of demand that falls within the coverage of the hubs increases. This in turn decreases emergency shipment costs. Furthermore, the transportation costs decreases because customers are located closer to a hub.

Moreover ***Figure 10*** shows that the costs function is decreasing over the total domain. This is caused by the use of lateral transshipments and the exclusion of opening costs of a hub. Lateral transshipments leads to inventory sharing and therefore inventory does not increase. Details about inventory are described in subsection 4, related to design question 4. Opening costs of hubs are fixed costs that have to be paid to open a hub for example costs of movement, organizing the lay-out or building the hub. In case the opening costs of a hub were added to the costs function, at a certain moment the costs function would increase. This moment depends on the opening costs. Therefore, the optimal number of hubs depends on the opening costs of a hub.

The opposite, positive, relation exists between the number of hubs and the service level: if the number of hubs increases, the service level also increases. This is caused by the fact that the coverage of all hubs increases or stays equal (in case the coverage area of the new hub is already covered by existing hubs). Furthermore, ***Figure 10*** shows that initially the addition of a hub causes a substantial increase in service level, but later the incremental increase in service due to addition of a hub is limited. This can be explained by the fact that the addition of a hub at the start does have a small overlap with the coverage of other hubs. The overlap will increase if the number of hubs increases.

<u>Conclusion number of hubs</u>

- Total costs are negatively related to number of hubs
- Total cost function is decreasing over the total domain due to lateral transshipments and exclusion of opening costs of hubs
- Service level is positively related to number of hubs

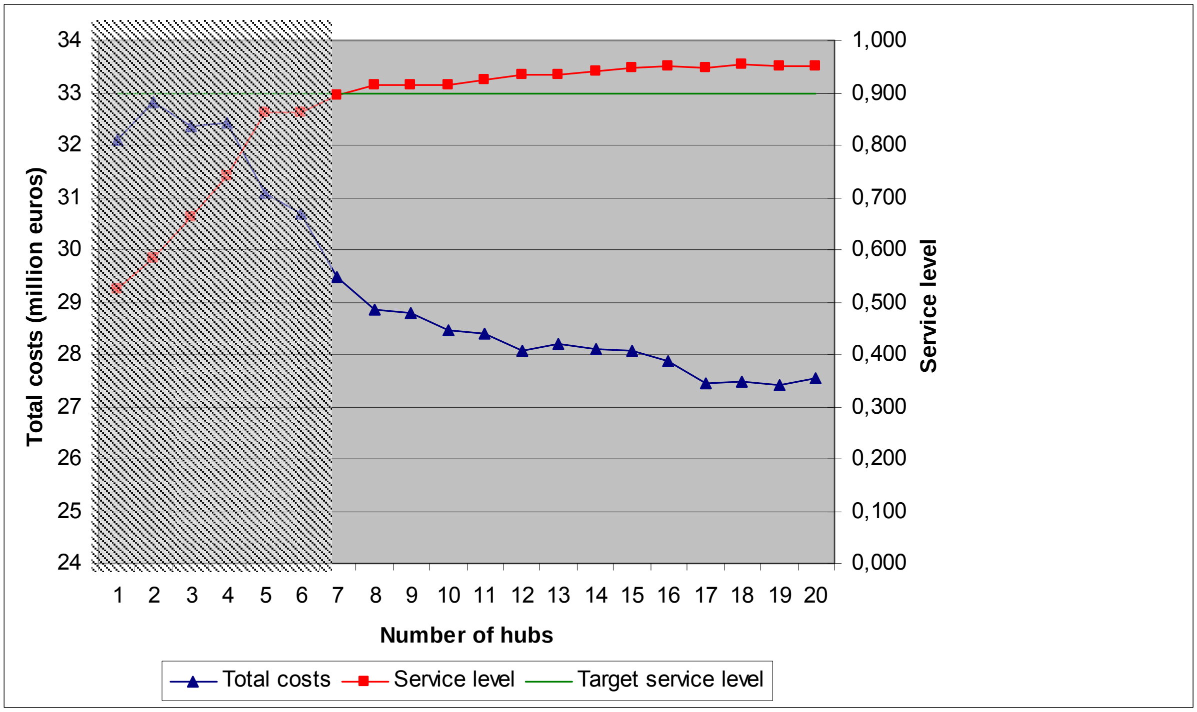


Figure 10: Influence of number of hubs on total costs and service level

4.2.2 Location of hubs

This subsection will describe the results of the analysis that investigates the location of hubs. To get insight in the optimal location and the influence of the location of hubs on the total costs, several scenarios for 6 to 14 hubs are investigated. The analysis starts at 6 hubs because of the service constraint that is first met at 6 hubs. ***Appendix M*** shows that the optimal set of locations is in most cases the set of locations that is geographically locally spread and are located in high density demand areas. The optimal set hereby is defined as the scenario that has, compared to the other investigated scenarios the minimal costs. Furthermore, ***Appendix M*** shows that the difference between the scenarios is minimal. The relative difference, which is calculated based on the maximum and minimal costs of the considered scenarios, varies from 1% (for 9 hubs) to 8% (for 6 hubs).

<u>Conclusion location of hubs</u>

- Optimal scenario shows geographically spread locations and locations in high density demand areas
- Location of hubs does not significantly influence total costs

4.2.3 Method to dispatch parts from hubs to customers

As already described, demand can be fulfilled by a regular, lateral or emergency (trans)shipment. The number of these shipments for a SKU follows from the base stock levels of the hubs. These base stock levels in turn depend on the number of hubs in the network. ***Figure 11*** shows that the number of lateral transshipments is positively related with the number of hubs. This is caused by the fact that addition of a hub increases the option to provide a lateral transshipment: The lateral transshipment array of a customer groups can increase with an extra hub. Moreover, an increase of the number of hubs does provide opportunity to decrease inventory of the hubs in case lateral transshipments are used. The relation between inventory and the number of hubs is investigated in the next subsection.

Furthermore, ***Figure 11*** illustrates that the number of emergency shipments is negatively related to the number of hubs. Recalling the relation between the service level and the number of hubs, initially the large part of the emergency shipments are used to fulfil demand of customers that are outside the coverage area of the hubs. This explains the extreme decrease in number of emergency shipments between 1 and 5 hubs. After this point, ***Figure 11*** illustrates a slowly decreasing line. There are still emergency shipments caused by either demand that falls outside the geographical coverage or by the fact that it is more beneficial than keeping more stock. This last benefit will occur due to the relatively low emergency shipment factor.

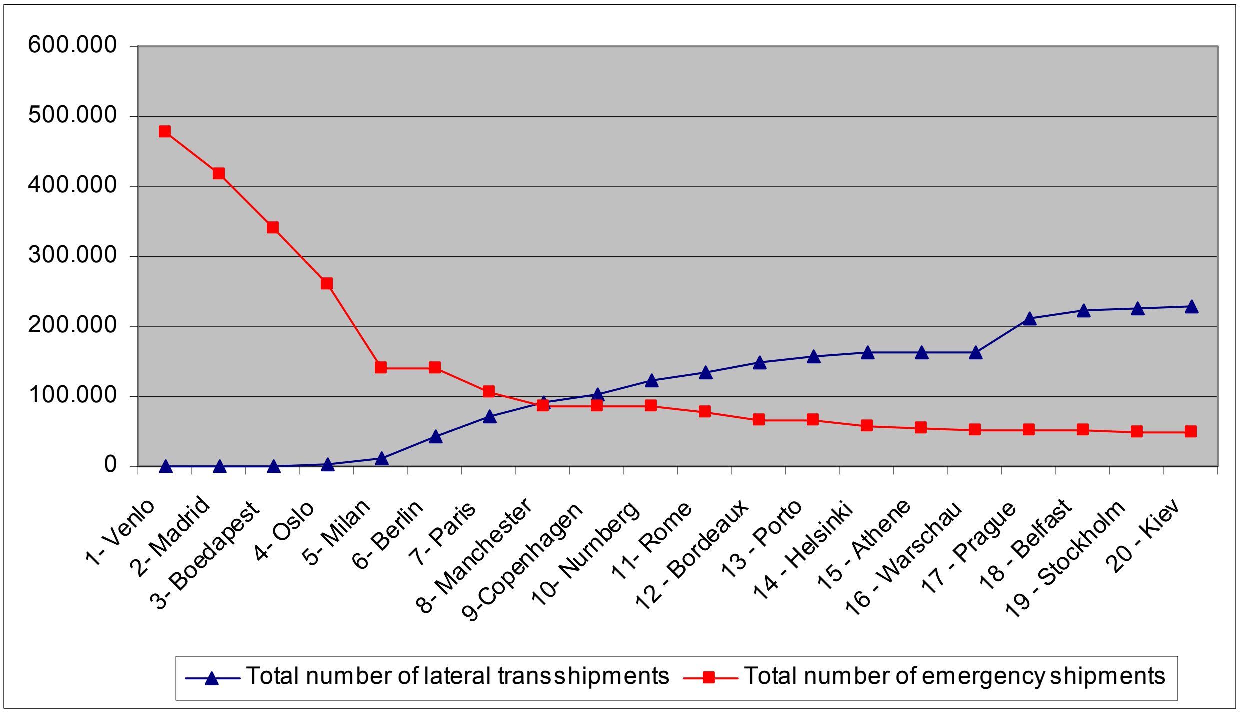


Figure 11: Influence of number of hubs on average number of lateral transshipments and emergency shipments

<u>Conclusion demand fulfilment of customers</u>

- Number of lateral transshipments is positively related to the number of hubs
- Number of emergency shipments is negatively related to the number of hubs

4.2.4 Inventory of hubs

The fourth research question concerning the inventory of the hubs is investigated by plotting the total amount of stock (in SKU units) against the number of hubs and by taking a snapshot of base stock levels of separate hubs.

The plot (***Figure 12***) of the total amount of stock against the number of hubs shows that the amount of stock first increases until the maximum is reached at about 5 hubs. The increase of inventory is caused by the addition of hubs that can fulfil all demand in their coverage area. This increases the service level significantly as shown in previous analysis. Regarding the service levels, these scenarios should not be considered because they are not feasible. After the maximum inventory point, the amount of stock gradually decreases until minimal stock at 12 hubs is reached. This decrease can be explained by the increase of lateral transshipments shown in ***Figure 11***. Lateral transshipments provide the opportunity to lower inventory because in case of a stock out occurs, there exist a relative low cost alternative. This effect is not noticeable earlier because in case of low number of hubs, the hubs are too far away (in distance) and are therefore not able to provide a lateral transshipment. After the minimal point of inventory is reached, a slow increasing line of inventory can be seen. This is a natural effect because each hub does have safety stock and therefore adding a hub has as a consequence more stock. However, it should be noted that the differences in amount of stock between 8 and 20 hubs are minimal; within 5% of the total amount of stock.

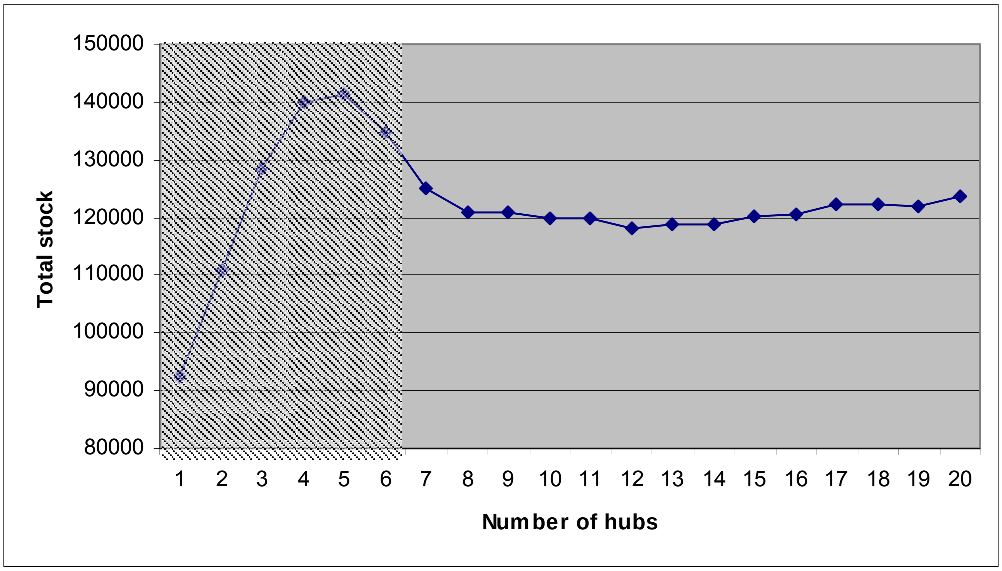


Figure 12: Influence of number of hubs on total stock

Figure 13 shows the spread of the amount of stock per hubs for the scenarios 5 to 11 hubs. One of the hubs does have the majority of the stock in all scenarios. This is Venlo, the current Central Buffer. The high amount of stock is caused by the extreme high demand areas in Germany that this hub covers. Furthermore ***Figure 13*** shows that addition of one hub does influence the amount of stock of the other hubs if they are geographically close to each other. A clear example is the addition of a hub in Paris (hub 7) which causes a significant decrease in stock of the hub in Venlo. Besides, the hub in Paris can also cover part of the south of France. This reduces the needed stock in Milan, which in the 6 scenario case had to fulfil this demand. The amount of stock of the hub in Budapest is not influenced by the addition of a hub in Paris due to the fact that the geographical coverage of the hub in Paris does not reach the geographical coverage of the hub in Budapest.

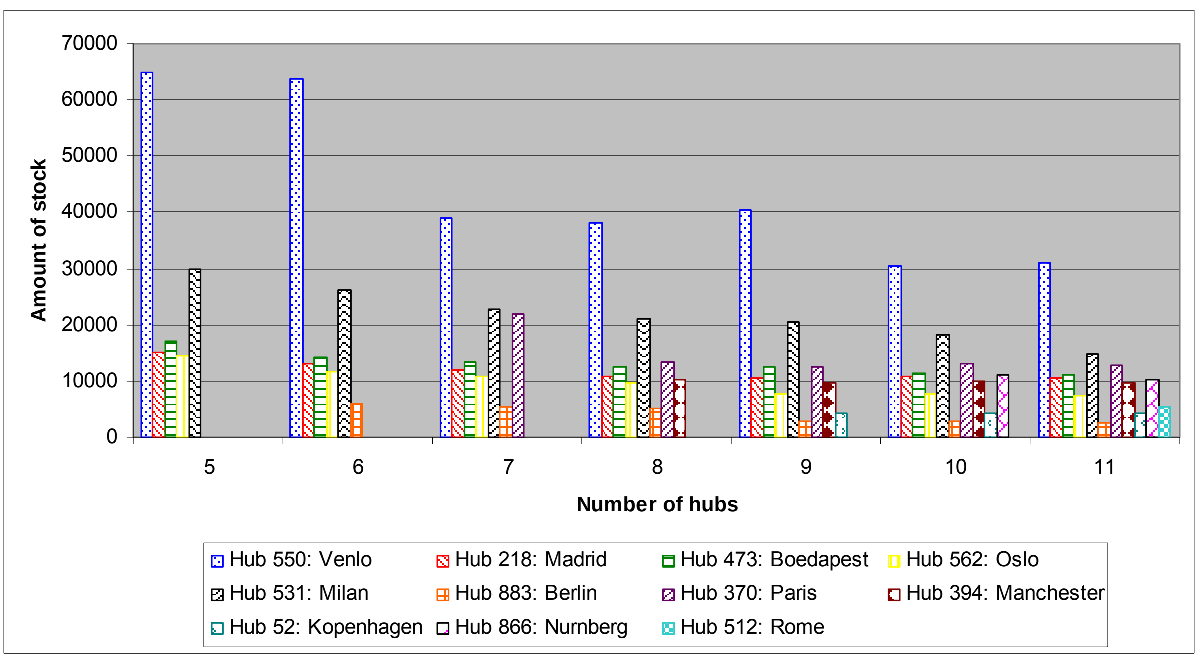


Figure 13: Influence of number of hubs on amount of stock per hub

Conclusion inventory of hubs

- The minimal amount of stock is reached at 12 hubs
- The inventory per hub depends on the inventory of other hubs that are within their geographical coverage
- In the regarded scenario the hub in Venlo has by far most stock

4.3 Differences between approach II and III

The effect of including lateral transshipments into the design approach is investigated by comparing the performance of this approach (III) with design approach II, the integrated design approach without lateral transshipments. In this section we will identify the differences in respectively total costs, amount of stock and structure of costs between the two design approaches.

4.3.1 Total network costs

Figure 14 shows the total costs and the service level for design approach II and III given a stepwise set of scenarios from 1 to 20 hubs. Based on this figure, some remarkable differences between the design approaches are noted:

1. The service level for design approach III is, for a high number of hubs, significantly higher than for design approach II. The service level of design approach II stays at the service target level which indicates that the service is 'forced' by the service constraint. So without the service constraint, the service would be lower than 90%. This also required that the difference between service levels in the third step of the optimization part of the greedy algorithm for design approach II had to be adjusted for the scenarios with more than 9 hubs because otherwise it would not meet the service target constraint.
2. The total costs of design approach III are significantly lower than the total costs of design approach II.
3. The total costs of optimal solution of the design approach II is higher than the total costs of all feasible solutions of design approach III.
4. The difference in total costs between design approach II and design approach III increases if the number of hubs increases.
5. The cost function has a different pattern can cause a difference in optimal number of hubs. Hereby the optimal number of hubs is defined as the point at which minimal costs given a set of scenarios is reached. Depending on the opening costs, the optimal number of hubs can be determined.

4.3.2 Inventory

The total cost line of design approach II is expected to increase significantly after the minimum costs are reached. Instead of this, a steady cost line can be identified from 10 to 20 hubs. This can be explained by the exclusion of opening costs of a hub. Moreover, it can be explained by the extremely high service target level that is aimed. A target service level of 90% means that almost all demand that lays within the geographical coverage has to be on stock at all times. Addition of one hub makes it possible to decrease the fill rate of the other hubs. Due to the exponential character of the fill rate related to the amount of stock and the fact that the fill rates are extremely high, a small decrease in fill rate of the other hubs saves a lot of stock. Therefore, we will analyze the amount of stock given a service target level of 80%.

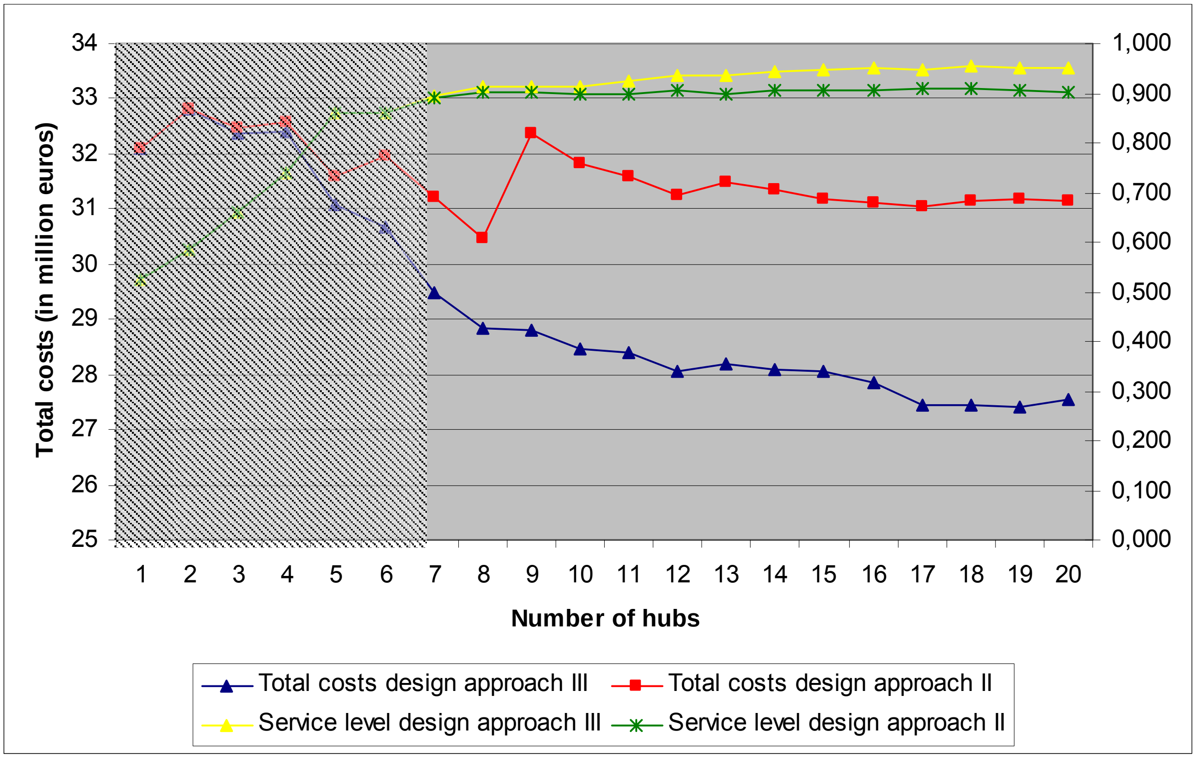


Figure 14: Influence of number of hubs on total costs and service levels design approach II and III

Figure 15 shows that the amount of stock for design approach II increases linear after the minimum is achieved at 8 hubs. Also a linear increasing line can be identified for design approach III. However, the minimum after which the linear increasing line starts is at 12 instead of 8 hubs. Furthermore, the linear increasing line has a significant smaller slope than the linear increasing line of the inventory function of the network without lateral transshipments.

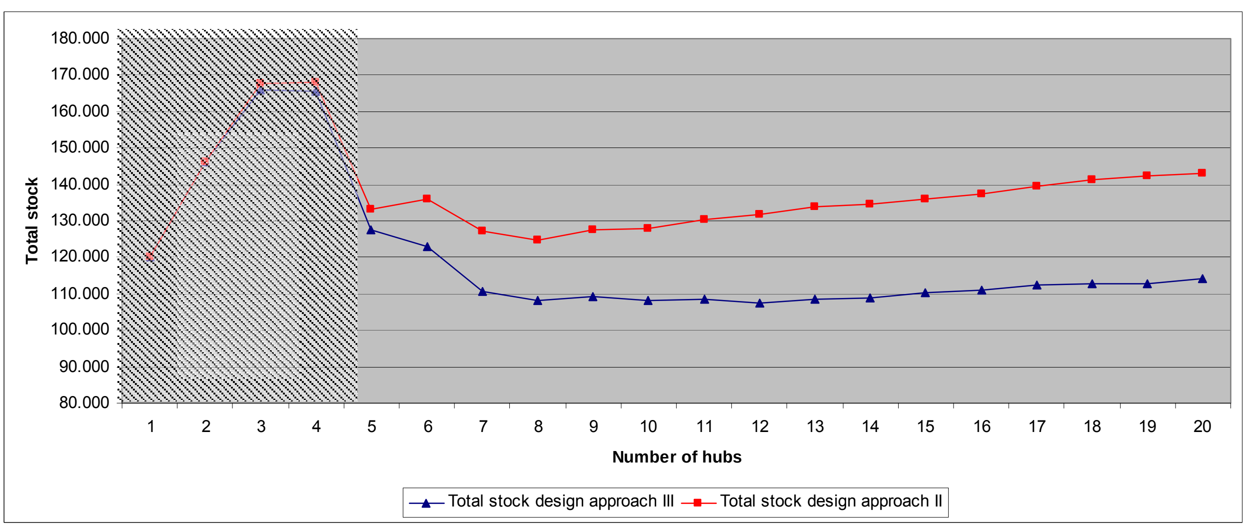


Figure 15: Influence of number of hubs on total stock target service level 80%

4.3.3 Structure of costs

To examine why the linear increasing line that is shown in the inventory plot, can not be identified in the total cost function, the structure of the total costs of design approach II and design approach III is displayed in ***Appendix N***. To provide a clear overview of the differences between the structure of transport and warehouse costs, ***Appendix N*** also shows the split of the total costs into transport and warehouse costs for both design approaches. These figures show that:

1. Total transportation costs of design approach II and design approach III are almost equal: This is caused by the fact that design approach II does not have lateral transshipment costs. However, the emergency costs of design approach II are higher, reflected in the earlier showed lower service level. The extra emergency shipment costs of design approach II are almost equal to the lateral transshipments costs of design approach III.
2. The handling costs of design approach II are lower due to the fact that these costs only have to be paid for regular or lateral (trans)shipments. Already reflected in the service level and the emergency costs, the amount of demand that is fulfilled by a regular or lateral (trans)shipments is lower for design approach III compared to design approach II.
3. The inventory costs are higher for design approach II than for design approach III. This relation was already illustrated in ***Figure 15***, the plot of the amount of stock against the number of hubs.

<u>Conclusion difference between design approach II and design approach III</u>

- Design approach III outperforms design approach II on total costs
- Design approach III outperforms design approach II on service
- Difference in total costs between design approach II and III is positively related to number of hubs
- The pattern of the cost function of design approach II and design approach III differ slightly which can lead, depending on the opening costs of a hub to a differ network design solution related to the number of hubs
- Main difference in costs between design approach II and III is caused by the difference in amount of stock and service level: design approach III has significantly less stock and difference in stock between design approach II and III is positively related to number of hubs

4.4 Differences between design approach I and design approach III

This section investigates the impact of the inclusion of the inventory costs into the design approach. Therefore, two analyses are executed. First the influence of inventory costs on the solution and total costs of design approach III is examined. The second sub section presents the differences between the output of a decoupled approach, the WSP and design approach III.

4.4.1 Impact inventory costs on total network costs

Appendix O illustrates the relation between the total costs and the total inventory and transportation costs. The figure shows that the minimal total costs are strongly related to the minimal inventory costs. So, the scenario that has minimal inventory costs, also minimal costs are reached. Hereby the minimal costs are determined by comparing scenarios with an equal number of hubs. Therefore, the inventory costs influence directly the solution of the network design.

4.4.2 Warehouse Site Planner

To investigate improvement of the current network, IBM uses the Warehouse Site Planner to determine the optimal network design. The customized tool for SPO EMEA determines the optimal network design by minimizing the total number of kilometres that have to be crossed to fulfil customer demand. Analysis of the output of this tool led IBM focus on the scenarios with 5, 6 and 7 hubs (illustrated in ***Figure 16***).

5 hubs	6 hubs	7 hubs
		

Figure 16: Optimal scenarios from the Warehouse Site Planner tool with 5, 6 and 7 hubs

To output of the WSP tool is put used as input for the scenarios to examine the difference in performance between design approach I and III. The total costs of design approach I are calculated with and without lateral transshipments. ***Figure 17*** shows that for 5 and 6 hubs design approach III outperforms design approach I. The scenarios containing 7 hubs shows that the total costs of the optimal scenario of design approach III and of design approach I with lateral transshipments are almost equal. For all scenarios there can be concluded that design approach III outperforms design approach I without lateral transshipments. Finally, there should be noted that the taking the target service level of 90% into account, not all presented solutions can be defined feasible. In general only the scenarios with 6 or 7 hubs are feasible. Moreover, the scenario with 6 hubs of design approach I is only feasible in case lateral transshipments are included.

<u>Conclusion differences design approaches I and III</u>

- Total inventory costs are strongly related to total costs. This supports use of design approach III
- Design approach III outperforms design approach I without lateral transshipments
- Design approach III outperforms design approach I with lateral transshipments for scenarios with 5 and 6 hubs
- Optimal scenario of design approach I and III for 7 hubs generates equal total costs

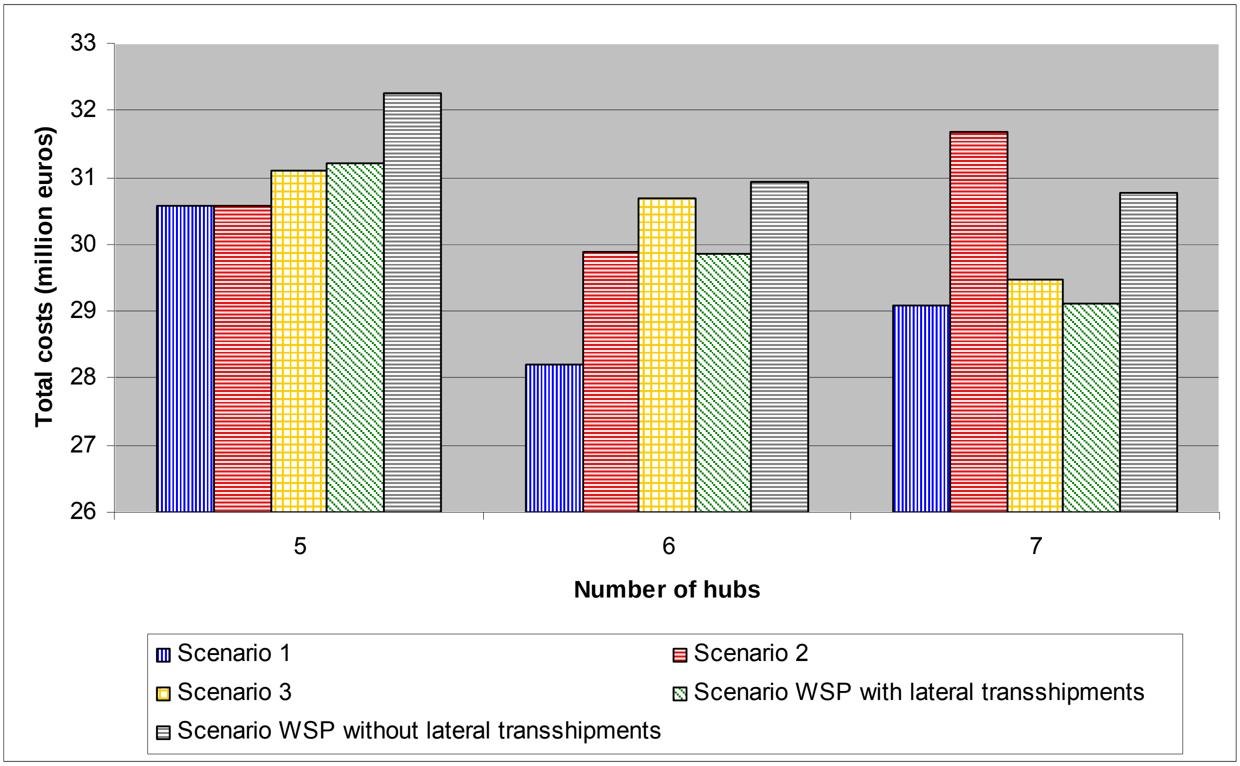


Figure 17: Influence of locations on total costs

4.5 Conclusion

The total network costs of design approach III are negatively related to the number of hubs. Moreover, the total costs function is decreasing over the whole domain. This is caused by the exclusion of opening costs of hubs and the inclusion of lateral transshipments. Due to lateral transshipments the inventory investment due to the addition of an extra hub is limited. The number and therefore impact of lateral transshipments on the total costs is positively related to the number of hubs.
The decreasing function of total costs over the total domain implies that extension of the network with hubs does not cause a huge inventory investment. This makes the network design flexible and robust for expansion of the customer area. Furthermore, it will be closer to the market that enhances the responsiveness to customer needs. Finally, this solution will be less sensitive to the network design of Logistic Service Providers.
The service level is positively related to the number of hubs. This is also reflected in the number of emergency shipments that are negatively related to the number of hubs. Furthermore, it is important to note that the location of hubs does not significantly influence the total costs.

Based on the comparison of design approach II and III, design approach III reduces total costs while increasing service level. The difference in total costs is positively related to the number of hubs which can be explained by the positively relation between the number of lateral transshipments and the number of hubs. The difference in costs is related to the amount of stock and the difference in service level.

Furthermore it is shown that design approach III benefits over design approach I because the inventory costs are strongly related to the total cost. Moreover, design approach III outperforms design approach I. Hereby the difference with design approach I without lateral is larger than with design approach I including lateral transshipments.

5. Extensions to the basic model

To reflect the IBM case even better, this chapter outlines 2 extensions that can be made to the basic model. The first section describes how a differentiation in service target level could be included into the model. Next, the option and method how to include an air hub in the model is discussed. Finally, a conclusion concerning the described extensions to the basic model is drawn.

5.1 Differentiation of service target level

Shown in the analysis, the target service level influences the base stock levels. The single-item model optimizes the base stock level for each SKU. However, the IBM business shows that the service target level does not have the same value for each SKU. There is not a clear guideline to determine the service target level but an important determinant of the service target level is the vitality level of the SKU. The vitality code of a SKU represents the criticality level of the SKU so to what extent the system fails if the SKU fails. IBM distinguishes vitality codes 1 to 5 with 1 indicating the highest criticality level. Therefore, the target service level of SKUs with a low vitality code will be higher than the target service level of SKUs with a high vitality code.
Besides the practicality of the model, a differentiation in service target levels can also be desirable based on financial considerations. Literature shows that the use of a system approach can create significant benefits. A **system approach** is a smart inventory control mechanism in that it focuses on the availability of technical systems instead of individual items. By putting something less on stock of very expensive items and something more of the cheaper items, in comparison to when single-product models are used, one can obtain the same target availability against lower costs (Rustenburg, 2000). The system approach does create most benefits in case there are significant differences in prices between SKUs. To imitate the system approach, the value of the target service level should be based on the value (in euros) of the SKU. In line with the system approach, SKUs with a high value should then have a lower target service level than SKUs with a low value. However, it should be noted that the optimization in the model is done per item and not over multiple items. Therefore, the trade off between whether a reduction of the service target level of more expensive items is beneficial is not made. This can lead to unfavourable decisions for example in case, due to the high value, the base stock level of a SKU with a low vitality level is reduced.
The differentiation of the service target level can be implemented in the model by making the service target level a characteristic of each SKU. As the SKU set is large, the determination of the target service can be based on a categorization of the SKUs. This categorization can, depending on the reasons for implementing the differentiation, be based on the vitality code or the value of the SKUs.

5.2 Air hub model

The assumption that only road transport is used for next day orders is justified because it is a flexible and low cost mode. However, in practice in some cases air transport can be beneficial as well. For example, the current Logistic Service Provider fulfils next day orders of customers in Portugal not from the hub in Madrid but from the Central Buffer in Venlo. Therefore, air transport can be interesting to take into account. The influence of air transport on the model parameters is that the value of the speed profile is much higher than of road transport. Therefore, the value of the speed profile should be adjusted in the model. As a consequence, the service time window will increase (recap: the service time window is a function of the distance and speed profile).
However, it would be unrealistic to assume that all hubs would use air transport because not all locations are near to an airport. Furthermore, the costs are based on road transport so only locations where air transport really would be beneficial (large base and flexibility) should be considered. The speed profile should thus only be changed for a specific location. As the total number of customer groups is nearly 1000 and the location of the air hub depends on the scenario, the speed profile should be changed individually in the distance matrix. This distance matrix is then scenario specific.

Another difference between the model and practice is the reach of the air hub: in the model the reach of the air hub is equal to a fixed distance but in practice distance is not the only driver for costs and availability of flights and therefore the reach will be variable and based on several factors.

5.3 Conclusion

This chapter showed that the basic model can be extended. First of all, by including target service levels that depend on the SKU, benefits could be gained. These benefits could be gathered financially by using a system approach idea. However, using a system approach could also harm the target service level for some SKUs significantly and a probably more suited option is to couple the target service level to the vitality code of the SKUs. The aim of the inclusion of differentiation of target service levels is in this case not driven financially but would increase the practicality of the model. The second described extension of the model is the inclusion of an air hub. Hereby it should be noted that the service time window is fixed and no specific air hub tariffs are taken into account. Furthermore, the air transport is constrained by many practical constraints and therefore can be seen as a coordination issue which should only be filled in for the near-to-optimal scenario. Moreover, it should be seen as a rough guideline.

6. Implementation

This chapter describes the issues that are related to the implementation of the model. It should be noted that we does not focus on the implementation of the network design model but on the implications that can be drawn from the model. Based on the steps for the design of an implementation plan by Van Aken et al. (2004) we starts with an outline of a delta analysis. This analysis identifies the main differences between the implications of the model and the current network. Next, the various stakeholders are outlined. For each of these stakeholders sources for resistance to change are described. Furthermore, the invention strategy that consists of actions that should be taken to avoid and overcome this resistance to change is presented. Finally, the practical issues of the implementation are addressed. The chapter ends with a conclusion.

6.1 Delta analysis

D**elta analysis** identifies the differences between the present situation and the designed one, so it is an analysis of what it is that has to be changed (Van Aken et al., 2004). As illustrated in ***Figure 18***, the structure between the current network and the designed network differs. The current network is hierarchical oriented and contains autonomous stocking points. The inventory of these autonomous stocking points is planned centrally. However, the country organizations do have influence on the amount of inventory that is actually present in the hubs. In the designed network not only central planning will be required but also the coordination of the inventory in the hubs has to be arranged centrally. The coordination control over the inventory of each country is therefore eliminated. Furthermore, lateral transshipments require that shipments are done cross border and so hubs are not autonomous anymore. Another difference between the current and proposed new network is that currently the base stock levels are determined for all types of orders together, whereas this study only considered (the base stock levels of) next day orders. Moreover, in the current situation the inventory of hubs is planned separately which differs from the simultaneously planning of inventory of hubs in the designed network.

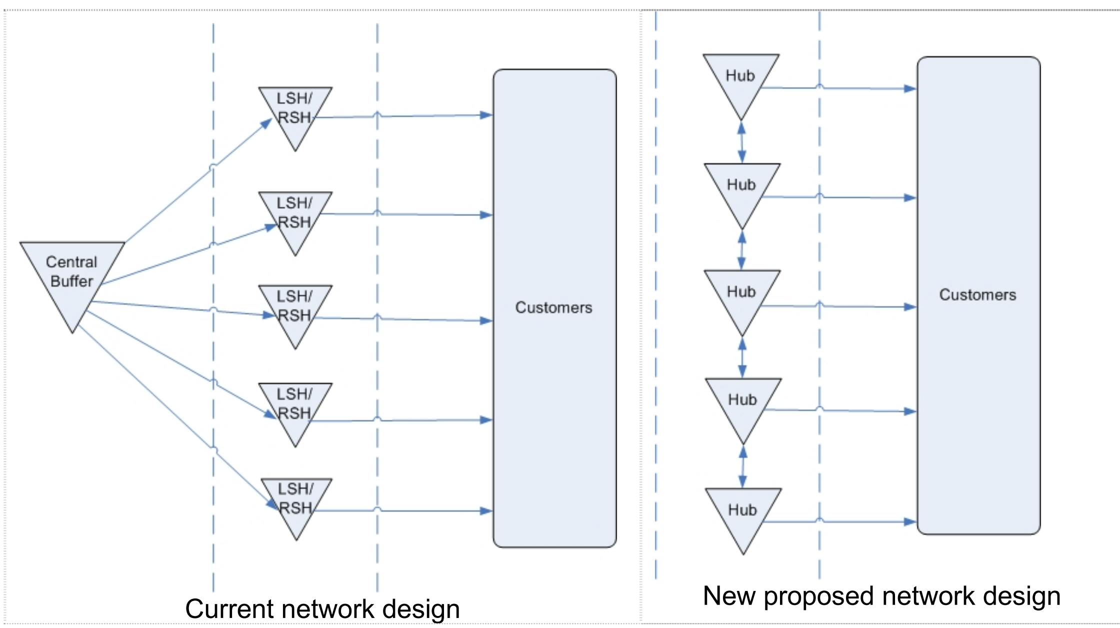


Figure 18: Differences between current and new proposed network design structure

6.2 Stakeholders

Both direct and indirect stakeholders are identified in this section. **Direct stakeholders** are defined as individuals or groups of which the work processes and/or organization roles will change (Van Aken et al., 2004). Subsequently, **indirect stakeholders** are defined as people who are to cooperate with direct stakeholders and therefore have to know about the changes in roles and processes of the latter (Van Aken et al., 2004).

- Planning department SPO EMEA: This department plans the inventory for EMEA for all types of orders. The planning method has to be changed in that it should include lateral transshipments. As already mentioned, this requires planning the inventory of all hubs simultaneously instead of planning the inventory of each hub separately. Not only the planning system has to be adjusted, also the planners should understand the concepts of the model.
- Country organizations of IBM: The country organizations are affected because the inventory has to be coordinated centrally. Furthermore, the inventory that is stocked in one country has to be available for other countries as well. Therefore, not the interest of the country has to be put on the first place, but the interest of IBM worldwide. Cooperation between country organizations has to be intensified to make sure that lateral transshipments are not blocked by political issues.
- Logistic Service Providers: Captured in the Control Tower Concept, warehousing and distribution activities are outsourced to Logistic Service Providers. Inclusion of lateral transshipments in the model requires that Logistic Service Providers are able to reach customers from different hubs. Therefore, the distribution network of the Logistic Service Provider should be flexible and the geographical coverage should be wide enough. Furthermore, it is necessary that accurate and up-to-date information about the inventory of the hubs is available. The inventory information of all hubs has to be coupled into one information system. Due to the links between the hubs that have to be strong and accurate, it might be more beneficial to use a single sourcing instead of a multiple sourcing construction. Furthermore, a contract with one Logistic Service Provider can intensify the cooperation which is essential in the highly intertwined network.
- Customers: The customers are indirect stakeholders in this implementation issue because they are in contact with the direct stakeholders. Furthermore, they will notice the performance of the network which should, based on the analysis, improve in terms of costs and service level. Moreover, it is expected that it leads to faster delivery of the parts, more flexibility and faster reaction to customers' needs.
- IBM (SPO) Management: These direct stakeholders have the decisional authority to approve or reject the design.

6.3 Sources for resistance to change

Naturally, people are resistance to change. Van Aken et al.(2004) identified 5 sources for resistance to change. These sources are first defined. Next, the sources for resistance to change per identified stakeholder (group) are described.

6.3.1 Types for resistance to change

A. Lack of understanding captures three issues: 1) one may not understand that there is a problem, or 2) one does not understand the new system or 3) one misunderstands the consequences of the change for one's own position and work processes.
B. Difference in opinion contains that one may understand the problem, but disagrees with the solution on technical and economic grounds
C. Lack of trust in the intentions or competences of the members that are part of the project to redesign the network structure

D. Low willingness to change captures that stakeholders may not want to change because they fear the unknown, or fear that they will not perform well in the new system or just don't want to loose a familiar organization environment.
E. Conflicts of interest can exist between material or immaterial interest of the various stakeholders.

6.3.2 Sources for resistance to change

This subsection will outline the sources for resistance to change that can occur at the various identified stakeholders.

Planning department SPO EMEA
The planning department may have a lack of understanding as source for resistance to change. This may be caused by the fact that the concept of lateral transshipments is relatively new and particularly known in the research area of spare parts. Currently, some companies that are also active in the spare parts business implemented this concept, but it is not a well-known concept in business. However, the aim of the planning department SPO EMEA is in line with the outcome of the analysis concerning lateral transshipments because it saves inventory costs and improves the service level.

Country organizations of IBM
Three types of sources for resistance to change are identified for the country organizations of IBM.
First of all, a lack of understanding that there is a problem and therefore the network has to be changed. Currently the performance of the network of IBM is in line with the management targets. Therefore, the country organizations may not see the need to change the network.
Secondly, the country organizations of IBM can have a lack of trust in the intentions of the members of the change project. This can be explained by the fact that the members of the change project are part of another level and part of the organizational structure of IBM. The redesign of the network encloses the whole geographical area and therefore does not take into account local details. This can be the base that the country organizations do not believe that the members of the change project do not take into account the interest of their country organization.
The third identified source of resistance to change is the conflict of interest. As already mentioned above, the country organizations are concerned about their country and not (directly) about the performance of the other country organizations of IBM. The central organization of SPO IBM that investigates the redesign of the network does take into account all countries in EMEA. Furthermore, this difference in interest is enhanced by the targets that are set by the IBM management. As the country organizations get targets for their geographical area, the use of lateral transshipments will be complicated. Lateral transshipments require intense cooperation between the different hubs in all countries. Therefore, the hubs in different countries have to collaborate and share inventory because the inventory of one country can, depending on the location, be used to fulfil demand of customers in other countries.

Logistic Service Provider(s)
The main source of resistance to change for Logistic Service Provider(s) is the conflict of interest. The aim of both IBM and the Logistic Service Providers is making profit by serving customers. However, IBM outsources the warehouse and transport activities to Logistic Service Providers. Therefore, the profit of one company can cost the other money. The new designed network would require a highly intertwined distribution network to deal with the different demand fulfilment options. This requirement causes a change in the operations of the Logistic Service Provider(s). Furthermore, the inventory level of the hubs will decrease on the short term. As IBM has to pay Logistic Service Provider(s) for holding inventory, the revenue of Logistic Service Provider(s) will also decrease on the short term. On the long term it can increase caused by the increase of the number of IBM customers.
Although the new network design may require business changes and affects the revenue of the Logistic Service Provider, the highly competitive business may force the Logistic Service Provider(s) to accept the

change. Furthermore, adapting to the needs of their customer IBM, may provide competitive advantage and therefore retrieve more customers.

Customers
It is expected that the resistance to change of customers is low because the only concern of the customers is the price of the Service Level Agreement and the quality of the service. The analysis showed that the designed network structure will increase the performance of the system. Therefore, the designed network structure only has positive influence on the issues that concern customers. The only source of resistance to change can be the low willingness to change because customers strongly rely on their information systems. Reliability of service is therefore very important to guarantee high uptime of their systems. Currently, the satisfaction of customers is on average high. A change in the service process can cause the fear at customers that it will harm the reliability of the service.

IBM (SPO) management
Two sources of resistance to change for IBM (SPO) management are identified. First of all, a lack of understanding due to the fact that the concept is new and not well-known as discussed before. This lack of understanding can lead to a misperception that the proposed network design does not contribute to the performance of the network. The second source of resistance to change is a difference in opinion. IBM (SPO) management may believe that this new design is not feasible because in practice currently cross border shipments are avoided. Furthermore, political issues like the separate country organizations could harm the use of lateral transshipments. However, if these issues are taken into account in the implementation process, the new design is feasible.

6.4 Intervention strategy

Design of the intervention strategy can be built on Tichy's TPC model (Van Aken et al., 2004). According to Tichy, one should manage organizational change processes simultaneously in three intertwined aspect systems: technical system, political system, cultural system. These three aspect systems will be elaborated below.

Technical system
The technical system is defined as the domain of technical and economic issues (Van Aken et al., 2004). In this case the technical system contains the integration of the developed tool with the current planning system. Furthermore, a good information system is required that includes all up-to-date and accurate information about the inventory levels of the hubs in the network.

To achieve this, documentation is needed that clearly defines the concepts of the new planning method. Furthermore, it should contain the input information that is needed for this. To create support for the new planning method within the planning department of SPO, training about the new planning method should be facilitated. The information system with inventory level information of all hubs in the network requires collaboration with the Logistic Service Provider(s) that perform warehouse activities. Therefore, it is essential that the Logistic Service Provider(s) participate in the project that aims to create the accurate and up-to-date inventory level information system. The participation of Logistic Service Providers creates support for the systems and during meetings information can be exchanged which vanishes the lack of understanding.

Political system
The political system is defined by Van Aken et al. (2004) as the domain of material and immaterial interests and of the formal and informal power individuals and groups may use to protect these interests. The main issue in the political system is the power of the management of IBM. They have to be convinced that the designed network is profitable and suited for IBM.

The lack of understanding that can exist at the management of IBM can be solved by technical interventions in the form of a report or presentation. However, to overcome the difference in opinion probably also political interventions are needed. Therefore, it is important that enough support is created within the organization. Furthermore, a change manager should be appointed within the organization that leads the implementation process.

Cultural system
The domain of corporate and departmental culture, of corporate, group and individual identity, and of the emotions connected with the close and repetitive interactions with other people within the organization is defined as the cultural system (Van Aken et al.,2004). As already mentioned, to implement the proposed network it is essential that the collaboration between the country organizations intensifies to enable lateral transshipments.

To overcome the sources of resistance to change for the country organizations of IBM, the benefits of the new network structure have to be made clear. This can be done by a presentation of a manager relatively high in the organizational structure. This does both includes a technical intervention because information is exchanged and a political intervention due to the power of the relatively high organizational positioned manager.

6.5 Practical issues

In this section two practical issues, the time frame of the implementation and the current inventory, are addressed.

Time frame
The contracts with the current Logistic Service Providers will expire at the end of 2009. This could be a good moment to implement the new proposed network design. In that case, one and a half year are left to prepare the implementation. This time will be needed to take the actions described in the intervention strategy.

Current inventory
The difference between the current and new proposed network design causes that inventory has to move. Furthermore, it is possible that inventory has to increase or decrease. Increasing the inventory can, depending on the lead time of the suppliers, take between 1 day and 1 year. To rebalance the inventory, the following steps should be followed:

1. Check the status of the parts in inventory: identify parts that are not ordered anymore due to a substitute part and scrap these parts
2. List for each part at each hub the surplus or shortage amount of stock
3. Move surplus inventory to hubs that have a shortage of inventory of this part
4. Do not place orders at suppliers for parts that have an surplus of total stock so the total amount of stock decreases for this part
5. Place orders at suppliers for parts that have a shortage of total stock

6.6 Conclusion

Due to the different structure between the current and new proposed network design, it is important that resources are put in the implementation of the new proposed network design. The intervention strategy that is based on a delta analysis, stakeholder analysis and the identification of sources of resistance to change can be used as a basis for the implementation. Lateral transshipments require that the country organizations collaborate strongly. Furthermore, an alliance with a LSP is desirable to make sure that all inventory information is up-to-date and accurate.

7. Conclusion and recommendations

This final chapter draws a conclusion of this project. Furthermore, recommendations are made. As this project creates value for both business practice and literature, a distinction in recommendations in these categories is shown.

7.1 Conclusion

According to the assignment, a design approach has been developed to provide insight in the optimal network design for next day orders. The network design is defined as optimal in case minimal costs are reached given that the solution is feasible, so predefined target service levels are met.
To have a benchmark for the results of the developed integrated network design approach with lateral transshipments (design approach III), two other design approaches are examined as well. These design approaches are the decoupled approach (I) and the integrated approach without lateral transshipments (II). The conclusions that can be drawn are now discussed per design question.

1. What is the optimal number of hubs?
Analysis of the results of design approach III shows that the total costs and service level are respectively negatively and positively related to the number of hubs. The total cost function is decreasing over the whole domain. This is explained by the exclusion of opening costs and the use of lateral transshipments. The optimal number of hubs for design approach II also depends on the opening costs. The total costs function does differ and the optimal number of hubs seems to be smaller than for design approach III.
Design approach I identified 5 to 7 hubs as options for a network design. Based on the total costs, 7 hubs are optimal. Using design approaches I and II as benchmark, we can conclude that the optimal number of hubs for design approach III is higher. This implies that the network design is more robust for expansion of customer area and less sensitive for the design of LSPs.

2. What is the optimal location of hubs?
Findings shows that the total costs of design approach III are not significantly influenced by the locations. In general, the scenario that contains geographically spread locations and locations in high density demand areas generates minimal costs. The locations generated by the WSP, based on a decoupled approach do not generate minimal costs.

3. What is the optimal assignment of customer groups to the hubs?
In design approach III the assignment of customer groups is based on the lateral transshipment array that is determined per customer group. Given the base stock levels of the hubs, the allocation fractions can be calculated. The number of lateral transshipments is positively related to the number of hubs because this increases the number of hubs in the lateral transshipment array. Contra dictionary, the number of emergency shipments decreases if the number of hubs increase which is reflected in a higher service level.
In design approach II, the lateral transshipment array of a customer group only contains the hub that has the minimal transportation costs to ship a part to that customer group. Compared to design approach III, design approach II has a lower service level reflected the higher amount of emergency shipments. Furthermore, the total costs of design approach II are significantly higher than of design approach III.

4. What are the optimal base stock levels of the hubs?
With a greedy heuristic the optimal base stock levels of the hubs are determined. Results of design approach III showed that the addition of a hub does not lead to huge inventory investments. The total inventory investment is significantly lower for design approach III than the inventory investment of design approach II. Hereby it should be noted that the service level attained in design approach III is significantly higher than in design approach II. So with significantly less costs, design approach III achieves significantly higher service level. Furthermore, the cost difference between design approach II and III

increases if the number of hubs increases caused by the positive relation between lateral transshipments and the number of hubs.

The main conclusions are:

- Based on the total costs, design approach III benefits over design approach I, the decoupled approach. This is supported by the finding that inventory costs are the main determinant of total costs
- The inclusion of lateral transshipments results in significantly lower network costs while achieving a significantly higher service level
- Solution of design approach III is, compared to design approaches I and II more robust for expansion of customer area and less sensitive for design of Logistic Service Providers

7.2 Recommendations

Based on the conclusions presented above, this section will outline the recommendations for IBM and recommendations for further scientific research.

7.2.1 Recommendations for IBM

Design approach for network design

Based on the performance of the three design approaches, we can conclude that the integrated design approach with lateral transshipments for network design is most beneficial. Therefore, we recommend IBM to use the developed design approach for their network designs.

Lateral transshipments

Analysis showed that the inclusion of lateral transshipments decreases inventory costs significantly. As the currently lateral transshipments are not used, we recommend IBM to include lateral transshipments in the planning of inventory by either using this model or adjusting their planning system.

Same day orders

The scope of the network design study is next day orders. Therefore, the same day contracts, orders and stores are considered out of scope. Research shows that lateral transshipments are beneficial in case of low demand and high value of parts. As same day orders do satisfy these characteristics, it is expected that including lateral transshipments in the planning of same day orders would gain inventory reductions. This network design study showed that lateral transshipments not only influence inventory costs but also influence the solution of the network design. Therefore, we not only recommend investigating the inclusion of lateral transshipments into the planning of same day orders, but also applying the integrated network design approach with lateral transshipments to the network of same day orders.

Evaluation of business cases

The developed design approach is useful to investigate the optimal network design for IBM. Furthermore, the model into which the design approach was implemented creates value in evaluation of business cases for a long time span. For example, it provides through the scenario analysis the opportunity to investigate the impact of a change in location of hubs or the impact of offers of Logistic Service Providers. Therefore, it is recommended to use the program in which the design approaches have been implemented.

Input data

As the input data determines the outcome of the model, it is very important that this information is accurate and kept up-to-date. Especially attention has to be paid to the demand information because it directly influences the base stock levels of the hub. It should be noted that all order data has to be stored accurate and complete.

7.2.2 Recommendations for further scientific research

Integrated design approach

Analysis of the results of the developed tool showed that the integrated design approach is valuable because the inventory costs are dominant for the optimal solution based on the minimal total network costs. As currently the literature about integrated design approaches is limited, it is recommended to extend research in this area.

Echelon structure

Currently the network structure of IBM is hierarchic with a Central Buffer and Local and Regional Stocking Hubs. In contra dictionary, the developed model sees all hubs in the same echelon and therefore the replenishment lead time to all hubs is equal. It is interesting to investigate the influence of variation of replenishment lead times of the hubs. This would extend the number of echelons considered in the model depending on whether a central buffer is included with 1 or 2 echelons. In practice, shipments are consolidated, so this should also be included in the replenishment decision.

Besides the replenishment of the hubs, the same day stores also have to be replenished. Due to the fact that currently these are in most cases replenished by the Central Buffer, this process was left out of scope. It is interesting to investigate whether it is beneficial to add one central buffer, select one of the hubs as central buffer, select a couple of hubs as central buffer, or to give all hubs have the central buffer function. Hereby, the effect on inventory should be taken into account.

Variation of service time window per type of orders

The model assumes one service time window for all orders that is based on a weighted average of all types of next day orders. However, this implies that for some orders the service time window is lower than their actual PDT and for other orders visa versa. To make the model and especially the base stock level determination more accurate, variation of the service time window per type of order should be included into the model. This does complicate the mathematically model significantly so research has to be done to investigate the options to include this variation. The variation of service time window could also be used to extend the model with same day orders. In this network design study same and next day day orders were decoupled. Interaction in inventory decisions between these types of orders could also be worthwhile for investigation.

Influence of part characteristic on profitability of lateral transshipments

The study shows the benefits of the inclusion of lateral transshipments on the network design solution and the total network costs. However, detailed analysis showed that the number of lateral transshipments fluctuates per SKU. This can be explained by the demand of the SKU because demand in concentrated customer groups does create more opportunity for lateral transshipments, than demand of wide spread customer groups. However, also SKU characteristics could be the cause of the fluctuations in the number of lateral transshipments per SKU. As the huge inventory reductions can be gained from lateral transshipments, it is recommended to research the influence of SKU characteristics on the profitability of lateral transshipments.

References

Alfredsson, P., and J. Verrijdt (1999). Modelling Emergency Supply Flexibility in a Two Echelon Inventory System, *Management Science* **45**, 1416-1431

Aken, J.E. van, Berends, H., Bij, H. van der (2004), *Handout Methodology for Business Problem Solving: a design-oriented and theory-based approach*, Eindhoven University of Technology, The Netherlands

Broekmeulen, R.A.C.M. (2004), *Handout Transport and Handling, course logistic 2*, Eindhoven University of Technology, The Netherlands

Candas, M.F. (2005), *Benefits of Considering Inventory in Logistics Network Design Problems with Time-Based Service Constraints*, Master thesis, University of Texas, United States

Candas, M.F., and E. Kutanoglu (2007), Benefits of Considering Inventory in Service Parts Logistics Network Design Problems with Time-Based Service Constraints, *IIE Transactions* **39**, 159-176

Cohen, M.A., Zheng, Y., and V. Agrawal (1997), Service Parts Logistics: A Benchmark Analysis, *IIE Transactions* **29**, 627-639

Corbijn van Willemswaard, P.H. (2006), *Multi-Item, Multi-Location Stock Control for the Service Parts Provisioning of DAF Trucks N.V.*, Master thesis, Eindhoven University of Technology, The Netherlands

Grahovac, J., and A. Chakravarty (2001), Sharing and Lateral Transshipment of Inventory in a Supply Chain with Expensive Low-Demand Items, *Management Science* **47**, 579-594

Huiskonen, J. (2001), Maintenance of Spare Parts Logistics: Special Characteristics and Strategic Choices, *International Journal Production Economics* **71**, 125-133

Jeet, V., and E. Kutanoglu (2005), Logistic Network Design with Inventory Stocking, Time-Based Service Allocation and Part Commonality, Working paper, University of Texas, United States, To appear

Koppes, M.M. (2007a), *Master Thesis Preparation - Literature Review in the Field of Network Design for Spare Parts*, Eindhoven University of Technology, The Netherlands

Koppes, M.M. (2007b), *Master Thesis Preparation - Research Proposal: Next Day Network Design for Spare Parts of IBM,* Eindhoven University of Technology, The Netherlands

Kranenburg, A.A. (2006), *Spare Parts Inventory Control under System Availability Constraints*, PhD Thesis, Eindhoven University of Technology, the Netherlands

Mason, S.J., Magableh, G.M., Ramasami, B., and E. Kutanoglu (2004), *Integrated Analysis of Transportation and Inventory in Intermodal Distribution Network*, Final report for MBTC Project 2041, University of Arkansas, United States

Montgomery, D.C., and G.C. Runger (2003), Applied Statistics and Probability for Engineers, New York, United States of America

Reijnen, I.C., Tan, T, and G.J. van Houtum (2008), Spare Parts Networks under Delivery Time Constraints, Working paper, Eindhoven University of Technology, The Netherlands, To appear

Rustenburg, W.D. (2000), *A System Approach to Budget-Constrained Spare Parts Management*, PhD Thesis, Eindhoven University of Technology, The Netherlands

Saccani, N., Johansson, P., and M. Perona (2007), Configuring the After-Sales Service Supply Chain: A Multiple Case Study, *International Journal Production Economics* **110**, 52-69

Sargent, R.G. (2005), Verification and Validation of Simulation Models, *Proceedings of the 2005 Winter Simulation Conference*

Shen, Z.M., and M.S. Daskin (2005), Trade-offs between Customer Service and Cost in Integrated Supply Chain Design, *Manufacturing and Service Operations Management* **7**, 188- 207

Silver, E.A., Pyke, D.F., and R. Peterson (1998), *Inventory Management and Production Planning and Scheduling*, New York, United States of America

Sommeren, R. van (2007), *Spare Parts Inventory Control for a Single-Echelon, Multi-Location model with Lateral Transshipments*, Master thesis, Eindhoven University of Technology, The Netherlands

Tan, T. (1999), *Slides Verification and validation, IE 403 SIMULATION, WEEK 9*, ATILI University, Turkey

Wong, H., Houtum, G.J. van, Cattrysse, D., and D. van Oudheusden (2005), Simple, Efficient Heuristics for Multi-Item, Multi-Location Spare Parts Systems with Lateral Transshipments and Waiting Time Constraints, *Journal of the Operational Research Society* **56**, 1419-1430

Wong, H., Houtum, G.J. van, Cattrysse, D., and D. van Oudheusden (2006), Multi-Item Spare Parts Systems with Lateral Transshipments and Waiting Time Constraints, *European Journal of Operational Research* **171**, 1071-1093

Other sources

[1] http://www.cprogramming.com/langs.html

List of concepts

Base stock policy is also known as the one-for-one replenishment policy, because once an SKU *k* is used to fulfil a customer order, immediately a new SKU is ordered to replenish the warehous**e**

Computerized model verification is defined as assuring that the computer programming and implementing of the conceptual model is correct

Conceptual model validation is defined as determining that the theories and assumptions underlying the conceptual model are correct and that the model representation of the problem entity is ‘reasonable’ for the intended purpose of the model

Control Tower Concept consist outsourced business processes due to the fact that these business processes are not seen as a core competence by management.

Data validity is defined as ensuring that the data necessary for model building, model evaluation and testing, and conducting the model experiments to solve the problem are adequate and correct

Decoupled design approach is defined as first making the strategic decisions of the network design and using this outcome as a fixed input for optimizing the tactical issues

Delta analysis identifies the differences between the present situation and the designed one

Direct stakeholders are defined as individuals of groups of which the work processes and/or organization roles will change

Echelon is a layer in the distribution network that consists of stocking points with the same function

Emergency shipment is defined as a shipment by the central source in case none of the hubs present in the lateral transshipment array of customer group *n* has SKU *k* on stock

Greedy algorithm is an approximation technique and known as the ´biggest-bang-for-the-buck´ method: it iteratively chooses the alternative that provides the ´biggest-bang-for the-buck´ until a certain stopping criterion is reached

Green field perspective means that current locations are not constraining the network design

Indirect stakeholders are defined as people who are to cooperate with direct stakeholders and therefore have to know about the changes in roles and processes of the latter

Integrated design approach is defined as simultaneously making the traditionally strategic decisions of network design and the tactical inventory decision

Internal validity is defined a measure for the amount of (internal) stochastic variability in the model

Lateral transshipment is defined as the provisioning of a part by a stocking point to a customer of another stocking point that is out of stock

Lateral transshipment array is defined as a customer group specific array that indicates the order of hubs that can fulfil demand of the customer group (and therefore satisfy the service window constraint) with the first position in the array equal to the first choice hub

Network design study is defined as designing the network with stocking points to fulfil customer demand. This requires determining the number and locations of stocking locations but also deciding inventory ordering policies and determining the method to dispatch the required parts from facilities to the customers in need (Candas and Kutanoglu, 2006).

Operational validation is defined as determining that the model's output behaviour has sufficient accuracy for the model's intended purpose over the domain of the model's intended applicability

Parts Delivery Time (PDT) indicates the percentage of customer deliveries that is delivered on time. The start of the time measurement is the call of the customer and the time stops if the requested part is delivered at the customers´ requested place.

Part Availability Level (PAL) indicates in how many cases a part request on a stock location is fulfilled (expressed in %).

Regular shipment is defined as the first option to fulfil demand by the hub that has the first position in the lateral transshipment array

Scenario analysis is defined as the process of analyzing the best suited network by considering alternative possible outcomes of scenarios. These scenarios consist of different number and locations of hubs. By using a logical scenario set, a near-to-optimal network design can be found.

Service Level Agreement (SLA) is the contract between the Original Equipment Manufacturer and the customer and states the level of service the Original Equipment Manufacturer has affirmed to its customer.

Service time window indexed by φ indicates the maximum time it can take to get the spare part from the hub to the customer

System approach is a smart inventory control mechanism in that it focuses on the availability of technical systems instead of individual items.

Verification is defined as the process to analyze whether the model is build right

Validation is defined as the process of reaching an acceptable level of confidence that the inferences drawn from the model are correct and applicable to the system and therefore it is the right model for the intended application environment

List of variables

Sets

N	Set of customer groups which centre is an option for locating a hub indexed by n
K	Set of Stock Keeping Units (SKUs), indexed by k
J	Set of locations that are selected from set N, indexed by j
R_n	Set of customers in customer group n, indexed by r

Input parameters

$P_{req,k}$	Required service level for SKU k
$\mu_{r,k}$	Demand rate of customer r for SKU k
c_j	Correction factor for shipment from hub j
h	Interest percentage for holding one unit of value/year (cost of capital:4%)
v_k	WAC value of SKU k (in euro´s)
u_j	Storage costs for storing 1 m^2 of storage space/year at location j (in euros/ m^2 /year)
o_k	Required storage space for SKU k (in euros/ m^2)
w_n	Handling costs for handling one SKU at potential location, the centre of customer group n
g_k	Weight SKU k
s	Speed profile (km/hour)
λ_{reg}	Uplift factor for regular shipments
λ_{lat}	Uplift factor for lateral transshipments
λ_{em}	Uplift factor for emergency shipments
$f\left(D_{j,n}\right)$	Fee value for a shipment which is a function of the distance between hub j and customer group n

Variables

$\mu_{n,k}$	Aggregated demand rate of customer group n for SKU k
$T_{j,n,k}^{reg}$	The transportation costs for a regular shipment from hub j to customer group n for SKU k
$T_{j,n,k}^{lat}$	The transportation costs for a lateral transshipment from hub j to customer group n for SKU k
$T_{j,n,k}^{em}$	The transportation costs for an emergency shipment from hub j to customer group n for SKU k
$\sigma_n\left(i\right)$	Lateral transshipment array for customer group n with the position in the array indexed by i
p_n	Number of locations that can reach customer group n within the service time window φ
C_k^t	Total transportation costs, consisting of regular, lateral and emergency shipments, for SKU k (in euros)
C_k^w	Total warehousing, consisting of inventory and handling, costs for SKU k (in euros)
C_k^g	Total costs for handling one SKU k (in euros)

$D_{j,n}$	Distance from hub *j* to customer group *n* (in km)
P_k	Service level of SKU *k*
S	Base stock level array
t_j^{reg}	Replenishment lead time from location *j*
$\hat{M}_{n,j,k}$	Demand from customer group n to warehouse *j* for SKU *k*
$\hat{M}_{j,k}$	Total demand at warehouse *j* for SKU *k*

<u>Output variables</u>

C	Total network costs of all SKUs (in euros)
$\beta_{j,k}(S)$	Fill rate of warehouse *j* for SKU *k*
$\alpha_{n,j,k}(S)$	Fraction of demand of customer group *n* that is fulfilled by location *j* at base stock level *S*
$\theta_{n,k}(S)$	Fraction of demand of customer group *n* that is fulfilled through an emergency shipment by the central warehouse at base stock level array *S*

<u>Decision variables</u>

Y_n	Binary decision variable that is equal to 1 if the centre of customer group *n* is selected as location of a hub indexed by *j*, and equal to 0 otherwise
$S_{n,k}$	Base stock level at potential location, the centre of customer group *n*, for SKU *k*

List of abbreviations

BO	Branch Offices
CB	Central Buffer
CE	Customer Engineer
CRU	Customer Replaceable Unit
EMEA	Europe, Middle-East and Africa
IBM	International Business Machines corporation
KPI	Key Performance Indicator
LSH	Local Stocking Hub
LSP	Logistic Service Provider
MVS	Multi-Vendor Systems
OEM	Original Equipment Manufacturer
PAL	Parts Availability Level
PCD	Personal Computer Devices
PDT	Parts Delivery Time
PSD	Printing System Devices
RO	Regional Offices
RSH	Regional Stocking Hub
RSS	Retail Storage Systems
SKU	Stock Keeping Unit
SLA	Service Level Agreement
SO	Support Offices
SPO	Service Parts Organization
SSD	Storage System Devices
WSP	Warehouse Site Planner

Appendices

Appendix A: Main environmental changes

This section describes developments in the IBM environment which cause reasons to revise the current network design. Although in reality several developments simultaneously take place, this section focuses on the two most important developments (for an extended version, the research proposal belonging to this thesis should be consulted (Koppes, 2007b). These developments refer to changes in the environment and the approaching end of the contract with current Logistic Service Providers (LSPs).
The content of this section is as follows. First, changes in the environment of the spare parts business of IBM are described. Next, the contract with the current Logistics Service Provider will be discussed.

A.1 Environmental changes

IBM is currently facing three types of environmental changes: 1) Shifting volumes in type of orders; 2) emerging markets; and 3) the trend that after-sales service is becoming a strategic asset.
In the following subsections these developments are each explained in detail.

Shifting volumes

As customers are getting increasingly demanding and systems are built more and more module-based, the critical components of high end systems are now built redundantly. These systems are called fault-tolerant systems. Fault-tolerant systems will not fail if critical parts fail, because in that case a redundant part takes over the operation. Therefore, it is less urgent to repair or replace the failed part indicated by the machine. The less urgent character of the orders increases the importance of next day orders.
However, as customers stay extremely demanding, this increase is not expected to lead to a high shift in the amount of same day and next day orders: on the short term these amount are thus expected to remain roughly the same as in the current situation.

Emerging markets

Worldwide emerging markets like China, Russia, Brazil and India are identified. Considering the region EMEA, beside Russia, also Eastern Europe is considered an emerging market. Emerging markets influence network design. Due to the fact that historical demand does not represent future demand, demand is hard to predict. Furthermore, extra stocking locations would be needed and the best suited network design, which is defined as a network at which customers are reached on time at minimal costs, could also change.

Strategic asset

A third trend is that after-sales service is becoming a strategic asset in companies' offerings (Saccani et al., 2007). IBM already recognized this: when looking at the organizational structure it becomes clear that it has a separate department for managing spare parts, the Service Parts Organization. However, competitors start to realize the importance of after-sales service too and therefore it is essential to create competitive advantage within the processes of after-sales service.

One of the core elements of after-sales service is a network to fulfil orders of customers at minimal costs and within the time constraint stated in the Service Level Agreements. Due to the fact that costs and time strike against each other, the optimal network is not straightforward. Moreover, outsourcing of activities within the network like warehousing and transport to Logistic Service Providers as already described by the Control Tower concept, diminish the transparency of the performance of the network. Therefore, it is essential but also complicated to identify the network that fits IBM spare parts environment best.

A.2 Re-contracting of network

The contract with the current Logistic Service Providers will end within one year. This creates the opportunity to outsource the activities in the network to other Logistic Service Provider(s). However, to judge the offers of Logistic Service Providers and to have a good negotiation position, it is essential to have insight into the best suited network design.

Appendix B: Geographical scope

Appendix C: General 'square' method to create customer groups

The starting point of the general 'square' method is the conversion of the postal codes of the customers that are available in the order data, into coordinates. Based on these coordinates, each country can be approximated with a rectangular and this country rectangular can be split into customer groups. The details of the creation of these rectangulars are described in the second subsection. The rectangulars are created per country because the postal code system differs per country. Finally, the assignment of customers to customer groups is described.

Conversion of postal codes to coordinates

After the gathering the demand per postal code (in case of availability), the postal codes are converted into coordinates. Converting the postal codes into coordinates had some complexities because some postal codes were not recognized and therefore not converted. This contained almost 33% of all postal codes and can be caused by either a new postal code or relative bad postal code system. This is country specific, for example France has a very complex and unclear postal code system whereas the postal code system of Germany is very clear and consistent.
In case more than 5% of the total orders of a country do have missing coordinate information, the coordinates are assigned to the order. This assignment is based on the postal code of the order with the missing coordinate information. From all orders, the coordinates of the postal code that is nearest to the postal code of the order that is not converted, is selected. Hereby it is assumed that postal codes that are closest in number are also located closest to that postal code. In case the percentage of orders with missing coordinate information is lower than 5%, the inability to convert the postal codes is considered as data invalidity and therefore this data is excluded.

Creation of customer groups

By calculating the minimal and maximum X- and Y- coordinate of each country, four points are retrieved. Drawing a horizontal line through the minimum and maximum Y- coordinate and a vertical line through the minimum and maximum X- coordinate, produces a rectangular (see ***Figure 3***).
Figure 3 also shows that the country rectangular is split in equal rectangulars. The numbers in these rectangulars correspond to the customer group number. The size of the rectangulars is at max 150 kilometres. The distance of 150 km is based on the fact that the time to reach the customers in the border of the group should be less than 3 hours which is assumed to be an acceptable time period regarding the service constraint of around 18 hours and a fixed speed profile of 50 km/hour.
The exact vertical size of the rectangulars is determined by dividing the vertical distance of the country rectangulars by the number of vertical customer groups and via versa for the horizontal size of the rectangular. The number of vertical customer groups is calculated by rounding up the division of the vertical country rectangular distance and 150 kilometres:

$K_{x,z}$ = horizontal distance of country rectangular, so maximum horizontal distance of country z

$K_{y,z}$ = vertical distance of country rectangular, so maximum vertical distance of country z

$N_{x,z}$ = horizontal number of customer groups of country z

$N_{y,z}$ = vertical number of customer groups of country z

$L_{x,z}$ = horizontal distance of customer rectangular of country z

$L_{y,z}$ = vertical distance of customer rectangular of country z

$$N_{x,z} = roundup\left(\frac{K_{x,z}}{150}\right)$$

$$N_{y,z} = roundup\left(\frac{K_{y,z}}{150}\right)$$

$$L_{x,z} = \frac{K_{x,z}}{N_{x,z}}$$

$$L_{y,z} = \frac{K_{y,z}}{N_{y,z}}$$

The calculation of the distance in kilometres for the country rectangular and for the distance between a customer group and a hub is done with the distance formula presented in section 3 of ***Chapter 2***.

Assignment of customers to customer groups

Based on the minimum and maximum X- and Y- coordinates of the country and the number of customer groups horizontally and vertically, the coordinates of the boundaries of the customer groups also can be determined. Given the coordinates of the postal codes of the customers, the customers are assigned to the customer group: The X- and Y coordinates of the customer should fall within the boundaries of the customer group.

Appendix D: Overview of customer groups

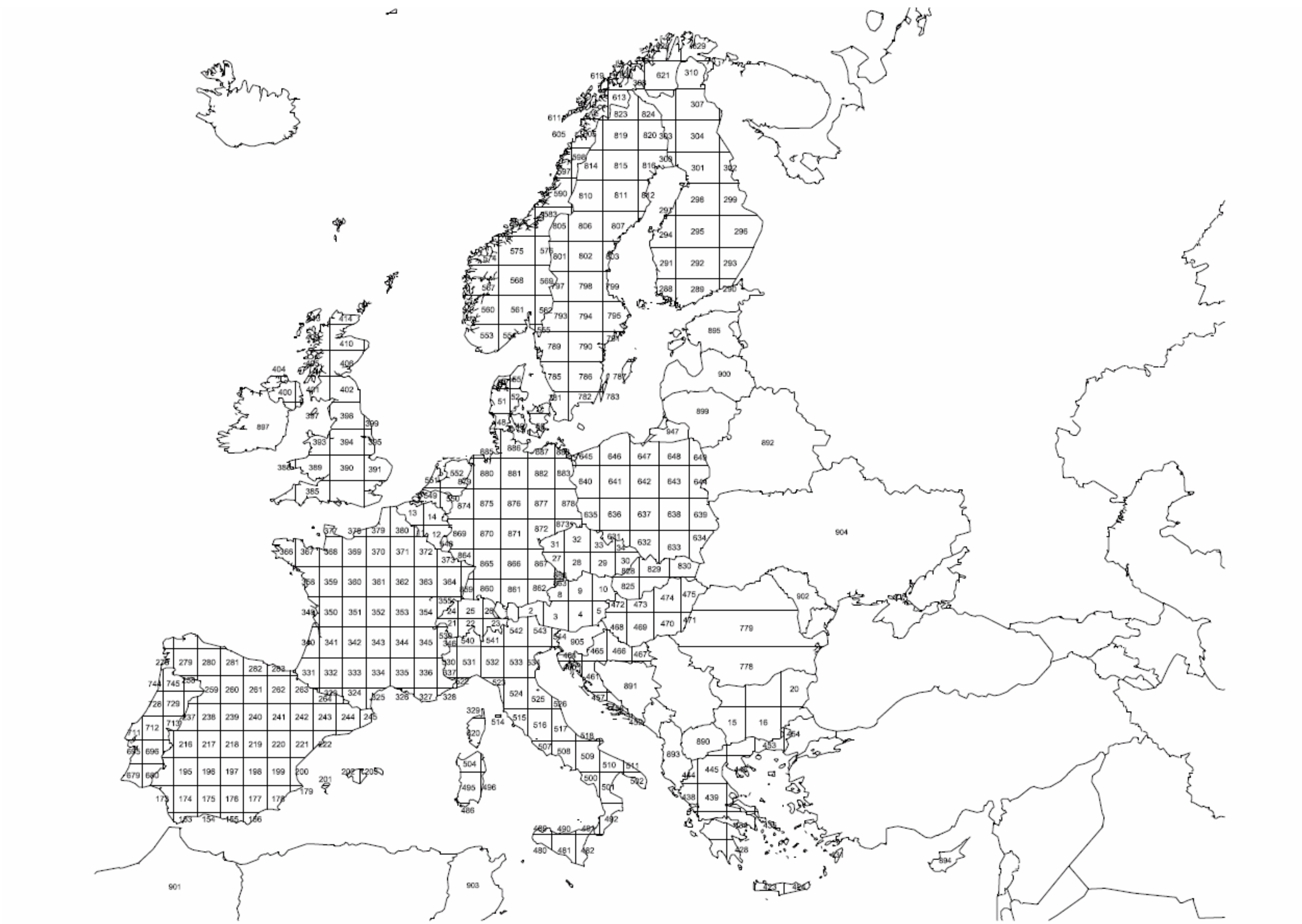


Figure 19: Overview customer groups

Appendix E: Data

Determination of centre of customer groups

In case the general ´square´ method is used to determine the customer groups, the centre of the customer groups is determined by taking the average of the minimum and maximum latitude and longitude coordinates. In case the demand is aggregated to country level, the capital of the country is used to represent the aggregated demand. The coordinates of the capitals are retrieved manually.

Customer demand data

$\mu_{r,k}$ Demand rate of customer *r* for SKU *k* per year (in number of demands/year)

As explained in ***Chapter 2***, the model assumes that the demand of SKU *k* can be described with Poisson rates. The demand rate of a customer group is equal to the sum of the demand rates of the customers that are assigned to that customer group. The demand rates of customer *r* for SKU *k* is based on historical order information of one year. From the IBM database the demand per SKU for emergency type 2 is extracted. Hereby it is assumed that next day orders are coupled to the characteristics emergency type 2. This is assumption is not completely correct, because same day orders can also have emergency type 2 code, but it does approximate the demand of next day orders best.

The aggregation of the demand of customer *r* to customer group level requires availability and conversion of postal codes to coordinates. However, for some countries the postal code availability is lacking. Analysis shows that more than 20% of the total orders in Greece, Hungary, Poland and Romania do not have postal code or address information. Therefore, it is essential to improve the completeness of all data. Recently the postal data coupled to orders has been improved for Italy because it is also an essential input for network neighbourhood, a same day demand fulfilment tool.

Set of customers

The set of customers is retrieved from order data of an historical period of 18 months. This order data contains the delivery addresses to which orders have been sent. Based on the postal codes that are part of the delivery address information, the distribution of locations of customers can be found. This method does not cover all customers because customers that did not report a failure in the last 18 months are not registered. However, from experience it is stated that each customers does have at least one order per 18 months. This does indicate that only new customers are neglected with this method. Nevertheless, it will cover most locations of customers taking into account that new customers can be located nearby ´old´ customers.

It was not possible to use information about the installed base instead of the order data because IBM only has accurate and complete installed base data of customers with same day contracts. There is a lack of installed base data of next day customers because next day contracts are not documented completely and because next day orders are also placed by customers without a service contract.

Therefore, the above described method represents the demand and thereby the installed bases of the next day order customers best.

Set of SKUs

Two criteria are used to select SKUs. The first criterion is that the SKU has a positive Weighted Average Cost (WAC) value. This WAC unit price is defined by IBM as the amount per unit which is the sum of all cost elements (like material and packaging) that becomes part of inventory costs divided by the financial quantity of a specific Weighted Average Cost level. The second selection criterion of SKUs requires that IBM has the ownership over the SKU. This is indicated in the database by the business group ‘MAI’ which is coupled to SKU identification number. The total number of SKUs that satisfy these criteria is 28.742.

Service level

$P_{req,k}$ Required service level for SKU k (in %)

The service is defined as the percentage of demand that is fulfilled by either a regular or lateral (trans)shipment. As already explained in ***Chapter 2***, this is in line with the current PDT measure. However, due to the fact that the echelon structure currently is hierarchical, it is difficult to determine the required service level for the echelon structure of the model. On average, over all product categories, the PDT measure is equal to 92%. However, this value is not set for difficult to reach countries like Cyprus. Therefore, in consultation with IBM business experts, the required service level is set equal to 90%.

Appendix F: Cost parameters

To obtain information about the cost parameters, IBM professionals are consulted. A categorization into transport and warehouse cost parameters is made.

Transportation costs

Due to the facts that no standard transport rate is available in the spare parts business and that the current transport rates are not transparent or have a clear consistent structure, it is hard to identify correct transport parameters. Furthermore, consulting Logistic Service Providers did not contribute because explaining the cost structure of their offerings harms their business. Therefore, the fee values of a shipment were calculated based on the contract of the Logistic Service Provider that serves the majority of the EMEA countries.

$f\left(D_{j,n}\right)$ Fee value for a shipment which is a function of the distance between hub j and customer group n (in euros)

The rates for next day orders depend on the weight of the shipment. The rates are therefore expressed in euro/kg. It should be noted that the minimal weight for the price of a shipment should be 2 kg, so even if the shipment is smaller than 2 kg, still 2 times the rates has to be paid for the shipment. Given the assumption that a shipment contains one SKU, the weight of the shipment is equal to the weight of one SKU. ***Table 1*** show that the transport fees do not only depend on the weight of the shipment, but also on the distance. Therefore, the transport fees in the categories (F1, F2 and F3) are based on different UPS tariffs.

To determine the transport fee for 0-200 kilometres, the average of the domestic rates of several (AT, BE, DK, FR, DE, NL, PT, ES, SE, NL) countries is taken. The transport fee of the second distance category is calculated by averaging rates of close neighbourhood shipment (NL-BE, DK-DE, BE-FR, BE-DE, PT-ES, SE-FI, SE-DK). Finally, transport fee of 400-800 kilometres is based on the average rates of shipments from Central Buffer to countries with average distance greater than 400 km (AT, DK, FI, FR, IE, IT, NO, PT, ES, SE, CH, GB).

	0-200	200-400	400-800
0-2 kg	F1 * 2	F2 * 2	F3 * 2
> 2 kg	F1 * weight item	F2 * weight item	F3 * weight item

Table 1: Transport fees

c_j Correction factor for shipment from hub j

The correction factor can be used to adjust the transport rates for changes in transport rates comparing different countries. However, as changes in rates between countries could not be made explicit, all correction factors are set equal to 1.

λ_{reg} Uplift factor for regular shipments

Regular shipments are as the name implicates the base case. Therefore the uplift of regular shipments is equal to 1.

λ_{lat} Uplift factor for lateral transshipments

As lateral transshipments are seen as a backup option because these are only used if the inventory of the first choice warehouse does not satisfy, the uplift is set equal to 1,2.

λ_{em} Uplift factor for emergency shipments

The uplift factor of emergency shipments is based on two facets. First of all, the cost of emergency shipments is the extra cost that has to be paid to execute an emergency instead of a regular shipment. Due to the large distance that has to be covered in a small time period, there can be assumed that emergency shipments are transported by plane. Therefore the difference between road and air transport is regarded. As the next available flight satisfies next day or (second) next day orders, the difference is quite small: air transport (from the current Central Buffer Venlo) is on average around 1,5 times higher than road transport. The second facet is the handling costs. The model does not calculate this cost for emergency shipments, because this demand is seen as 'lost'. However, because these costs occur in practice, the average handling costs, used for all other orders in the model, is taken into account calculating the uplift factor for emergency shipment. Based on the two mentioned facets, the value of uplift factor for emergency shipments is set equal to 2,5.

Warehouse costs

Inventory

h Cost of capital and obsolescence for holding one unit of value/year (in %/euro/year)

Similar to the numbers used in business cases of the finance department, the cost of capital is set equal to 10 percent. The obsolescence rate strongly differs per business unit and type of service. Based on the current business, a percentage of 10 percent seems a reasonable assumption for next day orders. So, in total the cost of capital and obsolescence is equal to 20% of the WAC value of SKU *k*.

v_k WAC value of SKU k (in euros)

The WAC unit price value of SKU *k* is defined by IBM as the amount per unit which is the sum of all cost elements (like material and packaging) that becomes part of inventory costs divided by the financial quantity of a specific Weighted Average Cost level. This value is like other SKU characteristics retrieved from the IBM database.

Storage space costs

u_j Storage costs for storing 1 m^2 of storage space/year at location *j* (in euros/ m^2 /year)

The storage space costs per m^2 are based on existing contracts with Logistic Service Providers. An average is taken over cost for Space in the UPS contract based on the "P13 rate per unique PN on stock". The space cost per square meter per year is calculated by dividing the total full year costs by the number of square meters in use.

A significant difference is found in tariffs between Western and Eastern European countries. Therefore, two different values of storage space costs are used. Countries that are defined as Western Europe are: Austria, Belgium, Denmark, Finland, France, Germany, Great Britain, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden and Switzerland. Eastern Europe contains the remaining countries: Belarus, Croatia, Cyprus, Czech Republic, Estonia, Greece, Hungary, Latvia, Lithuania, Macedonia, Moldavia, Poland, Romania, Slovak Republic, Slovenia and Ukraine.

o_k Required storage space for SKU *k* (in euros/ m^2)

The required storage space is determined by multiplication of the length and width of SKU *k*. These dimensions are retrieved like all SKU characteristics, from the IBM database.

Handling costs

w_n Handling costs for handling SKU *k* at potential location, the centre of customer group *n* (in euros)

Handling costs consists of both inbound and outbound tariffs for one product. As currently SKUs are consolidated, the current inbound tariffs have to be adjusted. Assuming on average three parts are replenished, the current inbound tariffs is divided by three to correct for the consolidated shipment. Therefore the existing tariffs of Logistics Service Providers for inbound divided by three and outbound is used to set the handling cost tariff, equal for each SKU k.

Appendix G: Differences between underlying model of evaluation algorithm of Kranenburg (2006) and Reijnen et al. (2008)

Kranenburg (2006) investigated the effect of lateral transshipments on inventory decisions. To quantify the effect, he developed a greedy algorithm to determine the base stock levels of stocking locations including lateral transshipments. As there exists some differences between the underlying model of Kranenburg (2006) and the model presented in this master thesis, Reijnen et al. (2008) adjusted the evaluation part of the algorithm to suite it to the model presented in this thesis. This appendix describes the differences between the underlying model of Kranenburg (2006) and the model presented in this master thesis and the basis of the evaluation algorithm of Reijnen et al. (2008).

The first difference is that in the model of Kranenburg (2006) each stocking locations has a fixed set of back up locations. So, in case a stock out occurs, a fixed set of back up locations will be checked. This set contains all stocking locations that are able to provide a lateral transshipment. The ability to provide a lateral transshipment is defined beforehand and defined as a 'main warehouse'. Contra dictionary, in the model presented in this master thesis, the set of back up locations is customer group specific indicated by the lateral transshipment array. Therefore, the order and number of back up locations can differ per customer group.

Furthermore, two smaller differences can be identified between the model of Kranenburg and the model presented in this master thesis. First of all, Kranenburg (2006) assumes that the costs parameters are equal for all stocking locations. As IBM experiences differences in costs regarding the geographical location, taking into account differences in costs parameters is essential for a network design. It is thereby expected that this difference in costs parameters influences the solution of the network design. The second small difference is that Kranenburg (2006) distinguishes between main and local warehouses where main warehouse are only able to provide lateral transshipments. In the model presented in this thesis no distinction is made between hubs: All hubs (within the service time window) are able to provide a lateral transshipment. Therefore, all hubs can, according to the model of Kranenburg (2006), be defined as main warehouses.

Appendix H: Verification and validation

To make sure that the proposed solution is correct, it has to be checked whether the model is build right (**verification**) and whether it is the right model for the intended application environment (validation). **Validation** is the process of reaching an acceptable level of confidence that the inferences drawn from the model are correct and applicable to the system (Tan, 1999). Therefore it has to be determined whether simplifications made in the model do not cause too large errors. Hereby the system is defined as the environment in which the model is intended to be used, so in this case the spare parts business of IBM. Sargent (2005) identified verification and several types of validation in the model development process. Illustrated in ***Figure 20*** a conceptual model is made based on the system. In this study the system represents the spare parts environment of IBM (***Chapter 1***) and the conceptual model is equal to the model presented in ***Chapter 2***. As described above, this model that consists of a main and sub model is programmed in Delphi. Furthermore, a simulation is added to the main model to imitate practice. This program is equal to the computerized model in ***Figure 20***. To make sure that the translation from the system to the conceptual model and to the computerized model is made correctly, the verification and validation process is necessary. The following subsections will present the different types of verification and validation shown in ***Figure 20***. Furthermore, the internal validity of the simulation tool is assessed.

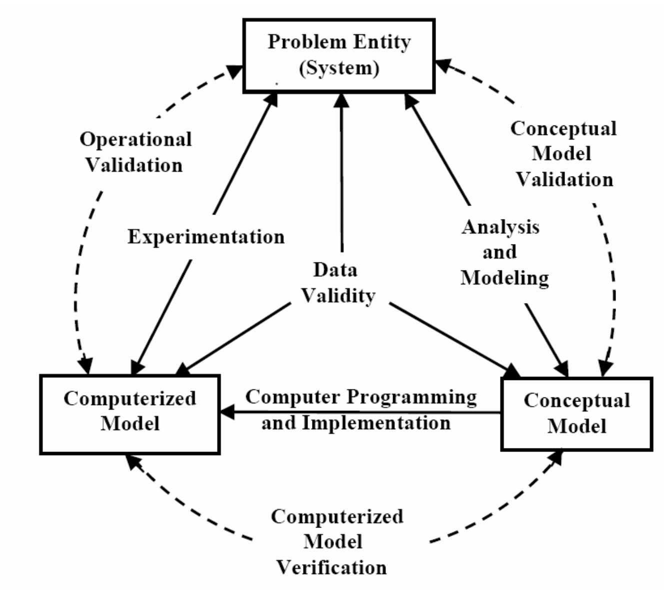


Figure 20: Verification and validation in the model development process

Verification

Computerized model verification is defined as assuring that the computer programming and implementing of the conceptual model is correct (Sargent, 2005).

The verification process started with a walkthrough and test runs regarding one SKU and one scenario. From here, the model was extended so more SKUs and scenarios were considered. Furthermore, the computerized model was tested through examination of the effects of extreme parameter settings. The simulation part of the program was verified by testing the difference between the generated demand rates of the simulation tool and the given input demand Poisson rates. Analysis showed a difference of less than 3%. Moreover, the difference between the costs per year calculated by the simulation and the costs based on the output of the evaluation part of the greedy algorithm are less than 2%. Finally, a parameter sensitivity analysis is conducted to test the actions and outcomes of the simulation tool (***Appendix I***). It shows that from this analysis there can be concluded that the deviation of the total costs is never higher than the deviation of the input parameters

Conceptual model validation

Conceptual model validation is defined as determining that the theories and assumptions underlying the conceptual model are correct and that the model representation of the problem entity is 'reasonable' for the intended purpose of the model. This is achieved by discussing the assumptions with IBM professionals.

Internal validity

Stochastic models possess variability due to the internal stochastic variability. To measure the amount of (internal) stochastic variability in the model, several replications called runs of the model are made. The runs are used in two different tests to assess the internal validity. Therefore **internal validity** is defined a measure for the amount of (internal) stochastic variability in the model. To assess to internal validity the influence of a warm up period on the results and the consistency between two similar runs is examined. ***Appendix J*** shows that the exclusion of a warm up period does not harm the results. Furthermore there can be concluded that the results of the runs are consistent and the variation less than 1% of the average total costs of the runs.

Data validity

Data validity is defined as ensuring that the data necessary for model building, model evaluation and testing, and conducting the model experiments to solve the problem are adequate and correct (Sargent, 2005). To avoid data inadequacy or incorrectness, the used input data are retrieved directly from the IBM database. In case the data was not available in the database, IBM professionals checked the used input data.

Operational validation

Operational validation is defined as determining that the model's output behaviour has sufficient accuracy for the model's intended purpose over the domain of the model's intended applicability.
To assess the operational validation, the current inventory investment is compared with the calculated inventory investment of the model. Furthermore, the base stock levels that are determined by the model are compared with the current base stock levels. To approach the current business as close as possible, the computerized model is run without lateral transshipments and as input the scenario containing the current LSH/RSHs is used. However, it should be noted that a one on one comparison is not possible due to:

1. Difference in network structure. The use of lateral transshipments can be excluded from the model but still no hierarchical structure is assumed. The model assumes that all hubs have the same function.
2. Difference in determination of base stock levels. The model determines the base stock levels for next day orders. However, in the current business one base stock level for each SKU is set so no distinction is made between same day and next day order demand.

<u>A. Comparison in inventory investment</u>

To enable a sanity check and get some feeling whether the model is in line with practice, the required inventory investment according to the model is compared with the current inventory investment. Hereby it is assumed that the ratio between same and next day demand is equal to 4:1. Therefore, the cost of capital and obsolescence of the model excluding lateral transshipments are multiplied by 5. This calculation of the total inventory investment of the model is around 19.4 million. The current inventory investment for next day demand is determined by selecting all urgency 2 and 3 orders. The calculated current inventory investment is equal to 15.5 million. However, the calculated inventory investment of 15.5 million excludes the base stock levels that are determined manually. The total inventory investment of manually set orders, of both same and next day orders, is around 55 million. So, the calculated inventory investment is higher than 15.5. The exact value is, according to IBM professionals hard to determine. Taking the manually determined base stock levels into account and the fact that the model does not completely reflect the current network design and planning method there can be concluded that inventory investment calculated by the model is realistic.

B. Comparison in base stock level values

The current base stock levels of 50 selected SKUs were compared with the base stock levels of the model excluding lateral transshipments. Analysis showed that differences occurred, but these can be explained by three main facets:

1. The requested quantity of a part is on average higher than 1. The assumption of the model is that a failure of 1 part causes a failure of a machine. Replacement of 1 part is therefore sufficient to repair the machine. However, in practice also demand of more than 1 part occurs. This can be caused by corrective maintenance demand. In this study only preventive maintenance is regarded.
2. The planning tool of IBM optimizes the base stock levels per country where this model makes the stock plan for the whole geographical area and does not differentiate per country
3. The planning tool of IBM takes into account other factors like the importance of the part or product group

Based on the complete comparison there can be concluded that the model does set logical base stock levels.

Conclusion

Based on the complete verification and validation process there can be concluded that the model can be considered robust and practical.

Appendix I: Verification - Parameter sensitivity analysis

Figure 21 illustrates the effect of a change of the input parameters on the total costs. The input parameters deviate from -100% to +200%. For example if the replenishment lead time is raised with 100%, the model reacts on this change with an increase of 10% of the total cost. Comparing the various input parameters, the total costs are most sensitive to the transport cost. Furthermore, ***Figure 21*** shows a positive relation between the deviation of the parameters and the deviation of the total costs. This illustrates that the model behaves naturally because if input cost decreases, it is logical that the total costs also decreases. Regarding the replenishment lead time, if the replenishment lead time decreases, less stock has to be kept which causes a decrease in inventory and in turn a decrease in total costs. Finally, there can be concluded that the model is quite robust because the deviation of the total costs is never higher than the deviation of the input parameters. Moreover, the highest deviation of total costs in equal to 85% by a deviation of the transportation costs with 200%.

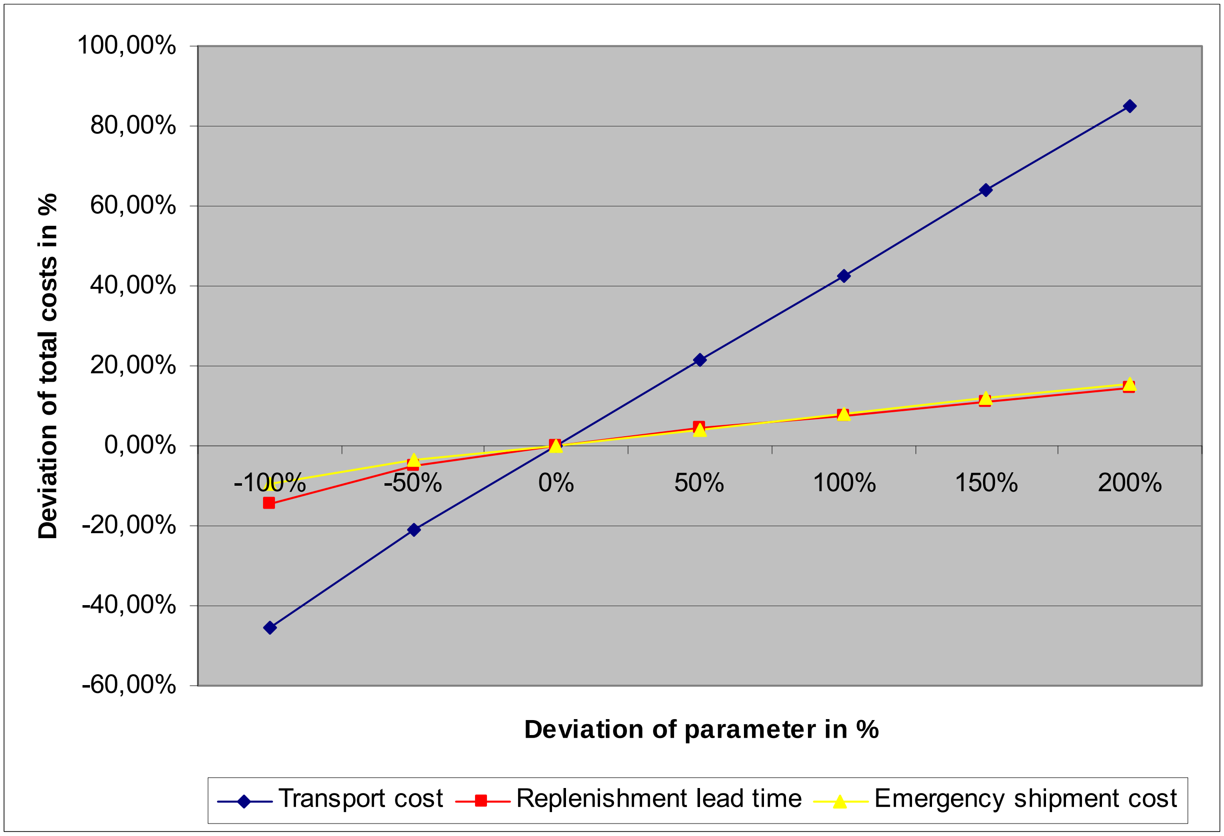


Figure 21: Parameter sensitivity analysis based on total costs

Appendix J: Validation - Internal validity

Internal validity is defined a measure for the amount of (internal) stochastic variability in the model. To assess to internal validity the influence of a warm up period on the results and the consistency between two similar runs is examined.

Influence warm up period

Figure 22 shows the impact of a warm up period on the deviation of the value of the outcomes. A warm up period is defined as the period from the start until a predefined time moment, with days as time unit. The values of the outcomes of the warm up period can differ due to the fact that the system is not yet in a steady state. However, ***Figure 22*** shows that no trend can be identified in the plot of the value of the outcomes and the sub-runs that follow after each other in time.

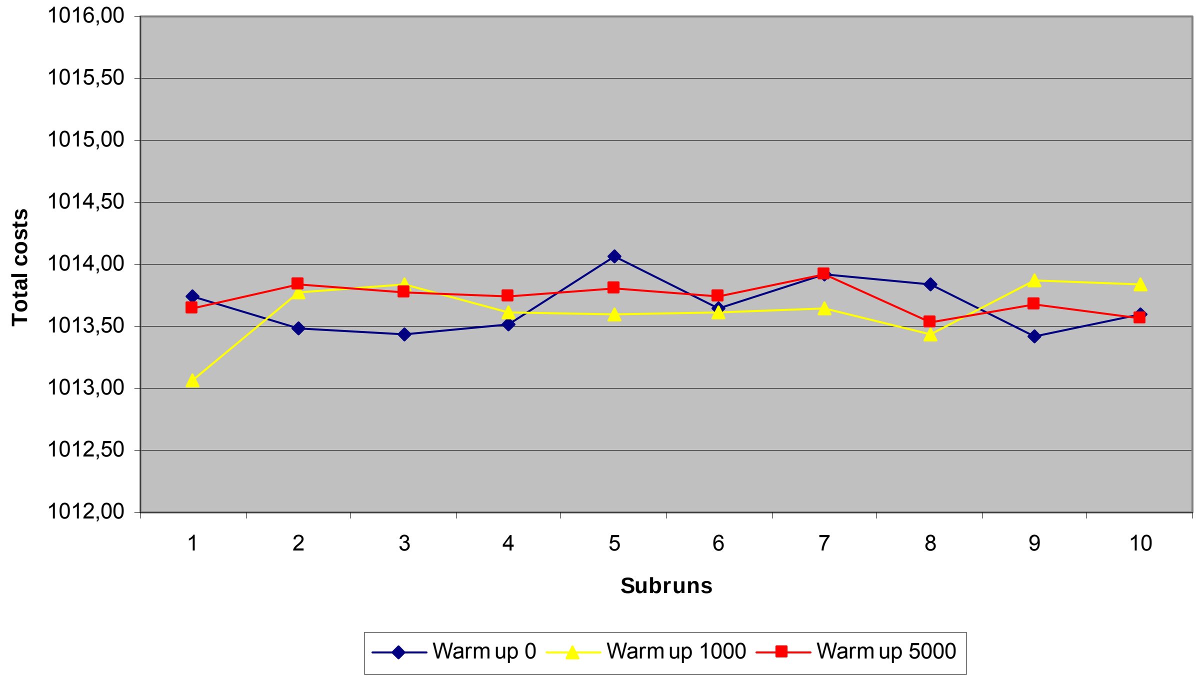


Figure 22: Impact of time and warm up period in simulation on results

To simultaneously illustrate the features of a data set, such as centre, spread, departure from symmetry, and identification of unusual observations or outliers, a box plot is used. ***Figure 23*** shows that a box plot displays three quartiles, the minimum, and the maximum of the data on a rectangular box, aligned vertically. When an ordered set of data is divided into four equal parts, the division points are called quartiles. The first or lower quartile, q_1, is a value that has approximately 25% of the observations below it and approximately 75% of the observations above. The second quartile, q_2, is a value that has approximately 50% of the observations below it and approximately 50% of the observations above. The second quartile is exactly equal to the median. The third or upper quartile, q_3, is a value that has approximately 75% of the observations below it and approximately 25% of the observations above.
The box indicates the variability by enclosing the inter-quartile range with the lower edge at the first quartile, q_1, and the upper edge at the third quartile, q_3. A line is drawn through the box at the second quartile (which is 50^{th} percentile or the median). A line, or whisker, extends from each end of the box. The lower whisker is a line from the first quartile to the smallest data point within 1,5 inter-quartile ranges from the first quartile. The upper whisker is a line from the third quartile to the largest data point within

1,5 inter-quartile ranges from the third quartile. Data further from the box than the whiskers are plotted as individual points. A point beyond a whisker, but less than 3 inter-quartile ranges from the box edge, is called an outlier. A point more than 3 inter-quartile ranges from the box edge is called an extreme outlier (Montgomery and Runger, 2003).

Figure 23 illustrates the median of the total costs increase if the warm up period increases. However, this difference is minimal showed by the fact that all inter-quartile ranges lays between the same lines. Furthermore, it should be noted that the box does contain less than 1 euro which is, related to the total costs, less than 1%. Besides the small increase in median, ***Figure 23*** also shows a decrease in variation when the warm up period increase. Hereby, it should be noted that the second sample with a warm up period of 1000 days has an outlier. Again, taking into account the range of the y-axis, there can be concluded that the variation is minimal. Moreover, there can be concluded that exclusion of a warm up period does not harm the results.

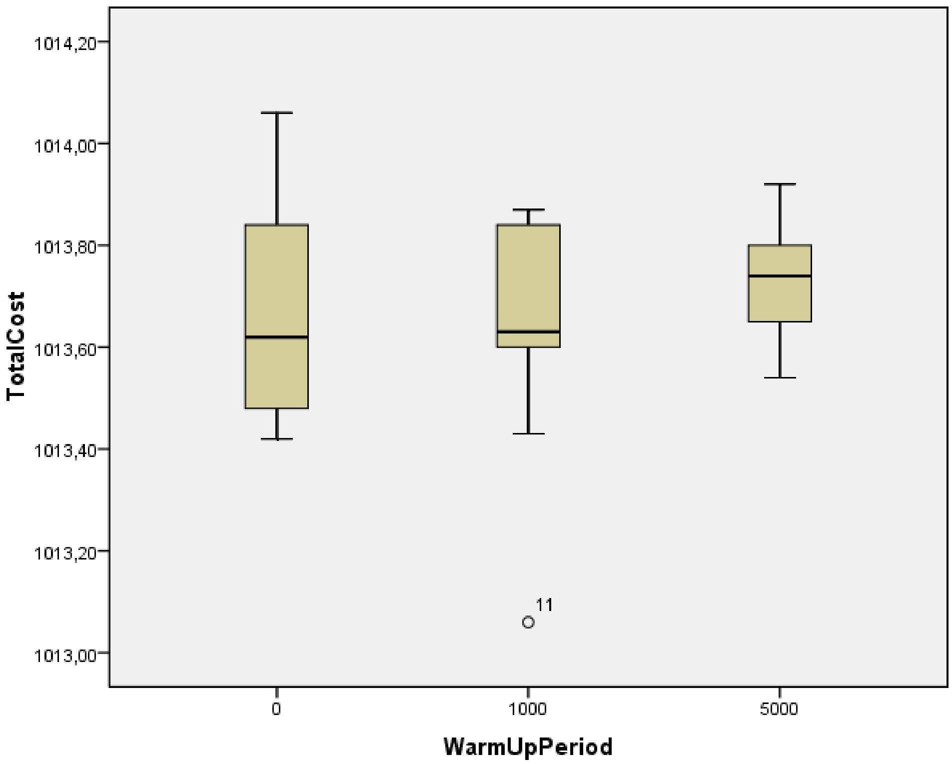


Figure 23: Results measured in total costs of runs with different warm up period

<u>Consistency between runs</u>
To check the consistency between runs, two runs are compared with each other. Each run is divided based on time into 10 sub-runs. The results of these sub-runs are displayed in the box plots in ***Figure 24***. It shows that the median of the second run lies lower than the first one, but falls within the inter-quartile of the first run. Another difference is that the box of the second run is smaller. However, taking the whiskers into account, the variation between the sub-runs is almost equal. Regarding the values on the y-axis there can be concluded that the results of the runs are consistent and the variation less than 1% of the average total costs of the runs.

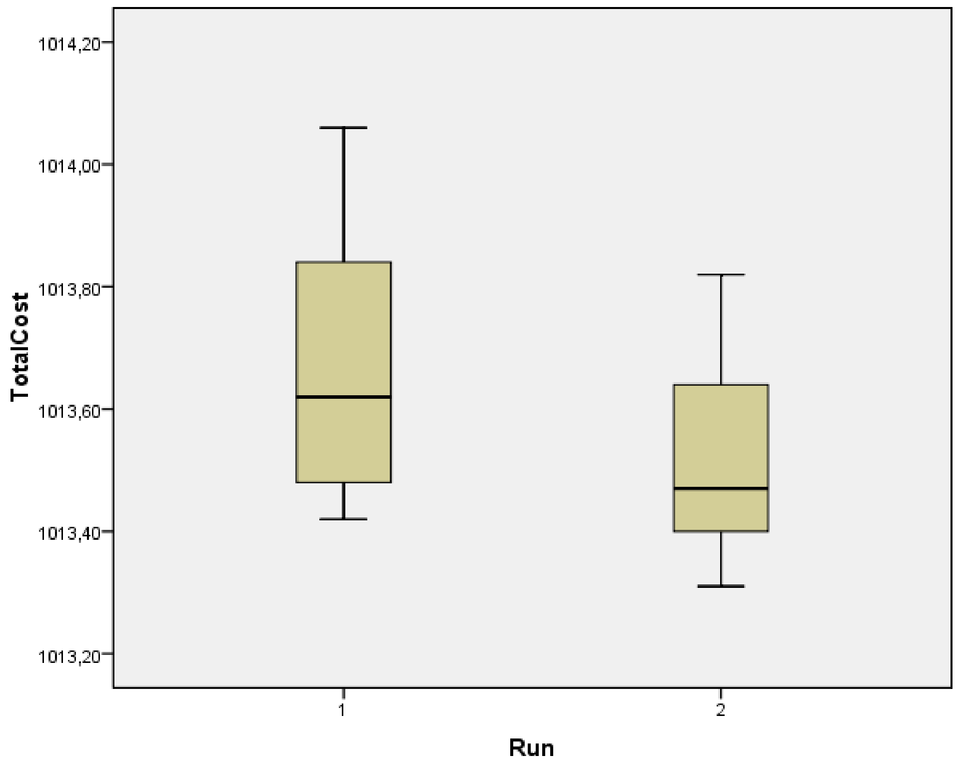


Figure 24: Difference in total costs between runs with no warm up period

Appendix K: Implementation of model in software

Both the main and sub model are implemented in software. To imitate practice for the scenario analysis, a simulation tool is created. This section first describes the choice of programming language. Next, it presents the structure of the program. The third section gives an overview of the input and output parameters. Finally, some usability characteristics of the program are addressed.

K.1 Programming language

Given the size of the problem (nearly 30,000 SKUs, 1000 customer groups and potential locations) a programming language is used to model the problem. Delphi, an object oriented programming language is used for several reasons. First of all, the greedy algorithm of Kranenburg (2006) is also programmed in Delphi which allows using part of the program. Furthermore, this provides the opportunity for further research at the TU/e regarding algorithms to approach problems in spare parts environments. The third reason is that the TU/e provides student assistance for programming in Delphi. The disadvantage is that IBM does not have a license to use Delphi and knowledge about this programming language. However, IBM SPO does have experience with the programming language Java. As both Java and Delphi are object oriented programming languages, there exists only a small difference between the two programming languages. The main difference is the traditional procedural language on which they are based on: Java on C and Delphi on Pascal [1]. Therefore, Java uses more symbols than the more textual based Delphi. Due to the small difference between the two programming languages there can be concluded that a lack of Delphi skills does not have to be a big concern for the implementation of the program.

K.2 Structure program

The structure of the program is similar to the structure of the model presented in ***Chapter 3***: it consists of 2 parts, the main and sub model. The results of the main model are simulated. The simulation uses the output of the greedy algorithm, the base stock levels of the hubs, are used as input. Therefore the initial base stock levels are set equal to the output of the greedy algorithm. Based on the Poisson demand rates per SKU and customer group, events are generated that represent the failure of a SKU at a customer. Next, the demand is fulfilled in order of the indicated lateral transshipment array including all hubs within the service time window of the customer group, so the closest hub is set at the first position in the array. In case the first choice hub does have inventory, a regular shipment takes place. Otherwise, if one of the other hubs in the lateral transshipment array does have inventory a lateral transshipment is initiated. The last demand fulfilment option is an emergency shipment if none of the hubs in the lateral transshipment array does have stock. The inventory levels of the hubs fluctuate due to fulfilment of demand and the simulation of a replenishment lead time of 7 days. Therefore, the option that is taken to fulfil demand can change in time.

Each generated demand event causes transportation costs (either for a regular, lateral or emergency (trans)shipments) and handling cost. The storage costs and cost of capital and obsolescence are, because these are dependent on the inventory, calculated per year. Hereby the assumption should be recalled that both physical and pipeline stock are defined as inventory over which costs have to be paid.

The simulation is run for a period of 100,000 days which can be considered acceptable regarding the average demand rate of 5 demand events per year. Furthermore the simulation period is justified by the verification and validation process described in ***Appendix H***. Details of the verification and validation process are presented in the next section.

K.3 Input and output parameters

Input data that have to be selected to run the program are files of SKU characteristics, demand per SKU per customer group, scenario, hub characteristics and distance matrix. Furthermore, some parameters can be changed in the input screen. This section starts with elaboration of this input data and ends with the output data.

The SKU characteristics and demand per SKU per customer group are text files. The SKU characteristics file is converted once to speed up the read in of the data.

Besides the SKU characteristics file, the distance matrix is also converted. The distance matrix is a huge Microsoft Excel file that contains all distances from a customer group to each customer group. Furthermore it contains the dummy variable that indicates whether the hub falls within the service time window of the customer group. The program converse this file once into a text file that contains the lateral transshipment array for each customer group. This lateral transshipment array contains all hubs that are reachable from the customer group. For each scenario a specific lateral transshipment array can be created that selects the hubs belonging to the scenario from this huge lateral transshipment array.

Hub characteristics like the storage space or handling costs and the scenarios are captured in an Excel file. It should be noted that the content of the files can be changed, but the lay-out of the files should stay the same.

As illustrated in ***Figure 25***, input parameters that can be changed in the input window are the uplift factor of lateral and emergency (trans)shipments, the percentage of cost of capital and obsolescence, replenishment lead time, transport fees, length of the simulation run, target service level and the difference in service levels (mathematical parameter used in optimization step of the algorithm). Furthermore, some checkboxes are included that are each described below.

The influence of the inclusion of step 3 of the optimization algorithm, the service constraint, can be investigated with the checkbox 'service level constraint'.

In case the checkbox 'fixed seed' is marked, the analysis is done with the same random numbers as the previous run. This reduces the fluctuations in results due to the generation of random numbers.

To enable analysis of the influence of lateral transshipments, the program has a check box as input value. In case the checkbox is marked, lateral transshipments are included in the model. Excluding the lateral transshipments from the model is achieved by jumping over step 3 to 5 of the evaluation part of the greedy algorithm. The result of step 6 for hubs with higher position in the lateral transshipment array than 1, the alpha value will be equal to 0.

The checkbox 'sub runs' is used for tests to verify and validate the computerized model. As it is shown in ***Appendix H*** that the model is verified and validated, this checkbox should not be marked during runs.

The output of the program consists of two text files. The first one shows the average costs per SKUs split into the different types of network costs. Furthermore, it shows base stock levels of the hubs, the fill rates of the hubs and the overall service level of the network. The second text file shows the demand fulfilment per customer group. First the number of the customer group and the total demand is displayed. Next, the hubs that are present in the lateral transshipment array are shown with the first choice hub put in front. Each customer group line ends with the code of the central source '0000' with the number of demand that has been fulfilled with an emergency shipment.

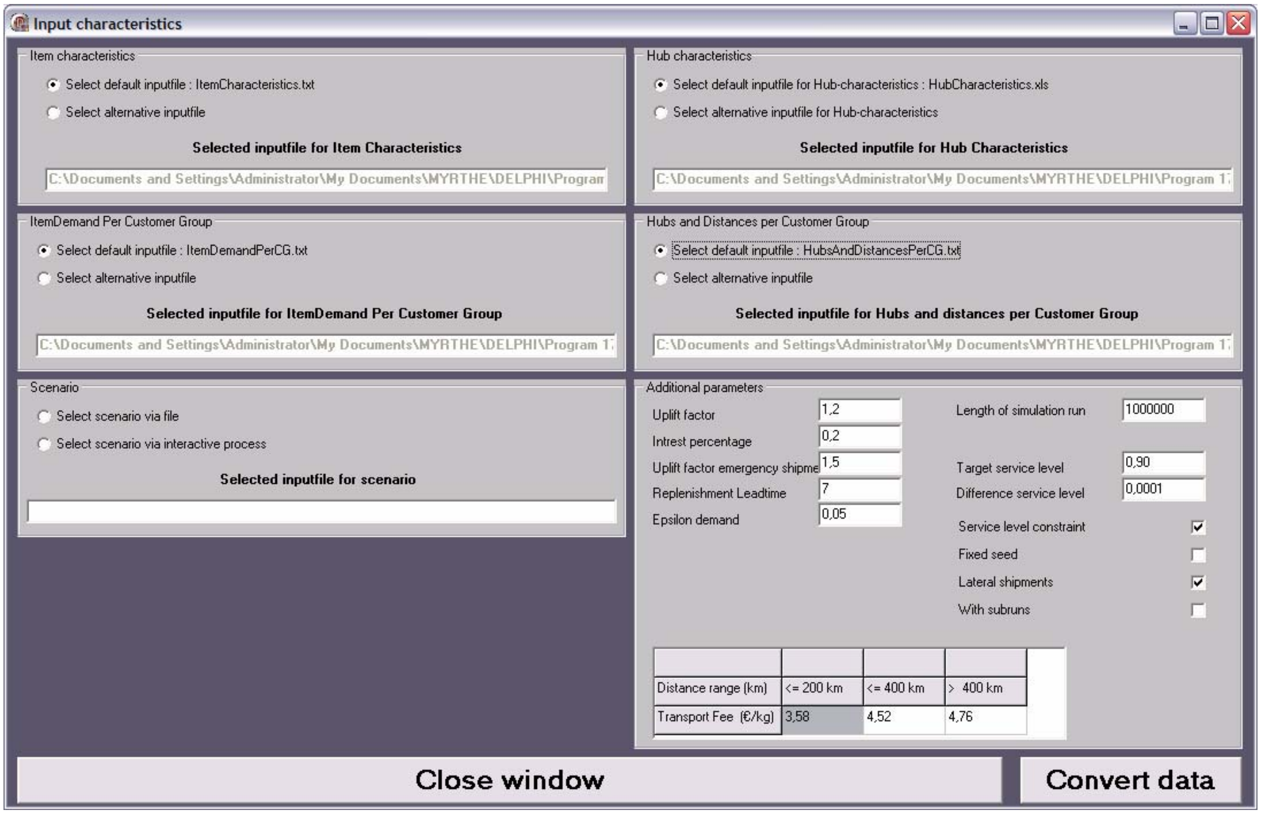


Figure 25: Input screen of tool

K.4 Usage

The program is offered in a standalone executable file. This file can also be used on computers without the program Delphi in which the code is programmed. The size of the file is limited and can therefore be run on standard computers.

The computation time strongly depends on the number of scenarios, the number of hubs per scenario and the number of SKUs. The computation time of several scenarios is equal to the sum of the run length of each single scenario. ***Figure 26*** shows the linear relationship between the number of hubs and the computation time. The variation in the computation time of the ´fast´ computer is caused by the use of other programs during the computation time. ***Figure 26*** also indicates that the specifications of a computer influence the run length significantly. The most important specification on which the 'fast' computer outperforms the 'standard' computer is the amount of memory and the speed of the processor. In this figure the run of the whole sample of SKUs (28.742) is taken into account. The run length of a single SKU is only a few seconds.

The output of the program is put in the folder that corresponds to the folder in which the input data that is selected last, is saved.

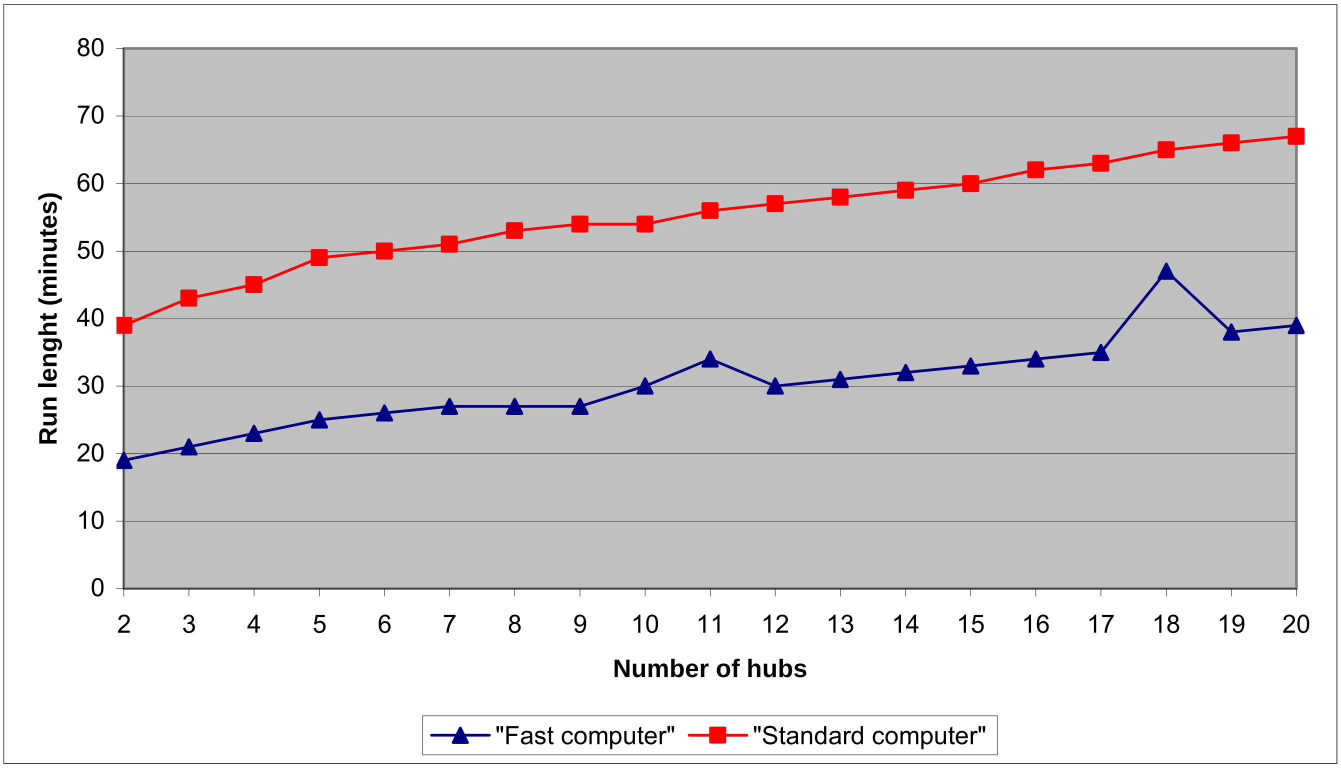


Figure 26: Run length of scenarios

Appendix L: Sensitivity analysis number of hubs design approach III

Figure 27 shows that the pattern of total costs related to the number of hubs is not influenced significantly by costs factors. The inventory and transportation costs are the parameters that have been changed to investigate the sensitivity.

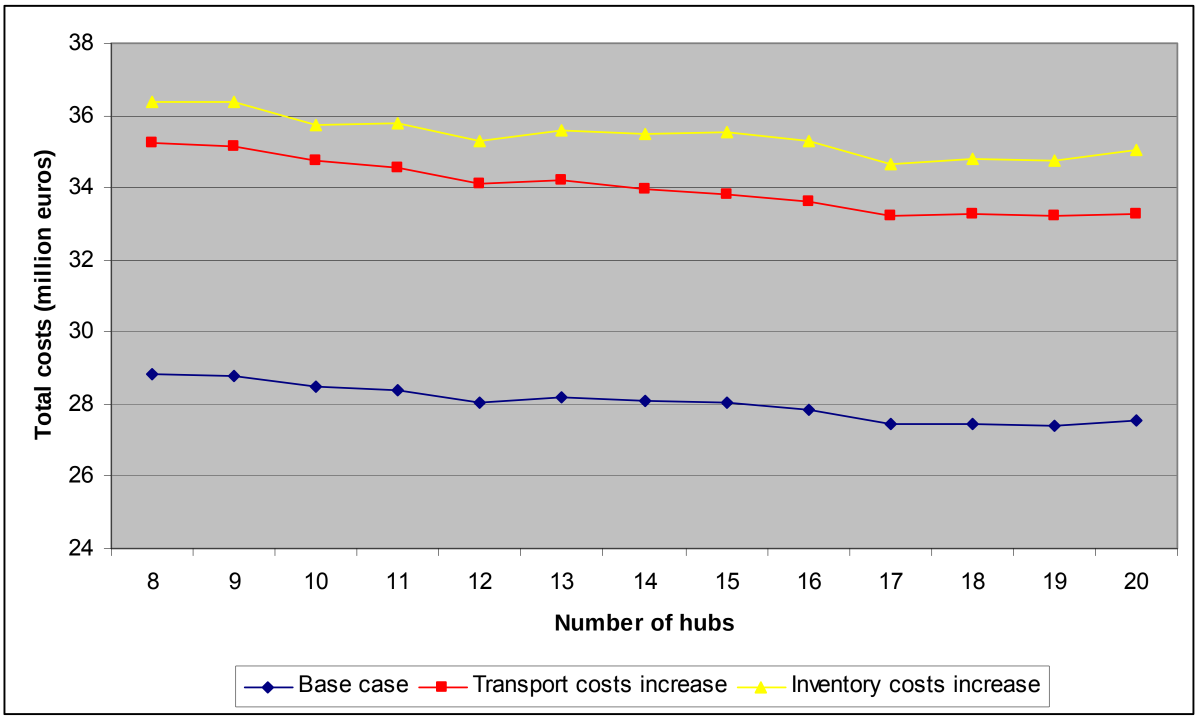


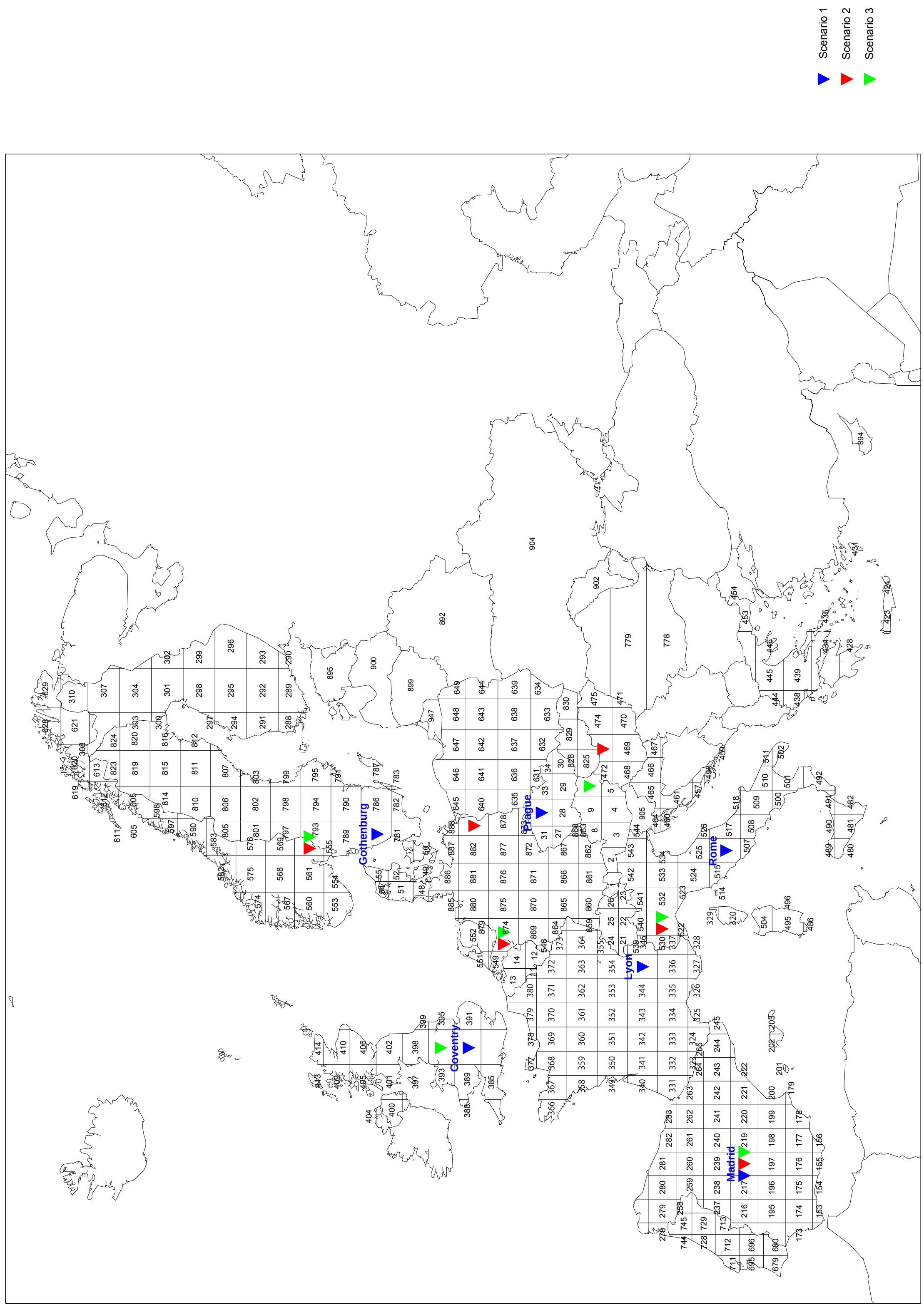
Figure 27: Sensitivity analysis of number of hubs on total costs

Appendix M: Optimal location of hubs for design approach III

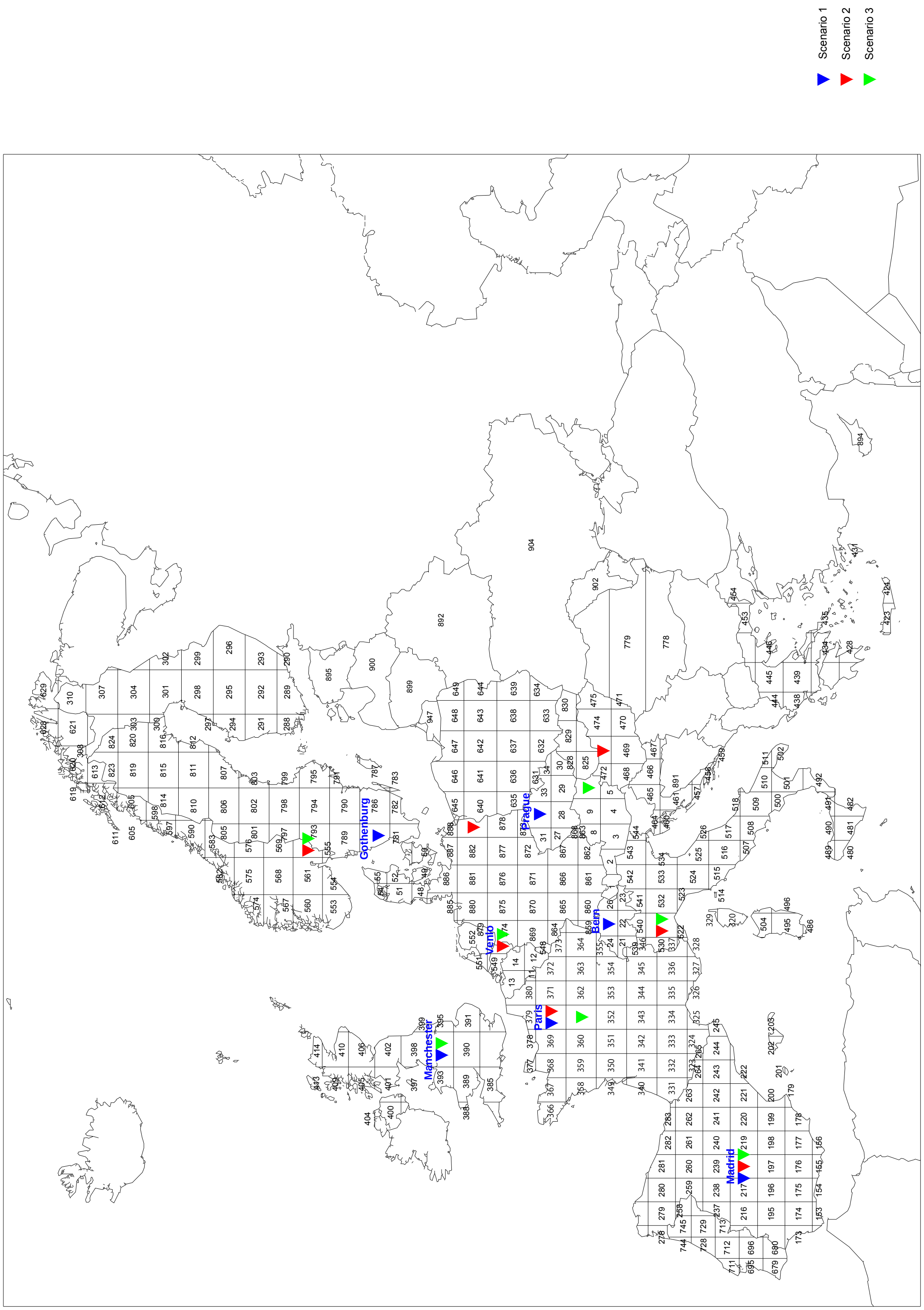
This appendix gives a graphical illustration of several scenarios. For 6 to 14 numbers of hubs, 3 different scenarios are shown. The blue indicated scenario, scenario 1, generates the minimal (related to the other illustrated scenarios) total network costs. Furthermore, this 'optimal' scenario is illustrated by the names of the hubs in the figure. Finally, the last part of this appendix contains an overview of the differences between the scenarios for every number of hubs.

Appendix M1: 6 hubs
Appendix M2: 7 hubs
Appendix M3: 8 hubs
Appendix M4: 9 hubs
Appendix M5: 10 hubs
Appendix M6: 11 hubs
Appendix M7: 12 hubs
Appendix M8: 13 hubs
Appendix M9: 14 hubs
Appendix M10: Relative costs differences between scenarios

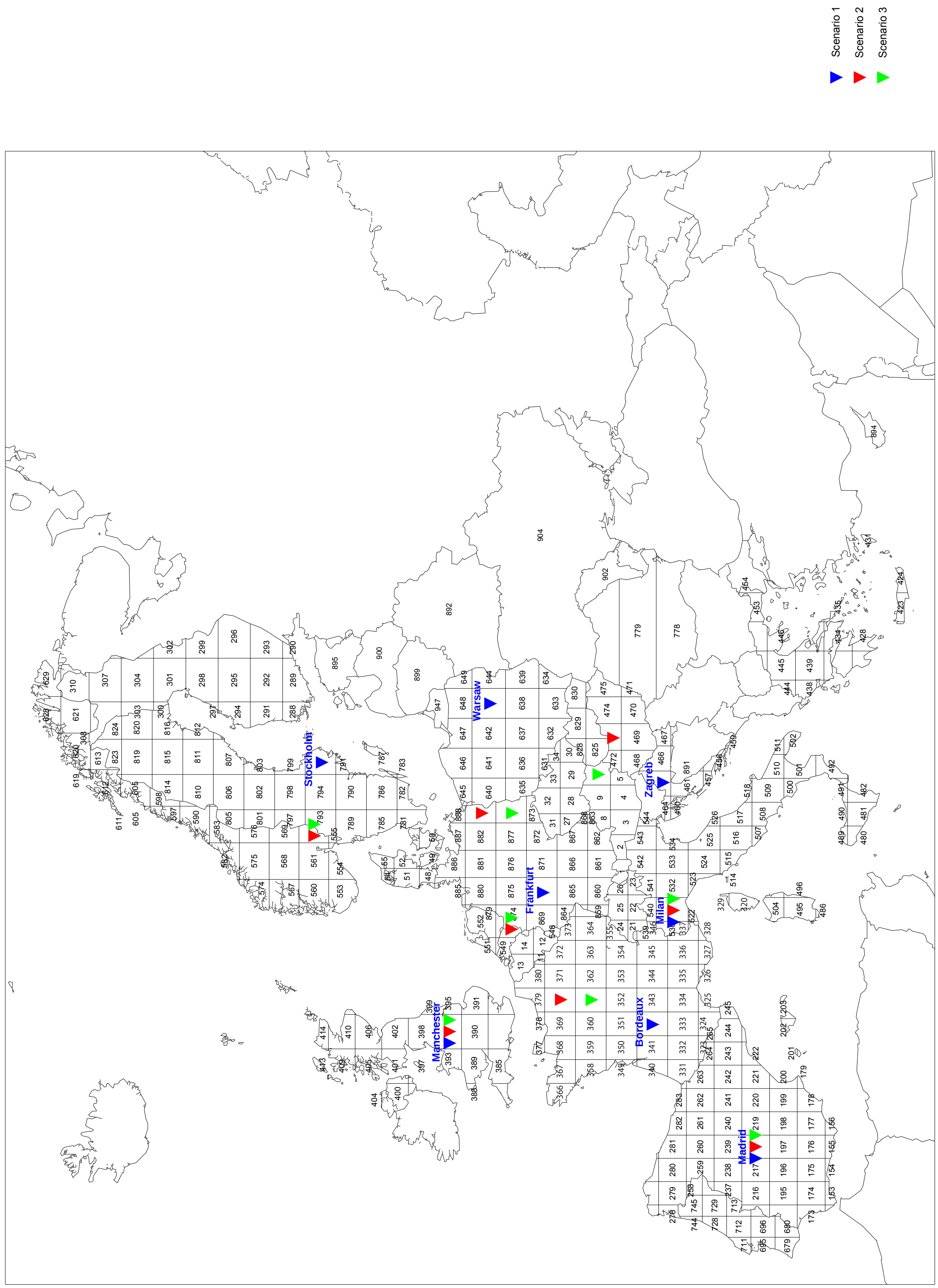
Appendix M1: 6 hubs



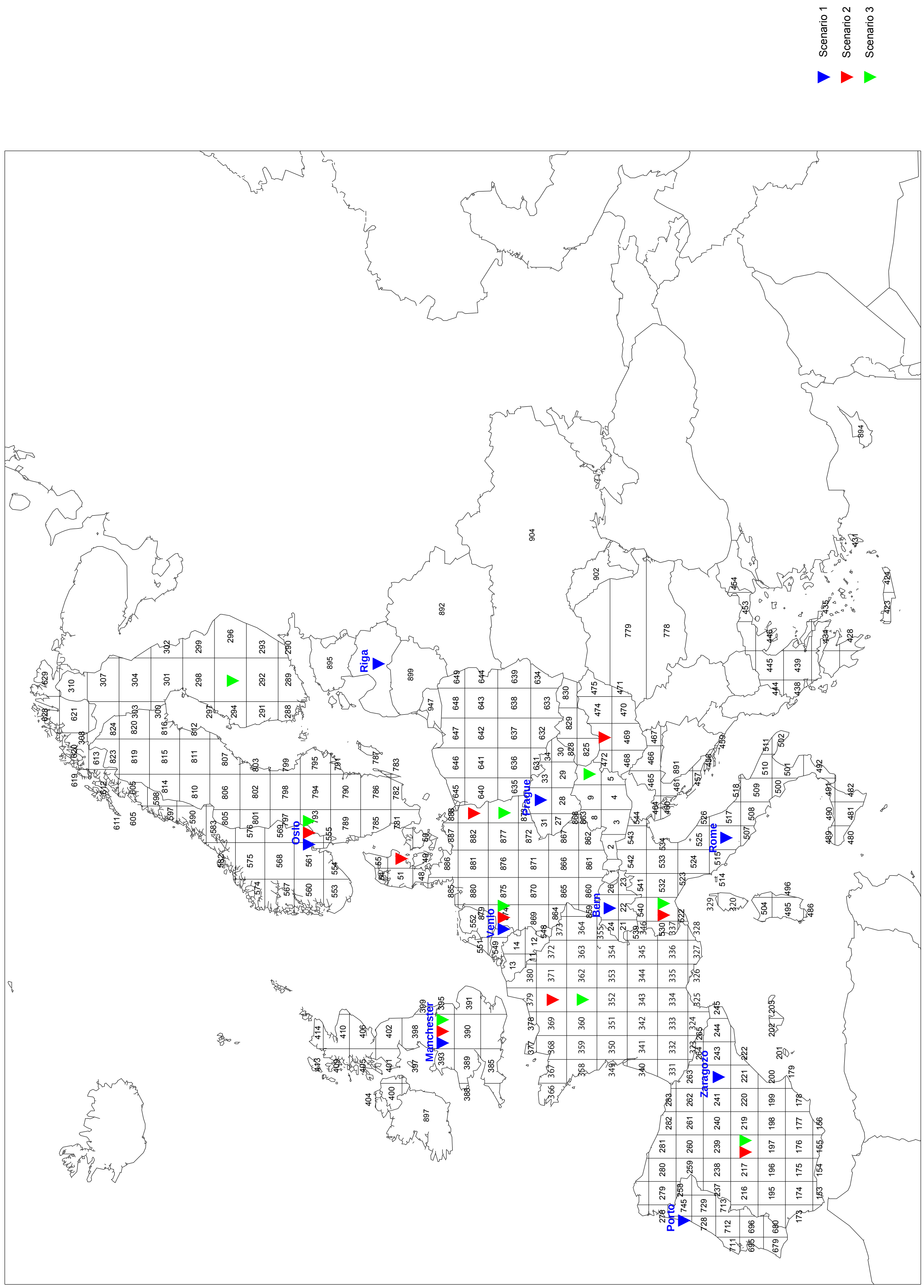
Appendix M2: 7 hubs



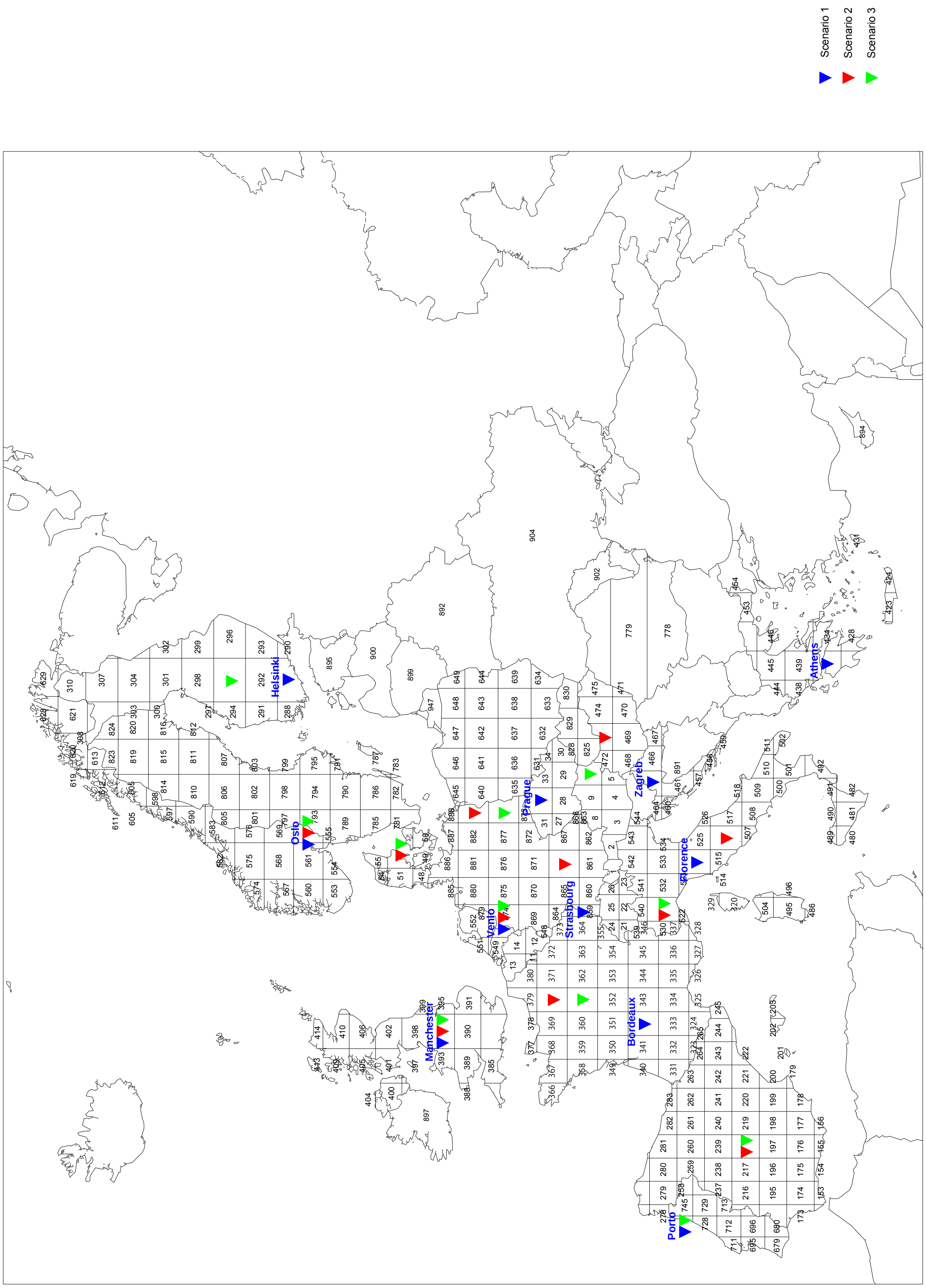
Appendix M3: 8 hubs



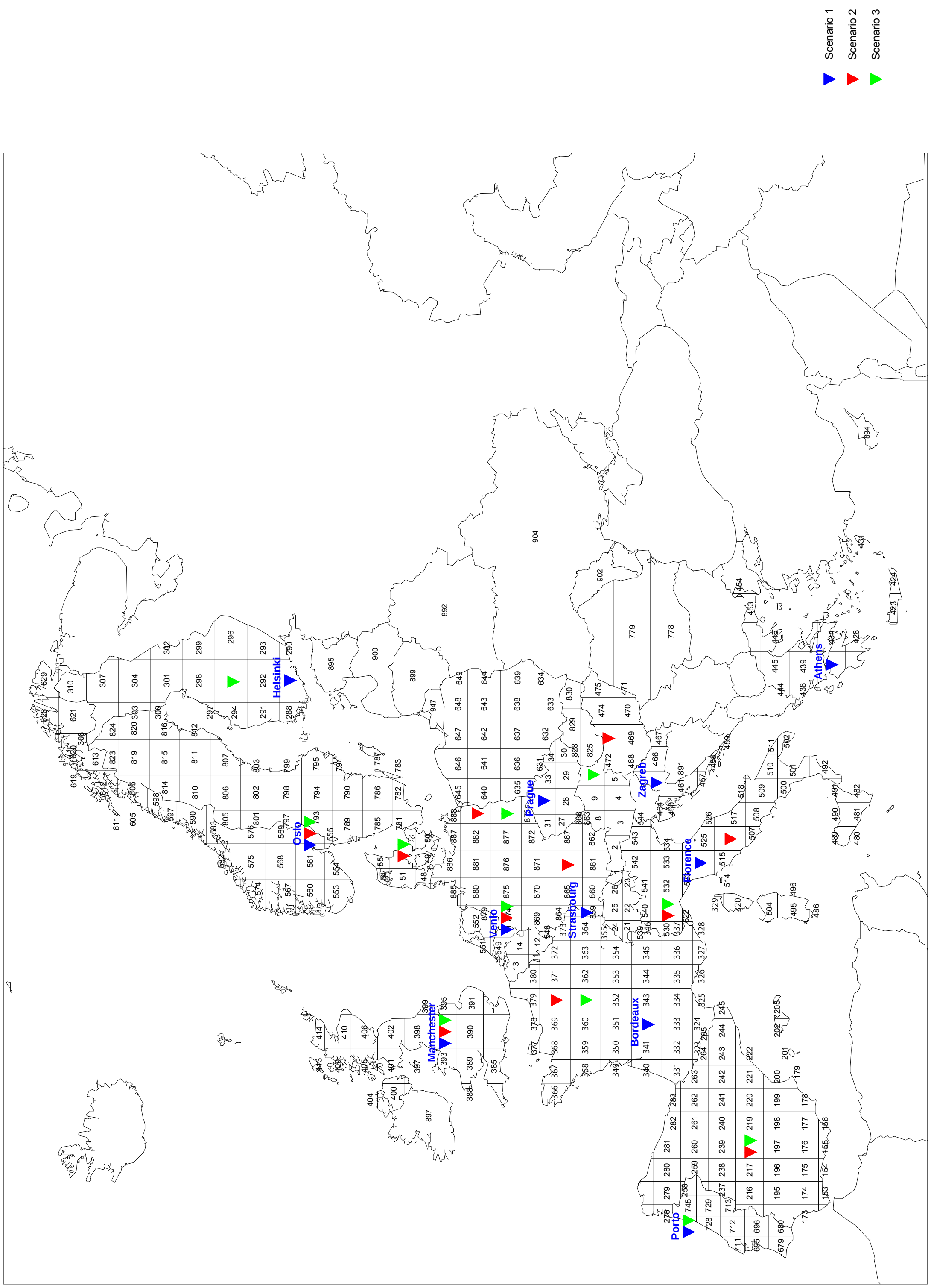
Appendix M4: 9 hubs



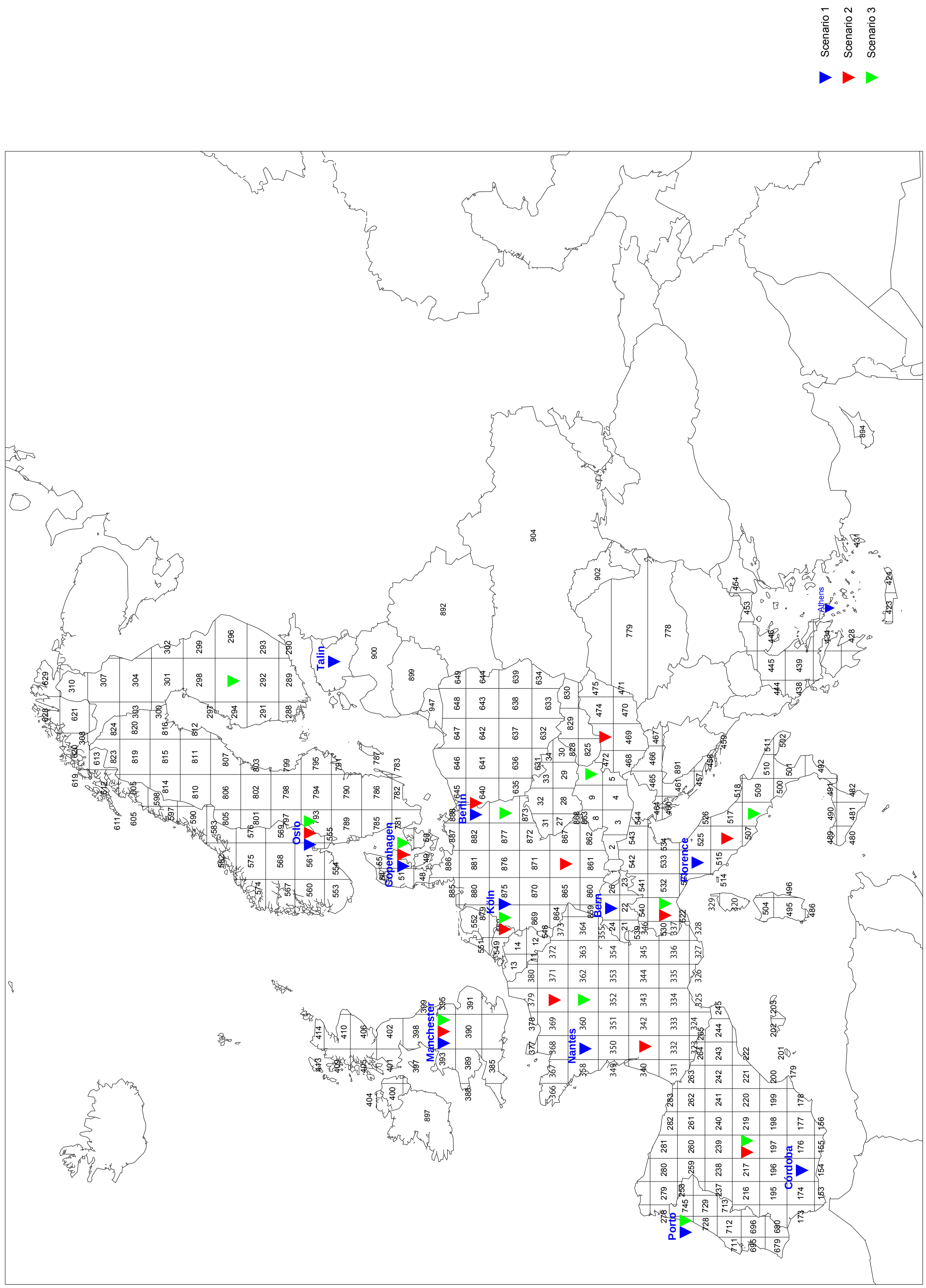
Appendix M5: 10 hubs



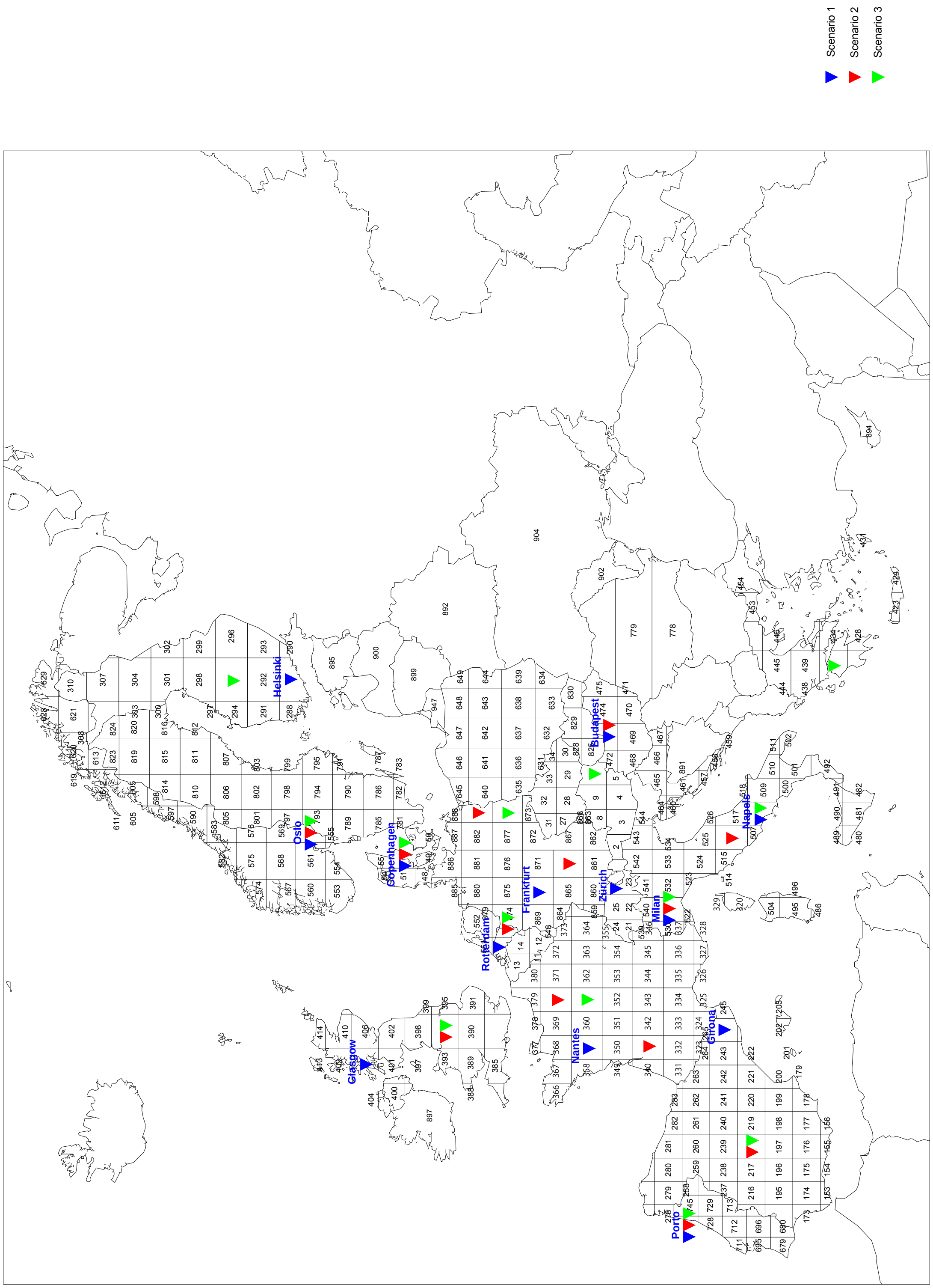
Appendix M6: 11 hubs



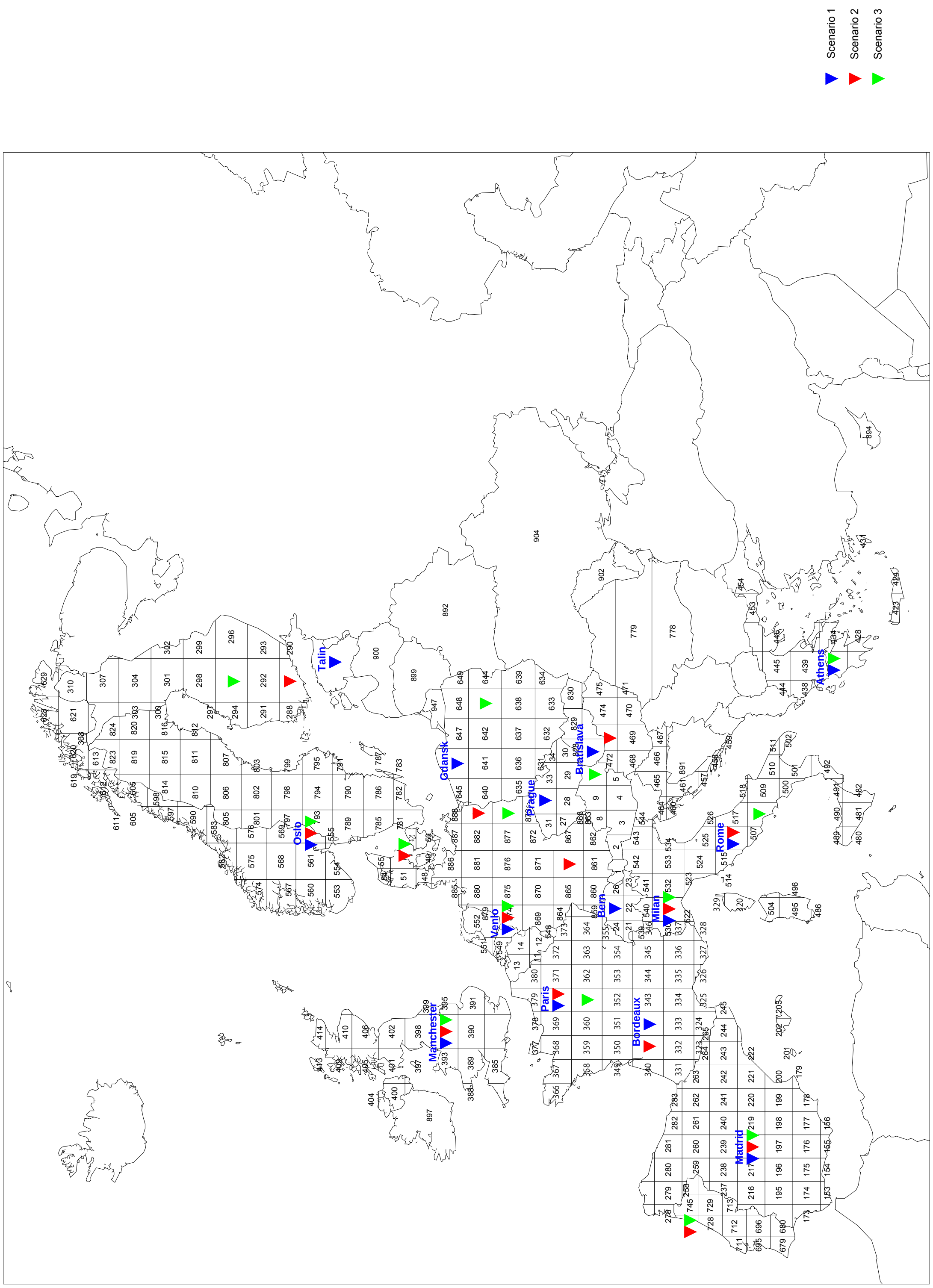
Appendix M7: 12 hubs



Appendix M8: 13 hubs



Appendix M9: 14 hubs



Appendix M10: Relative differences in costs between scenarios

Figure 28 shows the relative difference in costs between the scenarios that are illustrated in the previous part of this appendix. The difference is calculated by subtracting the minimal costs from the maximum costs regarding the three scenarios for each number of hubs. To change this absolute value into a relative value, this difference is divided by the minimal costs. By multiplying with 100% the relative costs difference is retrieved.

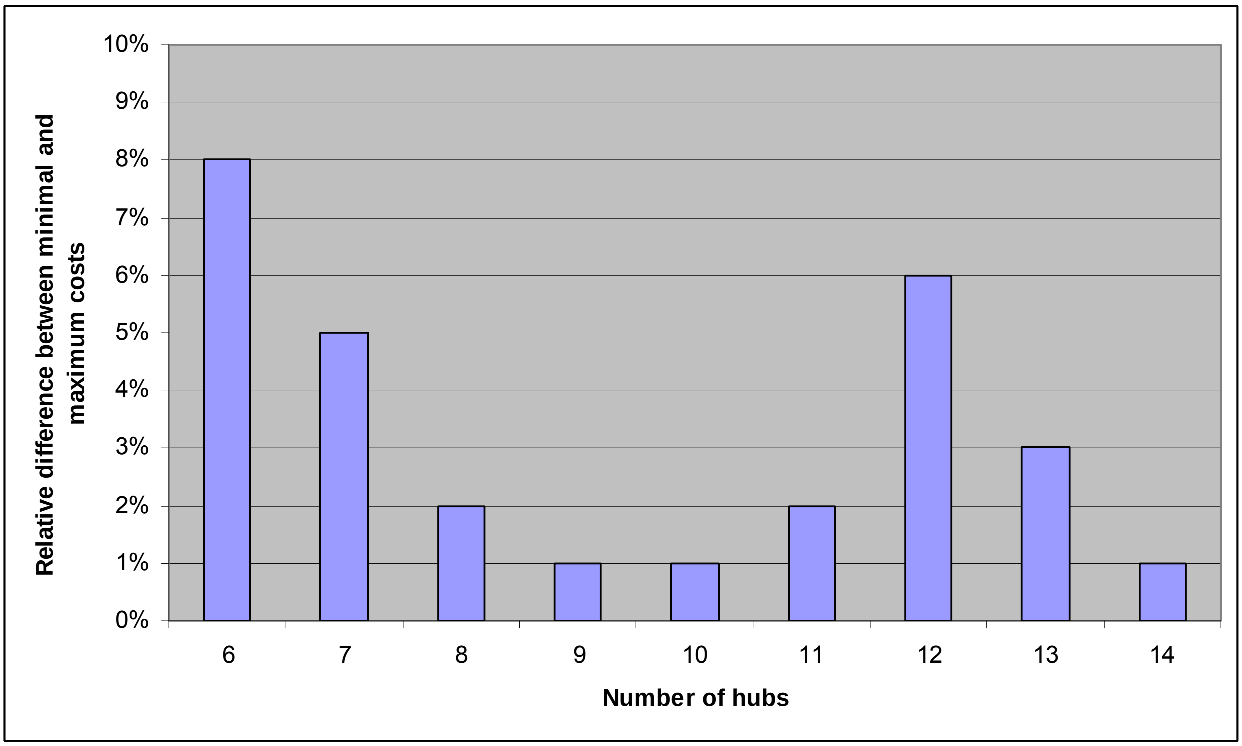


Figure 28: Relative difference in total network costs between scenarios

Appendix N: Difference in structure of costs between approach II and III

<u>Total cost structure</u>

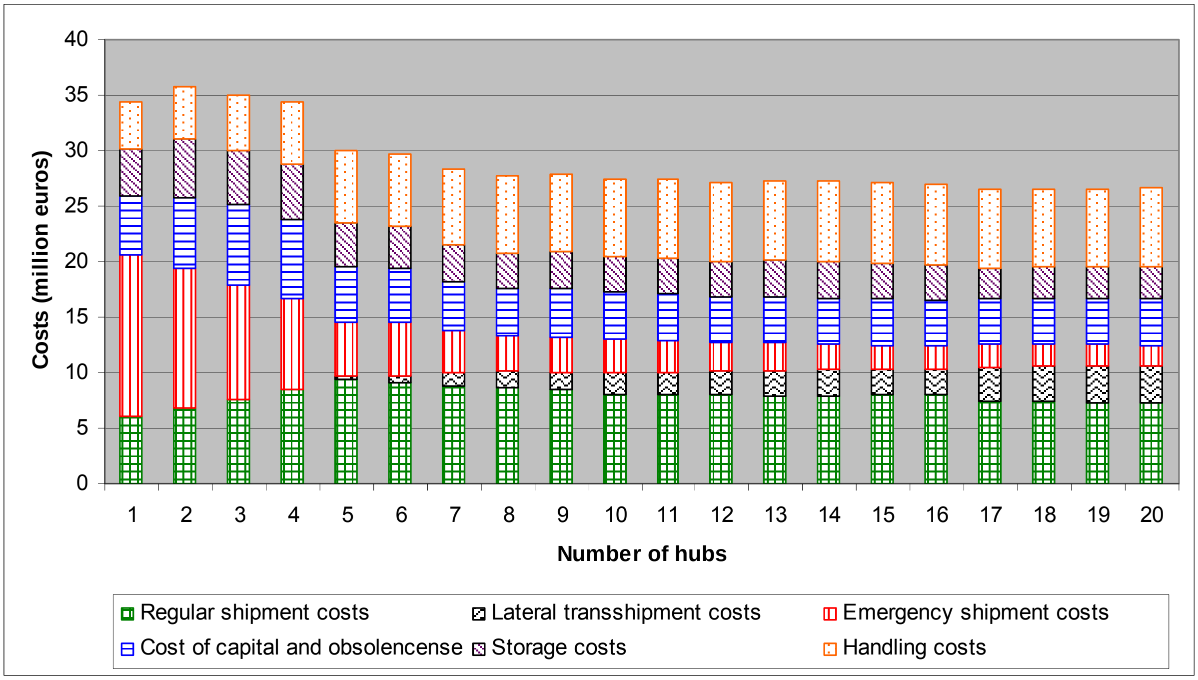


Figure 29: Cost structure of design approach III with service target level of 80%

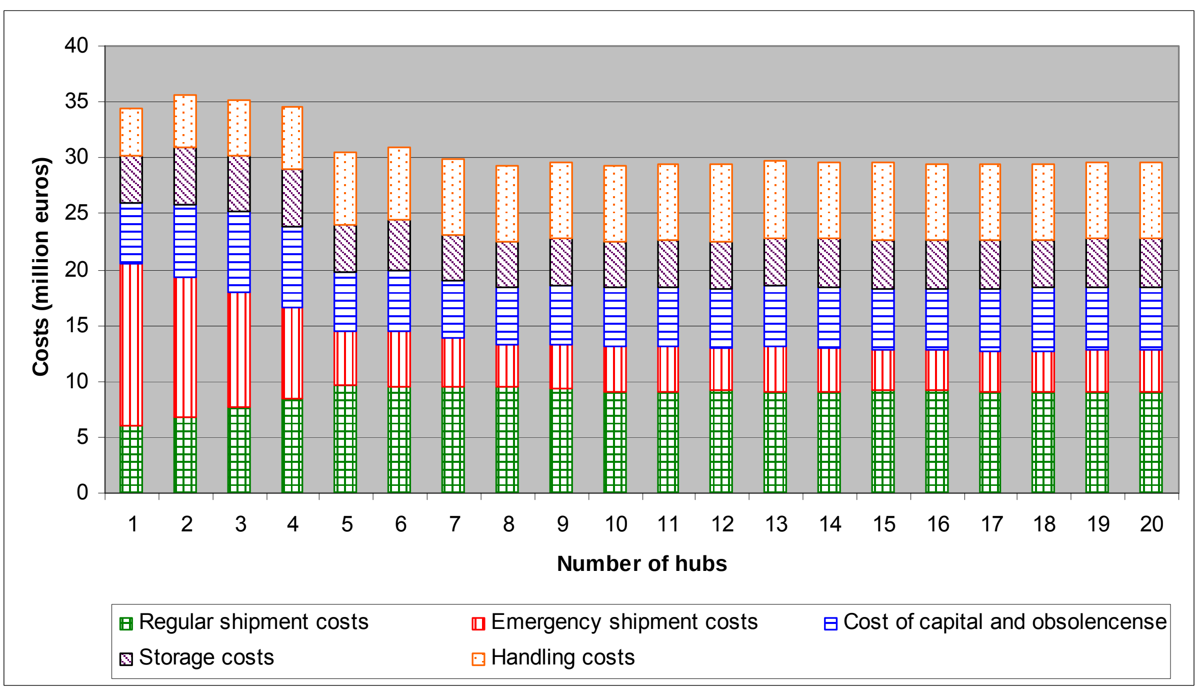


Figure 30: Cost structure of design approach II with service target level of 80%

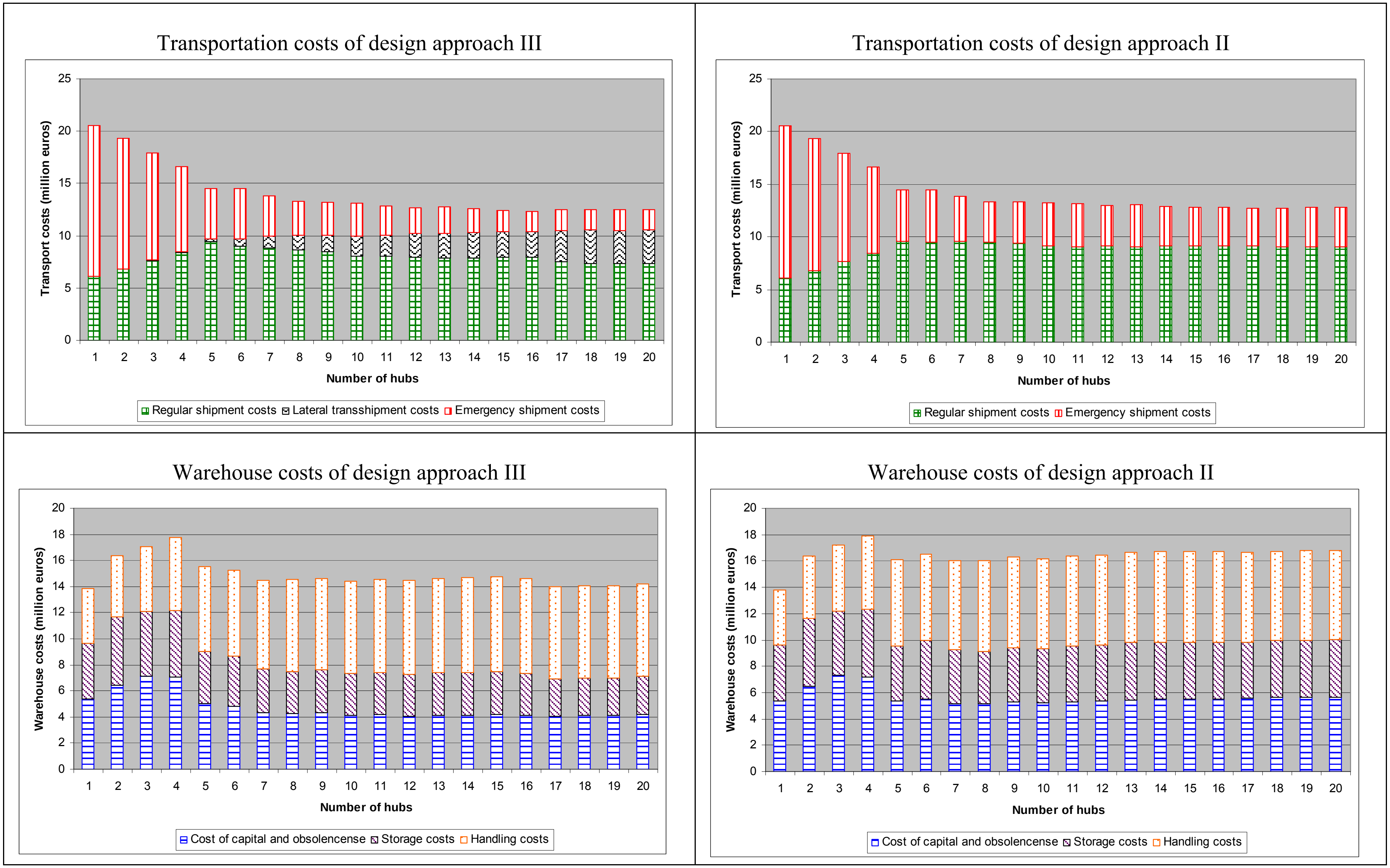


Figure 31: Difference in transport and warehouse cost structure between design approach II and III

Appendix O: Difference between design approach I and III – Impact of inventory costs on total costs

	Number of hubs	Service level	Total costs (million euros)	Total inventory costs (million euros)	Total transport costs (million euros)
Scenario 1	6	0,861	30,68	9,84	14,11
Scenario 2	6	0,903	29,88	9,49	13,14
Scenario 3	6	0,912	28,21	8,45	13,10

	Number of hubs	Service level	Total costs (million euros)	Total inventory costs (million euros)	Total transport costs (million euros)
Scenario 1	7	0,895	29,47	9,1	13,41
Scenario 2	7	0,912	31,67	8,9	12,86
Scenario 3	7	0,878	29,09	8,87	13,54

	Number of hubs	Service level	Total costs (million euros)	Total inventory costs (million euros)	Total transport costs (million euros)
Scenario 1	8	0,914	28,88	8,79	12,9
Scenario 2	8	0,914	29,07	8,88	12,85
Scenario 3	8	0,926	27,45	7,68	12,7

	Number of hubs	Service level	Total costs (million euros)	Total inventory costs (million euros)	Total transport costs (million euros)
Scenario 1	9	0,915	28,79	8,79	12,88
Scenario 2	9	0,919	29,09	8,96	12,75
Scenario 3	9	0,931	27,25	7,40	12,82

	Number of hubs	Service level	Total costs (million euros)	Total inventory costs (million euros)	Total transport costs (million euros)
Scenario 1	10	0,915	28,47	8,56	12,73
Scenario 2	10	0,920	28,95	8,89	12,67
Scenario 3	10	0,918	27,47	7,38	12,84

	Number of hubs	Service level	Total costs (million euros)	Total inventory costs (million euros)	Total transport costs (million euros)
Scenario 1	11	0,924	28,40	8,61	12,54
Scenario 2	11	0,920	29,04	8,97	12,69
Scenario 3	11	0,930	27,25	7,49	12,80

	Number of hubs	Service level	Total costs (million euros)	Total inventory costs (million euros)	Total transport costs (million euros)
Scenario 1	12	0,934	28,06	8,35	12,38
Scenario 2	12	0,931	28,81	8,85	12,50
Scenario 3	12	0,906	27,94	7,69	13,10

	Number of hubs	Service level	Total costs (million euros)	Total inventory costs (million euros)	Total transport costs (million euros)
Scenario 1	13	0,934	28,21	8,47	12,41
Scenario 2	13	0,933	28,77	8,89	12,41
Scenario 3	13	0,922	28,03	8,05	12,77

	Number of hubs	Service level	Total costs (million euros)	Total inventory costs (million euros)	Total transport costs (million euros)
Scenario 1	14	0,942	28,10	8,47	12,24
Scenario 2	14	0,939	28,42	8,65	12,34
Scenario 3	14	0,942	26,95	7,64	12,41

Figure 32: Influence of inventory costs on minimal total costs