Eindhoven, September 2017

Effects of configurations in a service network with lateral transshipments and emergency shipments

By E.V. Borghouts
Student id: 0908772

In partial fulfillment of the requirements for the degree of:

Master of Science
in Operations Management and Logistics.

Supervisors:
Dr. H. Blok, Eindhoven University of Technology
Dr. Ir. J.J. Arts, Eindhoven University of Technology
Dr. Ir. S.D.P. Flapper, Eindhoven University of Technology
Ir. J. van der Wal, PDEng, Royal Philips

Eindhoven University of Technology
School of Industrial Engineering and Innovation Sciences
Series Master Theses Operations Management and Logistics.

Abstract

In this master thesis project, we use mathematical inventory modeling to give Royal Philips strategic insights in their Japanese service network. The implemented model, based on the model from Kranenburg and Van Houtum (2009), considers multiple central warehouses and local warehouses and allows for lateral transshipments and emergency shipments. The model consists of an evaluation based on the Erlang loss model and a Greedy heuristic that optimizes stocking decisions while satisfying Material Availability targets.

The model is used to learn more about the sensitivity between service levels and the corresponding service network costs. Additionally, it gives insight in the effects of optimizing with an aggregate waiting time target, studies the dependency on one of the network's Regional Distribution Centers and shows the effects of decreasing replenishment leadtimes against higher transport tariffs.

Keywords: Inventory control, Spare parts, Service network, Lateral transshipments, Emergency shipments, Capital goods industry, Singe-echelon, Multi-item

Management summary

This master thesis is the result of a project at Royal Philips' Service Parts Supply Chain (SPS) department. As one of the world leaders in health technology solutions, Royal Philips performs maintenance for a large number of (medical) customer systems. This maintenance strongly relies on the availability of spare parts. For Royal Philips' product portfolio, such spare parts are typically expensive and have low demand rates. SPS is responsible for the total spare parts network from Royal Philips' factories and external suppliers to the customers. They aim to maximize the spare parts availability, minimize the costs of operations, and minimize the spare parts inventory levels. To achieve this, SPS operates an extensive and complex service network.

The project focuses on the Japanese service network, which is one of SPS' fastest growing service markets. The multi-echelon service network consists of six local warehouses that are used to accommodate quick demand satisfaction. Three Regional Distribution Centers (RDCs), located in Singapore, Roermond and Louisville, are responsible for the Japanese inventory replenishment. The Japanese service network is especially interesting due to its high service levels requirements. All demands either need to be delivered Same Business Day (SBD) or Next Business day (NBD). SPS sets local and aggregate Material Availability (MA) targets to make sure that a certain fraction of demand is satisfied quickly. The Japanese service network allows for regular shipments, lateral transshipments between local warehouses and emergency shipments from the RDCs.

Due to the rapid growth and many uncertainties, managing this service network is complex. SPS sees multiple possibilities to further improve the Japanese network performance, but due to this complexity, it is hard to oversee the full effect of such changes on the service network. In this master thesis, we perform a case study in which we model the Japanese service network with different network configurations. This contributes to SPS' strategic decisions regarding improvements in the Japanese service network.

We first help SPS to obtain better insight in the relationship between the MA targets and the corresponding costs for the current Japanese service network. The service network costs consist of the inventory holding costs and logistic costs for replenishment, lateral transshipments and emergency shipments. Later, we change our model to allow optimization with an aggregate waiting time service measure instead of MA service measures. Subsequently, we study the service network's dependency on RDC Singapore, by first leaving it completely out of scope, and by later only considering it for emergency shipments. Finally, we analyze whether it is beneficial for SPS to decrease replenishment leadtimes from RDC Roermond and Lousville to Japan against higher transport tariffs.

Results

To model the Japanese service network, we use a special case of the model introduced by Kranenburg and Van Houtum (2009), in which all local warehouses can accommodate and receive lateral transshipments. The model also allows emergency shipments from the RDC if demands cannot be satisfied from local inventory. We generalize the model with multiple RDCs and the dependency of transport tariffs on Chargeable Weight (CW). Due to the assumption that these RDCs have infinite stock, we can limit the model to a single-echelon. The model consists of an evaluation based on the Erlang loss model and a Greedy heuristic that optimizes stocking decisions while satisfying MA targets.

Running this model for the current service network shows the total service network costs required to achieve the MA targets. For the current target settings, the total costs consist of 48.3% holding costs, 29.3% replenishment costs, 16.3% emergency shipment costs and 6.1% lateral transshipment costs. Figure 1 gives more insight in the relationship between the MA targets and corresponding costs.

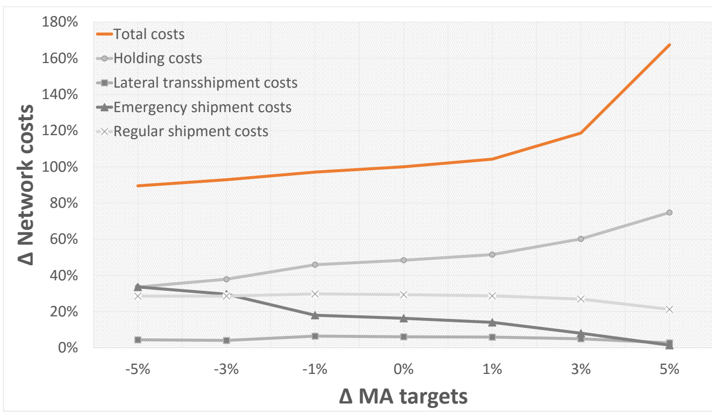

Figure 1: Total costs under different MA targets

By altering our Greedy heuristic, we are able to compare the results from optimizing over MA targets and aggregate waiting time targets. The results show that the same waiting times can be achieved under lower network costs. Alternatively, the aggregate waiting time model shows that the waiting time under the current budget can be reduced by 7.0%. Furthermore, we show that, if SPS considers waiting time to be important, the current MA targets should be reconsidered.

Our model shows that leaving RDC Singapore out of scope reduces the Japanese service network costs under the same service levels by 2.1%. This will, however, cause the aggregate waiting time to increase by 11.7%. We also test how the model performs if only the flow of replenishments is changed, while emergency shipments are still partially sourced from RDC Singapore. This results in 2.4% cost decrease compared to the original model, and does not negatively affect the aggregate waiting time.

Finally, decreasing the replenishment leadtimes reduces the Japanese inventory holding costs but leads to an increase of logistics costs. Decreasing the replenishment leadtime from RDC Roermond to Japan, from RDC Louisville to Japan or RDC Roermond and Louisville to Japan, will result in increased service network costs of 0.9%, 4.8% and 5.3%, respectively.

Recommendations

We recommend SPS to use our results if they are considering to increase the MA target levels. For each level of increase, our results give an indication of the increased Japanese network spendings. Our results show that the current local MA targets should be reconsidered if waiting time is considered to be important. We therefore recommend SPS to adjust the MA targets for Japan, Tokyo, Osaka, Sapporo, Fukuoka, Sendai and Okayama to {0.95, 0.92, 0.94, 0.75, 0.85, 0.76, 0.79}.

Secondly, instead of the current MA service measures, we recommend SPS to adopt an aggregate waiting time target in their planning tools. Such a target accounts for the delay caused by lateral transshipments or emergency shipments, and provides better financial performance in our model.

With regard to RDC Singapore, we recommend SPS to analyze the dependencies of other service demand areas on this distribution center. If RDC Singapore proves to be an important component in the service network of other demand areas, we recommend that the loss of pooling effects by changing the Japanese replenishment flow to RDC Roermond and Louisville are studied. We do in that case recommend that emergency shipments are still partially sourced from RDC Singapore. If RDC Singapore does not prove to be an important network component for other markets, SPS may consider removing the warehouse completely. The resulting increase in waiting times due to the increased emergency shipment leadtimes can then be limited by allowing cross-border lateral transshipments with neighboring markets.

Next, we recommend SPS not to use the decreased replenishment leadtimes against higher transport tariffs from Roermond and Louisville, if the reasoning to do so is purely cost based. Other reasoning to decrease replenishment leadtimes could be a slight reduction in the leadtime variability and a little bit more flexibility, thus reducing uncertainties. If SPS considers this to be worth a network cost increase of 0.9%, the faster replenishment leadtime from RDC Roermond to Japan can be chosen.

Finally, we have a recommendation that is beyond the scope of our research. We recommend SPS to involve the material breakdown structures in the planning system. Such an overview of components per installation enables better predictions of spare part failures. It could also keep SPS from placing parts on stock in markets that do not have the installation in place that may require the part.

Acknowledgments

This report is the result of my master thesis project at Royal Philips and marks the end of my Operations Management and Logistics master's program at Eindhoven University of Technology. A lot of people helped and supported me throughout the project, and I would like to thank everyone that has been involved. I also want to take this opportunity to express my gratitude to some people in specific.

First, I want to thank my supervisors at Eindhoven University of Technology. Herman, thank you for mentoring me throughout the project. I enjoyed our meetings, and your mathematical experience and eye for detail have really contributed to my project as well as my thesis. Secondly, I want to thank my second supervisor from the University. Joachim, your great knowledge in the field of service logistics helped me a lot while scoping and improving my thesis.

Next, I want to express my gratitude to Royal Philips, and especially to my company supervisor. Japke, thank you for giving me the opportunity to take on this project and for your weekly supervision at Philips. You gave me a lot of freedom in conducting the project, but were there for me when I needed feedback or guidance. I also want to thank to all colleagues from the Service Parts Supply Chain department and give a special thanks to Remco for enthusiastically discussing my project and giving me valuable input.

Finally, I feel like everyone will experience some ups and downs while working on a master thesis project. Nevertheless, I really enjoyed working on this project and I am proud of the results that are obtained. I want to thank my family and friends for their unconditional support, encouragement and distraction whenever I needed it.

Erik Borghouts,
Eindhoven, September 2017.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

Healthcare is becoming increasingly sophisticated and consequently, so are the systems used to aid in the diagnosis, monitoring or treatment of medical conditions. These complex systems can be marked as capital assets. Downtime of capital assets may, among other things, lead to lost revenues, customer dissatisfaction or public safety hazard and should therefore be minimized (Driessen, Arts, Van Houtum, Rustenburg, & Huisman, 2015). Because of the technical complexity of capital assets, the Original Equipment Manufacturer (OEM) often remains responsible for maintenance rather than the end user. Service level agreements between the OEM and the customer are used to negotiate how much downtime is allowed for a system.

Downtime of a capital asset is caused by one or multiple failing components. Typically in capital asset maintenance, a repair-by-replacement policy is followed. The failed part in the system is then replaced by an operational part, after which the failed part is scrapped or sent to repair. Maintenance of capital assets thus strongly relies on the availability of spare parts. Executing a repair-by-replacement strategy for a company with a broad product portfolio requires a large pool of spare parts inventory. Because capital asset spare parts are often expensive, keeping large volumes of inventory is costly. The central question of which parts to put on stock in which quantity and location is challenging due to uncertainty in leadtimes and demand, but a good service network and inventory policy can lead to significant cost reductions. Despite the fact that spare parts can tie up a lot of capital, maintenance of capital goods can actually be very profitable (Basten & Van Houtum, 2014).

As one of the world leaders in health technology solutions, Royal Philips, commonly referred to as Philips, performs maintenance for a large number of (medical) customer systems. This maintenance requires an extensive and complex service network, which makes it hard to oversee the full effects of network changes.

In this master thesis, we will focus on Philips' Japanese service market. Japan is one of the fastest growing markets, and is especially interesting due to the high service level requirements. Philips is interested in the sensitivity of the relationship between service levels and costs for this network. We first analyze the current service network configuration and then use mathematical modeling to gain more insight in the sensitivity between costs and service. Finally, we explore how several changes in the network influence this relationship.

1.1 Report structure

In Chapter 2, we give a brief description of Philips and its relevant departments to give a better understanding of the research environment. We end the chapter with the research assignment and research questions. In Chapter 3, we give more insight in the current service network, with special focus on the Japanese market. In Chapter 4, we introduce the model that we use to find optimal inventory levels. This model is then applied to a case study at Philips, which is discussed in Chapter 5. Finally, we give our conclusion and recommendations in Chapter 6.

Chapter 2

Research environment

This chapter describes the framework in which the research is conducted. First, we briefly describe the company and relevant department, followed by explanation of the inventory policy in §2.3. The final section describes the research assignment and research questions.

2.1 Royal Philips

Philips was founded in 1891 by Gerard Philips and his father Frederik Philips. A few years later, Gerard's brother Anton joined the firm. Philips started off producing incandescent lamps and other electrical products and in a few years, the company would grow out to be one of Europe's leaders in light bulb production. After opening a research center called Natlab in 1914, their product portfolio rapidly diversified. Some of the larger innovations are the Compact Audio Cassette, Compact Disc (CD) and the Digital Versatile Disc (DVD). In 2003, Philips opened the Natlab grounds to other technological companies, and so created the High-Tech Campus as it still exists today. By that time, the company was active in three main divisions: Healthcare, Consumer Lifestyle and Lighting.

After creating two stand-alone companies in 2014: Philips Healthtech and Philips Lighting, the company decided to further sharpen their focus in 2016. The lighting division was spun-off, enabling the company to completely focus on Health technology operations. Today, Philips employs approximately 105,000 people, has 95 production sites and has sales and service offices in more than 100 countries. In 2016, Philips realized total sales of €24.5 billion and a net income of €1.5 billion (Royal Philips N.V., 2017). Philips' mission is to improve people's lives through meaningful innovation. To achieve this, Philips developed the Health Continuum, which embeds all of their product segments. While Figure 2.1 shows that the Health Continuum focuses on both consumers and business, the service network only concerns the business to business operations.

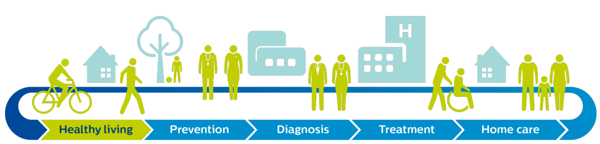

Figure 2.1: Health Continuum

This master thesis project is conducted at the Service Parts Supply Chain (SPS) department, which is partially located at the campus in Best. The company diagram in Figure 2.2 shows how SPS is positioned within Philips.

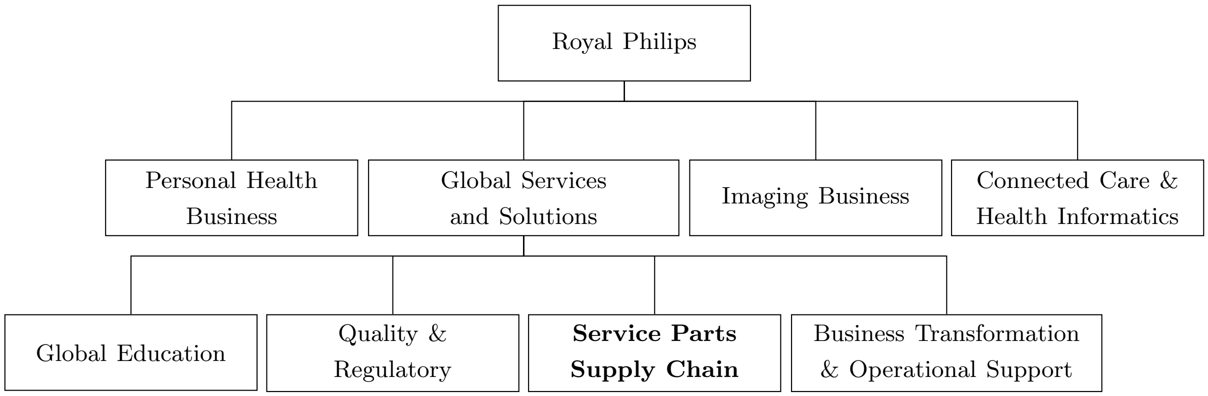

Figure 2.2: Position of SPS within Philips' organization

2.2 Service parts supply chain department

SPS is responsible for the total spare parts supply chain from Philips' factories and external suppliers to the customers. They aim to maximize the service parts availability, minimize the costs of operations, and minimize the service parts inventory levels. This requires a reliable and fast transportation network, an excellent inventory policy and accurate planning. To limit the costs, SPS aims to create pooling effects in centrally located warehouses by considering joint spare parts inventory for multiple demand areas. Annually, SPS moves around 1.8 million parts. Eight teams within the department, each responsible for a different part of the service parts supply chain, are constantly working to achieve their goals in an environment that is subject to many uncertainties. The teams with the highest involvement in this project are Customer Demand & Fulfillment and Strategic Planning & Supply. They take care of warehousing and transportation to fulfill customer demands, and optimizing the target stock levels to improve inventory planning, respectively.

To help support and improve their operations, SPS cooperates with multiple partners. The majority of warehousing activities and a large fraction of transport are outsourced to UPS. Sanmina is responsible for operations in the reverse flow, which may include testing, spare part repairs and disposal of defective parts. All of SPS' transactional activities are outsourced to Accenture.

2.3 Inventory policy

In each market, Philips employs Field Service Engineers (FSEs) to perform preventive and corrective maintenance. They place orders for required spare parts and repair the installation as soon as the parts have arrived. The demanded parts can be supplied from different stockpoints in the service network. SPS controls the inventory levels in those stockpoints with an automated planning system that follows an s, Q policy. As soon as the inventory position for a certain Stock Keeping Unit (SKU) falls below re-order level s, a re-order quantity Q is ordered. This re-order quantity is computed by following the the economic order quantity logic. Because capital asset spare parts are typically expensive and have low demand rates, batch ordering is only beneficial for a few SKUs.

The re-order levels are determined based on Material Availability (MA) targets. These MA targets are selected such that the service level agreements negotiated with customers are met. SPS sets aggregate MA targets per country, and for some countries local MA targets per warehouse are determined. The local MA targets are defined as: the fraction of demand satisfied directly by the warehouse that first sees the demand. The aggregate MA target is defined as: the fraction of demand that can be satisfied from any stockpoint in the country, which equals all deliveries from on hand stock and through lateral transshipments. The latter are demands that are fulfilled by another stockpoint in the same country. In literature, such MA targets are often referred to as fill rate targets.

Whenever the planning system sees an incoming demand, the system automatically checks the required service, and allocates the demand to a warehouse. The required service, or speed of delivery, for a demanded part may depend on multiple factors. Preventive maintenance can be scheduled in advance, so these orders are shipped with lower service. Parts ordered for corrective maintenance can have different levels of criticality. Based on this criticality, the required service is determined. Lower service demands require lower transportation speed and are therefore allocated to more centralized stocking points. Higher service demands need to be satisfied faster and are therefore often allocated to the nearest stocking point.

As soon as the demand is allocated, a rule based availability check is triggered. If the allocated warehouse has no on hand stock for the demanded part, there is a fixed sequence in which other stocking locations are checked for on hand inventory. In case another warehouse in the country has stock available, a lateral transshipment is applied. In case the part is not available at any warehouse in the country, the demand is satisfied through an emergency shipment from one of the Regional Distribution Centers (RDCs). The use of lateral transshipments and emergency shipments as visualized by Van Houtum and Kranenburg (2015), can be found in Figure 2.3.

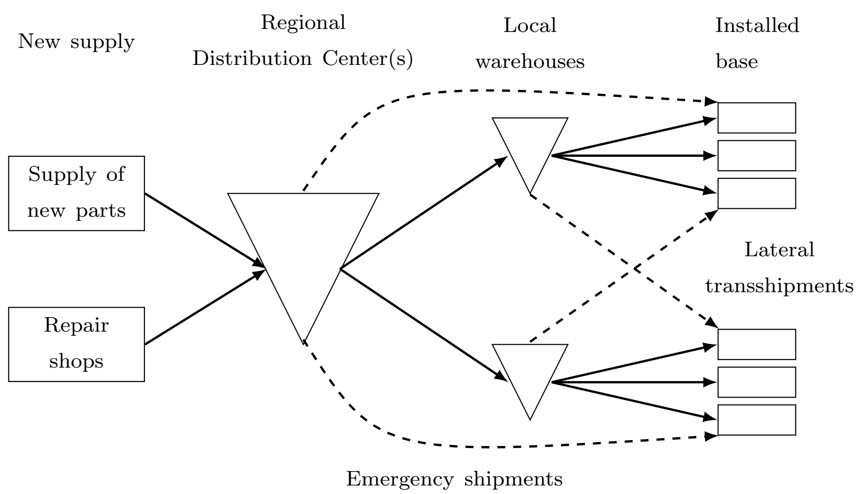

Figure 2.3: Lateral transshipments and emergency shipments

2.4 Research assignment

Due to uncertainties and the changing environment, managing large service networks can be very complex. Especially if the network consists of multiple echelon levels and allows for a variety in possibilities to fulfill demands in case of a stockout. SPS sees multiple possibilities to improve the Japanese network performance, but due to its complexity, it is hard to oversee the full effect of such network changes. In this master thesis, we help SPS to obtain better insight in the relationship between the MA targets and the corresponding service network costs for the Japanese market under different configurations. This will contribute to SPS' strategic decisions regarding improvements in the Japanese service network.

The Japanese service network consists of multiple local warehouses and satisfies inventory through regular shipments, lateral transshipments and emergency shipments. In order to obtain more insight, we modify a mathematical inventory model from literature and use it to perform a case study. The service network analysis and the case study give us the information to answer the following research questions.

Main research question:

How sensitive is the relationship between the Material Availability targets and the corresponding costs for the current Japanese service network; and how do changes in the network influence this relationship?

Sub research questions:

- *What is the current service network configuration for the Japanese market?*
- *How can the Japanese service network be captured in a mathematical inventory model available in literature, so that it yields the closest possible resemblance to reality?*
- *Under what costs can the Japanese Material Availability targets be achieved? And how much do target increases or decreases affect these results?*

- *Can the level of aggregate waiting time as obtained when satisfying the Japanese Material Availability targets be achieved under lower costs?*
- *How does the relationship between service and costs in the Japanese service network change if we remove the Regional Distribution Center of Singapore from our model?*
- *Can the Japanese service network costs be decreased by choosing faster replenishment modes?*

2.4.1 Scope

In order to reduce the complexity and increase the relevance, the following scope is defined.

- The master thesis project will focus on the Japanese service network.
- Demand data from 2016 is used.
- The reverse flow is left out of scope.
- Japanese office stock locations are left out of scope. Six Japanese warehouses fulfill 99% of all incoming spare part demands, while the remaining 1% is fulfilled by 20 office stock locations. In the remainder of this thesis, we only focus on the six warehouses.
- Field Change Order (FCO) demands are left out of scope. As these parts are pushed into the system, such orders cause a peak in demand. Because these FCO can be planned in advance, we do not want to take the demand streams into account in our model.

Chapter 3

SPS service network

In the previous sections, we explained that SPS uses a service network to satisfy the global demand for spare parts. This chapter gives more insight in the network structure, the flow of spare parts and the costs that are involved. In sections §3.1, 3.2 and 3.3, we address the global network, and in §3.4, we direct our focus to the Japanese network.

3.1 Network structure

The global service network is divided into three regions: Asia, Pacific (APAC), Europe, Middle East, Africa (EMEA) and North, Central and South America (AMEC). Each region has one RDC, located in Singapore, Roermond and Louisville. Each region contains a number of demand areas, which each have one or multiple Local Distribution Centers (LDCs) and Forward Stocking Locations (FSLs). These LDCs and FSLs are used to store inventory closer to customer installations, thus reducing the delivery time to customers. The FSLs are replenished by LDCs, which receive their replenishments from the three RDCs. The following sections elaborate more on how spare parts are moved through the network.

3.2 Spare parts flow

The service network accommodates movements of spare parts, which can be part of the forward flow or the reverse flow. The forward flow is discussed in §3.2.1 - §3.2.3, and the reverse flow is explained in §3.2.4. The full scheme of all possible movements is visualized in Figure 3.1.

3.2.1 Inbound logistics

All ready-for-use parts that enter the network are part of inbound logistics. These can be shipped from external suppliers, repair vendors or internal suppliers, also referred to as Business Innovation Units (BIUs). All parts that enter SPS' network are first shipped to one of the RDCs. To which RDC a part is shipped depends mainly on the part's network root. Such a network root indicates that the supplier is oriented in a certain area, and can only ship to the RDC in their region. SPS mainly sees US (United States), NL (Netherlands) and virtual network roots. A virtual network root indicates that the supplier is capable of shipping their parts to all three RDCs. SPS can then choose which RDC is favorable in terms of costs and delivery time.

3.2.2 In-network logistics

Inventory can be moved from one stockpoint to another through stock transfer orders. These may be triggered by two events: 1) a stockpoint has excess stock, which is periodically shipped to a more central stocking location or 2) the stock level reaches the re-order level, triggering a replenishment order. The TARN logic is applied to decide which source of spare parts is used for replenishment. This is an acronym for Transship, Allocate, Repair and New-buy. The first step is to check the possibility of Transshipments, in which the system will search for excess stock in the same echelon level warehouses. Secondly, Allocation is considered, in which the system takes stock from a higher echelon level. In case the higher echelon level is low on stock, the replenishment may be delayed as it can be beneficial to keep the part at a more central location in the system. The replenishment order will then be added to the re-order quantity for the higher echelon level warehouse. Whenever the part becomes available there, the replenishment will take place. In case no parts are available in a higher echelon level, the system considers which parts are being repaired or are incoming from repair. If no parts will become available at the repair shop, a new part will be purchased.

3.2.3 Outbound logistics

All shipments from SPS warehouses to the customer are part of the outbound flow. These are usually ordered by the FSE, but in some cases the customer can order the spare part and perform the repair independently. Such orders are referred to as No Engineer Material Orders (NEMOs). For reasons of timing or special material characteristics, e.g. very large parts, the customer can choose to have a part delivered at a Pick-Up Drop-Off point (PUDO). The FSE can then retrieve the part from the PUDO and install it at the customer installation. As explained in §2.3, demands can be satisfied through regular shipments from on hand stock, lateral transshipments and emergency shipments.

3.2.4 Reverse logistics

Whenever the FSE has performed the repair, he sends most failed parts back to SPS through the reverse spare parts flow. Spare parts are repairable or consumable, and the reverse flow consists of all repairable parts and some of the consumable parts. A fraction of the consumable parts can be scrapped at the customer, while others need to return to SPS for analysis or specified scrapping procedures. All parts in the reverse flow are first shipped to a Blueroom, which inspects whether incoming parts need to be scrapped, repacked or repaired. The Blueroom can handle most scrapping and repacking locally. Repacked parts that are ready-for-use can be shipped to the RDC. In case a part needs to be repaired, the Blueroom ships it to the repair vendor.

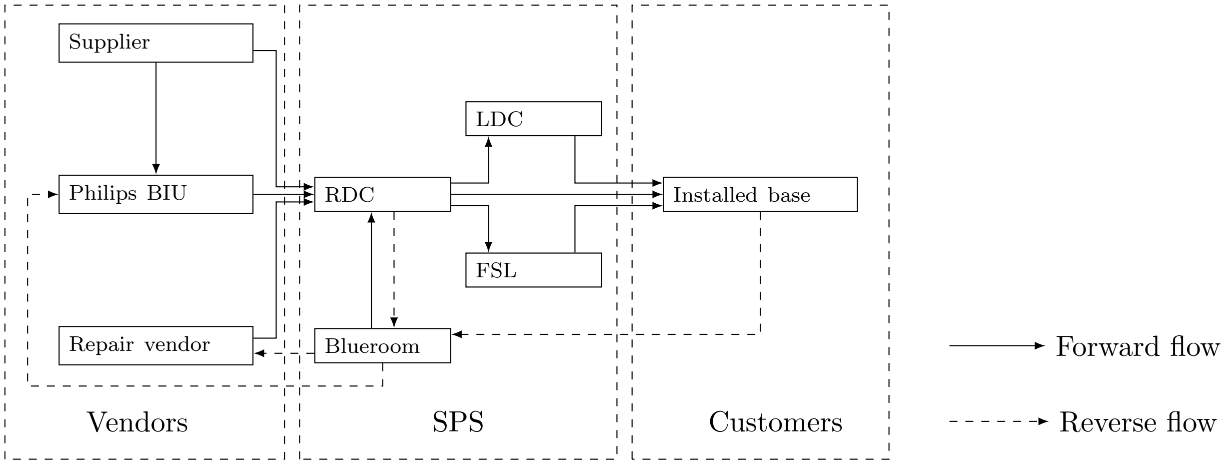

Figure 3.1: Spare parts flow

3.3 Network costs

SPS wants to maximize spare parts availability and minimize the costs of operations to achieve this. To obtain a better understanding of the cost structure, the following section briefly elaborates on the relevant service network costs.

3.3.1 Inventory holding costs

As spare parts are generally expensive, holding costs are an important factor in the costs structure. SPS assumes that annual holding costs are 20% of the part's value, which is based on an estimate for the cost of capital and cost of storage. In general, SPS leases the warehouses for a fixed rate for management and a certain amount of square meters of storage. Additional square meters and handling costs per part are charged with a variable rate.

3.3.2 Transport

SPS cooperates with multiple carriers to arrange transport throughout the service network. UPS is responsible for a large fraction of the transport, but for some lanes, other carriers offer better rates. Lanes can be domestic (shipments within country) or cross-border. In general, SPS uses two flavors of service: airfreight for lower service orders and parcel for higher service orders. For airfreight service, the carrier aims to consolidate shipments to reduce costs. It is therefore less flexible and slower, but also cheaper. For Parcel service, packages are shipped individually, which speeds up the process but is more expensive.

SPS and their logistic partners agreed upon fixed tariffs per lane. These tariffs are influenced by the requested service level, and the Chargeable Weight (CW). The CW is the maximum of the actual weight in kilograms and the volume weight. The volume weight is determined by dividing the package volume ($length \times height \times width$) in cm by 6000. CW is thus calculated as follows: $CW = \max(\text{actual weight}, \frac{l \times h \times w}{6000})$. To achieve consolidation, SPS' logistic partners set a minimum of 30 kilograms of CW for each airfreight shipment. If SPS wants to transport an airfreight shipment that sum up to a lower weight, they are still charged for 30 kilograms. In general, this threshold is always achieved. Figure 3.2 gives insight in the factor of cost difference between parcel and airfreight transport. For lower CW shipments, parcel shipments from Roermond to Tokyo are up to 31 times as expensive as airfreight shipments. As the part's weight increases, the cost difference between parcel and airfreight shipments becomes smaller. For SKUs with CW over 25 kilograms, parcel shipments from Roermond to Tokyo are only about three times as expensive as airfreight.

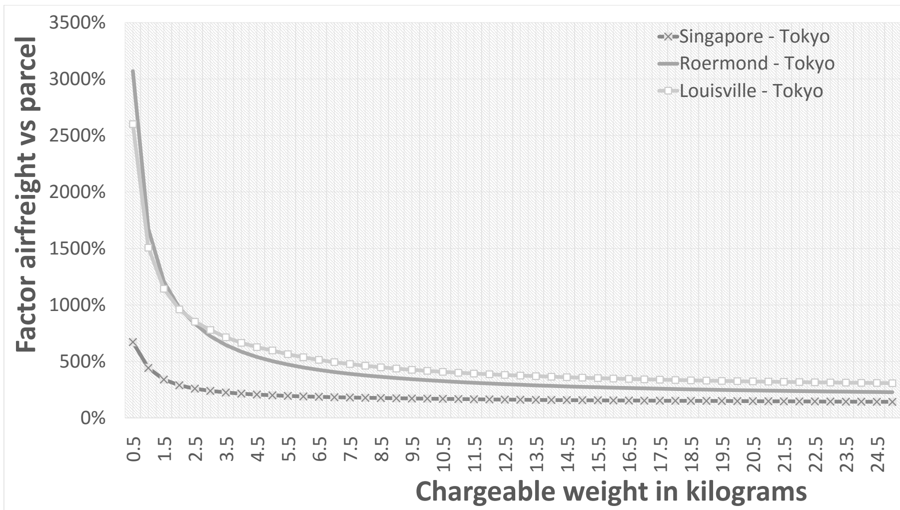

Figure 3.2: Factor of costs difference parcel vs airfreight per kg of CW

3.3.3 Customs

On cross-border lanes, SPS has to deal with customs clearance. This is the documented permission of the national customs authority that SPS is allowed to import their goods. Depending on the destination and the origin, the clearance involves paperwork or a physical inspection. In some countries, customs can be a bottleneck, whereas other customs authorities operate reliably and cause little delays. Costs involved with clearance are different per country and are in general based on the part's characteristics and value. Because paying customs clearance tariffs is inevitable when importing parts to Japan, no change in our model can affect the costs involved in customs. These costs are therefore left out of scope in the remainder of this report.

3.4 Japanese network

This research focuses on spare part deliveries to Japan, as a part of the APAC region. Around 7% of all global SPS spare part demands occur in Japan, which corresponds to around 29% of the APAC market. The Japanese service network is good for approximately €175 million of annual revenue. To keep all customer bases up and running, SPS employs 240 FSEs in Japan. In 2016, the Japanese service network placed orders for 5,649 unique SKUs, with 66,224 demanded parts in total. These orders were placed by 1,813 unique customers. Figure

3.3 shows how these demands were distributed over the six Japanese warehouses. Figure 3.4 gives insight in the distribution of demand rates SKU.

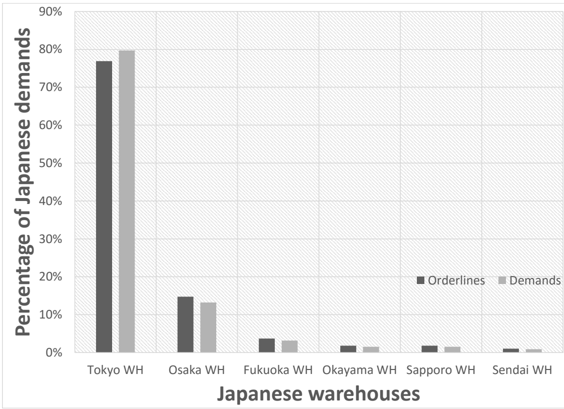

Figure 3.3: Demands per warehouse

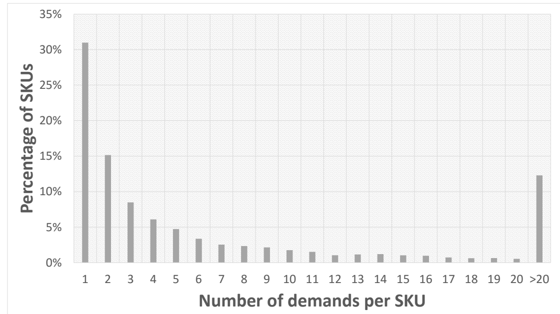

Figure 3.4: Annual demand rates

The Japanese service network consists of two LDCs in Tokyo and Osaka, and four FSLs in Sapporo, Fukuoka, Sendai and Okayama. Their geographical location is visualized in Figure 3.5. As elaborated on in §3.2.2, RDCs are used to replenish the LDCs and FSLs. Table 3.1 shows which RDCs account for which fraction of replenishments and emergency shipments.

For transport to the Japanese market, SPS cooperates with Nippon, UPS, FedEX and DHL as logistic partners. Seino is responsible for the domestic shipments. The Japanese customs authorities operate fast and reliably, which enables SPS to keep the delay on leadtimes limited.

Figure 3.5: Geographical location Japanese warehouses

Table 3.1: RDC allocation of Japanese demands

RDC	Replenishment	Emergency shipments
Singapore	16%	45%
Louisville	38%	23%
Roermond	46%	32%

The Japanese market is served with two main service levels: Same Business Day (SBD) and Next Business Day (NBD) delivery. In 2016, approximately 35% of demands required SBD delivery. All NBD Japanese demands can be satisfied by Tokyo and Osaka, while SBD demands are preferably shipped from the nearest stocking location.

Demands are thus allocated to a warehouse based on region and required service. In case the allocated warehouse has no on hand stock, the warehouse will request a lateral transshipment from one of the other five Japanese warehouses. In case none of the warehouses have available stock, an emergency shipment is triggered. Japan, has an ASAP (All Special Assistance Parts) location in the warehouse in Tokyo that receives and distributes all emergency shipments directly to the customer. Shipments to the ASAP location are supplied by one of the RDCs with parcel service. For such emergency shipments, RDC Singapore is always checked for available stock first. In case no stock is available, which most frequently occurs for US and NL rooted products, the part is shipped from the RDC in Roermond or Louisville.

The Japanese local warehouses are replenished by all three RDCs. Even though SPS follows an s,Q inventory policy, the re-order quantity Q is equal to one for almost 99% of the SKUs in Japanese warehouses. In practice, the Japanese service network replenishment is therefore similar to a basestock policy. Japanese inventory is planned with local MA targets per warehouse and an aggregate MA target for the country. The target levels can be found in Table 3.2. The local MA targets ensure that a certain fraction of orders with a SBD service requirement are delivered from the nearest warehouse, thus satisfying demand as soon as possible. The local MA targets per warehouse are presented in Table 3.2.

Table 3.2: Japanese service requirements

#	Warehouse	MA target
1	Tokyo	93%
2	Osaka	91%
3	Sapporo	70%
4	Fukuoka	70%
5	Sendai	60%
6	Okayama	60%
	Aggregate of Japan	95%

3.4.1 Pharmaceutical Affairs Law

The Japanese government has strict regulations regarding the quality of medical equipment. In order to safeguard the quality, each product in direct contact with patients needs to be tested to pass the Pharmaceutical Affairs Law (PAL). All PAL parts are tested in Tokyo, and can be distributed to the LDCs or FSLs after testing. This restriction needs to be kept in mind when modeling the service network.

Chapter 4

Inventory model

To be able to decide which parts should be placed on stock in which quantity and location, we consider several mathematical inventory models available in literature. In order to model the service network for Japan, we need a multi-item, two-echelon model that can deal with lateral transshipments and emergency shipments. Although various aspects of lateral transshipments have been researched, most of these studies consider a single-item system in a single-echelon environment (Wong, Van Houtum, Cattrysse, & Van Oudheusden, 2006). Wong et al. (2006) were one of the first to study a multi-item system with regular shipments, emergency shipments and lateral transshipments. Unfortunately, they use Markov processes to perform exact evaluation, making it hard to apply the model for larger, real-life problems.

Kranenburg and Van Houtum (2009) therefore introduce an approximate evaluation procedure, and prove that it provides accurate and fast results. They assume the same model as Wong et al. (2006), but consider multiple local warehouses that are replenished by a central warehouse or RDC. Due to the assumption that this central warehouse has ample stock, Kranenburg and Van Houtum (2009) can limit the system to one echelon. Within that echelon, a distinction is made between main and regular local warehouses. Main local warehouses are able to accommodate lateral transshipments, while regulars can only receive such transshipments.

To model a situation such as the Japanese service network, we use a special case of the model from Kranenburg and Van Houtum (2009) in which all local warehouses are main local warehouses. We generalize the model with multiple RDCs, which have different leadtimes to the local warehouses. Furthermore, we assume that transport costs strongly depend on the SKU's weight in stead of taking a fixed transport cost per lane. Finally, instead of the aggregate waiting time target that is used in Kranenburg and Van Houtum (2009), we use local and aggregate MA targets. In literature, these are commonly referred to as fill rate targets. These generalizations provide slightly more detailed results, but the general evaluation principle remains the same. The rest of this chapter explains the details of our model.

4.1 Model description

Let I denote a (non-empty) set of SKUs, numbered $i = 1, ..., |I|$ and J the (non-empty) set of local warehouses, numbered $j = 1, ..., |J|$. In case a part breaks down at the customer installation base, it is replaced by an operational part. We assume that such failures follow a constant Poisson process. The failure rate, or demand for SKU i at local warehouse j is given by $M_{i,j}$. Local warehouses are replenished by multiple RDCs.

In case a demand for SKU i and local warehouse j cannot be fulfilled from stock, the local warehouse will request a lateral transshipment. Other local warehouses $k \in J, k \neq j$ are then checked for on hand stock in a pre-specified sequence, in which the first local warehouse with on hand stock delivers the part. The sequence in which local warehouses are checked is denoted by vector $\sigma(j) = \sigma_1(j), ..., \sigma_{|J|-1}(j)$. Subset $J(k,j)(\subset J)$ gives all predecessors of local warehouse k in the pre-specified sequence for local warehouse j. The leadtime for a lateral transshipment from local warehouse $k \in J, k \neq j$ to local warehouse $j \in J$ is denoted by $t^{lat}_{j,k}$ (≥ 0), with corresponding costs $C^{lat}_{i,j,k}$ (≥ 0). Because CW is an important factor in transport costs, all transport costs are dependent on the SKU. Because each SKU is rooted to one RDC, the replenishment and emergency shipment leadtimes depend on SKU $i \in I$ and destination $j \in J$.

Whenever none of the local warehouses have on hand stock for the demanded SKU, an emergency replenishment from a RDC takes place. The emergency replenishment time and costs for SKU i to local warehouse j are denoted by $t^{em}_{i,j}(\geq t^{lat}_{j,k}, k \in J, k \neq j)$ and $C^{em}_{i,j}(\geq C^{lat}_{i,j,k}, k \in J, k \neq j)$ respectively. The replenishment leadtime for SKU i to local warehouse j is denoted by $t^{reg}_{i,j}(\geq t^{em}_{i,j})$, and replenishment costs are denoted by $C^{reg}_{i,j}(\leq C^{em}_{i,j})$.

Inventory at the local warehouses is controlled by a basestock policy. The holding costs for one unit of SKU i is denoted by C^h_i. For SKU i and local warehouse j, the basestock level is denoted by $S_{i,j}$. The overview of all basestock levels can be seen as a matrix.

$$S = \begin{pmatrix} S_{1,1} & S_{1,2} & \cdots & S_{1,|J|} \\ S_{2,1} & S_{2,2} & \cdots & S_{2,|J|} \\ \vdots & \vdots & \ddots & \vdots \\ S_{|I|,1} & S_{|I|,2} & \cdots & S_{|I|,|J|} \end{pmatrix}$$

$\beta_{i,j}(S_i)$ and $\theta_{i,j}(S_i)$ are introduced as the fractions of demands for SKU i seen by local warehouse j that are satisfied by on hand stock from warehouse j or by emergency shipments respectively. $\alpha_{i,j,k}(S_i), k \in J, k \neq j$ gives the fraction of demand for SKU i at local warehouse j that is satisfied by a lateral transshipment from local warehouse k, and the complete fraction of demand for SKU i at warehouse j fulfilled by lateral transshipments is found by $A_{i,j}(S_i) = \sum_{k \in J, k \neq j} \alpha_{i,j,k}(S_i)$. Note that all demands are satisfied through one of the three modes, so $\beta_{i,j}(S_i) + \theta_{i,j}(S_i) + A_{i,j}(S_i) = 1$, $\forall i \in I, j \in J$.

4.2 Assumptions

Because a model from literature can never account for all uncertainties and exceptions that occur in reality, some assumptions are required to obtain results.

1. **Poisson demand, for initial demand and for overflow demand processes.**
The assumption that failures follow Poisson processes is very common in spare parts literature. Additionally, in an earlier master thesis at Philips' SPS department, Huyps (2015) used a generic χ^2-test to validate the assumption that SPS' spare parts demands follow Poisson processes. To perform the tests, five unique SKUs were randomly

selected. Their observed demands are compared to demand generated by a Poisson distribution with the same mean as the observed demands. The hypothesis that demand data follows a Poisson distribution could not be rejected for any of the SKUs.

2. **Inventory is controlled through a basestock policy.**
Looking at the logic from the economic order quantity rule, a basestock policy makes sense for parts that have high inventory holding costs and low demand rates. As capital asset spare parts are in general expensive and have low demand rates, it makes sense to assume a basestock policy.
3. **The central warehouses have infinite stock.**
Assuming that the RDCs have infinite stock is equivalent to sourcing parts from outside the system or modeling a lost sales system (Alfredsson & Verrijdt, 1999). This assumption enables us to limit the system to one echelon and ensures that demands for replenishment and emergency shipments can always be satisfied with constant leadtimes.

4.3 Model objective

The objective for the model is to minimize the service network costs, while satisfying certain MA targets. These consist of the local MA targets per warehouse and an aggregate target for Japan. The service network costs consist of holding costs, replenishment costs, lateral transshipment costs and emergency shipment costs. Inventory holding costs per time unit can be computed as follows: $\sum_{j \in J} C_i^h S_{i,j}$. In case a part is delivered directly from stock, the replenishment costs for SKU i to warehouse j are charged. For lateral transshipments, costs consist of replenishment tariff for SKU i to warehouse k, and the transport tariff from warehouse k to warehouse j ($C_{i,k}^{reg} + C_{i,j,k}^{lat}$). For emergency shipments, the emergency tariff for SKU i to local warehouse j is charged. The total transport costs incurred for SKU i per time unit are found as follows:

$$\sum_{j \in J} M_{i,j} \left(C_{i,j}^{reg} \beta_{i,j}(S_i) + \sum_{k \in J, k \neq j} (C_{i,k}^{reg} + C_{i,j,k}^{lat}) \alpha_{i,j,k}(S_i) + C_{i,j}^{em} \theta_{i,j}(S_i) \right). \tag{4.1}$$

For SKU i, the expected total costs per time unit can be found by:

$$C_i(S_i) = \sum_{j \in J} C_i^h S_{i,j} + \sum_{j \in J} M_{i,j} \left(C_{i,j}^{reg} \beta_{i,j}(S_i) + \sum_{k \in J, k \neq j} (C_{i,k}^{reg} + C_{i,j,k}^{lat}) \alpha_{i,j,k}(S_i) + C_{i,j}^{em} \theta_{i,j}(S_i) \right). \tag{4.2}$$

Let the local MA levels for warehouses $j \in J$ be denoted by $MA_j^{local}(S)$. These levels can then be found as follows: $MA_j^{local}(S) = \sum_{i \in I} \frac{M_{i,j}}{M_j} \beta_{i,j}(S_i)$. The aggregate Japanese MA performance is found by $MA^{ag}(S) = 1 - \sum_{i \in I} \sum_{j \in J} \frac{M_{i,j}}{M} \theta_{i,j}(S_i)$. The following optimization problem P is formulated:

$$\begin{array}{llll}
\text{Min} & \sum_{i \in I} C_i(S_i) & & \\
\text{Subject to} & MA_j^{local}(S) & \geq MA_j^{local\,obj} & \forall j \in J \\
 & MA^{ag}(S) & \geq MA^{ag\,obj} & \\
 & S_{i,j} \in N_0, & i \in I, & j \in J\ .
\end{array}$$

4.4 Evaluation

To evaluate the system of multiple warehouses, the Erlang loss model can be used. Erlang loss is a formula for the blocking probability, that helps to determine which fraction of the demand can be satisfied from on hand stock under a certain inventory level. The Erlang loss model with c servers and load ρ is given by:

$$L(c,\rho) = \frac{\frac{\rho^c}{c!}}{\sum_{x=0}^{c} \frac{\rho^x}{x!}}, \quad \rho > 0. \tag{4.3}$$

The direct demand for each local warehouse is given by $M_{i,j}$, and each warehouse has basestock level $S_{i,j}$. $\theta_{i,j}(S_i)$ can be derived by looking at the aggregate system of all mains, and is constant for all local warehouses for SKU i. The steady state behavior of stock in the aggregate system can be modeled as an Erlang loss system with $\sum_{j\in J} S_{i,j}$ servers, and load $\sum_{j\in J} M_{i,j} t_{i,j}^{reg}$. The number of emergency shipments in the aggregate system is equal to the steady state behavior of the number of idle servers in an Erlang loss system:

$$\theta_{i,j}(S_i) = L\left(\sum_{j\in J} S_{i,j}, \sum_{j\in J} M_{i,j} t_{i,j}^{reg}\right), \quad i \in I \quad j \in J. \tag{4.4}$$

For each warehouse separately, fill rates $\beta_{i,j}(S_i)$ can be roughly approximated through the Erlang loss model. The fill rates cannot be exactly computed, as they are dependent on the demand for lateral transshipments. This approximation works as follows:

$$\beta_{i,j}(S_i) \approx 1 - L\left(S_{i,j}, M_{i,j} t_{i,j}^{reg}\right), \quad j \in J. \tag{4.5}$$

The remainder of demand for SKU i at local warehouse j is satisfied through lateral transshipments:

$$A_{i,j}(S_i) \approx 1 - \left(\beta_{i,j}(S_i) + \theta_{i,j}(S_i)\right). \tag{4.6}$$

The next goal is to improve the approximation for $\beta_{i,j}(S_i)$ by including the demands for lateral transshipments. These demand processes are referred to as the overflow demand processes, which are assumed to behave as Poisson processes. We use an iterative procedure to determine the updated demand $\hat{M}_{i,j}, j \in J$, the overflow demand $\hat{M}_{i,k,j}, k \in J, k \neq j$, the new fill rates $\beta_{i,j}(S_i), j \in J$ and the fractions of demand satisfied through lateral transshipments $A_{i,j}(S_i), j \in J$. This procedure starts with the assumption that no demand is fulfilled through lateral transshipments, so we set demand $\hat{M}_{i,j} = M_{i,j}, j \in J$. The Erlang loss model is used to find $\beta_{i,j}(S_i) = 1 - L(S_{i,j}, \hat{M}_{i,j} t_{i,j}^{reg}), j \in J$ and $A_{i,j}(S_i) = 1 - (\beta_{i,j}(S_i) + \theta_{i,j}(S_i)), j \in J$. We then determine $\hat{M}_{i,k,j'}, k \in J, k \neq j'$ for one main local warehouse j' via (4.7). The product term $\Pi_{\ell \in J(k,j)}(1 - \beta_{i,\ell}(S_i))$ is 1 if $J(k,j) = \emptyset$.

$$\hat{M}_{i,k,j} = \begin{cases} \frac{A_{i,k}(S_i) M_{i,k}}{1 - \Pi_{\ell \in J, \ell \neq k}(1-\beta_{i,\ell}(S_i))} \Pi_{\ell \in J(k,j)}(1 - \beta_{i,\ell}(S_i)), & \text{if } S_{i,\ell} > 0 \text{ for at least one } \ell \in J \setminus \{k\}, \\ 0, & \text{otherwise.} \end{cases} \tag{4.7}$$

These overflow demand processes are used to find the updated $\hat{M}_{i,j'}$ and fractions $\beta_{i,j'}(S_i)$ and $A_{i,j'}(S_i)$. After finishing this procedure for all local warehouses $j \in J$, we start with j'

again. The iterative procedure is continued until $\hat{M}_{i,j}$ does not change more than ϵ, which we set equal to a very small number ($\epsilon = 10^{-6}$). Finally, the fraction of demand at warehouse k fulfilled by warehouse j is found by: $\alpha_{i,j,k}(S_i) = \beta_{i,k}(S_i)\hat{M}_{i,j,k}/M_{i,j}$. The formal evaluation procedure can be found in Table 4.1.

Table 4.1: Algorithm A: Evaluation

	Algorithm A: Evaluation
Step 1	For all local warehouses $j \in J$, $\theta_{i,j}(S_i) = L\left(\sum_{j\in J} S_{i,j}, \sum_{j\in J} M_{i,j}t^{reg}_{i,j}\right)$.
Step 2	For all local warehouses $j \in J$, $\beta_{i,j}(S_i) = 1 - L\left(S_{i,j}, M_{i,j}t^{reg}_{i,j}\right)$, and
	$A_{i,j}(S_i) = 1 - (\beta_{i,j}(S_i) + \theta_{i,j}(S_i))$.
Step 3	For one local warehouse $j \in J$:
3-a):	Determine $\hat{M}_{i,k,j}$ with Eq. (4.7), and $\hat{M}_{i,j} = M_{i,j} + \sum_{k\in J, k\neq j} \hat{M}_{i,k,j}$.
3-b):	$\beta_{i,j}(S_i) = 1 - L(S_{i,j}, \hat{M}_{i,j}t^{reg}_{i,j})$ and $A_{i,j}(S_i) = 1 - (\beta_{i,j}(S_i) + \theta_{i,j}(S_i))$.
Step 4	Repeat step 3 for all other local warehouses $j \in J$.
Step 5	Repeat steps 3 and 4 until $\hat{M}_{i,j}$ does not change more than ϵ for each $j \in J$.
Step 6	For all local warehouses $j \in J$, $\alpha_{i,j,k}(S_i) = \beta_{i,k}(S_i)\hat{M}_{i,j,k}/M_{i,j}$, $k \in J, k \neq j$.

As mentioned earlier, we consider a system with main local warehouses only. Readers that are interesting in applying this model in an environment that contains regular local warehouses that can receive, but not accommodate lateral transshipments can find the extensions to this evaluation in Kranenburg and Van Houtum (2009).

4.5 Optimization

The previous section describes how our model performs approximate evaluation for a policy of basestock levels. We are now interested in finding the optimal policy for problem P. In order to do so, we use a slightly modified version of the Greedy heuristic as proposed by Kranenburg and Van Houtum (2009).

Instead of the aggregate waiting time target used in the Greedy heuristic by Kranenburg and Van Houtum (2009), we set targets for MA levels. Because our Greedy heuristic needs to satisfy local MA targets and an aggregate MA target, we also introduce an additional Greedy step. As both waiting time and MA levels are computed based on fractions $\beta_{i,j}(S_i)$, $\theta_{i,j}(S_i)$ and $A_{i,j}(S_i)$, our Greedy heuristic should provide similar performance.

Our heuristic consists of four steps. In the first (initialization) step, all basestock levels are set equal to 0. In the second step, an iterative procedure is used to increase basestock levels $S_{i,j}$ until costs cannot be further reduced by increasing any basestock level. As costs depend on S_i only, this step is executed for each SKU separately. In each iteration, the basestock level that yields the largest decrease in costs is increased by one unit. Let $e_j, j \in J$ be a row vector of size $|J|$, with the j^{th} element equal to 1 and all other elements equal to 0. While $\Delta_j C_i(S_i) \leq 0$, the basestock level that yields the largest cost decrease is raised. The change in costs $C_i(S_i)$ in case $S_{i,j}$ is increased can be found by:

$$\Delta_j C_i(S_i) = C_i(S_i + e_j) - C_i(S_i). \tag{4.8}$$

In the third Greedy step, we check whether the local MA constraints in problem P have been met. As long as at least one local MA level is lower than its target, basestock levels $S_{i,j}$ are iteratively increased until all local MA targets are met. For each basestock policy $S_{i,j}$, the distance $d_j^{loc}(S)$ to meeting the local MA targets is determined as:

$$d_j^{loc}(S) = \left(MA_j^{local\,obj} - MA_j^{local}(S)\right)^+. \tag{4.9}$$

For each iteration, we then look at the decrease in distance to meeting the local MA targets, relative to the increase in costs if a basestock level $S_{i,j}$ is increased by one unit. The costs increase by raising basestock level $S_{i,j}$ by one unit can be found by:

$$\Delta_{i,j}C(S) = \Delta_j C_i(S_i) = C_i(S_i + e_j) - C_i(S_i). \tag{4.10}$$

The decrease in distance to meeting the local MA targets is given by $-\Delta_{i,j}d^{loc}(S)$, in which $e_{i,j}$ is an $|I| \times |J|$ matrix with value 1 on position (i,j) and zeros on other positions. Due to the lateral transshipments, a dependency exists between all local warehouses for each SKU, but this does not influence the other SKUs. To limit computation time we therefore only update $j \in J$ for the SKU i that is increased.

$$\begin{aligned}\Delta_{i,j}d^{loc}(S) &= \sum_{j' \in J} d_{j'}^{loc}(S + e_{i,j}) - \sum_{j' \in J} d_{j'}^{loc}(S) \\ &= \sum_{j' \in J}\left(MA_{j'}^{local\,obj} - \sum_{i' \in I \setminus \{i\}}\left(\frac{M_{i',j'}}{M_{j'}}\beta_{i',j'}(S_{i'})\right) + \frac{M_{i,j'}}{M_{j'}}\beta_{i,j'}(S_i + e_j)\right)^+ - \\ &\sum_{j' \in J}\left(MA_{j'}^{local\,obj} - \sum_{i' \in I}\frac{M_{i',j'}}{M_{j'}}\beta_{i',j'}(S_{i'})\right)^+.\end{aligned} \tag{4.11}$$

Ratio $\Gamma_{i,j}^{loc}$ gives us the distance to meeting the local MA targets, relative to the costs increase. The combination of i and j that yields the largest $\Gamma_{i,j}^{loc}$ is raised by one unit.

$$\Gamma_{i,j}^{loc} = \frac{-\Delta_{i,j}d^{loc}(S)}{\Delta_{i,j}C(S)}. \tag{4.12}$$

In the final step in the Greedy heuristic, we check whether the aggregate MA target has been met. For lower service levels, this is often the case. If this is not the case, we will again iteratively raise stock until the constraint is satisfied. The computation of $\Delta_{i,j}C(S)$ remains the same, and to find the decrease in distance to meeting the aggregate MA target, we introduce distance $d^{ag}(S) = (MA^{ag\,obj} - MA^{ag}(S))$. The decrease in distance $\Delta_{i,j}d^{ag}(S)$ can then be found by:

$$\begin{aligned}\Delta_{i,j}d^{ag}(S) =& \left(MA^{ag\,obj} - \sum_{i' \in I \setminus \{i\}}\sum_{j' \in J}\left(\frac{M_{i',j'}}{M}\theta_{i',j'}(S_{i'})\right) + \sum_{j \in J}\frac{M_{i,j'}}{M}\theta_{i,j'}(S_i + e_j)\right)^+ - \\ &\left(MA^{ag\,obj} - \sum_{i' \in I}\sum_{j' \in J}\left(\frac{M_{i',j'}}{M}\theta_{i',j'}(S_{i'})\right)\right)^+.\end{aligned} \tag{4.13}$$

$$\Gamma^{ag}_{i,j} = \frac{-\Delta_{i,j} d^{loc}(S)}{\Delta_{i,j} C(S)}. \tag{4.14}$$

Ratio $\Gamma^{ag}_{i,j}$ gives us the distance to meeting the aggregate MA target, relative to the costs increase. The combination of i and j that yields the largest $\Gamma^{ag}_{i,j}$ is raised by one unit. When the aggregate MA target is met, the Greedy algorithm terminates. The formal description of our Greedy heuristic follows in Table 4.2.

Table 4.2: Algorithm B: Greedy heuristic

	Algorithm B: Greedy heuristic
Step 1	Set $S_{i,j} = 0, i \in I, j \in J$.
Step 2	For each SKU $i \in I$:
2-a):	Calculate $\Delta_j C_i(S_i) = C_i(S_i + e_j) - C_i(S_i), j \in J$.
2-b):	While $\min_{j \in J}\{\Delta_j C_i(S_i)\} \leq 0$:
	1) Determine j' such that $\Delta_{j'} C_i(S_i) \leq \Delta_j C_i(S_i), j \in J$;
	2) Set $S_{i,j'} = S_{i,j'} + 1$;
	3) Calculate $\Delta_j C_i(S_i), j \in J$.
Step 3	
3-a):	Calculate $\Delta_{i,j} C(S)$, $\Delta_{i,j} d^{loc}(S)$, and $\Gamma^{loc}_{i,j}, i \in I, j \in J$ using Eq. (4.10 - 4.12).
3-b):	While $d^{loc}_j(S) > 0$:
	1) Determine i' and j' such that $\Gamma^{loc}_{i',j'} \geq \Gamma^{loc}_{i,j}, i \in I, j \in J$;
	2) Set $S_{i',j'} = S_{i',j'} + 1$;
	3) Calculate $\Delta_{i,j} C(S)$, $\Delta_{i,j} d^{loc}(S)$, and $\Gamma^{loc}_{i,j}, i \in I, j \in J$.
Step 4	
4-a):	Calculate $\Delta_{i,j} C(S)$, $\Delta_{i,j} d^{ag}(S)$, and $\Gamma^{ag}_{i,j}, i \in I, j \in J$ using Eq. (4.10, 4.13 and 4.14).
4-b):	While $d^{ag}(S) > 0$:
	1) Determine i' and j' such that $\Gamma^{ag}_{i',j'} > \Gamma^{ag}_{i,j}, i \in I, j \in J$;
	2) Set $S_{i',j'} = S_{i',j'} + 1$;
	3) Calculate $\Delta_{i,j} C(S)$, $\Delta_{i,j} d^{ag}(S)$, and $\Gamma^{ag}_{i,j}, i \in I, j \in J$.

4.6 Model verification

We implement the model as introduced above in the programming language MATLAB. To check whether the model does what it is supposed to do, we perform model verification. The evaluation is verified by exactly reproducing numerical results from Kranenburg and Van Houtum (2009) before extending our model to six warehouses. Since this extension is relatively straightforward, we are confident that the extended evaluation also provides accurate results. Secondly, we want to check whether our Greedy heuristic makes logical stocking decisions. We verify the intermediate calculations of the decrease in distance for a handful of SKUs. Additionally, the final stocking decisions based on the maximum $\Gamma_{i,j}$ are checked, and these perform as expected.

Chapter 5

Case study

This chapter is concerned with the application of the model as introduced in Chapter 4 to a dataset from the Japanese service network at Philips. In §5.1, we summarize and elaborate on the relevant input to model the Japanese service network. §5.2 focuses on the current service network configuration, while in §5.3 - §5.5, we apply the model with different network configurations. These different network configurations sometimes require alternative input or parameter settings. The differences towards the original model are explained per section. Finally, in §5.6, we perform a sensitivity analysis, in which we test the sensitivity of several input parameters. All required computations to perform this case study are executed in the programming language MATLAB.

5.1 Model input

As input for our model, we use Japanese demand data from 2016. This involves 66,224 demands, which provides us with plenty of input data. In order to keep the computation time acceptable, we consider a subset of 500 SKUs out of the total 5,649 SKUs as model input. Depending on the target settings, the computation time for 500 SKUs varies around 15 minutes and increasing the subset size leads to a more than linear increase in computation time. To compose the subset, we randomly compile five subsets of 500 parts from the full dataset, and select the subset that has the closest resemblance to the full dataset in terms of average part weight, value and demands per SKU. We also look at the distribution of network roots, part segments and BIUs. When composing these five subsets, we find that the variation between subsets is limited and that all subsets show close resemblance to the averages of the full dataset. This indicates that a subset of 500 parts is large enough to give good representation of the full dataset. The data characteristic comparison of the subset and full dataset can be found in Appendix B. We consider Singapore, Roermond and Louisville as the three RDCs, and Tokyo, Osaka, Sapporo, Fukuoka, Sendai and Okayama as the six local warehouses in the Japanese service network. Our 500 unique SKUs and six unique local warehouses thus give us the model input of $|I| = 500$ and $|J| = 6$. For each of these SKUs, we see one or multiple demands at one or multiple local warehouses. The sequence of local warehouses in our model is given by $J = \{1, 2, 3, 4, 5, 6\} = \{$Tokyo, Osaka, Sapporo, Fukuoka, Sendai, Okayama$\}$.

5.1.1 Service allocation

Before using the demand data as model input, we perform some pre-processing to our dataset in MS Excel. A part of this pre-processing is the allocation of service levels to all demands. As the required service depends on the part's segment and the repair's criticality, one unique SKU does not necessarily always have the same service requirement. One hospital may for example solely rely on a single installation, while another hospital has backup installations. System downtime for the first case is likely to be more critical than the second. For modeling purposes however, we assume that each SKU has one fixed service requirement for all demands. As explained in §3.4, the required service for the Japanese network can either be SBD or NBD. Table 5.1 shows the ratio of SBD and NBD service requirements per segment for all Japanese demands in 2016. During pre-processing, each SKU is assigned a required service such that the distribution of SBD and NBD demands per segment averages the distribution as given in Table 5.1.

Table 5.1: Distribution of service requirements per segment

Segment	SBD	NBD
Customer critical parts	61%	39%
Tubes	60%	40%
Last time buy	58%	42%
High cost frequent mover	49%	51%
Slow mover	31%	69%
New product introduction	29%	71%
Low cost frequent mover	26%	74%
End of life	18%	82%
Tools	9%	91%

Warehouse allocation

Another part of pre-processing is the allocation of demands to warehouses. In case a demand requires SBD service, it is always allocated to the nearest warehouse to enable quick delivery. In case of NBD service, the demand is allocated to Osaka or Tokyo, depending on the region of the customer. NBD demands near warehouses {1,3,5} are assigned to Tokyo. NBD demands near warehouse {2,4,6} are assigned to Osaka. Due to the required inspection, all PAL parts are assigned to Tokyo, independent of the part's required service.

In case the assigned warehouse has no on hand stock, the warehouse requests a lateral transshipment. Our model assumes a predefined sequence in which other local warehouses are checked for on hand stock. Within SPS' rule based availability check, this sequence depends on multiple aspects and is not necessarily fixed. We, however, do define a fixed sequence. The logic of our predefined sequence is based on local warehouse's size and region. We distinguish the north region with Tokyo, Sapporo and Sendai in sequence of large to small, and the south region with Osaka, Fukuoka and Okayama in sequence of large to small. The logic is that other warehouses in the same region are checked in sequence of large to small, before checking warehouses in the other region in sequence of large to small. Tokyo and Osaka will first check each other before considering their own region. We obtain the sequence $\sigma(j)$, as displayed in Figure 5.1, for local warehouses $j \in J$. For convenience, we added the warehouse indices in Table 5.2.

$$\sigma(j) = \begin{pmatrix} 2 & 3 & 5 & 4 & 6 \\ 1 & 4 & 6 & 3 & 5 \\ 1 & 5 & 2 & 4 & 6 \\ 2 & 6 & 1 & 3 & 5 \\ 1 & 3 & 2 & 4 & 6 \\ 2 & 4 & 1 & 3 & 5 \end{pmatrix}$$

Figure 5.1: Sequence of $\sigma(j)$

Table 5.2: Warehouse indices

Index	Name	Region
1	Tokyo	North
2	Osaka	South
3	Sapporo	North
4	Fukuoka	South
5	Sendai	North
6	Okayama	South

After allocating Japanese demands to warehouses, we also allocate from which RDC the parts are sourced. This allocation may be different for replenishments and emergency shipments. For replenishment orders, RDCs are allocated based on the network root. NL rooted SKUs are always allocated to RDC Roermond, while US rooted SKUs are always allocated to RDC Louisville. All virtual network roots are allocated such that the distribution of replenishment network roots is equal to the RDC distribution in Table 5.3.

Whenever SPS needs to perform an emergency shipment in the real network, RDC Singapore is always checked for on hand stock first. Because our model assumes infinite stock at the RDCs, we make an allocation that is partially based on network roots and partially based on the current actual distribution. Table 5.3 shows the distribution, and because a larger fraction of the emergency shipments is allocated to RDC Singapore than for replenishments, we allocate all virtual network rooted parts to Singapore. Also, a fraction of the NL and US rooted parts are allocated to RDC Singapore to approach the current distribution of RDC allocation. All emergency demands are shipped to the ASAP location in Tokyo, and are further distributed to the customer from there.

Table 5.3: RDC allocation of Japanese demands

RDC	Replenishment	Emergency shipments
Singapore	16%	45%
Louisville	38%	23%
Roermond	46%	32%

Leadtimes

The model from Kranenburg and Van Houtum (2009) assumes that all local warehouses are directly replenished from the central warehouse. SPS currently ships parts to Tokyo or Osaka, and further distributes them from there. In our model, we therefore make an assumption on the leadtimes and replenishment tariffs to the other local warehouses. Since there is no data available regarding the replenishment leadtime and tariff to these local warehouses, we assume that these are equal to the leadtime and tariff for shipments to Osaka. To determine the replenishment and emergency shipment leadtimes to Tokyo and Osaka, we average the gap between the shipment date and receipt date in the 2016 demand data. The resulting leadtimes can be found in Table 5.4. All emergency shipments are shipped to the ASAP location in Tokyo, and are shipped to the customer from there. Because domestic shipments can always be satisfied NBD, we assume that lateral transshipments consume 1 day, so $t^{lat}_{j,k} = 1, \forall j \in J, \forall k \in J$.

Table 5.4: Replenishment and emergency shipment leadtimes in days

Lane	t^{reg}	t^{em}
Singapore - Tokyo	1.60	1.57
Singapore - Rest of Japan	3.15	
Roermond - Tokyo	7.78	3.07
Roermond - Rest of Japan	8.27	
Louisville - Tokyo	6.81	2.92
Louisville - Rest of Japan	6.60	

Costs parameters

For demands that are satisfied from stock, we charge the replenishment costs for SKU i to local warehouse j, $C^{reg}_{i,j}$. These are calculated with fixed tariffs per kilogram of CW per lane. The CW of SKUs in our subset ranges from 0.5 to 3,388 kilogram. As elaborated on in §3.3.2, the minimal consolidated weight of airfreight shipments is 30 kilograms of CW. We assume that this threshold is always met. The tariffs for replenishments from Singapore and Louisville are based on the airfreight rate card from UPS. For shipments from RDC Roermond, the Nippon airfreight rate card is used. The replenishment tariffs used in this master thesis are confidential. Costs for lateral transshipments, $C^{lat}_{j,k}$, are obtained from Seino's Japanese domestic shipment rate card. The rate card can be found in Appendix C. As mentioned earlier, emergency shipments to Japan are always satisfied through Tokyo. The emergency shipment costs for SKU i to local warehouse j, $C^{em}_{i,j}$, is therefore equal to C^{em}_{i} in this case study. These tariffs are based on the FedEx Parcel rate cards for shipments from RDC Louisville and Roermond. The emergency shipment costs from RDC Singapore is based on the DHL Parcel rate card. The overview of emergency shipment tariffs can be found in Appendix D. Finally, the holding costs are based on the part's value. If V_i gives the part's value in euros, the annual holding costs per part C^h_i are equal to $\frac{V_i}{5}$. The part values in our dataset range from €0.20 to €107,322.

Assumptions

Besides the general model assumptions as elaborated in Chapter 4, we make some additional assumptions to apply this model for the Japanese service network. The assumptions are summarized below:

1. Chargeable weight is rounded up to match a CW in the tariffs rate cards (e.g. CW 1.1 is rounded up to 1.5 in the emergency shipment rate card);
2. The minimal threshold of 30 kilograms CW per airfreight shipment is always met, so no additional transport costs are charged;

3. Leadtimes are independent of the part's characteristics. In special cases (heavy, oversized or hazardous parts), there may in reality be a dependency due to special transport;
4. All demands for a unique SKU have the same service requirement;
5. The sequence in which warehouses are checked to perform a lateral transshipment is fixed;
6. The transport rates and leadtimes to Sapporo, Fukuoka, Sendai and Okayama are the same as to Osaka.

5.2 Cost and service sensitivity under current configuration

In this section, we model the current service network with the input as specified in previous sections. We want to obtain more insight in the effects of the different cost components and the sensitivity between costs and MA targets. We first analyze the costs under the current Japanese service network configuration and MA targets. The aggregate target for Japan is $\{0.95\}$, and the local targets for Tokyo, Osaka, Sapporo, Fukuoka, Sendai and Okayama are $\{0.93, 0.91, 0.70, 0.70, 0.60, 0.60\}$. Our first goal is to gain insight in the cost components and how these components affect the total network costs. Next, we vary the MA targets from -5% ($0.95 \times MA_j^{loc}$, $\forall j \in J$ and $0.95 \times MA^{ag}$) to $+5\%$ ($1.05 \times MA_j^{loc}$, $\forall j \in J$ and $1.05 \times MA^{ag}$) to learn more about the sensitivity between service levels and total network costs. The local and aggregate MA targets under each percentage of increase or decrease can be found in Appendix E.

5.2.1 Results

To satisfy the MA targets as specified in Table 3.2, our model places 1,230 spare parts on stock and gives us the total service network costs. The inventory holding costs for these 1,230 parts equal 48.3% of the total service network costs. Replenishment costs, emergency shipment costs, and lateral transshipment costs equal 29.3%, 16.3% and 6.1% of the total service network costs respectively. If we vary the MA targets from -5% to $+5\%$, we obtain the sensitivity between MA targets and costs. The results for total costs are visualized in Figure 5.2. This figure also includes the cost components as a percentage of the total costs. Table 5.5 gives the same results for total costs in percentages of cost increase or decrease.

The behavior of the different cost components in Figure 5.2 can be explained intuitively. As the MA targets approach 1, almost all parts need to be delivered from stock, and thus the inventory levels significantly increase. Therefore, as target levels increase, the inventory holding costs account for a larger fraction of the total costs. Similarly, as more parts are delivered from stock, the replenishment costs increase, because fewer parts need to be satisfied through other modes. Since the replenishment tariffs are relatively low, and the increase is only caused by a shift in demand fulfillment, the replenishment costs are not rising very much.

On the contrary, when more parts can be satisfied from stock, the need for emergency shipments decreases. The emergency shipment costs therefore steadily drop as the target levels increase. The proportion of lateral transshipment costs slightly fluctuates around the same costs, and is less intuitive. For lower service levels, few parts can be satisfied directly from on hand stock, so the demand for lateral transshipments is high. Because the probability that the part is available in one of the other warehouses is limited, many of these unsatisfied

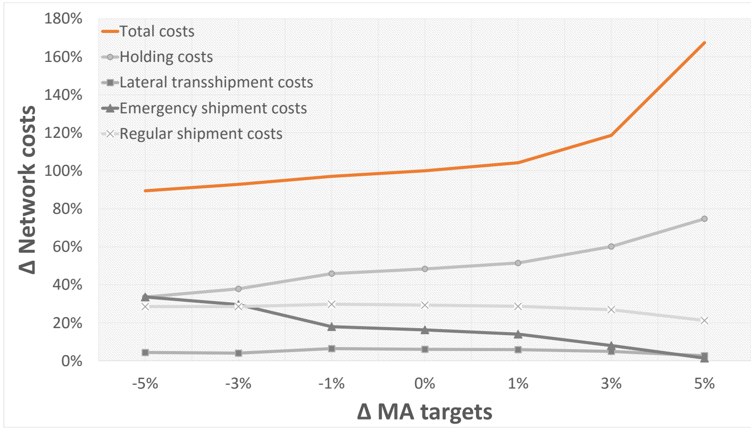

Figure 5.2: Total costs under different MA targets

Table 5.5: Total costs under different MA targets

Δ Service	Δ Total costs
-5%	-10.5%
-3%	-7.2%
-1%	-2.9%
0%	0.0%
1%	4.2%
3%	18.6%
5%	67.4%

demands are fulfilled by emergency shipments. When the demand for lateral transshipments decreases, because more parts are satisfied directly from stock, the probability that the part is available in another warehouse increases. The demand for lateral transshipments and availability in other warehouses approximately balance each other out, making the lateral transshipments a relatively steady cost component.

5.3 Optimization under waiting time constraint

SPS currently sets local MA targets to ensure that a fraction of demands can be satisfied from the nearest local warehouse, and an aggregate MA target to ensure that a certain fraction of demands can be satisfied from any Japanese local warehouse. The MA service measures thus contribute to quickly satisfying customer demands. In case a part is not available in stock, it is irrelevant for the MA service measures how long it takes to fulfill the demand through another lane. To account for the speed in which these other demands are satisfied, it may be beneficial to consider an alternative service measure.

Another way in which SPS can ensure that customer demands are satisfied quickly, is by setting an aggregate waiting time constraint. Let waiting time $W_{i,j}$ give the time it consumes to transport an operational spare part i to local warehouse j in days. When computing the waiting time, we can involve the standard delivery time from the local warehouse to the customer. As this can also be seen as a constant, we choose to leave it out the computation and only charge waiting time over the lateral transshipments and emergency shipments. The waiting time is equal to zero if the demanded part is already on stock in warehouse j, equal to $t^{lat}_{j,k}$ in case the part need to be shipped from another local warehouse k, and equal to $t^{em}_{i,j}$ in case the part needs to be shipped from one of the RDCs. To obtain an aggregate waiting time, we correct for the amount of demands per SKU per local warehouse. An aggregate waiting time service measure 'encourages' to distribute stock over FSLs and thus satisfy demands quickly, in a more natural way than the MA targets. Because the MA levels as well as waiting time are based on fractions $\beta_{i,j}(S_i)$, $\theta_{i,j}(S_i)$ and $A_{i,j}(S_i)$, the service measures can easily be compared, and the evaluation procedure remains unchanged.

Fortunately, it is possible to adjust our Greedy heuristic to an aggregate waiting time constraint, as the original Greedy heuristic from Kranenburg and Van Houtum (2009) also uses such a constraint. We introduce $\hat{W}^{obj}$ as the aggregate waiting time target and $\mathscr{S}$ as the set of all solutions. This set can be divided into a subset $\mathscr{S}^{feas} = \{S \in \mathscr{S} | \hat{W}(S) \leq \hat{W}^{obj}, \forall j \in J\}$ of feasible solutions and a subset $\mathscr{S} \setminus \mathscr{S}^{feas}$ of non feasible solutions. The aggregate waiting time is computed as follows:

$$\hat{W}(S) = \sum_{i \in I} \sum_{j \in J} \frac{M_{i,j}}{M} W_{i,j}(S_i). \tag{5.1}$$

Where:

$$W_{i,j}(S_i) = t^{em}_{i,j} \theta_{i,j}(S_i) + \sum_{k \in J, k \neq j} t^{lat}_{j,k} \alpha_{i,j,k}(S_i). \tag{5.2}$$

Trying to minimize the costs under an aggregate waiting time constraint, we use the following formulation for problem Q:

$$\begin{array}{lll} \text{Min} & \sum_{i \in I} C_i(S_i) & \\ \text{Subject to} & \hat{W}(S) \leq \hat{W}^{obj} & \\ & S_{i,j} \in \mathbb{N}_0, & i \in I, \quad j \in J\ . \end{array}$$

The distance to the set of feasible solutions $d(S)$ is then computed as:

$$d(S) = \left(\sum_{i \in I} \sum_{j \in J} \frac{M_{i,j}}{M} W_{i,j}(S_i) - \hat{W}^{obj} \right)^+ . \tag{5.3}$$

The costs increase by raising basestock level $S_{i,j}$ by one unit remains the same as in the original model and can be found by (4.10). The decrease in distance to the set of feasible solutions by increasing basestock level $S_{i,j}$ by one unit can be found by:

$$\begin{aligned} \Delta_{i,j} d(S) &= d(S + e_{i,j}) - d(S) \\ &= \left(\sum_{i' \in I \setminus \{i\}} \sum_{j' \in J} \frac{M_{i',j'}}{M} W_{i',j'}(S_{i',j'}) + \frac{M_{i,j'}}{M} W_{i,j'}(S_{i,j'} + e_j) - \hat{W}^{obj} \right)^+ - \\ & \left(\sum_{i' \in I} \sum_{j' \in J} \frac{M_{i',j'}}{M} W_{i',j'}(S_{i'}) - \hat{W}^{obj} \right)^+ . \end{aligned} \tag{5.4}$$

The decrease in distance relative to the increase in costs is then found by (5.5). We raise the inventory level of the combination of i and j that yields the highest level of $\Gamma_{i,j}$ by one unit.

$$\Gamma_{i,j} = \frac{-\Delta_{i,j} d(S)}{\Delta_{i,j} C(S)}. \tag{5.5}$$

To compare the optimization under a aggregate waiting time constraint to our original results, we compute the aggregate waiting times that correspond with the optimal solution in our original model by using (5.1) and (5.2). This yields the aggregate waiting times as displayed in Table 5.6. These results are then used as input values for the aggregate waiting

Table 5.6: Aggregate waiting time in days in original model

Δ MA targets	-5%	-3%	-1%	0	+1%	+3%	+5%
Corresponding $\hat{W}(S)$	0.270	0.233	0.193	0.170	0.145	0.090	0.033

time target $\hat{W}^{obj}$ in optimization problem Q to determine if these aggregate waiting times could be obtained under lower costs.

Additionally, we aim to optimize the aggregate waiting time in our new model under the same budget that was required to satisfy the current MA targets in our original model. We are interested in the optimal level of waiting time under this budget, and the corresponding MA levels.

5.3.1 Results

As the results in Figure 5.3 show, it is possible to obtain the aggregate waiting times from Table 5.6 under lower costs by optimizing with waiting time targets. Especially as the service requirements increase, the gap becomes larger. When optimizing the aggregate waiting time under the budget from our original model, we find that the aggregate waiting time can be reduced with 7.0%. If we compare the corresponding MA levels to the MA targets in the original model, all targets are satisfied, expect for the local target in Tokyo (91.9% instead of 93%) and the aggregate MA target (94.9% instead of 95%). The full comparison can be found in Table 5.7. These results indicate that, without specifying MA targets, but by optimizing the aggregate waiting time target with the same budget, very similar MA results and better aggregate waiting time results can be achieved. In case SPS wants to keep using MA targets instead of waiting time targets, the target settings can be reconsidered to obtain similar results. As Table 5.7 shows, a decrease of aggregate waiting time under the same budget can be achieved by redistributing the inventory over the Japanese local warehouses.

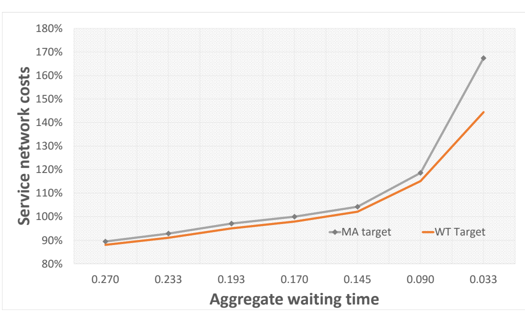

Figure 5.3: MA vs aggregate waiting time targets

Table 5.7: MA model vs waiting time model under fixed budget

	MA model	WT model
Waiting time	0.170	0.158
MA Tokyo	0.930	0.919
MA Osaka	0.917	0.939
MA Sapporo	0.718	0.750
MA Fukuoka	0.705	0.852
MA Sendai	0.601	0.763
MA Okayama	0.607	0.787
MA Japan	0.950	0.949
Parts on stock	1230	1218

5.4 Leaving RDC Singapore out of scope

Originally, SPS designed their network so that all flows to a region (APAC, EMEA, AMEC) were shipped through that region's RDC. Since freight transport possibilities have evolved and became much less expensive, it is now possible to ship all NL and US rooted parts directly from those countries to Japan, without using RDC Singapore as a hub. Because pooling effects in RDC Singapore are decreasing, and because the replenishment from RDC Singapore is fast but relatively expensive, SPS is now interested in the effects for the Japanese service network costs if RDC Singapore is completely left out of scope.

To model this scenario, we need to change the RDC allocation for replenishment and emergency shipments. For both replenishment and emergency shipments, the US and NL rooted parts are allocated to those RDCs. Virtual rooted parts are allocated to either RDC Roermond or Louisville in the same distribution as parts are currently allocated to the two RDCs. This means that we allocate 55% of replenishments and 58% of emergency shipments to Roermond, and the remainder to Louisville. This different allocation will have two significant effects towards the service network costs. First, as the replenishment leadtimes will increase, we expect the inventory levels in Japanese stockpoints to increase, which will cause inventory holding costs to rise. Secondly, the logistics costs will change but we are not yet certain whether they will increase or decrease. The replenishment costs will decrease due to the lower replenishment tariffs from Roermond and Louisville. The emergency shipment costs are expected to increase due to the higher emergency shipment tariffs from Roermond and Louisville. We are interested in the total effects of the replenishment, emergency shipments and holding costs.

Because this configuration does not only affect the costs, we are also interested in how the aggregate waiting time will change by removing RDC Singapore from our model. Because the emergency shipment leadtimes from RDC Roermond and Louisville are longer than from RDC Singapore, we expect the aggregate waiting time to increase. We compare the aggregate waiting time in this model to the original model to see how much the aggregate waiting time increases. As explained above, we expect this shift of spare parts flow to negatively influence the aggregate waiting time and emergency shipment tariffs. These negative effects are only caused by the fact that emergency shipments are no longer shipped from RDC Singapore. We therefore also run our model for a situation in which the replenishments are sourced from RDC Roermond and Louisville, but in which emergency shipments are allocated the same as in our original model.

5.4.1 Results

Due to the high difference in replenishment tariffs, removing RDC Singapore from our model surprisingly reduces the total service network costs. The holding costs and emergency shipment costs increase, but these increases are compensated by the decrease in replenishment costs. Table 5.8 and Figure 5.4 give insight in how the costs in this model change under different levels of MA targets. The total costs under the current service levels decrease by 2.1% compared to the original model. The aggregate waiting time for these service levels is 11.7% higher when RDC Singapore is left out of scope, due to the increased emergency shipment leadtimes. We can conclude that from a financial point of view, Singapore is not important for the Japanese service network replenishment, and changing the replenishment allocation and emergency decreases the service network costs.

We now run the model in which RDC Singapore is left out of scope for replenishment, but in which emergency shipments are still partially sourced from RDC Singapore. This results in 2.4% costs decrease compared to the original model, and does not negatively influence the aggregate waiting time compared to the original model. Changing the replenishment and emergency shipment allocation from RDC Singapore to the other RDCs will, however, result in loss of pooling effects in RDC Singapore. Before deciding if the replenishment allocation or replenishment and emergency shipment allocation should be changed, we recommend to study the effects of the loss of pooling effects in RDC Singapore.

Figure 5.4: Original model versus model without RDC Singapore

Table 5.8: Costs of model without RDC Singapore compared to the original model

Δ MA targets	-5%	-3%	-1%	0	+1%	+3%	+5%
Transport costs	-3.7%	-3.6%	-5.1%	-5.5%	-6.0%	-7.2%	-8.6%
Holding costs	+2.5%	+1.0%	+1.8%	+1.6%	+1.7%	+1.6%	+1.3%
Total costs	-1.6%	-1.8%	-1.9%	-2.1%	-2.1%	-1.9%	-1.2%

5.5 Decreasing replenishment leadtimes

SPS has been given the opportunity to speed up replenishment leadtimes from RDC Roermond and Louisville to the Japanese warehouses. While this change leads to an increase in the transport tariffs per kilogram of CW, it may result in lower inventory levels and thus lower inventory holding costs. We are interested in how the additional transport costs will weigh up against the increased inventory costs. Analyzing the scenarios to decrease replenishment leadtimes to Japan from RDC Roermond, from RDC Louisville or from both RDCs requires the change of input in our model as displayed in Table 5.9. For reasons of confidentiality, the rates are displayed as percentages.

Table 5.9: Replenishment leadtimes in days and tariffs in euros per kg of CW

Lane	Old leadtime	New leadtime	Old tariff	New tariff
Louisville - Tokyo	6.81	6	100%	168.4%
Louisville - Rest of Japan	6.60	6	100%	134.4%
Roermond - Tokyo	7.78	5	100%	119.0%
Roermond - Rest of Japan	8.27	5	100%	133.9%

5.5.1 Results

Looking at the changed input parameters, we expect to obtain much better results when decreasing the replenishment leadtime from RDC Roermond, than from RDC Louisville. Table 5.10 shows that decreasing the replenishment leadtime from Roermond leads to a total service network cost increase of 0.9%, while decreasing replenishment leadtime from Louisville leads to a 4.8% cost increase. Decreasing the replenishment leadtimes from both increases the costs with 5.3%. As expected, the inventory holding costs decrease, but the increased tariffs result in a larger increase in logistics costs. In terms of total service network costs, it is therefore not beneficial to decrease any on the replenishment leadtimes against higher transport tariffs. Intuitively, it is favorable to have shorter leadtimes, as this may result in slightly less variability on the leadtime and slightly more flexibility. If SPS feels like a decrease in variability and increase in flexibility may be worth a cost increase of 0.9%, it can be considered to decrease the replenishment leadtimes from RDC Roermond to Japan.

Table 5.10: Effects of decreasing replenishment leadtimes in euros

	Current leadtimes	Roermond	Louisville	Both
Total costs	100%	100.9%	104.8%	105.3%
Logistics costs	51.7%	56.4%	55.2%	59.8%
Holding costs	48.3%	43.6%	44.8%	40.2%
Parts on stock	1230	1173	1209	1139

5.6 Sensitivity analysis

To make sure that our model is not oversensitive to any of the parameters, we want to check several parameter settings. To analyze this, we change the parameter that we want to test, while keeping all other parameters constant. The results can then be compared to our original model. We perform a sensitivity analysis on C_i^h, $C_{i,j}^{reg}$, $C_{i,j}^{em}$, $t_{i,j}^{reg}$, $t_{i,j}^{em}$, $\sigma(j)$ and the subset $|I|$. Because the lateral transshipment account for only a small fraction of the satisfied demand, and the costs stay relatively constant as service levels are varied, we do not test the sensitivity for $C_{j,k}^{lat}$ and $t_{j,k}^{lat}$. The parameters of C_i^h, $C_{i,j}^{reg}$, $C_{i,j}^{em}$, $t_{i,j}^{reg}$, $t_{i,j}^{em}$, are varied with fixed percentages for all $i \in I$, $j \in J$ and $k \in J$.

As visualized in Figure 5.5, our model is quite sensitive for the parameter settings of holding costs. A 10% change of holding costs results in approximately 5% change in network costs. The parts' values therefore have a large impact on the total network costs. The sensitivity for replenishment costs is visualized in Figure 5.6. Changing the replenishment costs by 10% results in approximately 3.5% change in network costs. The model is less sensitive for the emergency shipment costs. The sensitivity can be found in Figure 5.7. Figure 5.8 shows that our model is not very sensitive for lower values of replenishment leadtimes. For replenishment leadtime increases, however, the sensitivity is higher.

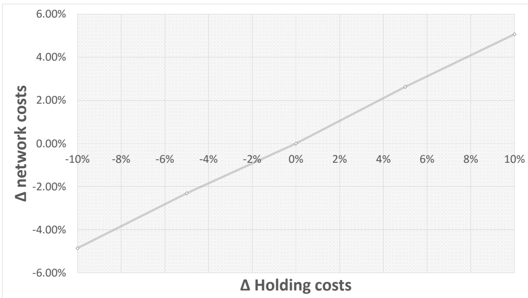

Figure 5.5: Sensitivity for C_i^h

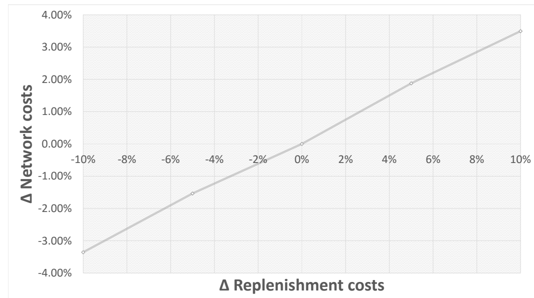

Figure 5.6: Sensitivity for $C_{i,j}^{reg}$

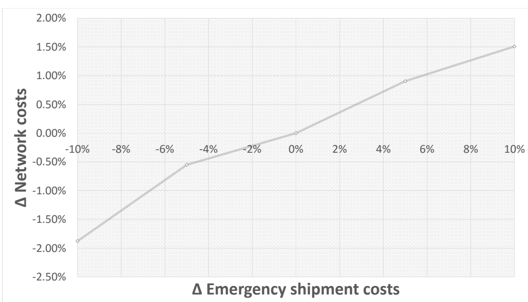

Figure 5.7: Sensitivity for $C_{i,j}^{em}$

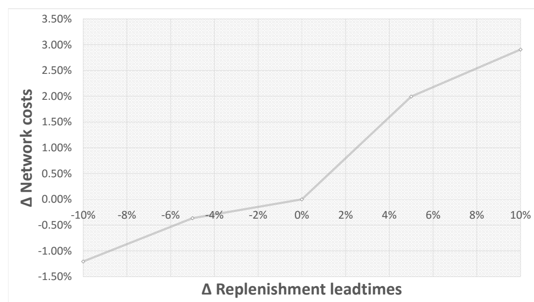

Figure 5.8: Sensitivity for $t_{i,j}^{reg}$

In terms of costs, the model is insensitive for the emergency shipment leadtimes $t_{i,j}^{em}$. Changing this parameter does however influence the aggregate waiting time, so the sensitivity between waiting time and leadtimes is considered. Figure 5.9 shows that 10% change of the emergency shipment leadtimes results in an aggregate waiting time change of almost 8%. The settings of $t_{i,j}^{em}$ is therefore important to obtain accurate results for the aggregate waiting time.

Next, the sensitivity of the sequence in which local warehouses are checked for lateral transshipments is tested. Instead of the logic as explained in §5.1.1, we look at the costs for domestic shipments. Ideally, we would also involve the replenishment costs to local warehouses $k \in J, k \neq j$ in choosing this sequence, but since the allocation to RDCs is different for each SKU, this is not straightforward. By sorting the sequence from low to high domestic shipment tariffs, we obtain the sequence as displayed in Figure 5.10. Running the model with this matrix as input for $\sigma(j)$ yields a maximum network costs difference of +0.69%, compared to the original sequence. We can conclude that the model is not very sensitive for the sequence of $\sigma(j)$.

To test if our model provides very different results when using another subset, we run one of the other randomly selected subsets of 500 SKUs. Due to slightly different averages of the part's value, demands per SKU and CW, the result for our alternative subset shows a slightly different curve. As the average part's value (€1025.16) is lower, the holding costs are also

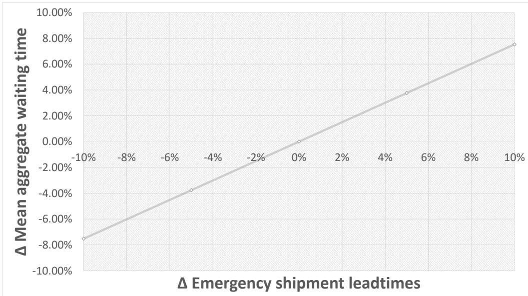

Figure 5.9: Sensitivity for $t_{i,j}^{em}$

$$\sigma(j) = \begin{pmatrix} 5 & 2 & 6 & 3 & 4 \\ 6 & 1 & 4 & 5 & 3 \\ 5 & 1 & 2 & 6 & 4 \\ 6 & 2 & 1 & 5 & 3 \\ 1 & 3 & 2 & 6 & 4 \\ 2 & 4 & 1 & 5 & 3 \end{pmatrix}$$

Figure 5.10: Alternative sequence for $\sigma(j)$

lower. This explains the less steep increase of network costs. Because the average weight is higher (16.9), the transport costs are slightly higher. Figure 5.11 shows that the alternative subset gives higher costs for lower MA targets, because holding costs play a less important role for such targets. As the target levels increase, so does the amount of inventory, which makes the original more expensive compared to the alternative subset.

Finally, we want to test the assumptions that involves the replenishment leadtimes and tariffs to Sapporo, Fukuoka, Sendai and Okayama. As these are currently replenished through Tokyo and Osaka, no actual tariffs and replenishment leadtimes are known. While we made the assumption that the costs and replenishment leadtimes are equal to the costs and replenishment leadtimes to Osaka, we might also assume that the replenishment leadtimes are longer and the tariffs are more expensive. We therefore perform a sensitivity analysis by increasing the leadtimes for these warehouses with one day, and increasing the transportation tariffs by 20%. We find that, although the model with this input obviously yields higher costs, the behavior upon changing the service levels remains the same. This can be explained by the fact that the demands that are allocated to these warehouses are not a very large fraction of the total Japanese demand.

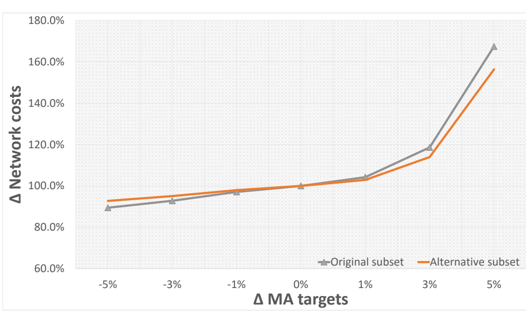

Figure 5.11: Subset sensitivity

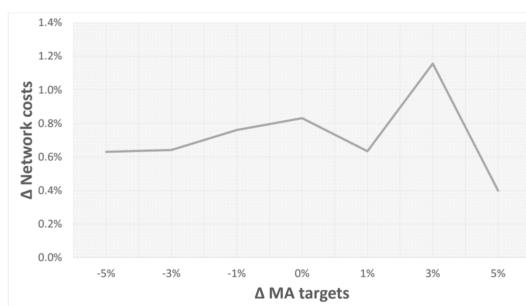

Figure 5.12: Increased leadtimes and costs for Sapporo, Fukuoka, Sendai and Okayama

Chapter 6

Conclusions and recommendations

In this chapter, we first draw conclusions by answering the research questions from §2.4. In the next section, we translate these conclusions to concrete recommendations for SPS.

6.1 Conclusions

We use mathematical inventory modeling to perform a case study for SPS' Japanese service network under different configurations. The findings from our service network analysis and the case study are used to answer the research questions. We answer the main research question as can be found below by discussing each sub research question.

Main research question: How sensitive is the relationship between the Material Availability targets and the corresponding costs for the current Japanese service network; and how do changes in the network influence this relationship?

What is the current service network configuration for the Japanese market?
The network contains six local warehouses that are geographically spread out over Japan to accommodate quick demand satisfaction. LDCs Tokyo and Osaka are directly replenished by the three global RDCs and further distribute inventory to the other local warehouses. SPS plans inventory based on local and aggregate MA targets to meet the service level agreements. In Chapter 2 and 3, we give more detailed information of the current Japanese service network configuration.

How can the Japanese service network be captured in a mathematical inventory model available in literature, so that it yields the closest possible resemblance to reality?
To model the Japanese service network, we use a special case of the model introduced by Kranenburg and Van Houtum (2009), in which all local warehouses can accommodate and receive lateral transshipments. The model also allows emergency shipments if demands cannot be satisfied from Japanese inventory. We generalize the model with multiple RDCs and dependency of transport tariffs on CW. We introduce MA targets that need to be satisfied. Due to the assumption that these RDCs have infinite stock, we can limit the model to a single-echelon.

Under what costs can the Japanese Material Availability targets be achieved? And how much do target increases or decreases affect these results?
We provide SPS with the total service network costs under which the current service levels can be achieved. The total service network costs consist of 48.3% holding costs, 29.3% replenishment costs, 16.3% emergency shipment costs and 6.1% lateral transshipment costs. Figure 5.2 gives more insight in the relationship between the MA targets and corresponding costs.

Can the level of aggregate waiting time as obtained when satisfying the Japanese Material Availability targets be achieved under lower costs?
By altering our Greedy heuristic in §5.3, we are able to compare the results from optimizing with MA targets and aggregate waiting time targets. The results show that the aggregate waiting times that correspond to optimal solutions in our original model can be obtained under lower costs. Alternatively, we use the network costs under the current service in our MA model with as input for a maximum budget in our waiting time model. We find that, under this budget, the aggregate waiting time can be reduced by 7.0%, and that this solution nearly satisfies all original MA targets. Furthermore, we show that, if SPS considers waiting time to be important, the current MA targets for Sapporo, Fukuoka, Sendai and Okayama should be reconsidered. We can conclude that planning with an aggregate waiting time target provides better financial and waiting time results than planning with local and aggregate MA targets.

How does the relationship between service and costs in the Japanese service network change if we remove the Regional Distribution Center of Singapore from our model?
Because of the increased distance from RDC Roermond and Louisville compared to RDC Singapore, we expect longer leadtimes and higher transport tariffs. The leadtimes and emergency shipment tariffs indeed increase, but surprisingly the replenishment tariffs from RDC Roermond and Louisville are lower than from RDC Singapore. Our model shows that leaving RDC Singapore out of scope actually reduces the Japanese service network costs under the same service levels by 2.1%. The reduced logistic costs compensate for the increase of inventory holding costs. The service-costs relationship is very similar as in our original model, but slightly shifted. The aggregate waiting time, however, will increase by 11.7%. We also test how the model performs if only the flow of replenishments is changed, while emergency shipments are still partially sourced from RDC Singapore. This results in 2.4% cost decrease compared to the original model, and does not negatively affect the aggregate waiting time.

Can the Japanese service network costs be decreased by choosing faster replenishment modes?
As we show in the case study in §5.5, decreasing the replenishment leadtimes reduces the Japanese inventory levels, which results in reduced inventory holding costs. The logistics costs, however, increase too much to achieve financial benefits. Decreasing the replenishment leadtimes from RDC Roermond to Japan, from RDC Louisville to Japan or RDC Roermond and Louisville to Japan, will result in increased service network costs of 0.9%, 4.8% and 5.3%, respectively.

6.2 Recommendations

Based on these conclusions, we can formulate multiple recommendations for SPS. First, we recommend SPS to use our results if they are considering to increase the MA target levels. If SPS is willing to increase network spendings for Japan with around 4% per year, the MA targets can be increased by 1%. The same holds for 11% cost increase and 2% MA targets increase and 19% costs increase for 3% MA targets increase.

Secondly, instead of the current MA service measures, we recommend SPS to adopt waiting time service measures. MA targets do not account for the delay caused by lateral transshipments or emergency shipments, while waiting time targets do. We find that especially for higher service levels, an aggregate waiting time target yields better financial results. Our model also shows that, if SPS prefers to keep planning on MA targets, the current local targets are far from optimal if waiting time is considered to be important. Changing the MA targets for Japan, Tokyo, Osaka, Sapporo, Fukuoka, Sendai and Okayama to {0.95, 0.92, 0.94, 0.75, 0.85, 0.76, 0.79} leads to aggregate waiting time reduction of 7% under the same budget. We therefore recommend SPS to adopt these target levels.

With regard to RDC Singapore, we recommend SPS to analyze the dependencies of other demand areas on this distribution center. If RDC Singapore proves to be an important component in the service network for other markets, we recommend that the loss of pooling effects when changing the Japanese replenishment flow to RDC Roermond and Louisville are studied. The Japanese emergency shipments will in that case still partially be sourced from RDC Singapore. If RDC Singapore does not prove to be an important component for other markets, SPS may consider removing the distribution center completely. The resulting increase in waiting times due to the increased emergency shipment leadtimes can then be limited by allowing cross-border lateral transshipments in the APAC market.

Next, we recommend SPS not to decrease the replenishment leadtimes from Roermond and Louisville against higher transport tariffs, if the reasoning to do so is purely cost based. Other reasons to speed up replenishment could be a slight reduction in the leadtime variability and a little bit more flexibility, thus reducing uncertainties. If SPS considers this to be worth a cost increase of 0.9%, the faster replenishment leadtime from Roermond to Japan can be chosen.

Finally, we have a recommendation that is beyond the scope of our research. SPS currently does not consider which installations are built in which countries, and has no insight in the material breakdown structure. We recommend SPS to gain insight in which components are part of which installation, to create such material breakdown structures. If these are linked to the planning system, it enables better predictions of spare parts failures. Also, this would avoid placing parts on stock in markets that do not have the installation in place that could require the part.

References

Alfredsson, P., & Verrijdt, J. (1999). Modeling emergency supply flexibility in a two-echelon inventory system. *Management science*, *45*(10), 1416–1431.

Basten, R., & Van Houtum, G. (2014). System-oriented inventory models for spare parts. *Surveys in operations research and management science*, *19*(1), 34–55.

Driessen, M., Arts, J., Van Houtum, G.-J., Rustenburg, J. W., & Huisman, B. (2015). Maintenance spare parts planning and control: a framework for control and agenda for future research. *Production Planning & Control*, *26*(5), 407–426.

Huyps, E. (2015). Integrated inventory and transportation mode selection at philips healthcare. *Master Thesis, School of Industrial Engineering*.

Kranenburg, B., & Van Houtum, G.-J. (2009). A new partial pooling structure for spare parts networks. *European Journal of Operational Research*, *199*(3), 908–921.

Royal Philips N.V. (2017, Mar.). *Financial results - investor relations. (n.d.).* Retrieved from `http://www.philips.com/a-w/about/investor/financial-reporting.html`

Van Houtum, G.-J., & Kranenburg, B. (2015). *Spare parts inventory control under system availability constraints* (Vol. 227). Springer.

Wong, H., Van Houtum, G.-J., Cattrysse, D., & Van Oudheusden, D. (2006). Multi-item spare parts systems with lateral transshipments and waiting time constraints. *European Journal of Operational Research*, *171*(3), 1071–1093.

Appendix A

List of Abbreviations

Abbreviation	Definition
AMEC	North, Central and South America
APAC	Asia and Pacific
ASAP	All Special Assistance Parts
AW	Actual Weight
BIU	Business Innovation Unit
CW	Chargeable Weight
EMEA	Europe and Middle-East
FCO	Field Change Order
FSE	Field Service Engineer
FSL	Forward Stocking Location
JP	Japan
LDC	Local Distribution Center
LVL	Louisville
MA	Material Availability
NBD	Next Business Day service
NEMO	No Engineer Material Order
NL	The Netherlands
OEM	Original Equipment manufacturer
PAL	Pharmaceutic Affairs Law
PUDO	Pick-up, Drop-off Point
RDC	Regional Distribution Center
RMD	Roermond
SBD	Same Business Day service
SGP	Singapore
SKU	Stock Keeping Unit
SPS	Service Parts Supply Chain
TARN	Acronym for Transship, Allocate, Repair, New-buy
US	United States
WT	Aggregate waiting time

Appendix B

Subset characteristics

Table B.1: Dataset characteristics

	Full dataset	Subset
Number of SKUs	5649	500
Total demand	66224	6754
Demands per SKU	11.72	13.51
Average part value (Euros)	1,043.07	1,039.64
Average Chargeable Weight	11.39	10.40
Demand allocated to:		
Tokyo	65.9%	66.8%
Osaka	25.3%	25.6%
Sapporo	1.6%	1.2%
Fukuoka	3.0%	2.7%
Sendai	1.6%	1.3%
Okayama	2.6%	2.4%
SKUs per segment:		
Slow Movers	41.1%	41.6%
Low Cost Frequent Movers	35.1%	33.4%
Customer Critical Parts	8.4%	9.0%
High Cost Frequent Movers	6.1%	6.2%
New Product Introduction	3.0%	3.6%
Last Time Buy	2.7%	2.8%
End Of Life	2.6%	3.0%
Tools	0.9%	0.2%
Tubes	0.1%	0.2%
Network roots:		
NL	43.7%	44.8%
Virtual	32.5%	32.4%
US	23.8%	22.8%

Appendix C

Domestic shipment rate card

The following tables show the Seino rate cards for Japanese domestic shipments per kilogram of chargeable weight. The empty columns represent the departure zone, and parts with CW above 500 kilograms are charged with a fixed rate per kilogram.

Appendix D

Emergency shipment rate card

The following tables show the relevant data from Nippon and UPS rate cards for Japanese domestic shipments per kilogram of CW. RMD, LVL, SGP and JP are abbreviations for Roermond, Louisville, Singapore and Japan.

Appendix E

Delta service

Table E.1: Changes MA targets

Δ MA targets	Japan	Tokyo	Osaka	Sapporo	Fukuoka	Sendai	Okayama
-5%	0.903	0.884	0.865	0.665	0.665	0.570	0.570
-3%	0.922	0.902	0.883	0.679	0.679	0.582	0.582
-1%	0.941	0.921	0.901	0.693	0.693	0.594	0.594
Current MA targets	0.950	0.930	0.910	0.700	0.700	0.600	0.600
+1%	0.960	0.939	0.919	0.707	0.707	0.606	0.606
+3%	0.979	0.958	0.937	0.721	0.721	0.618	0.618
+5%	0.998	0.977	0.956	0.735	0.735	0.630	0.630