Master of Science Thesis

MIOM05 Degree Project in Production Management

Comparing single- and multi-echelon methods for inventory control of spare parts at Volvo

Faculty of Engineering LTH

Division of Production Management

By:

Jakob Bengtsson & Alexander Larsson

Supervisor:

Johan Marklund

May, 2022

Abstract

Title: Comparing single- and multi-echelon methods for inventory control of spare parts at Volvo
Authors: Jakob Bengtsson & Alexander Larsson
Supervisors: Johan Marklund, Lund University | Christian Beckers & Johan Lidvall, Volvo
Background: Spare parts logistics is an area particularly concerned with challenges related to inventory control. The parts are vital to secure the functionality of critical equipment and the management of its supply chain tends to be complex. This thesis will focus around the inventory control of spare parts distribution within the Volvo group. Currently, the case company is using a single-node optimization approach in their inventory control process and the company is interested in improvements to their current process. Literature in the subject shows promising results when using multi-echelon approaches in inventory control. Thus, this thesis will study the effects of using a multi-echelon approach in the inventory control process at Volvo.
Purpose: The purpose of this thesis is to evaluate the use of a multi-echelon model in comparison to the currently used single-echelon model at Volvo.
Methodology: The thesis follows an operations research framework where first quantitative and qualitative data regarding the problem was gathered. Then, the inventory system was modeled in a computer program using a multi-echelon approach. Lastly, the performance of the model was evaluated through a numerical study using discrete-event simulation.
Conclusions: In this thesis, two sets of spare parts governed by two different single-echelon based models are examined. For the first set, using the multi-echelon model results in a reduction of average cost by 67 %. For the same set, the multi-echelon model also outperforms the single-echelon model in terms of meeting target service levels. For the other set, an average of 22 % cost reduction is achieved. For this set, the single- and multi-echelon model perform equally satisfactory in terms of meeting target service levels, however, the multi-echelon model provided reorder points that results in a lower total amount of stock-on-hand. The model mainly managed this by suggesting a set of reorder point allocating stock further downstream in the supply chain. Thus, the system faced a slight increase in stock at the dealers while experiencing a large reduction of stock at the RDC.

Keywords: Single-echelon, Multi-echelon, Spare parts, Inventory Control, Supply Chain Management

Preface

This M.Sc. thesis marks the end of the authors' 5 years of studies in Industrial Engineering and Management at the Faculty of Engineering LTH, Lund University, within the master's program in Supply Chain Management. This thesis has been carried out at the Division of Production Management in collaboration with Volvo Group.

We want to take the chance to send our greatest gratitude to the people who have helped us along the way in this thesis project. First of all, we would like to thank our supervisors at Volvo, Johan Lidvall and Christian Beckers, for helping us getting to know the organisation, and guiding us to the right people. We would also like to send our gratitude to Volvo-employees Max Engvall, Joakim Andersson, Niklas Samuelsson, Philip Mårtensson, and Marcus Bohman. Your assisstance has helped us gain valuable insights during the project. Lastly, we would like to thank our supervisor Professor Johan Marklund for the exceptional guidance and feedback you have provided. Your ideas and support have been of great value throughout this project.

— Jakob Bengtsson & Alexander Larsson
Lund, 2022

Contents

List of Figures

List of Tables

Abbreviations and symbols

	Definition
SKU	Stock-keeping Unit
SML	Service Market Logistics
GTO	Group Truck Operations
CDC	Central Distribution Center
RDC	Regional Distribution Center
VOR	Vehicle-Off-Road
TSL	Target Service Level
TFR	Target Fill Rate
VAU	Value of Annual Usage
EOQ	Economic Order Quantity
IMS	Inventory Management System
VMI	Vendor Managed Inventory
OWMR	One Warehouse Multiple Retailer
FCFS	First-Come-First-Served

	Definition
IL	Inventory level
$E[IL^+]$	Expected stock-on-hand
$E[IL^-]$	Expected number of backorders
IP	Inventory position
R	Reorder point
Q	Order quantity
$D(L)$	Demand during lead time
S_1	Cycle service level
S_2	Fill rate
S_3	Ready rate
μ	Mean demand during a time unit
σ	Standard deviation of demand during a time unit
μ'	Mean demand during the lead time
σ'	Standard deviation of demand during the lead time
β	Induced backorder cost
h	Holding cost
b	Backorder cost
SS	Safety stock

Chapter 1

Introduction

Section 1.1 provides introductory background for the master thesis. In Section 1.2 - Section 1.3 the case company and the distribution network studied is described. Lastly, the problem formulation with research questions and its delimitation is defined in Section 1.4 - Section 1.5.

1.1 Background

The supply chain management area is recognized by most top management to be crucial for a company's success. The control of in- and outflow of goods are of strategic importance for the company in order to achieve its purpose and further customer satisfaction. The area of inventory control has evolved during the 20th century as companies' have recognized the area to hold the potential of significant competitive advantage (Axsäter, 2006). Inventory managers are continuously confronted with the trade-off between inventory holding costs and customer service. On the one hand, a company should strive for minimizing the total tied up capital, nevertheless, there are also costs associated with unsatisfied customers not receiving their orders in time (van Donselaar et al., 2021). Advances in technology and progress in research during the past decades have enabled companies' to deal with this trade-off more efficiently. (Axsäter, 2006)

In recent years, an increasing number of companies have realized the value of approaching the trade-off between holding costs and customer service more holistically. Several inventory control techniques have evolved from research with the aim to reduce total inventory costs while reaching customer service targets for the entire supply chain, as opposed to optimizing inventory control routines at each warehouse one-by-one. One of the main difficulties to overcome in such an approach is to coordinate decisions at different warehouses within the supply chain with a limited amount of information. (Andersson et al., 1998)

An industry area especially concerned with challenges related to inventory control is the distribution of spare parts. Spare parts are often of high importance in the process of securing the functionality of critical equipment for many companies and, as a result, reliable supply of such goods is critical. Simultaneously, spare parts networks can retain tens of thousands of different spare parts which increases complexity of the inventory control process. Moreover, the distribution of spare parts is particularly challenging due to its variable demand, which can be of both slow-moving, lumpy and erratic character to name a few. (Turrini and Meissner, 2019)

This thesis will focus on inventory control of the distribution of spare parts at Volvo Group. More specifically, how a holistic inventory control method, e.g. a multi-echelon inventory control model, performs when used at the case company. The department at Volvo concerned with the distribution of spare parts, have historically used a single-node optimization method for their inventory control. As the company is currently developing a new centralized platform for their inventory control, there is an interest to understand to what extent a more holistic approach to their control could benefit their operations. With global reach and an international distribution network including many nodes and a wide range of customers, the inventory control process, inevitably, becomes a challenge to manage.

1.2 Description of the case company

Volvo Group is a global manufacturer of trucks, buses, construction equipment and marine industrial engines. The group has its headquarters located in Gothenburg, Sweden, and is divided into multiple divisions which are displayed in Figure 1.

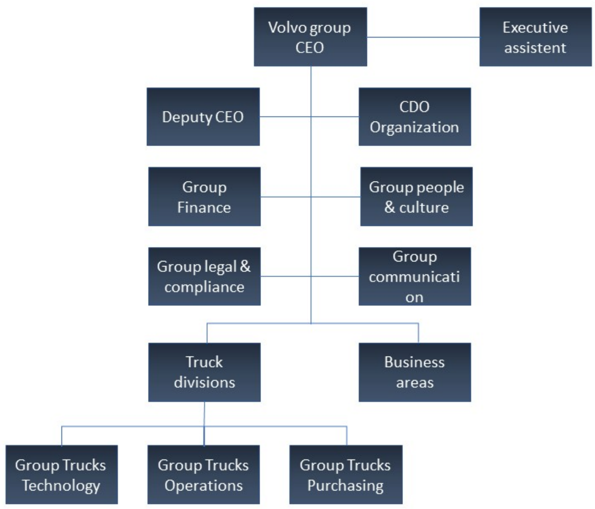

Figure 1: Volvo Group Organization (Volvo, 2022a).

This thesis will focus on the logistics of spare parts for different divisions within the Volvo Group, which is administered by the function called Service Market Logistics (SML) within Group Trucks Operations.

1.3 Current distribution network and inventory control process

The distribution network managed by the SML team is of global reach, high complexity, and handles about 700 000 articles.[1] It can conceptually be described as a set of separate 3-echelon networks. The first echelon is a single Central Distribution Center (CDC). The second echelon consists of Regional Distribution Centers, RDCs, and Supportive Distribution Centers, SDCs. The third, and final, echelon consist of a number of dealers.[2] A conceptual description of the network is illustrated in Figure 2.

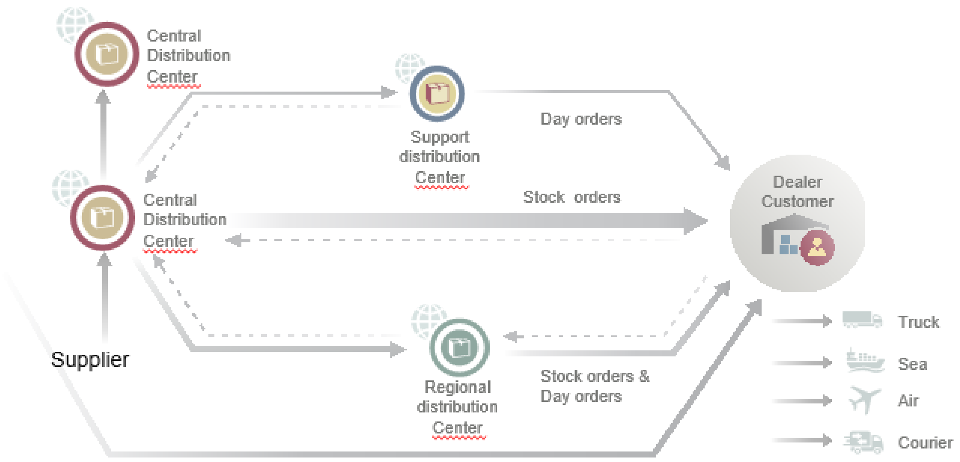

Figure 2: Volvo distribution network conceptual depiction. Figure included with the consent of Volvo Group (Volvo, 2022b).

Volvo Group has a few CDCs across the world supplying spare parts for around 40 RDCs and SDCs. Furthermore, many dealers in Volvo's network are independent but for several of these Volvo controls the inventory through a VMI (Vendor

[1]Interviewee 1, Supply Chain and Analytics Expert, Volvo GTO SML Operational Planning, interview February 24th 2022

[2]Interviewee 2, Supply Chain Data Modeling Expert, Volvo GTO SML Advanced Analytics, interview February 14th 2022

managed inventory) agreement.[3]

The network handles three separate types of orders: stock orders, day orders and Vehicle-Off-Road (VOR) orders. Stock orders refers to regular, planned orders, driven by the demand forecast. Day orders refers to emergency orders placed when stock is needed urgently, these orders are required to be shipped at the latest the morning after it is ordered. Lastly, the VOR-orders are placed in cases of severe emergency. If there is a SDC supplying the dealer, this warehouse handles the day- and VOR-orders while the RDC handles the day orders. In some areas, dealers are only supported by RDCs, in those cases the RDC handles all three types of orders.[4]

The current inventory control process at Volvo is described in Figure 3. The process can be divided into the four steps: (i) segmentation, (ii) TSL-optimization, (iii) Inventory modeling and, (iv) Real world order process. In step (i), all SKUs are divided into segments, then in step (ii) each segment is provided a target service level (TSL) by an optimization procedure aiming to minimize total holding- and back-order costs of the system. In step (iii) the inventory management system at Volvo (called MMI) mathematically models the distribution system in order to compute appropriate control policies aiming to minimize costs. Lastly, in step (iv) orders are requested from dealer to RDC and from RDC to CDC based on suggestions from the MMI-system as well as experience of Volvo employees. A detailed description of each step in the process is found in Section 2.

[3]Interviewee 2, Supply Chain Data Modeling Expert, Volvo GTO SML Advanced Analytics, interview February 14th 2022.

[4]Interviewee 2, Supply Chain Data Modeling Expert, Volvo GTO SML Advanced Analytics, interview February 14th 2022

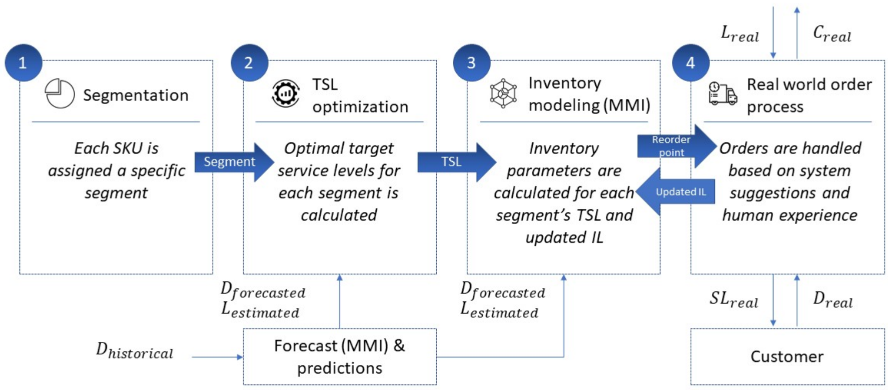

Figure 3: Volvo inventory control process.

1.4 Problem formulation

1.4.1 Inventory control process improvement areas

Volvo Group's complex supply chain and inventory control process provides numerous interesting areas to investigate. When defining the scope for this master thesis, various improvement areas for the different steps in the inventory control process as depicted in Figure 3 were considered.

Step (i), segmentation, offers a wide area of research. In literature, supply chain segmentation is widely studied and multiple different segmentation criteria and techniques have been suggested (see van Kampen et al. 2012, for an overview). There are numerous articles suggesting different dimensions and approaches to use in segmentation. Some examples include ABC-analysis (see e.g. Everette S. Gardner 1990; Flores and Clay Whybark 1986), FNS - analysis (see e.g. Cavalieri et al. 2008), Genetic Algorithms (see e.g. Guvenir and Erel 1998), decision trees (see e.g. Boylan et al. 2008) or neural networks (see e.g. Partovi and Anandarajan 2002).

In the TSL-optimization module (step ii) there are several potential areas of improvement to consider. First and foremost, when computing the optimal target service level in MMI, an optimization algorithm in order to find the most efficient TSLs is used. The research in this area is vast and there are numerous optimization algorithms available (see Böiers 2010 for an overview). Finding a more efficient algorithm for this optimization is an interesting subject of investigation.

Another area to consider in the TSL optimization step is the cost function. As the cost estimates act as inputs to the inventory modeling, it is of course important to validate the estimated costs. If the costs are inaccurate, the suggested inventory policies might be sub-optimal. Determination of holding costs is discussed by Berling (2008). However, no research describing a general way of determining backorder cost was found. The determination of these would be especially interesting to investigate as some of its cost components are difficult to quantify, e.g. badwill and future lost sales.

Another part of the process of interest is the demand forecast. The forecasting module provide inputs to both step (ii) and step (iii). When finding the optimal inventory parameters, the accuracy of the input data such as demand forecast is critical to obtain a reliable output (Basson et al., 2019). Thus, it is interesting to analyze the current forecasting method. Recent developments in forecasting technology have made good use of machine learning algorithms (see Feizabadi 2022).

Lastly, the inventory system modeling (step iii) is also an interesting area of research. An issue with the current system modeling at Volvo, is that the estimated service level for a set of inventory parameters often does not match the real situation observed in the system. This in turn, indicates improvement potential of the current model.[5] [6] The current inventory management system modeling uses a typical single-echelon method, such as in Axsäter (2006), to compute inventory parameters. Hence, the gap between model and reality could be a result of the single-echelon modeling which, e.g., uses the assumption of constant lead times. In reality at Volvo the lead times have been seen to have significant variation, sometimes because of stockouts at the upper echelons.[7]

With this issue in mind, the area of multi-echelon modeling is interesting. A multi-echelon model provides a holistic view of the inventory system, which would include the impact of delays caused by stockouts at the upper echelon. Thus, a multi-echelon model could provide a more accurate estimate of the real service levels. It is documented that for many use cases, a multi-echelon approach can produce inventory parameters that more often reach the targeted service level, or can do so at a lower cost (see e.g. Cattani et al. 2011).

Using an exact analytical multi-echelon model, like the one suggested in Axsäter

[5]Interviewee 4, Supply Chain and Analytics Expert, Volvo GTO SML Advanced Analytics, interview February 15th 2022

[6]Interviewee 3, Excellence Expert, Volvo GTO SML Advanced Analytics, interview February 23rd 2022

[7]Interviewee 3, Excellence Expert, Volvo GTO SML Advanced Analytics, interview February 23rd 2022

(2006) is not viable for a large network as the time requirements for the inventory parameter optimization would be unmanageable. Thus, approximative models are required and there are plenty of research articles describing suitable models for different network configurations (see de Kok et al. 2018 for an overview based on 394 unique papers). Even with approximation techniques the number of computations might be high, thus, clever optimization algorithms might be required, e.g., genetic algorithms (see Çelebi 2015).

A separate approach is to capture the holistic view of the network by using a simulation-optimization model to simulate the network dynamics (see e.g. Noordhoek et al. 2018). A simulation approach is capable of capturing more complex behaviors in comparison to an analytical approximative model. With such a model, different inventory parameters could be tested in the digital twin and with a smart optimization algorithm, optimal policies might be found. However, the drawback of simulation models for inventory control is that they require large amounts of development time as well as running time during optimization (Peidro et al., 2009). In conclusion, finding a proper way of modeling the network for Volvo Group Truck Operations - Service Market Logistics with a multi-echelon approach make for a challenging and exciting prospect.

1.4.2 Thesis scope

After due consideration it was decided that one of the most promising areas to analyze is the proposition of using a multi-echelon approach to model the inventory system in step (iii) of the inventory control process. Starting with investigating this step is a natural order as the value of optimizing the target service levels is highly dependent on whether the inventory modeling system can actually produce inventory parameters to realize these target service levels. Thus, improving the reliability of the modeling would also improve the value of the target service level optimization. In addition, the theoretical support for using multi-echelon approaches for modeling inventory systems is extensive and provides evidence of effects on the bottom line. Lastly, investigating a superior model is timely as it could be implemented in the new inventory management system currently being developed, *PlanIT*, and thus be useful to Volvo in the long run.

1.4.3 Research questions

Presently, Volvo is using a single-echelon optimization approach for the inventory system modeling, only focusing on a single node at a time when deciding upon optimal inventory parameters. However, the supply chain is a highly connected system where events at one node can affect others. For example, what if there is a systematic delay at the central distribution center resulting in orders being delayed to the regional distribution center, would the optimal inventory parameters still be the same or would they be different? In order to capture the system dynamics between nodes when optimizing the parameters, the system needs to be modeled as a whole - as a multi-echelon system. However, an exact analytical description of a multi-echelon system quickly becomes quite complicated, thus, clever model approximations are needed. Multi-echelon modeling is a popular research topic and there is a multitude of recent literature exploring it. Therefore, this master thesis will explore how a multi-echelon model could be implemented in the Volvo supply chain. Furthermore, the thesis will investigate what kind of improvements in terms of availability and cost efficiency that could be gained by a multi-echelon optimization approach. Thus, this master thesis aims to answer two research questions:

1. What is a suitable way of modeling the Volvo Group Truck Operations - Service Market Logistics supply chain with a multi-echelon approach?
2. What improvements in terms of spare parts availability and cost efficiency could be achieved by using a multi-echelon optimization approach?

1.5 Delimitations

Due to the limited time of the master thesis in addition to the vastness of the GTO - SML supply chain, it is not reasonable to explore the full extent of the supply chain. This master thesis will therefore focus on a smaller set of nodes and SKUs. The nodes and SKUs was chosen with the intention of being a good representation of the full network. A reasonable scope was decided to be the 2-echelon, One Warehouse Multiple Retailers (OWMR), distribution network of Volvo-Construction-Equipments located in South Africa. The network consist of a number of dealers spread through South Africa supplied by an RDC located in Johannesburg, which in turn is supplied from the central warehouse in Gent. Furthermore, Volvo is using several different sets of software, and with that, slightly varying single-echelon modeling setups to control the supply chain. As a consequence of the network delimitation this thesis will focus on the specific software used to control the South African network, which is called MMI.

1.6 Report disposition

In order for the reader to easier navigate through the thesis, this section presents the disposition of the report:

Section 1: Introduction
This chapter contains an introduction to the area of inventory control and spare parts logistics. Furthermore, a brief description of the case company and its current distribution network is included. This is followed by the problem formulation and the delimitation considered in this study.

Section 2: Mapping of the current inventory control process at the case company
In this chapter the reader is introduced to a mapping of the current inventory control process. This chapter is meant to provide the reader with a more solid understanding of the case company's inventory control process, and where in the process this study makes its contribution.

Section 3: Methodology
Initially, this chapter presents a framework for operations research projects that is used as an overall guideline throughout the report. This is followed by a description of the applied methodology focusing on more specific steps that was conducted to arrive at the final result.

Section 4: Theory
This chapter presents the theoretical concepts behind the models examined in the report. First, the reader is introduced to some basic concepts within the inventory control area which is followed by a description of a single-echelon optimization model. Eventually, the reader is introduced to multi-echelon theory together with a description of the multi-echelon method used to model the Volvo distribution network in this report.

Section 5: Numerical study
This chapter describes the numerical study used in this thesis. The multi-echelon model was evaluated with the use of a discrete-event simulation model. The reader is provided with explanations of the approach used when conducting the analytical modeling and simulation.

Section 6: Results & Analysis
This chapter presents and analyzes the results obtained from the numerical study.

Section 7: Conclusion
In this chapter the conclusions of the study are presented together with suggestion for further research.

Chapter 2

Mapping of the current inventory control process at Volvo

This section contains a mapping of the current inventory control process at Volvo. First an overview of the control process is presented in Section 2.1. Then each step of the process is described in Section 2.2 - Section 2.5.

2.1 Inventory control process overview

As of today, Volvo's organization uses several digital systems to control their vast supply chain. The four main systems are called MMI, GIM, DSP and PartsLinq. The structure of using several systems have emerged because of several restructurings of the handling of different spare parts over the years. All systems have similar processes and features and use a single-echelon modeling approach to calculate inventory parameters. Volvo has started the work to merge the different systems into one single system under the name PlanIT but at the time of the completion of this report this work is still under progress.[8] In this study the focus is on the system used for the South African network, namely, the MMI-system. Here, a thorough outline of the current procedure used with the MMI-system is presented which complements the one presented in 1.3.

The process can be described by the four steps: (i) Segmentation, (ii) TSL-optimization, (iii) Inventory modeling, and (iv) Real world order process. An overview of the procedure is seen in Figure 4.

[8] Interviewee 4, Supply Chain and Analytics Expert, Volvo GTO SML Advanced Analytics, interview February 14th 2022

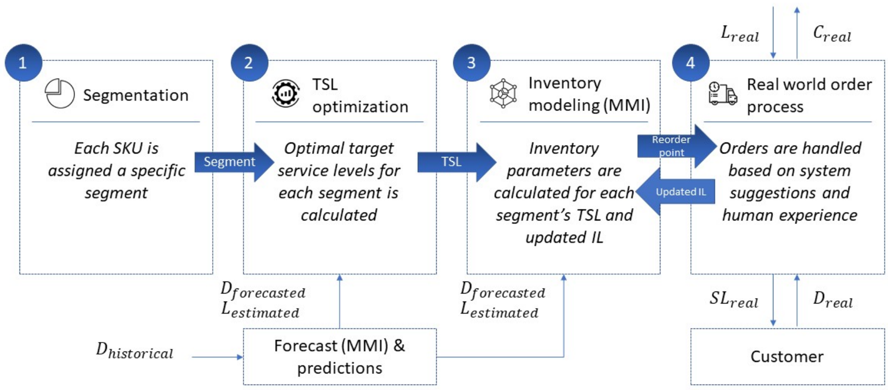

Figure 4: Volvo inventory control process.

2.2 Segmentation

The first step includes two separate segmentation processes conducted on a warehouse level. The first segmentation is done externally from the MMI system while the second segmentation is done within the MMI-system. The output of this step is that each SKU will belong to a unique segment for each warehouse. The different steps of the segmentation process are displayed in Figure 5.

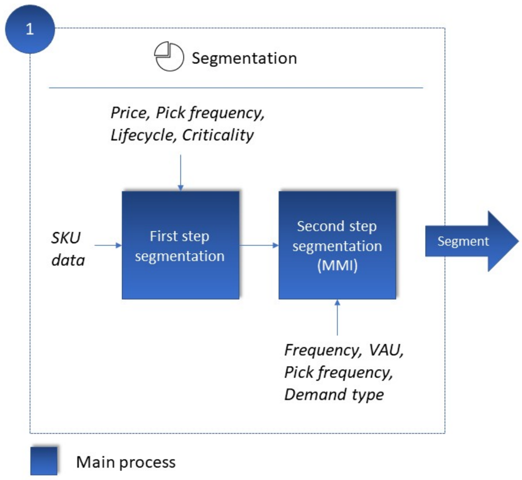

Figure 5: Segmentation process step.

First step segmentation

In the first step the SKUs are categorized into different segments by four different dimensions: price, order frequency, lifecycle and criticality. The price concerns the monetary value of the SKU and order frequency relates to how often a specific component is demanded. The lifecycle dimension is dealing with the maturity of the product whereas criticality refers to the impact on machine performance with regards to a breakdown of the particular part. These dimensions results in a segmentation chart where the SKUs are divided into 29 segments.[9]

Second step segmentation

The second step segmentation is done in MMI and is another categorization break-down. The break-down is based upon four measures: Value of Annual Usage (VAU), pick-frequency, demand type, and frequency. Noteworthy is that VAU and pick frequency are similar to dimensions price and order frequency used in the first segmentation. This indicates the existence of some overlaps in the process. The SKUs are divided into nine different pick-frequency classes based on the number of order lines per year. The VAU is calculated by multiplying price by forecasted yearly demand and is also divided into nine different classes. Furthermore, each SKU is, based on a forecast, assigned one of the three demand types: Normal, Poisson and Compound Poisson. Lastly, each SKU is also given one of two frequency classes, one-pick and multi-pick. In total, there are 14 094 possible

[9]Interviewee 4, Supply Chain and Analytics Expert, Volvo GTO SML Advanced Analytics, interview February 15th 2022

segments for each stock-keeping installation, a number that can be found by multiplying all the numbers of categories together. However, not all segments are necessarily in use. Eventually, each segment, and consequently all its allocated SKUs, is assigned a target service level.[10] This occurs in the next step in the inventory control process.

2.3 Target service level optimization

In order to decide upon target service levels for each segment, an optimization procedure which takes all SKUs within a segment into account is used. First, for each SKU, optimal inventory parameters are obtained for a set of 40 different target service levels in the range 50 % - 99 % using a single-echelon approach to model the system. Thus, the model assumes constant lead times and 100 % service level from the supplying warehouse in all installations. This is followed by an estimation of the total cost for each target service level using the calculated inventory parameters. When the total costs are calculated for each SKU in the segment, a large table is formed with the summarized costs in the segment associated with each of the 40 target service levels. Lastly, the target service level that minimizes the summarized costs of the segment is chosen.

As the algorithm finds a target service level for each segment rather than each SKU, it might not achieve the proposed minimum cost for each SKU, but this problem is theorized to be minor under the assumption that SKUs in the same segment have similar optimal target service levels. The algorithm is relatively complex and as the distribution network handles a vast number of SKUs and quite a large number of segments, it takes a significant amount of time to optimize a DC. Furthermore, the process includes a lot of manual assistance, such as data collection and qualitative assessments. The whole process takes about 30 hours to be completed. As a result, the target service levels are not updated frequently, approximately yearly, often in reaction to some change in costs or demand patterns. An overview of the TSL optimization process is displayed in Figure 6.[11] [12]

[10] Interviewee 4, Supply Chain and Analytics Expert, Volvo GTO SML Advanced Analytics, interview February 15th 2022

[11] Interviewee 4, Supply Chain and Analytics Expert, Volvo GTO SML Advanced Analytics, interview February 15th 2022

[12] Interviewee 3, Excellence Expert, Volvo GTO SML Advanced Analytics, interview February 23rd 2022

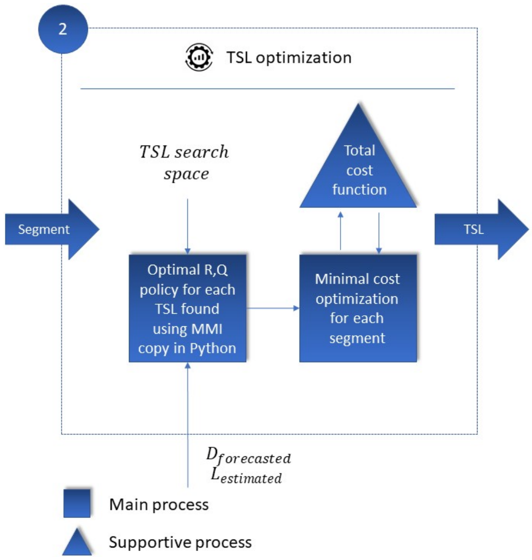

Figure 6: Target service level optimization process.

2.4 Inventory modeling

The MMI-system provided by a third party computes inventory policy parameters for each SKU based on its target service level. The inventory policy used is a continuous (R, Q) - policy, meaning that one should order Q units from the supplier when the inventory position (stock-on-hand and outstanding orders) reaches R. The order quantity, Q, is calculated by use of the EOQ - formula and then rounded to fit potential constraints like minimum and maximum possible quantity, or that Q needs to be a multiple of a specific amount. (Syncron, 2018)

When Q is set, the general system approach to find the optimal R is by iterating through increasing values of R until a service level equal to or above the target service level for the segment to which the SKU is associated is found. The search starts at $R = -Q$, or $R = E[D(L)]$ if the restriction that the safety stock needs to be greater than or equal to zero is used. The system uses a single-echelon optimization technique that finds the optimal reorder point when only the isolated single node is considered. Further, the system computes the safety stock, SS, from the reorder point with the definition $R = E[D(L)] + SS$. The safety stock is used in the next step of the process to generate orders. (Syncron, 2018)

As input for the algorithm the system thus requires a target service level associated with each segment and estimated lead times. The lead time estimation is based on the time required for different steps of a transportation route, e.g. picking, handling, transport to port, port-to-port, etc. The lead time entered into the system tries to capture the expected lead time, but is in the modeling treated as constant.[13] An overview of the inventory modeling process in MMI is shown in Figure 7.

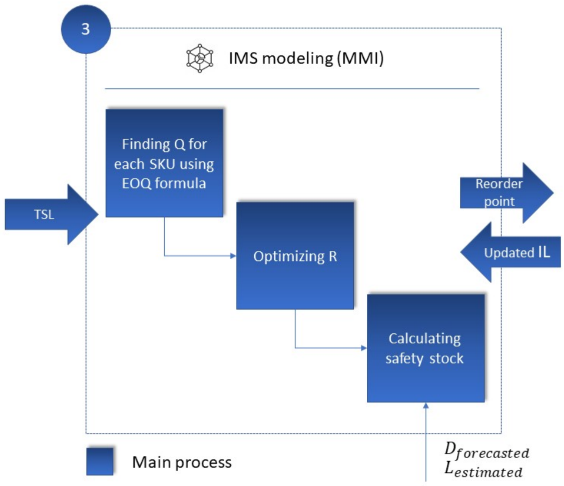

Figure 7: Processes in inventory management modeling.

2.5 Real world order process

Based on the parameters set in the previous steps, MMI will generate order suggestions for the SKUs, the system generates these suggestions once per week. During the weekly calculation, an order level is generated by adding the forecasted demand during the order lead time to the safety stock. If the inventory position is below this order level, then an order is generated. Based on the severity of the situation, an order can be generated as a regular order, or a rush order. In general, the rush order is faster but also more expensive. It might be that for a particular transport route the regular order is sent by ship, and the rush order by airplane.

When the orders are generated, an operator needs to accept the suggestions. However, the operator has the option to change order specifications, like quantities and transport mode. There are various reasons for these changes, e.g. consolidation of orders. However, many times the changes occur due to something that Volvo calls "Fair share" referring to when a spike in demand at a lower echelon results in a suggestion where a large portion of the available stock is sent. This rule is in place

[13]Interviewee 1, Supply Chain and Analytics Expert, Volvo GTO SML Operational Planning, interview February 24th 2022

to avoid stockouts at upper echelons, like CDC and RDCs, which could result in delayed orders. Thus, the order quantity is lowered so that the ordering installation only receives its "fair share".[14] The real world order processes are illustrated in Figure 8.

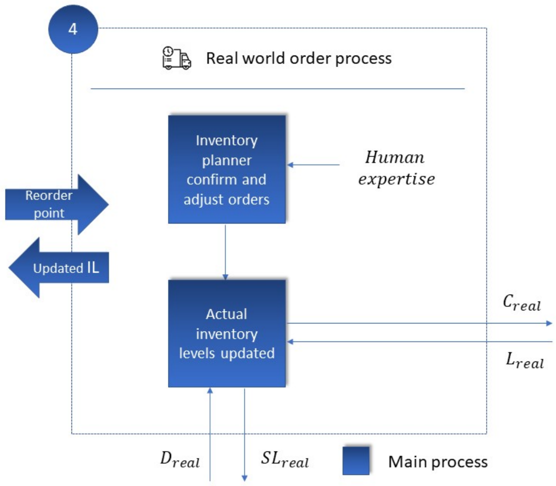

Figure 8: Real world order processes.

[14]Interviewee 1, Supply Chain and Analytics Expert, Volvo GTO SML Operational Planning, interview February 24th 2022

Chapter 3

Methodology

This section describes the scientific method applied in this thesis. The research framework used is described in Section 3.1 - Section 3.2. Its application to this thesis is delineated in Section 3.3.

3.1 Research design

This thesis is a study in two parts, the first part with the objective to understand the current inventory control process and explore improvement ideas. This part of the study falls in the category of an exploratory study see, for example, Höst et al. (2006). For an exploratory study Höst et al. (2006) suggests the method of case study, which is referring to the collection of data and information about a specific case or object, in our study, the inventory control process at Volvo. Further, as suggested by Höst et al. (2006), to gather data, the study made use of the following three techniques: (i) open interviews, which was held with selected Volvo employees; (ii) observations, the process steps was observed and analyzed; and (iii) analysis of documentation, technical documents describing the workings of the IT systems and process guides were reviewed. In the end, a mapping of the current inventory control process was created.

The second, and major, part of this thesis is the analysis of a multi-echelon modeling system for the Volvo distribution network. The purpose of this study is twofold, first, to find a suitable method to implement a multi-echelon modeling system, second, to analyze the potential improvements by employing such a system in comparison to the current process. This part of the study falls under the category of a problem solving study as described in Höst et al. (2006). For this part, an operations research framework was considered appropriate and the method used to conduct this is elaborated upon in Section 3.2 and Section 3.3 below.

3.2 Operations research framework

Several approaches have been proposed throughout the years within the operations research area. A framework describing the overall steps within a operations research study is proposed by Hillier and Lieberman (2010) and includes the following six steps:

1. Define the problem of interest and gather relevant data: This step sets an important foundation of the project and includes actions such as establishing a well-defined definition of the problem, determining appropriate objectives and constraints, and setting up a timeline for the execution of the project.

2. Formulate a mathematical model to represent the problem: Reformulate the problem into one convenient for mathematical analysis. This includes determining relevant parameters, decision variables, objective functions to optimize, and mathematical constraints to consider. As mathematical models reflecting the real world tend to become very complex, an appropriate approach is to start building a simple model and then expand it. An important trade-off in this step is the precision of the model and its tractability (the capability to efficiently produce a solution from the model).

3. Develop a computer-based procedure for deriving solutions to the problem from the model: The model is, inevitably, an approximation of reality, hence, the optimal solution of the model is not necessarily equivalent to the optimal solution for the real problem. Thus, post optimality (also called "what-if"-) analysis is important to find weaknesses of the model as well as suggesting improvements. Lastly, as the model is to be used as assistance for real-time decision making, the optimal model solution might be inconvenient to find due to time limitations and a good heuristic solution derived from an approximate optimization method might be more useful.

4. Test the model and refine it as needed: This step is also referred to as model validation and is highly dependent on the nature of the problem. One common way of validating the model is to use historical data and compare the model's outputs with the real world data.

5. Prepare for the ongoing application of the model as prescribed by management: This step refers to the extensive work of producing a computer based system and business process required for the implementation of the new model.

6. Implement: In this step the benefits of the new model is reaped and refers to the steps required for successful implementation of the new model, e.g. installing the new procedure into the computer systems and educating operators on changes in processes.

For the purpose of this thesis step 1-4 will be the main focus whereas step 5-6 is deemed out of scope for this thesis.

3.3 Applied Methodology

This section describes how the method presented in Section 3.1 and Section 3.2 was applied in this thesis work. Note that the steps have been altered slightly in order to more appropriately suit this study.

3.3.1 Step 1: Define the problem and gather relevant data

Developing a multi-echelon model that reflects the real inventory system is a challenging one. Sbai and Berrado (2019) suggests 3 actions relevant for this thesis on what data, policies, and criteria that need to be collected in order to arrive at an appropriate model. These actions are the following:

- Characterize the configuration of the distribution network
- Determine parameters relative how the system is modeled
- Decide on appropriate criteria for the network

In this report, in order to understand the context of the problem to be solved, the above actions were used to secure all necessary data needed to move on to *Step 2: Formulate a mathematical model to represent the problem* (Section 3.3.2). First, the configuration of the current distribution network was categorized by determining elements presented in Table 1.

Table 1: Network elements.

Element	Description
System specifications	• Structure of the distribution network • Number of echelons and nodes within each echelon • Relationship between nodes e.g. convergent or divergent distribution
Market specifications	• Market specific configurations e.g. what customers are served • Demand type for each SKU e.g. Normal, Possion, Compound Poisson • Reactions to disservice e.g. back-ordering, lost sales • Target service levels for each SKU
Resource specifications	• Configurations related to the capacity of the system. • Restriction on availability of resources e.g. bounded or infinite capacity • Lead times within the system • Approach to handle urgent orders e.g. using emergency shipments
Product specifications	• Type of products handled and their characteristics.

Once the configuration of the current network was completed, parameters and decision variables for modeling the system were identified. In this second action, review policy, ordering policy, prioritization system in case of stockouts, and method used to calculate the order quantity were identified. Following the second action point, key performance indicators, e.g. minimizing costs and reaching target service levels, and their importance for the company were identified for the system.

In order to attain an understanding of the current modeling process at Volvo, interviews with Volvo representatives were conducted using the interview guide found in Appendix A. Once the current assumptions and approach of modeling the system were identified, a literature review regarding multi-echelon literature was conducted in order to be able to move on to *Step 2: Formulate a mathematical model to represent the problem* (Section 3.3.2).

Interviews

The interviews held in connection with this thesis were conducted with Volvo employees with the purpose of exploring the current network. According to Höst et al. (2006) the proper structure to use for an explorative interview is an open structure. An open structured interview, as described by Höst et al. (2006), is based on questions from a prepared interview guide categorized in different areas to be explored. The order of the questions are determined by the development of the interview. The format also allows for follow-up questions conceived in the moment. The focus area of the interview will naturally depend on the expertise of the interviewee. In order to acquire a complete picture of the system, this thesis work includes interviews with people of various different roles with different responsibilities within the inventory control process at Volvo. The answers to the interviews were documented for future reference.

Literature review

A literature review with the purpose of finding material related to inventory control and especially multi-echelon modeling was conducted in this thesis work. The review was conducted according to the systematic approach presented in Höst et al. (2006): (i) search wide, (ii) select, (iii) search deep.

(i) Search wide: To find a wide sample of related articles to select from, various courses of action were taken. Sources for reference material were: Article suggestions by our supervisor, reference lists from related previous thesis work, the university database Lubsearch, reference lists from various relevant articles. The keywords employed for searching the databases were: *Multi-echelon, single-echelon, inventory control, inventory optimization, (R,Q) policy, order up to policy, stochastic demand, stochastic lead times, multi-echelon approximation, multi-echelon analytic models, spare parts, continuous review, periodic review.*

(ii) Select: The papers were selected based on their relevance for the thesis work, and the quality and credibility of the research. The relevance was based on the similarity of the network studied in the article to the Volvo network according to the dimensions presented in Table 1. When analyzing the quality and credibility of the research, focus was placed on the following areas of scrutiny as suggested by Höst et al. (2006): whether the article is peer reviewed, who is guaranteeing its credibility; is the methodology scientific; are the results conceived in a setting relevant for this thesis; and has the results been cited or confirmed in complementary work.

(iii) Search deep: Based on the findings in the papers reviewed, several areas were identified as important for further research, these were: lead time modeling, intermittent demand, and lot sizing. The Lubsearch database was used to find relevant papers as well as reference lists from the previously selected papers.

In the second step of the literature review, 82 abstracts were reviewed, 48 were selected for full paper review, and 35 were selected for inclusion in the thesis work. In the third step, six additional papers were reviewed and all six of them were chosen to be included in this thesis.

3.3.2 Step 2: Formulate a mathematical model to represent the problem

With the information and data collected about the inventory control process at Volvo during the interviews, a suitable inventory model for the distribution network was identified in the literature review. The multi-echelon approximation model suggested by Berling and Marklund (2013; 2014) was deemed suitable to represent the inventory control system for the studied distribution network at Volvo. The model chosen was hypothesized to provide a dependable balance between complexity, robustness and computational tractability for the Volvo group system. A thorough description of the model is found in Section 4.10.

3.3.3 Step 3: Develop a program for deriving solutions to the problem

The chosen mathematical model was programmed in Python (version 3.10). An important aspect when writing the code was first hand readability to ensure that the code can be easily understood in a future implementation at Volvo. As the model handles averages over discrete probability distributions defined over all positive integers, i.e. sums from zero to infinity, truncation of several sums was required. The methodology used was to iterate through the sums until reaching a cumulative probability close to one. The threshold value was chosen to be very close to one, $1 - 10^-6$, in order to avoid potential inaccuracies due to early truncation. The structure of the program is described in Appendix 7.2.

3.3.4 Step 4: Test the model

In this step, a numerical study was conducted. Following is a short description of the study, for a more detailed account see Section 5.

The purpose of the numerical study was: (i) to investigate and understand what systematic changes to expect if implementing the chosen multi-echelon model in the Volvo inventory control process and, (ii) to assess the potential of an implementation of the multi-echelon model at Volvo in terms of reaching target service levels, decreasing stock-on-hand as well as holding- and backorder costs.

The evaluation was performed by using a discrete-event simulation model to simulate the dynamics of the system in order to produce simulated values of service levels, inventory levels, and costs for comparison with the estimates provided by the two models under investigation.

The simulation model used was a discrete-event simulation model implemented in the software ExtendSim (version 9.2). The model was originally developed by researchers at the Department of Industrial Management and Logistics, Lund University, supporting a structure of one central warehouse in the upper echelon supplying ten installations facing customer demand in the lower echelon. In order to better suit the distribution network at Volvo the simulation model was expanded to support up to 15 installations at the lower echelon.

The simulation is constructed to mimic the real world behavior and is based upon historical data obtained from Volvo's inventory systems. While this evaluation will provide an initial insight to how an implementation of a multi-echelon model at Volvo will affect the outcome of the inventory control process, the real baptism by fire would be a pilot project. However, such a project would require several months of time, as well as high level decision making authority at Volvo, unfortunately, neither of which is possessed by the authors of this thesis.

Chapter 4

Theory

This section summarizes relevant theory regarding inventory control in general and multi-echelon inventory control in particular. The focus of Section 4.2 - Section 4.7 is on inventory control modeling in general. This will be followed by a description of the extension to a multi-echelon system in Section 4.8 - Section 4.9 where various models suggested in research will be presented. Lastly, the model presented in the articles Berling and Marklund (2013; 2014) is described in detail in Section 4.10.

4.1 Inventory control systems in general

An inventory control system is used to provide decision rules about how to control the inventory, e.g, when, and how much, to order at a given time when refilling stock at an installation in the distribution network. To find these decision rules, an analytical inventory model can be used which tries to mimic the behavior of the real system. The aim is to find a decision rule that minimizes total costs of inventory, while company targets regarding customer service are fulfilled. A basic inventory system modeling approach is to model each installation one-by-one according to Figure 9, which is referred to as a single-echelon model.

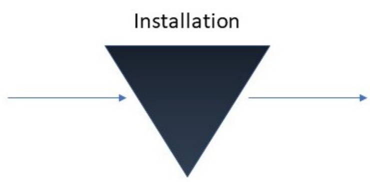

Figure 9: A single installation network.

4.2 Ordering systems

The term ordering system refers to some kind of system that tells the decision maker when and how much to order. It is common to use a policy that prescribes to order a certain *order quantity* when the *inventory position*, defined as in (1), of the system has declined to or below a certain amount, called the *reorder point*. The term *outstanding orders* in (1) refers to orders that have been placed but have not yet arrived at the installation. *Backorders* refers to customer orders that have been demanded at the installation but are not yet delivered due to a stockout. When designing an ordering system, an *ordering policy*, and a *review policy* are established. Together, these policies describe when, and how much to order. Below follows a description of these two concepts. (Axsäter, 2006)

$$\textit{inventory position} = \textit{stock-on-hand} + \textit{outstanding orders} - \textit{backorders} \tag{1}$$

4.2.1 Ordering policy

Two common ordering policies are the (R, Q) *- policy* and the (s, S) *- policy*. The (R, Q) *- policy* initiates an order with order quantity Q once the inventory position has declined to or below the reorder point, R. An alternative way of writing this policy is by using the notation (R, nQ). That is to emphasize that more than one order of size Q can be initiated when ordering. This notation can also be used in terms of Q being a fixed batch of units, e.g. a pallet, and n the parameter deciding how many batches to order. (Axsäter, 2006)

The (s, S) *- policy* initiates an order at reorder point, s, and orders enough units to reach the order-up-to level, S. The order quantity in the (s, S) *- policy* thus varies and the policy is sometimes referred to as *order-up-to policy*. A special case of the (s, S) *- policy* is the S *- policy* where $s = S - 1$, which orders up to S units whenever faced with demand. The S *-policy* is also called *base-stock policy*. (Axsäter, 2006)

The different decision variables, R, Q, s, and S are called *inventory policy parameters* and these are the decision-variables to determine when it comes to minimize costs in the systems. Cost minimization will be discussed in Section 4.6.1.

4.2.2 Review policy

Continuous review means that the inventory position is monitored continuously, and consequently, an order can be placed as soon as the inventory position reaches the reorder point. In a periodic review system, the inventory position is only reviewed at certain times, often with periodic intervals of T time units. This allows the inventory position to fall below the reorder point before the orders are placed. Thus, when using continuous review, the ordering system only needs to consider demand variations during the lead time when deciding upon parameters for reorder point and order quantity (lead time demand is elaborated upon in Section 4.5). When using periodic review, demand variations need to be considered for both the lead time and the review time. (Axsäter, 2006)

Both alternatives have their advantages, continuous review policies will reduce the required amount of safety stock and periodic review policies will allow for greater pooling of orders which can reduce transportation costs (Axsäter, 2006). From a modeling perspective, a periodic review system is slightly more complex. However, a period review system can be approximately modeled as a continuous review system; a common approximation is to add half a review period to the lead time in the continuous review model (Axsäter et al., 2013).

4.3 Inventory level and inventory position

While the decision of when and how to order depends on the inventory position, as explained in Section 4.2, the costs of the system depend on the *inventory level*. Hence, it is important to find expressions for the inventory level of a system. (Axsäter, 2006)

4.3.1 Inventory level and inventory position in steady state

The inventory level of a system in steady state of the system depends on the lead time demand and the inventory position. Let t be an arbitrary time and L the lead time, consider the time $t + L$ when all outstanding orders have arrived at the installation. The inventory level at time $t + L$, $IL(t + L)$, can be expressed as a function of the inventory position, $IP(t)$, and the demand during the time L, $D(t, t + L)$, according to (2). (Axsäter, 2006)

For the *(R, Q) -policy*, the inventory position in steady state is uniformly distributed according to (3) as shown by Axsäter (2006, p. 88). If the inventory system faces uncertain (stochastic) demand, which is most often the case when facing a competitive consumer market, both the inventory position and the inventory level are stochastic (Axsäter, 2006). To find the steady state distribution of the inventory level the distributions of the inventory position and the distribution of lead time demand is required. The estimation of a lead time demand distribution is elaborated upon in Section 4.5.

$$IL(t+L) = IP(t) - D(t, t+L) \tag{2}$$

$$IP \in U[R, R+Q] \tag{3}$$

4.3.2 Stock-on-hand and backorders

In the context of determining optimal inventory parameters, the stock-on-hand, IL^+, and number of backorders, IL^-, is of interest as these are related to the costs of holding inventory and keeping backorders, respectively. Consider the notation in (4) and (5), then the relation between inventory level, IL, stock-on-hand, IL^+, and backorders, IL^-, is as in (6). With this notation, the backorders can be expressed with a positive number, which will make the cost analysis more tractable.

$$x^+ = max(x, 0) \tag{4}$$

$$x^- = max(-x, 0) \tag{5}$$

$$IL^+ - IL^- = IL \tag{6}$$

When analyzing the system, the expected value of stock-on-hand and backorders will be of interest, these can be found using the definition of expected value in (7). For an (R, Q) policy, the maximum inventory level is $(R+Q)$ (occurring when just receiving a shipment and demand during the lead time was zero). This means that stock-on-hand takes on values between zero and $R+Q$. However, backorders can theoretically take on values between zero and infinity. With this in mind, using (6) is often advantageous when computing $E(IL^-)$.

$$E[X] = \sum_{k=-\infty}^{\infty} kP(X=k) = E[X^+] - E[X^-] \tag{7}$$

4.3.3 Backorders or lost sales

The company also needs to consider what happens in the real system once the inventory level reaches zero. There are two ways to model this situation when net stock-on-hand is zero depending on how the customers of the company react. Either, the customer is willing to wait for an order to arrive at the installation. Then, modeling the customer's order as a backorder with a corresponding backorder cost is a appropriate. Alternatively, the customer leaves without placing an order, maybe deciding to request the product from a competitor. Then, the company should consider all orders that cannot be satisfied due to stockout as a lost sale with a corresponding lost sales cost. The choice will further have an effect on the inventory level. For example, in a lost sales model the inventory level can never become negative, consequently, the optimal reorder point must be assumed to always be above 0. (Axsäter, 2006)

4.4 Service level definitions

Service levels are commonly used as key performance indicators of production and inventory systems. Service levels measures how well customers are served (i.e. receiving the right amount at the right time). From a practical point of view it is important to keep a common service level definition throughout the company that can be followed up using real data (Axsäter 2006, p.95). Furthermore, there are numerous articles discussing the importance of having centrally established objectives and targets as opposed to having different definitions throughout the company (see e.g. Rummler and Brache 1991; Shapiro 1977). Moreover, it is usually not suitable to assign the same target service level to all SKUs . To conveniently solve this problem, a segmentation on product level might be necessary to assign the same target service levels to SKUs with similar characteristics (Axsäter 2006). This section presents different service level definitions that are used in industry and methods for how to determine them.

It is up to the company and its customers to decide on what type of service level to be used but it should be noticed that although only one is implemented in the optimization of the inventory parameters, it is straightforward to adapt and analyze other key performance indicators that might be of interest (Wong et al., 2007).

4.4.1 Cycle service level, fill rate, and ready rate

Axsäter (2006) defines three different types of service levels, see Table 2. Out of all possible service level definitions, fill rate is the most common measure used in industry (Teunter et al., 2017). The mathematical formula to calculate the fill rate depends on the demand distribution. In Section 4.7 the fill rate for two different types of distributions, the Compound Poisson distribution and the Normal distribution, are calculated.

Table 2: Target service level definitions.

S_1 (Cycle service level)	Probability of no stockout per order cycle.
S_2 (Fill rate)	Fraction of demand that can be immediately satisfied from stock-on-hand.
S_3 (Ready rate)	Fraction of time with positive stock-on-hand.

4.4.2 Time based service levels

The above mentioned definitions do not take into account the time that customers facing a stockout have to wait for their order. An alternative set of service measurements that includes this consideration are time-based service levels. A common time based service level found in several real-world applications is time until the customer receives an order also called waiting time (Wheatley et al., 2015a).

4.5 Lead-time demand modeling

As explained in Section 4.3, the distribution of the lead time demand is required for determining the inventory level distribution. As the demand is unknown and uncertain, a suitable demand model needs to be estimated. When modeling the demand, a commonly used assumption is that demand during a period can be described as a non-decreasing stochastic process where increments are stationary and mutually independent. (Axsäter, 2006)

In a modeling context, it is mainly interesting to find a suitable distribution to describe the demand over a time period (the lead time). As demand often takes place in discrete amounts, rather than continuous, it makes sense to choose a discrete demand distribution for this purpose. However, using a continuous distribution, like the Normal distribution, could be more effective computationally. In this section some common choices of distributions are described.

4.5.1 Discrete demand: Compound Poisson distribution

A distribution family that is often used to model demand during a time unit is a Compound Poisson distribution. In such a distribution, the customers arrive at the installation according to a Poisson process, which means that the number of customers arriving during a time interval is described by a Poisson distribution according to (8). The λ is the expected number of customers during a time unit and t denotes the time period. Note that with this notation the time between customer arrivals can be found as $\frac{1}{\lambda}$.

Further, the demand size, j, i.e. the number of items demanded by a single customer, is an independent stochastic variable where f_j denotes the probability of a single customer demanding an order of size j. The probability that k customers demand j units, denoted by f_j^k, can be found as in (9). This means that the lead time demand is distributed according to (10). Remaining is the task of estimating a compounding distribution to describe the demand size of a single customer, f_j. A special case is to use the compounding distribution of $f_{j=1} = 1$, i.e. a customer always buys one item. In that case, the lead time demand is Poisson distributed, which is an attractive estimate due to its computational efficiency. (Axsäter, 2006)

$$P(k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, ... \tag{8}$$

$$f_j^k = \sum_{i=k-1}^{j-1} f_i^{k-1} f_{j-i}, \quad k = 2, 3, 4, ... \tag{9}$$

$$P(D(t) = j) = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} f_j^k \tag{10}$$

To find the mean and variance of demand during one time unit, (11) and (12) can be used. (Axsäter, 2006)

$$\mu = \lambda \sum_{j=1}^{\infty} j f_j \tag{11}$$

$$\sigma^2 = \lambda \sum_{j=1}^{\infty} j^2 f_j \tag{12}$$

4.5.2 Discrete demand: Logarithmic compounding distributions

Many different choices of compounding distributions describing the number of units a single customer orders have been suggested in literature, one example is the

Logarithmic distribution displayed in (13). Parameters α and λ are found according to (14) and (15), respectively. Here, μ' and σ' are estimates of mean and standard deviation of demand during the lead time. (Axsäter, 2006)

$$f_j = -\frac{\alpha^j}{ln(1-\alpha)j}, \quad j = 1, 2, 3, ... \tag{13}$$

$$\alpha = 1 - \frac{\mu'}{(\sigma')^2} \tag{14}$$

$$\lambda = -\frac{\mu'}{L}\frac{(1-\alpha)ln(1-\alpha)}{\alpha} \tag{15}$$

When using this compounding distribution, it can be shown that the lead time demand distribution is a Negative Binomial distribution (NBD). Then the lead time demand can be described as in (16), with parameters p and r according to (17) and (18), respectively. Using the NBD expression in (16) instead of the expression for the Compound Poisson distribution in (10) reduces the computational complexity of the model. (Axsäter, 2006)

$$P(D(t) = k) = \frac{r(r+1)...(r+k-1)}{k!}(1-p)^r p^k, \quad k = 1, 2, ... \tag{16}$$

$$p = 1 - \frac{\mu'}{(\sigma')^2} = \alpha \tag{17}$$

$$r = \mu'\frac{(1-p)}{p} \tag{18}$$

4.5.3 Discrete demand: Geometric compounding distribution

Another common assumption is that demand size, f_j, is distributed according to a delayed Geometric distribution as in (19) with parameter β found according to (20) (Axsäter, 2006). This distribution is sometimes referred to as stuttering Poisson distribution (SPD). (Turrini and Meissner, 2019)

$$f_j = (1-\beta)\beta^{j-1}, \quad j = 1, 2, 3, ... \tag{19}$$

$$\beta = 1 - \frac{2}{1 + (\sigma)^2/\mu} \tag{20}$$

In Figure 10 two examples of lead time demand distributions with Logarithmic- and Geometric compounding distributions, respectively, are shown. Here, mean and standard deviation of the distributions are equal. According to Axsäter (2006) the Logarithmic distribution is a better choice because of the simplified computations offered while the two distributions show significant similarities in the probability function.

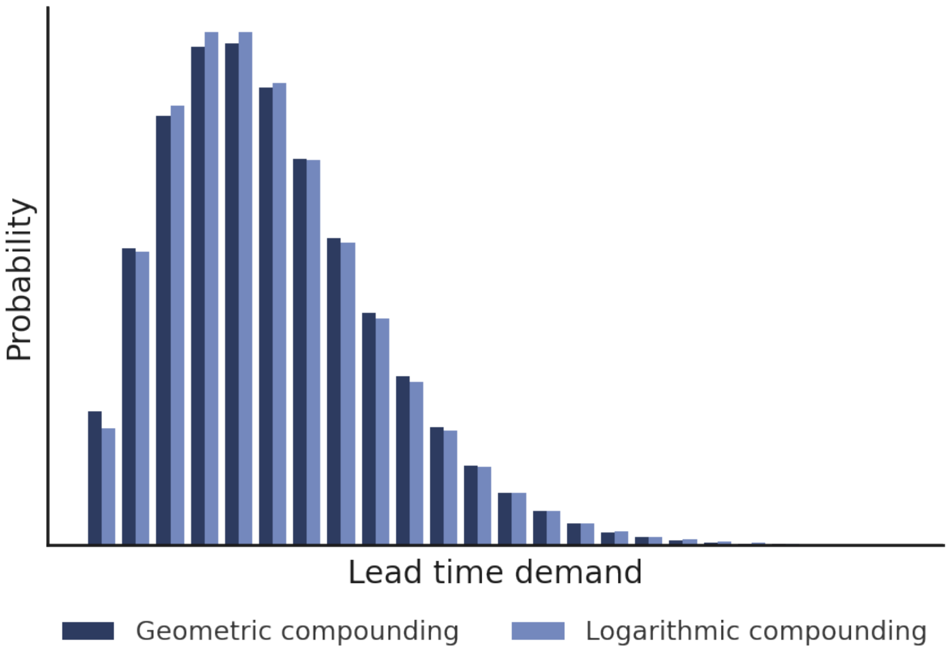

Figure 10: Examples of lead time demand from a Compound Poisson process with Logarithmic vs. Geometric compounding distribution.

4.5.4 Discrete demand: Empirical compound distribution

In practice, the distribution of the customer demand size may be subject to large and irregular variations resulting in difficulties to find a reasonable fit to any standard distribution. In this case, modeling the demand process as a Compound Poisson process with empirical compounding distributions could be appropriate. That is, the compounding distribution is estimated using the relative frequency for the real data set (i.e. based on orders history during a chosen time period, the probability of demand size j is set to the frequency of demand size j during the period). (Berling and Marklund, 2014)

4.5.5 Continuous demand: Normal distribution

If demand during the lead time is high, regardless of the real distribution, approximating the demand distribution with the Normal distribution could be suitable, according to the central limit theorem. For Normal demand, the lead time demand distribution is estimated with density function and distribution function according to (21) and (22), respectively. (Axsäter, 2006)

$$f(d)_{D(L)} = \frac{1}{\sigma'\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{d-\mu'}{\sigma'}\right)\right] \tag{21}$$

$$F(d)_{D(L)} = \int_{-\infty}^{d} f(x)_{D(L)}\, dx \tag{22}$$

While the Normal approximation grants convenient calculations, a drawback is that it allows negative lead time demand. Especially for smaller lead time demand, the probability of this might be impactful and the approximation will be poor. (Axsäter, 2006)

4.5.6 Continuous demand: Gamma distribution

To remedy the problem of of negative demand realisations experienced when using the Normal distribution, a Gamma distribution can be used. Then, the lead time demand is approximated with density function and distribution function according to (23) with parameters according to (24), (25) and (26). (Axsäter, 2006)

$$g(x) = \frac{\omega(\omega x)^{r-1} e^{\omega x}}{\Gamma(r)}, \quad x \geq 0 \tag{23}$$

$$\omega = \mu'/(\sigma')^2 \tag{24}$$

$$r = (\mu'/\sigma')^2 \tag{25}$$

$$\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} dx \tag{26}$$

In Figure 11 examples of lead time demand density functions where demand is modeled with Normal- and Gamma distribution are shown. The mean and variance of lead time demand is equal in the two examples. In the example the possibility of negative demand of the Normal demand distribution is clearly showcased.

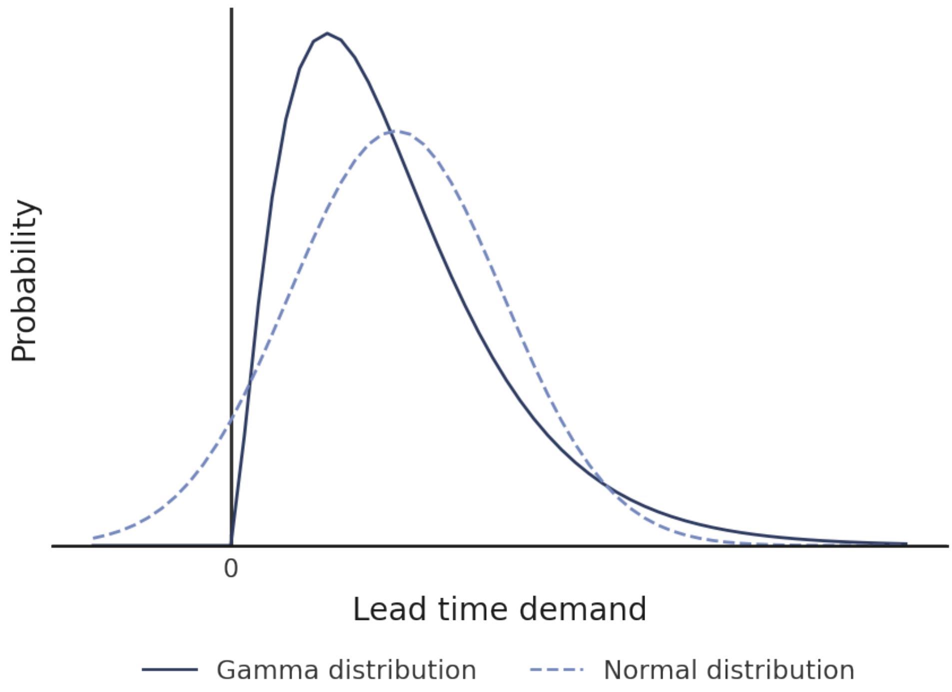

Figure 11: Normal- vs Gamma distribution.

4.5.7 Distribution fitting

The above mentioned distributions are common choices when estimating lead time demand, however, any distribution may of course be used. When trying to fit a distribution to demand, a common statistical approach is to perform goodness of fit tests on historical data, e.g. the Kolmogorov-Smirnov test. Turrini and Meissner (2019) used the Kolmogorov-Smirnov test and tested the five distribution types presented: Poisson, NBD, SPD, Normal, and Gamma on two large datasets containing spare part demands. Their work showed the Normal distribution to fit poorly to the datasets while the NBD and SPD showed most potential.

In practice, when large number of SKUs are handled, it might be more convenient to use some rule of thumb when choosing a demand distribution estimate instead of doing distribution fitting and goodness-of-fit tests for every single SKU. For discrete demand, Axsäter (2006) suggests using the variance-to-mean ratio (σ^2/μ). If the ratio is approximately equal to 1, he suggests using a Poisson distribution and if variance-to-mean is above 1.1, Compound Poisson is suggested. Furthermore, Axsäter (2013) concludes that the Normal distribution leads to large errors when used to model lead time demand of discrete nature and low mean, about 1-2 units during the lead time. For the use of the Normal distribution, Axsäter (2013) suggests a rule of thumb as follows: use the Normal approximation if expected lead time demand exceeds 10 and the variance-to-mean ratio is at most 2.

4.5.8 Estimating parameters from data

Regardless of the parametric family of distributions chosen, the parameters are unknown and need to be estimated. One could use the *method of moments* estimator with historical data, however, as the lead time demand under consideration is not necessarily described by the same parameters as historical demand, it might be more suitable to use some forecasting method (Altay and Litteral 2011). E.g. exponential smoothing, Box-Jenkins, etc. When designing the forecasting process it should be remembered that most distributions require at least two moments, the mean and the variance. Thus, both these estimates need to be produced in the forecasting process.

4.5.9 Stochastic lead times

When modeling lead time demand, the lead time (which refers to time between placement of the order until its arrival) is often estimated with a constant value. In reality, the lead time is rarely constant. More likely is that it holds variation from many sources, e.g. delays due to stockouts at supplying warehouse, variation in travel times, delays in handling, delays in re-packing locations. If (i) the orders are sequential, meaning that they can not pass each other in time, (ii) the lead time and demand during a time unit is independent, and (iii) the lead time mean and variance are known. Then, the mean and variance of lead time demand is found according to (27) and (28). (Axsäter 2006, p. 122)

$$E[D(L)] = \mu E[L] \tag{27}$$

$$VAR[D(L)] = \sigma^2 E[L] + \mu^2 VAR[L] \tag{28}$$

4.6 (R, Q) - policy cost minimization

The overall objective in inventory control is usually to minimize inventory related costs while adequately satisfying customer demand. This section will describe two general approaches regarding how to formulate the optimization problem to find optimal inventory policy parameters assuming a continuous review (R, Q) - policy. The first approach is an unconstrained minimization of holding- and backorder costs in the system whereas the second is a constrained optimization problem focusing on reaching target service levels while minimizing holding costs.

4.6.1 Cost minimization of (R, Q) - policies

If using the first approach, the optimization problem to find the R and Q that minimizes the total cost of the system can be formulated as (29) (Axsäter 2006, p.45). Here, the holding- and backorder costs are assumed linear in the inventory level. The holding cost per unit and time unit, h, refers to all costs related to holding an item in stock. It can be based on e.g. capital costs, storage space costs, and, tax and insurance costs. These types of costs are usually not too difficult for the company to determine. (Berling 2005, p.8-13)

The backorder cost per unit and time unit, b, represents the financial loss occurring when failing to meet demand and is usually more challenging to obtain. For example, the backorder cost could include a type of badwill cost which corresponds to the financial loss associated with customers' decrease in trust when the retailer cannot deliver. Estimation of these financial losses tend to become subjective and customer specific. (Axsäter 2006)

$$\min_{R,Q} C(R, Q) = \min_{R,Q}(hE[IL^+] + bE[IL^-]) \tag{29}$$

4.6.2 Cost minimization of (R, Q) - policies with service level constraints

An alternative to by-pass the problem of finding the backorder cost is by removing that part of the objective function and instead add a target service level constraint as in (30). Furthermore, for the single-echelon situation, this approach also allows for excluding the holding cost, h. In the single-echelon case in this approach, to minimize holding costs, is equivalent to minimizing stock-on-hand.

$$\min_{R,Q} C(R,Q) = \min_{R,Q}(hE[IL^+])$$
$$s.t. \quad SL \geq SL_{target} \tag{30}$$

While establishing a suitable target service level is a challenging task, it is often simpler and more practical compared to estimating a shortage cost (Axsäter 2006, p. 45). To our knowledge, the literature does not offer any extensive frameworks about determining appropriate service levels. However, this could be explained by the company-, customer- and SKU-specific problem that the determination of target service level boils down to.

If the fill rate or ready rate is used as service level definitions, then it is possible to evaluate defined service levels by translating them into shortage costs by the relations presented in (31) and (32), which hold for Compound Poisson demand and Normal demand, respectively. Here, R^* is the optimal reorder point for the installation. Moreover, if the lead time demand is Poisson, then (31) holds for the fill rate, S_2, as well (Axsäter 2006, pp. 103, 105). In a way, by setting a target service level, the decision-maker is determining a shortage cost indirectly. While the actual shortage cost may be hard to estimate, this method may provide managers with a tool to evaluate and find proper target service levels. The motivation being that it is easier to interpret and reason around a cost value compared to a service level percentage. (Axsäter, 2006)

$$S_3(R^*) \leq \frac{b}{h+b} \leq S_3(R^*+1) \quad (Ready\ rate) \tag{31}$$

$$S_2(R^*) = S_3(R^*) = \frac{b}{h+b} \quad (Fill\ rate) \tag{32}$$

Since appropriate target service levels might be easier to determine compared to backorder costs, this alternative to finding optimal inventory parameters could be advantageous (Axsäter, 2006). This approach is also preferred in industry where managers tend to set target service levels rather than specify backorder costs (Hopp et al., 1997).

4.7 Mathematical formulas for optimization of continuous (R, Q) - policy

As described in Section 4.6 there are two approaches of finding optimal inventory policy parameters, i.e. R and Q, for the continuous (R, Q) - policy controlled system. Preconditions for solving the optimization problems presented is to know what demand type the system is experiencing (Axsäter, 2006). Note that while the first, unconstrained, optimization approach does not require calculations of the service level, companies that use this approach may still want to compute estimates of the service level to use as a performance indicator.

4.7.1 Demand modeled as a compound Poisson process

When demand follows a compound Poisson process the fill rate, S_2, for the inventory system can be determined according to (33) (Axsäter 2006, p. 98).

$$S_2 = \frac{\sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \min(j,k) \cdot f_k \cdot P(IL = j)}{\sum_{k=1}^{\infty} k f_k} \tag{33}$$

In order to determine the item fill rate, S_2, $P(IL = j)$ is required for all positive values of j as can be seen from (33). However, as the maximum inventory level is $R + Q$ as described earlier in Section 4.3, the probabilities are only needed for the interval $1 \leq j \leq R + Q$. These probabilities are found according to (34) with lead time demand, $P(D(L) = d)$ as derived in Section 4.5.1, (10). Note that for $R = -Q$, there is no probability for positive inventory levels, according to (34). This means that for such reorder points, the fill rate, S_2, in (33) is equal to zero. (Axsäter, 2006)

$$P(IL = j) = \frac{1}{Q} \sum_{k=max(R+1,j)}^{R+Q} P(D(L) = k - j) \quad j \leq R + Q \tag{34}$$

With (34), $E[IL^+]$ and $E[IL^-]$ is found by using (6) and (7) presented in Section 4.3.2 earlier. Now, everything needed to solve both the unconstrained and constrained optimization problems for the situation of Compound Poisson demand is presented.

4.7.2 Normally distributed lead time demand

If the system instead is experiencing a normally distributed lead time demand (or approximated to do so), (35) is used to calculate the fill rate, S_2. Here, the fact that the fill rate, S_2, and the ready rate, S_3, are equal for continuous demand is used. (Axsäter, 2006)

$$S_2 = S_3 = 1 - P(IL \leq 0) = 1 - F_{IL}(0) = 1 - \frac{\sigma'}{Q}\left[G\left(\frac{R-\mu'}{\sigma'}\right) - G\left(\frac{R+Q-\mu'}{\sigma'}\right)\right] \tag{35}$$

In (35), $F(x)$ denotes the distribution function for the inventory level, which is found according to (36). The function $G(x)$ is called the *loss function* and is defined as in (37). φ and Φ are the density and probability functions of the $\mathcal{N}(0, 1)$ - distribution, given in (38). (Axsäter, 2006)

$$P(IL \leq x) = F_{IL}(x) = \frac{\sigma'}{Q}\left[G\left(\frac{R-x-\mu'}{\sigma'}\right) - G\left(\frac{R+Q-x-\mu'}{\sigma'}\right)\right] \tag{36}$$

$$G(x) = \int_x^\infty (v-x)\varphi(v)dv = \varphi(x) - x(1-\Phi(x)) \tag{37}$$

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{x^2}{2}\right) \qquad \Phi(x) = \int_{-\infty}^x u\varphi(u)du \tag{38}$$

To find $E[IL^+]$ and $E[IL^-]$, the density function of the inventory level, $f(x)$ as defined in (39), is used with the definition of expected value for continuous stochastic variables as stated in (40) and (41). Also here the relation in (6) can be used for more efficient computations. (Axsäter, 2006)

$$f_{IL}(x) = \frac{1}{Q}\left[\Phi\left(\frac{R+Q-x-\mu'}{\sigma'}\right) - \Phi\left(\frac{R-x-\mu'}{\sigma'}\right)\right] \tag{39}$$

$$E[IL^+] = \int_0^\infty u f_{IL}(u)du \tag{40}$$

$$E[IL^-] = \int_{-\infty}^0 u f_{IL}(u)du \tag{41}$$

4.7.3 Finding optimal inventory policy parameters for continuous (R, Q) - policy under service constraints

The minimization of the cost function under service level constraints, as presented in Section 4.6.2, can be approached in two ways. Either, the reorder point, R and order quantity, Q, are optimized simultaneously. Alternatively, the order quantities are pre-determined by some process and then the cost function is optimized by altering the reorder points. While the latter approach cannot guarantee to find the global optimum, it is computationally more efficient. Additionally, it has been shown that using the *EOQ* - formula, a commonly used deterministic method of lot sizing, the upper bound for the increased costs due to sub-optimal order quantity is small (Axsäter, 1996). Furthermore, many articles on the subject of inventory control modeling make the assumption of given order quantities (de Kok et al., 2018).

$$Q_{EOQ} = \sqrt{\frac{2A\mu}{h}} \tag{42}$$

The *EOQ* - formula is presented in (42), where A is the fixed setup cost for an order. Recall that μ represents the demand during one time unit and h is the holding cost per unit and time unit. The formula was first published by Harris (1913) and computes the optimal order quantities under several assumptions, e.g. the whole batch quantity is delivered at the same time, ordering and holding costs per unit are assumed to be constant, and, demand is deterministic and constant. Many other methods are available, some of them suggesting adjustments of the model proposed in 1913 in order to increase its reflection of reality. Continuous delivery of batches, and inclusion of quantity discounts are two examples of extensions to the model. It is also noteworthy that in practice the order quantity may also be constrained by factors such as the capacity of transportation mode or load carrier sizes. Also, the method does not include the environmental aspect of transporting where full load trucks are preferable. Thus, it becomes important for each company to identify what assumptions and constraints are plausible for their specific distribution network.

With the order quantity, Q, given, the procedure of determining the optimal reorder point, R^*, is remaining. Since both the holding cost and the service level are increasing functions of the reorder point, the optimal, R^* that minimizes the costs in (30), can be found as the minimal reorder point satisfying the service constraints. Finding R^* may be achieved by brute force, searching through all possible R - values, within a reasonable interval. However, a more refined algorithm may be used, e.g. the bisection search. (Axsäter, 2006).

Once the reorder point has been determined, the safety stock SS, defined as the average remaining stock just before an order is received, can be calculated using (43). Recall that μ' refers to the average demand during the lead time.

$$SS = R - \mu' \tag{43}$$

4.8 Multi-echelon inventory control

So far, the focus has been on inventory management for a single stock location, however, most inventory systems does not consists of a single installation, but contain multiple warehouses where stock moves between warehouses in different stages (echelons) before it finally reaches the customer. A system in multiple stages is called a multi-echelon system.

In the literature, a commonly studied multi-echelon system is structured as in Figure 12, this is a 2-echelon system with a central warehouse supplying several smaller warehouses which in turn service customers. This type of divergent setup is often found in distribution networks. The installations in the lower echelon which serve customers directly are often referred to as retailers (or dealers). To describe a multi-echelon system, more refined inventory control models are required in order to accurately capture the dynamics between stock in different locations. The theory of multi-echelon modeling is the area of researching and describing such models. (Axsäter, 2006)

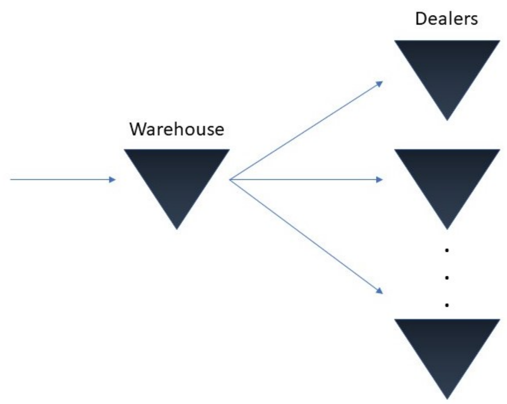

Figure 12: A divergent 2-echelon system, one-warehouse-multiple-retailers.

4.8.1 Multi-echelon inventory policies

Designing an optimal ordering system in multi-echelon systems is a very complex task as the optimal ordering policy for an installation may depend on the stock held at every location in the network. An approach for a centralized ordering policy, which means that the policy considers information from every installation in the network when suggesting an ordering decision, is to use the concept of *echelon stock inventory policy*. In such a policy, an installation orders when its echelon stock inventory position, IP_e which is defined as the inventory position for the installation plus all inventory positions of connected installations downstream, declines to or below a certain *echelon reorder point*, R_e. Axsäter (2006).

A centralized policy is attractive as every installation might react to any occurrence in any downstream installation. Consider Figure 13, with a centralized policy, an order could be triggered at the installation at echelon 1 directly by demand faced by the installation in echelon 4. Furthermore, the use of a centralized policy counteracts the bullwhip effect (see e.g. Lee et al. 1997 for details) as information about fluctuations in demand is seen and acted upon at all stages, which negates the possibility for fluctuations to increase as they move through the supply chain.

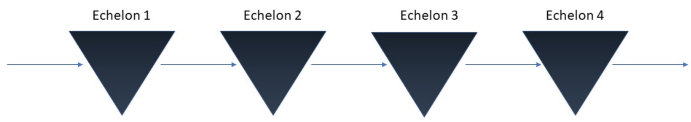

Figure 13: A 4-echelon inventory system.

However, in practice it might be complicated to use a centralized decision rule due to different entities in charge of inventory management at different locations, or that the information required for the centralized policy is not available at every installation. Thus, a de-centralized installation stock policy is a popular choice, referring to the use of a regular policy like the (R, Q) - policy as described in Section 4.2.1 at every installation, but with its parameters produced by a model considering all echelons. Again considering Figure 13, with an installation stock policy, an order can only be triggered at the installation in echelon 3 if an order is received from the installation in echelon 4 and so on. (Axsäter, 2006)

4.8.2 Implications on inventory modeling

As discussed in Section 4.7.3, the EOQ-formula is commonly used in the single-echelon setting to provide an optimal order quantity. In the multi-echelon case, the computation of optimal lot sizes can quickly become vastly complex as the network grows. The reason being that lot sizes do not only affect stock levels at the current installation, but also in upstream locations (Axsäter, 2006). An approximation model ensuring a close to optimal solution is presented by Muckstadt and Roundy (1993). However, in practice the lot sizes are often determined separately for each installation similarly to the procedure described for the single-echelon setting. (Axsäter, 2006)

Determining the optimal reorder points as discussed in Section 4.7.3 is also not directly applicable to the multi-echelon setting. As explained, dynamics in the system allows stock at one location to affect stock and service at another. E.g. if the safety stock is large at a downstream location, a lower service level (and thus a longer average delay before an order may be sent downstream) may be allowed for at the upper echelon (Axsäter, 2006). Moreover, the objective of minimizing costs under service constraints is no longer equivalent to minimizing stock-on-hand since holding costs may differ between different installations. These characteristics make the problem of finding optimal reorder points in a multi-echelon setting complicated and computationally costly in comparison to the single-echelon case.

There are no general solutions to describe the multi-echelon case, which means that the models used are required to be adapted to the system design in question, in order to produce a satisfactory solution (Axsäter, 2006). In Section 4.9 a number of multi-echelon models using different approximations and methods for determining near-optimal inventory parameters for different setups are presented.

4.9 Multi-echelon inventory modeling

In recent years, interest in multi-echelon theory has been substantial amongst researchers. A search for “multi-echelon inventory management” with an interval spanning the last 30 years (1992 - 2022) on Google Scholar produces 18 300 articles. Needless to say, there are numerous suggestions on decision rules, optimization models and heuristics to analytically describe and operate different types of multi-echelon networks. In this section a general overview of different modeling approaches, system setups, assumptions and approximation techniques will be described, all based on a thorough literature review.

4.9.1 Exact- vs approximative techniques

The task of analytically modeling multi-echelon networks is a tricky one, nevertheless, several exact models for different setups exist in literature. The models describe the behavior of a certain network setup under certain assumptions without any mathematical approximation. For example, Graves (1985) provides an exact model for a network of repairable parts under one-for-one replenishment, and Axsäter (2000) describes a model for exact evaluation of a 2-echelon network using continuous installation stock (R, Q) -policy and facing Compound Poisson demand.

de Kok et al. (2018) provides an extensive literature overview of the multi-echelon inventory control area. Of the 394 articles studied, 221 were classified as using approximative approaches. Likewise, a majority of papers studied in this thesis work also used approximative approaches. The motivation for using an approximative model rather than an exact model is to provide a more useful model, i.e. a model with less computational complexity (see e.g. Berling and Marklund 2013; Özkan et al. 2015; Yang et al. 2011). Oftentimes, companies control large networks of multiple stocking installations and a vast number of articles, especially in the spare parts business, which renders computationally complex models impractical.

4.9.2 Different types of system setups

As the inherent complexity of multi-echelon problems makes general solutions hard to find, research papers often focus on a specific setup and set of assumptions regarding the network structure, number of echelons, and policies, etc. While many of the models described in literature produce promising results in the associated numerical studies, few are feasible for use in practical applications. Generally, this is a consequence of quite restrictive assumptions, e.g. specific demand distributions, identical retailers in terms of policy and demand faced, or that the models are computationally too demanding Berling and Marklund (2013; 2014). It is not unusual that researchers build upon each-other's (or their own) previously published models to test different assumptions, and thus the robustness of the model, and to suggest extensions for a more general model setting.

A majority of the systems in the literature studied in this thesis investigates a divergent 2-echelon setup as depicted in Figure 12 in Section 4.8. This reflection is shared by the study conducted by de Kok et al. (2018) in which 244 out of 394 articles examine a system consisting of two echelons. Articles examining a general number of echelons motivates their choice by mentioning that many supply chain in today's world consist of more than 2-echelons but also due to the fact that they want to contribute with a more dynamic approach (see Caggiano et al. 2007 and Verrijdt and de Kok 1995). The choice of number of echelons studied has implications on the feasibility of different techniques, for example, Caggiano et al. (2007) mentions that lagrangian based techniques is less efficient with increasing number of echelons.

A way of making models more dynamic and robust is by considering the extension of non-identical retailers. In practice, it is quite obvious that in most cases, different retailers in a network are not identical, e.g. they do not face the same exact demand. Despite this, several papers such as Seifbarghy and Jokar (2006) and Andersson et al. (1998) assume identical retailers where the latter has been generalized to include non-identical retailers by Andersson and Marklund (2000). While the paper studying the case of identical retailers is important as it provides the basis for the model, clearly, the extension to non-identical retailers makes the model more suitable for use in practice.

Another common extension, seen in several papers, is the consideration of emergency- and lateral transshipments (i.e. shipments between retailers in the same echelon). In practice, emergency shipments are especially common within the distribution of spare parts, as the ability to react quickly to technical failures is important (Özkan et al., 2015). Examples of studies that include emergency shipments are Howard et al. (2015), Van den Berg et al. (2016) and Özkan et al. (2015). Emergency shipments are associated with extra costs but may also enable companies to deliver according to customer expectations. Therefore, an important factor to investigate is at what point in time emergency shipments should replace normal replenishment orders, as examined in e.g. Howard et al. (2015).

4.9.3 Model Generalization

The demand distribution is another crucial input to the model. Choosing a distribution reflecting the real world demand can be difficult, however, throughout literature there are a few proposed methods to find an appropriate one (see Section 4.5.7). In reality, it is a strict assumption that all products face the same type of demand. Thus, to make the model more useful in practice, it should be generalized to handle several different demand distributions. In Berling and Marklund (2014) the authors examine how the model initially developed in Andersson et al. (1998) performs with normally distributed demand. The authors also conclude that together with the result from Berling and Marklund (2013), which tests the same model with

Compound Poisson distributed demand, their model has shown to offer a flexible heuristic that can deal with different types of demand distributions at different retailers.

4.9.4 Ordering system and service objective considerations

Throughout industry, different types of order policies, review policies, and different definitions of service levels or measures, are used, and this is also reflected in the literature. In the research, there is a wide mix of different policies and service objectives used (see de Kok et al. 2018 for an overview). To name a few, Wheatley et al. (2015b) uses a base-stock policy with a time-based constraint, Minner et al. (2003) models a periodic, order-up-to-policy, while Axsäter et al. (2013) investigates a (R, Q) - policy with fill rate constraints, and Lagodimos and Koukoumialos (2008) examines the use of linear rationing rules to control the stock distribution. Regarding review period, most articles studied consider continuous review, like Axsäter et al. (2013). Even for practical situations where a periodic review policy is used, a continuous review model can still be accurate (Axsäter, 2006). However, for retailers with a longer and considerable review period the review time can have an impact. Axsäter et al. (2013) uses the approach of adding half of the review period to the lead time when modeling the lead time demand. This is a common method of compensating for the review period (Axsäter, 2006).

4.9.5 Decomposition techniques

A common modeling approach is to decompose the system into single-echelon systems, transforming the multi-dimensional multi-echelon optimization problem into a number of simpler, preferably one-dimensional, optimization problems like what was presented earlier in (29) and (30) in Section 4.6. The goal is to construct these simpler optimization problems so that their solutions constitutes a near-optimal solution to the original multi-dimensional optimization problem. This means that the techniques used is required to produce optimization problems that capture the dynamics of the full system appropriately. One common approach to capture upstream dynamics when optimizing policies in the downstream installations is by using a lead time estimate which depend on the upstream stock policy.

Several techniques to find such an estimate have been suggested, a famous example is the METRIC - approach where the lead time is assumed to consist of a constant part representing the transportation time between the stock locations considered and a stochastic waiting time due to delays at the warehouse during stockouts, see (44). The METRIC approximation is then to simply model the stochastic waiting time as its mean. This approach was introduced by Sherbrooke (1968), the article studies a system with base-stock policy and identical retailers with Poisson demand at all installations. For such a system the waiting time in (44) can be determined exactly. The approach has since been used in many articles (see e.g. Ahire and Schmidt 1996; Zang et al. 2017; Axsäter 2003). Two advantages of METRIC type approaches are that they are simple and computationally efficient for many practical applications. A drawback is however that they disregard the lead time uncertainty (Axsäter, 2006).

$$\bar{L}_i = L + E(W_0) \tag{44}$$

Berling and Marklund (2013) analyzed the performance of their multi-echelon decomposition model using two different approaches regarding lead time estimates. First by only estimating the lead time as its mean by (44). Second, by using the same estimate for the mean but also including the lead time uncertainty as described in (28)in Section 4.5.9. In their research, they use an estimate for the lead time variance as $V(L) = V(l + W) = V(W)$. They found that both approaches worked well with their model, however, only adjusting the first moment of lead time demand (its mean) was superior when the objective was to meet target fill rates with minimum stock. Adjusting for the lead time uncertainty performed better when the objective was to minimize the total expected holding and backorder costs.

Berling and Marklund (2013) explains this behaviour being due to a combination of three factors: (i) adjusting for lead time variance always produce lead time demand estimates with larger variance, (ii) reorder points are discrete which means that a value that produce a service realisation exactly on target is not always possible, and (iii) the service constraints force the optimization to find reorder points that produce service levels equal to or above the target. This means that the service will often be slightly over target and then the underestimation in lead time variance by the first method acts as compensation to reduce stock slightly (Berling and Marklund, 2013). When optimizing with backorder costs the optimization is not forced in either direction and thus, a more exact estimation of lead time demand ought to perform better. For further reading on the subject Andersson and Marklund (2000) provide both approximate and exact expressions for the lead time variance in a 2-echelon, non-identical retailers network setting.

4.9.6 Multi-echelon inventory parameter optimization

The lion's share of papers studied in this thesis work provide models under the assumption of given order quantities, that is, Q - values are pre-determined and only reorder points, R, is considered in the optimization. The motivation for this decision can be found in Section 4.7.3. In the single-echelon setting with an (R, Q) - policy, the optimization problem is then one-dimensional. In that case, a brute-force algorithm, as described in Section 4.7.3, is feasible.

In the multi-echelon setting, not only does the number of variables in the optimization problem increase with each node, the computations to find the service level and costs for a set of reorder points are more complex. Thus, a brute-force approach is not a practical option, this is the case for cost-minimization both with and without service constraints. Research in optimization theory provides many different techniques for minimizing functions that could be applied to the multi-echelon setting, e.g. genetic algorithms, as proposed by Çelebi (2015) or using the Lagrangian as in van Donselaar et al. (2021). However, many papers in multi-echelon theory often focus on providing more computationally tractable approaches to model the system. One such approach is the decomposition techniques as discussed in Section 4.9.5 above.

With a decomposition technique one might iteratively search through the solution space in order to find the optimal solution (as in Axsäter 2003). Berling and Marklund (2006) suggests a full decomposition method where first, induced backorder costs are used to find the optimal reorder point, R_0^* for the warehouse, and second, the optimal reorder points for the lower echelon are found. Thus, their approach ends up being not more computationally complicated than N+1 (N retailers and one central warehouse) single node-by-node optimizations. Their method is elaborated upon in Section 4.10 below.

4.10 Berling-Marklund multi-echelon inventory control model

In two recent papers by Berling and Marklund (2013; 2014), a decomposition-based approach is used to model a 2-echelon divergent OWMR inventory system. The two papers use the same modeling techniques but investigates different demand distributions. Berling and Marklund (2013) focuses on Compound Poisson customer demand (sometimes referred to as lumpy customer demand) whereas Berling and Marklund (2014) uses Normal customer demand. The model is based on the decomposition approach introduced in Andersson et al. (1998), where a penalty cost is used at the central warehouse to find its optimal inventory parameters. The penalty cost is intended to account for the extra holding- and backorder costs at the retailers caused by shortages at the central warehouse. Andersson et al. (1998) defines the penalty cost as the expected marginal cost per time unit at a retailer with respect to changes in the lead time, see (45). Andersson et al. (1998) then iteratively calculates retailer costs, lead time estimates, and penalty costs until convergence. Berling and Marklund (2006) presents a closed form estimate for the optimal penalty cost dependent on system parameters which the authors refer to as an induced backorder cost. Thus, the need for iteration is no longer there as the optimal induced backorder cost can be computed directly. The latter approach to find the induced backorder cost is what Berling and Marklund (2013; 2014) use in their models.

$$\beta_{i,Andersson} = \frac{dC_i(L_i(R_0))}{dL_i(R_0)} * \frac{1}{\mu_i} \tag{45}$$

Furthermore, the models assume all locations to use a continuous review, (R, Q), -ordering system, complete backordering at all locations, and that demand is served according to a First-Come-First-Served (FCFS) principle. This section will describe the model in detail, starting with a conceptual description followed by estimation of induced backorder costs, modeling of lead time demand at the central warehouse, estimates of retailer lead times, and the optimization of reorder points at the different locations. From here on, the model used in Berling and Marklund (2013; 2014) will be referred to as the BM-model.

4.10.1 Conceptual description of the BM-model

The cost minimization problem of the multi-echelon problem with one central warehouse (RDC) and N retailers (dealers) is described in (46). The index 0 refers to the central warehouse, and the indices $1, 2, ..., N$ refers to the N retailers in the system. TC is the total costs of the system, h_i refers to the holding cost per unit and time unit at installation i and R_i refer to the reorder point at installation i. $E[IL_i^+]$ is the expected stock-on-hand at installation i. Lastly, γ_i denotes the fill-rate of installation i and TFR_i the target fill rate at installation i. The BM-model decomposes this problem to N+1 single-echelon problems.

$$\begin{aligned} &\min_{(R_i \forall i)} TC = \min_{(R_i \forall i)} \left(h_0 E[IL_0^+(R_0)] + \sum_{i=1}^{N} h_i E[IL_i^+(R_0, R_i)] \right) \\ &s.t.\ \gamma_i(R_i, L_i) > TFR_i \quad i = 1, 2, ..., N \end{aligned} \tag{46}$$

The single-echelon problem used to compute optimal parameters at the central warehouse is an unconstrained optimization problem as described in Section 4.6.1. The holding cost per unit and time unit, h_0, is assumed known and the induced backorder cost at the central warehouse, per unit and time unit, β_{CW}, which acts the role of a penalty cost applied to the central warehouse, is computed as described in Section 4.10.2 below. The decomposition results in the optimization problem in (47). Recall that $E[IL_i^-]$ denotes the expected backorders at the installation. The cost function is convex, consequently, the solution to the one-dimensional optimization problem is easily found. (Berling and Marklund, 2013; 2014)

$$\min_{R_0} \tilde{C}_0(R_0) = \min_{R_0} \left(h_0 E[IL_0^+(R_0)] + \beta_{CW} E[IL_0^-(R_0)] \right) \tag{47}$$

The N single-echelon retailer problems regarding the retailers are either service-constrained optimization problems as described in Section 4.6.2 or minimization expected of holding and backorder costs as described in Section 4.6.1. The service measure used by Berling and Marklund (2013; 2014) is the item fill rate, as defined in Section 4.4. The dynamics between the echelons are here captured by the estimate of the stochastic lead time, $\hat{L}_i$, which depend on the delay at the central warehouse due to stockouts and thus depend on the choice of inventory parameters at the central warehouse. For the case of service constrained optimization, the N optimization problems are solved according to (48). The cost functions are now increasing with their respective reorder points, R_i, hence the solutions to these N, one-dimensional, optimization problems are also easily computed. (Berling and Marklund, 2013; 2014)

$$\begin{aligned} &\min_{R_i} C_i(R_i, L_i) = \min_{R_i} \left(h_i E[IL_i^+(R_0, R_i)] \right) \\ &s.t.\ \gamma_i(R_i, L_i) > TFR_i \quad i = 1, 2, ..., N \end{aligned} \tag{48}$$

To summarize, with the decomposition, as described above and illustrated in Figure 14, the original multi-echelon optimization problem in (46) is simplified to the problems presented in (47) and (48).

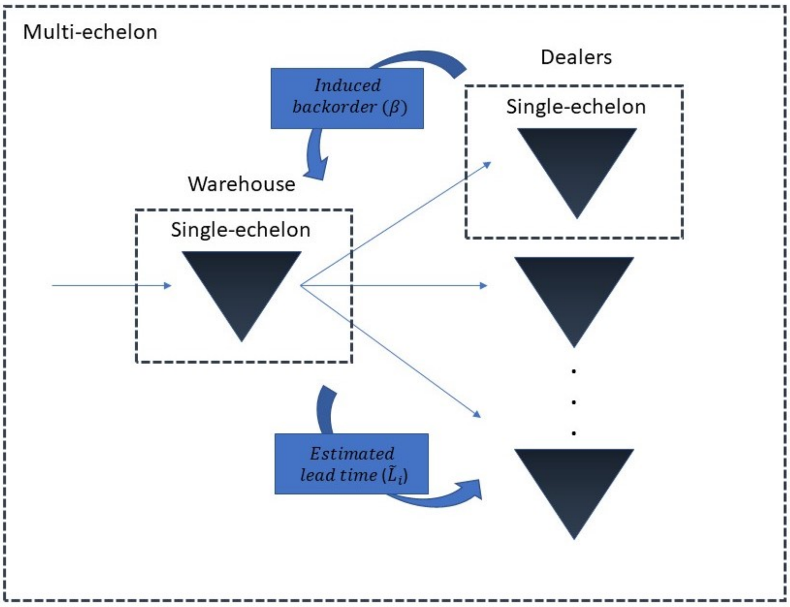

Figure 14: Decomposition procedure of BM-model.

4.10.2 Induced backorder cost

To increase the efficiency of the decomposition model described in Andersson et al. (1998), Berling and Marklund (2006) provide a closed form estimate of the optimal penalty cost, see (49), which they call induced backorder cost, denoted β_i. The calculations require the shortage costs at the retailers, p_i, however, when these are unknown, they have to be approximated. For cases when the system uses fill rate constraints instead of shortage costs to optimize the inventory parameters, the conversion formula presented in (50) is a viable approximation. (Berling and Marklund, 2014)

$$\beta_i = h_i * g(Q_{i,n}, p_{i,n}) * \sigma_{i,n}^{k(Q_{i,n}, p_{i,n})} \tag{49}$$

$$p_i = \frac{h_i * TFR_i}{(1 - TFR_i)} \tag{50}$$

Berling and Marklund (2006) show that the induced backorder cost, β_i, is mainly affected by the order quantity, Q_i, shortage cost, p_i, and standard deviation of customer demand, σ_i, their estimate is produced through linear regression with these attributes. The calculations for the induced backorder costs are based on a "normalized" system where a unit of demand equals 100, holding cost per unit and time unit equals 1 and transport time from warehouse to retailer equals 1. Any system's parameters can be scaled to this normalized model setting by using the conversion formulas in Table 3. Furthermore, the authors present two methods of computing the values for $g(Q_{i,n}, p_{i,n})$ and $k(Q_{i,n}, p_{i,n})$. They use a tabular approach where they provide tables for a set of combinations of inputs and then use interpolation between points in the table to find the estimates. They also provide closed form estimates, presented in (51) and (52) below.

$$\begin{aligned} g(Q_{i,n}, p_{i,n}) &= \min[g_a \cdot (Q_{i,n})^{g_b}, G], \\ g_a &= \min\left[0.015p_{i,n}, \max(\frac{0.65}{\sqrt{p_{i,n}}}, 0.05)\right], \\ g_b &= \max[-1.2, -2p_{i,n}^{-0.25}], \\ G &= \min[0.015, 0.005p_{i,n}^{0}.2] \end{aligned} \tag{51}$$

$$\begin{aligned} k(Q_{i,n}, p_{i,n}) &= \max[1, \min(k_a \cdot (Q_{i,n})^{k_b}, K], \\ k_a &= \max[0.7, \min(0.9, 0.6p_{i,n}^{0.075})], \\ k_b &= \min(0.2, 0.4p_{i,n}^{-0.35}], \\ K &= \max(1.3, \min(2, 2.5p_{i,n}^{-0.15}) \end{aligned} \tag{52}$$

Table 3: System parameter normalization conversions.

	Original system parameters	**Normalized system parameters**
Retailer order quantity	Q_i	$Q_{i,n} = 100Q_i/(\mu_i l_i)$
Central warehouse order quantity	Q_0	$Q_{0,n} = Q_0$
Retailer holding cost per unit and time unit	h_i	$h_i = 1$
Central warehouse holding cost per unit and time unit	h_0	$h_{0,n} = h_0/h_i$
Retailer shortage cost	p_i	$p_{i,n} = p_i/h_i$
Central warehouse lead time	L_0	$L_{0,n} = L_0/l_i$
Retailer transport time	l_i	$l_{i,n} = 1$
Expected demand per time unit at retailer i	μ_i	$\mu_{i,n} = 100$
Standard deviance of demand per time unit at retailer i	σ_i	$\sigma_{i,n} = 100\sigma_i/(\mu_i\sqrt{l_i})$
Induced backorder cost	$\beta_i = \beta_{i,n}h_i$	$\beta_{i,n}$

Lastly, the induced backorder costs in the case of non-identical retailers may differ between retailers. Thus, the induced backorder cost faced by the central warehouse is computed by a weighting scheme. Berling and Marklund (2006) examines several schemes but cannot find significant differences and thus in their latter work, Berling and Marklund (2013; 2014) use the weighting scheme as in (53) based on proportion of the total expected customer demand, due to its simplicity.

$$\beta_{CW} = \frac{\sum_i^N \mu_i\beta_i}{\sum_i^N \mu_i} \tag{53}$$

4.10.3 Lead time demand at central warehouse

The lead time demand at the central warehouse, $D_0(L_0)$, is in Berling and Marklund (2013; 2014) expressed in "subbatch" demand rather than a unit demand, where a subbatch, Q_s, is defined as the largest common divisor among the retailer order quantities, Q_i. The reason is that the warehouse can only face integer multiples of this quantity and for all other values of lead time unit demand, d, the probability is zero, i.e. $P(D_0(L_0) = d \mid d \neq nQ_s) = 0, \;\; n = 0, 1, 2, \ldots$. Considering only subbatch demand renders more computationally efficient formulas whilst also improving the accuracy of the demand approximations that will be described below. Furthermore, an assumption of the BM-model is that the lead time for an order to arrive at the warehouse from an outside supplier, L_0, is constant.

The derivation of the lead time demand is then as follows. First, $\delta_i(n)$, the probability of a retailer, i, ordering at most n times during the central warehouse lead time, L_0, is derived. The δ_i - function is found in (54) and (55) for retailers facing Compound Poisson customer demand (Berling and Marklund, 2013) and Normal customer demand (Berling and Marklund, 2014) respectively. Here, $x = IP_i - R_i$, and $D_i(L_0)$ refers to the customer unit demand at retailer i during the central warehouse lead time, L_0. Recall that φ and Φ denotes the density- and distribution function of the standard Normal distribution, $\mathcal{N}(0, 1)$. Also remember that Poisson- and Negative Binomial customer demand can be considered special cases of Compound Poisson demand, hence, (54) is valid in these cases of customer demand as well.

$$\delta_i(n) = \frac{1}{Q_i} \sum_{x=i}^{Q_i} P\left(D_i(L_0) \leq nQ_i + x - 1\right) \quad \text{for } n = 0, 1, 2, \ldots \tag{54}$$

$$\begin{aligned} \delta_i(n) = \int_0^{Q_i} \frac{1}{Q_i} \Phi\left(\frac{nQ_i + x - \mu_i L_0}{\sigma_i \sqrt{L_0}}\right) dx = \frac{1}{Q_i} \Bigg\{ Q_i \Phi\left[\frac{(n+1)Q_i - \mu_i L_0}{\sigma_i \sqrt{L_0}}\right] + \\ \sigma_i \sqrt{L_0} \left[\varphi\left(\frac{(n+1)Q_i - \mu_i L_0}{\sigma_i \sqrt{L_0}}\right) - \varphi\left(\frac{nQ_i - \mu_i L_0}{\sigma_i \sqrt{L_0}}\right)\right] + \\ (nQ_i - \mu_i L_0) \left[\Phi\left(\frac{(n+1)Q_i - \mu_i L_0}{\sigma_i \sqrt{L_0}}\right) - \Phi\left(\frac{nQ_i - \mu_i L_0}{\sigma_i \sqrt{L_0}}\right)\right] \Bigg\} \end{aligned} \tag{55}$$

With $D_0^i(L_0)$ as the subbatch demand faced by the central warehouse during the lead time from retailer i with $f_0^i(u)$ as its probability mass function. Now, $f_0^i(u)$ is found as in (56) where q_i is the ordering quantity of retailer i expressed in units of subbatches. (Berling and Marklund, 2013; 2014)

$$f_0^i(u) = P\left(D_0^i(L_0) = u\right) = \begin{cases} \delta_i(0) & \text{if } u = 0 \\ \delta_i(n) - \delta_i(n-1) & \text{if } u = nq_i,\ n = 1, 2, ... \\ 0 & \text{otherwise} \end{cases} \quad (56)$$

The exact lead time demand distribution at the central warehouse from all retailers can now be found by convoluting the probability mass functions of subbatch demand for all retailers over all possible values of u. However, this method is computationally costly. Berling and Marklund (2013; 2014) therefore approximate the exact lead time demand distribution with a known parametric distribution. First, the correct central warehouse lead time demand mean and variance are computed with (57) and (58), respectively, using the results from (56). Then finally, a parametric distribution is chosen.

$$\mu_0 = \mu_0^{(1)} + \mu_0^{(2)} + ... + \mu_0^{(N)} \text{ where } \mu_0^{(i)} = \frac{\mu_i L_0}{Q_s} \quad (57)$$

$$\sigma_0^2 = (\sigma_0^{(1)})^2 + (\sigma_0^{(2)})^2 + ... + (\sigma_0^{(N)})^2 \text{ where } (\sigma_0^{(i)})^2 = \sum_{n=0}^{\infty} (\mu_0^{(i)} - nq_i)^2 f_0^i(nq_i) \quad (58)$$

Berling and Marklund (2013; 2014) suggests the following rules of thumb for choosing which distribution to use for the modeling of central warehouse demand:

- If $\sigma_0^2/\mu_0 \geq 1$, use Negative Binomial distribution.
- If $\sigma_0^2/\mu_0 < 1$ and $\sigma_0/\mu_0 < 0.25$, use a discrete approximation of a Normal distribution.
- Else, use a discrete approximation of a Gamma distribution.

The discrete approximations of the continuous distributions suggested are produced according to (59), here, $F(x)$ is the cumulative probability function of the respective distributions.

$$P(D_0(L_0) = u) = f_0(u) \cong \begin{cases} F(0.5) & \text{for } u = 0 \\ F(u + 0.5) - F(u - 0.5) & \text{for } u = 1, 2, ... \end{cases} \quad (59)$$

The motivation for using these approximations is to maintain computational efficiency for all possible mean to variance ratios (Berling and Marklund, 2013; 2014). As described earlier, the Negative Binomial distribution is a Compound Poisson process with Logarithmic compounding distribution and may thus be most suitable to describe the discrete demand. A mean to variance ratio above one is however a requirement for this distribution to exist. For mean to variance ratios below 1, the Normal and Gamma distribution are used. The rule of thumb is constructed as a trade-off between: (i) the Normal distribution allows for demands less than 0, and (ii) the Gamma distribution is more computationally complex than the Normal distribution (see Section 4.5.5 - Section 4.5.6 for a comparison between the two distributions).

4.10.4 Optimal reorder points at the central warehouse

With the induced backorder cost, β_{CW}, computed, and the central warehouse lead time demand, $D_0(L_0)$, distribution approximated, all information required to solve the optimization problem in (47) is found. The expected stock-on-hand per time unit, $E[IL_0^+]$ and the expected backorders, $E[IL_0^-]$ are computed as in (61) and (62). Note that these are expressed in units of subbatches while cost values h_0 and β_{CW} are expressed per unit and time unit, why scaling with the subbatch, Q_s, is required. (Berling and Marklund, 2013; 2014)

$$\min_{R_0} \tilde{C}_0(R_0) = \min_{R_0} \left(Q_s h_0 E[IL_0^+(R_0)] + Q_s \beta^*_{CW} E[IL_0^-(R_0)] \right) \tag{60}$$

where

$$E[IL_0^+(R_0)] = \frac{1}{Q_0} \sum_{y=R_0+1}^{R_0+Q_0} E_{D_0(L_0)}[(y - D_0(L_0))^+] = \frac{1}{Q_0} \sum_{y=R_0+1}^{R_0+Q_0} \sum_{u=0}^{y} (y-u) f_0(u) \tag{61}$$

and

$$E[IL_0^-(R_0)] = \frac{1}{Q_0} \sum_{y=R_0+1}^{R_0+Q_0} E_{D_0(L_0)}[(y - D_0(L_0))^-] = \frac{1}{Q_0} \sum_{y=R_0+1}^{R_0+Q_0} \sum_{u=y}^{\infty} (u-y) f_0(u) \tag{62}$$

As the cost-function in (60) is convex in R_0, the optimum is easily found with a simple line search, an optimality condition is presented in (63). (Berling and Marklund, 2013; 2014)

$$R_0^* = \max[R_0 : \tilde{C}_0(R_0) - \tilde{C}_0(R_0 - 1) \leq 0] \tag{63}$$

4.10.5 Retailer lead time estimates

The BM-model uses the common METRIC-type approach of modeling the retailer lead times as a sum of an assumed constant transport time, l_i, and its stochastic delay caused by stockouts at the central warehouse, $W(R_0)$, i.e. $L_i = l_i + W(R_0)$. As described in Section 4.9, the literature provides plenty of methods for estimating this lead time expression. This thesis will focus on the methods explored in the papers connected to the BM-model. For the case of identical retailers, as in Andersson et al. (1998), the expected value of the lead time at an individual retailer is equal to the overall expected value of the lead time between central warehouse and retailers. Thus, the expected lead time at a retailer is found with the use of Little's law, see (64).

$$\tilde{L}_i = l_i + \frac{L_0}{\mu_0 Q_s} E[IL^-(R_0^*)] \tag{64}$$

For the case of non-identical retailers, the problem is more complicated. While Little's law provides the expected unit delay at the central warehouse, it is not certain that the units are uniformly distributed over the retailers. I.e. it is not certain that every individual retailer face the same expected delay at the central warehouse. This is due to the fact that the waiting time faced by an individual retailer is not only dependent on the chosen reorder point at the central warehouse (R_0), but also other parameters like order intensity and order quantity.

Berling and Marklund (2013) examines two estimates of the lead time for the case of non-identical retailers. The simpler of them was to approximate the mean lead time between the central warehouse and each individual retailer with the use of Little's law, according to (64). This approach was deemed preferable when applying the model with retailer parameter optimization under service level constraints.

4.10.6 Optimal reorder points at retailers

With lead time estimates established, the reorder point optimization at the retailers, i.e. the search for optimal $R_i \ \forall i$ in (48), is conducted as explained in Section

4.7.3. However, when the retailers face demand approximated by a Normal distribution, Berling and Marklund (2014) uses an adjustment to the item fill rate calculations.

A problem with the Normal distribution approximation for demand when computing the item fill rate is that it assumes that demand is continuous, which suggests that the inventory position never goes below the reorder point as the new order is placed as soon as it is reached. However, in reality, demand for most products is not continuous, it comes in integer numbers. For instance, consider the situation where a customer places an order of 2 items while the current inventory position is at 1 above the reorder point, i.e. $IP = R + 1$. Then the inventory position after the purchase will fall below the reorder point. In this example, the new order to the central warehouse is placed at $IP = R - 1$ and not at $IP = R$. Berling and Marklund (2014) calls the actual inventory position when the order is placed the *realized order point* and this behavior will result in the fill rate computed under the assumption of Normal demand (see (35) in Section 4.7.2) to overestimate the actual fill rate. They suggest two different adjustments, of which one is regarded in this thesis and presented below in (65) and (66) (Berling and Marklund, 2014; 2017).

$$\gamma_i = \sum_{u=0}^{\hat{u}} SERV_i(R_i - u)U_i(u) \tag{65}$$

where

$$U_i(u) = \left(\frac{1}{Q_i}\sum_{k=u+1}^{u+Q_i} O_i(k)\right) \Bigg/ \left(\sum_{j=0}^{\hat{u}} \frac{1}{Q_i}\sum_{k=j+1}^{j+Q_i} O_i(k)\right) \tag{66}$$

Here, $\hat{u}$ denotes the maximum undershoot, $SERV_i(r)$ is the fill rate of a realized reorder point, r, $U_i(u)$ represents the probability of an undershoot of size u, and $O_i(k)$ is the probability of a customer demanding an order of size k.

To conclude, using the steps presented in Section 4.10.2 - Section 4.10.6 will render near-optimal reorder points for the multi-echelon system. The use of a decomposition technique with induced backorder costs at the warehouse, as in the BM-model, results in computational efficiency, minimizing the number of iterations needed (e.g. compared to the model presented in Andersson et al. (1998)). In addition, the decomposition technique used in the BM-model allows the user with a single-echelon optimization model already in place, to reuse part of its logic and algorithms, making it easier to understand for the practitioner and, potentially, more efficient to implement.

Chapter 5

Numerical study

This section contains a detailed exposition of the execution of the numerical study. The objectives of the study is accounted for in Section 5.1. The origin of the data used is reported in Section 5.2. The structure of the analytical models are described in Section 5.3. Lastly, the composition of the simulation model is outlined in Section 5.4.

5.1 Objective and overview

A numerical study was performed with two objectives: (i) to investigate and understand what systematic changes to expect if implementing a multi-echelon approach (the BM-model) for optimizing the reorder point in the Volvo inventory control process and, (ii) to assess the potential of this method to achieve target service levels, decrease total inventory in the system as well as holding- and backorder costs. This numerical study was vital in answering the research questions of the thesis. The study encompass 52 items in a system consisting of one RDC and 15 dealers. It can be divided in three steps:

- Data was collected and analyzed to provide the inputs required.
- Based on the inputs, two sets of reorder points, as well as estimates regarding service level, stock levels and costs, were found analytically. One set was produced using a multi-echelon (ME) approach (the BM-model), the other using a single-echelon (SE) model made to resemble the current system.
- With the inputs and the two sets of reorder points, discrete-event-simulations were performed aiming to resemble a real-world setting and assess the accuracy of estimations produced by the analytical models.

In Figure 15 an overview of the numerical study, as well as inputs and outputs in each step is visualized. In Section 5.2, Section 5.3 and Section 5.4, each of the three steps in the numerical study is described in detail.

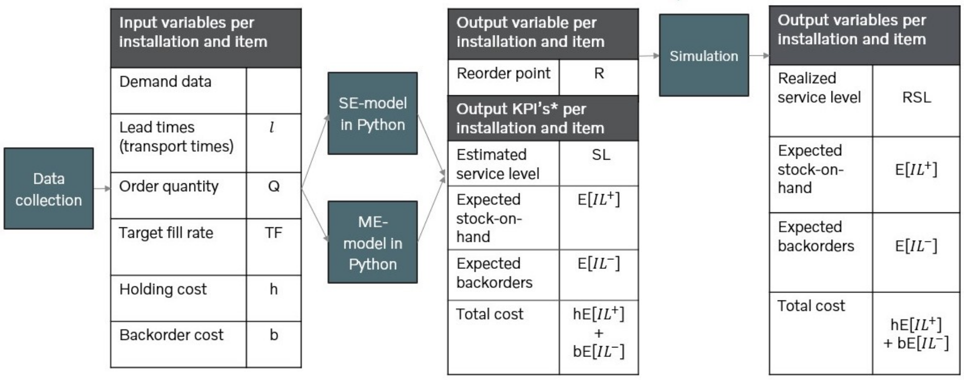

Figure 15: Overview of the numerical study.

5.2 Data collection

This section describes the data collection process. First, the network scope is exhibited, then the item scope is discussed, followed by an account of the required inputs and their origin. Lastly, the demand data and distributions used in the numerical study is presented.

5.2.1 Network scope

The inventory system under scrutiny in this numerical study is the 2-echelon OWMR divergent South African Volvo distribution network with one RDC in Johannesburg supplying 15 dealers within the country with spare parts of the VCE brand. The central warehouse in Gent supplying the RDC in Johannesburg is considered an "outside supplier" in this numerical study. This network was chosen in agreement with the case company, the major reasons were data accessibility and that in this network Volvo has VMI agreements in place with the dealers. This enables Volvo to centrally control inventory decision-making, which is a precondition for using a central inventory control policy as investigated in this thesis.

5.2.2 Item scope

Currently in the South African distribution system, two different methods of single-echelon control models are in place. One is characterized by a service level at the RDC close to 100 %, and follows the general case at Volvo when using the MMI system as analyzed and described in Section 2. This method is from here on referred to as the "regular policy single-echelon model", or "regular policy SE-model" for short. The second policy is to our knowledge specific to the South African network and was implemented ad-hoc in order to keep more stock at dealers and let the RDC only act as a cross-docking facility for SKUs on their way to the dealer. In order to put this policy into effect while still using the MMI system the target fill rate is set to 10 % at the RDC while also manually setting the safety stock at dealers to an amount equal to 22 – 30 days of demand forecast (Volvo, 2022c). This second method is from here on referred to as the "special policy single-echelon model" or "special policy SE-model".

While the special policy is specific to the South African network, approximately 20 % of the total number of items are currently controlled by the special policy (Volvo, 2022c). Therefore, it was decided to both examine items that are currently controlled by the regular policy and the special policy. In total, 28 items controlled by the regular policy and 24 items administered with the special policy were investigated.

While all the dealers have the possibility to order any item from the item catalog at the RDC, not all items face demand at every dealer. In those cases where no sales have been conducted of an item at a dealer, there is no registration of that item in the MMI inventory system. For this numerical study, it is therefore assumed that demand for these items is non-existent at these dealers. Furthermore, no item was found that was sold at every dealer. This means that the items analyzed are present at somewhere between 2 - 13 dealers in the network.

A dimension of segmentation at Volvo that is of interest in the item selection process is that the MMI system classifies the items according to one of eight demand types for each installation: *Fast, Erratic, Lumpy, Slow, New, Obsolete, Insufficient and Non-moving*. For this numerical study, only items with demand types *Fast, Erratic, Lumpy and Slow* are considered as for the others the data required for the system to make effective forecasts is insufficient. Analyzing such items was deemed outside of the scope of this thesis as results could be more heavily dependent on forecasting issues rather than model performance. Including the four different demand groups as described enhances generality of the study as the different demand patterns faced by the network is represented. Note that for all demand types, lead time demand is in this numerical study modeled according to a Compound Poisson process with empirical order sizes, more on that in Section 5.2.3 below. The four demand types in the study are characterized by intermittency (regularity of picks) and variation of demand (high or low). Examples of the four demand patterns is illustrated in Figure 16. In Table 4 a summary of the item scope is found.

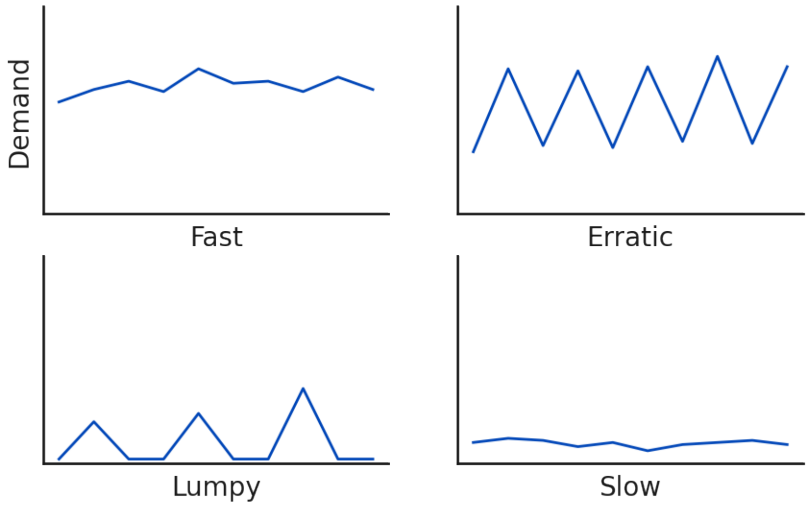

Figure 16: Examples of demand patterns for the four demand type segments.

Table 4: Summary of item scope.

Policy	**No of items**	**Average no of dealers per item**	**No of dealers facing demand type**			
			Fast	Erratic	Lumpy	Slow
Special	24	11.29	104	33	16	142
Regular	28	7.1	92	36	23	76

5.2.3 Model inputs and demand distribution fitting

In Table 5 the input data required for modeling is listed. The data points were extracted from the MMI inventory control system on the 19th of April 2022.

Table 5: Summary of model inputs.

Model input	Data source
Transport times, l_i	"Lead time" data field in MMI
Order quantities, Q_i	"Constr. opt. OQ" data field in MMI
Target fill rates, TFR_i	"Target srv. lvl." data field in MMI
Holding costs, h_i	"Unit cost" data field in MMI multiplied with capital cost rate for South Africa.
Demand mean during a time unit, μ_i	"EOD" data field in MMI, divided by 30 days.
Demand variance during a time unit, σ_i	"Std. dev." data field in MMI, divided by $\sqrt{30}$ days.
Order size sample	Historical order sizes collected from the two year period from the 19th of april 2020 to the 19th of april 2022.

The first four entries in Table 5 are directly entered into the analytical model and simulation model. Remaining is to use the mean demand during a time unit, demand variance and order size sample to produce demand distributions. Both the single-echelon and multi-echelon models in the study assume lead time demand to be drawn from stationary stochastic processes (see Section 4.5 for a more thorough explanation). It was therefore decided to perform this numerical study by fitting stationary stochastic processes to the real demand data collected, more precisely, Compound Poisson processes.

First, empirical order size distributions were constructed by using the relative frequency of order sizes in the historical order size data. As discussed in Section 4.5.4, this kind of distribution can capture order size distributions with large variability and irregularities where common parametric distributions make poor fits. Thus, it was deemed that these distributions were the most realistic description of the demand data and would thus provide the best basis to compare the two models.

The second step is to use (9), (10) and (11) (see Section 4.5 for reference) with the μ and σ forecasts collected from MMI, as described in Table 5, in order to compute the distribution of lead time demand for every combination of item and dealer in the item scope. With lead time demand distributions fitted, all information required to both analytically model the system and to perform discrete-event simulations is available.

5.3 Analytical modeling of the Volvo network

For the purpose of this thesis to compare a single- and multi-echelon inventory model, only the main inventory flows in the distribution network are modeled. Consequently, emergency shipments and reverse inventory flows are not taken into consideration. Currently, there is no explicit decision rule at Volvo regarding when to place an emergency order. Hence, modeling these order flows would first require the creation of an appropriate such decision rule. In addition, the volume of emergency shipments and back flows is correlated with an item's service level which has its target set prior to optimizing inventory parameters. For these reasons emergency shipments and reverse inventory flows falls outside the focus of this thesis.

In dialogue with Volvo the reorder points at the dealers were restricted to be equal to or above -1 in the SE-model. This was done due to the fact that currently this restriction is in place at Volvo, thus it was decided to have this restriction in both the multi-echelon and the single-echelon models. However, the reorder point at the RDC was in both cases unrestricted.

Furthermore, many items in the inventory system had the option of being "stocked" or "non-stocked". Table 6 shows the division of stocked versus non-stocked item locations as well as its relation to demand type. It was noticed that for the non-stocked items, the MMI-system enforced a target fill rate of 0 % and had set Q equal to one unit. By (33) a reorder point of $R = -1$ will give items where $Q = 1$ a fill rate of 0 %. Thus, a reorder point of $R = -1$ can in these cases be seen as equivalent to an item being "non-stocked". Consequently, the program has the option of indirectly setting an item as "non-stocked" while not explicitly having implemented such a function into the python program. This choice was made mainly due to time constraints in the project. It was later discovered that not all items set as "non-stocked" had Q-values of one, why this simplification did not fully hold. However, for the analysis of comparing the different models, this was not an issue.

Table 6: Summary of the dealers in the item scope.

Policy		**No of dealers facing demand type**			
		Fast	Erratic	Lumpy	Slow
Special	Stocked	104	33	9	58
	Non-Stocked	0	0	7	84
Regular	Stocked	91	35	18	32
	Non-Stocked	1	1	5	44
Total		196	69	39	218

5.3.1 Multi-echelon modeling with the BM-model

The multi-echelon modeling is made in accordance with the BM-model as described earlier, in this section it will be described exactly what equations and assumptions were used in this numerical study. For a detailed description of the model used, see Section 4.10. For the multi-echelon modeling, no distinction in terms of inventory settings was made between the two groups of items in the "regular" and "special" categories. For each of the five items, the following five steps were performed:

1. The optimal induced backorder cost at the RDC, β^*_{CW} is computed according to (49) and (53), see Section 4.10.2, with normalized values of Q_i, σ_i and p_i. These parameters were normalized according to the conversion formulas in Table 3. The shortage costs, p_i estimates were produced from the target fill rates by using (50).

2. The lead time demand at the central warehouse was modeled according to the approximations as described in Section 4.10.3. The mean demand during lead time, μ_0 and variance of demand during the lead time σ_0^2 was computed according to (57) and (58), respectively. The required components for these equations were found according to (56) and (54). Here the demand distribution of the dealers during the RDC lead time is required and as dealers were assumed to face Compound Poisson distributed demand, this was found according to what was described in Section 5.2.3 above.

3. (60), (61), (62) and (63) in Section 4.10.4 were used to find the optimal reorder point at the RDC.

4. The optimal retailer reorder points at the dealers were found by solving the

optimization problem as stated in (48) in Section 4.10.1. The retailer lead times, L_i, used were estimated as their transport time plus the mean stochastic delay at the warehouse according to (64), see Section 4.10.5. Estimates of fill rates were computed according to (33) with stock-on-hand, IL^+, distribution computed according to (34), see Section 4.7.1.

5. An estimate of average backorders in the system was computed with the use of the average stock-on-hand according to (7) in Section 4.3.2. An estimate of the total holding- and backorder costs of the system was produced by simply multiplying the average stock-on-hand and average backorders with the costs, respectively. The backorder costs used for these calculations correspond to the target fill rates at the specific installations and were obtained according to (50), see Section 4.10.2. Note that only backorders at the lowest echelon were considered as these installations are the ones interacting with end customers. Eventually, the costs were converted to an annual basis to enhance their interpretability.

Noteworthy is that when running the program, step 2 was the one mainly inhibiting scalability. As the lead time between central warehouse (located in Belgium) and the RDC in Johannesburg is 58 days, demand during lead time for certain items could become quite large. The calculations of the Compound Poisson distribution requires the convolution of all combinations of a certain number of customers buying a certain amount of items, see Section 4.5.1. For larger demand, this becomes time-consuming. As a result, outputs of certain items were computed almost instantaneously, while others required more time. However, running the program, computing outputs for all the 52 items, took less than five minutes on a regular laptop computer.

5.3.2 Single-echelon modeling with the regular policy SE-model

To model the regular policy, reorder points for both the dealers and the RDC were found by solving the service-constrained optimization problem according to what is presented in (30) in Section 4.6.2.

The reorder point optimization at the dealers is done as described in step four of the multi-echelon BM-modeling described in Section 5.3.1 above. The only difference being that lead time estimates are now the constant transport time as the SE-model does not make use of information from other nodes in the network. For reference the equations used are (33) and (34) with Compound Poisson distributed demand as described in Section 5.2.3. Note that also in the SE-models, for both regular and special policy, the Compound Poisson process with empirical order sizes is used to model lead time demand. This demand model is not currently in use at Volvo but it was used for this study in order to keep inputs to the multi-echelon and single-echelon models equal. As a result, the calculations of fill rate of the SE-models in this study is done more theoretically correct in comparison to what is used in the MMI-system.

For the RDC, a Normal demand distribution was assumed, as this was the case in the current system. The target service levels used was also obtained from the system and ranged between 90-99 %. The optimal reorder point was then found by solving (30), where the service level was computed by (35) and inventory level distribution as in (36), see Section 4.7.

Lastly, estimations of backorders and costs were computed in accordance with what was described in Section 5.3.1 regarding the BM-model.

5.3.3 Single-echelon modeling with the special policy SE-model

For the special policy the computations of reorder points at the RDC were analogous to the regular case as described above (see Section 5.3.2 above), however, with a target service level equal to 10 %. The reorder points for the dealers, however, were computed differently.

(43) in Section 4.7.3 displays the relationship between reorder point, safety stock and average lead time demand. With the lead time estimated as the transport time, l_i, which is known, and the mean demand during a time unit μ_i, the average lead time demand, μ' is easily found as $\mu' = \mu_i * l_i$. Moreover, by converting the safety stock from days to amount as $SS_{days} * \mu_i = SS$, all information to compute the reorder points according to (43) is available. With the reorder points computed, fill rate estimates, stock-on-hand estimates, backorder estimates and cost estimates are computed analogously to the regular case as presented in Section 5.3.2 above.

5.4 Evaluation by discrete event simulation

This section will describe the discrete event simulation model used to evaluate the accuracy of estimates of fill rate and costs provided by the two models. First, the purpose of the simulation will be discussed together with an explanation of how its outputs should be interpreted. Then a detailed explanation of the structure and function of the simulation model will be accounted for.

5.4.1 Purpose of simulation and interpretation of simulation output

The simulation model is constructed to mimic the Volvo distribution network in order to provide relevant information for the comparison of the currently used single-echelon models and the newly suggested multi-echelon model. The simulation model receives the same input as the analytical model, that is transport times, order quantities, holding- and backorder costs, demand processes. Moreover, the simulation model also receives as input the two sets of reorder points computed in the multi-echelon BM-model and the two single-echelon models (regular and special policy), respectively. Note that the simulation model does not perform any optimization of inventory decision variables, i.e. the reorder points, but only provides estimates of service levels, average stock-on-hand, average backorders and total costs for the system when the reorder points in the different analytical models are used.

The purpose of the simulation is two-fold, first, it is to analyze how well the different analytical models perform in terms of describing the system dynamics, i.e. providing accurate estimates of system performance indicators. As all of the models contain approximations and assumptions the simulation provides a reference for evaluating the accuracy of estimates of system performance indicators under these approximations and assumptions. Second, the discrete-event simulation provides a setting where performance indicators, such as fill rates, stock levels, and cost savings can be compared between the single-echelon and multi-echelon models without depending on accuracy of the estimations made in the different analytical models.

Furthermore, when interpreting the results it is important to understand that the simulation model will receive the same input values as the models, and this numerical study can not provide any measure of accuracy or performance of these inputs. While this may seem like a limitation, it is also a necessity. The simulation is supposed to describe the real world as accurately as possible, thus it requires the most accurate inputs present. The same argument holds for the analytical models, why it would not make sense to use different inputs in the analytical setting and the discrete-event simulation. Furthermore, using this setting also results in that the differences in accuracy of estimations can be directly tied to the difference in the models, which thus provide a solid basis for evaluation and comparison of the models.

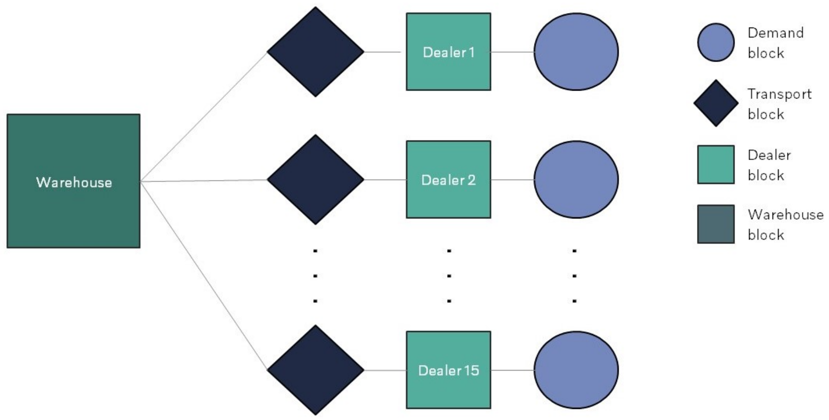

Figure 17: Conceptual setup of the simulation model.

5.4.2 Discrete event simulation model setup

Figure 17 shows the conceptual setup of the simulation model. It was built according to the OWMR structure with up to 15 dealers (the model allowed for readjustment of the amount of dealers between 1-15). The actual model is built in the software ExtendSim which uses a variety of blocks to represent different processes in the system. For example, there is one block to generate demand, one to represent a dealer, and one to represent the warehouse. Noteworthy, is that as opposed to the model in Python, the simulation can only handle order sizes up to 50 units. Thus, only items with maximum order size equal to or less than 50 was included in this study. During the data collection only one item displaying order sizes above 50 was found, and consequently discarded, so this restriction only had minor impact on the study.

In the model, demand is generated by two stochastic variables drawn from two different distributions. One depicts interarrival times between customers which is drawn from an exponential distribution and the other represents the demand size of a customer which is drawn from the order size distributions generated as explained in Section 5.2.3 above. Customer demands are then generated according to the drawn interarrival times and order sizes. This setup is using exactly the same Compound Poisson demand with order sizes based on historical frequencies as described earlier for the analytical models.

The demand reaches the dealer-block which keeps track of its internal inventory and when the reorder point is reached, an order of, Q_i, units is sent to the RDC. The RDC-block works similarly to the dealer-block and delivers orders when they are demanded by the dealer-block as well as sending orders to the outside supplier (representing the warehouse in Gent). Both the RDC-block and the dealer-block act according to a FCFS-policy, partial deliveries and full backordering. Which means that in case of a stockout, when the next shipment is received, all reserved backorders will be serviced first. Technically this is done by allowing the stock level to become negative, which can be interpreted as if there are customers waiting for orders.

During the simulation data is collected regarding stock-on-hand, backorders and fill rate at every installation. At the end of the run the data is collected and average stock-on-hand, backorders, fill rates as well as total costs are calculated.

Each item is simulated two times, one time with the reorder points computed by the BM-model, and one time with reorder points computed with its respective current control method, either regular- or special policy SE-model. All of the data collected was averaged over 30 runs of 25200 time units each. Every run was given a new pseudo-random seed to use in the sampling for the demand generation processes, this is done to reduce the risk of bias in the demand generation. At the start of each run all installations started with stock-on-hand equal to $R_i + Q_i$, and before each run there was also a warm-up period of 3000 time units in order to allow the system to reach stationarity before starting data collection.

Chapter 6

Results & Analysis

In this part of the thesis, the results from the numerical study is presented. The differences in proposed decision rules between the different models are analyzed in Section 6.2. The performance in terms of customer service, stock levels, backorders, and holding- and backorder costs are evaluated in Section 6.3 - Section 6.6. Lastly, the results are summarized in Section 6.7

6.1 Introduction to results

The analysis of the results will focus on the comparison of the decision rules (i.e. the reorder points) proposed by the multi-echelon BM model compared to those proposed by the single-echelon (SE) models currently in use at Volvo (see Section 5.3 for a detailed exposition on the determination of these). The comparison will focus on performance indicators of attained customer service levels, average stock levels, average backorders, and holding- and backorder costs. The performance was evaluated with the use of the described discrete event simulation model (see Section 5.4 for further explanation).

The results are presented both per item and per demand category from an aggregated system viewpoint. Moreover, the items are split into two sets according to the two different policies (regular and special, see Section 5.2.2). The results are presented separately for these two item groups as the output when changing from the old decision rules, according to the single-echelon models' suggestions, to the ones proposed by the BM-model differs depending on which SE-model (regular- or special policy) is in use.

6.2 Reorder points

In this section the proposed decision parameter at the different installations, i.e. the reorder points, are studied. This will give an understanding of what systematic changes in decision-making can be expected when introducing the multi-echelon BM-model in the Volvo distribution process.

The reorder points at the RDC are displayed in Figure 18 and Figure 19, while the reorder points at the dealers can be studied in Figure 20 and Figure 21.

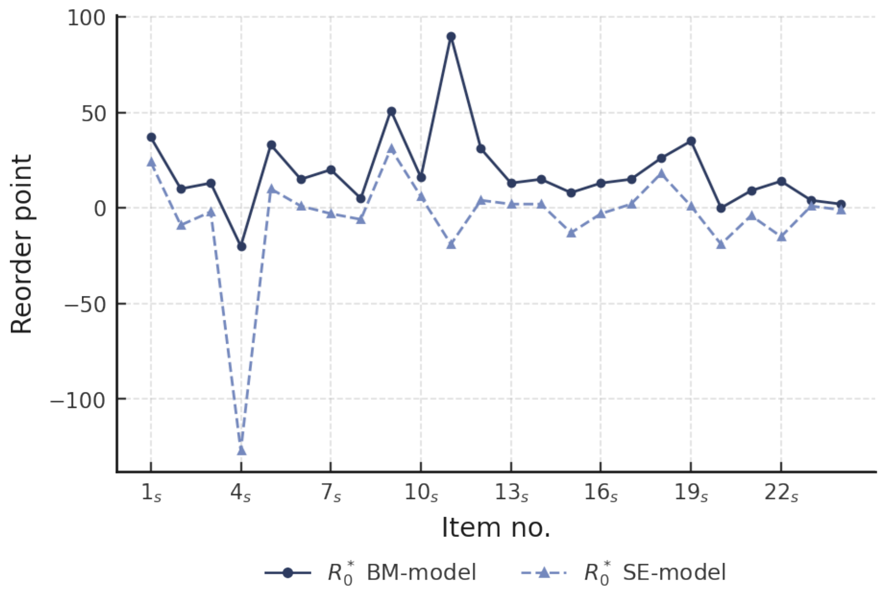

Figure 18: Optimal reorder point at RDC, R_0^*, per item where the special policy is used in the SE-model.

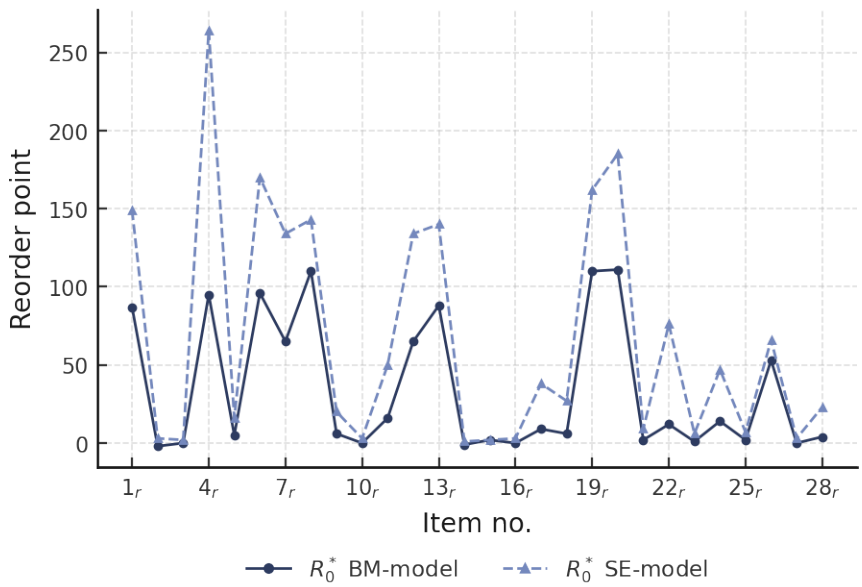

Figure 19: Optimal reorder point at RDC, R_0^*, per item where the regular policy is used in the SE-model.

Figure 18 and Figure 19 show the optimal reorder point for each item at the RDC proposed by the BM-model and the special- and regular policy SE-models, respectively. Regarding the items following the special policy, the optimal reorder points at the RDC increase when controlling the system as proposed by the BM-model. The opposite pattern is seen when comparing the regular policy SE-model to the multi-echelon BM-model, here, the reorder points decrease. This behaviour is not unexpected, the two policies represent two extremes as the special policy aims for a fill rate of 10 % at the RDC while the regular policy uses a target fill rate close to 100 % in all cases. In Section 6.3 below, it is shown that using the reorder points determined by the BM-model results in an RDC fill rate above 10 % for special policy, and, below 100 % for the regular policy case. Recall that the fill rate at an installation increases with the reorder point (see Section 4.6.1).

Figure 20 and Figure 21 displays the mean reorder point at the dealers proposed by the different models. As can be seen in the figures, the reorder points at the dealers as proposed by the BM-model generally increase in comparison to the SE-models under either policy. Only four items 14_s, 15_s, 19_s and 21_s, the result shows slightly reduced reorder points when the BM-model is used.

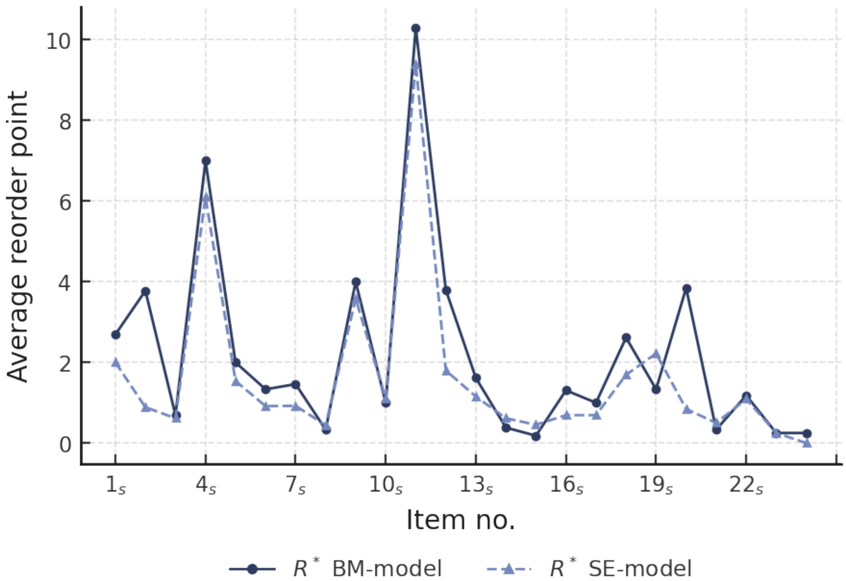

Figure 20: Average optimal reorder point at the dealers, $\bar{R}^*$, per item where the special policy is used in the SE-model.

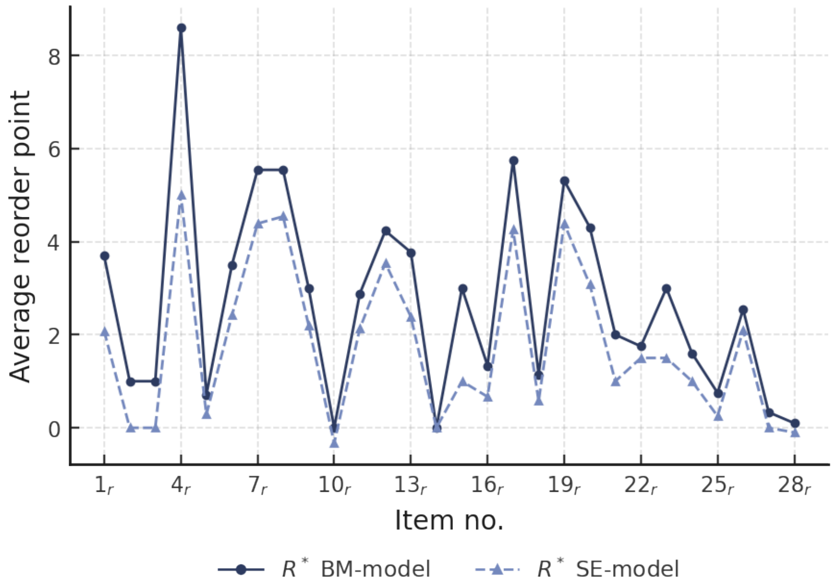

Figure 21: Average reorder point at the dealers, $\bar{R}^*$, per item where the regular policy is used in the SE-model.

Studying the changed reorder points when shifting to the BM-model from the regular policy SE-model, (see Figure 19 and Figure 21) the decrease in reorder point at the RDC in combination with the increase in reorder point at the dealers indicates that the BM-model pushes stock closer to the end customer in the distribution chain. This effect will be explored further in Section 6.4 below.

Figure 22 and Figure 23 exhibits the mean reorder point per demand classification. The changes in optimal reorder point when comparing the BM-model to the SE-models are similar regardless of which model policy is used. Items whose demand was classified as *Erratic* according to the MMI system show the largest increase in reorder point. This suggests that for items facing *Erratic* demand the BM-model suggests a more severe shift in stock from RDC to Dealers.

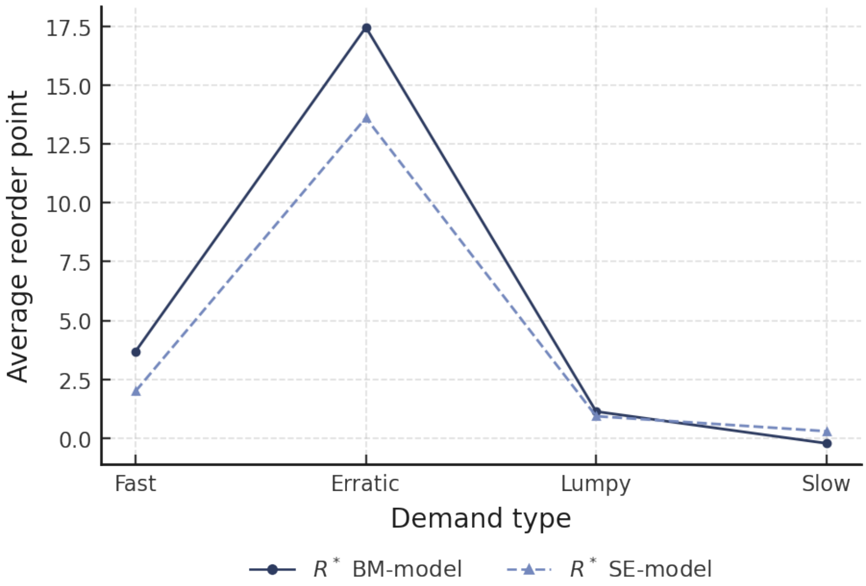

Figure 22: Average optimal reorder point at the dealers, $\bar{R}^*$, by demand type classification where the special policy is used in the SE-model.

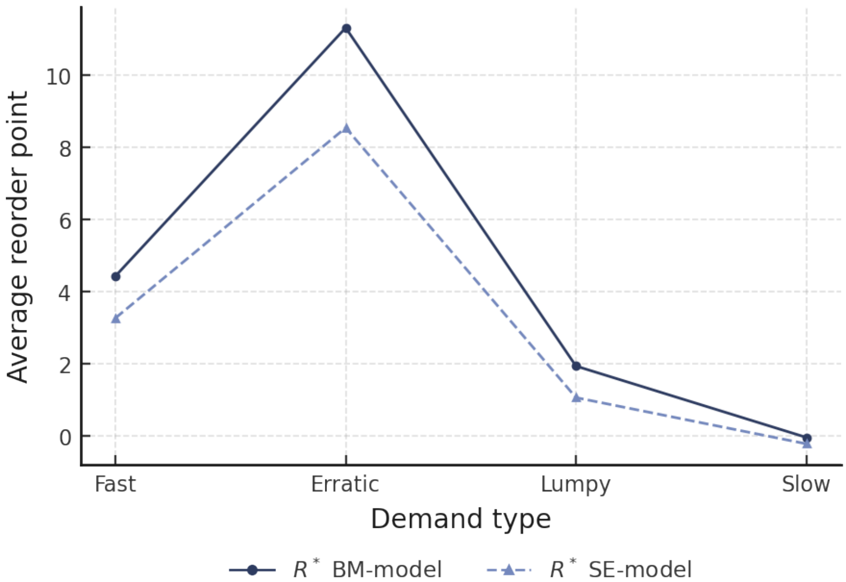

Figure 23: Average optimal reorder point at the dealers, $\bar{R}^*$, by demand type classification where the regular policy is used in the SE-model.

6.3 Fill rates

In this section, the fill rates estimated by the different models as well as the fill rates simulated with the different reorder points suggested by the models are presented. The analysis here is two-fold, first, it is interesting to compare the fill rates between the SE-models and BM-model to assess if one has an advantage over the other in terms of reaching targets. Second, it is interesting to compare model estimates to those produced by simulation, as this gives a measure of the accuracy of the mathematical approximations.

First, the fill rates at the RDC is studied in Figure 24 and Figure 25, then the fill rates at dealers are examined in Figure 26, Figure 27 and Figure 28. Finally, the dealer fill rates are studied by demand type in Figure 29 and Figure 30.

The fill rate estimates at the RDC produced by the different models are displayed in Figure 24 and Figure 25. Recall that while both the SE-models, using the two different policies, used a target fill rate at the RDC when finding optimal reorder points, the BM-model only considers service requirements at the dealers. Consequently, the fill rate at the RDC estimated by the BM-model should not be evaluated in terms of reaching the target at the RDC.

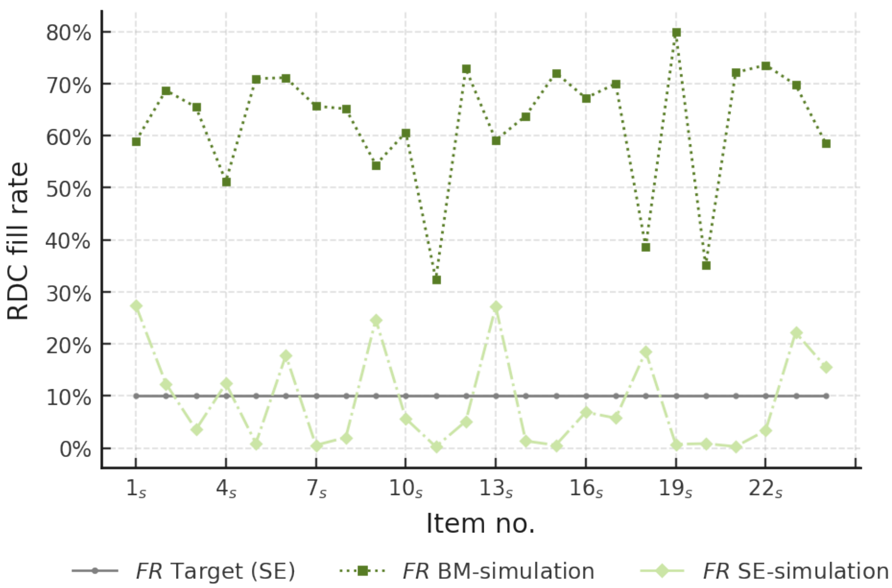

Figure 24: Fill rate estimates at RDC per item where the special policy is used in the SE-model.

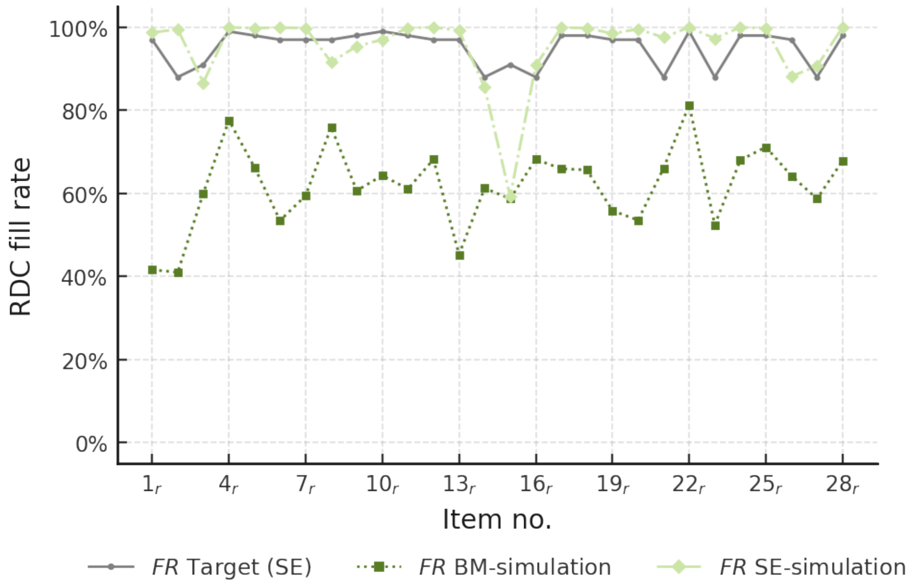

Figure 25: Fill rate estimates at RDC per item where the regular policy is used in the SE-model.

The simulation shows that the reorder points proposed by the BM-model produces a fill rate at the RDC in the range of 35 – 80 %. While this is quite a large range, the fill rate for most items is above 60 % and on average, the simulated fill rate for the BM-model reorder points is ∼ 68 %. In conclusion, the BM-model allows for quite a large share of backorders at the RDC, but avoids any extreme fill rates i.e. close to 0 or 100 %. The two SE-models however, have fill rates closer to these extreme values and the patterns are similar to what was seen in Section 6.2 regarding reorder points above.

As seen in Figure 24, the special policy SE-model both over- and undershoots the target fill rate. Noteworthy is that in 8 cases (items 5_S, 7_S, 11_S, 14_S, 15_S, 19_S, 20_S and 21_S) the fill rate drops to almost 0 %. In these cases, indicating a severe lack of stock in the system. Regarding the regular policy SE-model, as shown in Figure 25, the simulated fill rates reaches the targets in most cases. While this is desirable it should be noted that most of the items overshoot this target. With targets close to 100 %, the stock requirements are already high, which means that this could be a costly overshoot. In other words, this indicates high stock levels for the regular policy SE-model. The stock levels are further examined in Section 6.4 below.

Figure 26 and Figure 27 display the mean deviation from target fill rates for items under the special policy. For this policy, the dealers are separated into stocked and non-stocked. The reason is that the mean deviations tend to vary significantly between the two sets which is apparent in the figures.

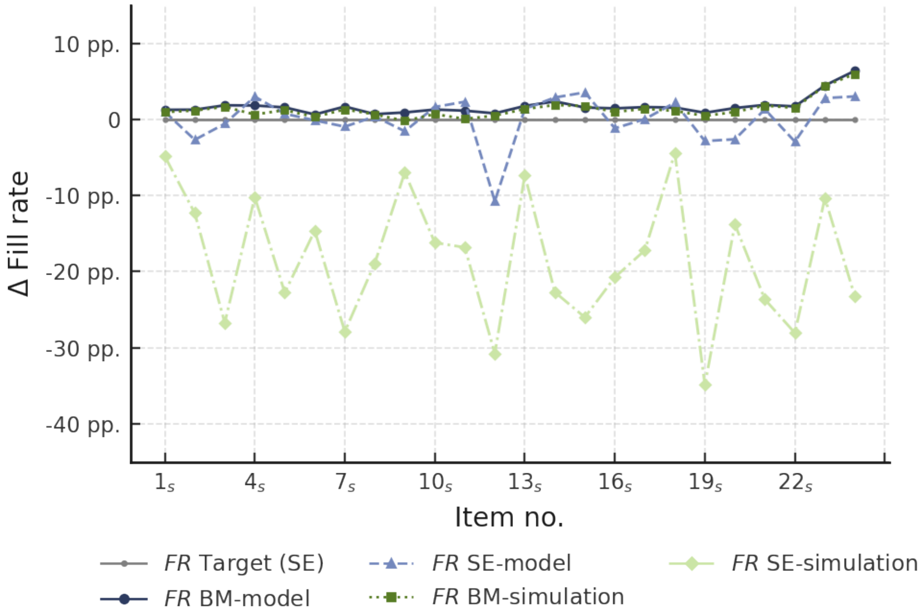

Figure 26: Mean deviation from target fill rate at dealers with the item stocked and the special policy used in the SE-model.

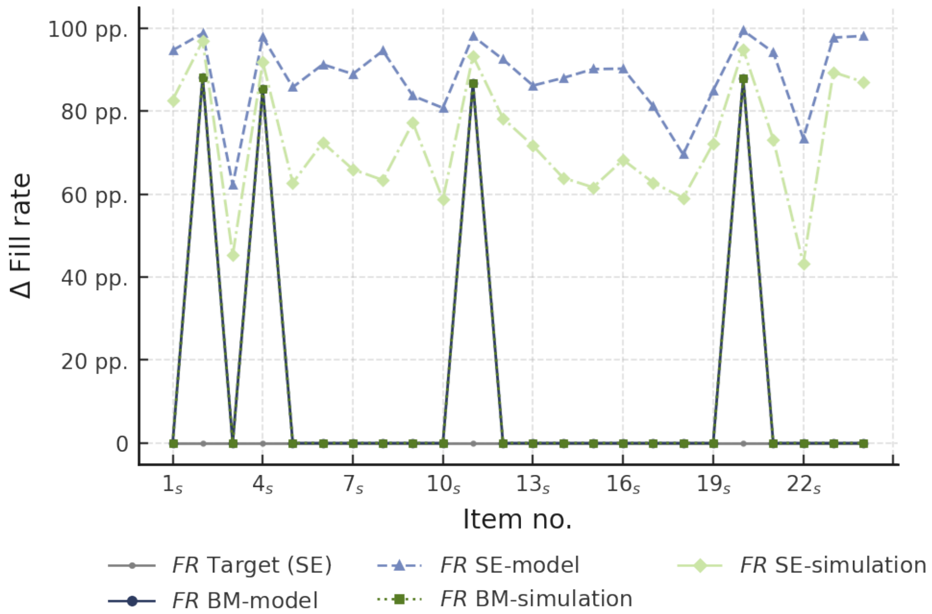

Figure 27: Mean deviation from target fill rate at dealers with the item not stocked and the special policy used in the SE-model.

In Figure 26, it is shown that the special policy SE-model overestimates the fill rate for dealers where the items are stocked. The model estimates the fill rate to reach the target in almost all cases, but, the simulation shows that all items are in fact catered for a much lower percentage of orders. The simulated fill rates in some cases were as low as an average of 35 percentage points below target. This is a consequence of the model not accounting for the significant waiting time affecting the supply to dealers. Sequentially, this is a result of the very low fill rate of $\sim 10\ \%$ at RDC (as shown earlier in Figure 24) when, in fact, the model assumes $100\ \%$ fill rate at the RDC. The simulation also shows that the BM-model is more reliable both in terms of providing accurate estimations of the fill rate, as well as producing reorder points capable of reaching target fill rates. Figure 26 also shows a slight overshoot of the target fill rate for all items when using the BM-model. However, this is no surprise, recall from Section 4.7.3 that the optimal reorder point, R^*, is the smallest reorder point to produce a fill rate equal to or above the target. Thus, a slight overshoot is to be expected since the reorder points are integer valued.

In Figure 27 the fill rate estimates and simulated values are shown for the BM-model and special policy SE-model for the non-stocked dealers. Note that in the figure, the fill rates produced by the BM-model and the BM-simulation are directly on top of each other. Similarly to the stocked dealers, the special policy SE-model overestimates the fill rates compared to the realised fill rates from the simulation. Again, the BM-model provides more accurate estimates in comparison to the simulation. However, more interesting is the severe overshoot of target fill rate, almost up to 100 percentage points over target, seen by the SE-model as well as the four spikes in the BM-model estimates.

Recall that all "non-stocked dealers" had a target fill rate equal to 0 % (see Section 5.3). Theoretically, a target of 0 % is equivalent to a reorder point of $R = -Q$, as discussed in Section 4.7.1 earlier. However, the special policy SE-model cannot produce a reorder point less than 0, as the formula in (43) can only generate reorder points $\geqslant 0$ (see Section 5.3.3 for more details). As a result, all the dealers were assigned the lowest possible reorder point for the special policy, that is, $R = 0$. Moreover, as the demand for the items at the non-stocked dealer is very low, the reorder point of $R = 0$ generates fill rates far above 0 % for all items. The BM-model however, is only constrained to reorder points of $R \geq -1$. Thus, for item-dealer combinations where $Q = 1$, the BM-model can satisfactory reach appropriate reorder points for the intended target fill rates. The spikes in deviation, however, are the result of order quantities being greater than one, as an example, $Q = 10$ for item 2_s resulting in fill rates above 0 %.

Figure 28 shows the mean deviation of fill rates for items currently administered by the regular policy. Here, the results are not split in stocked and non-stocked as both the BM- and regular policy SE-model are restricted to $R \geq -1$, and there was no systematic difference in the deviation when inspecting stocked versus non-stocked dealers.

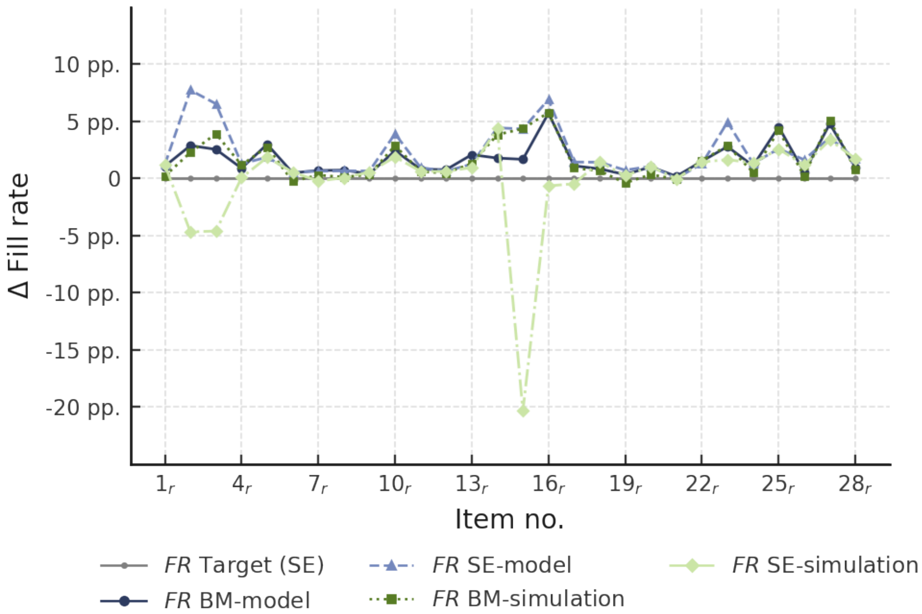

Figure 28: Mean deviation from target fill rate at dealers where the regular policy is used in the SE-model.

In Figure 28, it can be seen that the realized fill rates generally are slightly overshooting the targets. This happens both when using reorder points proposed by the regular policy SE-model as well as when using those proposed by the BM-model.

A significantly deviating data point is item 15_r's negative deviation from target for the SE-simulation. The large deviation from the target fill rate might be the consequence of the low demand for this item. Low demand will result in low reorder points (in this case the two dealers handling this item was given reorder points equal to 2 and 0 by the regular policy SE-model). With such low demand, an increase or decrease by one unit of the reorder point may have large effects for the fill rate. The BM-model however, seem more suited to handle this scenario as it reaches targets for all items investigated.

Figure 26, Figure 27, and Figure 28 show that while the BM-model provides reorder points that reach targets in all cases, the reliability of the SE-models depend on whether the regular or special policy is used. A likely explanation for these observations is found by scrutinizing the assumption of a constant lead time in the SE-models.

The constant lead time used is an estimated transport time between RDC and dealer. Under the regular policy the RDC had a fill rate close to 100 % (as shown in Figure 25) and can almost always deliver as soon as an order is received. Thus, using the transport time between RDC and dealer as lead time estimate is probably accurate, and in turn renders good fill rate estimations. In the case of special policy however, the RDC has fill rates between 0 - 30 % (see Figure 24). Thus, it can only dispatch an order immediately at most three times out of ten, and sometimes never. Consequently, the lead time faced by dealers is somewhere in between the transport time between RDC and dealer (5-10 days for the examined items and dealers) and this transport time plus a waiting time up to the full lead time between central warehouse and RDC (58 days for the examined RDC). In conclusion, the lead time estimate used is probably quite poor for these items which in turn renders worse fill rate estimations by the special policy SE-model. As explained in the Section 4.10 the BM-model takes this waiting time at the warehouse into account when estimating fill rates and determining reorder points.

Figure 29 and Figure 30 display the deviations from target fill rate by demand type for items controlled by the special and regular policy, respectively. Note that the deviation is larger for the items controlled by the special policy.

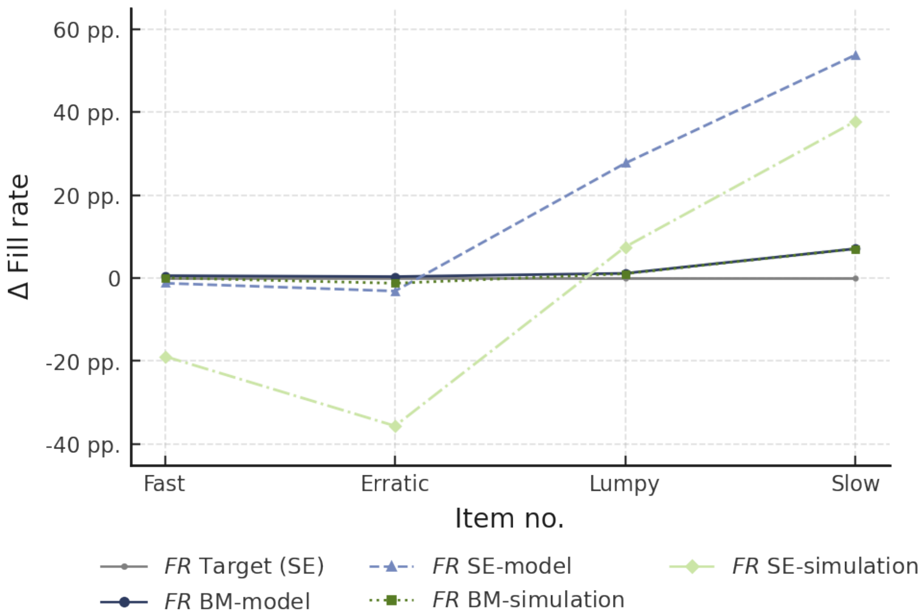

Figure 29: Fill rate estimates at dealers by demand type classification where the special policy is used in the SE-model.

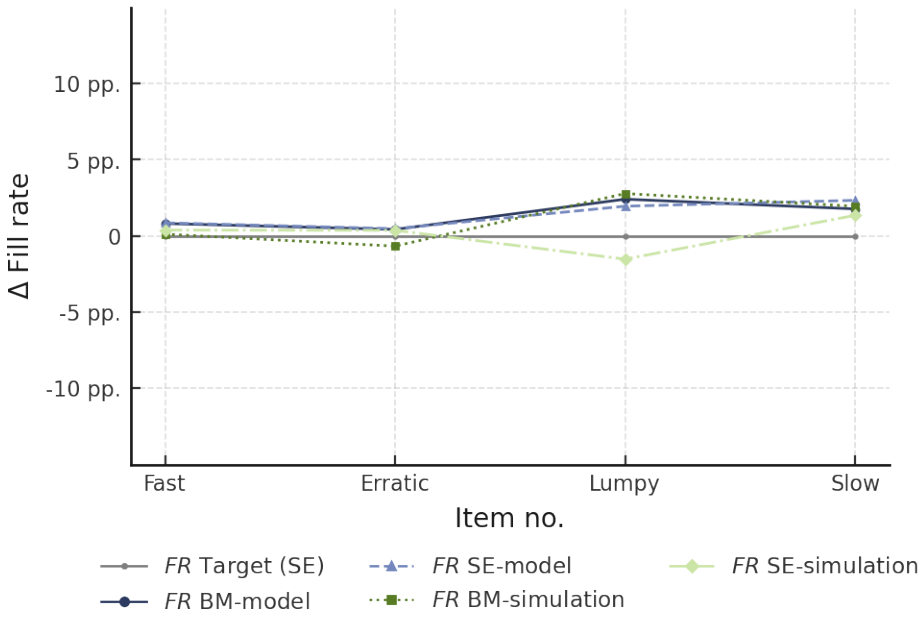

Figure 30: Fill rate estimates at dealers by demand type classification where the regular policy is used in the SE-model.

In Figure 29, the reorder points proposed by the BM-model produce fill rates close to target for all demand types. Moreover, the simulation shows that the special policy SE-model provides reorder points which results in fill rates under target for items facing demand classified *Fast* and *Erratic*. However, for low demand items classified as *Lumpy* and *Slow*, the fill rates reach far above targets. This pattern is likely explained by the fact that non-stocked items are often categorized with demand type *Slow* or *Lumpy* while items with *Fast* and *Erratic* demand types are all stocked. The pattern thus follows the pattern for stocked and non-stocked items as discussed above and shown in Figure 26 and Figure 27.

For items currently administered under the regular policy, the fill rates per demand type are displayed in Figure 30. Here, the deviations are smaller (note the different scales on the y-axes in Figure 29 and Figure 30). In conclusion, there does not seem to be any significant differences in terms of reaching target fill rates between different demand types.

6.4 Stock-on-hand

This section investigates the differences in stock-on-hand. Here, the differences between expected stock under the BM-model compared to the two single-echelon models in place are examined. Furthermore, the estimates accuracy in comparison to the simulation is evaluated.

In Figure 31 and Figure 33 the total system stock per item is displayed. Figure 32 and Figure 34 exhibits the increase in stock-on-hand per item based on simulated values. Lastly, Figure 35 and Figure 36 show the average stock-on-hand at the RDC and the dealers for each item.

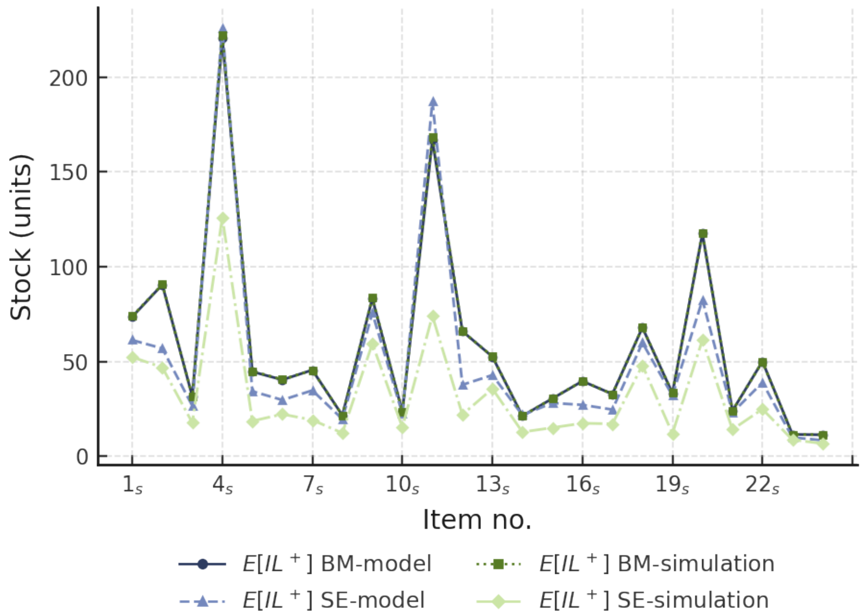

Figure 31: Average system stock-on-hand, $E[IL^{+}]$, per item where the special policy is used in the SE-model.

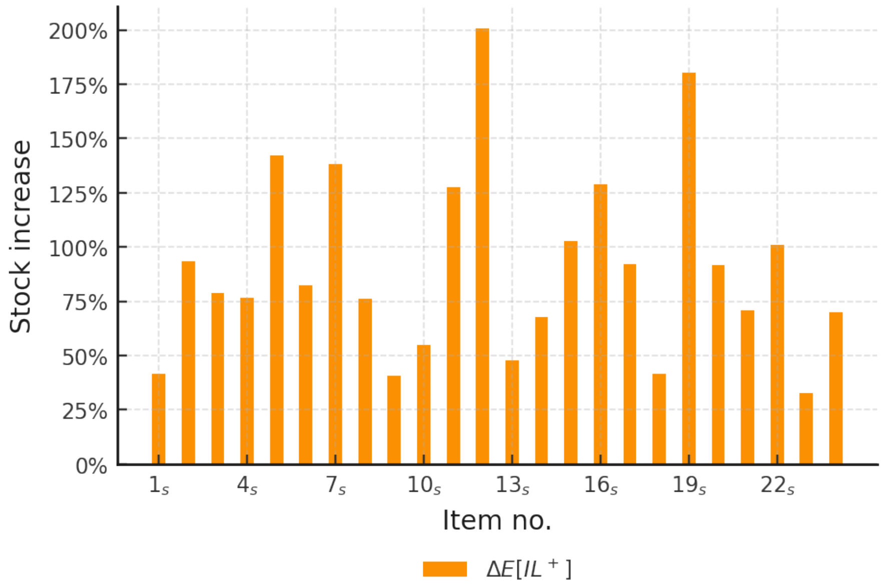

Figure 32: Simulated increase in average system stock-on-hand, $E[IL^+]$, per item using the BM-model compared to the special policy SE-model.

In Figure 31 it is shown that the total stock for most items in the system is increasing slightly under the BM-model in comparison to using the reorder points produced by the special policy SE-model. Noteworthy is also that the simulation shows that the SE-model heavily overestimates the stock levels at these location. These results cohere with the finding that fill rates were generally underestimated for stocked items as seen earlier in Figure 26. Consequently, there is an increase in stock on hand for all items when using the BM-model which more accurately estimates the fill rate, as can be seen in Figure 32.

Regarding the regular policy, the numbers are reversed, the stock-on-hand generally decrease when using the BM-model, as is displayed in Figure 34.

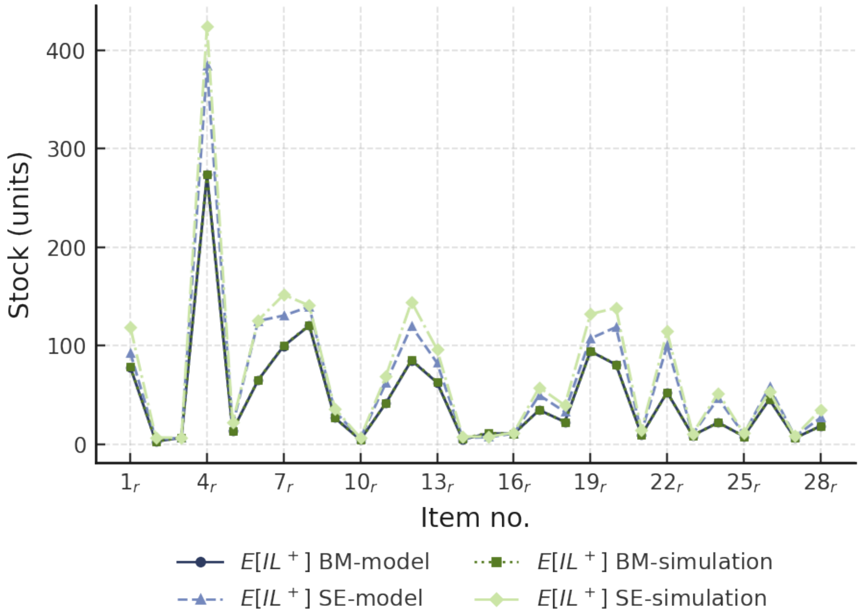

Figure 33: Average system stock-on-hand for the system, $E[IL^+]$, per item where the regular policy is used in the SE-model.

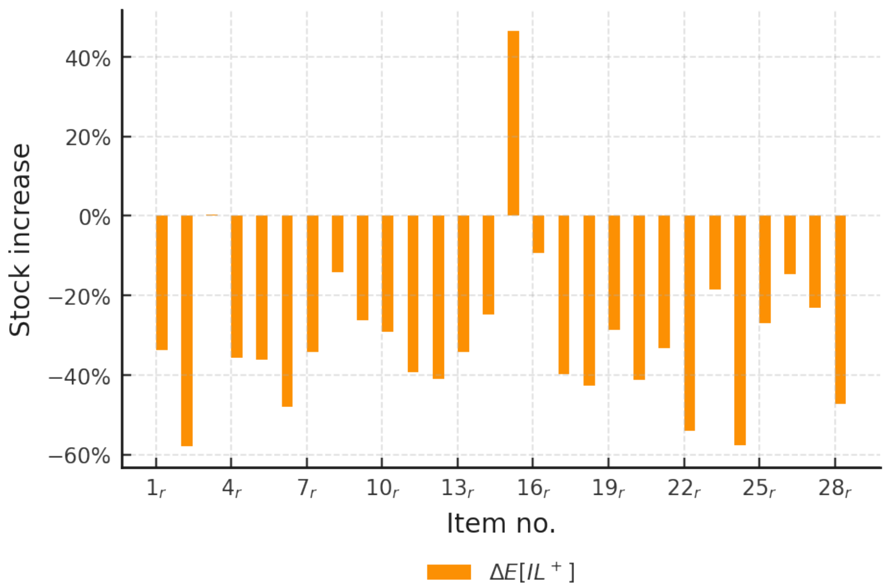

Figure 34: Simulated increase in average system stock-on-hand, $E[IL^+]$, per item using the BM-model compared to the regular policy SE-model.

Regarding the items currently using the regular policy SE-model, only item 15_r experiences an increased stock under the BM-model. The results thus indicate a general decrease in stock-on-hand when using the BM-model in comparison to the regular policy SE-model. The increase in stock for item 15_r can be explained by the fact that this item was the one where the regular policy SE-model had significant issues in reaching target service levels. In Figure 28 it is shown that this item experienced an average deviation of 20 *pp.* below target. The reorder points suggested by the BM-model however, reached target fill rates. With this in mind, the increased stock for this item is not unreasonable. Furthermore, the $\sim$ 40 % increase in expected stock-on-hand (see Figure 33) is in absolute terms only a couple of units, 3.64 to be exact .

Noteworthy is that for all items, the BM-model provides accurate estimations of expected stock-on-hand, which can be seen in Figure 31 and Figure 33. The SE-model only provides accurate estimates under the regular policy, as was also the case when studying fill rate estimations.

Figure 36 and Figure 35 exhibit the total system average stock-on-hand divided on RDC and dealers for items governed by the special and regular policy, respectively.

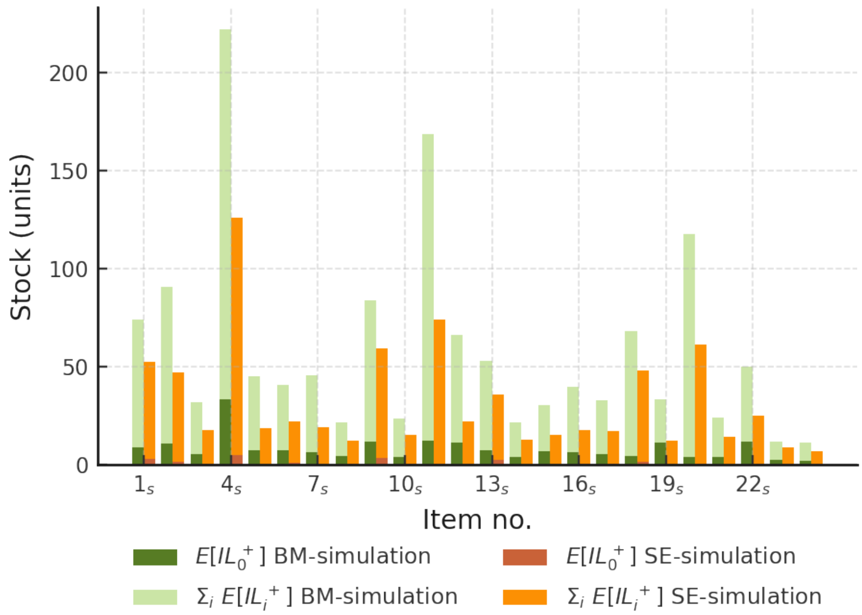

Figure 35: Average system stock-on-hand, $E[IL^+]$, allocated to RDC and dealers where the special policy is used in the SE-model.

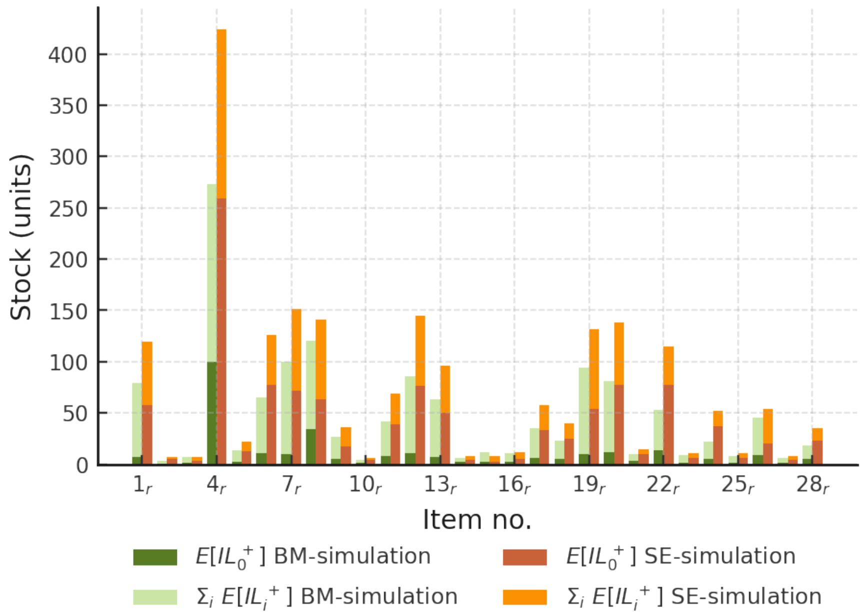

Figure 36: Average system stock-on-hand, $E[IL^+]$, allocated to RDC and dealers where the regular policy is used in the SE-model.

Starting with studying the allocation of stock in the supply chain for the regular policy in Figure 36, it is clear that when comparing the multi-echelon BM-model to the more traditional single-echelon approach there is a prominent pattern. The BM-model suggests that a larger portion of the stock is located at dealers while less is to be held at the RDC. The regular policy SE-model suggests a lot more stock to be held at the RDC, which is not surprising as the fill rate requirements at the RDC are close to one hundred percent.

Figure 35 shows the stock allocation for the items regulated by the special policy. In comparison to the special policy SE-model, the BM-model suggests more stock to be held at the RDC. It is however remarkable that the increase in stock held at the RDC is only slight. This suggests that the goal of the special policy, namely to keep the larger share of stock at the dealers rather than the RDC, is quite well fulfilled by the BM-model as well.

6.5 Backorders

In this section the average amount of backorders in the system will be studied. As it is only backorders at dealers that incurs a penalty cost, it is interesting to investigate the backorders at dealers and RDC separately.

Figure 37 and Figure 38 display the average backorders for items currently controlled by the special policy while Figure 39 and Figure 40 exhibit results for the items presently controlled by the regular policy.

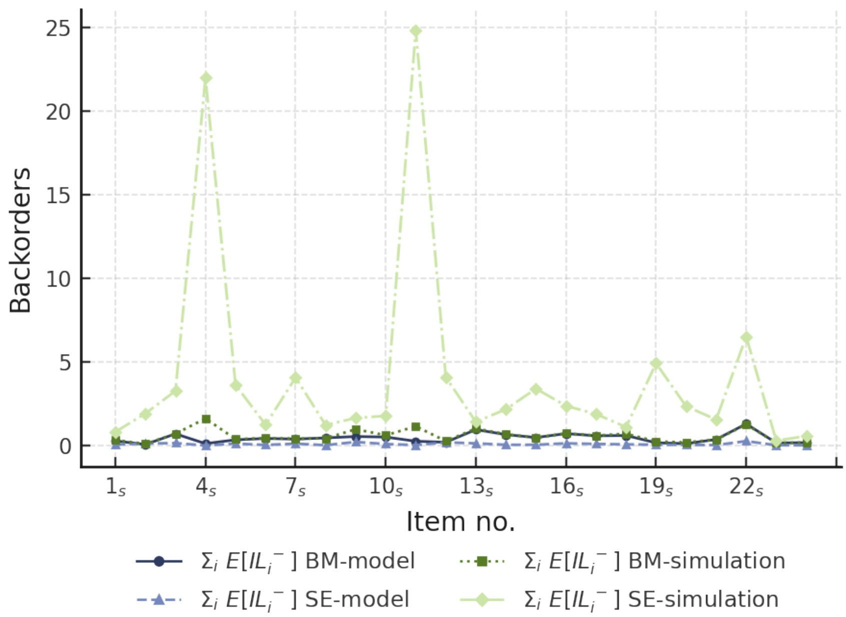

Figure 37: Average backorders at dealers where the special policy is used in the SE-model.

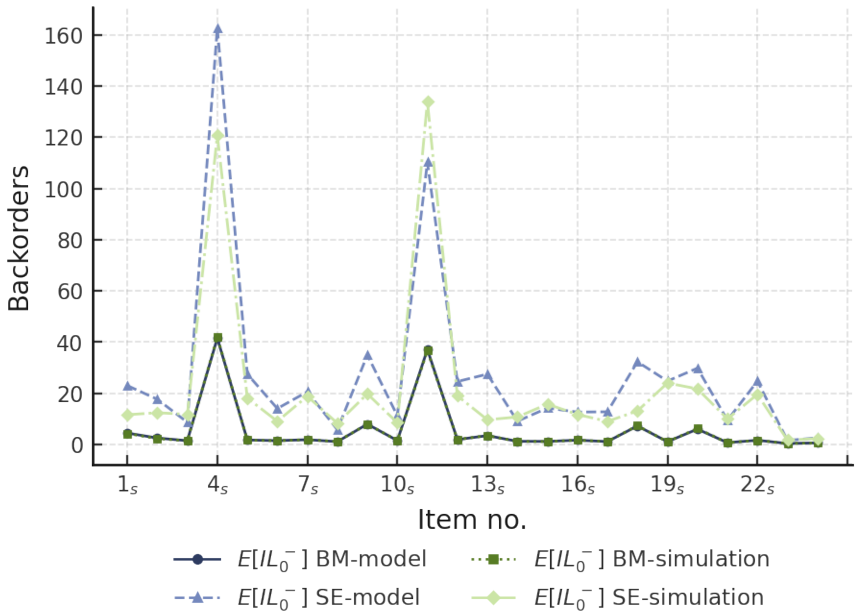

Figure 38: Average backorders at RDC where the special policy is used in the SE-model.

The simulation indicates that the actual number of backorders greatly surpass the special policy SE-model's estimates for most items, as seen in Figure 37. This underestimation of backorders is connected to the special policy SE-model's overestimation of fill rates (see Figure 26). The BM-model in turn, accurately estimates the backorders at the dealers.

The two spikes in Figure 38, i.e. the high number of backorders for item 4_S and 11_S, is a consequence of a relatively high demand at one or a couple of dealers compared to other items in the sample. This high demand in combination with long delays at the RDC, cause spikes in backorders for these items when simulated. Noteworthy is that the deviation from target fill rate for these items do not differ significantly from the other items in the sample as can be seen in Figure 26 and Figure 27.

Figure 39 and Figure 40 present the expected number of backorders for items administered by the regular policy at dealers and RDC, respectively.

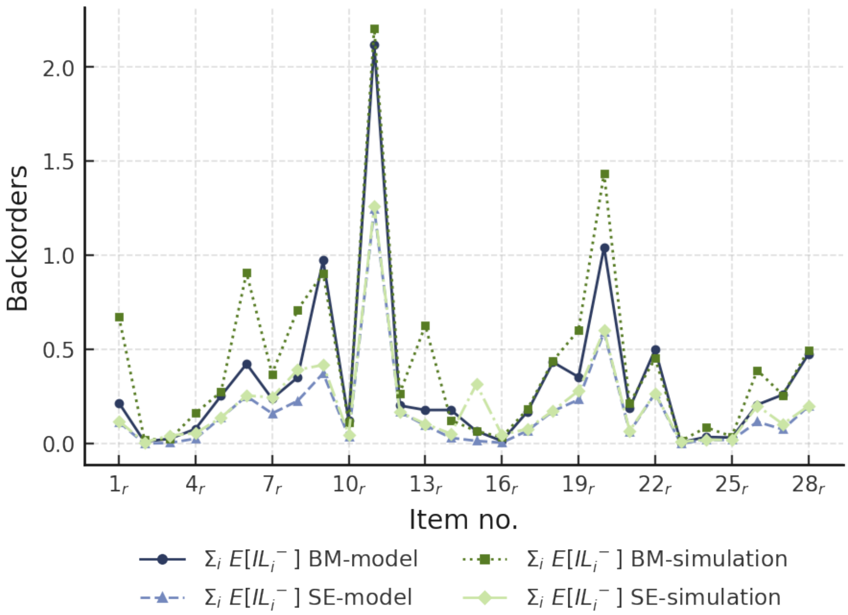

Figure 39: Average backorders at dealers where the regular policy is used in the SE-model.

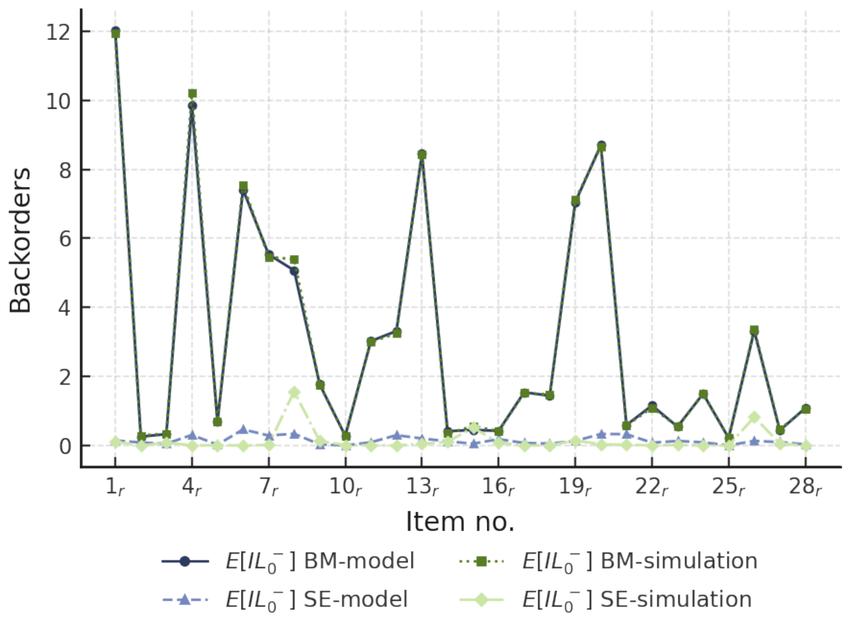

Figure 40: Average backorders at RDC where the regular policy is used in the SE-model.

Studying the backorders at the dealers, the multi-echelon BM-model's performance in terms of accuracy is similar to the regular policy SE-model. Both models allow similar amounts of backorders at the dealers. This correspond well with the fact that the two models both provided reorder points reaching target fill rates, as seen in Figure 28.

Studying the average backorders at the RDC (see Figure 40) the multi-echelon BM-model behaves rather irregular between items. While it is difficult to distinguish any patterns, the graph shows an interesting aspect of the multi-echelon control system. While the regular policy SE-model is quite static and kept at the RDC's target service level, the multi-echelon is allowed more freedom with the inventory levels.

6.6 System holding- and backorder costs

This section studies the total cost that can be expected when using the SE-models compared to using the BM-model.

Figure 41 and Figure 42, show the holding- and backorder costs in terms of absolute values and relative change between models for the items governed by the special policy. Figure 43 and Figure 44 display corresponding results for the items controlled by the regular policy.

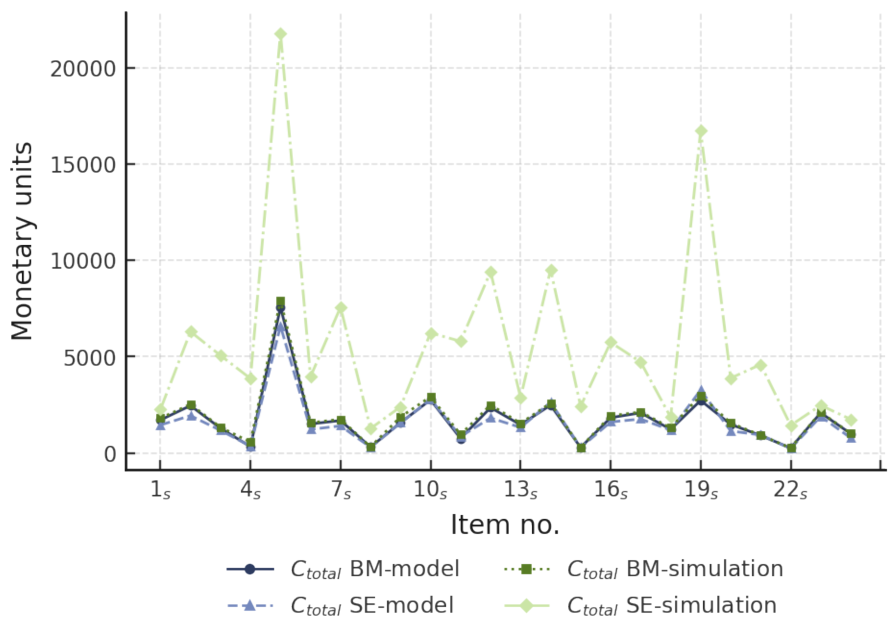

Figure 41: System holding- and backorder costs per item where the special policy is used in the SE-model.

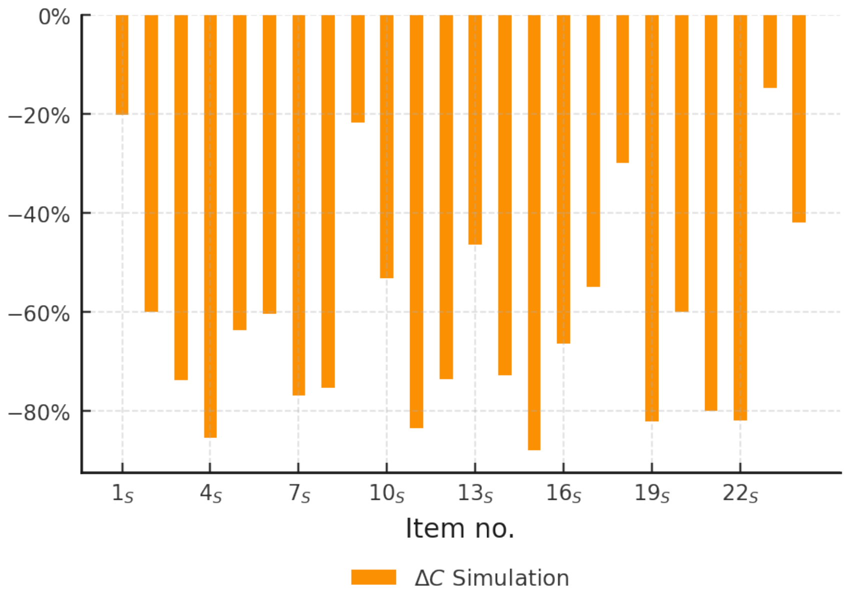

Figure 42: Cost difference per item where the special policy is used in the SE-model.

In Figure 41 it can be observed that estimations obtained from the special policy SE-model deviates largely from the total cost provided by the simulation. Unlike the single-echelon approach, the BM-model accurately estimates the total holding- and backorder costs when compared to the simulation. This correspond to the results of the special policy SE-model underestimating the amount of backorders (see Section 6.5).

In Figure 42, the cost reduction per item is shown for each item. It is evident by this graph that the BM-model provides a more cost-effective system setup. Five items (4_S, 11_S, 15_S, 19_S, 22_S) reduce costs with as much as 80 %

Figure 43 displays the total annual system holding- and backorder costs for the items controlled by the regular policy. The cost estimates produced by the models is generally quite accurate, which has been the case for the regular policy items for the other performance indicators as well.

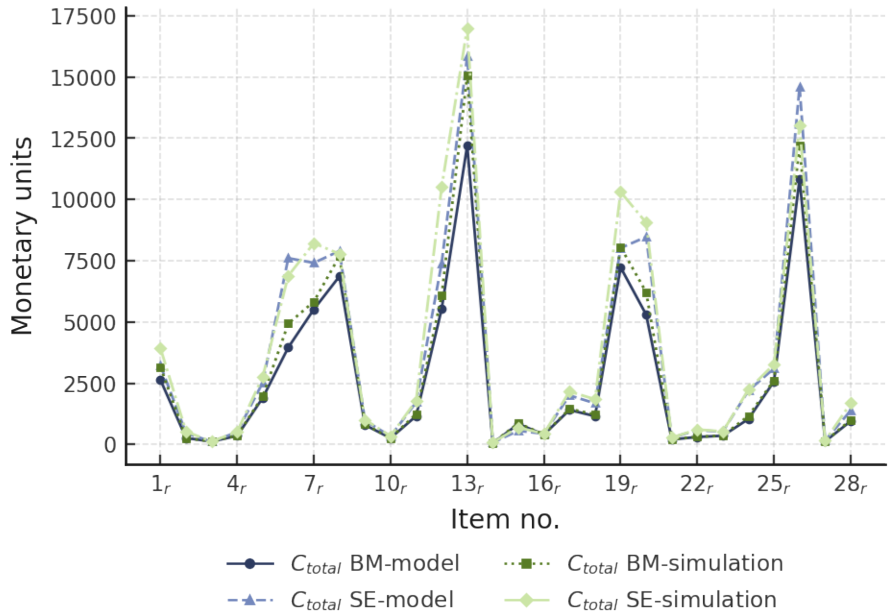

Figure 43: System holding- and backorder costs per item where the regular policy is used in the SE-model.

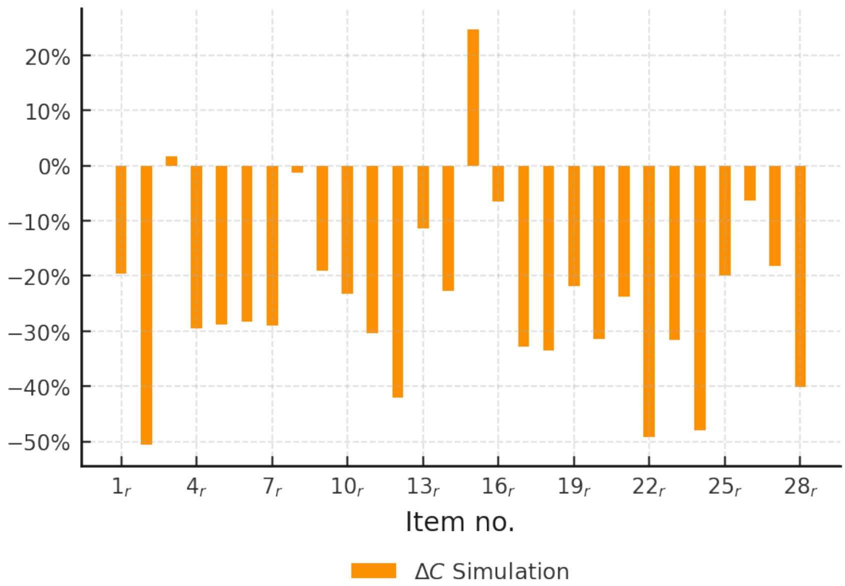

Figure 44: Cost difference per item where the regular policy is used in the SE-model.

In Figure 44 the cost reduction using the BM-model is obvious, reaching up to

50 % total system holding- and backorder costs reduction for item 2_r. The outlier that can be found in item 15_r is the only item that significantly increased costs out of all the items in the study. This increase is due to the increased stock-on-hand as explained in Section 6.4 (when discussing Figure 34).

6.7 Result summary

In Table 7 and Table 8 a summary of the results can be found. The values in the table are based of the simulated values from discrete-event simulation model for the reorder point setups as suggested by the BM-model and the SE-model (regular policy and special policy), respectively.

Table 7: Result summary for items 1_s - 24_s, where the special policy is used in the SE-model.

	BM-model	**SE-model, special policy**
Avg. dealer IFR dev.	Stocked: 1.35 *pp.* Non-stocked: 14.50 *pp.*	Stocked: −18.39 *pp.* Non-stocked: 72.29 *pp.*
Avg. exp. stock-on-hand	58.50 units (+ 85.3 %)	31.58 units
Avg. annual holding- and backorder costs	1862.98 (− 66.6 %)	5584.85

The numerical study indicate that the special policy SE-model has difficulties in meeting target fill rates. For the stocked items fill rate targets are on average 18.39 *pp.* below targets. Moreover, realised fill rates for non-stocked items greatly surpass targets, measuring on average 72.29 *pp.* above targets. The study also indicate that introducing the BM-model would increase the stock for these items. The average system stock-on-hand of the items studied increased by 85.3 %. However, holding- and backorder costs decreased due to the better performance in terms of reaching fill rate targets, thus decreasing backorder costs. The average holding- and backorder costs decreased by 66.6 %.

Table 8: Result summary for items 1_r - 28_r, where the regular policy is used in the SE-model.

	BM-model	**SE-model, regular policy**
Avg. dealer IFR dev.	1.67 *pp.*	−0.08 *pp.*
Avg. exp. stock-on-hand	47.05 units (− 35.6 %)	73.06 units
Avg. annual holding- and backorder costs	3004.67 (− 22.0 %)	3850.24

In the case of the items currently administered by the regular policy SE-model, the BM-model and the regular policy SE-model performed similarly in terms of reaching target fill rates. As shown in Table 8 the BM-model provided a set of reorder points resulting in a slight overshoot, on average 1.67 *pp.* above targets. Likewise, the reorder points suggested by the regular policy SE-model also resulted in realised fill rates close to targets. On average, the use of the regular policy SE-model resulted in fill rates 0.08 *pp.* below targets. Moreover, the average system stock decreased by 35.6 % when using the reorder points suggested by the BM-model, compared to the regular policy SE-model. Lastly, the average annual holding- and backorder-costs of the system was reduced by 22.0 % when controlling the system by the BM-model.

Noteworthy is that the reduction in inventory displayed in Table 8 is similar to the results presented by Berling and Marklund (2014) in their case study. Their study showed a reduction of average system inventory by 30 % - 38 % when using the BM-model. In their analysis they studied a case company of similar nature which used an inventory management system from the same software supplier as what is used by Volvo Group.

Chapter 7

Conclusion

Section 7.1 aims to answer this master thesis' research questions as well as to summarize key takeaways from the results and analysis section. Finally, areas of future research within the domain studied is suggested in Section 7.2.

7.1 Research questions

The problem formulation presented in the beginning of this report focused around two main research questions:

1. What is a suitable way of modeling the Volvo Group Truck Operations - Service Market Logistics supply chain with a multi-echelon approach?
2. What improvements in terms of spare parts availability and cost efficiency could be achieved by using a multi-echelon optimization approach?

Based on the mapping of the inventory control process, the thorough literature study of multi-echelon theory conducted, and results from the numerical study, the authors suggest the Berling-Marklund (BM) model to be a suitable choice for modeling the VGTO-SML supply chain.

The scaling potential for this model is promising. This is an important attribute for the inventory control model as the Volvo supply chain consists of a large number of nodes and SKUs. Already, the single-echelon optimization models used today is computationally demanding.

Furthermore, the BM-model decompose the network into single-echelon problems. This should increase compatibility with the current system in use, making the model an attractive choice. In addition, it has been shown in the numerical study that the model performs well with a shifting number of dealers. The model also allows for optimization of reorder points using both backorder costs or fill rates constraints, providing extra flexibility.

To answer the second question, the results from the numerical study is vital. The study showed that using the BM-model produce significant cost reductions for almost every item in the study. Using the BM-model, 66.6 % cost reduction of the average holding- and backorder costs is exhibited for the items currently controlled by the South African special policy. Furthermore, a 22.0 % reduction of average holding- and backorder costs is displayed for the items governed by the regular policy.

Regarding the comparison between the BM-model and the special policy SE-model, the BM-model also renders satisfying results in terms of spare parts availability. The study showed that reorder points produced by the special policy SE-model was unable to reach target fill rates for dealers with stocked items, on average being 18.4 *pp.* below target. The BM-model, on the other hand, confidently reached target fill rates for each of these items, on average achieving 1.35 *pp.* above target. However, the suggested reorder points by the BM-model also resulted in more stock than before. Average expected stock-on-hand was increased by 85.3 % over all the items associated with the special policy. Nevertheless, the extra stock-on-hand reduces backorders, and thus backorder costs. This reduction is greater than the increase in holding costs resulting in the total cost reduction mentioned above.

The main reason for superior performance of the BM-model compared to the special policy SE-model can be attributed to a superior lead time estimate. The optimal reorder points at dealers obtained from the special policy SE-model assumes no delayed orders due to stockouts at the RDC. When in reality, the low fill rate at the RDC results in large delays. The BM-model handles this both by including the delay in its lead time estimates as well as proposing an RDC reorder point providing higher service.

For items governed by the regular policy, the BM-model performed equally satisfactory to the regular policy SE-model in terms of reaching target fill rates. The BM-model was on average 1.67 *pp.* above target while the regular policy SE-model was slightly below target at a 0.08 *pp.* undershoot. However, the BM-model provided reorder points that could reach target fill rates while keeping a lower amount of stock-on-hand. On average, the expected stock-on-hand of the items currently controlled by the regular policy was reduced by 35.6 %. The BM-model managed this by suggesting a set of reorder point so that stock was allocated further downstream in the supply chain. Thus, the system faced a minor increase in stock at the dealers while experiencing a large reduction of stock at the RDC.

7.2 Future research

This study has been focused around a limited number of items. A larger study, or one focusing on a different range of items in another part of the supply chain network could increase the generality of the results. The number of echelon could also be extended in order to better reflect the structure of the case company's 3-echelon supply chain setup. Furthermore, a sensitivity analysis of the different input variables would also strengthen the reliability of the model while also provide an increased understanding of its limitations, e.g. the importance of accurate demand forecasts.

Further research could also aim to evaluate the performance of different model extensions. Possible extensions could be, but is not limited to: implementing other service constraints, such as waiting time constraints, instead of target fill rates, establishing and modeling a policy to control rush orders, or analyzing the use of capacity restrictions at the installations.

References

S. L. Ahire and C. P. Schmidt. A model for a mixed continuous-periodic review one-warehouse, n-retailer inventory system. *European Journal of Operational Research*, 92(1):69–82, 1996.

N. Altay and L. A. Litteral. *Service Parts Management. Demand Forecasting and Inventory Control.* Springer London, 2011. ISBN 9780857290380.

J. Andersson and J. Marklund. Decentralized inventory control in a two-level distribution system. *European Journal of Operational Research*, 127(3):483–506, 2000.

J. Andersson, S. Axsäter, and J. Marklund. Decentralized multiechelon inventory control. *Production and Operations Management*, 7(4):370–386, 1998.

S. Axsäter. Using the deterministic eoq formula in stochastic inventory control. *Management Science*, 42(6):830–834, 1996.

S. Axsäter. Approximate optimization of a two-level distribution inventory system. *International Journal of Production Economics*, 81:545–553, 2003.

S. Axsäter, C. Howard, and J. Marklund. A distribution inventory model with transshipments from a support warehouse. *IIE Transactions*, 45(3):309–322, 2013.

S. Axsäter. Exact analysis of continuous review (r, q) policies in two-echelon inventory systems with compound poisson demand. *Operations Research*, 48(5): 686 – 696, 2000. ISSN 0030364X.

S. Axsäter. *Inventory Control. [electronic resource].* International Series in Operations Research & Management Science: 90. Springer Science+Business Media, LLC, 2006. ISBN 9780387332505.

S. Axsäter. When is it feasible to model low discrete demand by a normal distribution?. *OR Spectrum*, 35(1):153 – 162, 2013. ISSN 01716468.

L. M. Basson, J. Peter Kilbourn, and J. Walters. Forecast accuracy in demand planning: A fast-moving consumer goods case study. *Journal of Transport and Supply Chain Management*, 13(0):e1 – e9, 2019. ISSN 2310-8789.

P. Berling. On determination of inventory cost parameters. 2005.

P. Berling. Holding cost determination: An activity-based cost approach. *International Journal of Production Economics*, 112(2):829 – 840, 2008. ISSN 0925-5273.

P. Berling and J. Marklund. Heuristic coordination of decentralized inventory systems using induced backorder costs. *Production and operations management*, 15 (2):294–310, 2006.

P. Berling and J. Marklund. A model for heuristic coordination of real life distribution inventory systems with lumpy demand. *European Journal of Operational Research*, 230(3):515 – 526, 2013. ISSN 0377-2217.

P. Berling and J. Marklund. Multi-echelon inventory control: an adjusted normal demand model for implementation in practice. *International Journal of Production Research*, 52(11):3331 – 3347, 2014. ISSN 00207543.

P. Berling and J. Marklund. Corrigendum. *International Journal of Production Research*, 55(17):5183–5183, 2017. doi: 10.1080/00207543.2017.1322671.

J. E. Boylan, A. A. Syntetos, and G. C. Karakostas. Classification for forecasting and stock control: A case study. *The Journal of the Operational Research Society*, 59(4):473 – 481, 2008. ISSN 01605682.

L.-C. Böiers. *Mathematical methods of optimization. Lars-Christer Böiers.* Studentlitteratur, 2010. ISBN 9789144070759.

K. E. Caggiano, P. L. Jackson, J. A. Muckstadt, and J. A. Rappold. Optimizing service parts inventory in a multiechelon, multi-item supply chain with time-based customer service-level agreements. *Operations Research*, 55(2):303–318, 2007.

K. Cattani, F. Jacobs, and J. Schoenfelder. Common inventory modeling assumptions that fall short: Arborescent networks, poisson demand, and single-echelon approximations. *Journal of Operations Management*, 29(5):488 – 499, 2011.

S. Cavalieri, M. Garetti, M. Macchi, and R. Pinto. A decision-making framework for managing maintenance spare parts. *Production Planning and Control*, 19(4): 379–396, 2008. ISSN 09537287.

T. de Kok, C. Grob, M. Laumanns, S. Minner, J. Rambau, and K. Schade. A typology and literature review on stochastic multi-echelon inventory models. *European Journal of Operational Research*, 269(3):955 – 983, 2018. ISSN 0377-2217.

J. Everette S. Gardner. Evaluating forecast performance in an inventory control system. *Management Science*, 36(4):490 – 499, 1990. ISSN 00251909.

J. Feizabadi. Machine learning demand forecasting and supply chain performance. *International Journal of Logistics: Research & Applications*, 25(2):119 – 142, 2022. ISSN 13675567.

B. E. Flores and D. Clay Whybark. Multiple criteria abc analysis. *International Journal of Operations & Production Management*, 6(3):38 – 46, 1986. ISSN 0144-3577.

S. C. Graves. A multi-echelon inventory model for a repairable item with one-for-one replenishment. *Management Science*, 31(10):1247 – 1256, 1985. ISSN 00251909.

H. Guvenir and E. Erel. Multicriteria inventory classification using a genetic algorithm. *European Journal of Operational Research*, 105(1):29–37, 1998. ISSN 03772217.

F. W. Harris. *How many parts to make at once*, volume 10. The magazine of management, 1913.

F. S. Hillier and G. J. Lieberman. *Introduction to operations research, 9th ed.* McGraw-Hill, The McGraw-Hill Companies, Inc., 1221 Avenue of the Americas, New York, NY 10020., 2010. ISBN 978-0-07-337629-5.

W. J. Hopp, M. L. Spearman, and R. Q. Zhang. Easily implementable inventory control policies. *Operations Research*, 45(3):327–340, 1997.

C. Howard, J. Marklund, T. Tan, and I. Reijnen. Inventory control in a spare parts distribution system with emergency stocks and pipeline information. *Manufacturing & Service Operations Management*, 17(2):142–156, 2015.

M. Höst, B. Regnell, and P. Runeson. *Att genomföra examensarbete*. Författarna och Studentlitteratur, 2006. ISBN 978-91-44-00521-8.

A. Lagodimos and S. Koukoumialos. Service performance of two-echelon supply chains under linear rationing. *International Journal of Production Economics*, 112(2):869–884, 2008.

H. L. Lee, V. Padmanabhan, and S. Whang. The bullwhip effect in supply chains. *Sloan management review*, 38:93–102, 1997.

S. Minner, E. Diks, and A. G. de Kok. A two-echelon inventory system with supply lead time flexibility. *IIE Transactions*, 35(2):117–129, 2003.

J. A. Muckstadt and R. O. Roundy. Analysis of multistage production systems. *Handbooks in operations research and management science*, 4:59–131, 1993.

M. Noordhoek, W. Dullaert, D. S. Lai, and S. de Leeuw. A simulation–optimization approach for a service-constrained multi-echelon distribution network. *Transportation Research Part E*, 114:292 – 311, 2018. ISSN 1366-5545.

F. Y. Partovi and M. Anandarajan. Classifying inventory using an artificial neural network approach. *Computers & Industrial Engineering*, 41(4):389 – 404, 2002. ISSN 0360-8352.

D. Peidro, J. Mula, R. Poler, and F.-C. Lario. Quantitative models for supply chain planning under uncertainty: a review. *International Journal of Advanced Manufacturing Technology*, 43(3/4):400 – 420, 2009. ISSN 02683768.

G. A. Rummler and A. P. Brache. Managing the white space. *Training*, 28(1): 55–70, 1991.

N. Sbai and A. Berrado. A strategic framework for multi-echelon inventory system selection. In *2019 International Colloquium on Logistics and Supply Chain Management (LOGISTIQUA)*, pages 1–8. IEEE, 2019.

M. Seifbarghy and M. R. A. Jokar. Cost evaluation of a two-echelon inventory system with lost sales and approximately poisson demand. *International journal of production economics*, 102(2):244–254, 2006.

B. P. Shapiro. Can marketing and manufacturing coexist?. *Harvard Business Review*, 55(5):104 – 114, 1977. ISSN 00178012.

C. C. Sherbrooke. Metric: A multi-echelon technique for recoverable item control. *Operations research*, 16(1):122–141, 1968.

Syncron. *Inventory Calculations and Stock Replenishments Computing [Internal material]*, 2018.

R. Teunter, A. Syntetos, and M. Babai. Stock keeping unit fill rate specification. *European Journal of Operational Research*, 259(3):917 – 925, 2017. ISSN 0377-2217.

L. Turrini and J. Meissner. Spare parts inventory management: New evidence from distribution fitting. *European Journal of Operational Research*, 273(1):118 – 130, 2019.

D. Van den Berg, M. van der Heijden, and P. Schuur. Allocating service parts in two-echelon networks at a utility company. *International journal of production economics*, 181:58–67, 2016.

K. van Donselaar, R. Broekmeulen, and T. de Kok. Heuristics for setting reorder levels in periodic review inventory systems with an aggregate service constraint. *International Journal of Production Economics*, 237:108137, 2021.

T. J. van Kampen, R. Akkerman, and D. Pieter van Donk. Sku classification: a literature review and conceptual framework. *International Journal of Operations & Production Management*, 32(7):850 – 876, 2012. ISSN 0144-3577.

J. Verrijdt and A. de Kok. Distribution planning for a divergent n-echelon network without intermediate stocks under service restrictions. *International Journal of Production Economics*, 38(2):225–243, 1995. ISSN 0925-5273.

Volvo. *Volvo SML intranät [internal material]*, 2022a.

Volvo. *Volvo SCP Training [internal material]*, 2022b.

Volvo. *South African Setup PowerPoint presentation [Internal material]*, 2022c.

D. Wheatley, F. Gzara, and E. Jewkes. Logic-based benders decomposition for an inventory-location problem with service constraints. *Omega*, 55:10 – 23, 2015a. ISSN 0305-0483.

D. Wheatley, F. Gzara, and E. Jewkes. Logic-based benders decomposition for an inventory-location problem with service constraints. *Omega*, 55:10–23, 2015b.

H. Wong, B. Kranenburg, G.-J. Houtum, and D. Cattrysse. Efficient heuristics for two-echelon spare parts inventory systems with an aggregate mean waiting time constraint per local warehouse. *OR Spectrum*, 29(4):699 – 722, 2007. ISSN 01716468.

L. Yang, J. Yang, G. Yu, and H. Zhang. Near-optimal (r, q) policies for a two-stage serial inventory system with poisson demand. *International Journal of Production Economics*, 133(2):728 – 735, 2011. ISSN 0925-5273.

Y. Zang, J. Gao, Z. Xing, and J. Long. Inventory management model of metro vehicle repairable spare parts based on metric. In *2017 36th Chinese Control Conference (CCC)*, pages 2956–2960. IEEE, 2017.

D. Çelebi. Inventory control in a centralized distribution network using genetic algorithms: A case study. *Computers & Industrial Engineering*, 87:532 – 539, 2015. ISSN 0360-8352.

E. Özkan, G.-J. Houtum, and Y. Serin. A new approximate evaluation method for two-echelon inventory systems with emergency shipments. *Annals of Operations Research*, 224(1):147 – 169, 2015. ISSN 02545330.

Appendix A: Interview guide

Name:
Title:
Department:
Date:

Company general
Tell us more about Volvo spare parts, can we access a company description overview?
What does the organization look like, is it possible to access an organization tree?
What type of products/vehicles are you supplying spare parts to?
Who are your end customers?
What characterizes the supply chain of spare parts?
What characterizes the demand of spare parts?

Current distribution network
What does your current distribution network look like?
What types of warehouses are you using?
Central warehouses/retailers/supporting warehouses?
What is the function of the supporting warehouse?
Do you divide your operations into different geographies?
Do you supply different geographies with different spare parts?
Based on these geographies, which are your main markets?
Does the activity differ between different geographies?
Do you use any priority between different geographies/markets?
Does the demand differ between different geographies?
Do you have a map of your distribution network e.g. OWMR?
Do you have control over the inventory management of all warehouses in the chain (including outsourced ones)?
Do you use any VMI (vendor managed inventory) at local warehouses?
What are your main challenges with your current distribution network?
What type of information systems are you using, continuous/periodic?

Current inventory control system
What kind of inventory policy are you using?
Are you using the same policy at all warehouses?
Are you mostly using continuous review or periodic?
If periodic - what periodicity?

Do we have any restrictions on Q? (E.g. fixed, fixed batches = nQ, minimum, maximum)
Do we have any cost data associated with different order quantities?
Are you using collective deliveries, and is this something we need to take into account in the inventory model?
With what time interval is the policy updated, and with what time interval do you want it to be updated?
What is restricting the update-time interval?
Are you updating the policy with regards to only historical data or with regards to a more complicated prognosis?
In case of a stockout, do you backorder the goods or do you count it as a lost sale?
When delivering backorders, are you using a FCFS system, or do you have key customers who get prioritized?
If so, how many are these compared to all customers?
What are the main challenges with your current inventory control system?
Do you use any policy for the use of the supporting warehouse?

Current optimization algorithm
How does the current optimization algorithm work conceptually?
What type of tools are you using (e.g. Python)?
What assumptions or approximations on the inventory system is implied in your algorithm (e.g. demand assumed to be Normal)?
What are the current challenges with the algorithm?
Do you have any concrete improvement ideas on the current algorithm?
Do you have any concrete ideas on 'new' algorithms to be tested?

Available data
On which aggregate levels do you have demand data stored (e.g. SKU-level, SKU at warehouse-location - level, etc.)?
On which time-interval do you store demand data (e.g. weekly, monthly, daily)?
Do we have access to past deliveries with order-time and amount?
Do we have access to delivery times between the different warehouse locations (both transportation times and actual delivery times, i.e. time between order and delivery.)?
Do we have access to stock-level data over time?
Do we have access to past inventory policy-parameters?
Have you seen any differences in the data in the presence of Covid?
Do you think that pre-Covid data will render a better result?

Performance measures

Which are Volvo's KPIs and targets for their inventory management?
What are Volvo's KPIs for the inventory management optimization model?
What are the exact definitions of these KPIs?
If you are using a service-level, what is your definition?
Do you have an estimate of a lost-sales cost (or back-order cost) that we can use in our thesis?

Appendix B: Python model overview

Figure 45, Figure 46, and Figure 47 display a conceptualization of the different maps used in the Python model for both the SE- and BM model. Each figure corresponds to a map containing different modules. The user should read Figure 45 and Figure 46 accordingly:

- The top name, e.g. "single_echelon_utils" in Figure 45, is the name of the folder in Python.
- The blocks to the left, e.g. "dealer_optimization" in Figure 45, corresponds to modules containing different functions. The top block uses functions from the downstream blocks, e.g. "dealer_optimization" uses functions from "service_level_computation" which in turn uses functions from the "demand_models" and "Inventory_levels_computation" blocks.
- To the right in the figures is a visualization of the type of systems modelled in the map. In the bottom right is a list of all the test modules included in the map.
- The optimization approach differs between the two maps. The "dealer optimization" block in "single_echelon_utils" uses service level constraints when minimizing total costs whereas the "warehouse_optimization" block in "warehouse_modeling" uses a backorder cost when minimizing costs, refer to Section 4.6.1 for a more thorough description of the two methods. The expressions for each optimization can be found next to the two blocks in each figure.

Lastly, Figure 47 includes a conceptualization of the main program. This folder uses the method presented in Section 4.10 and builds upon functions from both "single_echelon_utils" and "warehouse_modeling".

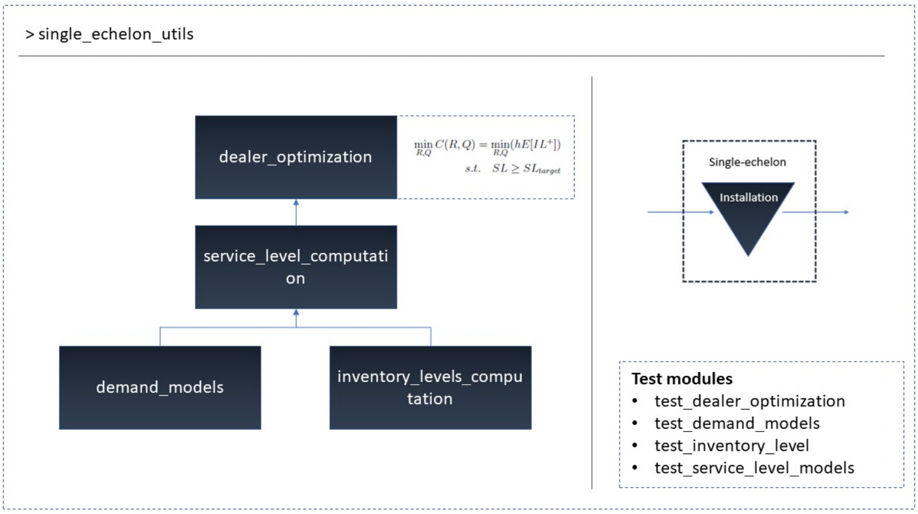

Figure 45: Conceptual overview of single-echelon module.

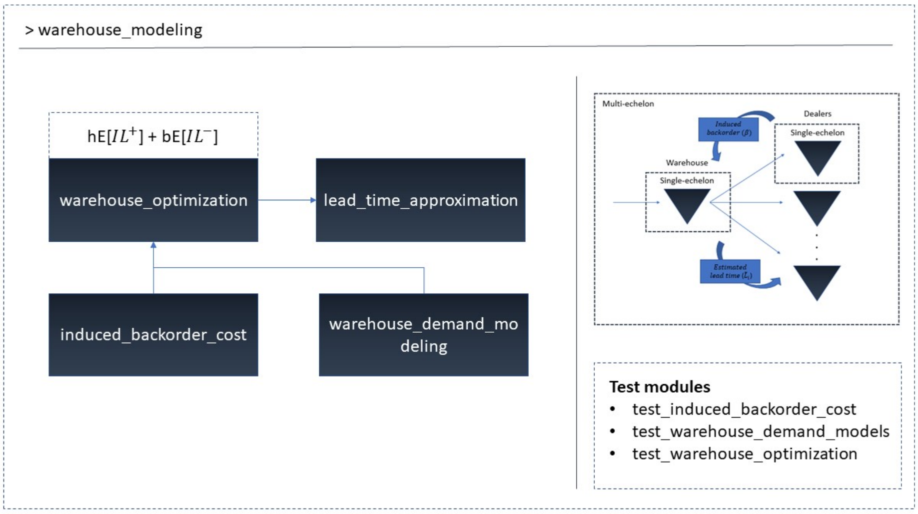

Figure 46: Conceptual overview of modules and functions from the BM-model.

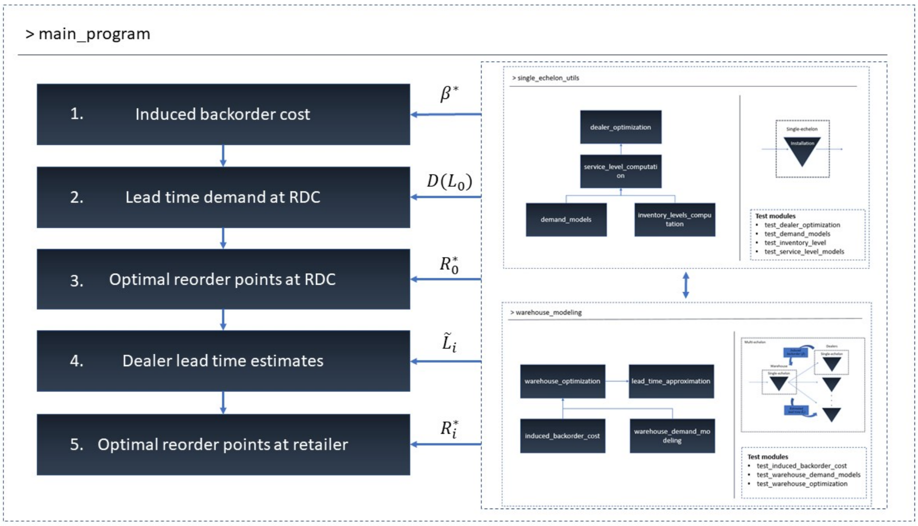

Figure 47: Conceptual overview of main program.