Multi-echelon inventory control in a high complexity high value, and low volume environment

a case study in an assemble-to-order system

Mantje, T.

Award date:
2020

Eindhoven University of Technology
Department Industrial Engineering & Innovation Sciences

Multi-echelon inventory control in a high complexity, high value, and low volume environment

A case study in an assemble-to-order system

by T. (Teun) Mantje

May 3, 2020

Master Thesis

Operations Management and Logistics

University supervisors:
Prof. dr. A.G. (Ton) de Kok – Operations Planning Accounting & Control, TUe
Dr. W. (Willem) van Jaarsveld – Operations Planning Accounting & Control, TUe

Company supervisors:
Dr. M. (Mehmet) Atan – Supply Chain Planning Excellence, ASML
Dr. S. (Bas) van Velzen – Supply Chain Planning Excellence, ASML

Eindhoven University of Technology
School of Industrial Engineering

Series Master Theses Operations Management and Logistics

Key words: multi-echelon, inventory control, hedging, high-tech industry, low volume environment, demand uncertainty, supply chain flexibility, forecast-driven supply chain

Abstract

In the high value and low volume semiconductor equipment manufacturing industry, high demand uncertainty exists. Moreover, the time to increase capacity and source key (sub)modules is long, as well as the cycle times to build modules and assemble systems. Due to these characteristics, planning and control of end items is of high complexity. A strategy for dealing with uncertainty in planning is buffering. This master thesis project is conducted at ASML and examines the applicability of buffering against demand uncertainty. A new buffer planning model is introduced by combining the hedging tactic with the echelon base-stock control policy in a serial system. For each end item, a specific hedging scenario can be defined based on the supply lead times, commitment value curve, and customer order lead time. A simulation model is developed to simulate a serial system with base-stock policy control. Shang & Song heuristic and Diks & De Kok algorithm are able to determine near-optimal echelon base-stock levels for a set target service level. The performance of both analytical methods are evaluated and compared under different hedging scenarios. Furthermore, the current situation at ASML is analyzed and advice is provided on the quantity and position of inventory allocation in the supply chain in order to obtain the desired service level. Finally, trade-off curves are developed which can be used for decision-support as insights are provided in the impact of demand parameters, customer order lead time, costs, and service level.

Management summary

This report presents a master thesis project on multi-echelon inventory control in a high complexity, high value and low volume environment. A case study is performed within ASML, the largest supplier of lithography systems for the semiconductor industry operating with the manufacturing system strategy assemble-to-order.

Problem statement

ASML operates within an industry where high demand uncertainty exists, the time to increase capacity and source key (sub)modules is long, as well as the cycle times to build modules and assemble systems. For each lithography system, up to 100 supply chains of 10-tier deep are required to source thousands of components (ASML SCM, 2019). Due to these characteristics and uncertainties, planning and control of (sub)modules and, systems is of high complexity. Demand uncertainty is experienced as the most important uncertainty in tactical planning due to the impact on the deployment of the integral supply plan. A strategy for dealing with this uncertainty in planning is buffering.

Based on stakeholders' perceptions and judgements, it is identified that within the department Supply Chain Planning, the challenge is to determine the buffer levels in order to maximize customer service level against affordable costs and applicable risks. There are several reasons to motivate the existence of this challenge. For example, the process of defining the type, number, and position of buffers is insufficient. This means that the procedure on when to use which buffer type, insight in where the buffers are positioned, and concrete agreements about buffer levels are unclear. In order to address this challenge, the following research question was formulated:

"Which buffer planning model and workflow could address demand uncertainty in order to maximize service level against affordable costs and applicable risks?"

Uncertainties and buffering concept

The position of master planning is firm-wide with a mid-term horizon. It focuses on the problem of what, where and how to produce on a product type or family aggregated level. The most important uncertainty in master planning for ASML is demand uncertainty, which is experienced in terms of quantity due to frequent up- and downturns of the lithography market and customer requirements. Based on the coefficient of variation, it is observed that there is high variability in ASML's demand figures. The risk of order commitment is high due to the fact that product costs are extremely high, 90% of the product costs are sourced from OEM suppliers, and supplier lead times are long (ASML SCM, 2019).

Buffering is a common strategy for dealing with uncertainty. According to Nahmias & Olsen (2015), buffering is defined as "maintaining excess resources to cover for fluctuations in supply or demand". The buffer strategy determines the way how customer's demand will be met regardless of the variations across the supply chain, this means that buffers should be placed strategically in order to meet customer service requirements. Pipeline safety stock could be utilized to hedge against stockouts or undesirable backlogs. Hedging is a master scheduling tactic to decide on order commitment based on uncertainty and cost commitments. Two phenomena that have led to the hedging concept are that a demand forecast in the far future is less reliable and the commitment costs are time-phased.

Buffer planning model and workflow formulation and current situation

A buffer planning model and workflow results in managing uncertainty in a systematic manner and hence prepare ASML for dealing with future scenarios. The buffer planning model should be applicable for the assemble-to-order production environment, relatable to the current way of planning and buffering (hedging tactic) end items, and connected to the current sales & operations planning (S&OP) structure. The model should be able to make a trade-off between service level and costs in order to use it for decision-support. Finally, a workflow aligned with the S&OP meeting structure should be determined.

At ASML, system buffers are used. These are system starts with a reserved start slot, planned above the firm (expected) demand. To determine the system buffer levels, a tool is used which generates a demand plan and evaluates different replenishment strategies. However, drawbacks of this tool exist. For example, it does not reflect reality as there is no option to evaluate scenarios with high or low demand variability, and it is overdesigned. Therefore, the buffer levels are mainly determined based on mental models of employees. For some system types there is a buffer level agreement. System buffers are removed from the supply plan and canceled at the suppliers or can be pushed out to a different start week if no customer is allocated in a certain time window before final assembly (FASY) start. In both situations upstream inventory is created at suppliers.

Buffer planning model design

In the supply plan, the final assembly of systems are planned in a specific week in the future. This could be translated to a number of planned systems in the pipeline towards the FASY start. Therefore, ASML's system planning on tactical level could be seen as a serial system with a number of stages. In the serial system, the stages could be seen as hedging positions. At these hedging positions, the number of planned systems could be controlled in terms of regular and buffer system starts, i.e. a hedge could be made at that position to plan system buffers to cope with variability in demand. The inventory control policy that can be applied in a serial system is a base-stock policy.

To determine the number and place of the hedging positions, different aspects should be considered. First, the positions in the serial system derived from the assembly system represent when ordering decisions should be made for items and it creates the timing of assembly starts. Second, the commitment value curve visualizes the jumps in cumulative added value of component procurement commitments and costs related to manufacturing activities. Before large value jumps, hedging positions could be placed in order to create a buffer before a high value should be committed. Finally, the customer order lead time determines the decoupling point between the forecast-driven supply chain and customer-order-driven supply chain. This means that the last hedging position is always just before the customer order lead time upstream in the supply chain.

Experimental analysis

A simulation approach is developed to generate hedging scenarios and simulate a serial system with base-stock policy control. The Shang & Song heuristic and Diks & De Kok algorithm are able to determine near-optimal echelon base-stock levels for a set target service level. In this experimental analysis, both analytical methods are evaluated and compared under different hedging scenarios for a target ready rate and a target fill rate.

The Shang & Song heuristic achieves the lowest mean absolute percentage errors (MAPEs) for a ready rate target; therefore, it is able to determine echelon-base stock levels which results in a ready rate closer to the set target than Diks & De Kok algorithm. However, the Diks & De Kok algorithm determines better cost optimal echelon base-stock levels as the required costs to increase the ready rate by 1% compared to ready rate of the other analytical method are significant lower. When evaluating the performance of the analytical methods for a target fill rate, the Diks & De Kok algorithm outperforms the Shang & Song heuristic significantly. Hence, the Diks & De Kok algorithm is able to determine better cost optimal echelon base-stock levels which could achieve a fill rate closest to the set target, the MAPE for a 95% target is only 0.9%. The difference between the target service level and actual service level provided by simulation can be explained by different reasons. Shang & Song (2003) evaluated the heuristic only for high demand rates and short lead times with different demand distributions. For the Diks & De Kok algorithm, the fact that in the simulation there is a low demand per time unit combined with discrete demand is a source of deviations for the algorithm as this uses a continuous demand distribution. When the service level decreases, the effect of discrete demand is strengthened.

Case study

The simulation model is used in order to evaluate the base-stock levels in the current situation at ASML and performance in practice. Whereas, the desired situation is obtained by applying the Diks & De Kok algorithm as fill rate is the most common service level used in practice. The percentage error between the current and desired fill rate of system type A is -24.3%. Furthermore, an additional inventory investment of factor 2.1 with respect to the current investment is required to obtain the desired fill rate. This inventory investment is especially required at the last hedging position. However, the local base-stock levels should primarily be increased at the first and last hedging position to obtain the desired fill rate.

To make optimal buffer decisions, it is important that the relationship between customer order lead time, fill rate, demand variability, and costs are analyzed. Trade-off curves are developed by defining multiple scenarios. The impact of customer order lead time on the costs is high, i.e. the total supply chain costs decrease significantly when the customer order lead time increases. When fill rate decreases, less buffers are required due to the allowance of backorders, and hence lower inventory investments have to be made. Furthermore, as the standard deviation determines the variability to buffer against, it has a large impact on the costs. When a specific scenario in terms of customer order lead time and required fill rate is chosen from the trade-off curves, the corresponding echelon base-stock levels are checked and translated to echelon buffer base-stock levels. These are decreasing over time, which means that uncertainty decreases over time.

An iterative process is required to ensure that the demand uncertainty is addressed at the correct positions and with the correct quantity in order to achieve a desired service level against the lowest costs. The first step in the workflow is that the business objectives and market insights should be clear. Second, the commitment value curve and hedging scenarios per lithography system platform are defined. Then, the buffer planning model should be applied. Informative graphs and trade-off curves between inventory investment, fill rate, customer order lead time, and demand variability can be created, which support decision-making. The jointly decided buffer base-stock levels are guidelines for the end item planners to create the supply plan. After deployment, this plan is reviewed in order to track the performance. The buffer planning workflow should be performed with a quarterly frequency to ensure that buffer levels are gradually adjusted.

Preface

This report is the result of my master thesis project conducted at ASML in order to fulfill my education at Eindhoven University of Technology (TUe). After obtaining my Bachelor Industrial Engineering and Management at Amsterdam University of Applied Sciences, my next adventure started at the TUe with a Premaster's program followed by a Master's program. This report marks the end of my master Operations Management and Logistics, from which I acquired multidisciplinary academic knowledge in the field of improving efficiency and effectiveness of operations in businesses. In this preface, I would like to take the opportunity to thank a number of people who had a supportive role in this research project.

First of all, I would like to thank my first supervisor Ton de Kok. I was surprised, but honored, when I received your invitation at 07:11 AM on the day supervisors were able to match with students. You are a professor with interests in the application of mathematical optimization models to describe and support businesses. Your extensive academic knowledge and industry experience helped me during this master thesis project to formulate the problem and to develop a business-oriented model. In my opinion, the progress meetings at TUe and ASML were intensive and effective. After a meeting I had to digest the useful information. These meetings were not taking place without your assistant José van Dijk-Kok, thank you for the quick responses to my meeting requests. Secondly, I would like to thank my second supervisor Willem van Jaarsveld. Without your and Zümbül Atan's lectures of the course Multi-Echelon Inventory Management, I was not able to conduct this master thesis project this way. I was glad that you, with your in-depth inventory management knowledge, confirmed my progress and provided me with useful feedback.

Furthermore, I am really thankful for the opportunity to conduct my master thesis project at ASML. This opportunity was created by the informative and inspiring interview with Jelto Bijlsma during the European Supply Chain Forum (ESCF) speed dates. Next, I would like to thank Mehmet Atan for the supervision in this project. You guided me through the organization and quickly learned me most ins and outs regarding demand and planning processes. I felt treated as an equivalent colleague, you gave me responsibility, and the freedom to explore ASML. Furthermore, our weekly meetings provided me with guidance and improved my work each time. At a later stage, Bas van Velzen joined our meetings to have in-depth discussions about the current planning processes and how literature matches and could contribute. Thank you Bas for your time and inventory management knowledge. Next, I would like to thank Vincent Eveleigh who joined the progress meetings biweekly at a later stage. You provided me with valuable feedback from a business perspective. Furthermore, I would like to thank my team and other colleagues of ASML for creating a warm and friendly environment, providing me with input and data, and having nice lunch breaks together.

Finally, thanks to my parents, family, friends and girlfriend for their support and interest in my research project.

Teun Mantje

Table of contents

List of abbreviations

ASSY	assembly
ATO	assemble-to-order
ChainScope	supply chain optimization planning engine
CLT	customer order lead time
CODP	customer order decoupling point
CTO	configure-to-order
CV	coefficient of variance
DUV	deep ultraviolet
EUV	extreme ultraviolet
FASY	final assembly
FR	fill rate
MAPE	mean absolute percentage error
MPS	master production schedule
RR	ready rate
SBS	synchronized base-stock (policies)
SCP	Supply Chain Planning (department)
S&OP	sales and operations planning

List of tables

List of figures

1. Introduction

This report presents a master thesis project on multi-echelon inventory control in a high complexity, high value, and low volume environment. This study is conducted at ASML, the largest supplier of lithography systems for the semiconductor industry. In the semiconductor equipment manufacturing industry high demand uncertainty exists, the time to increase capacity and source key (sub)modules is long, as well as the cycle times to build modules and assemble systems. For each lithography system up to 100 supply chains of 10-tier deep are required to source thousands of components, where around 90% of the product costs are sourced from OEM suppliers (ASML SCM, 2019). Due to these characteristics and uncertainties, planning and control of (sub)modules and, systems is of high complexity. A strategy for dealing with these uncertainties in planning is buffering. In this study a buffer planning model is developed for the multi-echelon supply chain by combing hedging and multi-echelon inventory literature to address demand uncertainty.

This chapter is structured by first describing the research methodology, which will be the structure of this thesis. Second, ASML is introduced with general facts about customers, product portfolio, and main figures. Hereafter, ASML's supply chain and planning processes are described. Fourthly, the problem context and problem definition are stated. Next, the design of the research is determined in terms of research goal, scope, and research questions. Finally, the outline of this master thesis project is described.

1.1 Research methodology

According to van Aken & Berends (2018), there are two process structures for conducting research in the business and management field. These structures are defined as the empirical cycle and problem-solving cycle. As this research is driven by a business problem, the problem-solving cycle is applicable. The most important characteristics of this cycle are that the problem mess in an organization is the starting point for research and that the methodology is theory-informed. As can be seen in Figure 1, the first step of the cycle is to identify and structure the problem mess, which result in a clear problem definition. This problem context and description is defined in section 1.3. This research project will continue following the steps of the problem-solving cycle. In the analysis and diagnosis, the problem will be quantified. Furthermore, the problem and context are analyzed in order to identify the cause and effects. Based on this diagnosis, a solution will be designed that tackles the most important causes. To design this solution, a systematic literature review should be performed to identify possible concepts that can contribute to solve the problem. This literature review will be combined with knowledge of involved people to design a solution. The last part of this research project is to provide recommendations regarding the implementation process. The intervention, evaluation and learning steps are out of scope in this research project.

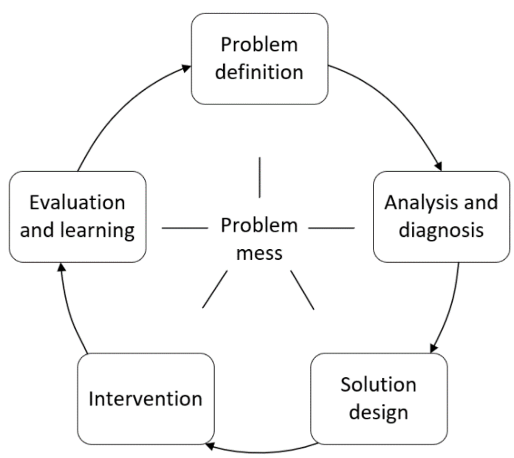


Figure 1. Problem-solving cycle (van Aken & Berends, 2018)

1.2 Company introduction

ASML is established in 1984 by a joint venture of Philips and Advanced Semiconductor Materials International (ASMI) and is currently the largest supplier of lithography systems for the semiconductor industry. ASML employs around 25,000 people across locations in 16 countries, with corporate headquarters located in Veldhoven, the Netherlands (ASML, 2019). In 1995, ASML became a fully independent public company with shares traded on the stock exchanges NASDAQ and Euronext Amsterdam. The capital raised by this initial public offering was required for future growth. From the establishment until now, ASML is growing at a fast pace due to the increasing demand of microchips. ASML is continuously innovating in order to meet the accelerating demands of their customers, where microchips must be smaller, faster, and be produced more cost efficiently.

ASML provides customers with lithography systems, metrology & inspection systems, and customer support. Lithography systems can be distinguished in extreme ultraviolet (EUV) and deep ultraviolet (DUV). ASML is organized in five different business lines, namely: EUV, DUV, Applications, Cymer Light Sources, and Mature Products & Services. The DUV net system sales contributed for around 66% to the total net system sales (ASML, 2019). These figures imply that DUV systems are an important driver of ASML's growth, and it is expected that over the next 5 to 7 years DUV systems remain important (ASML, 2018)

According to the annual report 2019 the net sales, gross profit, and R&D expenses in 2019 were €11.8 billion, €5.2 and €1.9 billion respectively (ASML, 2019). The growing opportunities for 2025 in annual net sales are €15 billion with low market scenario and €24 billion with high market scenario (ASML, 2019). According to The Information Network (2018), the revenue market share in 2017 of ASML in the semiconductor lithography market was 85.4 percent.

1.2.1 ASML's supply chain

The production process of ASML within the integral supply chain is shown in Figure 2. In the assembly (ASSY) process, the different produced and sourced submodules are assembled into modules in specific work centers, e.g. wafer stage, laser, and lens. The produced and sourced modules are assembled in the final assembly (FASY) process. The cycle times of ASSY and FASY are long and vary between the lithography (sub)module and system type. The customer order decoupling point (CODP) is an important point in the supply chain and indicates where a lithography system is related to a specific customer order. This point is located before a part of ASSY and the complete FASY, which means that some modules are made to stock and others are assembled to order. Therefore, the supply chain activities upstream from the CODP are forecast-driven. On the other hand, the downstream supply chain activities are customer-order-driven. The manufacturing situation which correspond with this CODP is assemble-to-order (ATO). However, as ASML offers different main configurations for each system type to enable customer specific lithography systems, the production situation can be defined as configure-to-order (CTO). In a CTO environment, customers are able to select components from predefined subsets, which is defined as a configuration. This enables companies to deliver more customized products to their customers. The difference between ATO and CTO is minor in terms of operational level; however, important for the demand information level (Song & Zipkin, 2003). In addition to main configurations, customer can request factory options. These options are additional upgrades in hardware and software to enhance speed or quality.

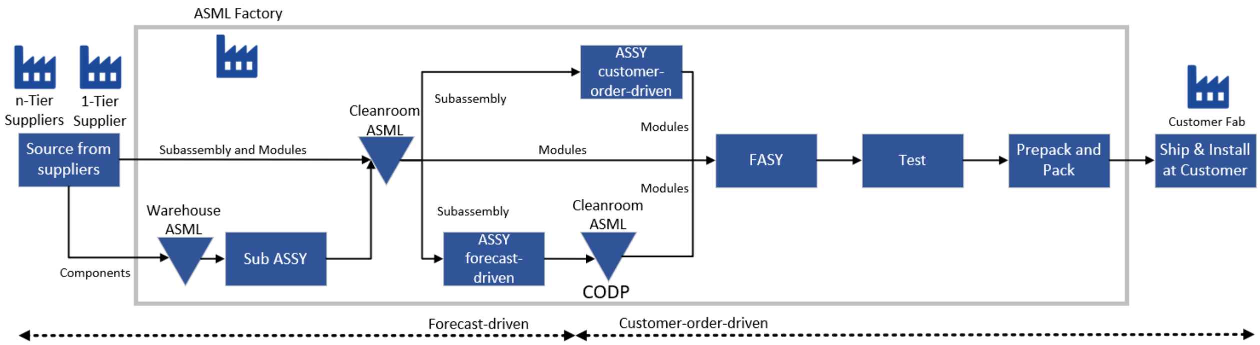


Figure 2. ASML's production process and CODP

1.2.2 ASML's supply chain planning and control processes

The case study in this research project is conducted in the Supply Chain Planning (SCP) department which is part of the Supply Chain Management sector. SCP focuses primarily on delivering the best integral supply plan for customers, factories, and suppliers. The supply plan should be feasible in terms of capacity and material constraints. For this, conscious decision-making is required where business line requirements have to be taken into account.

Three different planning levels are distinguished within ASML, namely strategic, tactical and operational. The strategic planning is focused on long-term market review and capacity planning, and is reviewed quarterly. Tactical planning is focused on mid-term planning with a monthly plan cycle. The plan cycle starts with a meeting with all stakeholders to discuss constraints and agreements of business lines' financial targets with required output to support these goals, opportunities and risks in demand, supply constraints, and risks of new product introductions. This meeting concludes with an agreement on the planned output capability that should be met. This planned output capability defines how much lithography systems could be produced by taking into account the agreements and constraints. The planned output capability, demand plan, and planning parameters (e.g. lead times of critical parts and buffer level agreements) are used as input to determine the integral supply plan. This plan is created by end item planners from the SCP department and should maximize the future expected system outputs with a customer order. It determines and secures the output dates of systems with a horizon of six quarters. System planning is of high complexity due to uncertainty in demand, supply and technology maturity. The final step of the tactical planning is agreement on the integral supply plan, which is performed in the S&OP deployment meeting. After agreement, the supply plan will be implemented. Operational planning focuses on the deployment of the integral supply plan. The review of this planning level is on a weekly and daily basis. Operational planning consists of material checks, decisions on FASY system start dates, critical part allocations, and production progress monitoring. An overview of the different planning levels and related activities at ASML are shown in Figure 3.

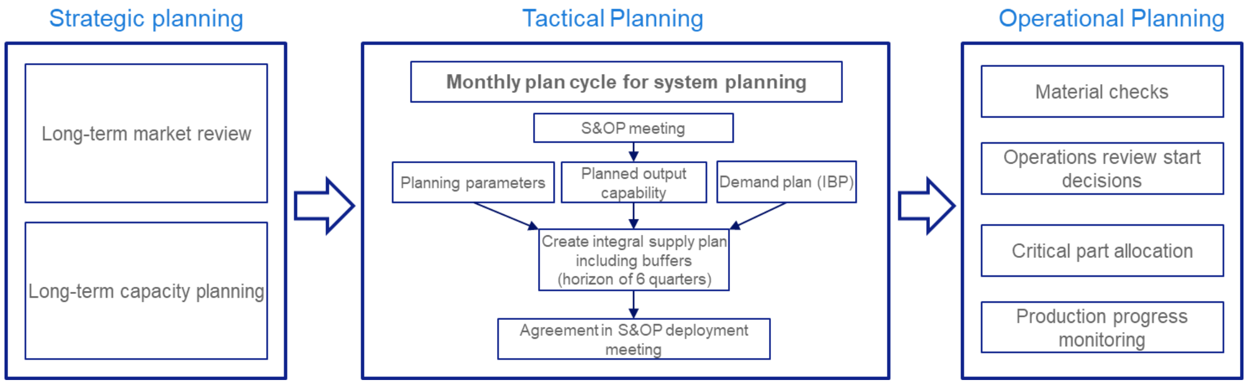


Figure 3. ASML's planning levels and activities

1.3 Problem description

1.3.1 Problem context

The focus of this research project will be on the DUV business line because DUV lithography systems are mature and are produced in volume, i.e. 203 DUV systems out of 229 total lithography systems are produced in 2019 (ASML, 2019). Within the DUV business line a main distinction can be made between immersion and dry systems, which is the Twinscan NXT and Twinscan XT system family respectively. Both system families have four types: NXT (2000, 1980, 1970 and 1965) and XT (1460, 1060, 860 and 400) (ASML, 2019). However, in the future, systems types could be upgraded, or new system might be introduced.

The semiconductor equipment manufacturing industry, where ASML operates in, can be characterized as a low-volume, high-value business, i.e. in 2019 ASML sold 229 lithography systems with a net system sales of €9 billion (ASML, 2019). In this industry, time to increase capacity and source key (sub)modules is long, as well as the cycle times to build modules and assemble systems. For each system up to 100 supply chains of 10-tier deep are required to source components, where around 90% of the product costs are sourced from OEM suppliers (ASML SCM, 2019). Furthermore, ASML is facing different uncertainties related to demand, supply, technology maturity, material quality, wear and tear, and engineering changes. According to interviews with stakeholders, the most important uncertainties for tactical planning are demand, supply and technology maturity. The focus will be on demand uncertainty as this is within ASML the most significant uncertainty that affect the deployment of the integral supply plan and could be mitigated by using buffering techniques.

1.3.2 Problem definition

Due to the semiconductor equipment manufacturing industry characteristics and uncertainties, planning and control of (sub)modules and systems is of high complexity and one of the key processes. A strategy for dealing with these uncertainties in planning is buffering. Based on stakeholders' perceptions and judgements, it is identified that within SCP the challenge is to determine the buffer levels in order to maximize customer service level against affordable costs and applicable risks.

There are several reasons to motivate the existence of this challenge. Firstly, the process of defining the type, number and position of buffers is insufficient. This means that the procedure on when to use which buffer type, insight in where the buffers are positioned, and concrete agreements about buffer levels are unclear. Secondly, not for all DUV system types a model is used. Finally, the impact of buffer decisions in terms of customer service level, operational, and inventory (commitment) costs of different scenarios is not insightful, which result in a difficult decision-making process. These inefficiencies in planning buffers might impact the stability of the supply plan, customer service level, inventory (commitment) and operational costs.

A cause and effect diagram is visualized in Appendix I. As can be seen, the two starting points in this diagram are “insufficient buffer model and workflow” and “demand uncertainty”. On high level, the fact that the buffer model and workflow is insufficient, leads to suboptimal buffer decisions regarding type, position, and number. These suboptimal decisions result in integral supply plan changes, e.g. if there is an excessive system buffer number, the starts in the plan will be postponed or canceled. These changes will affect the production planning of the factory and suppliers. Demand uncertainty could result in expected future demand, unexpected future demand, or customers that cancel, postpone, or bring forward orders. Furthermore, customers could switch system type or main configuration. These situations result in changes in the integral supply plan, i.e. push out, pull in or remove systems in the plan. The final effects in this diagram are that the production plan is instable, customer service level is decreased, and inventory and operational costs are increased.

In conclusion, the problem statement is as follows: *“The model and workflow to define the type, number, and position of buffers in the supply plan to address demand uncertainty is insufficient, which result in suboptimal buffer decisions that are affecting service level and costs.”*

1.4 Research design

1.4.1 Research goal

Due to the current problem as described in section 1.3, the SCP DUV business line is interested in a buffer planning model and aligned workflow that could maximizes customer service level against affordable costs and applicable risks. A trade-off should be made between these three aspects. The model should determine what, where/when and how much to buffer, whereas the workflow should describe how this model is used in planning activities.

In conclusion, the research goal is: *“Design a buffer planning model and workflow that addresses demand uncertainty in order to maximize service level against affordable costs and applicable risks.”*

1.4.2 Scope

This research is mainly focused on ASML’s tactical planning level and to some extend to the operational planning level. The buffer planning model focuses on the forecast-driven supply chain and should be developed on system level to be applicable in the process of creating the system supply plan. The focus is on new system demand, which means that demand for refurbished systems, and Development & Engineering testing and training systems are excluded. Second-tier and further upstream suppliers are out of scope in this research.

1.4.3 Research questions

Which buffer planning model and workflow could address demand uncertainty in order to maximize service level against affordable costs and applicable risks?

1. What is the concept "buffering" and why is it required?
2. What is the uncertainty in demand and risks for tactical planning at ASML?
3. How to formulate a buffer planning model and how to implement this with a workflow and IT?
4. How is the model and workflow organized in the current situation?
5. Which model could address demand uncertainty and enables to make a trade-off between customer service level, costs and risks?
6. How can the performance of the model be measured?
7. What are the optimal buffer decisions regarding the type, position, and number of buffers for a specified service level objective?
8. How should the workflow be organized?

1.5 Thesis outline

The remainder of this thesis is described according problem-solving cycle. In the next chapter the different types of uncertainties according to literature and within ASML are described. Furthermore, the concept of buffering is explained and multiple buffer mechanisms to address the uncertainties will be elaborated. Chapter 3 describes how a buffer planning model should be formulated and implemented with a workflow and IT. ASML's current way of working regarding buffer planning is discussed in Chapter 4. The solution design starts in Chapter 5, where a buffer planning model to address demand uncertainty is developed based on the hedging and multi-echelon inventory control literature. In Chapter 6 it is explained how the buffer planning workflow should be organized. The final conclusions and recommendations of this research are discussed in Chapter 7. This thesis will be concluded with the discussion in Chapter 8.

2. Uncertainties and buffering concept in master planning

2.1 Master planning

The position of master planning is firm-wide with a mid-term horizon. It focuses on the problem of what, where and how to produce on a product type or family aggregated level. The output will be planned production and inventory quantities specified for each product type family in each period. The main objective is to effectively coordinate the material flows across multiple stages in the supply chain to meet customer demand (Mönch, Uzsoy, & Fowler, 2018). For the semiconductor industry a long planning horizon is required since at least the upstream supply chain cycle time should be covered. This ensures to address the production of components and (sub)modules that are required for end items in order to guarantee material availability. Demand planning delivers input to master planning in terms of firm orders and demand forecasts. Master planning combined with demand planning is sales and operations planning (S&OP).

The S&OP provides important visibility into the interaction between sales, marketing, production and finance (Jacobs, Berry, Whybark, & Vollmann, 2011). In this process, top management is able to control the business because important trade-off decisions are made. In the traditional way, first the forecast quantities are determined by demand planning and subsequently master planning use this as input to determine the master plan (Albrecht, Rohde, & Wagner, 2015). However, performing these activities subsequent could result in a plan where costs are more expensive and suboptimal from an organizational perspective. This suboptimal organizational perspective refers to the conflicting goals of the supply and demand side parties, e.g. production, sales, product development, procurement. Therefore, it is important that all stakeholders are involved in the S&OP process to align on those goals, be familiarized with the impacts of the master plan and perceive accountability for the results.

Master planning decision is a negotiation process between stakeholders and should be supported by designed mathematical and computational tools. The S&OP process drives the master production schedule (MPS), where is stated which CODP items at which point in time will be manufactured to meet the demand. The MPS is the disaggregated version of the sales and operations plan.

Master planning is related to the Supply Chain Operations Planning (SCOP) function of De Kok & Fransoo (2003) in terms of the coordination of material flow in a supply network. The objective of this coordination is to minimize the inventory and backorder costs over a time horizon by taking into account material availability and capacity constraints. However, SCOP can be distinguished from master planning in terms of the specification of the release quantities of materials and resources. Therefore, the time frame of SCOP overlaps mid- and short-term planning, i.e. mid-term planning decisions are translated into short-term planning decisions.

2.2 Uncertainties in supply chains

Due to the semiconductor equipment manufacturing industry characteristics and uncertainties, the planning and control of (sub)modules and systems is of high complexity and one of the key processes. ASML is facing different uncertainties related to demand, supply, technology maturity, material quality, wear and tear and engineering changes. Interviews with stakeholders are conducted to identify the uncertainties for tactical planning. The uncertainties for tactical planning are demand, supply and technology maturity because these uncertainties affect the deployment of the integral supply plan. These uncertainties are discussed in more detail below after the description of the different uncertainties in literature. The focus in this research is on demand uncertainty, as this is the most important uncertainty that should be addressed by using a buffer planning model and workflow.

First, the different uncertainties according to literature are described. Wazed, Ahmed, & Yusoff (2009) states that with stochastic uncertainty the future system state is unknown due to lack of information and vagueness of events. Uncertainty can be categorized in two different groups: environmental uncertainty and system uncertainty (Ho, 1989). Environmental uncertainty includes uncertainties outside manufacturing systems' boundary, e.g. demand and supply uncertainty. Whereas system uncertainty are uncertainties associated with the production process, e.g. production lead time and quality uncertainty (Wazed, Ahmed, & Yusoff, 2009).

In tactical planning, significant sources of environmental and system uncertainty are faced, where corresponding risks should be managed and/or mitigated. Nahmias & Olsen (2015) distinguish uncertainty in four different types, namely variation, foreseen uncertainty, unforeseen uncertainty and chaos. Variation is defined as anything that affects the manufacturing system due to regular and predictable behavior, e.g. variability in customer orders and lead times. Variation can be described by a probability distribution. Risks that can be identified beforehand and planned for are defined as foreseen uncertainty, e.g. supply breakdowns. This uncertainty has larger impact than variation and is often described by occurrence probability. On the other hand, the risks of unforeseen uncertainty cannot be predicted ahead and are very unlikely. Finally, chaos is the same as unforeseen, but the impact is on operations and company level. The focus in this research will be on the uncertainty type variation, as these can be mitigated by developing strategies. The other uncertainty types require contingency plans or cannot be mitigated.

Hopp & Spearman (2000) distinguish variation in two types: controllable variation and random variation. Controllable variation is a consequence of decisions, e.g. variability of physical dimensions and manufacturing time for different products. Whereas random variation is a consequence that is beyond companies' control, e.g. timing of customer demand. The focus in this research will be on random variation of environmental uncertainties as the effects are larger and more complicated models are required to mitigate these.

2.2.1 Demand uncertainty ASML

ASML distinguish demand in expected (firm) demand and upside demand. The lithography system demand volume directly depends on the investments of chip manufacturers. Therefore, frequent up- and downturns are present in the lithography market. These demand signals could be unnoticed for ASML in the short-term. In the long-term, large fluctuations exist due to the business cycles, e.g. technological innovation of 5G network.

Furthermore, there is uncertainty in customer requirements, which means that it is not always transparent which lithography system customers prefer and when this system should be delivered. This result in situations where customers might bring forward or postpone the requested delivery date. Customers could also switch configuration. The reasons for uncertainty in customer requirements could be due to the fact that customers' semiconductor fabrication plants are not ready, new electronic products releases are delayed, or enhanced system specifications are required.

Variability could be measured in terms of absolute (variance) and relative variability (coefficient of variation). Relative variability is important to make comparisons. The coefficient of variation (CV) can be used as measure of the relative variability of a random variable [1]. There is low variability if CV < 0.75, moderate variability if 0.75 ≤ CV ≤ 1.33, and high variability if CV > 1.33 (Hopp & Spearman, 2000). According to this calculation, it is observed that there is a high variability in ASML's demand figures.

$$Coefficient\ of\ variation\ (CV) = \frac{\sigma}{\mu} \qquad [1]$$

2.2.2 Supply uncertainty ASML

As sourced (sub)modules are complex, there exist uncertainty in the supply lead time of sourcing these. This uncertainty is faced regarding quantity and timing. Quantity uncertainty means that the supplier is not capable of producing the requested quantity, where timing uncertainty means that there is a probability that the actual supply lead time deviates from the agreed supply lead time. The agreed lead time of (sub)modules could increase up to a high number of months. When suppliers are not able to meet the order due date, this could have impact on the FASY start dates. Furthermore, in ASML's factory there could exist variability in cycle times of module build (ASSY) and lithography system assembly (FASY). Uncertainty in supply could result in not meeting the customer requested delivery date.

2.2.3 Technology maturity

Risk exists in technology maturity of components or (sub)modules that are sourced. It can be the case that the supplier is not able to finish and test the component or (sub)module to ensure a qualified buy to ASML. It can be decided to purchase the component or module and produce and/or test at an ASML location. This means there is uncertainty in technology performance and therefore a backup plan is required. This plan could be to additionally purchase the mature component or submodule to be able to use that in the module build. On system level, if you are not able to deliver the new system type, then you could propose to customers to deliver the mature one. However, this means lower specifications for the customer with a risk of not acquiring a purchase order for the new system type.

2.2.4 Risks

ASML faces risks in terms of capability, capacity and commitment. Firstly, the capability of suppliers to produce the requested number of (sub)modules could be restricted due to for example capacity or sourcing issues. Increasing the plan is difficult due to the fact that the suppliers have supply constraints, especially within order lead time. Secondly, ASML's factory has a maximum capacity due to constraints related to for example work centers, tools, engineers, and inventory space. For ASML, the utilization of capacity is extremely important. This is a high risk for the business line and is connected with the capability of suppliers. Capability should not be lost and leveling of the factory is required. Furthermore, if a FASY start slot will not be used this will result in operational costs for the cabin, tools, and engineers.

Lastly, there is commitment risk at suppliers when components and (sub)modules are ordered. This means that if orders are purchased, ASML is committed to pull these orders on stock at some point in time. When orders are delayed, suppliers experience plan changes as well which is not desired. If orders are not called off and ASML wants to pull another sub (assembly) in, suppliers could be restricted in space capacity. Commitment to inventory could also lead to obsolescence at a certain point in time. As around 90% of the product costs are sourced from OEM suppliers and components and (sub)modules should be ordered a long time in advance, commitment costs are extremely important (ASML SCM, 2019).

As this research is focused on demand uncertainty, the risk of order commitment for components and (sub)modules are taken into account and mitigated with the buffer planning model. However, the capability and capacity risks are not taken into account as these are associated with supply uncertainty.

2.3 Buffering concept

In the S&OP, the highest level of tactical planning, effectively managing uncertainties is a challenge (Nahmias & Olsen, 2015). Manufacturing firms should be responsive and able to address uncertainties quickly in order to sustain competitive with other companies (Wazed, Ahmed, & Yusoff, 2009). Various techniques can be used to address these uncertainties in order to minimize the effects on customer service, it is important to understand which uncertainty to address with which technique.

Buffering is a common strategy for dealing with uncertainty. According to Nahmias & Olsen (2015), buffering is defined as "maintaining excess resources (inventory, capacity and/or time) to cover for fluctuations in supply or demand". According to literature review of (Wazed, Ahmed, & Yusoff, 2009) a distinction is made between buffering and dampening. Buffering refers to tangible techniques (e.g. safety stock), whereas dampening refers to intangible techniques (e.g. safety lead time). However, the definition of Nahmias & Olson (2015) is followed, where buffering encompasses inventory, capacity and time. Hopp & Spearman (2000) name this buffering concept as variability buffering law, which means that variability in a production system will be buffered by some combination of inventory, capacity, and time. The main idea of this buffering law is the concept "pay me now or pay me later", which indicate that if you do not pay to reduce variability, you will have to pay in different ways. These ways are lost throughput, wasted capacity, inflated cycle times, larger inventory levels, long lead times, and/or poor customer service (Hopp & Spearman, 2000). The buffering strategy is dependent on the business strategy and production environment.

According to Nahmias & Olsen (2015), operations analysts and managers commonly deal with the uncertainty type variation by developing strategies to mitigate this variation. Within all production environments variability in customer requirements and operations exist, and should be mitigated with the use of buffers. The buffer strategy determines the way how customer's demand will be met regardless of the variations across the supply chain, this means that buffers should be placed strategically in order to meet customer service requirements. The buffer strategy is dependent on the business strategy and production environment and it is important to identify where in the supply chain which flexibility is required and how to create that flexibility in order to respond to fluctuations in the market. A buffer strategy can consist of three different buffer mechanisms, namely safety stock (inventory buffer), safety lead time (time buffer), and safety capacity (capacity buffer). When the buffer strategy is inconsistent, inventory costs can be higher and customer service lower.

2.4 Buffering mechanisms

2.4.1 Decision buffering mechanism

As stated, a buffer strategy can consist of the buffer mechanisms safety stock, safety time and safety capacity. Hopp & Spearman (2000) defined three examples to clarify the buffering mechanism decision. The first example assumes a retailer which sells inexpensive ball pens and operates under unpredictable demand. This retailer is not able to buffer in time and capacity as customers buy ball pens somewhere else if out of stock. Therefore, the production system is a make-to-stock environment and an inventory buffer is required to stock ball pens. A second example is an emergency service, where there is variability in the requested demand for a fire engine or ambulance. It is not possible to buffer inventory (buildings or people that are in an emergency state). Furthermore, a time buffer is not applicable as this is the key performance measure of an emergency service. Therefore, the buffer that should be applied is capacity which can be seen in practice as the utilization of emergency transport is low. The last example is regarding organ transplants, where variability in demand and supply exists. A capacity buffer is ethically not allowed as the supply is determined by the number of donor deaths. Moreover, buffering with inventory is not applicable due to the short usability time of the organ. In conclusion, the buffering mechanisms should be time, which reflect reality as there are long waiting times for organ transplants.

Safety stock, inventory buffer, is suggested to apply when dealing with quantity uncertainty, and safety lead time when dealing with timing uncertainty within an MRP environment (Wazed, Ahmed, & Yusoff, 2009). Based on simulation experiments, Jacobs, Berry, Whybark & Vollmann (2011) states that safety stock is outperforming safety lead time in order to cope with uncertainty in demand and supply quantity. This means safety stock achieves a higher service level than safety lead time when average inventory is equal. The use of safety stocks outperforms safety lead time in terms of costs when demand variability is high and lead time variability low. The time buffer, safety lead time, is suggested to apply when dealing with timing uncertainty (Wazed, Ahmed, & Yusoff, 2009). Safety lead time outperforms safety stock when coping with uncertainty in demand and supply timing. This means safety lead time achieves a higher service level than safety stock. The lowest costs are obtained with safety lead times when both variabilities are high (Wazed, Ahmed, & Yusoff, 2009). High variability in lead time and demand strongly affect the optimal level of safety lead times and safety stocks.

Safety capacity and rescheduling are also common buffering techniques (Wazed, Ahmed, & Yusoff, 2009). Safety capacity is the ability to expand the production capacity by using for example mid-term capacity expansion investments, overtime production or additional production shifts. Rescheduling refer to delivery flexibility, which means that ability to modify the amount and/or moment of delivery such as the option to backlog (Esmaeilikia, et al., 2016).

In this research the focus will be on inventory buffer to cope with uncertainty in demand quantity and because the buffer planning model should be applicable in the current way of working, hedging the supply plan. When using safety time, the model is not applicable in practice. Safety capacity is not applicable due to the fact that this is more production oriented, whereas in this research the focus is mainly on the mid-term supply planning.

2.4.2 Inventory buffer

Four inventory functions can be distinguished, namely safety stock, anticipation stock, transit stock, and cycle stock (Jacobs, Berry, Whybark, & Vollmann, 2011). Transit stock is equivalent to pipeline stock; however, transit stock is more associated with transport environments. Whereas pipeline stock is a more generic term as this covers operational activities as well, e.g. assembly activities. In this thesis transit stock (inventory in-transit) and pipeline stock (pipeline inventory) are used interchangeably due the inconsistency in used references.

Safety and anticipation stock will be elaborated because these stocks are related to buffer management. Moreover, pipeline stock can also be used to buffer, as this is purchased inventory that is somewhere in the supply chain, but not yet at the final stage where this inventory is required to assemble the end item (Miller, 1979). Cycle stock is not applicable as the lot size is one instead of higher economic quantities.

The quantity of raw material stock is influenced by batching, variability, and obsolescence. Inventory is work-in-progress (WIP) when it is in one of the following states: queueing, processing, waiting for batch, moving, or waiting to match. Finished goods inventory occur or is used to be responsive to customers, batch production, forecast errors (released jobs without customer order), when products cannot be shipped earlier, and build-up inventory to meet peaks in demand (Hopp & Spearman, 2000).

Safety stock

Safety stock is defined as protection against irregularities or uncertainties in the demand or supply of items, i.e. in situations when the demand exceeds the forecast or when the time to resupply is longer than expected (Jacobs, Berry, Whybark, & Vollmann, 2011). The right safety stock guarantee that customer demand can be satisfied without backlog. The quantity of safety stock determines the extend of irregularity or uncertainty protection. However, a trade-off should be made between this protection and investment in safety stock. According Mönch, Uzsoy, & Fowler (2018), one of the inventory management problems in the semiconductor industry is about: “how to determine the levels and locations of safety stocks throughout the supply chain to maintain an appropriate trade-off between customer service and costs”. This problem also occurs in many other industries.

Anticipation stock

To deal with seasonal demand patterns, anticipation stock can be used which is also known as seasonal stock or pre-built stock. This stock is built in advance and depleted during demand peak periods, an example is shown in Figure 4. The stock contributes to reduce lost sales, costs for overtime, and opportunity costs for unused machines and tools (Fleischmann, Meyr, & Wagner, 2015). However, there will be risk due to forecast uncertainty and involved holding costs. The trade-off that should be made is between regular capacity level (option to investigate expansion costs), overtime capacity, and anticipation stock costs and risk. Customer service level could also be incorporated in the safety and anticipation stock trade-off. This means that customer service level objectives are translated to the required inventory quantity, i.e. x units of stock are required to achieve y percentage customer service level.

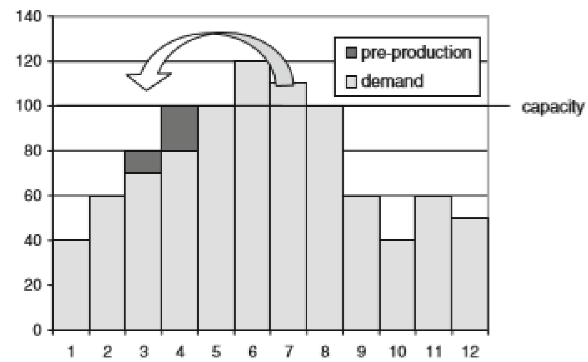
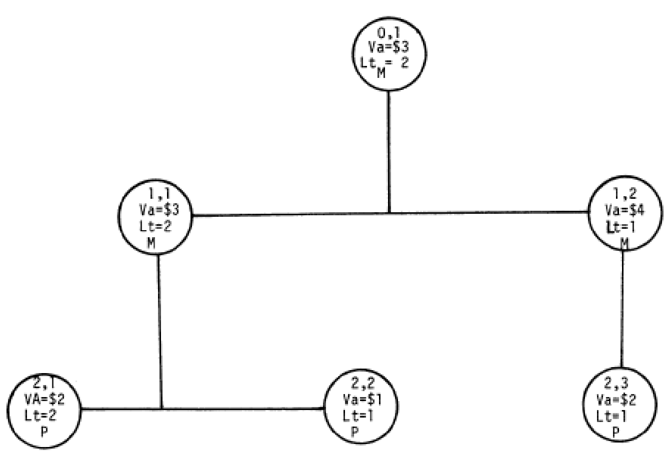


Figure 4. Example anticipation stock (Fleischmann, Meyr, & Wagner, 2015)

Pipeline stock

Pipeline stock is purchased inventory that is somewhere in the supply chain, but not yet at the final stage where this inventory is required to assemble the end item. This inventory type could be utilized to hedge against stockouts or undesirable backlogs, which is called pipeline safety stock (Miller, 1979). The word hedging is defined as "to do something to protect yourself against problems, especially against losing money" (Ofxord University Press, 2019). Hedging in the production environment is defined by Miller (1979) as "a master scheduling tactic which allows variations in the level of uncertainty and cost commitments over time to be taken into consideration in ordering decisions". Pipeline safety stock is created when the output of finished goods in the master production schedule is manipulated by hedging on future forecasted end item outputs. This results in lower inventory investments than on-hand stock investments.

Hedging can only be implemented when demand dependent logic in the product structure is used. This means that demand for items at lower levels in the product structure are derived from successor items at a higher level. An example of the product structure is shown in Figure 5.



The first set of numbers denotes the item number, e.g. level 0 item 1, level 1 item 1

Va = Value added at each stage of manufacture or purchase denoted by M or P (level 2 items are all purchased, level 1 items are submodules, and level 0 is the finished end product).

Lt = the procurement or manufacturing lead time

Figure 5. Example product structure (Miller, 1979)

Two phenomena have led to the hedging concept (Miller, 1979). The first phenomena is that the reliability of demand forecasts in the far future are lower than near future forecasts, which is visualized in Figure 6. This figure shows that in the future there are more forecasted orders. However, these will be consumed by customer orders when time to current date is decreasing, e.g. the point in time where the red vertical line is positioned (part above the curve are forecasted orders and below are customer orders). This means that there is more uncertainty for items that are in the bottom of the product structure because of the cumulative lead time. However, the decision to assemble the end item can be extended until the lead time of the assembly operation in advance of the required completion date. This decision can be made based on more a reliable demand estimation.

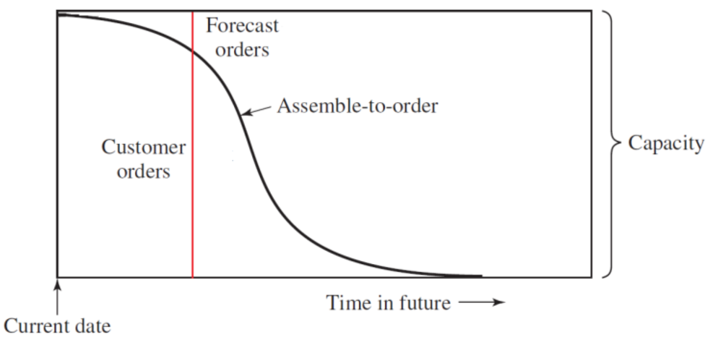


Figure 6. Forecasts consumed by customer orders over time (Jacobs, Berry, Whybark, & Vollmann, 2011)

The second phenomena is that commitments to suppliers and manufacturing costs are time-phased. A commitment profile could visualize this time-phased characteristic by showing the cumulative added value of component procurement commitments and costs related to manufacture modules and end items. An example of a commitment profile and different types are shown in Figure 7.

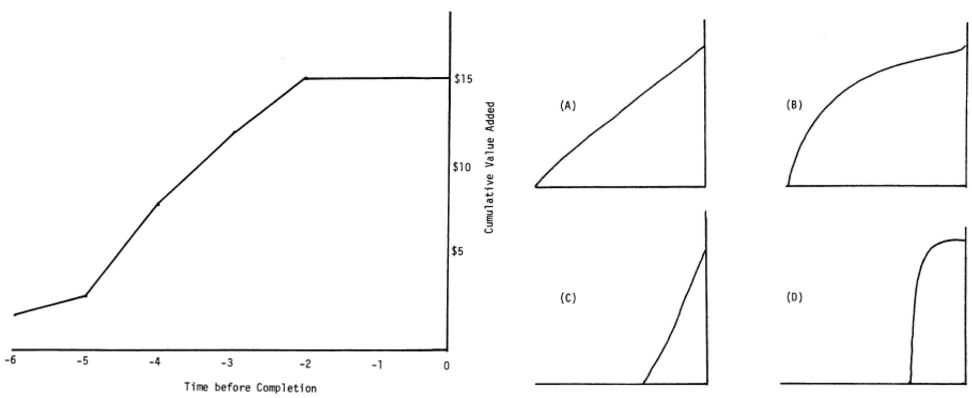


Figure 7. Commitment profile types (Miller, 1979)

When applying the hedging tactic without taking assembly operations into account, there are no inventory costs involved because the items are not physically in stock, but only on a purchase order. The risk of hedging is that the items are purchased and should be paid and pulled in at stock at a certain point in time, which is dependent on agreements between the company and supplier. Hedging is effective when the company is willing to adjust the master production schedule when time passes, and new information is available. When the commitment profile is shaped similar as (B), major risks are involved in the hedging decisions as the commitment value is rapidly increasing long in advance of completion. The hedging decisions are made by higher management level because of the major risks (Miller, 1979).

There are two forms of hedging, namely static and dynamic hedging. Static hedging focuses on a one-time hedge for a non-standard product for a specific customer or new product introduction. On the other hand, dynamic hedging is oriented on the standard product portfolio which are offered to all customers. For dynamic hedging, the company should identify how uncertainty varies over time. The hedges should be continuously made, reviewed and updated when it moves in time due to the fact that uncertainty and added value changes over time as well.

The concept of hedging is simple and logical; however, it is complex to manage and should be reviewed constantly. Priority control is a challenge because orders reflect the demand requirement for firm orders and safety stock in the pipeline. To deal with this challenge in the factory different coding can be used for firm orders and hedging. The constant change in number and position of hedges result in excessive number of rescheduling messages. As MRP logic uses dependent demand logic, all components in the product structure experience these rescheduling messages, which result in increased administrative costs. Therefore, hedging should only be used when products have high values and long lead times.

There are some advantages and disadvantages when adopting the hedging tactic (Bartezzaghi & Verganti, 1995). The advantages are that inventory costs and parts obsolescence is reduced as stock are at lower levels of the product structure, and uncertainty variability over time is tracked closely. The disadvantages are that there are unrealistic planned orders, component commonalities could not fully exploit, and higher organizational costs to control information.

2.5 Conclusion and insights

In this research the focus is on demand uncertainty. According to literature, demand uncertainty is positioned as environmental uncertainty with random variation as demand uncertainty is outside the manufacturing systems' boundary and beyond companies' control. ASML experiences demand uncertainty in terms of quantity due to frequent up- and downturns of the lithography market and customer requirements. Based on the coefficient of variation, it is observed that there is high variability in ASML's demand figures. The risk of order commitment of components and (sub)modules are high due to the fact that product costs are extremely high, 90% of the product costs are sourced from OEM suppliers, and supplier lead times are long. To address demand uncertainty and commitment risk buffering could be used.

A buffer strategy can consist of three different buffer mechanisms, namely inventory, time and capacity. The focus in this research will be on inventory buffer as the buffer planning model should be applicable in ASML's current way of working. There are four inventory functions, namely safety stock, anticipation, pipeline stock and cycle stock. The quantity of safety stock determines the extend of irregularity or uncertainty protection. Pipeline safety stock could be utilized to hedge against stockouts or undesirable backlogs. Hedging is a master scheduling tactic to decide on order commitment based on uncertainty and cost commitments. Two phenomena that have led to hedging concept are that a demand forecast in far future is less reliable and the commitment costs are increasing time-phased.

3. Formulation buffer planning model, workflow and IT

It is important to identify where in the supply chain which flexibility is required and how to create that flexibility in order to respond to fluctuations in the market. A buffer strategy could provide this flexibility and consists of a buffer planning model and a workflow to describe the position in the planning process and frequency of the model.

Companies that operate in an assemble-to-order (ATO) environment, assemble components into end items when customer demand confirmation is received. However, decisions about the procurement or production of these components should be made in advance. The manufacturing process consists of component procurement and/or production, and assembly. Operating policies in ATO systems are required to minimize the echelon costs with a service level constraint. These operating policies are defined as: “a combination of component replenishment/production policy and an inventory allocation policy of these components” (Atan, Ahmadi, Stegehuis, De Kok, & Adan, 2017). ATO models can operate with periodic review models and continuous review models, where a distinction is made between single end item and multiple end item models. As ASML offer a product portfolio with multiple system platforms and system type, the focus will be on multiple end item models with periodic review. In a multi-period setting components are ordered at the beginning of every period and demand arrives at the end of every period. The inventories and backlog at the end of a period are transferred to the subsequent period.

At ASML, there are different meetings and procedures where information is exchanged, and decisions are taken. As shown in Figure 3, there is a long-term market review and capacity planning on strategic planning level. On tactical planning level, there is a monthly S&OP meeting and S&OP deployment meeting. In the S&OP meeting the capabilities are determined how many systems the factory should to output per week. Whereas in the S&OP deployment meeting the master production schedule (MPS) is adjusted and new systems are planned in the new horizon. Finally, on operational planning level there is a weekly operation review start decision meeting. In this meeting, decisions are made regarding the FASY start dates of systems. These meeting procedures are described in more detail in the next chapter.

A buffer planning model and workflow results in managing uncertainty in a systematic manner and hence prepare ASML for dealing with future scenarios. The model requires information technology (IT) to be able to calculate the buffer levels with a specified recurrence. The input parameters of the model should be adjustable for each system type as the strategy for system types differ. Different factors can determine this strategy, namely market share ambition, cost reduction targets, growth strategy and expected demand growth with associated market volatility.

As stated in the company introduction, ASML’s revenue market share in 2017 in the semiconductor lithography market was 85.4 percent. This means that there still exists competition for specific lithography system types. For system types with high competition, a strategy could be to ensure a very high customer service level to avoid backorders or lost sales, which result in higher buffer levels. On the other hand, if the market share is extremely high, it could be a strategy to focus more on commitment and inventory cost reduction. Furthermore, the growth strategy of ASML is important to determine the number of outputs per system type. It can be the case that the strategy is to boost the output of a specific system type. The expected growth in customer demand is an important factor as this could be an indicator what the expected number of sales per period is. Moreover, the volatility that exist in the market is important to take into account when determining buffer levels as the volatility has direct impact on your customer service level. The decisions about both factors influence the service level requirement and costs. Therefore, it is important that a trade-off could be made between service level and costs.

In conclusion, the buffer planning model should be applicable for the ATO production environment and periodic review with multi-period characteristic of ASML. The model should be relatable to the current way of planning end items and should be connected to the current way of buffering, namely the application of hedging tactic. The model should be executed in an IT environment and programming language which could be supported and maintained. Furthermore, the model should be able to make a trade-off between service level and costs in order to use it for decision-support. Finally, a workflow should be determined how and when the buffer planning model should be applied. This workflow should be aligned with the different meetings and procedures of ASML.

4. Current buffer planning model and workflow

To cope with uncertainty in demand and supply, the hedging technique in the master plan is applied in tactical planning. This technique manipulates the output of finished goods in the master production schedule by hedging on future forecasted end item outputs in order to create pipeline safety stock. Pipeline stock is purchased inventory that is somewhere in the supply chain, but not yet at the final stage where this inventory is required to assemble the end item. When applying the hedging tactic without taking the assembly operations into account, there are no inventory costs involved because the items are not physically in stock, but only on a purchase order. The risk of hedging is that the items are purchased and should be paid and pulled in at stock at a certain point in time, which is dependent on agreements between the company and supplier. In this chapter the current buffer planning model and related workflow is described. Moreover, ASML's S&OP meeting structure and planning procedures are explained.

4.1 Buffer planning model

On end item level a hedge could be made, which is called "system buffer". This system buffer is a system with a reserved FASY start slot which is planned above the firm (expected) demand. All components and (sub)modules for this buffer are ordered based on MRP and addresses uncertainty in demand. The position of this system buffer (hedge) in the supply plan determines where the buffer is created, i.e. which components, (sub)modules are buffered. A customer order could be allocated to a system buffer when it is activated, this is a management decision. There is a tool to determine the system buffer levels. This tool generates a demand plan, and based on this plan MRP logic states when to replenish materials. The assumption is that lead times are deterministic. As demand is not stable mismatches between demand and supply will occur within lead time which are addressed by buffers. A large number of buffer strategies on raw material, semi-assembled, and end item level are determined to create a trade-off curve between customer service level and inventory value. One of these buffer strategies can be chosen. However, there are several drawbacks of this tool. Firstly, the assumed way of working in the tool does not reflect the reality of creating a plan. Next, it does not reflect reality, there is no option to evaluate scenarios with high and low demand variability as the model is based on average demand only. Furthermore, the tool is overdesigned, and the main parameter forecast accuracy is hard to control. Lastly, it is labor-intensive to obtain results from the tool and the buffer strategy does not reflect reality in terms of implementation. Therefore, the buffer levels are mainly determined based on mental models of employees. This means that the current tool is not related to a recurring process and rarely used in practice.

4.2 Buffer planning workflow

Figure 8 shows the process of how the system buffers are determined. System starts are planned for firm demand and buffer systems are added to cover upside demand. For some system types there is a buffer agreement within integral cycle time (= longest lead time + FASY cycle time), which means that if buffers are above this agreement, the buffer level is decreased. When there is no customer order allocated to a system buffer *x* weeks before FASY start, there is a meeting to decide on how to continue. This process is shown in Figure 9. It can be possible that system buffers are removed from the supply plan and canceled at supplier or can be pushed out to a different start week. In both situations extra upstream inventory is created at suppliers.

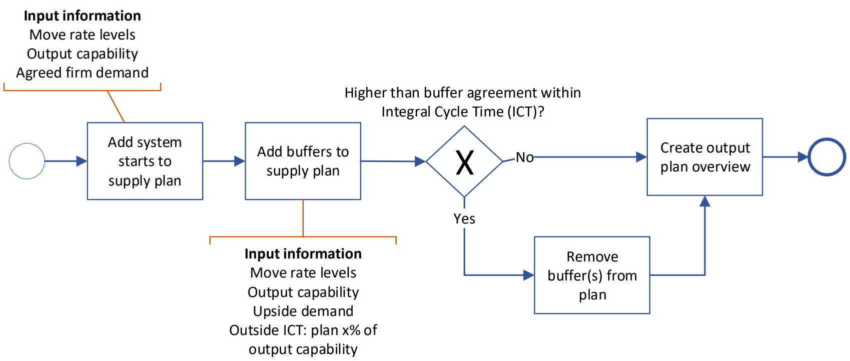


Figure 8. Process to determine system "WIP" and "supply chain" buffers

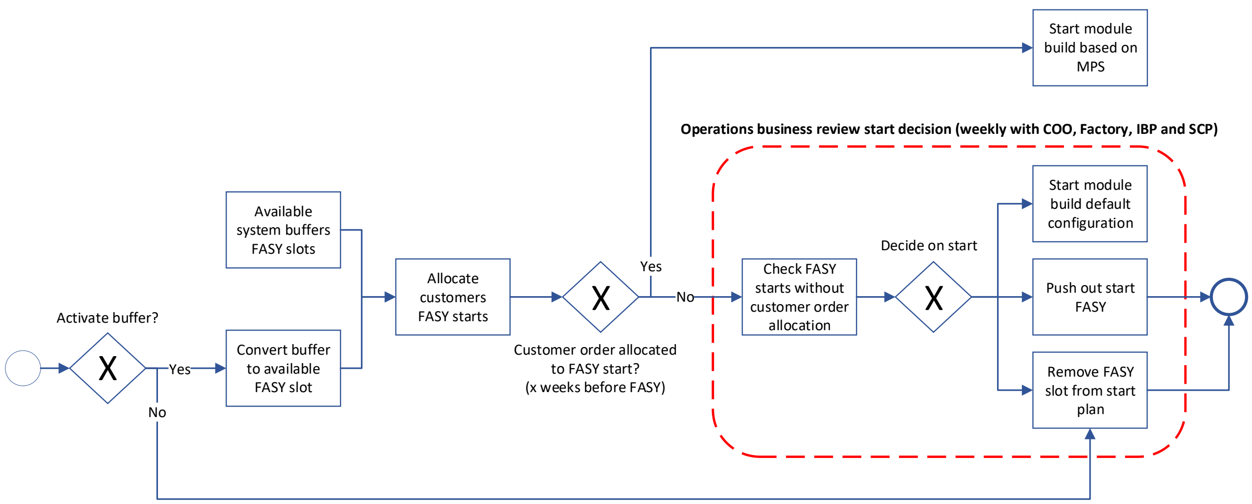


Figure 9. Process of "supply chain" buffer activation and customer order allocation

In addition to system buffers, "mix flexibility buffers" are planned, which is relatable to bill of material (BOM) planning. Mix flexibility buffers are material packages that are ordered on top of a planned system. This buffer can be used to switch between system type and/or configuration to provide flexibility in serving customers. For system type flexibility, this is a material package (BOM) that is defined in MRP and covers the delta materials to switch between two system types. Whereas for configuration flexibility, different modules are ordered to enable a switch in configuration. At a certain point in time a decision should be made because the flexibility is not required anymore, or the costs are too high. The flexibility of system types and configurations is important for ASML in order to maintain a high service level towards customers. The processes to determine the flexibility buffers are shown in Figure 10 and Figure 11.

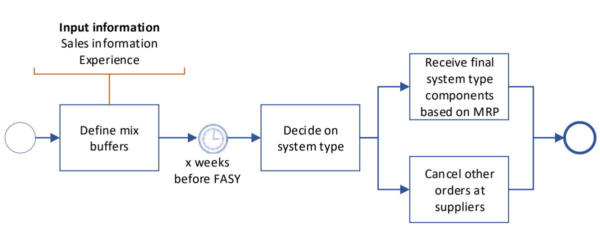


Figure 10. Process to determine flexibility mix buffer

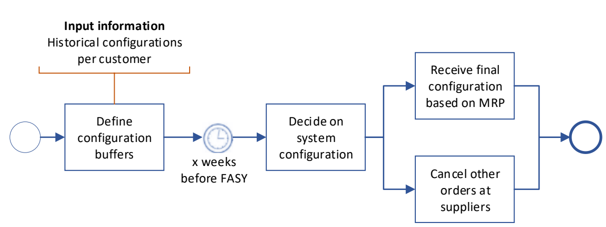


Figure 11. Process to determine flexibility configuration buffer

Currently, there are no module buffers regularly used with the intention to address uncertainty. Based on the start plan, the factory will start building modules. This means that modules and end items are not decoupled. Some modules can only start when there is a positive decision in the operations start meeting. When a system start is pushed out, then modules that are in stock are used for other FASY starts based on first come first serve. In the factory there is a limit of storage capacity to have modules on stock. There are a few reasons why modules on stock are increasing. Firstly, when there is no customer allocated to a system and it is agreed that the module build will be started. Secondly, when the supply plan is instable and systems are pushed out, the module build could already have been started. Secondly, some modules are required later in the FASY process than others. However, all modules should be ready the day before FASY start date, which means that some modules are held on stock until required in the FASY. This is a well-considered choice in order to be flexible to switch FASY starts of critical systems in the planning. Finally, delays could occur during build modules which might impact the FASY start date. This means that other modules are waiting longer on the stock position.

At ASML, there are different meetings and procedures where planning information is exchanged and planning decisions are taken. On strategical planning level, the long-term capability is determined by the board of directors. These decisions are based on business opportunities, market forecasts, and required investments of tooling and factories. Appendix II visualizes the monthly tactical planning process in more detail than presented in Figure 3. As can be seen in this overview, various departments are involved in the plan cycle. The swimming lane process overview provides insight in the timing and responsibilities of the processes. For each process step details are shown according to the IDEF method, which defines the input, resources, control and output of each process step.

In the S&OP meeting, the capabilities are reviewed, and adjustments are proposed if required. Decisions are made about how many systems should be built by taking into account factory capacities and supply chain capabilities. Moreover, information about service level targets and available budget is collected from the business line and operations. This means that the number of buffers (budget) are determined based on a trade-off curve or sometimes only an available budget is stated as trade-off curves are not available. The role of Integrated Business Planning (IBP) in this meeting is to connect with sales and marketing to develop an integrated business plan that is aligned with the financial objectives of ASML. These capability guidelines are used in the S&OP cycle to determine a supply plan including the implementation of the buffers. In the S&OP deployment meeting the details are discussed and an investigation is performed if the supply chain is able to meet this determined number of lithography system outputs. The final decisions are made by COO and business lines' Vice President. The capability is not the maximum capacity; however, the purpose of capability is to regulate the normal flow condition to determine what the sales and cost levels will be. When the supply plan is multiple times above capability, feedback is provided to the S&OP meeting in order to adjust the capability decisions, i.e. investment decisions regarding engineers, physical capacity and tools.

5. Buffer planning model design

Stochastic inventory models explain quantitative relationships between material requirements, material supply, inventory, and customer service level. For single-item single-location, state-of-the-art inventory models can be used to understand this relationship. However, when large networks of inventory locations are analyzed caution should be taken with current multi-item multi-echelon inventory systems (Wiers & De Kok, 2018). This caution is related to structural complexity, which means the number of items, resources, item-item, and item-resource relation. This result in many variables and constraints that should be taken into account by the mathematical model. As a result, it is complicated to find a solution that is optimal or even feasible. These relationships of complexity are taken into account when designing a buffer planning model. First, ASML's system planning is related to a known inventory policy. Hereafter, the approach to determine the number and place of the hedging positions is discussed. Third, the methods to calculate the base-stock levels are elaborated. Next, a simulation is performed to validate the base-stock policy. Finally, the optimal buffer levels are visualized by a trade-off curve with varying service level, customer order lead time, and demand uncertainty.

5.1 Relation ASML's lithography system planning to a known inventory policy

The supply plan of ASML is an overview of how many lithography system FASY starts are planned in each week of the planning horizon, this includes system buffers. Figure 12 shows an example of a supply plan focused on the FASY starts per week for lithography system type A with fictitious numbers, where the blue bars are regular system starts for expected (firm) demand and green bars are system buffer starts for upside demand.

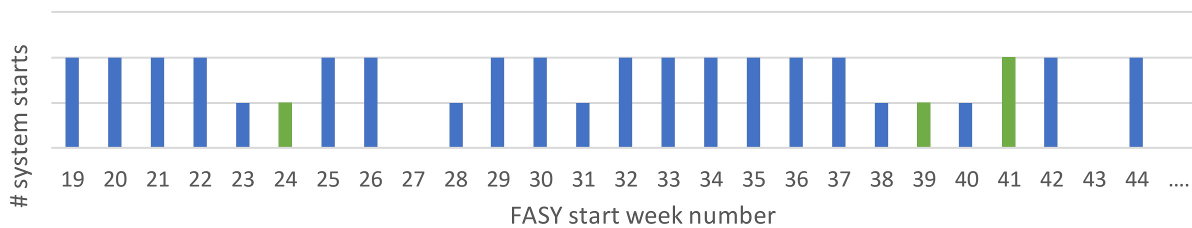


Figure 12. Example supply plan system type A with fictitious numbers

As the bars represent a number of planned systems in a specific week, this could be translated to a number of planned systems in the pipeline towards the FASY start. Therefore, ASML's system planning on tactical level could be seen as a serial system with a number of stages. Within ASML, MRP logic is used as product structure which enables the application of the hedging technique. Hedging is applied because system buffers are planned to cover variation in demand. The high value of lithography systems and long lead time to purchase and build them outweigh the increase in administrative costs due to applying the hedging tactic. In the serial system, the stages could be seen as hedging positions. At these hedging positions, the number of planned systems could be controlled in terms of regular and buffer system starts, i.e. a hedge could be made at that position to plan system buffers to cope with variability in demand. This control over the commitment costs of suppliers is extremely important for ASML as 90% of the product costs are sourced from OEM suppliers (ASML SCM, 2019). Hedging decisions should be made when a system start date enters the integral cycle time (longest lead time item plus assembly lead time). If planned system starts are outside the integral cycle time, no components will be ordered, but prepares suppliers for an output capability level. The horizon where buffer decisions should be taken is within the integral cycle time of a lithography system.

The assumptions in this serial system are: 1) demand arrive with a specified rate and occurs only at the last position J, 2) periodic review policy (for this research the review period equals one week) 3) supplier lead times and ASSY durations are deterministic, 4) suppliers have infinite supply (suppliers are able to deliver components and (sub)modules when ordering the lead time upfront), 5) no fixed ordering costs (no batching), 6) costs per item on-hand and in-transit are calculated per time unit, and 7) backordering costs are taken into account and occur when demand is not satisfied at the last hedging position.

Figure 13 shows a serial system with J number of hedging positions, where the positions j are numbered by following the flow of goods. Each hedging position represent a fixed point in the upstream supply chain from FASY start. This fixed point is associated with a time before FASY start, e.g. start ASSY, finished ASSY, order commitment to a specific part or (sub)module. The lead time (L'j) to a hedging position represents to what extend components and sub(modules) are ordered or modules are assembled. ASML is able to control the complete serial system as this represents ASML's supply plan.

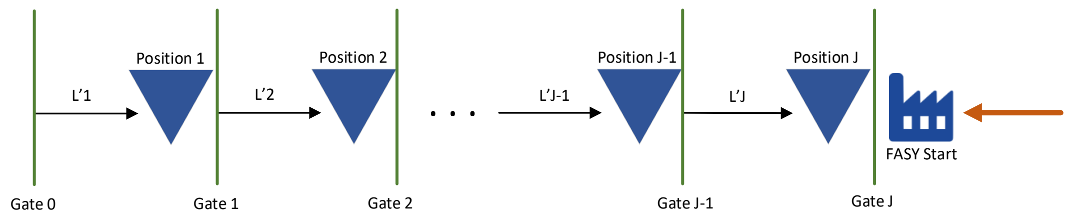


Figure 13. Serial system

Hedging position J faces external demand of customers, whereas the other hedging positions face demand from successor position j+1. The first hedging position orders from an external supplier, whereas hedging positions j>1 orders from predecessor position j–1. Orders that are placed will be L'j time units in-transit. Each position is allowed to have inventory or backorders. The inventory control policy that can be applied in a serial system is a local base-stock policy or echelon base-stock policy. These policies ensure that system buffers do not have to be canceled or postponed anymore in the supply plan as system buffer are positioned at different fixed hedging positions. At every point in time the inventory on-order position should be equal to the base-stock level. In the next two subsections, the local and echelon base-stock policy are elaborated following the notation and terminology applied in Atan & Jaarsveld (2019).

5.1.1 Local base-stock policy

The local base-stock policy has decentralized control where each hedging position monitors its own inventory position, i.e. place orders at predecessor and deliver orders from successor.

Remark: variables related to local base-stock policy are marked with a prime: ' (echelon not)

Local base-stock policy notation, where time t ≥ 0:

$I'_j(t) = local\ inventory\ at\ hedging\ position\ j$

$B'_j(t) = local\ backorders\ at\ hedging\ position\ j$

$IN'_j(t) = I'_j(t) - B'_j(t) = local\ net\ inventory\ at\ hedging\ position\ j$

$IO_j(t) = IT_j(t) - B'_{j-1}(t) = inventory\ on\ order\ at\ hedging\ position\ j$

$$IOP'_j(t) = IN'_j(t) + IO_j(t) = local\ inventory - order\ position\ at\ hedging\ position\ j$$

$$IT_j(t) = inventory\ in\ transit\ to\ hedging\ position\ j\ (\#units\ in\ j\ 's\ supply\ system)$$

$$ITP'_j(t) = IN'_j(t) + IT_j(t) = local\ inventory - transit\ position\ at\ hedging\ position\ j$$

$$s'_j = IOP'_j(t) = local\ base - stock\ level$$

$$h'_j = \sum_{i=1}^{j} h_i = \ local\ inventory\ holding\ cost\ at\ hedging\ position\ j\ (real\ value\ in\ position\ j)$$

As backorders can occur at all hedging positions except hedging position j = 1, a distinction should be made between the inventory-order and inventory transit position. Figure 14 shows an example of the event logic when a demand will arrive at the last hedging position. As can be seen, the demand is cascaded upstream through the system to the first hedging position. In the starting situation as well as in the end situation, the local inventory-order positions of all hedging positions are equal to the corresponding local base-stock levels.

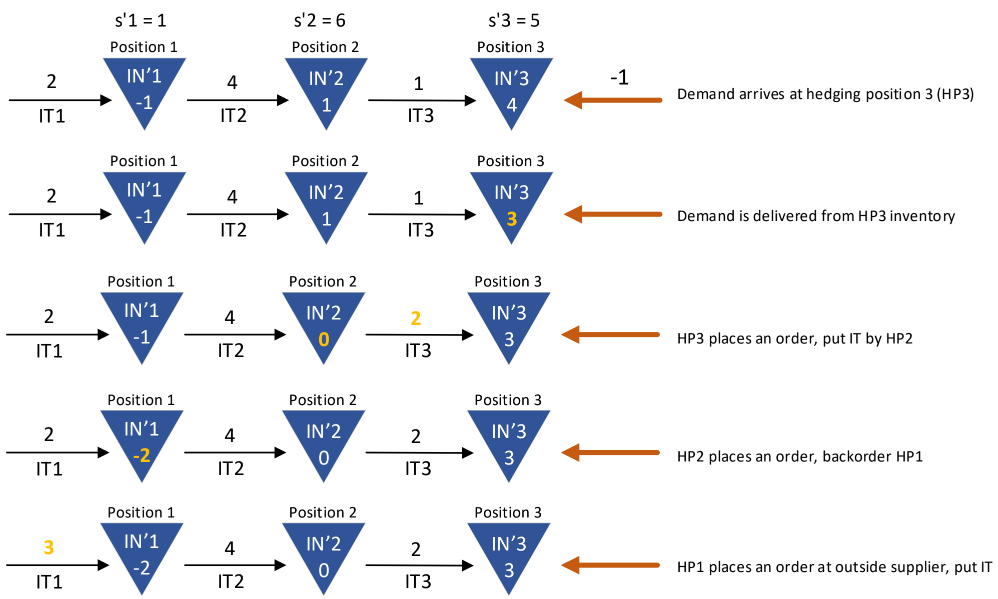


Figure 14. Example standard serial system logic (numbers on arrow = IT and numbers in rectangle = IN'j)

5.1.2 Echelon base-stock policy

In the local base-stock policy, each hedging position monitors its own position. However, the echelon base-stock policy has centralized control and therefore information from downstream hedging positions can be used. Therefore, an echelon is defined as a hedging position including the pipeline to that positions and all hedging positions with pipelines downstream.

Echelon base-stock policy notation, where time t ≥ 0:

$$I_i(t) = I'_j(t) + \sum_{i>j} [IT_i\ (t) + I'_i(t)]\ =\ echelon\ inventory\ at\ hedging\ position\ j$$

$$B(t) = B'_J = system\ backorders$$

$$IN_j(t) = I_j(t) - B(t) = echelon\ net\ inventory\ at\ hedging\ position\ j$$

$$IO_j(t) = IT_j(t) - B'_{j-1}(t) = inventory\ on\ order\ at\ hedging\ position\ j$$

$IOP_j(t) = IN_j(t) + IO_j(t) = echelon\ inventory - order\ position\ at\ hedging\ position\ j$

$IT_j(t) = inventory\ in\ transit\ to\ hedging\ position\ j\ (\#units\ in\ j\ 's\ supply\ system)$

$ITP_j(t) = IN_j(t) + IT_j(t) = echelon\ inventory - transit\ position\ at\ hedging\ position\ j$

$s_j = IOP_j(t) = echelon\ base - stock\ level$

$h_j = h'_j - h'_{j-1} = echelon\ inventory\ holding\ cost\ at\ hedging\ position\ j\ (added\ value\ by\ j)$

Note that the IOj and ITj are equivalent to the local base-stock policy notation. In order to keep IOPj equal to sj, hedging position j will order when the IOPj falls below the echelon base-stock level sj. The echelon base-stock policy is equivalent to the local base-stock policy and can be translated by using the following principle:

- from local to echelon: $s_j = \sum_{i=j}^{J} s'j$
- from echelon to local (for nonincreasing sj): $s'_j = s_j - s_{j+1}$

To show the equivalence of the local and echelon base-stock policy, Figure 15 shows a serial system with local parameters.

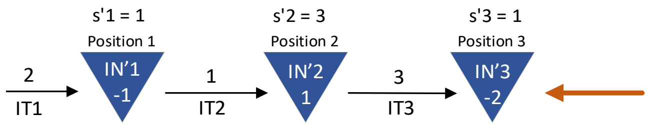


Figure 15. Example serial system local parameters

If the local parameters are translated to echelon parameters, the following numbers are obtained:

- Ij = physical inventory downstream of j: I3 = 0, I2 = 4, I1 = 5 and B (t) = 2
- INj = physical inventory downstream of j minus BJ: IN3 = -2, IN2 = 2, IN1 = 3
- ITPj = all inventory downstream from j plus IT to j minus BJ: ITP3 = 1, ITP2 = 3, ITP1 = 5
- IOPj = all inventory downstream from j plus IOj minus BJ IOP3 = 1, IOP2 = 4, IOP1 = 5

The difference between ITP2 and IOP2 is due to the fact that HP1 has a backorder of one. This backorder is not added in the ITP as it is not in transit yet. However, it is on order and therefore taken into account in the IOP.

Based on the IOPj, the echelon base-stock levels are: s1 = 5, s2 = 4, s3 = 1. To verify this, the translation formula (sj is nonincreasing): *s'j = sj – sj+1* is used. s'1 = 5 – 4 = 1, s'2 = 4 – 1 = 3, and s'3 = 1 – 0 = 1. These local base-stock levels are equivalent to the levels in Figure 15.

5.2 Determine number and place of hedging positions

The product structure of ASML is a demand dependent logic, which means that demand for items at lower levels in the product structure are derived from successor items at a higher level. To determine the number and place of the hedging positions different aspects should be taken into account. First, the product structure and supply-demand network topology are discussed. Hereafter, the phenomena time-phased commitment to suppliers and manufacturing costs are explained and visualized. Finally, other factors are discussed that can influence the decision about hedging positions.

5.2.1 Supply-demand network topology and product structure

To build one system, thousands of customized components are required which are sourced through 100 supply chains of 10-tier deep (ASML SCM, 2019). This results in an extremely complex product structure. ASML's supply-demand network topology is an assembly system. In an assembly system, there is one final end item which is produced by preceding assembly activities and procurement activities in order to source components for these assembly activities. A simplified assembly system is shown in Figure 16, where the vertical lines between numbers represent an assembly operation. The assembly operation of 1 and 2 with a duration of 1.5 time units could represent the assembly operation of a module by integrating two sourced items (components or submodules). Whereas, the assembly operation of 3 and 4 with a duration of one time unit could represent the final assembly operation of a system by integrating one sourced module and one assembled module.

Rosling (1989) has proven that an assembly system is equivalent to a pure serial system. This proof is based on the idea that the assembly system in the long-run balance can be interpreted as a serial system. Thus, an assembly system can be translated to an equivalent serial system, note that these two systems are only equivalent in echelon terms. This means that all parameters (including arrival rate and backorder costs) are equal except the local inventory holding costs and local lead times. Figure 16 shows an assembly system with translation to equivalent serial system. The positions in the serial system derived from the assembly system represent when ordering decisions should be made for items and it creates the timing of assembly starts. These serial system's positions could be used as hedging positions. However, a large extent of hedging positions are possible when the product structure complex. Therefore, additional aspects should be taken into account to decide on the best position, which will be discussed in the next subsections.

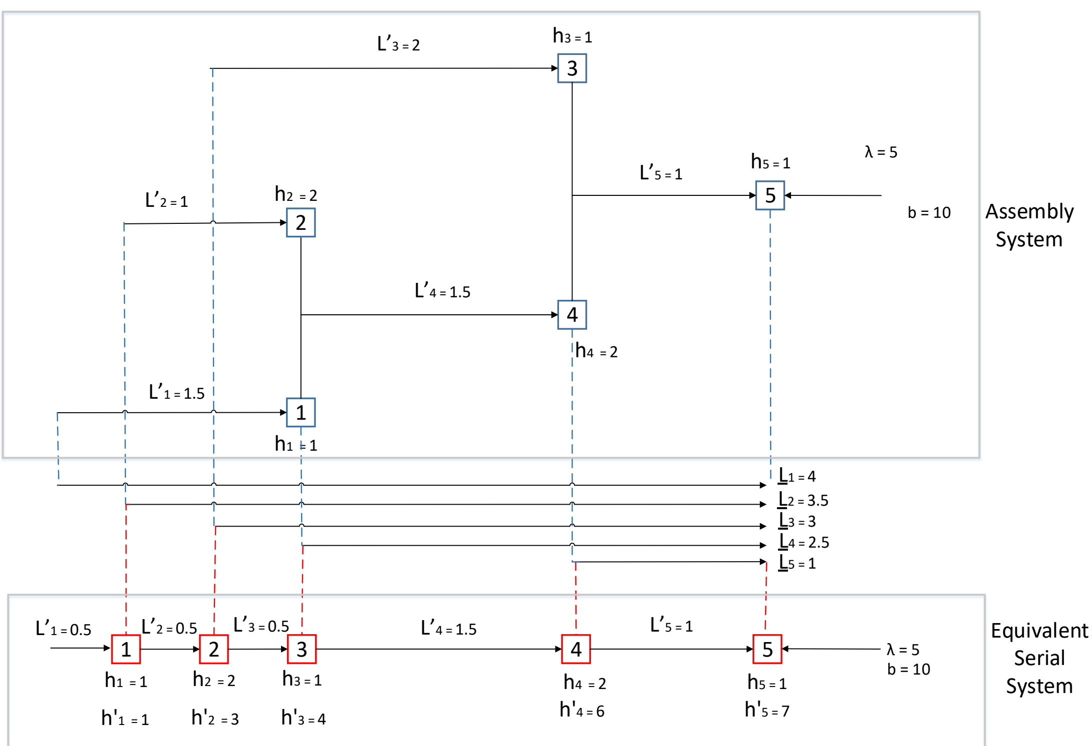


Figure 16. Example of an assembly and equivalent serial system ($\underline{L}j$ = forward echelon lead time)

5.2.2 Cumulative value curve

As hedging is based on the phenomena that commitments to suppliers and manufacturing costs are time-phased, i.e. the commitment value curve increases over time. Control over the commitment costs are extremely important for ASML as 90% of the product costs are sourced from OEM suppliers (ASML SCM, 2019). The commitment value curve visualizes the cumulative added value of component procurement commitments and costs related to manufacturing activities. Figure 17 shows an example of a cumulative value curve of a lithography system with fictitious lead times and costs.

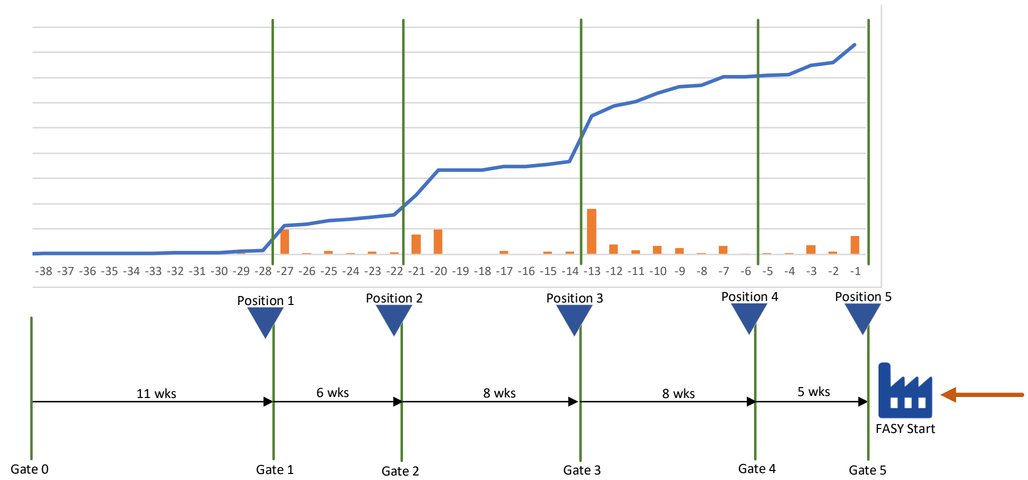


Figure 17. Example lithography system fictious cumulative value curve: commitment and manufacturing costs

Based on this cumulative value curve, the large jumps in value could be identified. In this example these jumps are 27, 21, and 13 weeks before FASY start. Before these large value jumps, hedging positions could be placed in order to create a buffer before a high value should be committed. Furthermore, due to the assumption that buffering is not possible in-between assembly operations, a hedging position is placed before ASSY. The ASSY lead time in this example is 5 weeks. This hedging position ensures that buffering is only possible before or after ASSY. All hedging positions corresponds to specific positions when the assembly system for this product structure is created and translated to a serial system. For example, the item(s) that are ordered 27 weeks before FASY start have an echelon forward lead time of 27 weeks. The remaining positions in the translated serial system are eliminated in the serial system.

5.2.3 Supplier decoupling points and customer order lead time

The other factors that could influence the decision about the number and place of hedging positions are decoupling points of suppliers and customer order lead time. The first factor of decoupling points or intermediate stocking points of suppliers could be important to take into account in order to create alignment between the buffer positions of ASML and their suppliers. The intermediate stocking points of suppliers are real positions where intermediate or final buffers can be positioned. When decisions are made without taking this into account, the disadvantage is that the buffer position defined by ASML can mismatch with the exact buffer positions in practice at suppliers. This is especially important for the suppliers associated to the large jumps in the cumulative value curve.

The second factor that influence the place of the hedging positions is customer order lead time. The customer order lead time determines the decoupling point between the forecast-driven supply chain and customer-order-driven supply chain. The last hedging position is always just before the customer order lead time upstream in the supply chain. This is visualized with an example in Figure 18. As can be seen, the last hedging position moves upstream in the supply chain. Furthermore, the hedging positions which are within the customer order lead time vanishes as these are not required. As lead times are deterministic, the last hedging position represents the point where customer orders arrive. When these orders could be delivered directly from "stock" (ASML or supplier at that point in time), the order can be delivered on time as the activities after this hedging position are deterministic. This concept is in line with the definition of De Kok & Fransoo (2003) where CODP, the last hedging position, is the distinction between customer order-driven and forecast-driven parts in the supply chain for a specific product market combination. At the CODP, inventory is held based on forecasts due to the fact that the demand is unknown, or partly unknown when the material orders are released. There is no inventory for items downstream of the CODP because the demand is known when orders are released. The challenge is to determine the base-stock levels at each hedging position to achieve a service level target. This will be explored in the next section 5.3.

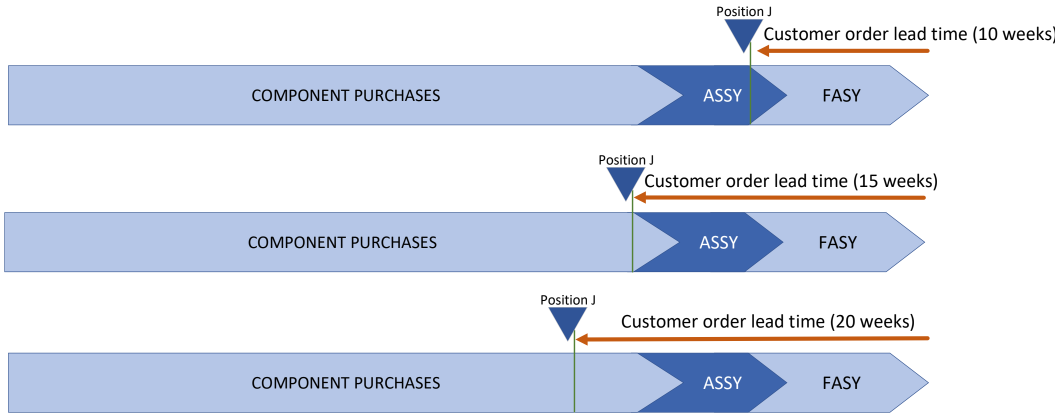


Figure 18. Example with fictitious lead times of the impact of customer order lead time on hedging positions

5.3 Finding near-optimal base-stock levels

After deciding the number and place of the hedging positions, the base-stock levels should be determined in order to achieve a service level target. Clark & Scarf (1960) started with the research on models for multi-echelon inventory systems. In this research an optimal inventory control policy for a N-echelon serial system is determined. However, this procedure is impractical to use. Based on this preliminary research, two analytical methods were developed to calculate the near-optimal base-stock levels for a N-echelon serial system. In this section, the Shang & Song heuristic is explained. Hereafter, the synchronized base-stock (SBS) policies to control general inventory systems and Diks & De Kok algorithm are discussed.

5.3.1 Shang & Song heuristic

Shang & Song (2003) developed a heuristic to find near-optimal base-stock levels that are required parameters for the base-stock policy to control assembly and serial systems. This heuristic can be applied for assembly and serial systems due to their equivalence that is proven by Rosling (1989) described in section 5.2.1. The main focus is on continuous-review and compound-Poisson-demand systems; however, for period-review models with independent and identically distributed (i.i.d.) demands the results are applicable. Based on extensive numerical experiments, the average relative cost error of the heuristic compared to optimal is 0.24% with a maximum error of 1.5% (Shang & Song, 2003).

This heuristic calculates an upper (sj^u) and lower (sj^l) echelon base-stock level bound, subsequently the near-optimal echelon base-stock level is approximated by the average of the two bounds. The heuristic uses the fact that if there are two positions (j-1 and j) where the local inventory holding costs are equal, position j-1 will immediately put items in transit to position j. This is logical as there are no costs involved when transferring inventory from position j-1 to j and the inventory can be closer positioned to the customer. When two or more positions have this characteristic, the positions can be merged to a single-position system. Figure 19 shows an example of a serial system with equivalent merged single-position system. The backorder costs are defined by 'b'.

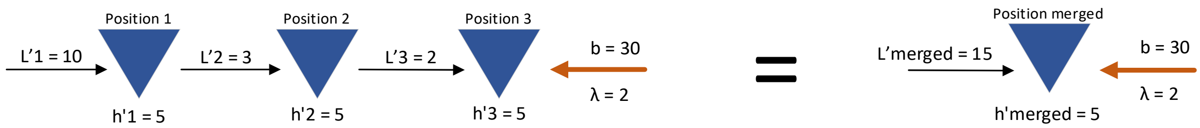


Figure 19. Serial system and equivalent single-position system (only difference between costs for inventory IT)

The optimal base-stock level of a single-position system should minimize the variable total cost function C(sJ) [2], which excludes the costs for inventory IT as these are constant.

$$C(s_J) = E[h'_J I_J + bB_J] = h'_J E[(s_J - \widetilde{D}_J)^+] + bE[(\widetilde{D}_J - s_J)^+] \quad [2]$$

The demand $\widetilde{D}_J$ is distributed according to a specified distribution with corresponding parameters. The optimal echelon base-stock level for a single-position system can be calculated by using a Newsvendor equation [3].

$$s_J = F_{\widetilde{D}j}^{-1}\left(\frac{b}{b+h'_J}\right) \quad [3]$$

The optimal echelon base-stock level s_J is the smallest s with the constraint $P(\widetilde{D}j \leq s) \geq \frac{b}{b+h'_J}$.

However, in practice the inventory holding costs will not be equal at all positions. Therefore, Shang & Song (2003) developed a heuristic to approximate the base-stock levels by bounding the serial system. To obtain the upper bound of the first position a serial system, the local holding costs h'_j for all j are adjusted to the holding costs of the first position, namely $h'_1 = h_1$ (local = echelon due to the fact that it is position 1). By adjusting the local holding costs to the lowest local holding costs in the system, the overall system holding costs decrease and thus a higher echelon base-stock level is calculated. To obtain the lower bound of the first position in a serial system, the local holding costs h'_j for all j are adjusted to the local holding costs of last position J, namely $h'_J = \sum_{j=1}^{J} h_j$. By adjusting the local holding costs to the highest local holding costs in the system, the overall system holding costs increase and thus a lower echelon base-stock level is calculated.

The upper and lower echelon base-stock levels for each position can be calculated with equation [4] and [5]. These formulas can be explained by the following logic. The upper and lower bounds of a serial system can only be calculated for the first position in a serial system. However, in order to calculate the bounds for the other positions, the number of positions in the serial system should be reduced in order to change the first position. The positions that should be excluded in order to determine the upper and low bound of position j in the original serial systems are all upstream positions from j. This means that the number of positions in the adjusted serial system are given by J + 1 – j. Figure 20 shows an example of creating the different serial systems in order to calculate the upper and lower bound of each position, which are systems Zi and Zi' respectively. The near-optimal echelon base-stock level is the average of the two bounds [6]. If the average of the two bounds is not an integer, the number can be rounded up or down to an integer value. Based on numerical experiments, a slightly better approximation is achieved when rounding down; therefore, this procedure will be followed. This heuristic is implemented in Excel and validated with the solutions of Shang & Song (2003) for specified scenarios with Poisson demand.

$$sj^{u} = F_{\widetilde{D}j}^{-1}\left(\frac{b^{Y}}{b^{Y}+h1^{Y}}\right) = F_{\widetilde{D}j}^{-1}\left(\frac{b+\sum_{i=1}^{j-1} hi}{b+\sum_{i=1}^{j-1} hi+hj}\right) \quad [4]$$

$$sj^{l} = F_{\widetilde{D}j}^{-1}\left(\frac{b^{Y}}{b^{Y}+\sum_{i=1}^{J^{Y}} hi^{Y}}\right) = F_{\widetilde{D}j}^{-1}\left(\frac{b+\sum_{i=1}^{j-1} hi}{b+\sum_{i=1}^{j-1} hi+\sum_{i=j}^{J} hi}\right) \quad [5]$$

$$F_{\widetilde{D}j}^{-1} = inverse\ cumulative\ distribution\ function\ of\ random\ variable\ Dj$$

$$\widetilde{D}j = \sum_{i=j}^{J} Di \sim Distribution\ (distribution\ parameter * \sum_{i=j}^{J} L'i)$$

$$sj^{a} = \frac{(sj^{u}+sj^{l})}{2} \quad [6]$$

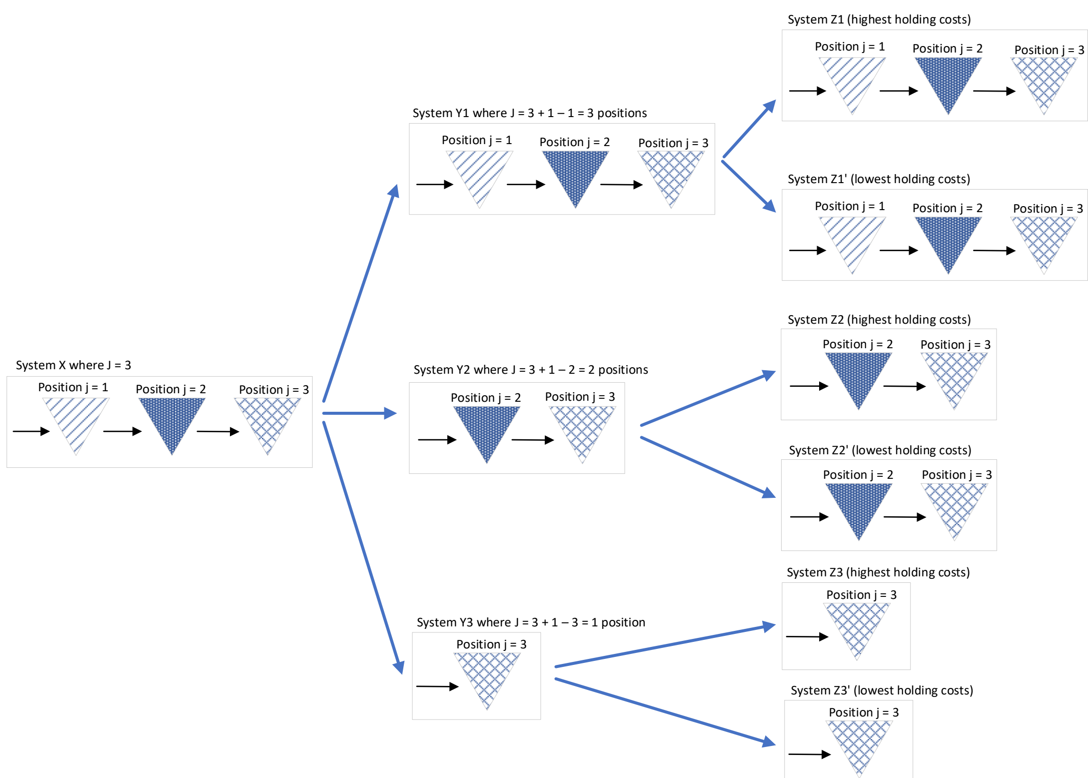
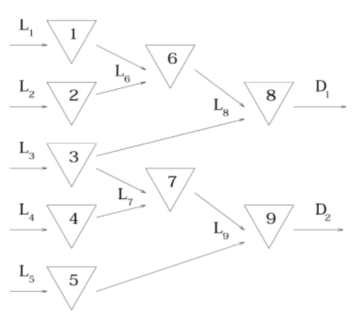


Figure 20. Example creating serial systems to calculate upper and lower bound for each position

5.3.2 Synchronized Base Stock policies and Diks & De Kok algorithm

The synchronized base stock-policies (SBS) are a class of strategies to control general inventory systems including multi-item multi-echelon systems. Diks & De Kok (1999) analyzes a divergent N-echelon inventory system with the allowance of inventory at each stockpoint. A stockpoint is equal to a hedging position. The assumptions for this system are equal to those stated in section 5.1. SBS policies consist of on the one hand creating a structure and apply base-stock policies with echelon base-stock levels. First, the method of De Kok & Visschers (1999) to create a divergent N-echelon inventory system is explained. Hereafter, the algorithm of Diks & De Kok (1999) that determine the near-optimal echelon base-stock levels for minimal long-run average cost is described.

De Kok & Visschers (1999) developed a method to decompose the assembly system into divergent a multi-echelon system where the pre-allocation of common components of end items is the key idea. The decomposition method is based on the equivalence of the assembly system and serial system researched by Rosling (1989) and described in section 5.2.1. The first step in the method is to create for every end item in the assembly network the equivalent serial system by using the cumulative lead times (forward echelon lead times). For convergent inventory systems, the SBS polices are optimal as these are equivalent to the transformation from assembly to serial system by Rosling (1989). Hereafter, the difference between the assembly system and multiple serial systems should be minimized. Therefore, the hedging positions of common components used in end items should combined in order to create the portfolio effect, i.e. demand fluctuations for end items decrease which smoothen the demand for common components. In a serial system, the decision to allocate is made when ordering components. Due to this decomposition method, the allocation decision is postponed until the serial inventory system should order an order component that should be assembled with the common component. In Figure 21, Figure 22 and Figure 23 an example of the decomposition method is shown. As can be seen, the assembly system is decomposed in individual serial systems for each end item. Hereafter, the common component at hedging position 3 is combined. This result in a three weeks postponed decision of allocation as this is the first moment that end item 2 requires an assembly operation with component 4. If the divergent system in Figure 23 is used to control the system, the order release decisions in Figure 21 are always feasible. However, the advantage of the divergent system is that decision nodes are created. In pure base-stock policies there is a base-stock level defined per single item. However, in SBS policies multiple base-stock levels, each associated with a decision node, can be used for a single item. This implies that in the SBS policies more degrees of freedom are provided, which enables the allowance of a combination of target customer service levels (De Kok & Fransoo, 2003). This can be seen in the example as well, in Figure 21 there are 9 decision nodes, whereas there are 10 decision nodes in Figure 23.



L1 = 4
L2 = 2
L3 = 7,
L4 = 2
L5 = 2
L6 = 1
L7 = 5
L8 = 3
L9 = 1

Figure 21. Example assembly system with two end items (De Kok & Visschers, 1999)

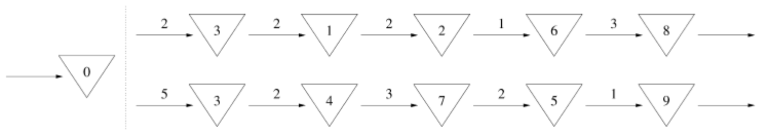


Figure 22. Serial system for each end item in Figure 21 (equal to Rosling (1989)) (De Kok & Visschers, 1999)

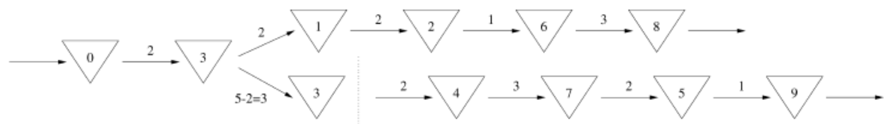


Figure 23. Divergent multi-echelon system after decomposition method (De Kok & Visschers, 1999)

To determine the near-optimal echelon base-stock levels and allocation fraction for a divergent N-echelon inventory system, an algorithm is developed by Diks & De Kok (1999) that minimizes long-run average total costs. The optimal allocation functions are approximated with two linear allocation functions, equation [7] and [8]. Below the notation is explained and within brackets the equivalent notation to the echelon base-stock policy described in previous section.

$S_i = order - up - to - level \left(= echelon\ base - stock\ level\ s_j\right)$

$h_i = additional\ holding\ costs\ due\ to\ value\ added\ in\ hedging\ position\ i\ (= hj)$

$p_i = penalty\ costs\ for\ each\ backlogged\ product\ at\ end\ hedging\ positions\ i\ (= BJ)$

$\alpha_k^i = non - stock\ out\ probability\ (ready\ rate)\ of\ end\ hedging\ position\ k\ in\ divergent\ echelon$

$system\ in\ which\ the\ most\ upstream\ hedging\ position\ i\ uses\ order - up - to\ policy\ with\ S_i$

$U_i = set\ of\ hedging\ positions\ on\ path\ from\ supplier\ to\ hedging\ position\ i$

$V_i = all\ hedging\ positions\ which\ are\ supplied\ by\ i$

$q_i = allcation\ fraction\ to\ hedging\ position\ i$

The echelon base-stock level of an end hedging position i (Si) satisfies equation [7].

$$h_i - (h_i + \sum_{j \in U_i} h_j + p_i)\ (1 - \alpha_i^i\ (S_i)) = 0 \qquad [7]$$

The echelon base-stock level of an intermediate hedging position i (Si) and its downstream allocation fraction satisfy equation [8].

$$h_i - \sum_{i_{n-1} \in V_i} q_{i_{n-1}} [h_{i_{n-1}} + \sum_{i_{n-2} \in V_{i_{n-1}}} q_{i_{n-2}} [\ \ldots + \sum_{i_1 \in V_{i2}} q_{i_1} [\ h_{i_1} - (h_{i_1} + \sum_{j \in U_{i1}} h_j + p_{i_1}) * \ (1 - \alpha_i^i (S_i))] \ldots\ ]] = \ 0 \qquad [8]$$

$$\alpha_k^i = \frac{\sum_{j \in U_i} h_j + p_k}{h_k + \sum_{j \in U_k} h_j + p_k} \qquad [9]$$

The algorithm starts to determine at each end hedging position (low-level code = 1) the echelon base-stock level in order to that the newsboy-styled equation [9]. Hereafter, the echelon base-stock level and allocation fractions of each hedging position with the next low-level code are determined such that equation [8] is satisfied. The echelon base-stock level is adapted until the sum of the allocation fraction is close to 1. Next, an adaption procedure should be performed if echelon base-stock level of hedging position i is lower than the sum of echelon base-stock levels of its successors. The echelon base-stock levels from most upstream to most downstream hedging position should be nonincreasing. The detailed steps of the algorithm can be found in Diks and De Kok (1999).

In conclusion, the algorithm determines the near-optimal control parameters of a N-echelon inventory system, namely the echelon base-stock levels and allocation fractions. This algorithm is applicable for serial and assembly systems and assumes continuous time. The algorithm is implemented by prof. dr. A.G. de Kok from Eindhoven University of Technology in a tool, named supply chain optimization planning engine (ChainScope). In the tool a supply chain model can be build which can be evaluated with the current parameters and optimized to a desired state. In this tool, the decision node structure of De Kok & Visschers (1999) as described above can be created. The optimization function is able to determine the optimal allocation of inventory capital in the supply chain. This result in situations where in the current state the amount of inventory captital invested is equal tothe desired state; however, in the desired state the service level is higher. The input requirements for the optimization function is a supply chain network with associated customer order lead time, review period, demand (μ and σ), lead times, added inventory values. The average demand over a time period is equal to the μ. The standard deviation of the demand, σ_D, represent the uncertainty in a demand sample and can be calculated with equation [10].

$$\sigma_D = \sqrt{\sum_{i=1}^{n} \frac{(D_i - \mu_D)^2}{n-1}} \qquad [10]$$

In the algorithm, based on the first two moments, the Gamma distribution is fitted, which is a continuous distribution with the property of positive numbers. Furthermore, the target service level should be given. The target service level could be a target fill rate or a target ready rate, the definitions are explained in the next chapter. From now on, “Diks & De Kok algorithm” refers to the algorithm developed by Diks & De Kok (1999), which is implemented in ChainScope.

5.4 Simulation echelon base-stock policy

In this section, the simulation approach including performance measures are discussed first. Hereafter, the echelon base-stock policy is simulated with the echelon base-stock levels of the two analytical methods, Shang & Song heuristic and Diks & De Kok algorithm, in order to evaluate and validate. Third, the practical situation is simulated to validate the applicability. Finally, a simulation is performed to evaluate the current situation.

5.4.1 Simulation approach

A general simulation approach is required in order to evaluate the two analytical methods, validate the echelon base-stock policy concept, and evaluate the buffer planning at ASML. Figure 24 shows the overview of the simulation approach. This approach will be discussed step by step. The formulas of the serial system described in section 5.1 are implemented and coded in Excel VBA to develop the simulation model.

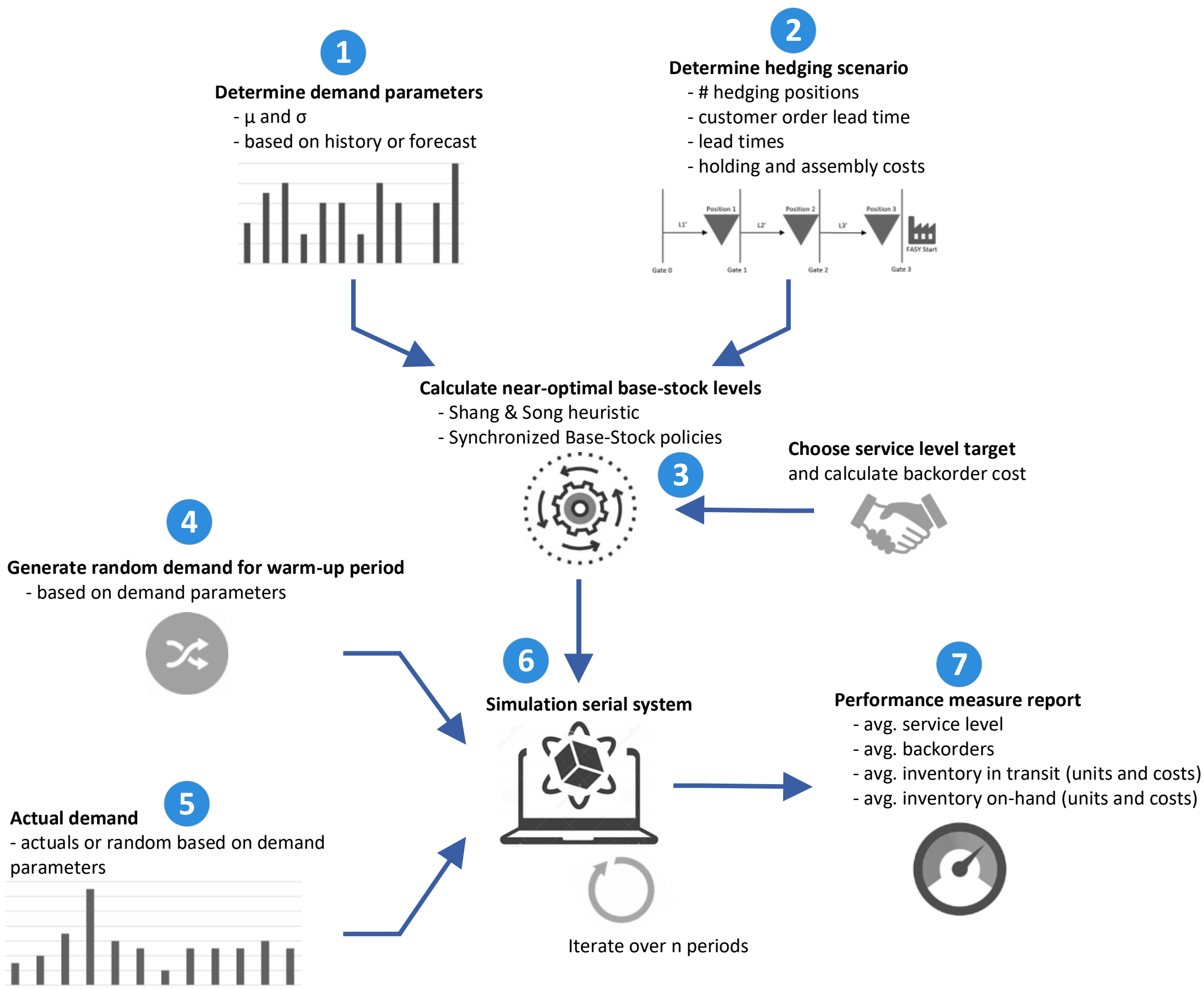


Figure 24. Simulation approach

Step 1: determine demand parameters

In the first step, the expected demand parameters for the upcoming time horizon for which buffers are required should be determined. Stochastic demand could be distributed differently. For example. the advantages of a Normal or Gamma distributed parameters μ and σ, compared to the Poisson distribution where $E[D] = \lambda$ and std $D = \sqrt{\lambda}$, is that it is possible to model more or less variation. However, the disadvantage is the continuous distribution, whereas the Poisson distribution is a discrete distribution. As ASML is operating in a low volume environment, models are required where the output values are integers. Adan, Eenige & Resing (1994) developed and proofed a method to fit a discrete distribution on a given distribution. Four classes of discrete distributions (Poisson, Binomial, Negative Binomial and Geometric) are used in order to select one of these distribution with μ and CV (coefficient of variation) that are equal to the μ and CV of an arbitrary discrete distribution with non-negative integers. Broekmeulen & Donselaar (2013) implemented this method in Excel.

As stated in section 2.2.1, ASML's demand is observed as demand with high variability due to the high coefficient of variation. For high coefficients of variation the Normal distribution is not appropriate due to the fact that significant negative demand occurs (Hillier, 2000). Therefore, when dealing with higher coefficients of variation the Gamma distribution should be used. Based on experience of De Kok (2020), the Gamma distribution has a good fit with practical demand data.

However, the Gamma distribution is a continuous distribution. The continuous random demand can be transformed to discrete demand in order to be applicable for the low volume environment of ASML. Therefore, a discretization process is required to transform the continuous Gamma distributed demand into discrete numbers. First, based on the input parameters μ and σ, the α and β are calculated with equations [11] and [12]. With these parameters, random continuous Gamma demand is generated. Next, the first random demand number is rounded to zero decimals. The leftover or additional value to increase the number to an integer is added to or subtracted from the subsequent demand number. This process is repeated until the last random demand number. There is a small effect on the standard deviation of the set of random demand numbers, i.e. small increase in standard deviation.

$\mu = \alpha * \beta$ leads to $\beta = \frac{\mu}{\alpha}$ [11]

$\sigma^2 = \alpha * \beta^2$ plugging in β and simplify leads to $\alpha = \frac{\mu^2}{\sigma^2}$ [12]

Based on the three years demand data in weeks (156 data points) of a lithography system type A, the μ and σ is calculated. Furthermore, a frequency table is created in order to calculate the probability mass function (PMF). The tool of Broekmeulen & Donselaar (2013) is used to fit one of the four classes of discrete distribution to the demand data by using μ and σ in order to calculate the PMF. The best fit according to Adan, Eenige & Resing (1994) is the Negative Binomial distribution. Moreover, 1,000,000 random generated continuous demand numbers are discretized and based on these numbers a frequency table is create in order to calculate the PMF. Figure 25 visualized the PMF of the different distributions (P(Demand = X)). Due to confidentiality, the probabilities and X-values are not shown; moreover, the order of the 20 X-values are randomly chosen. Based on the PMF, the summed square of residuals (S) is calculated with formula [13]. By applying this formula, the summed square of residuals are [0.0070, 0.0028] for [Gamma (discretized), Negative Binomial] respectively. As can be seen, the Negative Binomial is the best fitted distribution. However, the Gamma (discretized) distribution performs also fairly good. In conclusion, as the Negative Binomial Distribution is not a common used distribution in an experimental setting and the Gamma (discretized) distribution performs good, this distribution is chosen in the simulation approach applied in the next sections.

$S = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$, where y_i is the data and $\hat{y}_i$ the fit [13]

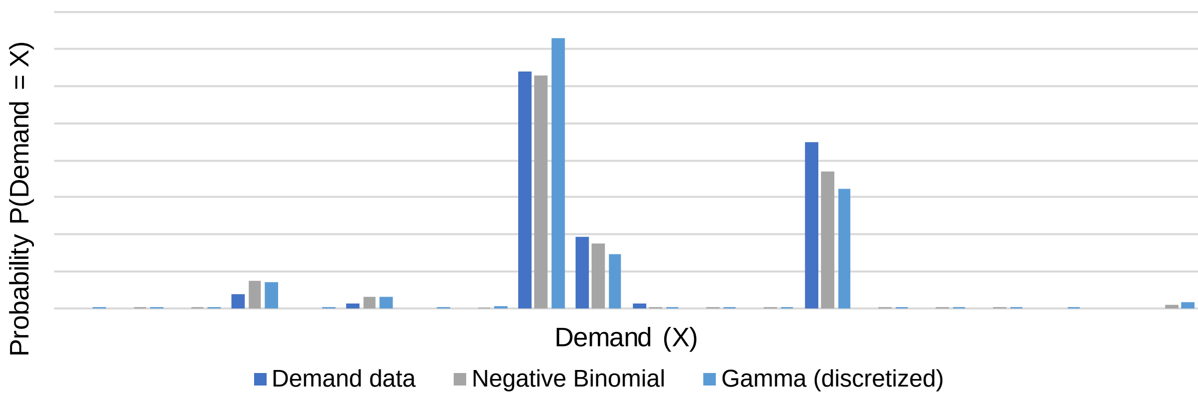


Figure 25. PMF comparison of demand distributions

Step 2: determine hedging scenario

The next step is to determine a hedging scenario. This means that the customer order lead time and number of hedging positions with associated lead times and holding and/or assembly costs should be chosen.

Step 3: calculate near-optimal base-stock levels

In the third step, the target service level should be determined. There are three service level measures used in inventory management, namely cycle service level, fill rate, and ready rate.

According to Silver, Pyke & Peterson (1998) these service levels are defined as follows:
- Cycle service level (P1): "fraction of cycle in which a stockout does not occur"
- Fill rate (P2): "fraction of customer demand that is met routinely; that is, without backorders or lost sales" to be satisfied routinely from the shelf"
- Ready rate (P3): "fraction of time during which the net stock is positive"

The net stock is equivalent to the local net inventory (IN'j) as described in section 5.1. In this research the time unit is weeks and the review period is set to one week; therefore, the cycle service level is equal to the ready rate. Practitioners are mostly interested in the fill rate. The ready rate could be more relevant for example in case of emergency equipment, which should be always available. The ready rate is also defined as the non-stockout probability (De Kok, 2015).

According to Silver, Pyke & Peterson (1998), the ready rate is equivalent to the fill rate when demand is Poisson distributed. However, no detailed information is provided about circumstances. According to De Kok (2020), this holds only for continuous time inventory models or a compound renewal process where the interarrival times are exponentially distributed and demand per time unit is deterministic.

The Diks & De Kok algorithm can be used to obtain the near-optimal echelon base-stock levels for a specific hedging scenario and target fill or ready rate. For the Shang & Song heuristic, backorder costs are required to determine the echelon base-stock levels. In case the backorder costs b are unknown, the Newsvendor fractile [14] can be used to determine these costs (De Kok, 2015). This Newsvendor fractile can only be used for the ready rate (non-stockout probability). The cost structure defined in the previous step can be used to calculate the local holding costs at the last hedging position.

$$P\{IN_J \geq 0\} = \frac{b_J}{b_J + h'_J} \quad \text{can be rewritten to} \quad b_J = \frac{P\{IN_J \geq 0\} * h'_J}{1 - P\{IN_J \geq 0\}} \qquad [14]$$

The parameters defined for the hedging scenario and backorder costs are used to calculate the echelon base-stock levels with the Shang & Song heuristic [4], [5], and [6].

<u>Step 4 and 5: generate random demand for warm-up period and actual demand</u>

In the fourth step, random demand is generated for a warm-up period based on the demand parameters from step one. Furthermore, the actual demand is random demand generated with the parameters defined in step one or the actuals to evaluate. The demand is stationary, i.e. a stochastic process where demand parameters are not changing over time. Non-stationary demand can be analyzed by adjusting the demand parameters and rerun the simulation.

<u>Step 6: simulation serial system</u>

The sixth step is to simulate the serial system defined in the hedging scenario with the echelon base-stock levels from step 3. The simulation could be iterated over n number of periods. Based on experience in simulating inventory control models, 20,000 iterations are required to obtain a reliable point estimate (De Kok, 2020).

<u>Step 7: performance measure report</u>

After the iterations, a performance measure report is created. The performance measures are average service level (ready rate and fill rate), average backorders, average inventory in-transit in terms of units and costs, and average inventory on-hand in terms of units and costs. Furthermore, insights in the variation of these performance measures could be provided.

5.4.2 Experiment performance Shang & Song heuristic and Diks & De Kok algorithm

In this subsection, the proposed echelon base-stock levels of the Shang & Song heuristic and Diks & De Kok algorithm are evaluated with the simulation of a serial system. With this simulation, the actual service level could be compared for each analytical method with the set target.

As described in subsection 5.3.2, based on the first two moments, the Gamma distribution is fitted in Diks & De Kok algorithm. The gamma distribution is a continuous distribution which result in continuous base-stock levels. However, the simulation is a discrete time model and is built to evaluate discrete base-stock levels under a discrete demand distribution. This is to ensure that output values are integers in order that this simulation is applicable for ASML's low volume environment. In order to make an equal comparison, the Shang & Song heuristic is performed under Gamma distributed demand as well to determine the discrete echelon base-stock levels. The echelon base-stock levels from Diks & De Kok algorithm will be rounded to the nearest integer number in order to create discrete values.

Diks & De Kok algorithm is able to determine near-optimal base-stock levels for a target ready and fill rate. However, the Shang & Song heuristic is only applicable for a target ready rate. Therefore, in the first experiment, the performance of the two analytical methods is evaluated based on a target ready rate. Whereas, in the second experiment the Shang & Song heuristic is adjusted in order to compare the performance with Diks & De Kok algorithm for a fill rate target.

Experiment for target ready rate

In this experiment, the echelon base-stock levels of the two analytical methods are evaluated in a four-echelon serial system under different hedging scenarios. For each hedging scenario, the echelon base-stock levels per method are determined based on the target ready rate. The lead times and echelon holding costs remain equal in each hedging scenario and are the following fictitious numbers: (L1, L2, L3, L4) = (20, 10, 5, 5) and (h1, h2, h3, h4) = (5, 150, 200, 100). Figure 26 visualizes the serial system hedging scenario. The scenarios are shown in Table 1 where the target ready rate varies from [85, 90, 95], μ from [1, 2] and σ from [0.5, 1, 1.5]. The backorder costs, required for Shang & Song heuristic, are calculated with [14] such that the ready rate (RR) is equal to the set target.

Table 1. Fictitious scenarios for comparison Shang & Song heuristic and Diks & De Kok algorithm – ready rate

Scen.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
RR %	85	85	85	85	85	85	90	90	90	90	90	90	95	95	95	95	95	95
μ	1	1	1	2	2	2	1	1	1	2	2	2	1	1	1	2	2	2
σ	0.5	1	1.5	0.5	1.0	1.5	0.5	1	1.5	0.5	1	1.5	0.5	1	1.5	0.5	1	1.5

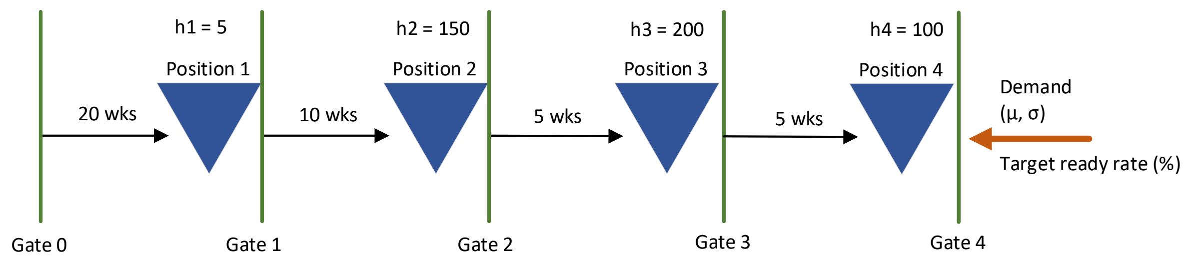


Figure 26. Hedging scenario for comparison

The discretization process to discretize the Gamma distributed random numbers is explained in section 5.4.1 step 1. In order to generate the required standard deviation of [0.5, 1, 1.5], the input standard deviation is lowered. Trial and error is used to obtain the required standard deviation of the set of random demand numbers. For example, an input standard deviation of 0.33 is required to obtain a standard deviation of 0.5, whereas for target standard deviation 1.5, the input standard deviation was 1.46. However, this is different for each trial due to the impact of the randomness. To ensure an equal comparison between the analytical methods, the same random generated demand numbers are used per scenario.

The results of the evaluation of the two analytical methods is shown in Table 2, where the signs above parameters ^ refer to Shang & Song heuristic and * to Diks & De Kok algorithm. It is interesting to observe that the Diks & De Kok algorithm always allocate more inventory to the first hedging position. Whereas, the Shang & Song heuristic always allocate equal or more inventory to the second and third hedging position. C(s) are the total costs, this includes the average long-run costs for inventory in-transit and inventory on-hand. The average long-run inventory in-transit costs are equal per scenario for both methods, as these are dependent on the lead time and μ.

To compare the methods, for each method it is examined whether there are scenarios in which the ready rate is higher, but total costs are lower than the other method. For Shang & Song heuristic there no scenarios were found. However, for Diks & De Kok algorithm in scenario 16 the ready rate is 0.2% higher and total costs are 4 lower, shown in Table 2. This difference can be experienced as small; however, when the holding costs are in millions of euros this difference is large. This difference occur due to the different inventory allocation of the methods described above.

For the other scenarios the total cost per 1% deviation (M) are calculated for each scenario where ready rate and total costs are higher compared to the other method. This will result in a comparison by magnitude, shown in Table 2. For example, when M^ (Shang & Song heuristic) is 27.1, this means that on average in this scenario 27.1 extra total costs are required to increase the ready rate by 1% compared to the ready rate of Diks & De Kok algorithm*. The average total costs increase per 1% for each target ready rate are for Shang & Song heuristic: (85%, 90%, 95%) = (64.1, 129.4, 260.7) respectively. Whereas, for the Diks & De Kok algorithm these are: (85%, 90%, 95%) = (33.8, 20.1, 22.0) respectively. These significant differences could indicate that Diks & De Kok algorithm determine better cost optimal echelon base-stock levels than Shang & Song heuristic when evaluating the ready rate.

Table 2. Results evaluation ready rate Shang & Song heuristic (^) and Diks & De Kok algorithm (*)

Target RR	Scen.	μ	σ	s_1^/s_1*	s_2^/s_2*	s_3^/s_3*	s_4^/s_4*	RR^/RR*	C(s)^/C(s)*	M^/M*
85.0	1	1	0.5	48/49	24/24	14/13	9/8	82.4/79.6	5675/5599	27.1/-
	2	1	1	54/57	27/27	16/16	12/11	80.3/81.2	6820/6849	-/32.2
	3	1	1.5	60/66	31/30	19/18	15/14	83.1/81.3	8247/8005	134.4/-
	4	2	0.5	89/90	45/45	25/24	15/14	40.8/38.7	10347/10343	1.9/-
	5	2	1	95/98	48/48	27/27	17/17	67.8/69.1	11348/11394	-/35.4
	6	2	1.5	101/106	52/51	29/29	20/20	77.0/76.0	12609/12516	93.0/-
90.0	7	1	0.5	48/50	25/25	14/14	9/9	92.0/92.9	5934/5961	-/30.0
	8	1	1	55/59	29/28	17/17	12/12	87.9/72.7	7370/6480	58.6/-
	9	1	1.5	62/70	33/32	20/20	16/15	87.7/87.4	8842/8750	306.7/-
	10	2	0.5	89/91	46/46	25/25	15/15	60.2/63.6	10475/10510	-/10.3
	11	2	1	96/99	50/49	28/27	18/17	80.1/74.9	11853/11602	48.3/-
	12	2	1.5	103/109	54/53	31/30	21/20	85.2/82.9	13312/13073	103.9/-
95.0	13	1	0.5	49/51	26/26	15/14	10/9	97.4/96.8	6352/6185	278.3/-
	14	1	1	57/62	31/30	18/18	13/13	93.0/92.7	7990/7873	390.0/-
	15	1	1.5	65/74	36/36	23/22	18/17	93.9/93.7	10041/9951	450.0/-
	16	2	0.5	90/92	46/46	26/25	15/15	63.7/63.9	10520/10516	-/-
	17	2	1	98/101	51/51	29/29	19/18	87.0/87.5	12254/12265	-/22.0
	18	2	1.5	106/112	56/56	33/32	22/22	91.1/90.7	14111/14029	205.0/-

As can be seen in Table 2, the ready rate in the simulation for both analytical methods deviates from the target ready rate. In order to evaluate if this deviation is on average too high or low, the mean error (ME) (bias), is calculated by using the actual errors. Moreover, the mean absolute percentage error (MAPE) is calculated by using the absolutes errors in order to indicate on average what the deviation percentage of the simulation ready rate to target ready rate is. Table 3 shows the MPE [15] and MAPE [16] for both methods per target ready rate by evaluating the ready rate difference. The ME for both methods indicate that in two out of three target ready rates, the simulation ready rate were too low. The MAPE indicates that on average in two out of three cases the simulation ready rate of the Shang & Song heuristic outperforms the Diks & De Kok algorithm. For the other case, Diks & De Kok algorithm performs equal. The detailed MAPE for each scenario is visualized Figure 27. Remarkable is that for target ready rate scenario with μ =

2 and σ = 0.5 (scenario 4, 10 and 16), the MAPE is extremely high. This might be caused due to the relatively low standard deviation compared with the mean. Therefore, the echelon base-stock levels are determined too low in the two methods, which result in more periods of negative net inventory. To investigate this, 100,000 random Gamma distributed demand are generated with μ = 2 and σ = 0.5. These demand numbers are discretized according to the procedure described in subsection 5.4.1. For the Gamma distributed demand, a frequency table is created with a bin size of 0.1. This table is combined with the frequency table of the Gamma (discretized) demand and visualized in Figure 28. As can be seen, there is a large difference in the frequency per demand number. In conclusion, the Shang & Song heuristic is able to determine echelon-base stock levels which result in a ready rate closer to the set target than Diks & De Kok algorithm. However, Diks & De Kok algorithm determine better cost optimal echelon base-stock levels.

$$ME = \frac{\sum_{i=1}^{n}(simulation\ RR_i - target\ RR_i)}{n} \quad \text{i = scenario number and n = number of scenarios} \qquad [15]$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{simulation\ RR_i - target\ RR_i}{simulation\ RR_i}\right| \qquad [16]$$

Table 3. Error measures simulation RR versus target RR Shang & Song heuristic (^) and Diks & De Kok algorithm (*)

Target RR	ME^	ME*	MAPE^	MAPE*
85%	-13.1%	-14.0%	15.4%	16.5%
90%	-7.8%	-10.9%	9.4%	13.0%
95%	7.3%	7.5%	8.5%	8.5%

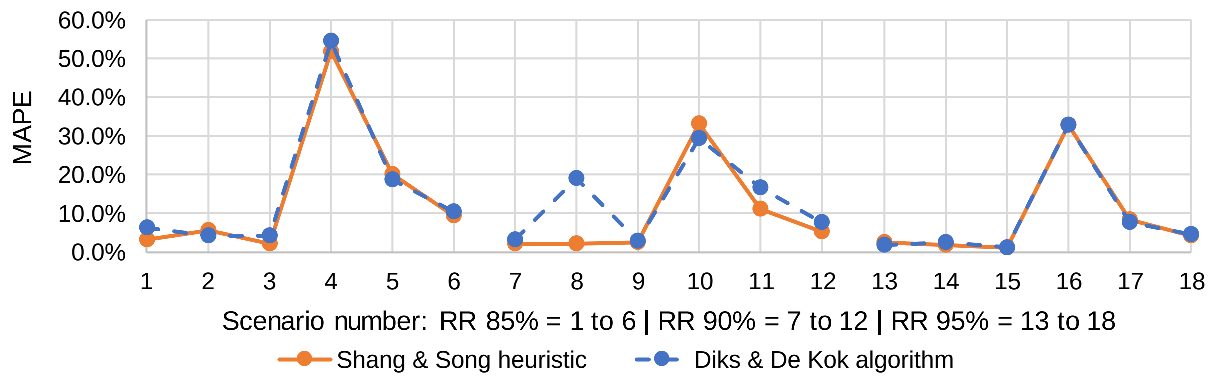


Figure 27. MAPE ready rate per scenario for target ready rate

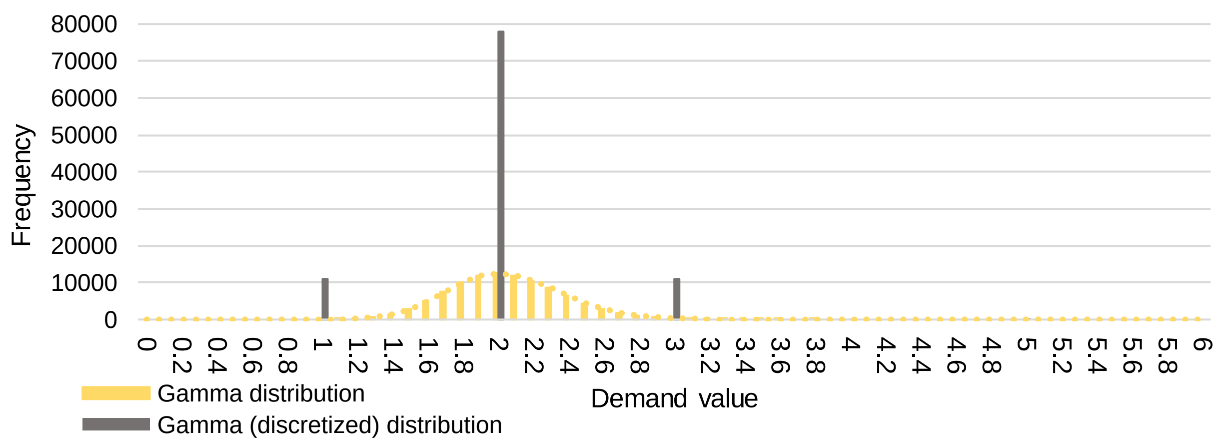


Figure 28. Frequency table for Gamma and Gamma (discretized) distribution (100,000 random values)

A small additional experiment is performed to identify the difference between the obtained ready rate and fill rate in the simulation with a target ready rate. Table 4 shows the average deviation of the different scenarios for a target ready rate specified per analytical method. It is remarkable that the deviation between these two service levels varies between 8% and 18.2%, and decrease when the target ready rate is set higher. This differs from the continuous inventory models, where the ready rate is always really close to the fill rate.

Table 4. Deviation simulation ready rate and fill rate Shang & Song (^) and Diks & De Kok algorithm (*)

Target RR	(FR^-RR^)	(FR*-RR*)
85%	17.7%	18.2%
90%	10.9%	12.1%
95%	8.0%	8.0%

Experiment for target fill rate

The same analysis could be performed to evaluate and compare the obtained fill rate to the target ready rate of the two methods. However, the Shang & Song heuristic is not able to determine echelon base-stock levels for a set target fill rate. Therefore, the heuristic should be adjusted. For this experiment in the Shang & Song heuristic it is assumed that the target ready rate is equal to the target fill rate. Remark, this holds only for the Shang & Song heuristic and not for Diks & De Kok algorithm. In this target fill rate experiment the same four-echelon serial system and hedging scenarios are used as in the previous ready rate experiment (L1, L2, L3, L4) = (20, 10, 5, 5) and (h1, h2, h3, h4) = (5, 150, 200, 100), shown Figure 26 and Table 5.

Table 5. Fictitious scenarios for comparison Shang & Song heuristic and Diks & De Kok algorithm – fill rate

Scen.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
FR %	85	85	85	85	85	85	90	90	90	90	90	90	95	95	95	95	95	95
μ	1	1	1	2	2	2	1	1	1	2	2	2	1	1	1	2	2	2
σ	0.5	1	1.5	0.5	1.0	1.5	0.5	1	1.5	0.5	1	1.5	0.5	1	1.5	0.5	1	1.5

Table 6 shows the results of the fill rate evaluation. The average total costs increase per 1% fill rate increase (M) compared with the other analytical method for each target fill rate scenario are for Shang & Song heuristic (M^): (85%, 90%, 95%) = (98.8, 158.4, 303.2) respectively. Whereas, for the Diks & De Kok algorithm (M*) these are: (85%, 90%, 95%) = (81.3, 132.0, 266.3) respectively. These significant differences could indicate that Diks & De Kok algorithm determine better cost optimal echelon base-stock levels than Shang & Song heuristic when evaluating the fill rate target.

De Kok & Fransoo (2003) indicates by extensive numerical studies two import learnings of supply chain design. First, inventory is primary concentrated at the CODP. The reasoning behind this is because the CODP is the only point where inventory directly contribute to customer service level. Therefore, it can be concluded that optimal solutions of inventory allocation are characteristics by items in the supply network flowing towards the CODP. However, the second learning states that items which are cheap and have a long lead times, a buffer of these items are held in decoupling stocks. In Table 6 the echelon base-stock levels are shown for each scenario and the obtained fill rate and total costs C(s) from simulation. Both learnings can be identified in the results of the Diks & De Kok algorithm. As can be seen, little inventory is allocated to hedging position 2 (s'2^) and 3 (s'3^) as the local base-stock levels (s'j^) are almost equal to the average demand over the lead time in each scenario. However, for hedging position 4 (s'4^) there is an additional buffer level to cope with the variation in demand. When the standard deviation and/or fill rate target is increasing, the inventory at hedging position 4 (s'4^) is increasing as well. Furthermore, inventory is increasing at hedging position 1 (s'1^) as the lead time to this position is long (20 weeks) and the commitment costs of these items are extremely low compared to the other positions. For example, the local base-stock levels of scenario 2 are (s'1^, s'2^, s'3^, s'4^) = (30, 11, 5, 11). The demand over the lead time of position 2 and 3 are 10 * 1 = 10 and 5 * 1 = 5 respectively. Only 1 extra buffer is planned for position 2. However, for position 1 and 4 there are 10 and 6 buffers respectively planned above the demand over the lead time.

Table 6. Results evaluation fill rate Shang & Song (^) and Diks & De Kok algorithm (*)

Target RR	Scen.	μ	σ	s_1^/s_1*	s_2^/s_2*	s_3^/s_3*	s_4^/s_4*	FR^/FR*	C(s)^/C(s)*	M^/M*
85.0	1	1	0.5	48/48	24/24	14/13	9/8	94.1/93.1	5675/5599	-/78.4
	2	1	1	54/57	27/27	16/16	12/11	87.3/87.8	6820/6849	63.0/-
	3	1	1.5	60/68	31/31	19/19	15/15	85.2/86.2	8247/8383	134.7/-
	4	2	0.5	89/89	45/44	25/23	15/14	91.9/90.1	10347/10318	-/15.9
	5	2	1	95/97	48/47	27/26	17/16	89.8/87.6	11348/1156	-/88.9
	6	2	1.5	101/106	52/51	29/28	20/19	89.3/87.7	12609/12390	-/142.2
90.0	7	1	0.5	48/49	25/24	14/13	9/8	97.2/93.5	5934/5706	-/61.0
	8	1	1	55/59	29/28	17/17	12/12	92.0/91.3	7370/7229	-/188.0
	9	1	1.5	62/70	33/33	20/20	16/16	88.8/89.7	8842/8983	158.4/-
	10	2	0.5	89/90	46/45	25/24	15/14	94.1/91.7	10475/10343	-/54.8
	11	2	1	96/98	50/48	28/27	18/17	93.5/90.3	11853/11394	-/143.4
	12	2	1.5	103/108	54/53	31/30	21/20	93.0/91.8	13312/13063	-/212.8
95.0	13	1	0.5	49/50	26/25	15/14	10/9	99.0/97.4	6352/5961	-/238.4
	14	1	1	57/62	31/30	18/18	13/13	95.4/95.1	7990/7873	-/468.0
	15	1	1.5	65/75	36/37	23/23	18/18	94.0/95.3	10041/10420	303.2/-
	16	2	0.5	90/91	46/46	26/24	15/14	94.6/93.4	10520/10461	-/53.2
	17	2	1	98/101	51/50	29/28	19/17	95.6/94.2	12254/11936	-/233.8
	18	2	1.5	106/112	56/55	33/32	22/21	95.7/94.9	14111/13827	-/338.1

The mean error (ME) and mean absolute percentage error (MAPE) for the different target fill rates are shown Table 7. The simulation fill rate is on average slightly too high in all cases. The MAPE is relatively small and is decreasing when the target fill rate is increasing, thus higher echelon base-stock levels. For a target fill rate of 95% the MAPE in the Diks & De Kok algorithm is only 0.9%. It can be concluded that the Diks & De Kok algorithm outperforms the Shang & Song heuristic significantly as the MAPE for each target fill rate is lower. Figure 29 shows the MAPE for each scenario. Remarkable is that for most scenarios where σ is 0.5 (scenario 1, 4, 7, 10, 13), the MAPE is much higher than the other scenarios. This in line with observations and reasons discussed previously for the target ready rate. In conclusion, the Diks & De Kok algorithm is outperforming the Shang & Song heuristic with respect to base-stock levels closer to the set fill rate target and better cost optimal. When the target fill rate increases, the actual simulation fill rate error will decrease. Furthermore, a fill rate closest to the target can be obtained on average when μ is higher and σ is lower because the continuous demand assumption in Diks & De Kok algorithm performs better in these cases. The average MAPEs for scenarios with μ = (1, 2) are (2.5%, 2.1%) respectively and for with σ = (0.5, 1.0, 1.5), the MAPEs are (4.2%, 1.5%, 1.2%).

Table 7. Error measures simulation FR versus target FR Shang & Song (^) and Diks & De Kok algorithm (*)

Target FR	ME^	ME*	MAPE^	MAPE*
85%	4.6%	3.8%	5.4%	4.4%
90%	3.1%	1.4%	3.9%	1.6%
95%	0.7%	0.0%	1.2%	0.9%

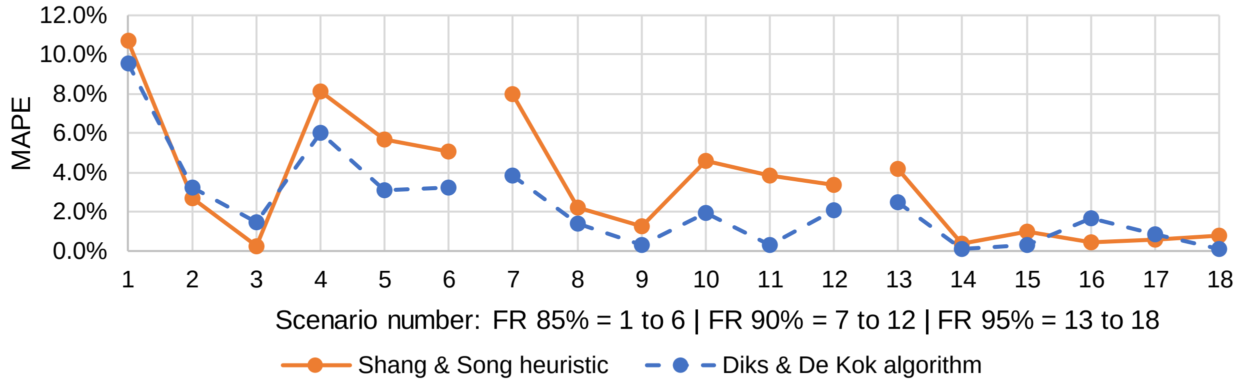


Figure 29. MAPE fill rate per scenario for target fill rate

Reasoning for difference between target and simulation service level

The difference in target service level and actual service level provided by simulation can be explained by different reasons, specific for each analytical method. For the Shang & Song heuristic, the main numerical experiment of Shang & Song (2003) examined 32 scenarios with equal Poisson demand rate and lead times; however, backorder costs and echelon holding costs vary to obtain a ready rate of 90 and 99%. In this experiment lead times were L1 = L2 = L3 = L4 = 0.25 periods and Poisson demand rate λ = 16. Besides this main experiment, additional experiments with a few scenarios were examined. First, a single scenario with L1 = L2 = L3 = L4 = 1.5 and λ = 4. Next, a system is evaluated with normal distributed demand with μ = 50 and σ = 10 and L1 = L2 = L3 = L4 = 2. Furthermore, a system is evaluated with negative binomial demands with μ = 50 and $σ^2$ = 10 and L1 = L2 = L3 = L4 = 0.25. As can be seen, the demand rates and means used in the numerical experiments of Shang & Song (2003) are extremely high compared with the scenarios evaluated with the simulation model namely μ varies from [1, 2] and σ from [0.5, 1, 1.5]. Moreover, the lead times are extremely low in comparison with all scenarios (L1, L2, L3, L4) = (20, 10, 5, 5). Furthermore, the Gamma distributed demand is not used in the experiments of Shang & Song (2003). For Diks & De Kok algorithm, the difference in target service level and actual service level in the simulation can be explained by the fact that in the simulation there is a low demand per time unit. This in combination with the discrete demand is a source of deviations for the algorithm in ChainScope as this uses a continuous demand distribution. When the service level decreases, the effect of discrete demand is strengthened. Moreover, the base-stock levels determined by the algorithm are rounded to the nearest integer in order to use them in simulation.

5.4.3 Simulation to validate practical situation

In order to validate the practical situation of planning in a low-volume and high-value environment, it is important that the focus is on fill rate. According to Hopp & Spearman (2002), fill rate is the performance indicator that is used by most practitioners, including ASML. In the previous subsection, based on the experiment it is shown that the Diks & De Kok algorithm performs determines the best cost-optimal echelon base-stock levels for a target fill rate and can obtain an actual fill rate close to this target. Especially for higher fill rate targets, as the MAPE for fill rate target 95% is only 0.9%.

For this simulation all the steps from the simulation approach described in subsection 5.4.1 are followed. A warm-up period of 30 weeks is used to create a balanced serial system, 30 weeks is chosen to ensure that the longest lead time of 20 weeks is covered. Hereafter, 52 weeks of actual demand are simulated. The total of 30 + 52 weeks can be seen as one iteration. After each iteration, the start situation, where inventory at each hedging position equals local base-stock level, is restored. This means that the simulation is transient as it stops after a set time. In total 10,000 iterations of 30 + 52 weeks will be simulated to obtain a reliable one-year estimation.

The hedging scenario that is analyzed is a specific scenario from the previous subsection 5.4.2. However, the demand parameters differ as these are actual past year demand figures of a lithography system. ASML's demand can be characterized as high variability, as stated in subsection 2.2.1, which means that the orderbook visibility vary and therefore managing the buffers based on unreliable forecasts creates nervousness. Therefore, data from the past is used. Due to confidentiality, the mean and standard deviation are not given and shown as x and y. The hedging scenario is shown in Figure 30. In Diks & De Kok algorithm, the echelon base-stock levels are calculated for this hedging scenario by using a fill rate target of 95%. As the echelon base-stock levels are continuous numbers, these are rounded to the nearest integer value.

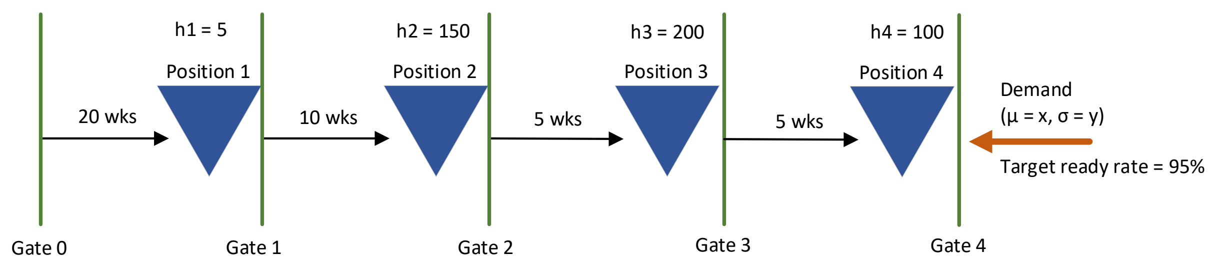


Figure 30. Hedging scenario for simulation to validate practical situation

Table 8 shows the results of the simulation in average values, be aware that these are the averages of the iterations (one iteration is 52 weeks). The average fill rate of one year demand is 95.1%. Moreover, insights in the variation per year of the fill rate and inventory on-hand at each position can be given. These insights are visualized in Figure 31 and Figure 32. As can be seen, the randomness does play a role as this determines for example the variation in fill rate from 29% to 100%, with most of the cases within the range [93%,100%]. This is obvious as the average is 95.1%. Furthermore the inventory on-hand at the last hedging position (HP4) vary between around 8 and 12, and has some outliers.

Table 8. Results validation practical situation

	Echelon base-stock level (s1, s2, s3, s4)	**Avg. IOH_1**	**Avg. IOH_2**	**Avg. IOH_3**	**Avg. IOH_4**	**Avg. fill rate**
Simulation	(86, 42, 26, 19)	18.9	4.6	2.2	10.3	95.1%

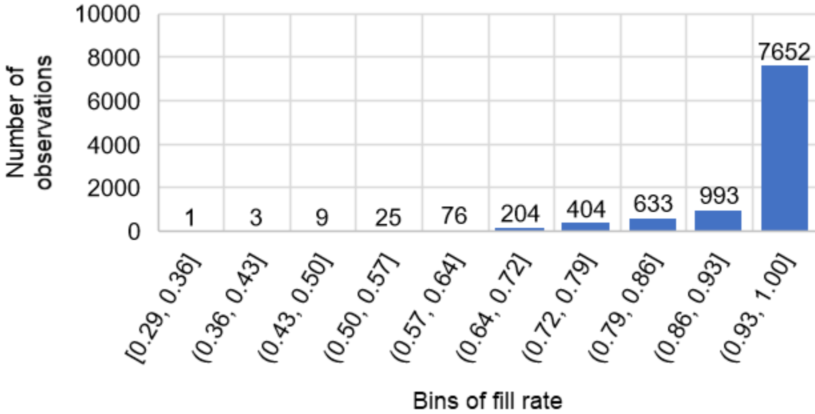


Figure 31. Variation in fill rate level

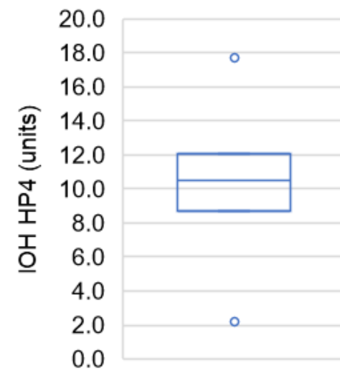


Figure 32. Variation HP4 IOH

5.4.4 Simulation to evaluate current situation

In this section, the current situation at ASML is analyzed with the simulation model in order to evaluate the base-stock levels and the performance in practice. Moreover, the desired situation is obtained by applying the Diks & De Kok algorithm. First, the actual commitment value curve and lead times of a lithography system type A are analyzed to determine where to place the hedging positions. Furthermore, for this hedging scenario a customer order lead time is assumed. Due to confidentiality, it is not possible to state the actual numbers. To evaluate the current situation an assumption has to be made about demand parameters at a specific point in time. To determine the demand parameters, a snapshot is made from the demand plan where all forecasted demand figures of a specific future year were shown. It is assumed that the total year firm demand can be used to determine the number of lithography system demand in weeks (μ). Whereas, the upside demand can be used to determine the variability in this demand (σ).

To obtain the base-stock levels at each position, the supply plan is analyzed. The sum of the lithography system FASY starts in a specific time window (lead time to hedging position) is assumed to be equal to the local base-stock level of that hedging position. When the local base-stock levels are obtained, the echelon base-stock levels can be calculated.

By using these echelon base-stock levels and hedging scenario in the simulation model, results can be obtained regarding fill rate and ready rate, average inventory on-hand with associated value and inventory in-transit with associated value. This current situation performance measures could be compared with the desired situation. The desired situation is obtained by determining the echelon base-stock levels with Diks & De Kok algorithm for a target fill rate. Due to confidentiality, it is not possible to provide detailed information. The percentage error between the current and desired fill rate of system type A is -24.3%. Furthermore, an inventory investment of factor 2.1 with respect to the current investment is required to obtain the desired fill rate. This inventory investment is especially required at the last hedging position. Finally, the local base-stock levels should primarily be increased at the first and last hedging position to obtain the desired fill rate.

5.5 Optimal buffer decision for a specified service level objective

In this section trade-off curves will be visualized of different scenarios to enable an informed decision of the required buffer levels. In the previous subsection the two analytical methods were evaluated with the simulation. For this subsection, the Diks & De Kok algorithm will be used to create the trade-off curves. This analytical method, implemented in the tool ChainScope, can do it in a more efficient way than using simulation. Note that there is a small deviation in target and actual fill rate, which is stated in subsection 5.4.2. It is important that the relationship between fill rate, inventory commitment costs, and customer order lead time is shown. Customer order lead time can be useful in order to determine how much buffers are required to achieve a performance in fill rate to customers based on an agreed lead time. Furthermore, the impact of demand variability can be visualized. The trade-off curves visualize the scenarios if the demand parameters in the future are exactly the same as the input parameters. In reality there will be always a difference; however, the insights in costs and flexibility, and where decisions in the supply chain should be made is useful for management and supports decision-making.

First, the initial hedging scenario is decided. Hereafter, the impact of customer order lead time, fill rate, and demand variability on costs are analyzed. The initial hedging scenario is visualized in Figure 33 and is equal to one of the scenarios used in section 5.4 with $\mu = 2$ and $\sigma = 1.5$.

When the customer order lead time is increasing, the position of the last hedging position is shifting upstream and hedging position(s) might completely vanishes. This means that components with a shorter lead time than the customer order lead time should not be purchased before a customer order arrives. In order to perform this analysis, assumptions have to be made about how the added value at the hedging positions are divided over the corresponding lead time, i.e. at which time which proportion of the value is added. This proportion is determined randomly for L2, L3 and L4 as these are influenced when the CODP shifts upstream. The random cost determination for each period in the lead time is shown in Table 9. The overview of the customer order lead times scenarios with associated real inventory values (h'j) at the positions are shown in Table 10.

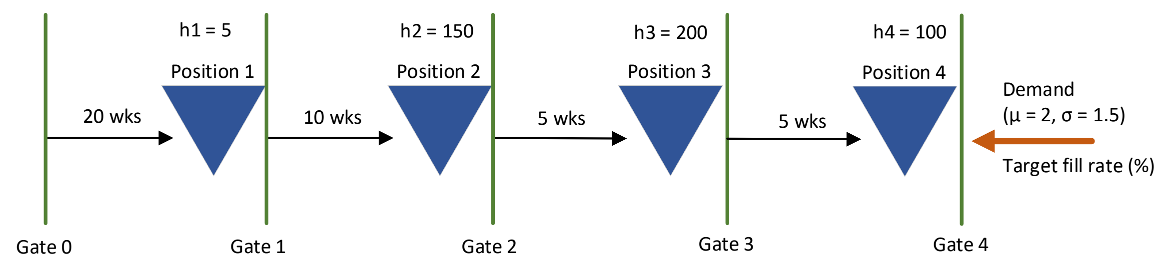


Figure 33. Initial hedging scenario optimal buffer decision

Table 9. Distribution added value over lead time

Lead time	1	2	3	4	5	6	7	8	9	10	Total
HP2				15		5	10		40	80	150
HP3	10	20		60	110						200
HP4	15	5	15		65						100

Table 10. Overview of customer order lead time scenarios and associated local holding costs

Scenario CLT	L1	h'1	L2	h'2	L3	h'3	L4	h'4
0	20	5	10	155	5	355	5	455
1	20	5	10	155	5	355	4	390
2	20	5	10	155	5	355	3	390
3	20	5	10	155	5	355	2	375
4	20	5	10	155	5	355	1	370
5	20	5	10	155	5	355		
6	20	5	10	155	4	245		
7	20	5	10	155	3	185		
8	20	5	10	155	2	185		
9	20	5	10	155	1	165		
10	20	5	10	155				
11	20	5	9	75				
12	20	5	8	35				

In Figure 34 the trade-off curves are shown of the total supply chain costs based on different customer order lead times and fill rates. These trade-off curves can also be named as efficient frontiers, since the curve represent the lowest inventory investment for a fill rate combination (Hopp & Spearman, 2000). The total supply chain includes the commitment costs (inventory in-transit and inventory on-order) upstream from the CODP and the costs for inventory in-transit downstream from the CODP up to FASSY (customer-order supply chain). According to De Kok & Fransoo (2003), a large impact on stock investment occur when CODP is shifted upstream. This impact can be seen in Figure 34, as the total supply chain costs decrease when customer order lead time increases.

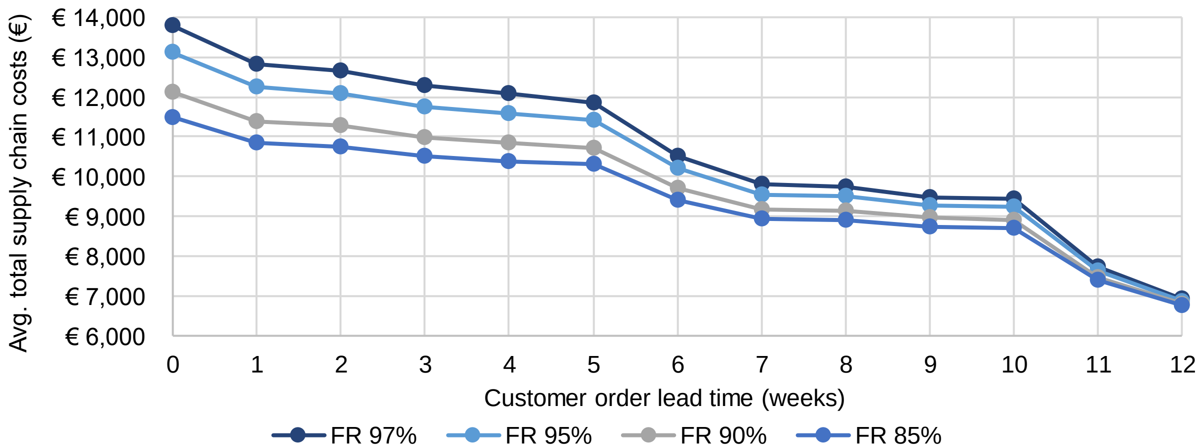
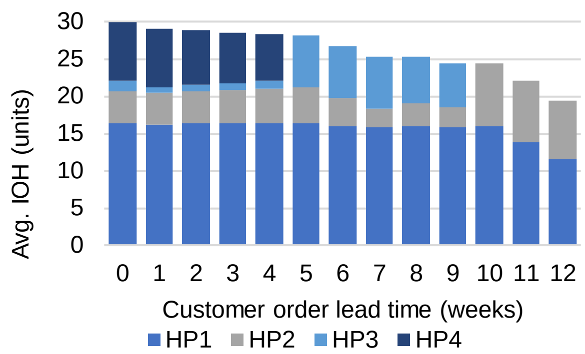


Figure 34. Trade-off curves total supply chain costs

As the total inventory in-transit and review period are equal for each scenario, there are only two reasons to explain the decreasing cost curve. The first factor is the inventory on-hand (IOH). As can be seen in Figure 35, the total average inventory on-hand is decreasing when customer order lead time increase. The second factor is the decrease in added value per hedging position, shown in Figure 36. When the customer order lead time increases, orders can be purchased and assembly activities can start after the customer order is received. The inventory on-hand is multiplied with the total added value of an end item in that specific hedging position to obtain the total inventory on-hand costs.



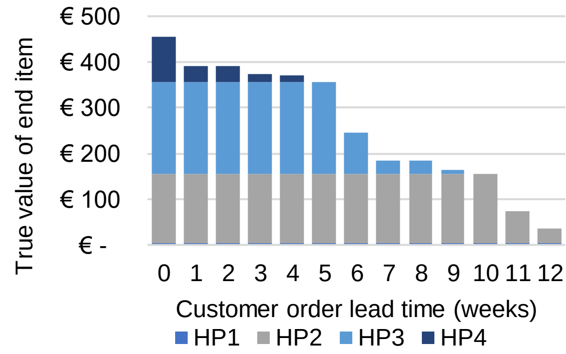


Figure 35. On-hand inventory with fill rate 95%

Figure 36. Added value of a system per hedging position

Figure 37 shows the trade-off curves of the inventory on-hand costs, these are similar to the costs for buffers because this on-hand inventory is required to cover demand variation. The demand during the review period of one week is included in these buffer costs. In Figure 38 the impact of varying the standard deviation for the scenario fill rate is 95% and customer order lead time is 4 weeks is shown. The average inventory on-hand costs vary between 573 euro and 6341 euro with standard deviation between 0.5 and 2.5 respectively.

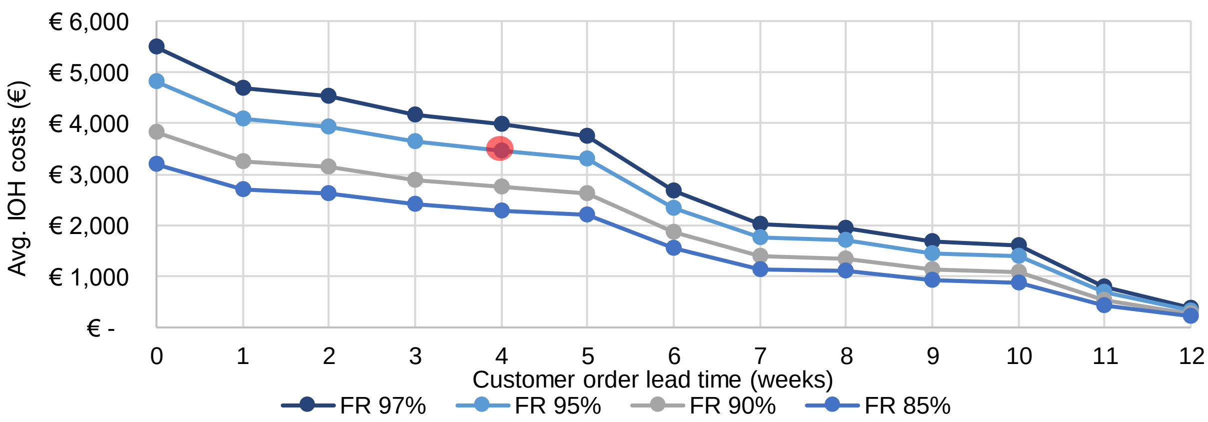


Figure 37. Trade-off curves inventory on-hand costs (buffer level costs) (μ = 2, σ = 1.5)

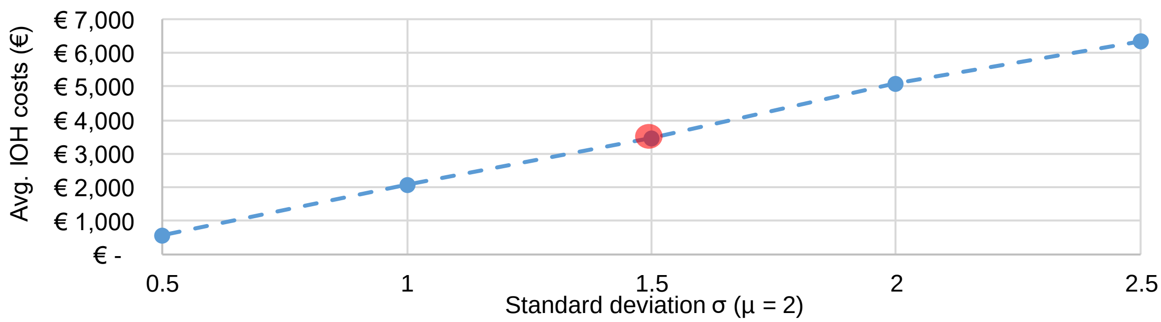


Figure 38. Impact standard deviation on IOH costs: scenario fill rate 95% and customer order lead time 4 weeks

When a specific scenario in terms of customer order lead time and required fill rate is chosen from the trade-off curves, the corresponding echelon base-stock levels are checked. For example, when customer order lead time is 4 weeks and fill rate is 95% the echelon base-stock levels based on Diks & De Kok algorithm are (s1, s2, s3, s4) = (102.1, 45.8, 21.7, 11.9). As shown in Figure 39, the local base-stock level is used cover demand over the lead time and variation in demand.

The buffer to cover the review period is also included in the stock for demand variation to avoid large negative local base-stock levels after subtracting the demand over the lead time. This enables to make conclusion about how many buffers should additional planned above the expected demand over the lead time.

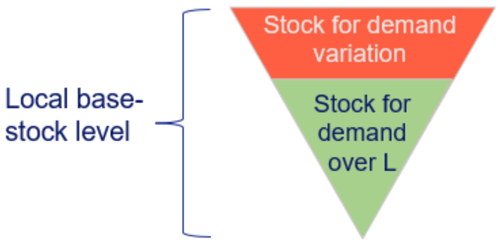


Figure 39. Local base-stock level coverage

The echelon base-stock levels should be translated to local base-stock levels with the following formula (for nonincreasing sj): $s'_j = s_j - s_{j+1}$. When the echelon base-stock levels are increasing, the downstream base-stock level should be adjusted to the previous echelon base-stock level, which is lower. After obtaining these local levels, the demand over the lead time should be subtracted in order to create the local base-stock levels for demand variation. When the local base-stock level is negative, it should be adjusted to 0. The local buffer base-stock levels can be translated back to echelon buffer base-stock levels. The translation from echelon base-stock levels to echelon buffer base-stock levels is shown in Table 11. In Figure 40 the buffer base-stock levels are visualized in echelon perspective. The echelon base-stock levels decrease over time, which means that uncertainty is decreasing over time. For this same scenario, the impact of the standard deviation on the local buffer base-stock levels are analyzed. The local buffer base-stock levels and value rapidly declining when the standard deviation decreases, which is visualized in Figure 41 and Figure 42.

Table 11. Translation echelon base-stock levels to echelon buffer base-stock levels: fill rate 95% and CLT 4 weeks

Hedging position	**1**	**2**	**3**	**4**
Echelon base-stock levels (s1, s2, s3, s4)	102.1	45.8	21.7	11.9
Local base-stock levels (s'1, s'2, s'3, s'4)	56.3	24.1	9.8	11.9
Demand over lead time (µ * Lj)	40	20	10	2
Local buffer base-stock level for demand variation	16.3	4.1	0 (-0.2)	9.9
Echelon buffer base-stock level for demand variation	30.3	14	9.9	9.9

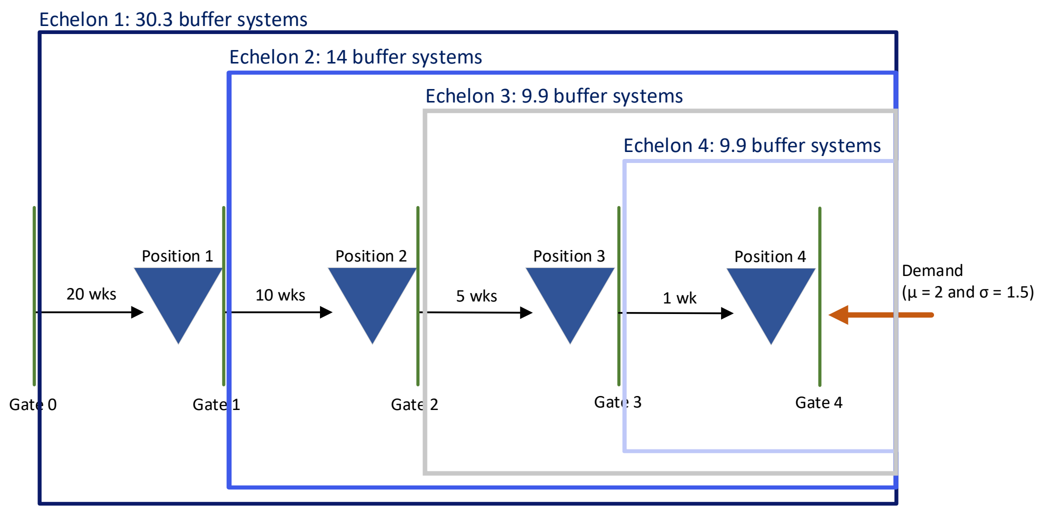


Figure 40. Echelon buffer base-stock levels

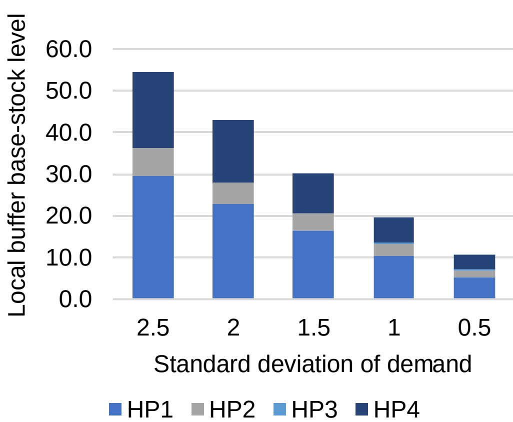


Figure 41. Impact standard deviation on buffers

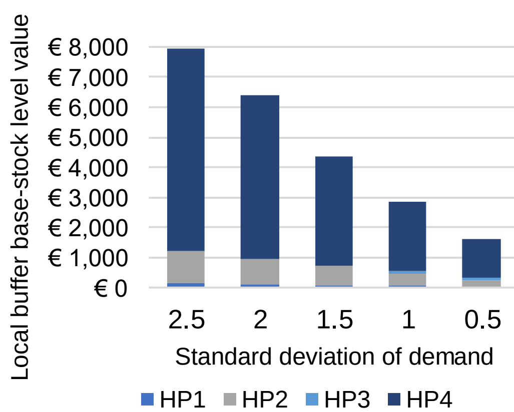


Figure 42. Impact standard deviation on buffer values

6. Buffer planning workflow

This chapter describes how the developed buffer model in the previous chapter can be implemented in the buffer planning workflow and is stated which adjustments should be made. The buffer planning workflow consist of the following steps: 1. Defining business objectives, 2. Analysis hedging scenarios, 3. Decision hedging scenario, 4. Deployment buffers, and 5. Review buffers. Each step will be explained in more detail below, with associated S&OP meeting structure. This structure and planning processes are explained in more detail in section 4.2.

<u>1. Defining business objectives</u>

The business objectives should be clear in order to make a decision that is aligned with this strategy. Business objectives could be related to the ambition to increase market share for a certain lithography system type. This market share ambition is important when there is market competition. On the other hand, when market competition exists, costs are also an important factor. This information can be acquired from the long-term market review meeting. Furthermore, the volatile market has impact on the market share and costs ambition as well. Finally, when the market for lithography systems grows, the number of system outputs should increase as well. This information can be acquired from the long-term capacity planning meeting. All these aspects have influence on which service level (fill rate) is required and which costs could be spent. It is important that these business objectives are known for each lithography system platform.

<u>2. Analysis hedging scenarios</u>

The commitment value curve is analyzed, and hedging positions are determined per lithography system platform. The input parameters μ, σ, customer order lead times and fill rate targets are collected in order to create different scenarios. With these scenarios, the buffer model is applied to develop informative graphs and trade-off curves between inventory investment, fill rate, customer order lead time, and demand variability. Examples are shown in section 5.5. This new buffer model and performed analysis differs from the current situation explained in section 4.1.

<u>3. Decision hedging scenario</u>

In the S&OP meeting, information about service level targets and available budget is collected from the business line and operations. This means that the point on the trade-off curve is jointly decided, i.e. fill rate and associated total supply chain costs, demand parameters, and customer order lead time. The trade-off curves and informative graphs developed in step 2 enables this decision-making process and result in concisions decisions. Based on the decided point on the trade-off curve, the number of system buffers are determined.

<u>4. Deployment buffers</u>

When the point on the trade-off curve is decided, the associated buffer base-stock levels are determined and discussed in the S&OP deployment meeting. These buffer base-stock levels are guidelines for the end item planners at which quantity and position of buffer systems in the supply plan should be implemented. These buffer systems should be planned above the expected demand over the lead time.

5. Review buffers

Based on the developed supply plan, the echelon base-stock levels could be checked at each position and reviewed with the echelon base-stock levels that are required to obtain the decided fill rate with associated costs. This is required to track if the guidelines are followed, and if there are deviations from the guidelines it is important to identify the reasons. Furthermore, it is important to calculate the performance of the system buffers after the actual demand occurred. This information could be used as feedback to stakeholders and decision-makers when the buffer planning workflow is repeated.

The buffer planning workflow is an iterative process. This iterative process should be performed with a certain frequency. To define this frequency, a trade-off should be made between the effort required to develop the trade-off curves and the frequency of changes in business objectives and input parameters. The most important parameters are regarding demand, where μ can change, σ can change, or both μ and σ can change. The effort for the analyses is considered as reasonable. The business objectives and input parameters are defined for a medium to long horizon. Moreover, the effectuation of the buffers levels takes a long time due to the long lead times. Therefore, the advice is to perform the buffer planning workflow on a quarterly basis. This is to ensure that the buffer levels are gradually adjusted. Whereas, step 5, review buffers, can be repeated each monthly plan cycle.

7. Conclusion

In this chapter the conclusion of this master thesis project on multi-echelon inventory control in a high complexity, high value and low volume environment is presented. In the first chapter of this thesis the research design is formulated based on the problem description. In order to achieve the research goal and to solve the problem, the following research question is formulated: *"Which buffer planning model and workflow could address demand uncertainty in order to maximize service level against affordable costs and applicable risks?"*. In order to provide an answer to this research question, the conclusion of each sub question is described below.

1. What is the concept "buffering" and why is it required?
Buffering is a common strategy for dealing with uncertainty. According to Nahmias & Olsen (2015), buffering is defined as "maintaining excess resources to cover for fluctuations in supply or demand". Hopp & Spearman (2000) name this buffering concept as variability buffering law, which means that variability in a production system will be buffered by some combination of inventory, capacity, and time. The main idea of this buffering law is the concept "pay me now or pay me later", which indicate that if you do not pay to reduce variability, you will have to pay in different ways. The buffer strategy determines the way how customer's demand will be met regardless of the variations across the supply chain. Safety stock provides protection to irregularity or uncertainty. Whereas, pipeline safety stock could be utilized to hedge against stockouts or undesirable backlogs. Hedging is a master scheduling tactic to decide on order commitment based on uncertainty and cost commitments. Two phenomena that have led to hedging concept are that a demand forecast in far future is less reliable and the commitment costs are increasing time-phased.

2. What is the uncertainty in demand and risks for tactical planning at ASML?
Master planning focuses on the problem of what, where and how to produce, on a product type or family aggregated level. The output will be planned production and inventory quantities. The main objective is to effectively coordinate the material flows across multiple stages in the supply chain to meet customer demand (Mönch, Uzsoy, & Fowler, 2018). According to literature, demand uncertainty is positioned as environmental uncertainty with random variation due to the fact that demand uncertainty is outside the manufacturing systems' boundary and beyond companies' control. ASML experiences demand uncertainty in terms of quantity due to frequent up- and downturns of the lithography market and customer requirements. Based on the coefficient of variation, it is observed that there is high variability in ASML's demand figures. The risk of order commitments are high due to the fact that product costs are extremely high, 90% of the product costs are sourced from OEM suppliers, and supplier lead times are long.

3. How to formulate a buffer planning model and implement this with a workflow and IT?
A buffer planning model and workflow results in managing uncertainty in a systematic manner and hence prepare ASML for dealing with future scenarios. The buffer planning model should be applicable for the assemble-to-order production environment and relatable to the current way of planning and buffering (hedging tactic) end items. Furthermore, it should be connected to the current S&OP meeting structure. The model should be able to make a trade-off between service level and costs in order to use it for decision-support. Next, a workflow should be determined how and when the buffer planning model should be applied.

4. How is the model and workflow organized in the current situation?
On end item level a hedge in the supply plan could be made, which is called “system buffer”. This system buffer is a system start with a reserved start slot which is planned above the firm (expected) demand. The position of this system buffer (hedge) in the supply plan determines where the buffer is created, i.e. which components, (sub)modules are buffered. There is a tool to determine the system buffer levels; however, drawbacks of this tool exists. For example, it does not reflect reality, there is no option to evaluate scenarios with high or low demand variability, and it is overdesigned. Therefore, the buffer levels are mainly determined based on mental models of employees. Lithography system starts are planned for firm demand and system buffers are added to cover upside demand. When there is no customer order allocated to a system buffer *x* weeks before FASY start, there is a meeting to decide how to continue. It can be possible that system buffers are removed from the supply plan and canceled at supplier or can be pushed out to a different start week. In both situations extra upstream inventory is created at suppliers. There are different meetings and procedures where planning information is exchanged and planning decisions are taken, e.g. S&OP meeting where capability is reviewed and information about service level targets and available inventory budget is collected from the business line.

5. Which model could address demand uncertainty and enables to make a trade-off between customer service level, costs and risks?
In the supply plan, the final assembly of systems are planned in a specific week in the future. This could be translated to a number of planned systems in the pipeline towards FASY start. Therefore, ASML’s system planning on tactical level could be seen as a serial system with a number of stages. Hedging is applied because system buffers are planned to cover variation in demand. In the serial system, the stages could be seen as hedging positions. At these hedging positions, the number of planned systems could be controlled in terms of regular and buffer system starts, i.e. a hedge could be made at that position to plan system buffers to cope with variability in demand. The inventory control policy that can be applied in a serial system is base-stock policy. At every point in time, the inventory on-order position should be equal to this base-stock level.

To determine the number and place of the hedging positions, different aspects should be taken into account. First, Rosling (1989) has proven that an assembly system is equivalent to a pure serial system. The positions in the serial system derived from the assembly system represent when ordering decisions should be made for items and it creates the timing of assembly starts. Second, the commitment value curve visualizes the cumulative added value of component procurement commitments and costs related to manufacturing activities. Based on this cumulative value curve, the large jumps in value could be identified. Before these large value jumps, hedging positions could be placed in order to create a buffer before a high value should be committed. Finally, the customer order lead time determines the decoupling point between the forecast-driven supply chain and customer-order-driven supply chain. This means that the last hedging position is always just before the customer order lead time upstream in the supply chain.

In this research, two analytical methods that are able to calculate near-optimal base-stock levels to control serial systems. The first method is the Shang & Song heuristic and is developed by Shang & Song (2003). This heuristic calculates an upper and lower echelon base-stock level bound, subsequently the near-optimal echelon base-stock level is approximated by the average of the two bounds. The bounds are calculated based on the fact that if there are two positions (j-

1 and j) where the local inventory holding costs are equal, position j-1 will immediately put items in transit to position j. This is logical as there are no costs involved when transferring inventory from position j-1 to j and the inventory can be closer positioned to the customer. On the other hand, a multi-item multi-echelon system can be controlled by the synchronized base stock-policies (SBS). De Kok & Visschers (1999) developed a method to decompose the assembly system into a divergent multi-echelon system where the pre-allocation of common components of end items is the key idea. For convergent inventory systems, the SBS polices are optimal as these are equivalent to the transformation from assembly to serial system by Rosling (1989). Diks & De Kok (1999) analyzes a divergent N-echelon inventory system and developed an algorithm that determines the near-optimal echelon base-stock levels for minimal long-run average total costs Diks & De Kok algorithm approximates the optimal allocation by two linear allocation functions, and is implemented in the tool ChainScope.

6. How can the performance of the model be measured?
A general simulation approach is developed in order to evaluate the two analytical methods stated in the previous sub question, validate the echelon base-stock policy concept, and evaluate the buffer planning at ASML. First, the demand parameters are determined. Hereafter, the hedging scenario is determined. Third, the echelon base-stock levels are calculated with one of the analytical methods. Subsequently, random demand for warm-up period and random actual demand is generated. Next, the serial system with base-stock control policy is simulated. Finally, the performance measure report is generated which includes average fill and ready rate, average backorders, and average inventory in-transit and on-hand inventory in terms of units and costs.

Based on extensive experiments with different hedging scenarios, it can be concluded that the Shang & Song heuristic is able to determine echelon-base stock levels which results in a ready rate closer to the set target ready rate than Diks & De Kok algorithm. However, Diks & De Kok algorithm determines better cost optimal echelon base-stock levels as the required costs to increase the ready rate by 1% compared to ready rate of the other analytical method are significant lower. Furthermore, for a fill rate target Diks & De Kok algorithm outperforms the Shang & Song heuristic significantly on both aspects. Hence, the Diks & De Kok algorithm is able to determine better cost optimal echelon base-stock levels which could achieve a fill rate closest to the set target, the mean absolute percentage error for a 95% target is only 0.9%. The difference between the target service level and actual service level provided by simulation can be explained by different reasons. Shang & Song (2003) evaluated the heuristic only for high demand rates and short lead times with different demand distributions; therefore, the performance is decreased for extremely low demand per time unit. For the Diks & De Kok algorithm, the fact that in the simulation there is a low demand per time unit combined with discrete demand is a source of deviations for the algorithm as this uses a continuous demand distribution. When the service level decreases, the effect of discrete demand is strengthened.

When analyzing ASML's supply plan, the percentage error between the current and desired fill rate of system type A is -24.3%. Furthermore, an additional inventory investment of factor 2.1 with respect to the current investment is required to obtain the desired fill rate. This inventory investment is especially required at the last hedging position. However, the local base-stock levels should primarily be increased at the first and last hedging position to obtain the desired fill rate.

7. What are the optimal buffer decisions regarding the type, position, and number of buffers for a specified service level objective?

To make optimal buffer decisions, it is important that the relationship between customer order lead time, fill rate, demand variability, and costs are analyzed. Trade-off curves for multiple scenarios are developed in order to visualize the impact on costs, example Figure 37. The impact of customer order lead time on the costs is high, i.e. the total supply chain costs decrease significantly when the customer order lead time increases. This is due to the fact that inventory allocation is shifting upstream and the added value of a single system. When fill rate decreases, less buffers are required due to the allowance of backorders, and hence lower inventory investments have to be made. When analyzing the impact of standard deviation for a scenario with 95% fill rate and 4 weeks customer order lead time, the average inventory on-hand costs vary between 573 euro and 6341 euro with standard deviation between 0.5 and 2.5 respectively. This shows that the standard deviation has large impact on the costs, as this determines the variability to buffer against. When a specific scenario in terms of customer order lead time and required fill rate is chosen from the trade-off curves, the corresponding echelon base-stock levels are checked and translated to echelon buffer base-stock levels. These are decreasing over time, which means that uncertainty decreases over time.

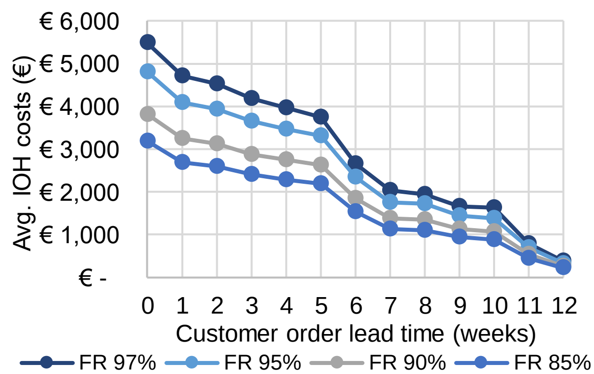


Figure 37. Trade-off curves buffer level costs (σ = 1.5)

8. How should the workflow be organized?

The buffer planning workflow consist of the following steps: 1. Define business objectives, 2. Analysis hedging scenarios, 3. Decision hedging scenario, 4. Deployment buffers, and 5. Review buffers. First, the business objectives and market insights should be clear in order to make a decision that is aligned with this strategy. These aspects have influence on which fill rate is required and which costs could be spent. Second, the commitment value curve is analyzed, and hedging scenarios per lithography system platform are defined. Then, the buffer planning model is applied to position inventory investments at the best position in the supply chain to cope with demand variability. Informative graphs and trade-off curves between inventory investment, fill rate, customer order lead time, and demand variability can be created to support decision-making. Hereafter, information about service level targets and available budget is collected from the business line and operations. This means that the point on the trade-off curve is jointly decided and corresponding buffer base-stock levels are determined. These buffer base-stock levels are guidelines for the end item planners to create the supply plan. After deployment, this plan is reviewed in order to track the performance. The buffer planning workflow is an iterative process and should be performed with a quarterly frequency to ensure that buffer levels are gradually adjusted.

Practical learnings

In this research, which is focused on multi-echelon inventory control in a practical application, multiple complexities were experienced. Understanding the system and buffer planning that is used in practice, and translating that to a known inventory policy was the most complex. Especially due to the fact that changes in the supply plan are required because of the high demand uncertainty, i.e. unexpected future demand, or customers that cancel, postpone or bring forward orders. The changes in the supply plan to deal with this can be pushing out, pull in or removing systems in the supply plan. Moreover, limited research is conducted that focuses on buffering in an environment with high demand uncertainty, complex product structure, high commitment costs, and long lead times. Therefore, the translation to a known inventory policy was of complex nature.

However, when it was identified that the hedging concept was used to create buffers in the supply chain and that the position of the hedge determines the items that are buffered, the combination with a system topology could be made. The serial system concept is commonly applied in supply chains where each stage represents a company where inventory is kept. For example, a supply chain with factory, factory warehouse, wholesaler, and retailer. However, this concept is now applied within one company to control the order commitments at different hedging positions, which is an important learning from this research. Finally, understanding the characteristics of the forecast- and customer-order-driven supply chain and impact of customer order lead time was important to understand supply chain learnings. The first supply chain learning is that inventory is primary concentrated at the customer order decoupling point because of the direct contribution to customer service level. Moreover, the second learning is that cheap items with long lead times should be buffered. It was challenging to understand and communicate these two most important supply chain learnings.

8. Discussion

8.1 Scientific contribution

In literature, the applicability of the serial system concept is commonly in supply chains where each stage represents a company where inventory is kept. For example, a supply chain with factory, factory warehouse, wholesaler, and retailer, where each stage operates under a local base-stock policy. In this case, no extra value might be added to the product, but the holding costs increases. In this research a translation has been made from a master production schedule to the concept of serial systems by using the assembly system. Furthermore, hedging could be applied to manipulate on future forecasted end item outputs. The hedging concept was conceived in the late 70s and has been evolved in the 80s and 90s. However, from that time the hedging tactic has been rarely discussed in literature. The combination of a serial system and hedging technique has led to a multi-echelon inventory control model that lowers inventory investments while dealing with demand uncertainty. For this multi-echelon inventory control model, two analytical models from literature that determines the echelon base-stock levels has been evaluated and compared.

It can be concluded that the Shang & Song heuristic is able to determine echelon base-stock levels that result in a ready rate closer to the ready rate target. However, the Diks & De Kok algorithm determine better cost optimal base-stock levels due to inventory allocation. Furthermore, the Diks & De Kok algorithm are able to determine echelon base-stock levels that result in a fill rate close to the fill rate target. In literature, limited papers (based on the forward snowballing technique) are available that apply and evaluate of the Shang & Song heuristic. For example, Sari (2007) applied the Shang & Song heuristic to calculate the echelon base-stock level for a two-echelon system, retailer and distributor; however, no results were shown regarding performance. Furthermore, generic algorithms are evaluated focused on a system with supplier, manufacturer, warehouse, and retailer with short lead times and high demand rates (Daniel & Rajendran, 2006). Moreover, Gupta (2006) extended the heuristic with a production rate for a situation with short lead times and only linear increasing holding costs. Next, the serial system heuristic is applied and hereafter decomposed into a two-echelon distribution system with a high demand rate and short lead times (Rong, Atan, & Snyder, 2017). No papers are focused on serial systems with a low demand rate, long lead times, and large value jumps.

A proof of concept is built for this new multi-echelon inventory control model where the hedging tactic is combined with a serial system for end items. This proof of concept is demonstrated with actual data from ASML. It is experienced that in practice it is possible to determine the required parameters, namely cost structure and supplier lead times. Furthermore, it is able to fit actual low volume demand data to a theoretical demand discrete distribution. The multi-echelon inventory control model enables to evaluate the current situation and determines the desired situation or generates trade-off curves to decide on this situation. Furthermore, the developed model is customized for the low value production environment as the base-stock, levels and therefore output values are integers. This feature and low demand numbers resulted in interesting results regarding the ready rate and fill rate. The ready rate was in most of the analyzed scenarios significantly lower than the fill rate. However, when the service level increases, the gap between ready rate and fill rate is becoming smaller. This is different from the continuous inventory models, where the ready rate and fill rate are always close together.

8.2 Limitations

In this research, a scope has been set and several assumptions have been made to complete this master thesis project within the specified time horizon. However, setting a scope and making assumptions leads to research limitations. The developed buffer planning model focuses on demand uncertainty for the forecast-driven supply chain and it is assumed that the lead time of suppliers and operational activities are deterministic. Furthermore, the assumption is that suppliers of components and (sub)modules have enough production capacity, i.e. when an order is placed at a supplier, the order can be delivered the lead time later. In this research, decoupling points of suppliers are not taken into account to define the hedging positions. Hence, inventory cannot exactly be positioned at the hedging positions and have to flow downstream to the first suppliers' decoupling point. Moreover, the main serial system used in this research for comparison of the two analytical methods, evaluation and analyses was four-echelon. To a minor extend, the two- and three-echelon serial system have been analyzed to identify the impact on costs when increasing the customer order lead time. The lead times and cost structure remains equal in this research, whereas the target service level and demand parameters (σ, μ) are varied. The number of hedging positions, cost structure, demand parameters and target service level is for each end item different. Variety in parameter values could result in echelon base-stock levels that achieve a service level closer to the target.

8.3 Future research

As this multi-echelon inventory control model takes into account demand uncertainty while assuming deterministic supply lead times, it is interesting to investigate how this model can be adjusted in order to deal with supply uncertainty. The model for ASML is evaluated with new lithography system demand. However, extra demand streams will be added when the level is changed from system to (sub)modules. These demand streams arise due to refurbished systems, system upgrades, service and spare parts. Supply uncertainty consist of variation in the supply lead time and yield uncertainty (defects). Additional analysis could be performed to identify this uncertainty and to determine the hedge that should be made to deal with this uncertainty in addition to the demand uncertainty hedge. Hence, the echelon buffer base-stock levels will increase.

However, when the translation is made from system level to (sub)module level, the commonality in (sub)modules of different lithography systems types can result in risk pooling. Risk pooling result in less inventory that is required due to demand aggregation. When commonality exists between system types, the network topology can be defined as divergent systems. For this system topology, the algorithm developed by Diks & De Kok (1999) is able to determine the near-optimal echelon base-stock levels. Risk pooling will result in a decrease of the echelon base-stock levels as the variance in demand or supply are aggregated and therefore decreased. These commonality buffers are called flexibility buffers and the function is defined according to Hopp & Spearman (2000) as: "flexibility reduces the amount of variability buffering required in a production system". The postponement strategies of suppliers are another example of the flexibility buffer, and should be investigated. When suppliers are able to postpone the point at which the personality of the product is configured, flexibility to change the demand for multiple products is created (Atan & Jaarsveld, 2019). When this flexibility exists, less buffers are required.

Furthermore, it should be investigated if the model could take the capacity constraints of suppliers into account, as these suppliers are dependent on their production facility and n-tier suppliers. In the current model, decisions are based on central information. When information is exchanged and shared, better decision and collobariation might be achieved. However, as there are up to 100 supply chains per lithograpy system required, is it not feasible to create individual alignment with each supplier. Therefore, priority has to be set to the most critical modules and corresponding suppliers. The echelon base-stock policy and hedging tactic will coordinate and control the other supply chains. To cope with demand fluctuations while dealing with supply capacity constraints, it is important to examine the use of anticipation stock. This means when there is underutilization in the factory, lithography systems should be prebuilded in order to cope with the overutilization in the future.

One important parameter to fill in the model is demand, which are the expected demand and demand variability. However, when the forecast is not reliable, historic demand parameters should be used or an estimation should be made. Therefore, future research should be focused on the demand distribution, demand patterns, demand signals, and forecast methods. Finally, research can be conducted on the product lifecycle, i.e. how to deal with phase-in and phase-out of lithography systems and the impact on buffers.

Bibliography

Adan, I., Eenige, M. v., & Resing, J. (1994). Fitting discrete distributions on the first two moments. *Memorandum COSOR, 9436*.

Albrecht, M., Rohde, J., & Wagner, M. (2015). Master Planning. In H. Stadtler, C. Kilger, & H. Meyr, *Supply Chain Management and Advanced Planning – Concepts,* (pp. 155-176). Berlin: Springer.

ASML. (2018). DUV Products and Business Opportunity. Eindhoven. Opgehaald van https://www.asml.com/-/media/asml/files/investors/financial-calender/inverstor-days/2018/asml_20181108_06_asml_investor_day_2018_duv_products_and_business_opportunity_rkool.pdf

ASML. (2019). *ASML at a glance*. Opgehaald van ASML: https://www.asml.com/en/company/about-asml/asml-at-a-glance

ASML. (2019). *Integrated Report.* Veldhoven. Opgehaald van https://www.asml.com/en/investors/annual-report/2019

ASML. (2019). *Products*. Opgehaald van ASML: https://www.asml.com/en/products

ASML SCM. (2019, September). *ASML - Supply Chain Management*. Opgehaald van YouTube: https://www.youtube.com/watch?v=7JwSF_6sdyo

Atan, Z., & Jaarsveld, W. v. (2019). Multi-echelon inventory management. Opgehaald van TUe lectures notes course 1CM100 Multi-Echelon Inventory Management

Atan, Z., Ahmadi, T., Stegehuis, C., De Kok, A., & Adan, I. (2017). Assembly-to-order systems: A review. *European Journal of Operations Research, 261*, 866-879.

Bartezzaghi, E., & Verganti, R. (1995). Managing demand uncertainty through order overplanning. *International Journal of Production Economics*, 107-120.

Broekmeulen, R., & Donselaar, K. (2013). DoBr tool. Eindhoven.

Clark, A., & Scarf, H. (1960). Optimal Policies for a Multi-Echelon Inventory Problem. *Management Science, 6*, 475-490.

Daniel, J., & Rajendran, C. (2006). Heuristic approaches to determine base-stock levels in a serial supply chain with a single objective and with multiple objectives. *Europoean Journal of Operational Research, 175*(1), 566-592.

De Kok, A. (2015). Buffering Against Uncertainty in High-Tech Supply Chains. *Winter Simulation Conference*, 2991-3000.

De Kok, A. (2020, March). Interview ChainScope and simulation. (T. Mantje, Interviewer)

De Kok, A., & Fransoo, J. (2003). Planning Supply Chain Operations: Definition and Comparison of Planning Concepts. *Handbooks in Operations Research and Management Science: Supply Chain Management*, 597-675.

De Kok, A., & Visschers, J. (1999). Analysis of assembly systems with service level constraints. *International Journal of Production Economics, 59*, 313-326.

Diks, E., & De Kok, A. (1999). Computational results for the control of a divergent N-echelon inventory system. *International Journal of Production Economics, 59*, 327-336.

Esmaeilikia, M., Fahimnia, B., Sarkis, J., Govindan, K., Kumar, A., & Mo, J. (2016). Tactical supply chain planning models with inherent flexibility: definition and review. *Annals of Operations Research, 244*(2), 407-427.

Fleischmann, B., Meyr, H., & Wagner, M. (2015). Advanced Planning. In S. Hartmut, C. Kilger, & H. Meyr, *Supply Chain Management and Advanced Planning* (Fifth ed.). Berlin: Springer.

Gupta, D. (2006). Performance Evaluation and Stock Allocation in Capacitated Serial Supply Systems. *Manufacturing & Service Operations Management, 8*(2), 169-191.

Hillier, M. (2000). Component commonality in multiple-period, assemble-to-order systems. *IIE Transactions, 32*(8), 755-766.

Ho, C. (1989). Evaluating the impact of operating environments on MRP system nervousness. *International Journal of Production Research, 27*(7), 1115-1135.

Hopp, W., & Spearman, M. (2000). *Factory Physics.* New York: McGraw-Hill Higher Education.

Jacobs, F., Berry, W., Whybark, D., & Vollmann, T. (2011). *Manufacturing Planning and Control for Supply Chain Management* (Sixth ed.). New York: McGraw-Hill Irwin.

Miller, J. (1979). Hedging the master schedule. In L. Ritzman, L. Krajewski, W. Berry, S. Goodman, S. Hardy, & L. Vitt, *Disaggregation: Problems in Manufacturing and Service Organizations* (pp. 237-256). Dordrecht: Springer Science Business Media.

Mönch, L., Uzsoy, R., & Fowler, J. (2018). A survey of semiconductor supply chain models part III: master planning, production planning, and demand fulfilment. *International Journal of Production Research*, 4565-4584.

Nahmias, S., & Olsen, T. (2015). *Production and Operations Analysis.* United States of America: Waveland Press.

Ofxord University Press. (2019, October). Opgehaald van Oxford Learner's Dictionaries: https://www.oxfordlearnersdictionaries.com/definition/english/hedge-against

Rong, Y., Atan, Z., & Snyder, L. (2017). Heuristics for Base-Stock Levels in Multi-Echelon Distribution Networks. *Production and Operations Management, 26*(9), 1760-1777.

Rosling, K. (1989). Optimal Inventory Policies for Assembly Systems under Random Demands. *Operations Research, 37*(4), 565-579.

Sari, K. (2007). Exploring the benefits of vendor management inventory. *International Journal of Physical Distribution & Logistics Management, 37*(7), 529-545.

Shang, K., & Song, J.-S. (2003). Optimal Policies in Serial Supply Chains. *Management Science INFORMS*, 618-638.

Silver, E., Pyke, D., & Peterson, R. (1998). *Inventory Management and Production Planning and Scheduling* (3rd edition ed.). Willey.

Song, J., & Zipkin, P. (2003). Supply Chain Operations: Assemble-to-Order Systems. In A. de Kok, & S. Graves, *Handbooks in OR & MS* (Vol. 11, pp. 561-596). Elsevier B.V.

The Information Network. (2018, February 12). *ASML increases dominance of lithography market*. Opgeroepen op September 2019, van eeNews: https://www.eenewsanalog.com/news/asml-increases-dominance-lithography-market

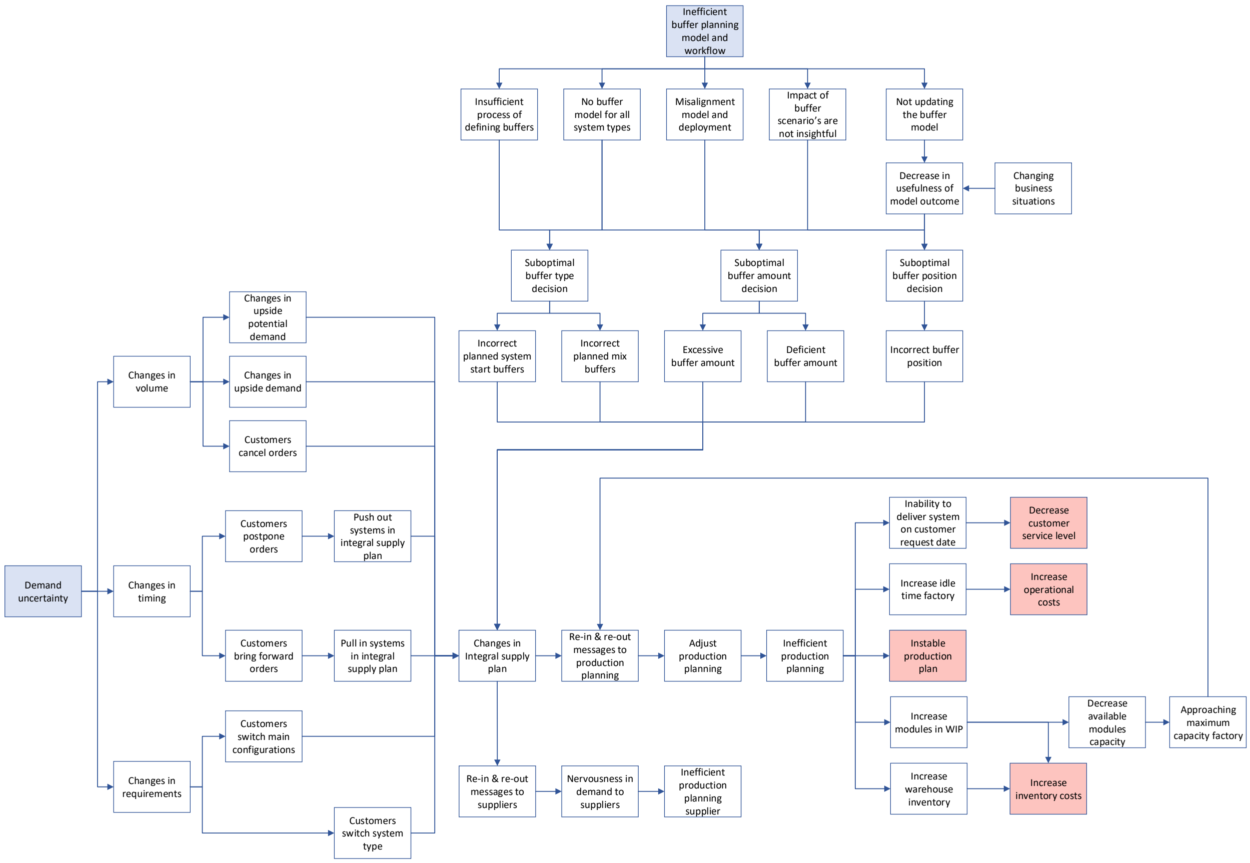
van Aken, J., & Berends, H. (2018). *Problem Solving in Organizations: A Methodological Handbook for Business and Management Students.* United Kingdom: Cambridge University Press.

Wazed, M., Ahmed, S., & Yusoff, N. (2009). Uncertainty Factors in Real Manufacturing Environment. *Australian Journal of Basic and Applied Sciences, 3*(2), 342-351.

Wiers, V., & De Kok, A. (2018). *Designing, Selecting, Implementing and Using APS Systems.* Cham: Springer.

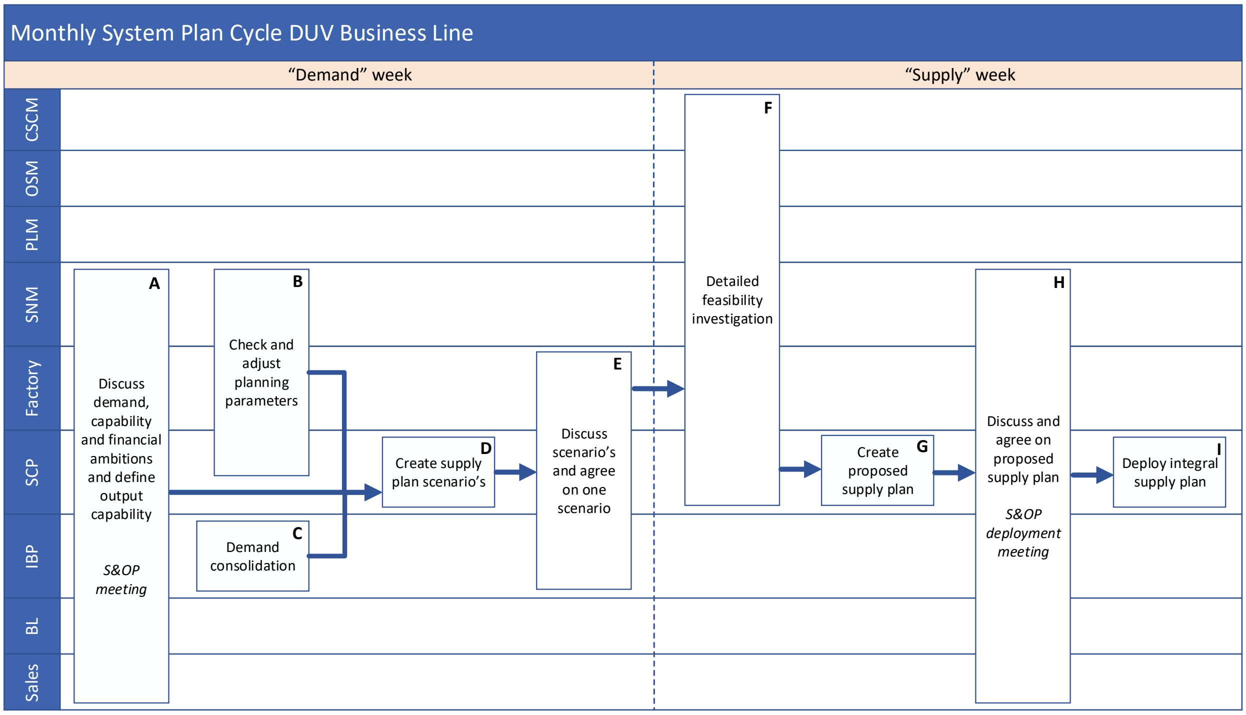
Appendices

Appendix I: Cause and effect diagram

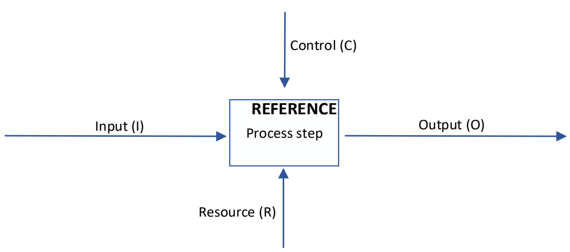


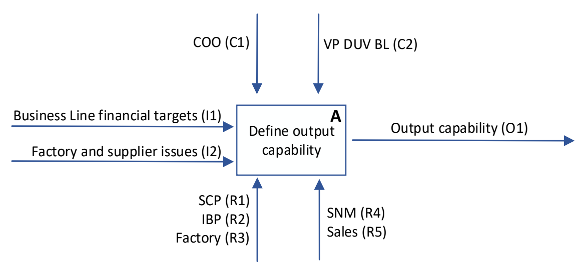
Appendix II: Tactical planning process and IDEF schemes

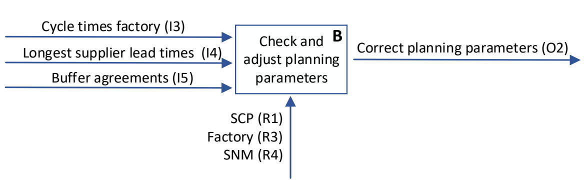
The IDEF schemes that correspond to the steps in the plan cycle overview can be found with a reference letter.

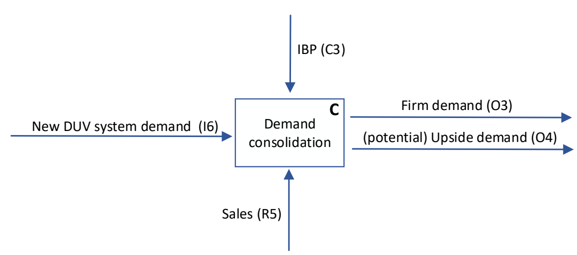


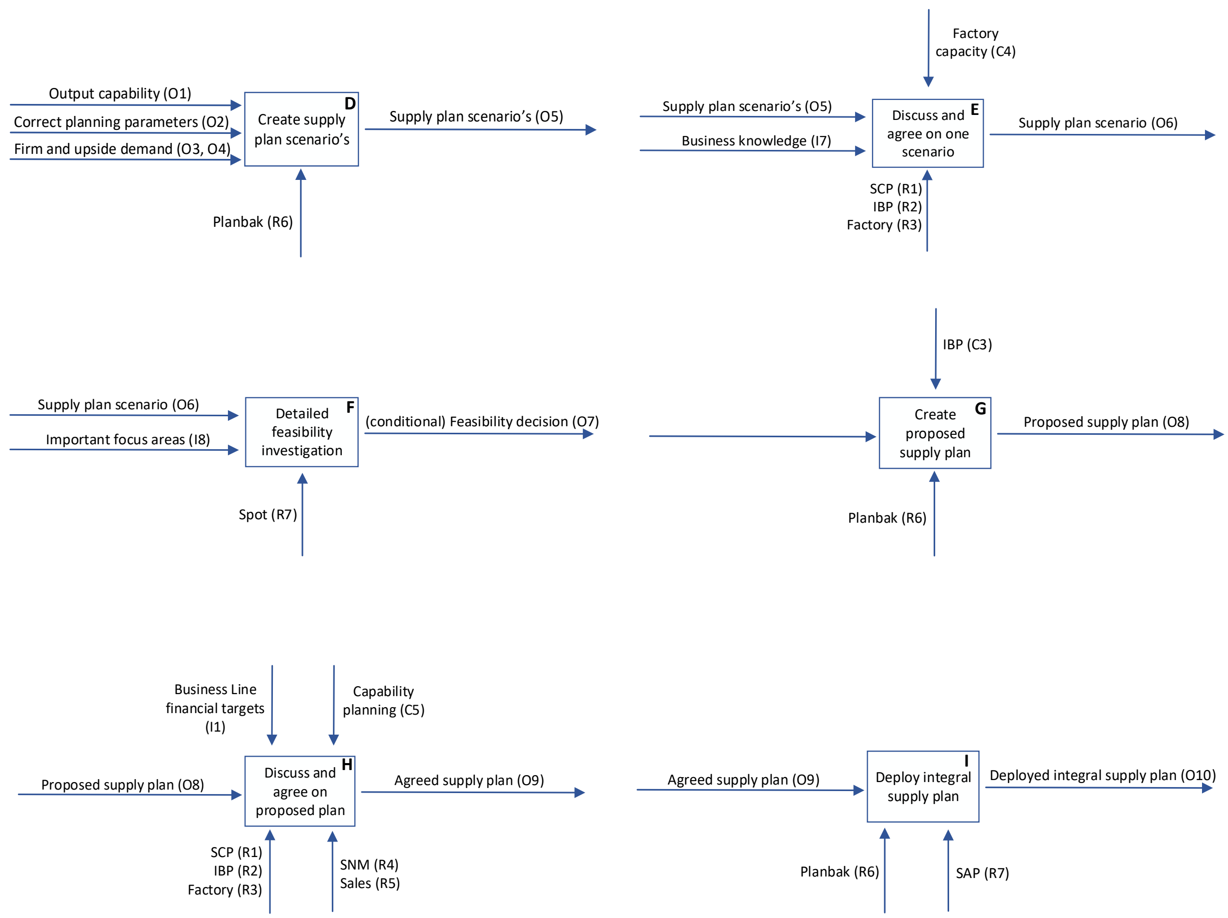
Legend abbreviations:
CSCM = Customer Supply Chain Management
OSM = Operational Supplier Management
PLM = Product Lifecycle Management
SNM = Supplier Network Management
BL DUV = Business Line DUV
Planning parameters = cycle times FASY, buffer agreements
ICT = Integral Cycle Time = longest lead time + production cycle time










Output capability (O1)
Correct planning parameters (O2)
Firm and upside demand (O3, O4)
Create supply plan scenario's
D
Supply plan scenario's (O5)
Planbak (R6)
Factory capacity (C4)
Supply plan scenario's (O5)
Business knowledge (I7)
Discuss and agree on one scenario
E
Supply plan scenario (O6)
SCP (R1)
IBP (R2)
Factory (R3)
Supply plan scenario (O6)
Important focus areas (I8)
Detailed feasibility investigation
F
(conditional) Feasibility decision (O7)
Spot (R7)
IBP (C3)
Create proposed supply plan
G
Proposed supply plan (O8)
Planbak (R6)
Business Line financial targets (I1)
Capability planning (C5)
Proposed supply plan (O8)
Discuss and agree on proposed plan
H
Agreed supply plan (O9)
SCP (R1)
IBP (R2)
Factory (R3)
SNM (R4)
Sales (R5)
Agreed supply plan (O9)
Deploy integral supply plan
I
Deployed integral supply plan (O10)
Planbak (R6)
SAP (R7)