Tampereen yliopisto

Juho-Pekka Mattila

SIMULATION OF INVENTORY INVESTMENT AND DELIVERY RELIABILITY

Comparison of configure-to-order and make-to-order

Faculty of Engineering and Natural Sciences
Master's thesis
Examiner: Jussi Heikkilä
Examiner: Mohammad Moshtari
April 2020

ABSTRACT

Juho-Pekka Mattila: Simulation of inventory investment and delivery reliability – Comparison of configure-to-order and make-to-order
Master of Science Thesis, 90 pages, 9 appendices
Tampere University
Master's Degree Program in Industrial Engineering and Management
April 2020

This thesis studied the differences between a configure-to-order (CTO) and a make-to-order (MTO) approach to order fulfillment. The aim of the study was to identify and compare the different processes' characteristics in terms of the investment in inventories and delivery reliability. The goal was to investigate what impact demand uncertainty, capacity level and supply lead time have on the processes' inventory investment and delivery reliability.

The product studied in this process is an automated material-handling solution produced by the case company. The product utilizes a modular product structure which enables customizing the solution to different customer needs. The manufacturing process currently in use is the CTO approach. It was deemed valuable to study another possible approach to identify and evaluate the advantages and disadvantages of both approaches.

The study was conducted by creating a simulation model which modelled two processes. The first mimicked the current CTO process in use at the case company and the second how the process would operate if the company utilized an MTO approach. The simulation was run in multiple scenarios which differed from each other by demand, capacity level and component supply lead times. The results of the scenarios were then analyzed to discern answers to the research questions.

The results of the study show that the inventory investment needed in CTO is many times more than in an MTO environment. This is because more buffer inventories are kept in the CTO process both at component and at the module level. However, the delivery reliability was found to be better in the CTO process. The positioning of the order penetration point in the CTO process enables a more flexible production schedule and the increased buffering means that deliveries can be fulfilled quicker. This makes the CTO process less reliant on capacity buffers and more robust against demand fluctuation.

The most important factor when determining the inventory investment needed was clearly the supply lead time. Delivery reliability was mostly dictated by the capacity level, but it was noted that increasing the capacity level over a certain threshold quickly starts to provide diminishing benefits to delivery performance.

Keywords: configure-to-order, make-to-order, simulation, delivery reliability, inventory investment

TIIVISTELMÄ

Juho-Pekka Mattila: Simulaatio pääoman sitoutumisesta ja toimitusvarmuudesta – Vertailu tilauksesta valmistamisen ja tilauksesta konfiguroimisen välillä.
Diplomityö, 90 sivua, 9 liitettä
Tampereen yliopisto
Tuotantotalouden diplomi-insinöörin tutkinto-ohjelma
Huhtikuu 2020

Tämä tutkimus tutki tilauksesta valmistavan (make-to-order, MTO) ja tilauksesta konfiguroivan (configure-to-order, CTO) toimitusprosessin eroja. Työssä tunnistettiin ja vertailtiin näiden prosessien eroja varastoihin sitoutuneessa pääomassa ja toimitusvarmuudessa. Toinen tavoite oli tutkia, miten kysynnän epävarmuus, kapasiteettipuskuri ja komponenttien toimitusaikojen muutokset vaikuttavat näihin kahteen prosessiin.

Tuote, jota tutkimuksessa tutkittiin, on automaattinen materiaalinhallintajärjestelmä, joka voidaan konfiguroida erilaisiin asiakastarpeisiin sen modulaarisen tuoterakenteen avulla. Tällä hetkellä tuotetta valmistetaan CTO-periaatteella siten, että moduuleja valmistetaan varastoon. Kun toimituspäivä lähestyy, moduulivarastosta keräillään tarvittavat moduulit lähetyksiin, mikä laukaisee täydentävien moduulien kokoamisen. Kohdeyrityksessä todettiin, että olisi mielenkiintoista tutkia vaihtoehtoisen prosessin toimintaa ja verrata sitä nyt käytössä olevaan prosessiin.

Tutkimus toteutettiin luomalla simulaatiomalli, joka jäljitteli kahta tutkittavaa prosessia. CTO-prosessi mallinnettiin vastaamaan nykyistä prosessia ja MTO-prosessilla yritettiin imitoida sitä, miten yrityksen prosessit toimisivat, jos ne toteutettaisiin MTO-periaatteiden mukaisesti. Simulaatiomallilla luotiin useita eri skenaarioita, jotka erosivat toisistaan kysyntätasojen, komponenttien toimitusaikojen, sekä kapasiteettipuskurien osalta. Skenaarioiden tuloksista pyrittiin analysoimaan vastauksia tutkimuskysymyksiin.

Tutkimuksen tulokset osoittavat, että CTO-prosessin vaatima varastopääoma on moninkertainen MTO-prosessiin verrattuna. Tämä johtuu isommista varmuusvarastoista sekä komponentti- että moduulitasolla. CTO-prosessin toimitusvarmuuden todettiin kuitenkin olevan huomattavasti parempi simulaatiossa käytetyillä muuttujilla. Sijoittamalla tilauksen kohdentamispiste lähemmäs toimitusta, voidaan osa tuotantoprosessin vaiheista tehdä ennusteperusteisesti, mikä mahdollistaa joustavamman tuotantoaikataulun. Lisäksi puskurivarastot prosessin eri vaiheissa nopeuttavat toimituksia. MTO-prosessin todettiin olevan riippuvaisempi kapasiteettipuskureista ja heikompi vastaamaan kysynnän epävarmuuteen.

Isoin vaikutus varastoon sitoutuneeseen pääomaan oli komponenttien toimitusajoilla. Toimitusvarmuuteen eniten vaikutti kapasiteettipuskurit. Toimitusvarmuuden tuloksista huomattiin kuitenkin, että tietyn raja-arvon jälkeen kapasiteettipuskurin lisäämisen luomat hyödyt vähentyvät merkittävästi.

Avainsanat: tilauspohjainen valmistus, tilauspohjainen konfigurointi, simulaatio, toimitusvarmuus, varastoon sitoutunut pääoma

PREFACE

Finishing this thesis marks the end of my studies in Tampere University. The years spent studying at TTY have been the best in my life so far, but as this part of my life comes to a close, I'm enthusiastically awaiting to see the experiences and challenges up ahead.

Writing the thesis was quite a journey. The beginning was quite slow but after several months of gradual preparative work the process started gaining speed fast. After the writing got going, it was quite smooth sailing right on to the finish despite the turbulent times the world is going through at the moment.

First and foremost, I would like to thank Konecranes for the opportunity to conduct this thesis. I also want to express my gratitude towards my coworkers and my instructor Veli-Pekka Vuoti for their patient assistance during the research. I would like to thank my family, friends and especially Lea for the encouragement and support during the process. Lastly, special thanks to my examiner Jussi Heikkilä for the guidance throughout the process. Now I would like to end this preface with a quote which fits this thesis rather well.

"All models are wrong, but some are useful."

- George E. P. Box 1976

Tampere, 1 April 2020

Juho-Pekka Mattila

CONTENTS

1. INTRODUCTION 1
1.1 Research background and motivation 1
1.2 Research questions and goals 2
1.3 Research Approach 3
1.4 Research limitations 4
1.5 Thesis structure 5
2. THEORY BACKGROUND 7
2.1 Supply chain uncertainty 7
2.1.1 Demand uncertainty 9
2.2 Mass customization 9
2.2.1 Modular product architecture 10
2.2.2 Delayed differentiation 13
2.3 Different approaches to order fulfillment 14
2.3.1 The significance of positioning the order penetration point 16
2.3.2 Make-to-order strategy 17
2.3.3 Configure-to-order strategy 18
2.4 Theory summary 19
3. OVERVIEW OF THE AGILON-SYSTEM 21
4. THE PROCESSES 25
4.1 Configure-To-Order Process 27
4.2 Make-To-Order Process 30
4.3 Summary 31
5. SIMULATION 33
5.1 Limitations 33
5.2 Modelling uncertainty 35
5.3 Modelling the bills of material 39
5.4 Calculating key variables in the simulation 42
5.4.1 Module demand 42
5.4.2 Inventory control 47
5.5 Inputs & Outputs 48
5.6 Flow logic 49
5.6.1 Initialization & Scenario cycle 50
5.6.2 Week cycle 51
5.7 Validation 54
5.7.1 Comparison to the real system 55
5.7.2 Comparison to a deterministic model 56
5.7.3 Consistency 58
5.7.4 Stability 60
5.7.5 Module demand approximation 61
5.8 Implementation 64

5.8.1 Main simulation 64
6. RESULTS 66
6.1 Tied-up capital 66
6.1.1 Regression analysis method 69
6.1.2 Regression analysis results 72
6.1.3 Inventory investment result summary 75
6.2 Delivery reliability 76
6.2.1 Main simulation 76
6.2.2 Supplementary simulation 77
6.2.3 Capacity utilization 81
6.2.4 Summary 83
7. DISCUSSION 85
7.1 Conclusions 85
7.2 Error evaluation 87
7.3 Managerial implications 88
7.4 Further research topics 89
REFERENCES 91

APPENDIX A: THE AGILON MODULES
APPENDIX B: AGILON SYSTEM EXAMPLES
APPENDIX C : INITIALIZATION AND SCENARIO CYCLE FLOWCHART
APPENDIX D: COMPLETE CTO FLOWCHART
APPENDIX E: COMPLETE MTO FLOWCHART
APPENDIX F: SYSTEM CONFIGURATIONS AND PROBABILITIES
APPENDIX G: TABLE OF THE CORRESPONDENCE BETWEEN SYSTEM AND ROBOT DEMAND AND THE TOTAL PROBABILITIES FOR ROBOT DEMAND
APPENDIX H: EXAMPLE REGRESSION RESIDUAL ANALYSIS RESULTS
APPENDIX I: MAIN SIMULATION DELIVERY RELIABILITY RESULTS GROUPED BY THE INPUT VARIABLES

LIST OF FIGURES

Figure 1: *The order fulfillment process* ... 5
Figure 2: *Illustration of the hourglass shape created by a modular product structure (adapted from: Heikkilä et al. 2002, pp. 13)* ... 11
Figure 3: *The three types of modularity (adapted from: Ulrich & Eppinger 2012, pp. 186)* ... 12
Figure 4: *Position of the order penetration point in the different order fulfillment strategies (adapted from: Olhager 2003)* ... 15
Figure 5: *The applicable order fulfillment processes in terms of demand volatility and P/D ratio (adapted from: Olhager 2003)* ... 17
Figure 6: *An Agilon warehouse* ... 21
Figure 7: *The different modules in an Agilon system* ... 22
Figure 8: *A standard Agilon system* ... 23
Figure 9: *A retail Agilon* ... 24
Figure 10: *The factors influencing the processes (adapted from: Angkiriwang et al. 2014)* ... 25
Figure 11: *Positioning of the OPP in the processes* ... 27
Figure 12: *The outline of the CTO process* ... 28
Figure 13: *The outline of the MTO process* ... 30
Figure 14: *All possible Agilon configurations in the simulation* ... 38
Figure 15: *The distribution of components in a robot* ... 41
Figure 16: *Probability distribution of robot demand in the example scenario* ... 46
Figure 17: *Cumulative robot demand distribution function of the example* ... 46
Figure 18: *The simplified week cycle* ... 51
Figure 19: *Incoming order handling* ... 52
Figure 20: *Production & purchasing sequence* ... 52
Figure 21: *End-of-the-week sequence* ... 54
Figure 22: *Consumption & Replenishment of the example items* ... 57
Figure 23: *Consumption & Replenishment with variation* ... 58
Figure 24: *Distribution of average inventory values in the CTO consistency test* ... 59
Figure 25: *Distribution of average inventory values in the MTO consistency test* ... 59
Figure 26: *Theoretical robot demand distribution with the different system demand levels differentiated by colors* ... 63
Figure 27: *Average inventory investment comparison* ... 68
Figure 28: *Example regression model compared to the observed values* ... 72
Figure 29: *Extended delivery reliability results grouped by capacity* ... 79
Figure 30: *Impact of demand deviation on delivery reliability* ... 79
Figure 31: *Impact of the supply lead time modifier on delivery reliability* ... 80
Figure 32: *Capacity utilization in the high capacity scenarios example* ... 82
Figure 33: *Capacity utilization in the low capacity scenarios example* ... 82

LIST OF TABLES

Table 1: *The research questions* ... 2
Table 2: *Summary of differences in the processes* ... 31
Table 3: *Dimensions and probabilities of the frame types* ... 36
Table 4: *Probability of the number of access points depending on the frame size* ... 36
Table 5: *Probability of the number of robots depending on the number of access points* ... 37
Table 6: *Example calculation for configuration probability* ... 39
Table 7: *PBOM multipliers for robots and access points* ... 39
Table 8: *The component decision variables for the modules* ... 42
Table 9: *Average weekly module demand levels in a scenario where average system demand is 4.* ... 43
Table 10: *Data from the example simulation of weekly robot demand* ... 45
Table 11: *Example of the maximum dependent demands for all the modules when the maximum independent system demand is 8* ... 47
Table 12: *Comparison of the simulation with the real values.* ... 56
Table 13: *Key variables calculated by the simulation* ... 56
Table 14: *Inventory variables from the CTO stability test* ... 60
Table 15: *Inventory variables from the MTO stability test* ... 60
Table 16: *Binomial distributions for each system demand level* ... 61
Table 17: *Distributions from table 10 adjusted to consider the probability of the system demand level* ... 62
Table 18: *The level of robot demand each cell represents* ... 62
Table 19: *Comparison between the numerical and theoretical distributions* ... 63
Table 20: *Demand scenarios* ... 64
Table 21: *The limitations of secondary input variables* ... 65
Table 22: *Tied-up capital results for the CTO process* ... 66
Table 23: *Tied-up capital results for the MTO process* ... 67
Table 24: *Average inventory investment needed for one system* ... 68
Table 25: *Example regression model summary* ... 70
Table 26: *Example regression model coefficients* ... 70
Table 27: *Model characteristics and standardized coefficients for the CTO regression models* ... 73
Table 28: *Model characteristics and standardized coefficients for the MTO regression models* ... 74
Table 29: *Unstandardized coefficients for the CTO regression models* ... 74
Table 30: *Unstandardized coefficients for the MTO regression models.* ... 75
Table 31: *Delivery reliability results from the main simulation grouped by demand level* ... 76
Table 32: *Delivery times for the late deliveries in the main simulation* ... 77
Table 33: *Delivery reliability results grouped by demand level for the extended simulation* ... 78
Table 34: *Delivery times for the late deliveries in the extended simulation* ... 81
Table 35: *Key values from the capacity utilization examples* ... 81

LIST OF ABBREVIATIONS

AP	Access point
ATO	Assemble-to-order
BOM	Bill of materials
CODP	Customer order decoupling point
CTO	Configure-to-order
ERP	Enterprise resource planning
ETO	Engineer-to-order
FOQ	Fixed order quantity
MC	Mass customization
MTO	Make-to-order
MTS	Make-to-stock
OPP	Order penetration point
PBOM	Planning bill of materials

1. INTRODUCTION

1.1 Research background and motivation

As Cristopher and Holweg (2011) state, it has become almost a cliché to begin a research paper by stating how the business world is becoming increasingly global, customers are more demanding than ever and uncertainty in the markets is increasing. These changes create challenges to which companies must respond. (Christopher & Holweg 2011)

Mass customization (MC) is one answer and it has been a very popular subject for research and in practice since the late 1980s. (Fogliatto et al. 2012) The aim of mass customization is to try and gain the benefits of high volume production while providing a customized product offering (Duray et al. 2000). Common ways to achieve the benefits are for example a modular product architecture and postponing the differentiation of products (MacCarthy 2013). These methods allow differentiating products after the customer order which reduces the risk of inventory obsolescence (Yang & Burns 2003).

This thesis was done to the Agilon-department of Konecranes. Agilon is a material handling system which can be configured to fit each customers' need to provide more transparent and smooth material flows in e.g. manufacturing or maintenance equipment. The configurability makes the order fulfillment very complex as each delivery is likely to differ from the earlier deliveries and the number of needed resources varies a lot.

The current order fulfillment process in place is a type of configure-to-order (CTO) process. A CTO keeps semi-finished products or modules in inventory to provide more standardized manufacturing processes, better customer delivery speed and variety in the offering at the cost of the increased inventories (Cheng et al. 2002). The CTO process has worked well due to the dilemma of the speed of customer delivery as a marketing asset and the rather long manufacturing lead time needed to assemble the modules.

However, due to the relatively large amount of capital the CTO-process ties up in inventories, it was deemed important to compare this process to a different approach. The competing approach is a make-to-order (MTO) approach. This approach can reduce inventories because work in progress is eliminated, but the delivery speed and reliability might suffer (Vollmann et al. 2005, pp. 456-457).

It was decided that an appropriate way of researching the differences in the approaches is a simulation study. A simulation can give powerful insights into the problem at hand. It enables studying the performance of the system in multiple scenarios and in a short amount of time. Most importantly the simulation study doesn't interfere with the daily operations of the company but still offers approximate results on how the different approaches behave. (Laguna & Marklund 2013)

While the research provides useful insights into the processes of the case company, it also falls quite well into the suggested future research topics of a couple of prior researches. For example, Su et al. (2010) suggest studying the impact of demand, production capacity and degree of delaying differentiation on lead-time in a configure-to-order environment. Nyaga et al. (2007) studied the effects of demand uncertainty and configuration capacity in a configure-to-order environment and stated that similar research in a make-to-order environment could give more insights.

1.2 Research questions and goals

The aim of the thesis was to challenge and evaluate the CTO order fulfillment strategy in use at the company. The competing approach is an MTO process where no work in progress is kept in inventory. The research was based on one primary research question with one secondary research question focusing the direction of the study.

The primary research question was: How does a CTO strategy differ from an MTO strategy in terms of capital tied up in inventories, capacity utilization and delivery reliability? The secondary question was: How do different levels of capacity buffers, supply lead times and demand uncertainty affect the processes' inventory investment and delivery reliability? The research questions are shown in table 1.

Table 1: *The research questions*

Primary	**How does a CTO process differ from an MTO process in terms of inventory investment and delivery reliability?**
Secondary	How do different levels of capacity buffers, supply lead times and demand uncertainty affect the processes' inventory investment and delivery reliability?

The research questions enforced the goal of evaluating the company's processes well. The main research question was worded quite loosely as it aimed to get a broad understanding of the differences in the processes and to evaluate them from multiple angles.

The focus was on scenarios which resemble the current situation of the Agilon business at the start of the year 2020, as well as some possible future scenarios.

The secondary question was more focused to gain insights of more specific interactions in the processes. This was done to attain an understanding of the impact of these factors and their relative strength in determining the performance of the processes'.

1.3 Research Approach

The aim of the research is to understand and evaluate the constructs currently in place at the company which determine the way the company goes about its daily business. Therefore, the purpose of this study is exploratory. An exploratory study strives to understand the current situation and gain some new insights into the problems at hand. Also, as the aim is not to incur changes in the processes but rather study them, the study was conducted from a regulatory perspective. (Saunders et al. 2009, pp. 119-120, 139)

In this study, the fundamental nature of the phenomena under scrutiny was interpreted to be independent of social actors and determined by the observable data and processes. Therefore, the ontology of this research was positivism. A positivistic approach sees that perceivable phenomena can be explained by the data they provide and seeks to reduce them to their simplest components. The purely quantitative nature and objectivistic point of view of the study enforces the adoption of positivism. (Saunders et al. 2009, pp. 119) This is particularly true in the creation of a simulation model. Simulations are always just an approximation of the actual situations they emulate (Klee & Allen 2011, pp 3; Laguna & Marklund 2013, pp. 99). However, they can provide insights into the problem and in this study the simulation is seen as a valid means to turn the collected data into legitimate findings which can be objectively analyzed.

Due to the objectivistic nature and regulatory perspective of the research, the research paradigm used was the functionalist paradigm. A research which operates with the functionalist paradigm seeks to give rational explanations to the studied phenomena. The case company was seen as a rational entity which has rational problems that can be solved with rational explanations. (Saunders et al. 2009, pp. 120-121)

The time horizon of the research is mostly cross-sectional with some longitudinal elements. The simulation model was based on the processes of the company in early 2020, which enforces a snapshot-based perspective. However, simulations are often used in compressing time to model the continuous behavior of a process in a short time-frame (Laguna & Marklund 2013, pp. 254). To create a truly longitudinal time-horizon to the research would require the assessment of the results against the actual performance of the company, which is not feasible due to the narrow timeframe of the study.

The study was conducted as a mono-method simulation study. The primary data used was the product structure of the Agilon system, which included all of the possible configurations of the modules and components which make up the systems. The data was processed through the pre-set simulation model with the chosen scenario variables.

The simulation approach was chosen as its advantages were enabled by the scope and time-horizon of the study. A simulation can experiment a large number of scenarios over long time-periods in a matter of seconds and give concrete estimates of differences in the performance of processes without interrupting the operations or needing changes in infrastructure (Robinson 1994, pp. 7; Laguna & Marklund 2013, pp. 100-101).

The simulation model was created by mimicking the current order fulfillment process in use at the company, and then approximating how the other approach would be operated with the resources available. The information needed to create the model was gathered through researching the appropriate equations and laws from literature, observations of daily processes and conversations with different employees at the case company.

The simulation model was programmed with the Python programming language and without a ready-made simulation library. This was done to achieve the most freedom of design and simplicity for the model, as the use of an actual simulation software can be more expensive, less flexible and harder to use. These advantages were seen to outweigh the disadvantages of self-programmed simulations, which in addition to the aforementioned are mainly the possibility of more time spent on creating and validating the model. (Robinson 1994, pp. 11-12) The limiting factors of the simulation are discussed more in the next section and the more in-depth decisions made when creating the model are elaborated on in chapter 5.

1.4 Research limitations

The focus of this study was the order fulfillment process of the case company in the Agilon factory in Tampere. This contains the processes from the customer order until the order is delivered to customer site. As the study focused on the operations in the Agilon factory, the actual installation was excluded as it is performed at the customer site often by external technicians. Figure 1 depicts the steps of the process. The steps marked in blue are included in the scope of the study. The simulation focuses on the steps inside the light blue box. The order processing and delivery are outside the box since they were only the connecting points between the simulation and the outside operators in the supply chain. The customer order and installation are performed by outside actors and therefore are not a part of the study.

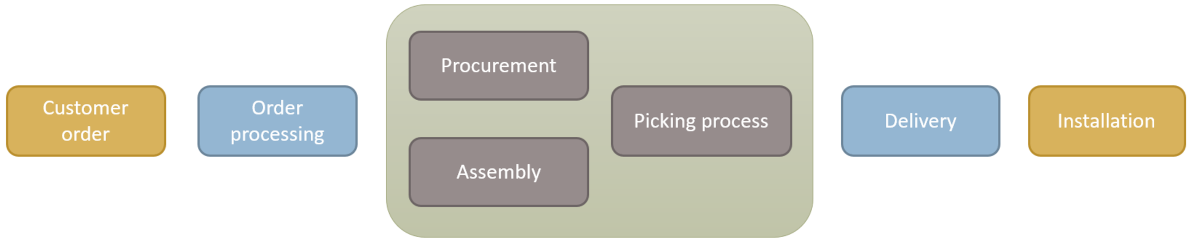


Figure 1: *The order fulfillment process*

As seen in figure 1, the focus is on procurement, assembly and the picking process. For the purposes of the study, procurement encapsulates all the daily purchasing and inventory handling procedures. Assembly represents the actual assembly of the modules which are needed in each delivery. The picking process includes the gathering and packing all the modules and needed components and preparing them for delivery.

It is highly impractical and time-consuming to try and make a simulation track every single detail of a real-life process (Laguna & Marklund 2013). This means that only parts of the process should be simulated in detail. The scope of the simulation is also so wide that some of the simpler details must be compressed to get a good approximation of the performance measures in the overall process.

The main parts of the process which were included in the simulation in detail are the above-mentioned production of modules, purchasing and picking the deliveries. The ordered component deliveries are tracked in daily amounts and the production of modules and delivering systems in weekly amounts. Many things are also not tracked in the simulation. For example, order processing, inventory handling and delivering systems were made to happen instantaneously. The BOMs of the modules were also flattened to just one layer for simplicity. A complete list of assumptions and generalizations can be found in chapter 5.

1.5 Thesis structure

The thesis will start with a literature review of the main topics relevant for this research. The literature review focused on explaining the key points of related theory having to do with different aspects of and external factors affecting the order fulfillment processes of a company. The first topic is an overview of the uncertainties a company faces in its operations. Demand uncertainty is elaborated on as it is the type of uncertainty the simulation imitates. The second topic is mass customization, where the main concepts under scrutiny are a modular product architecture and delayed differentiation. Lastly, an overview is given on the different order fulfillment strategies a company can utilize. Make-to-order and configure-to-order strategies are explained in more detail as they are the approaches which are compared in the simulation.

The Agilon system is introduced in chapter 3. The aim of the chapter is to create an understanding of the modularity of the systems and the extensive number of configurations which can be created from the modular product structure. Understanding the product structure is important to clarify the concepts and operations modelled in the simulation. The chapter consists of two sections. The first section lists the different modules that make up the Agilon units. The second section showcases the types of systems Agilon can be configured.

The two material handling concepts are presented in the fourth chapter. The chapter first explains some common things and compares e.g. the order penetration points in the different approaches. The CTO process emulates the actual operations in use at the case company while the MTO process tries to approximate the operations as they would be operated if the company utilized an MTO process.

The 5th chapter first explains the basis on which the simulation was created and then explains on its limitations and inner workings further. Key things are demand and system modelling, key variable calculation and the simulation's flow logic. An important aspect of creating the model was also to validate its results. This was done by running the simulation multiple times with different test-variables and analyzing whether it created consistent results. Also, it was analyzed if the results adhere to the real-life performance of the process. Lastly, the 5th chapter elaborates on the structure of the simulation.

The sixth chapter lists the results of the simulation. The results are first analyzed from the point of view of inventory investment. First the section gives an overview of the inventory investment and compares the processes' results. After that the individual impact of the simulation's input variables is analyzed by conducting a regression analysis. The second part of the chapter analyzes the delivery reliability results in terms of percentage of late deliveries, delivery times and capacity utilization.

The seventh chapter consists of the evaluation of the results and discussion how the results match up to the theories and concerns in the literature review. The discussion with the literature aimed to identify notions and ideas which enforce or supplement the prior research done on the subject. The research was then critically evaluated to identify possible shortcomings. Lastly, the section provides managerial implications and directions for future research.

2. THEORY BACKGROUND

This chapter consists of a literature review of the most significant theory concepts behind this research. The intention was to present topics which have an impact on a company which produces complex configurable systems in an environment with a high degree of demand uncertainty. This was carried out by going through the topics in a progressive manner, starting from the more abstract external forces and proceeding to the more concrete actions and strategies for a company in this type of environment.

In practice this means that first section 2.1 gives an overview of the variation and uncertainty a company faces in its business environment. The subsections give more insights about demand uncertainty, which is the main type of variation occurring in the simulation model. The details of the simulation are discussed in chapter 5. Section 2.2 then explains the concept of mass customization, which is a way to competitively provide variable products. The subsections 2.2.1 and 2.2.2 dive deeper into the tactics of achieving successful MC through lowering costs and dealing with uncertainty. Section 2.3 goes through the different strategies by which a company strives to fulfill customer demand. The first subsections elaborate on how positioning the order penetration point affects a company's operations. The subsequent subsections describe the theory behind the order fulfillment strategies used in the simulation. A summary of the theory background is synthesized based on the main principles of this research in section 2.4. Lastly, section 2.5 discusses the implications of the theory from the point of view of the case company.

2.1 Supply chain uncertainty

A company faces multiple types of variation in its operations. Simangunsong et al. (2011) describe supply chain uncertainty as the uncertainties and risks which can take place at any point in the global supply chain network. This section will first give an overview of the types of uncertainty a company has to deal with and how this uncertainty can be handled. First a broad overview is given, and the scope is then narrowed down to the factors which are the most relevant for this research.

In their literature review Simangunsong et al. (2011) give an extensive view on the uncertainties a company faces. They conclude that supply chain uncertainty is caused by the following three types of factors:

- Internal organization uncertainties
- Internal supply-chain uncertainties

- External uncertainties

The internal organization uncertainties consist of uncertainties originating from the company in question, e.g. the company's manufacturing processes and product characteristics. The internal supply-chain uncertainties are uncertainties that arise from things that are controlled by the company or its supply chain stakeholders. These include for example end-customer demand and supplier performance. Lastly the external uncertainties are factors from outside the supply chain such as government regulation or natural disasters. (Simangunsong et al. 2011)

A pioneering article written by Davis (1993) identified that the three main sources of uncertainty in a company's supply chain are: suppliers, manufacturing and customers. These can be derived to the three main types of uncertainty: supply, process and demand uncertainty. (Davis 1993) The first two are part of the internal supply-chain uncertainties and therefore are only partially in control of the focal company. The process uncertainty represents the internal uncertainties.

The uncertainty in a company's processes is detrimental because it creates costs in the form of stock-outs, excess capacity and inventory buffers (Christopher & Holweg 2011). Uncertainty is also the main culprit behind the late deliveries, machine breakdowns and order cancellations a company faces in its daily operations. Basically, inventories are kept only to account for uncertainty (Davis 1993)

Angkiriwang et al. (2014) divide the means to mitigate uncertainty into two categories: proactive and reactive strategies. Simangunsong et al. (2011) categorized them into strategies which try to cope with uncertainty and strategies which attempt to reduce uncertainty. These approaches have a similar ideology behind them: Some strategies try to proactively reduce uncertainty and other strategies strive to improve the ability to live with and react to the effects of uncertainty. The strategies have an interesting antagonistic relationship. The proactive strategies can be linked to the waste-eliminating philosophies of Lean and JIT, while the reactive strategies are trying to increase buffers and waiting periods into the processes to create stability.

The proactive strategies try to redesign an organization's operations to reduce the encountered level of variation. Some examples of uncertainty reducing strategies are: Lead time reduction, setup time reduction, postponement, outsourcing and subcontracting. The reactive strategies are often trying to add buffers to a company's processes to account for variability. Classic examples of reactive coping strategies include safety stock, capacity buffer and safety lead times. (Angkiriwang et al. 2014) These strategies don't come without a cost, however. Battling uncertainty increases costs and therefore a good compromise between the costs, actions and the benefits is needed to succeed (Davis 1993)

2.1.1 Demand uncertainty

The operations and processes of a company are dictated to a very high degree by the demand for its products. The demand patterns steer the production planning and purchasing and variations in demand make planning operations increasingly difficult. (Nyaga et al. 2007)

Balancing supply and demand is a considerable problem for all manufacturing companies. Companies need to prepare for future demand quite a bit in advance. This creates risks if e.g. demand forecasts differ from the actual demand. (Heikkilä & Ketokivi 2005, pp. 117, 119) Deviations from the forecasts cause for example increases in inventories, increased risk of component or end-product obsolescence (Zäpfel 1998)

The variation in demand can present itself in multiple ways. The mean volume of demand, the variability between the highest and lowest demand and the demand mix of the end products can differ greatly from period to period. (Harrison & Skipworth 2007) In the case of configurable products the uncertainty regarding the product mix can be defined as configuration uncertainty as the product is the same but the configuration is different (Chen-Ritzo et al. 2010).

As forecast-errors are a big factor in demand uncertainty, a good way to increase robustness against demand variation is to increase the accuracy of forecasts. This can be achieved by reducing the time-horizon of forecasts by postponing the differentiation of products. (Yang et al. 2004) Another approach to reducing demand uncertainty is postponing the manufacturing processes until the customer has placed an order. Customer-order based delivery processes can eliminate almost all the demand uncertainty. (Zäpfel 1998) However, this often comes at the cost of longer delivery times which the customer doesn't always tolerate. (Yang & Burns 2003) There are many ways of trying to cope with uncertainty. The next subchapter discusses some of these answers in more detail.

2.2 Mass customization

Stevenson (2011, pp. 150) states that companies like producing standardized products because they enable high product volumes with relatively low costs. However, because customers prefer a wide variety of products with low prices, standardization doesn't always cut it. Therefore, manufacturing companies need to try and solve this dilemma. For some companies, the answer is mass customization. (Stevenson, W. J. 2011, pp 149-150)

It is generally accepted that the concept of mass customization was first coined by Stanley M. Davis in 1987. (e.g. Duray et al. 2000; Partanen & Haapasalo 2004; Guo et al. 2019) It has since been researched and utilized extensively from multiple perspectives. The literature review done by Fogliatto et al. (2012) identified the fundamental principle in mass customization as the co-existence of customized products and economies of scale in production.

As stated above, the basic need for mass customization arises from the need to simultaneously provide customized products and to keep the company's operations efficient. While there are multiple ways to define it, MC is a way to provide customized products that fit individual customers' needs while taking advantage of the benefits of mass production. (Partanen & Haapasalo 2004; Haug et al. 2009)

There are many ways to achieve mass customization. Generally, the actions have something to do with modularity, postponement or both. (e.g. Duray et al. 2000; Salvador et al. 2002; Stevenson, W. J. 2011, pp 150; MacCarthy 2013) While there are many types of modularity and postponement, for the purposes of this research the focus will be on a modular product architecture and delayed differentiation of a product. These actions are elaborated on in the following sections.

2.2.1 Modular product architecture

A modular product architecture enables manufacturers to create a large variety of products from a smaller set of parts (Duray et al. 2000). This is achieved by creating modules by grouping components into subassemblies. The modules can then be combined in different ways to create end products with different capabilities and appearances. (Stevenson, W. J. 2011, pp 150) Figure 2 shows how a modular product architecture creates an hourglass shape as the variation is smaller on the module level than on the component or product variant levels.

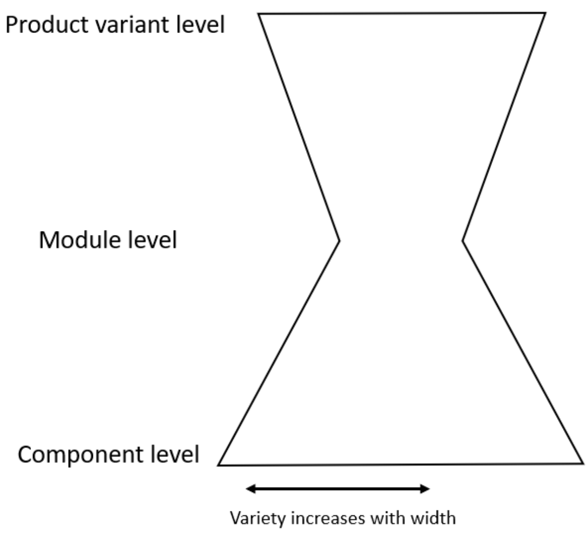


***Figure 2:** Illustration of the hourglass shape created by a modular product structure (adapted from: Heikkilä et al. 2002, pp. 13)*

Ulrich (1995) describes that product architecture is defined by the following three elements:

- The arrangement of functional elements
- The mapping from functional elements to physical components
- The specification of interfaces among interacting physical components

As the name implies, the functional elements combine to create the function of the product. The arrangement of the functional elements means the way they are connected to perform the tasks they are intended to perform. Some elements only interact with different functional elements while others interact with external entities, such as the environment in which the product is used. (Ulrich 1995)

The mapping of functional elements to physical components determines the sets of components which implement the different functional elements of the product (Ulrich 1995). These sets of components will make up the modules which are then combined to create the product.

Lastly, the specification of the interfaces shows how the modules interact with each other (Ulrich 1995). The interfaces in a product determine how the modules fit together, connect and communicate. (Baldwin & Clark 1997) These interfaces are essential for realizing the benefits of modularity as they make the mixing and matching of different modules possible. (Sanchez & Mahoney 1996)

Ulrich (1995) lists three different types of modularity: slot-modular, bus-modular and sectional-modular. These types differ from each other by the way the interactions between modules is coordinated (Ulrich 1995). The three different types are illustrated in figure 3.

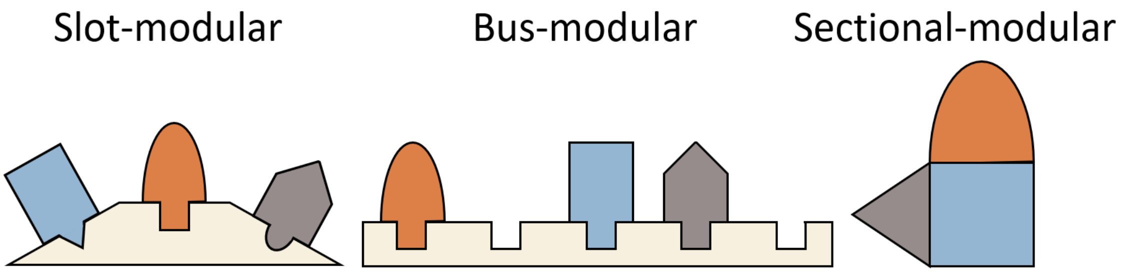


Figure 3: *The three types of modularity (adapted from: Ulrich & Eppinger 2012, pp. 186)*

In the Slot-modular architecture the different functional elements have a slot in which they can be set into. The modules can't be interchanged due to the slots having different types of interfaces, as seen in figure 3. (Ulrich & Eppinger 2012, pp. 186) The slots can accommodate different styles of modules if the interface is the same. For example, a computer can have different graphic cards as long as the connection to the rest of the system is standardized.

The bus-modular architecture connects the modules with a common bus. The connecting interfaces are uniform so the modules can be attached to the bus in any arrangement. An example of this is an expansion card for a pc. (Ulrich & Eppinger 2012, pp. 186) The card can have e.g. several USB-ports to which external USB-devices can be attached in any order.

In a sectional-modular architecture there is no single connecting module as in the earlier two types. The modules are connected to each other by identical interfaces which allows for multiple end-product variants. Piping systems are an example of sectional modularity. (Ulrich & Eppinger 2012, pp. 187)

Ulrich and Eppinger (2012, pp. 187) state that the slot modular architecture is the most common one as most of the time the individual modules demand different styles of interfaces. This is because of the unique interactions the different functional modules are meant to perform. The other two types are better when the end products need to have wide variety but can be connected uniformly. (Ulrich & Eppinger 2012, pp. 187)

The modular architecture also comes with many benefits. Product modularity has been found to reduce costs, improve quality and make production more flexible (Jacobs et al. 2007). As different end products share some common modules, the modules can be produced in bigger volumes. (Ulrich & Eppinger 2012, pp. 189) This standardization and its effects is at the core of mass customization as companies strive to achieve the benefits of mass producing.

While it enables standardization of processes while providing variety, the modular product structure can technically restrict the variety when compared to products which are assembled from the component level, as there are less building blocks from which the product is assembled (Stevenson, W. J. 2011, pp 150-151) Another possible downside of modular product architecture is that the system-level design phase is of utmost importance (Ulrich & Eppinger 2012, pp. 191). If something goes wrong there, the company might not be able to realize the benefits.

2.2.2 Delayed differentiation

Delayed differentiation, also called form postponement, is a strategy which allows greater standardization of operations and reduces uncertainty in a company's manufacturing process. (Su et al. 2005) It means processing some of the uniform steps in a product family's manufacturing process first and delaying the differentiating steps until later parts in the supply chain. (Stevenson, W. J. 2011, pp. 150)

Delaying differentiation is in part enabled by a modular product structure. Ulrich and Eppinger (2012, pp. 199-200) give a fitting example of this. In the case of an electrical device marketed to different countries a different power supply is needed for some countries. If the power supply is integrated to the design, every country needs a different manufacturing process for the product. If the power supply is a module in a modular product architecture, the whole manufacturing process can be standardized to a point in which only the power supply is added prior to final packaging. (Ulrich & Eppinger 2012) MacCarthy (2013) also notes that if the degree of customization wanted by the customer can be achieved through modular product structure and postponement, it can be a very good operational strategy to implement.

As Pagh and Cooper (1998) state, a classic example of delayed differentiation is painting the products only after the customer has decided which color is the best. Therefore, all the other production steps can be done before this in a standardized way. (Pagh & Cooper 1998) If the components would be painted before assembling the final product, the number of different components would be considerably larger as there would be a need for each component in each color.

Harrison and Skipworth (2007) list three things that should be considered when planning a manufacturing process which implements delayed differentiation to get the best results:

- Previously added value should not be decreased during the delayed differentiation process.
- The postponed processing time should be short compared to the total manufacturing process time.
- The number of generic products should be kept to a minimum and each variant should have high volume demand and low demand variability.

The first argument implies that if reworking is needed due to the postponement the benefits are easily lost. Secondly, if the differentiation processes are very long and diverge from each other a lot, the benefits of the common processes can turn out to be too small to be profitable. The third argument stems from the fact that the target is to benefit from extensive standardization and if the number of generic products becomes too large, the level of standardization might remain too low to be beneficial enough. (Harrison & Skipworth 2007)

As stated before, postponing the differentiation of product reduces the number of different components which need to be kept in inventory. For producing a large variety of products, delaying the differentiation has been shown to reduce the capital tied-up in inventories greatly. (Pagh & Cooper 1998; Appelqvist & Gubi 2005)

In addition to reducing inventories and providing a more standardized way to produce variety, delayed differentiation can also fulfil another basic customer need, the speed of delivery. Performing production steps in advance will naturally decrease the time a customer has to wait after placing an order. Delaying differentiation can likely decrease the delivery time to just a fraction of what a non-postponed manufacturing process can achieve. (Appelqvist & Gubi 2005; Harrison & Skipworth 2007)

2.3 Different approaches to order fulfillment

Order fulfillment is the process of responding to and delivering customer orders. Stevenson (2011, pp. 682-683) lists the following four approaches as some of the most common strategies:

- Engineer-to-order (ETO)
- Make-to-order (MTO)
- Assemble-to-order (ATO)
- Make-to-stock (MTS)

The key difference between these approaches is the point at which the demand becomes defined by the customer rather than a forecast (Olhager 2003). This point is usually called either the order penetration point (OPP) or customer order decoupling point (CODP), and it is the point where the customer is involved in the process (Vollmann et al. 2005, pp. 20). The OPP positions of the different strategies are illustrated in figure 4. The dotted lines represent forecast based operations and the continuous lines operations that are based on customer orders.

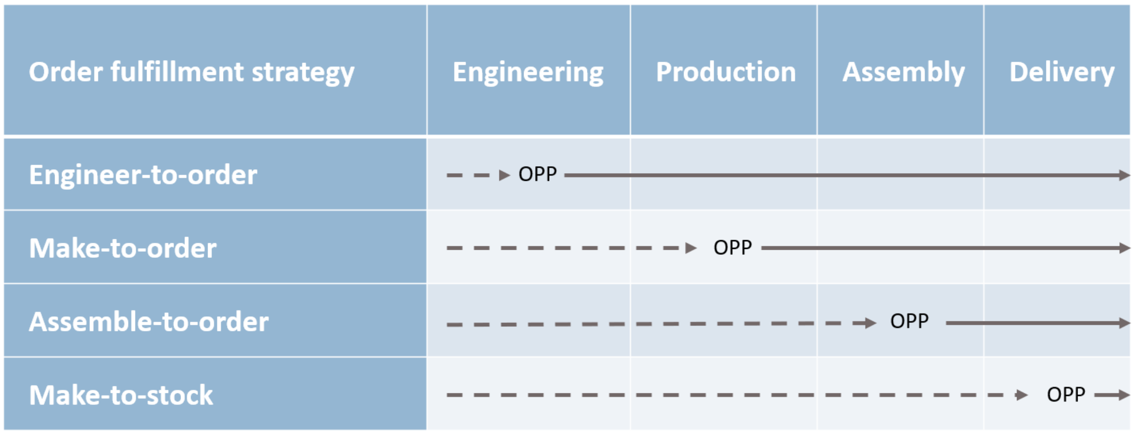


Figure 4: *Position of the order penetration point in the different order fulfillment strategies (adapted from: Olhager 2003)*

The ETO strategy involves the customer all the way from the engineering phase. The products are designed and produced to fit the customers' specifications. (Stevenson, W. J. 2011, pp. 682) The ETO processes often deal with large and complex project deliveries, such as construction projects. As with often with large project-based products, the common characteristics of an ETO-environment are a rather long delivery time, high level of difference between deliveries and a need for a flexible delivery process. (Gosling & Naim 2009) In an engineer-to-order the materials needed for the end-product are not clear at the start of the project and they need to be defined as the design process progresses (Vollmann et al. 2005, pp. 23)

In the make-to-order approach the customer order launches the production of a product. While the products can range a wide variety, the processes of procuring the needed materials and producing the product are defined before the customer order. (Vollmann et al. 2005, pp. 23) As such in the MTO-process, forecasting involves the types of products the customers might want and how they are produced.

An assemble-to-order process is a hybrid between the make-to-order and make-to-stock strategies (Cheng et al. 2002). It is often associated with delayed differentiation. The end-products are assembled to customer order from a prefabricated set of subassemblies. ATO-processes often take advantage of a modular product structure to achieve this. The production steps done before the OPP allow for a compromise between variety and speed of delivery. (Olhager 2003; Atan et al. 2017)

The make-to-stock process is often used for standardized products with predictable demand patterns. The OPP is at the point of delivery so the products are kept at the finished state and the service level is defined solely by whether the item is in stock or not, as the delivery time is usually negligible. (Vollmann et al. 2005, pp. 21-22) As an MTS-process often focuses on a rather narrow product range and the focus of the production is to compete with cost-efficiency as the process can be optimized to a high degree in terms of capacity and inventories. (Olhager 2003)

2.3.1 The significance of positioning the order penetration point

As stated before, the order penetration point separates decisions based on speculation on future demand from commitment against actual orders (Wikner & Johansson 2015). Positioning the OPP has a big impact on the attributes and requirements of a company's order fulfillment process.

The positioning of the OPP is often dictated by the requirements in the markets. If customers require a short delivery time, the OPP cannot be shifted too far backwards. On the other hand, if the customers demand a big variety of products, shifting the OPP too far forwards can result in huge investments in inventories to keep the required variety in stock. (Olhager 2003)

The commanding factor ruling the OPP position is the ratio between the supply lead time and delivery lead time (P/D-ratio). If the delivery lead time requirements are shorter than the supply lead time the supply chain is capable of, some process steps need to be done in advance of the customer order. This can occur if the products have a long manufacturing lead time or if the components have a long delivery time. The component-based difference can be accounted for by keeping component inventories, but the longer manufacturing lead time requires some manufacturing steps to be done in advance. (Wikner & Johansson 2015)

Olhager (2003) states that the choice of fulfillment strategy is also influenced by the volatility of demand, as in how much the demand tends to deviate from the mean demand. Figure 5 depicts how the different fulfillment strategies fit in terms of demand volatility, P/D-ratio. Low demand volatilities can be seen to fit MTS strategies, while higher demand volatilities need more customer-based approaches. The P/D-ratio of less than one indicates that a MTO approach could be beneficial while, as mentioned before, P/D-ratio over one requires that part of the production is done based on speculation.

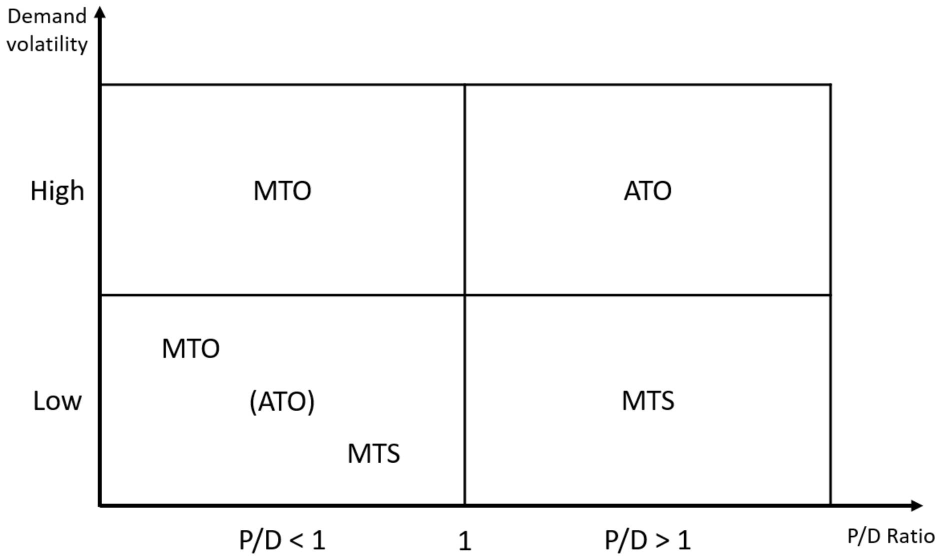


Figure 5: *The applicable order fulfillment processes in terms of demand volatility and P/D ratio (adapted from: Olhager 2003)*

Positioning the OPP has a big impact on the delivery process. Shifting the OPP forward can reduce the customer lead time and enable better optimization of processes. The downsides are that production will rely more on forecasts, product customization can become more expensive and a probable increase in work-in-process inventory. Alternatively, shifting the OPP backwards has the opposite effect. Lead times might increase, and optimizing processes is harder. However, reliance on forecasts is reduced, easier customization is enabled, and inventories reduced. (Olhager 2003)

The buffer-inventories in a company's supply chain are usually kept at the OPP. This is due to the downstream operations being driven by the customer orders and therefore don't need to buffer against demand variation. (Saeed et al. 2016) Also, the inventory at the OPP enables the upstream operations to focus on efficiency as it buffers against demand fluctuations and therefore the upstream processes don't have to drastically react to them and risk compromising performance (Wikner & Johansson 2015).

Before the OPP the aim is to optimize and stabilize the process. However, after the OPP the goal is to offer high variety and customization. (MacCarthy 2013) The findings of Harrison & Skipworth (2007) add to this notion as they conclude that if the downstream activities don't have enough excess capacity for responding to the demand mix uncertainty, the delivery reliability of the whole process will suffer.

2.3.2 Make-to-order strategy

Due to the increasing demand for customized products many companies have their focus to producing products after a customer order rather than producing to stock (Chen-Ritzo et al. 2010). As discussed in the previous chapter, this is accomplished e.g. by utilizing a make-to-order strategy. This section will elaborate on the concepts and dimensions of this approach.

The MTO strategy is characterized by wide variety of end-products, production orders being launched by the customer order and often rather long delivery times (Vollmann et al. 2005, pp. 456) As the products are largely customized according to the customers specifications, the MTO strategy is quite similar to the ETO strategy. Usually companies don't rely only on one or the another, but often provide a set of MTO products while designing some bigger projects with customers. (Willner et al. 2014)

However, in MTO the products are configured from a pre-determined set of designs, and therefore there is no need for as much engineering and design work. The design and development phase for MTO products is conducted beforehand according to market forecasts. The customer involvement is usually interactions also limited to only the order specification phase and delivery. (Willner et al. 2014)

While the designs of the products are pre-determined, the processes of producing them are often not standardized. The customization leads to different material and production requirements and different job-routings on the shop-floor. (Stevenson, M. et al. 2005) As the demand is quite well defined at the production state, the uncertainty in MTO is often not the quantity or timing of the demand but rather the resources needed to produce the specified products (Vollmann et al. 2005, pp. 26)

While the delivery lead times can be long, they can be improved by overlapping schedules due to the different requirements and timetables of the delivery projects. Excess capacity in different parts of the process can also help in scheduling the deliveries to make the required delivery dates. (Vollmann et al. 2005, pp. 456-457)

Reliability of delivery is quite hard due to the customized products and long delivery projects. The delivery accuracy is quite important as lateness in deliveries might cause extra costs and early deliveries can be quite inconvenient. (Vollmann et al. 2005, pp. 456-457).

2.3.3 Configure-to-order strategy

Configure-to-order (CTO) is a slight variation of the assemble-to-order strategy. The difference to ATO is that instead of choosing from a standardized list of products, the customer can configure a customized product which consists of the available modules in arbitrary multiples. (Cheng et al. 2002; Chen-Ritzo et al. 2010) The differences between CTO and ATO are quite small. This makes many of the attributes of ATO environments applicable to CTO environments.

In CTO, inventory is kept at the module level. This can enable fast customer delivery times and variable products. CTO is most suitable when the end-product assembly time is short compared to either the production time or the replenishment time of individual components. (Cheng et al. 2002)

Due to the final assembly process being relatively short, the delivery reliability can often be better than in MTO. This is due to some of the production steps being standardized and done in advance. The delivery speed can be improved by reducing the lead times of the final assembly steps. (Vollmann et al. 2005, pp 456-457)

As mentioned in the OPP section, the flexibility of processes downstream of the OPP is important. In the case of CTO, it is often called configuration capacity. In high demand uncertainty situations, the capacity to produce the complex products needed by the customer is crucial. (Closs et al. 2010) Another important aspect is the processing and configuration of the orders. In the face of variable demand, the quick processing of orders is needed to keep the customer service level high. (Nyaga et al. 2007)

The customization options cause the product structure of CTO products to vary between deliveries. This results in an uncertain BOM and more difficult planning of purchasing. The accuracy of forecasts and planning can be done by implementing variable quantities for components in the bills of material and thereby aggregating the demand to be on the level of product families instead of fully configured end-products. (Chen-Ritzo et al. 2010) The final assemblies are determined only after the customer demand is known (Vollmann et al. 2005, pp. 448-449).

2.4 Theory summary

Companies' face a lot of uncertainties in their processes. The uncertainty affects all aspects of the operations ranging from supply, internal processes and demand (Davis 1993). The uncertainty creates risks which differ a lot in their gravity. Some can be very unlikely but have great implications such as natural disasters, while others can be something as usual as the fluctuations in some manufacturing processes lead time (Simangunsong et al. 2011).

The type of uncertainty studied in this thesis is demand uncertainty. It can appear as quantity and mix uncertainty. The quantity of customer demand varies and because a company often offers multiple products the shares of the individual products varies in the overall demand. (Harrison & Skipworth 2007) In the case of configurable products the mix uncertainty is often referred as configuration uncertainty (Chen-Ritzo et al. 2010).

To respond to the variations in demand, a company can apply both reactive and proactive measures. The reactive strategies involve buffering against the variation in some way, such as adding a safety stock to inventories. The proactive strategies include postponing differentiation of products until customer demand is known. (Angkiriwang et al. 2014) While these actions help in mitigating uncertainty, they also come with downsides. Buffer inventories lead to higher inventory costs and postponement can e.g. lengthen the customer delivery time. The response to uncertainty is always a compromise.

A well-researched approach to providing customized products with close to mass production performance is called mass customization. The ways to implement mass customization provide ways to respond to highly variable demand with relatively low costs and in an adequate time. (Duray et al. 2000) The two main ways of mass customizing products for this research are a modular product architecture and delayed differentiation. The modular product architecture and delayed differentiation often go hand-in-hand. The key is to enable production of standardized products and differentiate them only after the customer has placed an order to ensure that the right number of products is produced. This allows for standardized manufacturing processes and wide product offering. (Harrison & Skipworth 2007)

The order fulfillment process of a company is dictated by the requirements of the customers as well as the production process. If the demand is simple and can be easily forecasted, it is often best to try and take advantage of producing standardized products to stock with an MTS approach. The other extremity is an ETO process customers require unique systems which call for different processes and materials for each project. (Vollmann et al. 2005, pp. 21-23)

The strategies used in this research are located between the extremities and include some parts of both standardization and customization. MTO processes create custom products when the customer orders a product. All the production steps are done after the customer order. However, the product is not truly customized as the attributes are chosen from a set of pre-designed options, which were engineered beforehand according to a forecast. This allows for a high level of customization and reduced inventories but entails a longer customer lead time. (Vollmann et al. 2005, pp. 23; Willner et al. 2014) A CTO process tries to provide shorter delivery times by producing standardized modules in advance from which the final product is configured according to the customers' needs. The lead times can be shortened quite a lot, but the company will have to rely more on forecasts and invest more into work-in-progress inventories. (Cheng et al. 2002)

3. OVERVIEW OF THE AGILON-SYSTEM

Agilon is an automated material-handling solution which gives the customer a transparent view of the material flows in the system in real time. Agilon consists of an automated warehouse system and the supporting web-applications. This study focuses on the manufacturing process of the warehouse systems.

An example of an Agilon warehouse system is shown in figure 6. The main elements of the system are the robot which can be seen in the cutoff at the top of figure 6 and the access point, located in the bottom-right corner of figure 6.

Figure 6: *An Agilon warehouse*

Agilon systems are configured to all customers individually. The product is based on a modular structure which allows for a high level of versatility. The amount of access points and robots as well as the dimensions of the system can be determined to fit the customer's needs. This makes it so that each Agilon-delivery uses the same basic modules but the contents of each delivery changes depending on the configuration.

The modular product structure of an Agilon system is shown in figure 7. The green boxes indicate that the module has a fixed bill of material (BOM) and it doesn't change depending on the dimensions of the system. The blue boxes are modules that have differing bills of material based on the dimensions and features of the system.

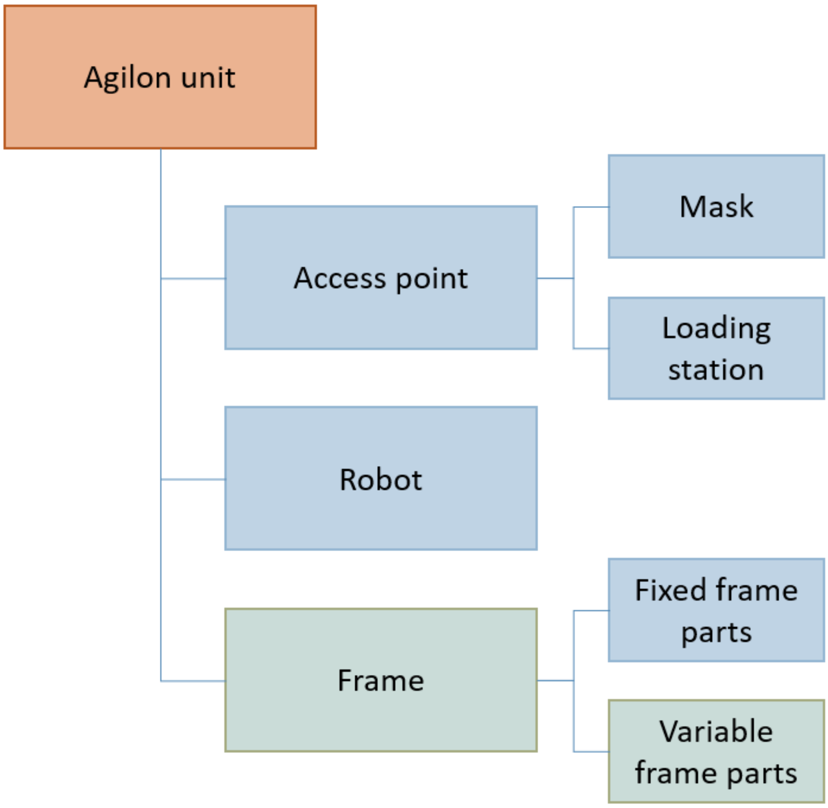


Figure 7: *The different modules in an Agilon system*

While the access point and robot have fixed product structures, the number of these modules can vary between systems. On the other hand, each Agilon system only has one frame module, which includes some fixed and some variable elements. An Agilon frame is shown in figure 8 and the rest of the modules are displayed in Appendix A.

The robot is the most complex module in an Agilon system. It consists of a lifter and a capsule. The lifter moves on the drive rail in the ceiling of the Agilon and adjusts the height of the capsule to access the items in the warehouse. An agilon can have either 1 or 2 robots. The robots handle the packages inside the Agilon.

The access point works as the interface between the robot and the user. It is used to access the packages inside the Agilon. An access point consists of a loading station and a mask. The mask and loading station can be changed individually, but the whole module doesn't work without both. For the purposes of this study, the access point is handled as a single module.

As seen in figure 7 some of the contents of the frame module depend on the configuration of the system, while the others are fixed. The fixed frame parts are the components that an Agilon frame always has. These include the winch rail and its components, and the power supply for the conductor rail from which the robot draws electricity. However, the fixed parts of the frame make up quite a small percentage of the total parts in a frame, as most of the components naturally depend on the size of the frame. These include for example the vertical frame poles, cover sheets and the shelves. The cover sheets are the sheets which create the outside walls. The number and configuration of the shelves depends on what type of items the customer is going to store in the system. An Agilon frame with 1 access point is shown in figure 8.

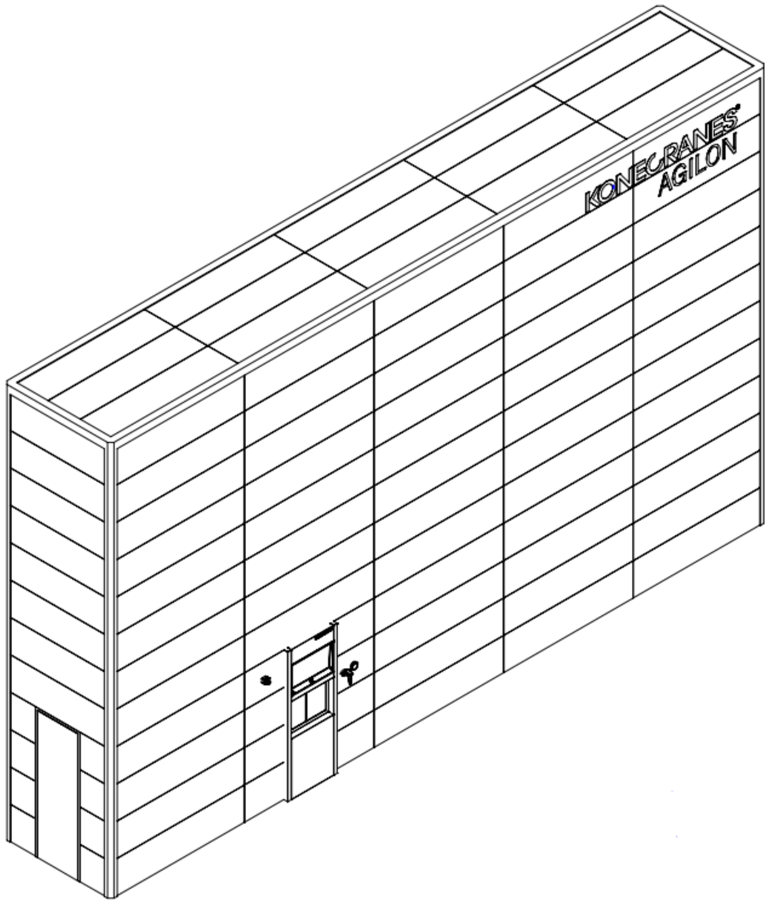


Figure 8: *A standard Agilon system*

Retail units tend to be a lot smaller than industrial units. The frames in retail units are also a bit sturdier than in their industrial counterparts to protect from vandalism but most of the functional elements are very much the same. The access point mask is designed to allow usage by people with disabilities. A retail unit is show in figure 9.

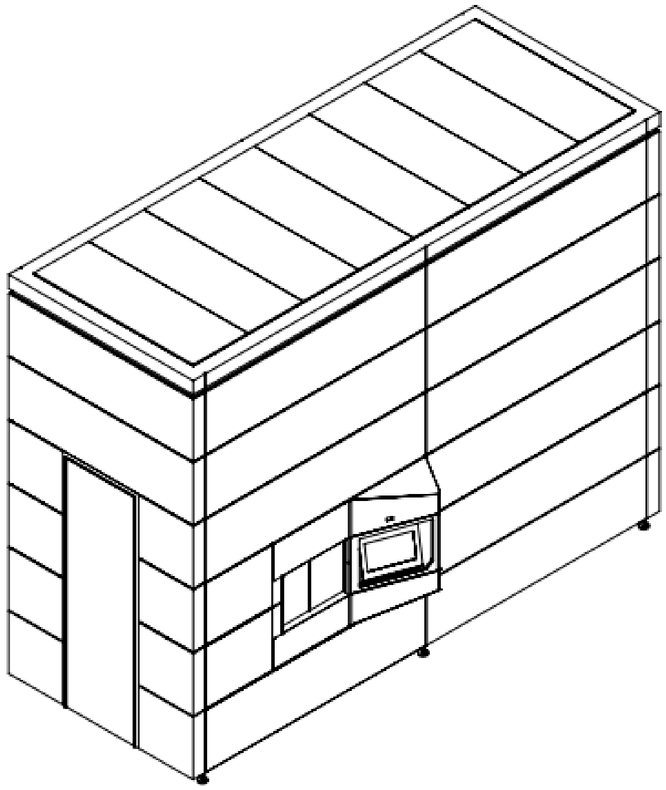


Figure 9: *A retail Agilon*

The Agilon can be configured to meet the customer needs with some additions to the frame module. For example, when two or more parallel units next to each other connected by a 180-degree rail curve, it is called a stacked Agilon. The stacking allows for more efficient usage of floor space. An Agilon can also have an underpass which can get the system over e.g. walkways or conveyor systems while utilizing the empty space above those for storage space. Different Agilon storage units can also be linked with connection tubes. The tubes can be fitted with curves and slopes to allow the connection even in complicated situations. Example pictures of these additions can be found in appendix B.

4. THE PROCESSES

The two different order fulfillment processes are explained in this chapter. First, a brief overview of the processes and the factors influencing them is given. The following subsections describe the processes more closely and explain the policies in place in the purchasing of components. The main factors, strategies and performance metrics are depicted in figure 10. The blue boxes are present in both approaches while the orange box, postponement, only affects the CTO approach. The performance metrics will be elaborated on in chapter 5.

The external effects influencing the processes in the simulation are the variation in quantity and configuration of the demand. This is the only type of uncertainty present in the simulation as supply and manufacturing lead times are set to be constant.

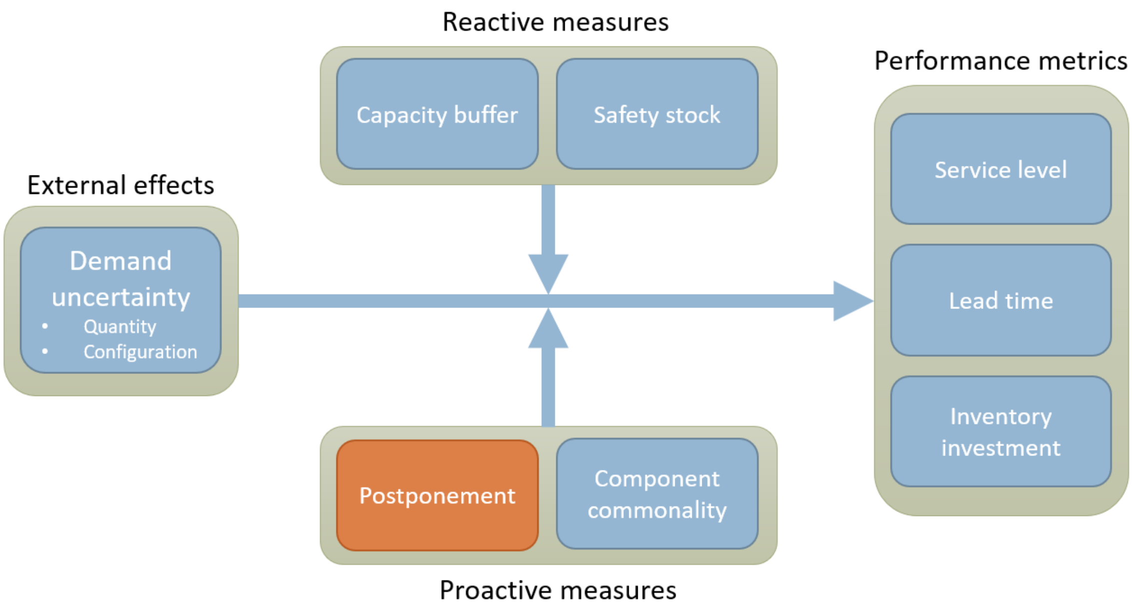


Figure 10: *The factors influencing the processes (adapted from: Angkiriwang et al. 2014)*

The proactive measures to account for the demand uncertainty stem largely from the modular product architecture. In both approaches the sharing of components and modules helps in standardizing the processes and reduces the number of components needed. The main difference in the proactive strategies is that the CTO approach utilizes delayed differentiation while the MTO approach does not. The common modules between pretty much all the systems, robots and access points, are manufactured before the customer order according to a forecast. For the terms of this simulation it should help especially if the assembly processes would face congestion due to large orders in subsequent weeks.

The main reactive strategies are capacity buffering and safety stocks. The capacity buffering is done similarly in both strategies. Determining the size of the buffers is done similar to the approaches of Nyaga et al. (2007) and Closs et al. (2010). They set the lowest level at the capacity to meet the mean demand, and two levels at set intervals higher. Closs et al. (2010) set the buffers as follows: first 100% of mean demand, second at 150% and third at 200% of the capacity needed to satisfy the mean demand. Nyaga et al. (2007) set the levels at 100%, 125% and 150% respectively.

The appropriate buffer-levels for this simulation were seen to be in between those examples, and only the lower and upper limit were needed as the simulation chooses random variables between the limits. The lower limit was set at 120% and the upper limit at 200% of the capacity needed to meet the mean demand. Later in this study the capacity levels are marked as percentages of the maximum demand, which means they are 60% and 100% respectively, as the maximum demand is double the mean demand in this research.

Keeping safety stock is a good method to account for uncertainty in demand quantity (Vollmann et al. 2005, pp. 487). While both approaches utilize them, the safety stock policies differ a lot. The MTO process only keeps buffers for the items which have a longer supply delivery time than the promised customer delivery time. This means that all the items with a shorter supply lead time are ordered only after the customer order is set. The MTO process does not keep modules in inventory. The CTO process keeps buffer stocks for the modules and all the components, except the highly variable frame components with a relatively short lead time. This is done to keep the module inventories at the wanted levels to provide delivery reliability.

As probably evident from the safety stock policies, the OPP is positioned rather differently between the processes. Figure 11 depicts the positioning of the OPPs in the processes. The grey arrow indicates that the process steps are done based on a forecast, while the blue arrow shows the steps which are done according to the actual customer order. The OPP is set at purchasing phase in the MTO process. This means that most of the purchasing and all of the manufacturing is done after the customer order.

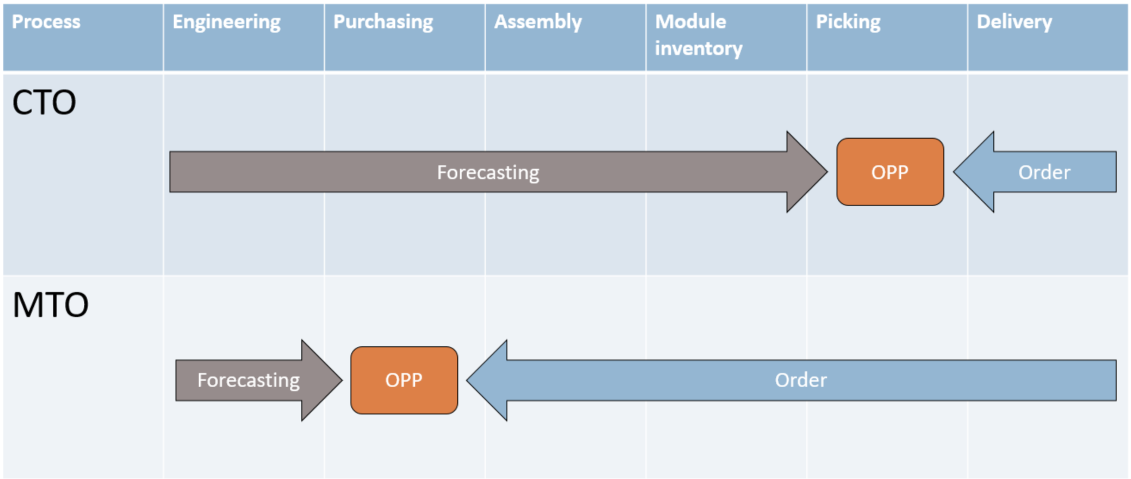


Figure 11: *Positioning of the OPP in the processes*

The CTO process has the OPP a lot more downstream. The OPP is set at the picking phase, which means that all the assembly phases which are done at the factory should be done before the customer order and the order is configured and picked from the component and module inventories. However as mentioned before, some high variance frame parts are purchased after the order. Also, if there is a high peak in demand, the OPP might creep upstream towards the module assembly phase due to module inventories decreasing below the set levels.

4.1 Configure-To-Order Process

The CTO process follows the actual processes of the case company quite closely. The process buffers against the demand variance with component and finished goods inventories which hold a set number of modules and components which are then assigned to deliveries. The finished module buffer can be set depending on the forecasted demand levels and thus leveling out spikes in demand that would otherwise strain the production. The basic outline of the process is shown in figure 12.

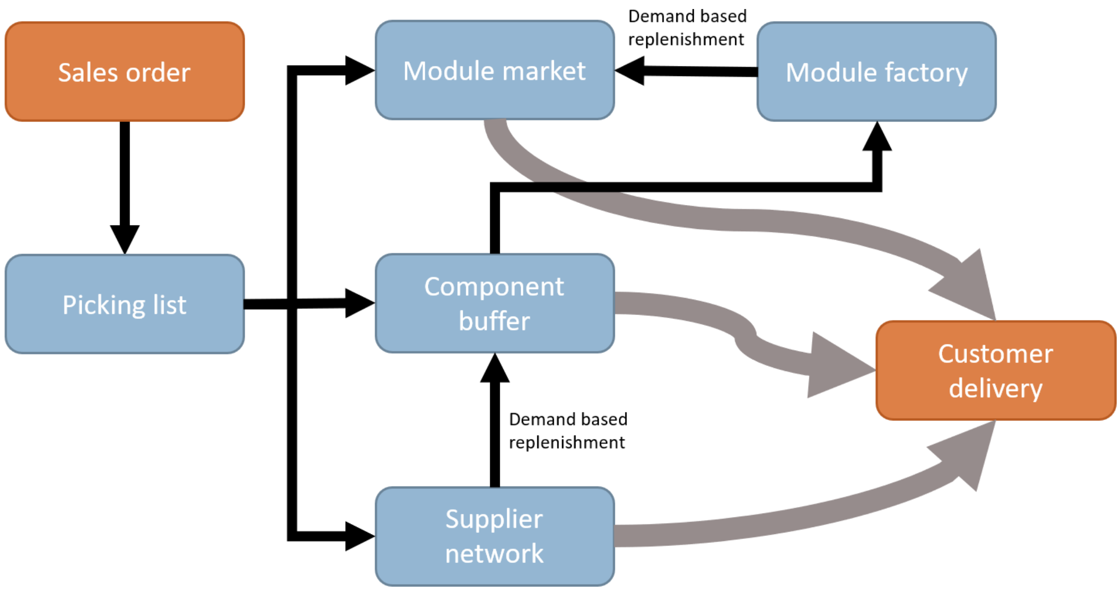


Figure 12: *The outline of the CTO process*

As seen in picture 12, the first thing that launches the process is the sales order. The sales order is then processed into a picking list, which has the BOM of the ordered system. The items needed in the picking list come from three different sources. The first is the module inventory, which is the inventory for the robots, loading stations and masks. Component buffer is the actual component inventory. Lastly, the supplier network provides the parts which vary a lot depending on the configuration of the system e.g. the frame poles and shelves.

When an order is placed, the needed modules and components are taken from the component and module inventories which triggers replenishment orders and production suggestions upstream. This can be seen in figure 12 as the lines which have demand-based replenishment next to them. The component buffer is replenished from the supplier network as items are used up in assemblies and deliveries. The module factory also replenishes the module inventory as modules are assigned to deliveries. As there usually is no need to assemble modules for deliveries, the order can be picked and ready to deliver to the customer site for installation quite quickly.

The maximum level of module inventory is tied to the weekly production capacity of a module. This makes the size of the module inventory consistent with the input variables of the simulation scenarios. It also means that less module inventory is kept in the scenarios with less capacity. The value of the module inventory is calculated by values which include the work needed in the assembly process.

The basic concept of statistical inventory management was created in the classic article written by Wilson in 1934. He broke the inventory control issue into two parts: order quantity and reorder point. (cited by: Hopp & Spearman 2000, pp. 64) This is a simple and effective way to handle the inventory control problem, and therefore the CTO process the orders for the buffered components are done with a reorder point approach. Also, this type of rate-based material planning works well upstream of the OPP, where the production is standardized with a limited set of end-products. (Vollmann et al. 2005, pp. 457-459; Saeed et al. 2016)

For the purposes of the simulation the reorder point, safety stock and lot size calculations were kept as simple as possible, disregarding the possible differences inventory holding costs and setup costs make. The more in-depth calculations for these variables can be found in chapter 5.

When using the reorder point method, a batch of items is ordered when the inventory level of an item drops below the set reorder level, which is called the reorder point. The goal is to place an order when the inventory at hand is enough to satisfy the demand for the components supply lead time. (Stevenson, W. J. 2011, pp. 578)

The reorder point calculation also includes the safety stock. The safety stock is used to maintain an adequate service level and to reduce the possibility of a component stocking out. (Vollmann et al. 2005, pp. 145) The service level in the simulation is set at 95%.

The lot-sizing policy for the buffered components is a fixed reorder quantity (FOQ). The fixed order quantity approach can yield optimal results if the demand for components is quite uniform. This can happen if it is used to order lower level components which can be used in a variety of end-products, as is the case in this process. (Stevenson, W. J. 2011, pp. 524-525)

In practice the FOQ approach means that when the inventory level drops below the reorder point, a pe-determined quantity is ordered. The order frequency will fluctuate depending on the variations in demand. (Vollmann et al. 2005, pp. 144-145) If the safety stock has been utilized to satisfy the demand in one period, the deficiency needs to be added to the lot quantity. This is done to replenish the safety stock back to the determined level. (Hopp & Spearman 2000, pp. 129)

The non-buffered items are ordered with a lot-for-lot policy. They have a rather short delivery time and due to their high variability keeping them in stock is not sensible. The lot-for-lot ordering policy minimizes the inventory kept due to the exact order sizes making sure that no extra inventory is kept (Hopp & Spearman 2000, pp. 125).

4.2 Make-To-Order Process

The MTO process aims to reduce the inventory needed by delaying the ordering of the components needed until the customer order has been received. The process is not a pure MTO process as it does buffer some components to keep the customer delivery time reasonable, but otherwise it resembles a classic MTO process very closely. The process is presented in figure 13.

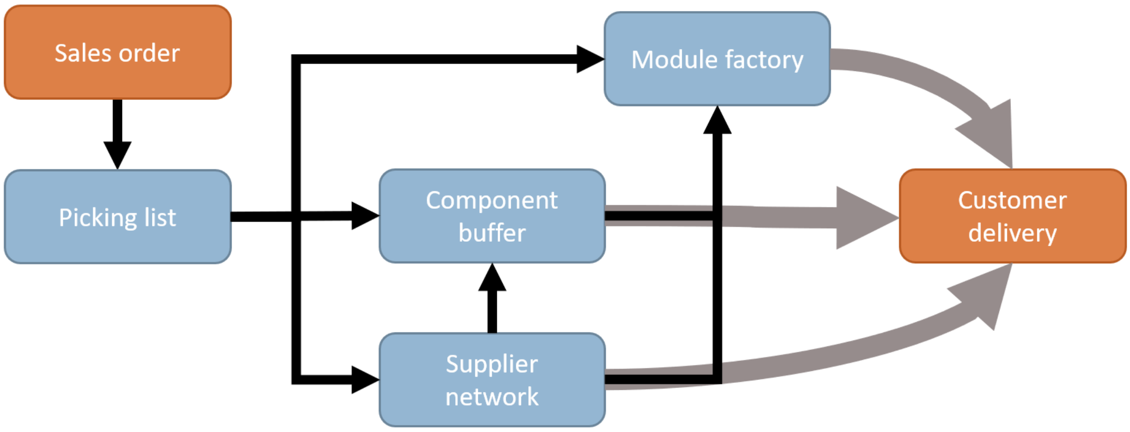


Figure 13: *The outline of the MTO process*

This process has the same elements as the CTO with one key part missing as in this process the modules are assembled after the order has been placed there is no module inventory. The process starts similarly with the sales order and picking list which state which components and modules are needed to fulfil the order.

The picking list triggers the component orders for the parts which have a shorter delivery time than the promised system delivery time. It also launches production orders for the different modules, but the assembly is not done before the ordered parts arrive. To keep the system lead time reasonable, the parts which have a longer delivery time than the promised system delivery time need to be buffered. Some components have quite long delivery times, so the component buffer needs to be large enough to account for that. After all the modules are assembled and the other components arrived, the system can be delivered to the customer for installation.

In the MTO-process, no module inventory is kept, and all the components are ordered with a lot-for-lot policy. This enables the minimizing of inventory as the exact need for the components is known when the orders are placed (Hopp & Spearman 2000, pp. 125; Stevenson, W. J. 2011, pp. 524).

The component buffer inventories are kept at levels which will enable an adequate service level. They are calculated in a similar way to the CTO process with the goal to reduce the probability of stocking out. The buffer inventory is essentially set at a level which the process is likely to use up during the difference between the supply lead time and the delivery lead time.

4.3 Summary

This chapter summarizes the differences between the processes. The main difference is that the CTO process keeps buffer inventories at both the component level and finished module level, while the MTO process manufactures everything only after the customer order. The differences are summarized in table 2.

Table 2: *Summary of differences in the processes*

	CTO	**MTO**
OPP	Picking	Purchasing
Purchasing	Reorder point	Customer order
Lot-sizing (normal components)	Fixed order quantity	Lot-for-lot
Buffer inventories	Most components, Shuttles, APs	Components with long lead times.
Production strategy upstream of OPP	Pull. FG deficiencies launch production orders for assemblies.	None
Production strategy downstream of OPP	Push. Customer orders launch order picking from inventories.	Push. Customer orders launch module assemblies and delivery picking.

In the MTO process, the OPP is set at purchasing. The customer order launches the purchasing of the needed components. In the CTO process, the OPP is set at picking. At this point the variable frame parts are ordered and the other items are picked from the inventories, if available. The purchasing of the rest of the components is triggered by the reorder points set for the components. If after a module assembly the inventory level of a component dips below the reorder point, a replenishment order is placed.

The lot-sizing is fixed order quantity for CTO and lot-for-lot for the MTO process. The fixed order quantity was chosen due to its advantages in rate-based operations upstream of the OPP. In MTO, the lot-for-lot was chosen to minimize inventories. The variable frame components are ordered lot-for-lot in both processes as it is the most sensible option for items which have high variance and short supply lead time.

Both processes keep buffer inventories at the component level. On the MTO side the buffer inventories are held to enable shorter customer lead times. The CTO process keeps them to enable pull-based module inventory replenishment. The CTO process also keeps module inventories for smoother capacity utilization and faster delivery times.

The CTO process utilizes rate-based pull production strategy upstream of the OPP and a push strategy downstream. Upstream the assemblies are triggered by the need for module inventory replenishment and downstream the customer order pushes the picking of deliveries into the schedule. The customer order launches both the assemblies and picking of deliveries in the MTO process.

5. SIMULATION

The simulation was implemented by creating a model that emulates the processes of the Agilon factory. The simulation model was created with a python script. The script simulates a set number of weeks during which the demand of Agilon systems is randomly generated within the set minimum and maximum demands. During the simulated weeks the Agilon factory works to meet the demand by assembling the needed modules and keeping inventories at set levels. The factory has a predetermined capacity to produce different modules each week.

The configurations of the systems that the demand consists of are based on the currently installed Agilon systems and predictions of future demand. The different modules that make up the systems will have simplified one-layer BOMs. The components in the BOMs are the most used and valuable items to get an approximation of the capital investment needed in different situations.

The simulation comprises of two different processes which were introduced in the previous chapter. The simulation includes several sets of starting variables that predict the possible future scenarios that the Agilon unit might face. Each scenario will be executed several times with both processes to get an appropriate approximation of the circumstances and possibilities (Robinson 1994, pp. 169).

The script keeps track of the inventory levels, produced modules and delivery reliability. It saves the accumulated data to Microsoft Excel-workbooks which will be used to analyze the results of the different scenarios. This chapter describes the different aspects of the simulation.

The inventory values in this chapter were substituted with ratios by dividing the inventory values with an arbitrary number. This provides a way to compare the values shown in this research. The ratios are consistent throughout the thesis. This was done to exclude any confidential information from the final version of the thesis.

5.1 Limitations

Because simulating the real world with perfect detail is nearly impossible, several simplifications and assumptions need to be made to make the simulation practical (Klee & Allen 2011, pp 3; Laguna & Marklund 2013, pp. 99).This chapter goes through these to create an understanding of the scope of the simulation.

The data required to accurately model a probability distribution for demand is usually not readily available, as is the case in this study (Chen-Ritzo et al. 2010). Due to this, the demand was set to be randomly distributed between the minimum and maximum demands for each week. The data of the past demand simply is not comprehensive enough to give a good enough approximation of the demand distribution, especially when it would need to be projected to the scenarios with vastly higher maximum demand.

The simulation assumed that the suppliers have unlimited capacity and supply lead times remain constant. The simulation focused on the operations in the Agilon factory, and therefore the supplier's performance was not relevant. This was due to the research focusing on demand-variability.

The manufacturing lead time of the different modules was also constant. There were no complications in the production apart from delays caused by possible stockouts. If the needed components were available and there was capacity left on the week, the module could be produced. If not, the module was not manufactured and the eligibility for assembly was checked again the following week.

Because the frames are assembled after the delivery at the customer site, the frame-modules were handled a bit differently to the access points and robots. The weekly manufacturing capacity of all the different frame types was infinite as they are only picked from inventory and little to no other production steps are needed.

The module assemblies were set to happen at the end of the week and as such the inventories changed on a weekly basis. This enabled benefiting from the advantages of discrete-event nature of the simulation, e.g. simpler structure and time-compression (Laguna & Marklund 2013, pp. 257-259).

The order processing, inventory handling and picking was assumed to happen instantly. For example, when a lot of items arrives, it is immediately placed into available stock. Similarly, the deliveries that have all the needed modules are instantly taken away from the finished goods inventory and marked to be delivered.

The promised delivery times might change depending on the current order situation or for example if the customer orders the system to be delivered a couple of months from now. In the simulation the customer delivery times were constant, and the system was marked as late if this time is exceeded.

The inventory capacity of the Agilon factory is assumed to also be infinite. This is done because the aim of the simulation is to see how the inventory value will fluctuate in the different scenarios and restricting the inventory levels might affect the results.

5.2 Modelling uncertainty

This chapter goes through how the simulation mimics demand uncertainty. The systems the simulation creates for the demand are based on the currently installed Agilon systems. The demand uncertainty can be divided to demand mix and demand quantity uncertainty.

The quantity of weekly system demand is based on a uniform distribution ranging from zero to the maximum demand in each scenario. In practice it means that if the maximum weekly demand of a scenario is for example 4, the weekly demand can be 0, 1, 2, 3 or 4. The possibility for each is therefore 20%, as there are five options. This is a simple yet effective way to emulate the fluctuation of weekly demand, as there is not enough data to project past demand to the higher demand scenarios in a reasonable way.

In this simulation the mix uncertainty derives from the way the system configurations are modelled. Simplifications were made to decrease the complexity of the system modelling. The more unusual features of the installed systems such as tubes and underpasses were deemed redundant for the purposes of this study.

Chen-Ritzo et al. (2010) note in their simulation study that a relatively reliable distribution for the product configurations can be determined with data from the previous deliveries. This method was utilized to calculate the probability distributions which determine how the simulation creates configurations for the systems.

The two Agilon product types are industrial and retail. Based on the number of installed systems, the ratio of product type demand was determined to be roughly 75% to 25%. This means that about three quarters of the systems are likely to be industrial systems and the rest retail systems.

Considering the simplified structure for the systems in the simulation, the three distinct factors making up an Agilon system are the dimensions of the frame and the number of robots and access points. The calculations for these factors are shown next.

It should be noted that because there is much less variation in the configuration of retail units the simulation only creates retail units have 1 robot and 1 access point. This is true for all but one installed retail unit. The size of the average retail frame was calculated from all of the installed retail units.

The industrial frame sizes were generalized into three groups. These three are small, medium and large frames. The dimensions for them were determined by dividing the previously installed industrial systems into three roughly equal-sized groups by their length and calculating the average height and length in each group. One type of retail system was decided to be enough for the simulation as there is a smaller number installed and less variation in the configurations. The dimensions and probabilities of the frame types are shown in table 3.

Table 3: *Dimensions and probabilities of the frame types*

	Height (m)	Length (m)	Probability
Small	5,6	8	25,75%
Medium	5,6	14	25,75%
Large	5,1	24	23,5%
Retail	3,1	4	25%

The dimensions were used to define the number of the variable components needed to build them. The probabilities for the different industrial frames were calculated from the number of units in each group. The percentages are not the same due to the number of installed units not being divisible by 3. They add up to the 75% of total systems, which was the probability of a system being an industrial system.

The access point count in currently installed systems ranges from 1 to 5, but the range was limited from 1 to 3 as systems that have more than 3 are rare with less than 10% of the systems having more than 3 access points. Due to this, the systems which have more than 3 access points were seen as having 3 for the purposes of the calculations in this section. The distributions for the number of access points in different frame types are shown in table 4.

Table 4: *Probability of the number of access points depending on the frame size*

Access points	1	2	3
Small	74 %	26 %	0 %
Medium	39 %	43 %	17 %
Large	52 %	33 %	14 %
Retail	100 %	0 %	0 %

Table 4 shows that small systems can only have 1 or 2 access points, while the larger units can have any of the three options. The probabilities for the number of access points in medium and large systems are quite similar. This is due to the number of access points generally being dictated by the intended use of the system rather by its size.

The robot count in the installed systems can be either 1 or 2. The likelihood of having two robots increases in the systems with more access points. As before, the probabilities were calculated from the installed systems. The distributions of robots in the systems by access point count can be seen in table 5. The total values are used in calculating the module demand for robots.

Table 5: *Probability of the number of robots depending on the number of access points*

Robots / APs	1	2
1	100 %	0 %
2	69 %	31 %
3	43 %	57 %
Total	88,7%	11,3%

As table 5 shows, systems with only one access point can only have 1 robot as it would be rather nonsensical since adding more would not increase the speed of the system as the robots would need to wait for each other to clear the access point area before getting to drive past or access the access point.

With the parameters showcased in this section, there are 13 possible end-product configurations. The configurations are shown in figure 14. The possible configurations can be derived by following the arrows. For example, a small industrial unit can only have one or two access points while the larger industrial frames can have 1, 2 or 3. The same can be seen from table 4.

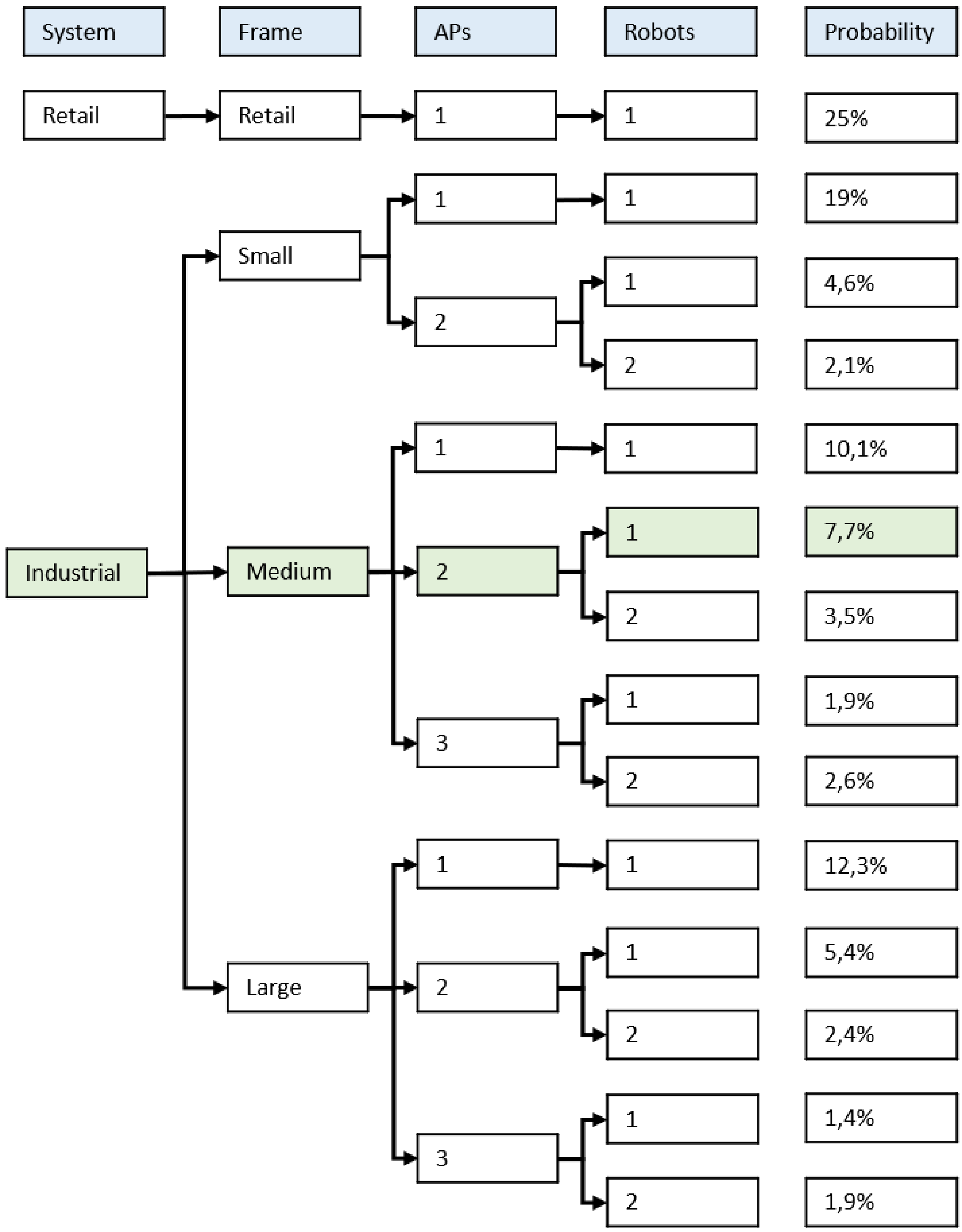


Figure 14: *All possible Agilon configurations in the simulation*

The total probabilities in figure 14 are calculated with the probabilities from tables 3, 4 and 5. The same graph with the possibilities next to the choices can be found in appendix D. The probabilities of the choices are multiplied to get the result. An example calculation for an industrial unit, which has a medium frame, two access points and one robot, is shown in table 6. The example route is highlighted in figure 14 with a green color.

Table 6: *Example calculation for configuration probability*

Choice	Probability
Medium frame in an industrial system (Table 3)	25,75%
Two access points in a medium frame (Table 4)	43%
One robot in a system with two access points (Table 5)	69%
Total probability	7,7%

To generate the average industrial system models for capacity and purchasing planning, the previously presented distributions were used to calculate the average amounts of robots and access points for an industrial and a retail system. The average amounts are shown in table 7. These numbers work as a pseudo-planning BOMs as the aggregate system demand forecasts are multiplied by these factors to get predictions for the corresponding number of robots and access points. The actual calculations for the capacity and purchasing can be found in section 5.4.2.

Table 7: *PBOM multipliers for robots and access points*

System	Robots	Access points
Industrial	1,15	1,55
Retail	1	1
Total	1,11	1,41

As seen in table 7, the average robot count for the industrial Agilon units is 1,15. Most of the systems have only one robot. The average for access points if 1,55 which means that the systems are likely to have either one or two access points. Due to the 25%-75% split between retail and industrial system demand, the weighted averages of the total multipliers are closer to the industrial values.

5.3 Modelling the bills of material

This section describes how the Agilon BOM was simplified to make it practical to manage and how the components were rated in order to pick the most impactful ones to be tracked in the simulation. The item and BOM data used in the simulation was gathered mainly from the company's ERP-system. For each component in the modules' BOMs, the price and delivery time were collected. Also, the values for assembled APs and robots were drawn from the ERP-system. The assembled frame values were estimated from different sized frames, due to no exact precedents being available.

Due to the ERP-system being quite new, all the items didn't have up-to-date information yet. The missing information was gathered either from recent purchase info records and the system used before the current ERP or generalized from the supplier's usual price level and delivery performance.

The BOM of each module consists of multiple subassemblies. Simulating the production of each subassembly is not practical and the scope of the simulation calls for a more simplified approach. Therefore, the BOMs were flattened to just a list of all the components used in assembling the module and its subassemblies.

To keep the simulation as simple as possible, the BOMs assigned to the modules needed to be fixed BOMs. The robot and the two access point types have a fixed BOM, which could be used without any changes.

The frames require variable quantities of components depending on the dimensions, which made determining the BOM harder. An existing BOM with the exact dimensions detailed in table 3 was utilized for the retail frame. There were no exact matches for the dimensions of the industrial frame types, so a variable BOM was approximated from previously delivered systems' frames. Required quantities of variable components were averaged for each frame size to turn the variable BOM into the fixed form needed in the simulation.

It has been shown that the Pareto principle, also called the 80-20 rule, can be applied to predict the distribution of many things. The division of invested capital between different components in a company's inventories is no exception. Basically, it means that a small fraction of the components held in inventories make up almost all the value. (Hopp & Spearman 2000, pp. 587) The Agilon factory is also not an exception, as all the modules have dozens, even hundreds, of components, most of which are quite insignificant to the overall cost of the module and the total capital tied up in inventories. Therefore a reasonable approach is to focus the inventory control efforts based on the relative importance of the items (Stevenson, W. J. 2011, pp 563).

The lists were refined with a method closely resembling the multiple-criteria ABC analysis method detailed by Vollmann et al. (2005, pp. 157-159). Two criteria were used: price and supply lead time. Supply lead time was decided to be the most pragmatic choice for the second factor as using other non-cost criteria can be quite complex. Also, the cost and the order interval of a component dictate the average investment in inventories, which fits the focus of the study well. (Vollmann et al. 2005, pp. 140, 158)

When refining the item lists, first all the bulk items such as screws and bolts were discarded. These components fall into the C-category, which makes up for most of the actual number of components, but only a fraction of the cost. They are also managed in bulk, which means that individual delivery times and prices were not readily available.

All the items were evaluated as they appeared in the data gathered from the ERP, with one exception. The cover sheets used in covering the frames come in a multitude of sizes and shapes in addition to the basic sheet. Individually the less used cover sheets would not be significant, but together they make up a big portion of the total cost due to the high amount needed in a system. Also, they all come from the same supplier with the same supply delivery time. They were decided to be clumped together with the basic sheet with an average price to get a better estimation for the impact of the total investment in cover sheets.

After this the actual component selection was performed for each module. To consider both criteria, the prices and supply lead times were multiplied together to get a single value which indicates the components relative importance. If the module required more than one component of the same item, the price was multiplied with the required number. In the case of the variable frame components, the average quantity needed was used for this selection.

The values were then sorted in ascending order. A threshold value was determined for each module individually to select an adequate number of components which resemble most of the value. The 80-20 rule worked as a rough guideline for deciding the threshold.

Figure 15 shows the distribution of robot components according to the decision criteria. The values have been normalized by dividing with the mean of each category. Majority of the components is clumped at the bottom of the graph close to the origin of the graph. It also shows that the decision clearly manages to select the items with either a long supply lead time, a high price or both, while excluding the least significant items.

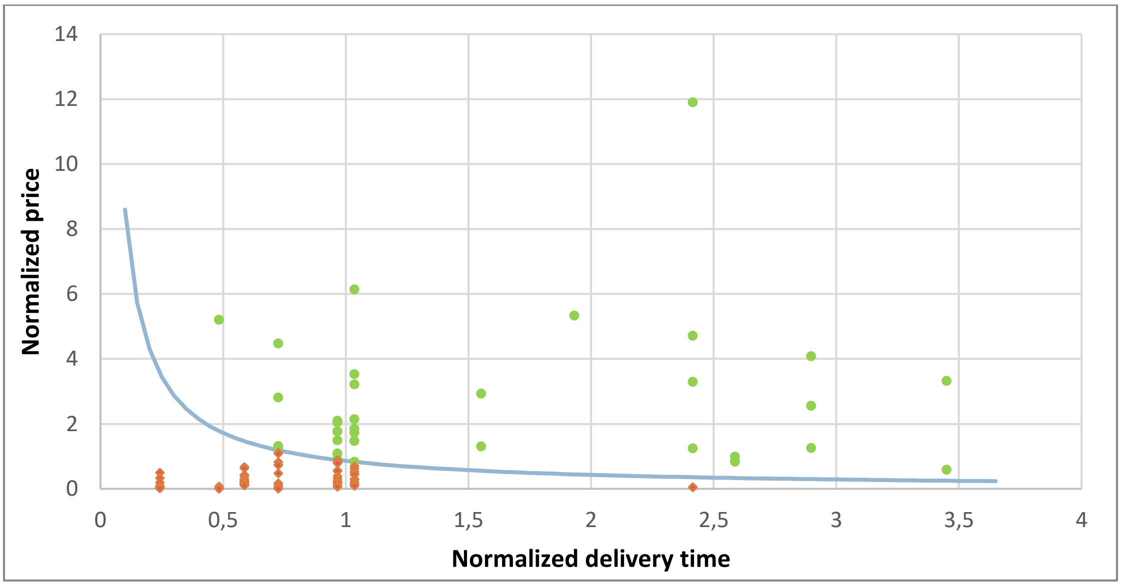


Figure 15: *The distribution of components in a robot*

The threshold values were determined individually instead of a single value for all modules to make sure that all modules got a decent sized list of components to be tracked. Table 8 shows that the percentage of number of components chosen and the percentage of the total value of the components roughly fits the pareto principle for each module.

Table 8: *The component decision variables for the modules*

	% of items chosen	% of total value
Robot	23 %	75 %
AP	25 %	75 %
Retail AP	23 %	79 %
Retail frame	30 %	76 %
Industrial frames	-	-

The industrial frames don't have the percentages marked in table 8 due to the BOM being an approximation and therefore an unambiguous value could not be determined. Most of the total value was still definitely tracked.

5.4 Calculating key variables in the simulation

This section goes through how the simulation calculates the reorder points, lot sizes, safety stock levels for the individual components. However, before explaining them the preliminary demand calculations are explained as they are needed to calculate the inventory control variables as they are calculated differently in each scenario based on the demand level.

5.4.1 Module demand

This chapter explains the concepts of average and maximum module demand as they play a major role in the subsequent calculations. The simulation uses the weekly demand values for its calculations. In this chapter the maximum demand means the service level adjusted maximum demand.

An important notion in this section is the difference between independent and dependent demand. The independent demand is the demand which comes from the customer while the dependent demand is the demand for the components which make up the end products (Hopp & Spearman 2000, pp. 110). In this study the system demand is the independent demand for whole Agilon systems and module demand is the dependent demand which is derived from the system demand.

As the independent system demand distribution is uniform, the average system demand is half of the maximum system demand. The average dependent demand for the modules can be derived by multiplying the average system demand with the probabilities of the system including the module, which can be found in tables 3 and 7. Table 9 shows an example of the average module demand for a scenario in which the weekly max demand is 8 and the average system demand is therefore 4.

Table 9: *Average weekly module demand levels in a scenario where average system demand is 4.*

Module	Multiplier	Average demand
Robot	1,11	4,44
AP	0,75*1,41 = 1,0575	4,23
Retail-AP	0,25	1
Frame-S	0,2575	1,03
Frame-M	0,2575	1,03
Frame-L	0,235	0,94
Frame-R	0,25	1

For example, the industrial access points multiplier is the product of the probability of an industrial system, which is 75%, and the average number of access points an industrial system has (Table 7). Similarly, as every system has at least one robot, the average system demand is multiplied with the average number of robots in a system (Table 7).

The maximum dependent module demand used in the calculations is adjusted to meet the service level of 95%. The method used to set the maximum demand is derived from the stockout probability method presented by Vollmann et al. (2005, pp. 144-145) Service level is used to create a compromise between the probability of stocking out and inventory investment. Buffering against the full variation can be very impractical and costly. (Vollmann et al. 2005, pp. 144-145) The basic idea of the method is to utilize a maximum demand value for which there is a 95% probability for the weekly demand not to exceed it.

The demand probability distribution is needed to adjust the maximum demand (Vollmann et al. 2005, pp. 144). The module demand distributions can be derived analytically, but it is quite laborious. Calculating the distribution can be done in a similar way for all the modules except the industrial access point. The reason for this is that the other modules have two options in any one system, 1 and 2 for robots and 0 and 1 for frames, while the number of access points can range between 0 and 3.

Due to these reasons the probability distributions used in the simulation were decided to be approximated with a numerical method. This method could be generalized to fit all the modules and all demand scenarios, and it gives an approximation which very closely resembles the corresponding theoretical distribution.

This section focuses on describing the numerical method and only a brief overview of the analytical method is given. The way to derive the analytical solution and the comparison between the numerical and analytical methods can be found in section 5.7.5.

The total probability distribution for a module's weekly demand can basically be calculated as the sum of (n+1) equally probable binomial distributions, where n is the maximum system demand in the scenario. A binomial distribution is formed by a series of yes/no questions where the probability of a yes-answer is constant. The probability for each number of yes-answers can be calculated which forms the distribution. (Florescu & Tudor 2013, pp. 92)

In the case of industrial access points there are more than two possible outcomes, as the number of access points can range between 0 and 3. Due to this the distributions used are not binomial but rather multinomial. A multinomial distribution is a generalization of the binomial distribution where the number of dimensions can be more than 2. They work with the same principle but with a different equation to calculate the result as the set of possible outcomes is larger than 2. (Florescu & Tudor 2013, pp. 219) For the purposes of the numerical method this doesn't make a difference, since the weekly demand can be drawn from 4 choices as easily as 2.

To mimic the distribution, the simulation calculates an arbitrarily large set of weekly demand patterns, which can then be used to determine the probability for each eventuality. Table 10 represents an example distribution of dependent robot demand in a scenario with the maximum system demand is 8.

The total demand was calculated total of 117000 times, 13000 for each independent system demand level. This means e.g. that there were 13000 pairs of values for each demand level which show how many systems had 1 robot and how many had 2 robots that week. The probability of a system having 1 or 2 robots is 88,7% and 11,3% respectively, as shown in the total value row in table 5.

Table 10: *Data from the example simulation of weekly robot demand*

Demand	Systems	Robots	Probability	Cumulative
0	13000	13000	11,11 %	11,1 %
1	13000	11566	9,9 %	21,0 %
2	13000	11565	9,9 %	30,9 %
3	13000	11849	10,1 %	41,0 %
4	13000	11576	9,9 %	50,9 %
5	13000	11707	10,0 %	60,9 %
6	13000	11603	9,9 %	70,8 %
7	13000	11699	10,0 %	80,8 %
8	13000	11738	10,0 %	90,9 %
9	0	7298	6,2 %	97,1 %
10	0	2636	2,3 %	99,3 %
11	0	656	0,6 %	99,9 %
12	0	100	0,1 %	99,99 %
13	0	7	0,01 %	100 %
14	0	0	0 %	100 %
15	0	0	0 %	100 %
16	0	0	0 %	100 %

Few examples to clarify the results in table 10:

- As there was 13000 weeks where the system demand is 0, there was also 13000 weeks where the robot demand is 0. The probability of robot demand being 0 is therefore 11,11%.
- 13000 weeks where the system demand is 1 were also simulated, but there was only 11566 weeks where the robot demand was 1. This occurs because the single system which needs to be delivered on that week can also have 2 robots instead of just 1.
- The theoretical maximum demand of robots with this level of system demand is 16, but the probability of that is very low. That's why the simulation did not create even one week with that level in the whole run. The largest number of robots needed in a single week was 13, with 7 occurrences and a probability of one-hundredth of a percent.

Figure 16 represents the probability distribution. The distribution is quite uniform until 8, and then drops of quickly. This is also evident from the data in table 8.

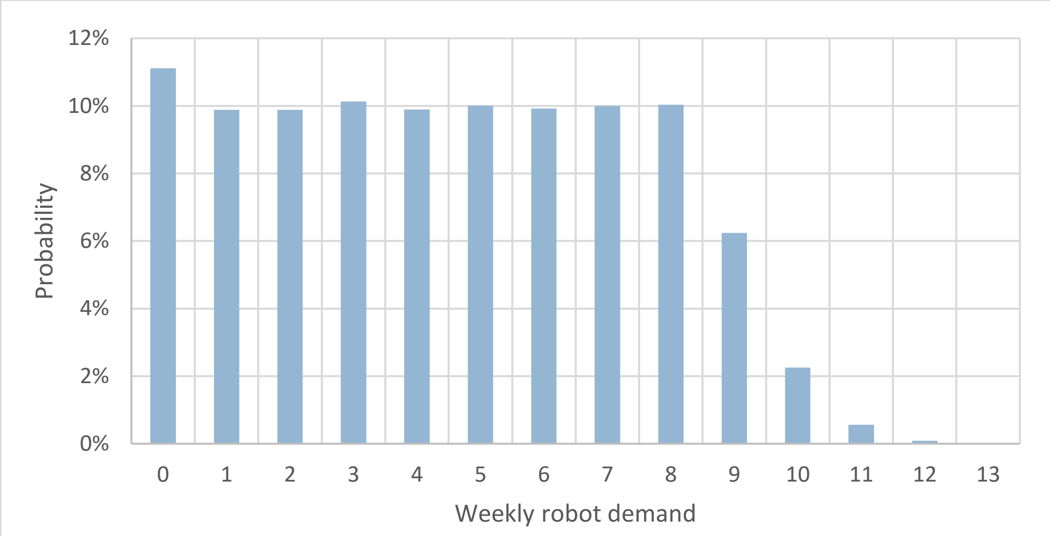


Figure 16: *Probability distribution of robot demand in the example scenario*

The way to determine the service level adjusted maximum demand is shown in figure 17. It shows the data from table 10 as a cumulative function. The service level adjusted maximum demand is met when the cumulative line crosses the orange line which represents the 95% mark.

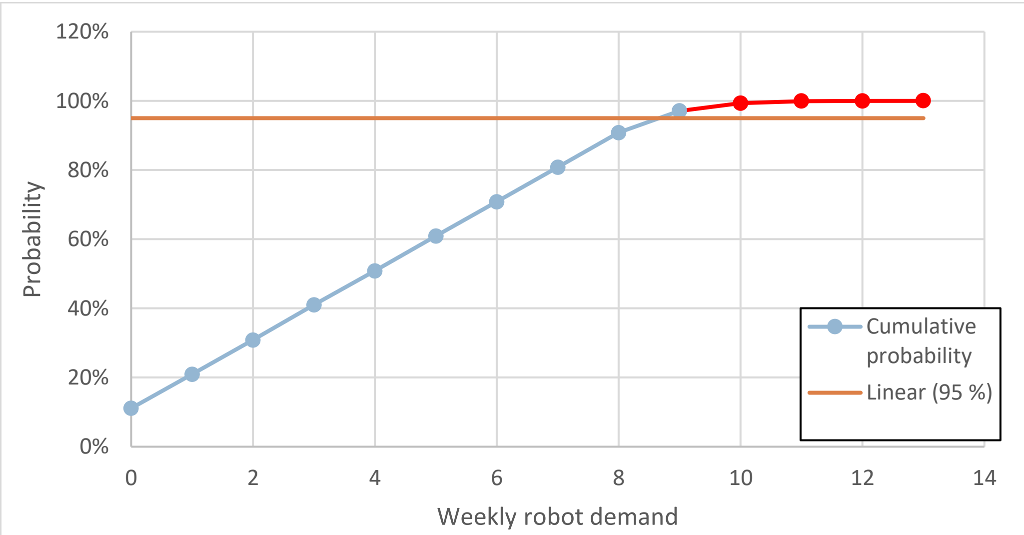


Figure 17: *Cumulative robot demand distribution function of the example*

The lines cross between 8 and 9. This means that according to this approximation, there is less than 5% chance of the demand being higher than 9. Therefore, the maximum robot demand that needs to be considered is 9. Table 11 shows the service level adjusted and theoretical maximum demand levels the other modules get assigned in a scenario with maximum demand of 8.

Table 11: *Example of the maximum dependent demands for all the modules when the maximum independent system demand is 8*

Module	Theoretical	Service level adjusted (95%)
Robot	16	9
AP	24	11
Retail-AP	8	3
Frame-S	8	3
Frame-M	8	3
Frame-L	8	3
Frame-R	8	3

The similar service level adjusted values for retail access points and the frames are explained by the rather similar probabilities for a system to have one, roughly one fourth for each. Because there is the most variation in the number of industrial access points in a system, it is assigned the highest maximum dependent demand.

5.4.2 Inventory control

Calculating the inventory control variables is quite simple after the average and maximum demands have been determined. The variables are the lot sizes, reorder points and safety stocks for the individual components.

The weekly average and maximum demands for a component is derived from the demand of the modules which include the component. Multiple modules can require the same component, so the individual quantities required are added together to get the total number needed.

The reorder point is only used in the CTO process. The reorder point consists of two components: safety stock and the required level of inventory to meet the average demand in a period.

The safety stock is used to account for the possibility of a higher demand level in a period (Vollmann et al. 2005, pp. 144). In this case the reorder point is set at the point where the maximum module demand calculated in the last section will be satisfied, which means that there is only a 5% probability for a stockout in the reorder period.

$$ROP = D * t, \quad (1)$$

where D is the service level adjusted maximum demand for the component and t is the component's supply lead time.

The lot sizes in lot-for-lot approach are the exact number of components needed for the ordered system and therefore don't require calculation. The fixed order quantity lot sizes are calculated as follows:

$$LS = (d * t) + SS_d, \quad (2)$$

where d is the average demand for the component, t is the supply lead time and SS_d is the deficiency from the safety stock. The safety stock deficiency means the number of components the current stock level is less than the set safety stock level. If the stock is higher than the safety stock, then the deficiency is 0.

Safety stock is used to buffer component inventories in both the CTO and the MTO processes. The safety stock for a component is calculated from the lot size and the reorder point as follows:

$$SS = \mathrm{ROP} - (\mathrm{d} * \mathrm{t}), \quad (3)$$

where ROP is the reorder point, d is the average demand and t is the supply lead time for the component. (Vollmann et al. 2005, pp. 144) While it is not actually used in calculating the reorder point, calculating the safety stock is relevant in finding out whether the safety stock is deficient. All the values calculated with the formulas shown in this section are rounded up to the closest integer.

5.5 Inputs & Outputs

A simulation takes a set of input variables and predicts the corresponding level of output variables (Robinson 1994, pp. 6). This chapter goes through the inputs and outputs of the simulation used in this research.

The main input variables are the modules' BOMs and the scenario variables. For the purposes of this research the BOMs stay mostly constant but allowing for changes in the BOMs enables the studying of the impact of single components to the overall system. The scenario consists of three variables: maximum demand, capacity modifier and supply lead time modifier. The different scenarios are elaborated on in chapter 5.8.

The level of module inventory can also be adjusted, but it is not regarded as a full-blown input variable. This is due to it being tied to the weekly capacity in most of the simulations. As mentioned in chapter 4.1, the default value for the module inventory for robots and access points is the weekly production capacity. This will always be the case if not otherwise mentioned.

The simulation tracks multiple aspects of the Agilon factory's operations and therefore produces a lot of data. The tracked variables act as the outputs for the simulation. The main measures which are tracked are the inventory levels, delivery information and the weekly production variables.

The weekly inventory levels are tracked for each component and each module individually. This can be derived to the total weekly inventory values for the total value or the modules and components separately. This enables the estimation of the total capital tied up in inventories for either the whole inventory or a subset of the inventory. It also makes identifying the most critical components possible.

The production variables are tracked weekly for each module. They are the production amount and the demand level. These variables enable studying the utilization in the different scenarios, which gives insights into the delivery process and its performance. The delivery information is essential for studying the performance and delivery reliability of the order fulfilment process. The configuration and delivery time of each system is saved. The overall service level and tardiness can be calculated from this data.

5.6 Flow logic

This chapter goes through the simulation flow logic. The first section explains the initialization and how the scenarios are cycled through. The key points are the creation of demand and calculating the key variables according to the rules distributions and formulas presented in the previous sections of this chapter. They explain how the week cycle is modified to match each scenario.

Section 5.6.2 dissects the week cycle into smaller sequences for more understandable explanation. Disconnecting the pieces of the flowchart into separate entities can be confusing at times, so it is recommended to occasionally check how the parts combine from the full flowchart. The full flowcharts for the CTO and MTO week cycles can be found in appendices D and E respectively.

The markings for the flowchart are done in a uniform way:

- The blue boxes represent a point in which an action is done where the state of one or multiple variables changes.
- The yellow diamond-shaped boxes indicate a choice. The flow can divert into two different paths depending on the answer.
- The orange boxes are a way to compress a complex pattern into a single box to maintain the scope and simplicity of the flowchart. An example of this can be seen in appendix C, where the whole week cycle is compressed into one box to keep the scenario cycle flowchart simple and readable.

- Lastly, the grey boxes mark points which connect the parts to the higher-level orange boxes. An example of this is the grey boxes in figure 18 which connect to the orange box in appendix C. These connections are not always obvious, but they are always detailed in the paragraphs before or after a figure containing an orange or grey box.
- The white boxes with grey outlines are comments and not actually a step in the flowchart. They mark some important pieces of information for the step which they connect to.

5.6.1 Initialization & Scenario cycle

The initialization sequence flow chart is shown in Appendix A. The simulation starts by reading the BOM data from an excel file. The data consists of a list of items which have an ID, price, supply lead time in days and a list of modules it is used in with the individual requirements for each.

When the simulation begins, it is given a set of scenarios to be run. The scenarios are combinations of the demand level, supply lead time modifier and the level of capacity buffering. It then creates lists of demands for each scenario. The lists are created with the demand level and the rules and probabilities from section 5.4. Each scenario can have an arbitrary number of demand lists. The demand patterns are created in advance to enable the use of the same lists for both CTO and MTO simulations which are going to be carried out. The implementation logic is elaborated on in section 5.8. While the lists are created in advance, the actual “customer orders” become available for the Agilon factory periodically throughout the week cycle.

After the demand patterns are created, the simulation calculates the key variables needed for the scenario in question. The service level adjusted max demand levels are simulated as explained in section 5.4.1 and after that the corresponding safety stock is set for each component with formula 3.

The supply delivery times of each component are adjusted according to the delivery time modifier of the scenario. The starting stock is determined differently in CTO and MTO. The MTO starting stock is only the safety stock of the buffered items. In CTO the starting stock also includes a cycle stock with enough components to respond to demand during the different components supply lead times. This means that the inventory starts in a state in which every component is at the reorder point.

After the initialization is done, the week cycle is carried out. When the week cycle has run its course, the data is saved. If there are more scenarios to run, the variables and data structures are reset. If the demand level doesn't change, the previously calculated demand values can be used again, otherwise they are calculated again. Then the cycle continues until each scenario has been run. After everything the last things are written to file and the simulation ends.

5.6.2 Week cycle

The week cycle is the part of the simulation which mimics the operations of the Agilon factory. A simplified illustration of the week cycle is shown in figure 18. The week cycle consists of three parts: Incoming order handling, production & purchasing and the end-of-the-week sequence. The cycle continues until the end of the week sequence decides that there are no more weeks left and ends the scenario.

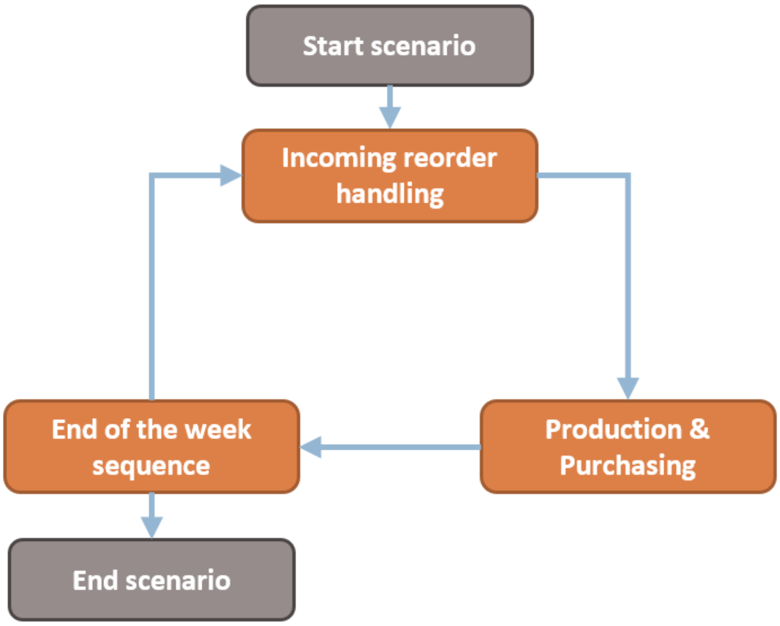


Figure 18: *The simplified week cycle*

The incoming order handling tracks incoming orders daily and adds them to stock if some arrive. The days are incremented by one after the incoming orders are checked. When seven days have passed, the production & purchasing phase starts. The incoming order handling is illustrated in figure19.

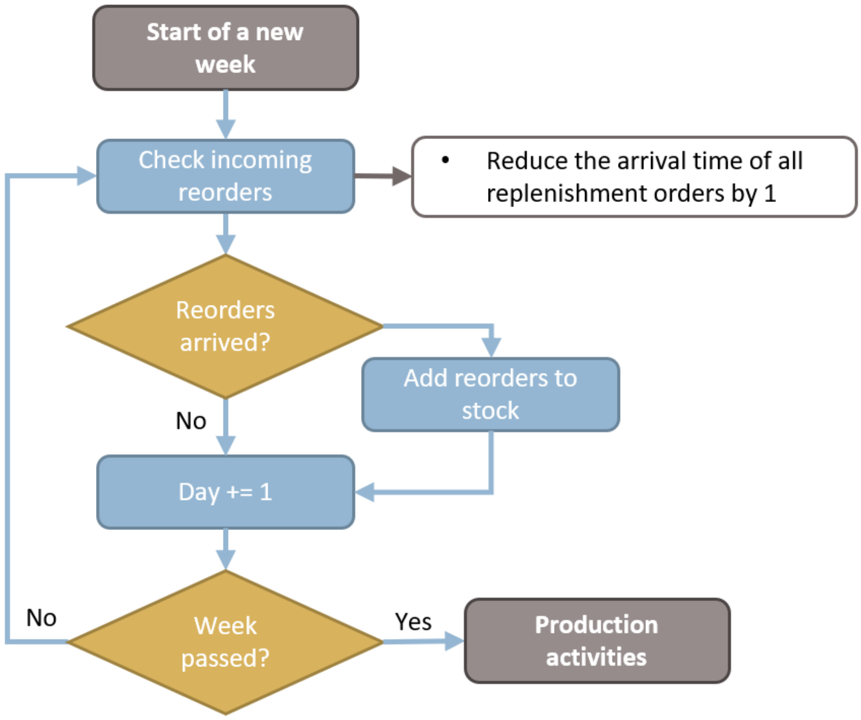


***Figure 19**: Incoming order handling*

Production & purchasing sequence, shown in figure 20, conducts the production activities for the week and records the changes in inventory levels and weekly capacity. The weekly need is calculated by checking which systems in the demand list are ordered. These are basically the production orders for the week. A system is applicable to be produced when its delivery week is less than the delivery time away from the current week. This enables the tracking of the customer delivery time and tardiness.

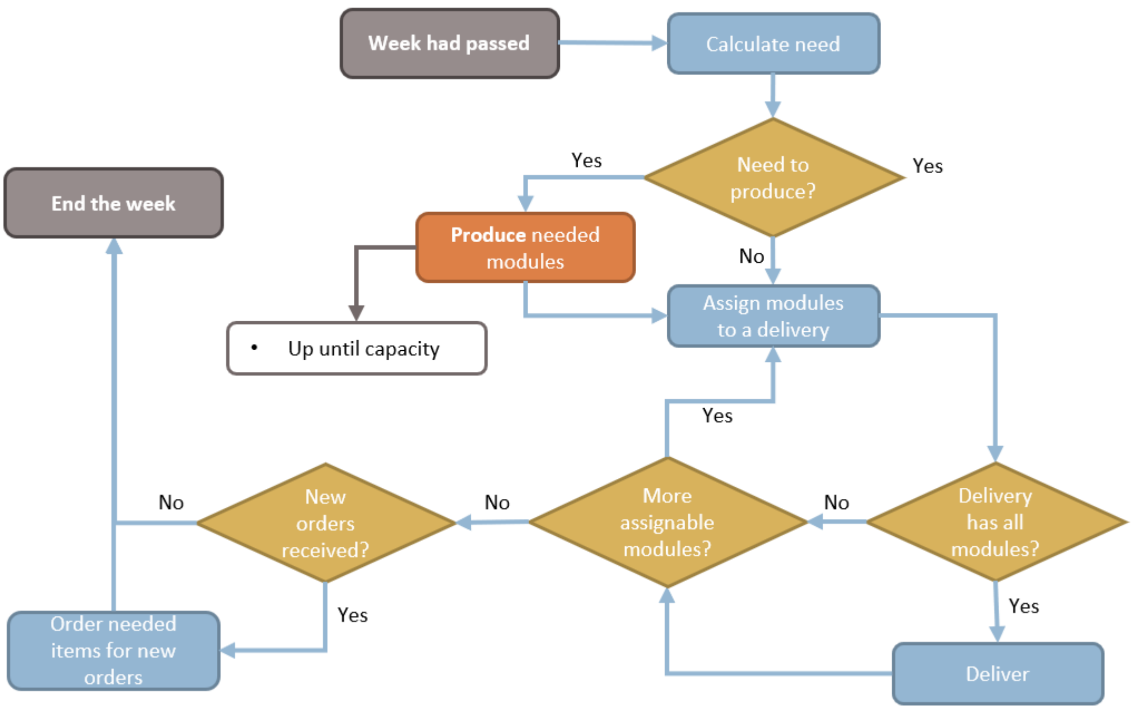


***Figure 20:** Production & purchasing sequence*

In the CTO process, the module inventory is replenished before the need is calculated. To fulfill the weekly need, the modules are taken from the module inventory first, and then remaining capacity is then used to produce modules. This can be seen from the slight differences between figure 20 and the same part of the CTO flow chart in appendix D.

There orange produce tile in figure 20 represents the production tasks portrayed in the lower right corners of appendices D and E. This means that for example when the CTO-process reaches the orange box in figure 20, the steps depicted inside the corresponding production task -box are taken next, before returning to follow the flow in figure 20. The production tasks differ a little due to the purchasing practices. In CTO, reorders are sent when a component's stock drops below the reorder point. The reorder point and lot sizing are done with formulas 1 and 2 from section 5.4.2.

The production is basically done by checking if there are enough components in stock and then reducing the stock with the corresponding quantities. The module's weekly production capacity is also reduced by 1 and the module is placed into the module inventory. This is repeated until all the module production orders, which are listed in the calculate need -phase.

After production the finished modules are assigned to deliveries and deliveries are delivered if they have all the needed modules. Then new orders for the week are checked. If there is new due four weeks from the current week, orders are placed for them. In CTO only the variable frame items are ordered while in MTO the order includes every component needed in the system.

When all the production is done and orders are placed, the end-of-the-week sequence commences. It is detailed in figure 21. It begins with calculating the warehouse value. This includes calculating the total value and the values of the component and module inventories. The data of the week's production, deliveries, inventory levels and inventory values are written to an excel file at the end of the week. The week counter is then incremented by 1. If there are still weeks left, a new week is started. Otherwise the scenario is ended, and the flow continues as shown in appendix C.

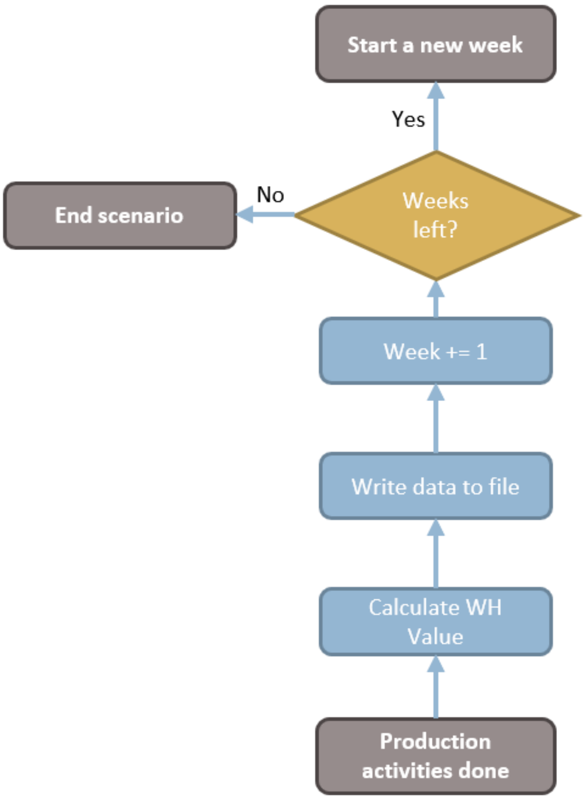


Figure 21: *End-of-the-week sequence*

The flow of the simulation is quite complex. Therefore, making sure it works correctly and consistently is important. The different aspects of the simulation are critically evaluated in the next section.

5.7 Validation

This chapter evaluates the validity and the performance of the simulation. This is not an easy task. For example Sterman (2000, pp. 846, 850) states that validating or verifying a simulation model is impossible, as all models are inherently flawed and limited representations of reality. He does add that this does not mean that efforts to validate a model shouldn't be undertaken. When using a model to help in decision making, rigid validation can help in showing the model's usefulness and revealing the limitations it has. This information can be then utilized when analyzing the results of the model. (Sterman 2000, pp. 846, 850)

While the validating the simulation is not easy as there is little applicable data available to compare the simulation results to, there are still ways to evaluate the validity of the simulation model (Robinson 1994, pp. 142). As stated by Sterman (2000, pp. 850), a model should be tested by seeking ways to compare the model to the real world and utilizing a wide range of test.

This chapter strives to provide a brief overview of the tests conducted in validating and evaluating the usefulness of this simulation model. The first section in this chapter compares the simulation's inventory investment results to a recent inventory conducted in the case company. The second section evaluates the correctness of the inventory calculations against the theoretical values. Next, the consistency of the simulation results is evaluated by running the simulation a large amount of times and studying the spread of the results. Then the stability of the processes is studied. Both methods are complemented throughout with the face validity analysis. Lastly, the validity of calculating the service level adjusted maximum demands with the simulation method is evaluated against the analytical method.

In addition to the tests described in this chapter, the model was carefully scrutinized at each step of the model creation process, testing each implemented functionality thoroughly. The model was run countless times with different sets of input variables and BOM structures to seek out deficiencies and to adjust the operation to provide useful insights into the problems laid out by the research questions.

5.7.1 Comparison to the real system

A good way to validate a simulation is to compare it to the real system it is trying to mimic (Robinson 1994, pp. 143). Comparing the simulation with the real system is not easy, as there is no continuous data of the past component stock levels available. However, the Agilon factory recently carried out an inventory which gives a snapshot estimation of the actual component stock levels.

This comparison can't give an absolute answer to the accuracy of the simulation, but it provides some evidence that the scale of the results compares adequately to gain practicable information. Therefore, it is quite essential to the validation of the simulation.

10 simulation runs were carried out with the characteristics of the current situation. To compare the value given from the simulation to the real one, the inventory levels need to be adjusted as not all the items are tracked. This can be done by calculating an estimation of the distribution of how much the different modules' components take up in the inventory. This estimation can then be divided by how much of the value in a module is tracked, which can be found in table 8. A dummy value of 0.75 was used for industrial frames.

A more detailed breakdown of the results of this test was available during the assessment of the thesis but it was redacted from the final version. Table 12 shows the differences between the simulation test runs and the real system.

Table 12: *Comparison of the simulation with the real values.*

	Difference
Total component value	2 %
Total inventory value	-1 %
Module inventory %	-2 %

As table 12 shows, the values are exceptionally close to the real values. This definitely doesn't imply that the simulation is only wrong by a couple of percent every time. For example, the inventoried value of the real system is only a snapshot, so the value of the real average inventory might differ from it. However, with the current demand levels the difference is likely to not be drastic.

5.7.2 Comparison to a deterministic model

Another way to evaluate a simulation model is to compare it against a deterministic mathematical model. This means that the performance of the model is evaluated with minimal variability. This method is useful in validating the model and making sure that it works in the intended way when there is little to no variation included. (Robinson 1994, pp. 144)

To achieve this, the model was run with the same minimum and maximum demand. The test was run using the CTO process as it utilizes all the equations shown in section 5.4. The weekly maximum system demand was set as 4, and therefore the minimum system demand was also set at 4, instead of the usual 0.

Table 13 shows the key variables which the simulation calculated for two example items. The variables match the theoretical values calculated with formulas 1, 2 and 3 exactly. Some values were redacted in table 13 from the final version of the thesis.

Table 13: *Key variables calculated by the simulation*

	Item 1	Item 2
Maximum demand (System)	4	
Maximum demand (Robot)	6	
Average demand (Robot)	4,44	
SS	15	101
ROP	56	387
Lot size	41	286

Figure 22 shows the weekly inventory levels for the example items. The graph is consistent with the variables shown in table 13. It also shows that the simulation works as expected with stable demand. Exactly 4 robots are assembled every week, so the component usage is constant. The supply lead time is much shorter for item 1, so the replenishment interval is a lot shorter. This is also consistent with the principles of a similar example shown by Vollmann et al. (2005, pp. 140)

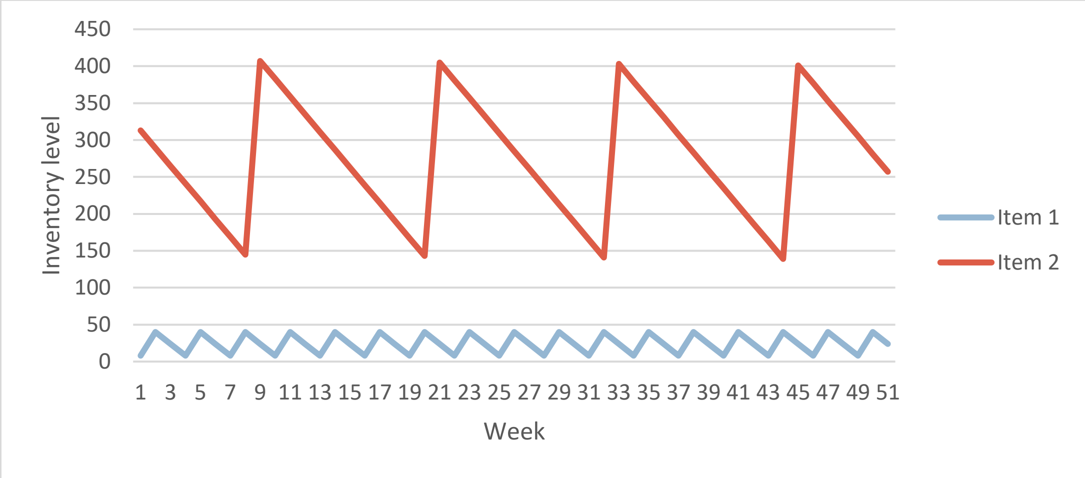


Figure 22: *Consumption & Replenishment of the example items*

While it looks like the lower limit of Item 2 is way higher than the calculated safety stock of 101, the graph is slightly deceiving. The inventory level would meet the safety stock level, but because the lot arrives before the weekly assemblies are done, the inventory level seems to stay higher. This can be seen also from when the lot arrives the inventory is roughly at 150, and when it is added it is just above 400. The discrepancy of almost 30 when compared to the lot size comes from the assemblies done that week and explains why the inventory level seemingly stays above the safety stock.

Figure 23 shows the consumption of the same components when there is variability included. The values are from a simulation run with otherwise the same starting variables but with a demand set created with the minimum demand being 0, as it will be in the actual simulation runs. Therefore, the weekly production numbers fluctuate, and the component usage is not consistent. This creates differences in the reorder interval and shows the importance of safety stock.

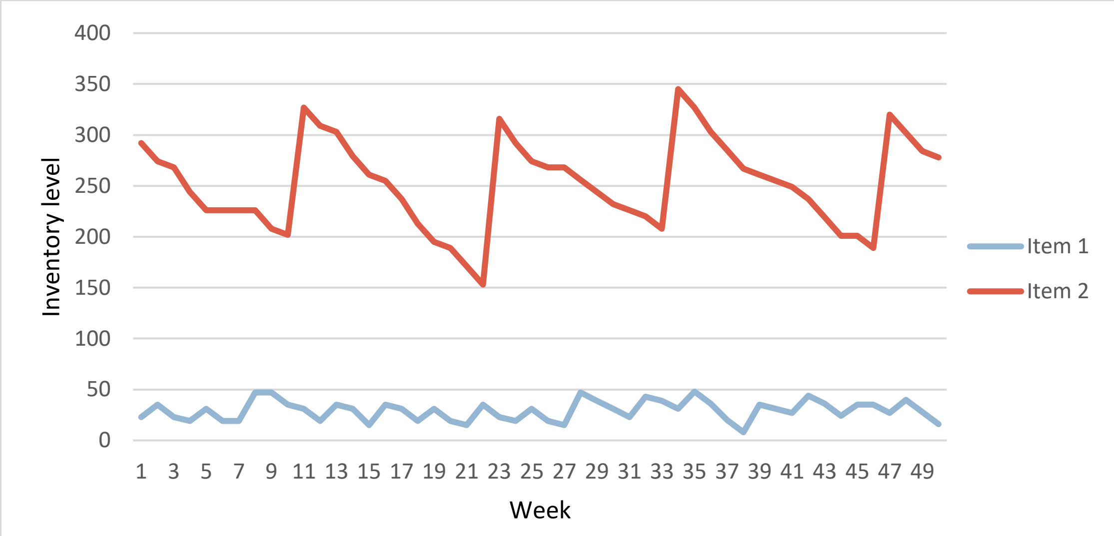


Figure 23: *Consumption & Replenishment with variation*

Figure 23 shows that the model works consistently with variability too. The inventory levels for item 2 stay quite consistently high due to the buffering warranted by the possibility of much higher consumption with a low probability.

5.7.3 Consistency

To evaluate the consistency of the results, a simulation was run was conducted which ran the simulation a thousand times with the same starting variables. This was done two times for both CTO and MTO, first with the same demand pattern each time and the other with differing demand patterns. In the scenario used in these runs the demand was 8, capacity modifier 0,9 and supply lead time modifier 1.

When the simulation is simulated with the same demand pattern as well as the other starting variables, the results are always the same. This proves that the simulation works the same way every time and that it doesn’t break down or change the variables in a way which is not wanted.

When a different demand pattern was created for each of the thousand runs, the results naturally varied. This is expected as the level of uncertainty in the simulation creates different circumstances. The spread of average inventory values is shown in figure 24. The average inventory value was chosen because it indicates the performance of the overall process well.

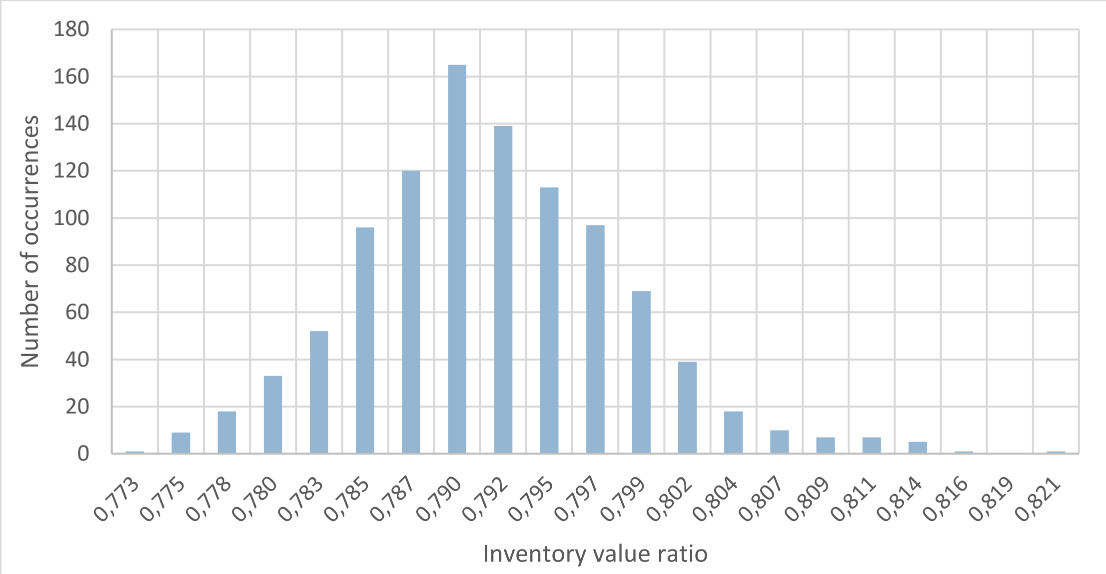


Figure 24: *Distribution of average inventory values in the CTO consistency test*

The distribution in figure 24 indicates that while there is variation, the simulation results are quite consistent. Over half of the average inventory values land into the four catego-ries closest to the total average values. The results of the MTO consistency test, shown in figure 25, indicate a similar pattern to a bit smaller extent.

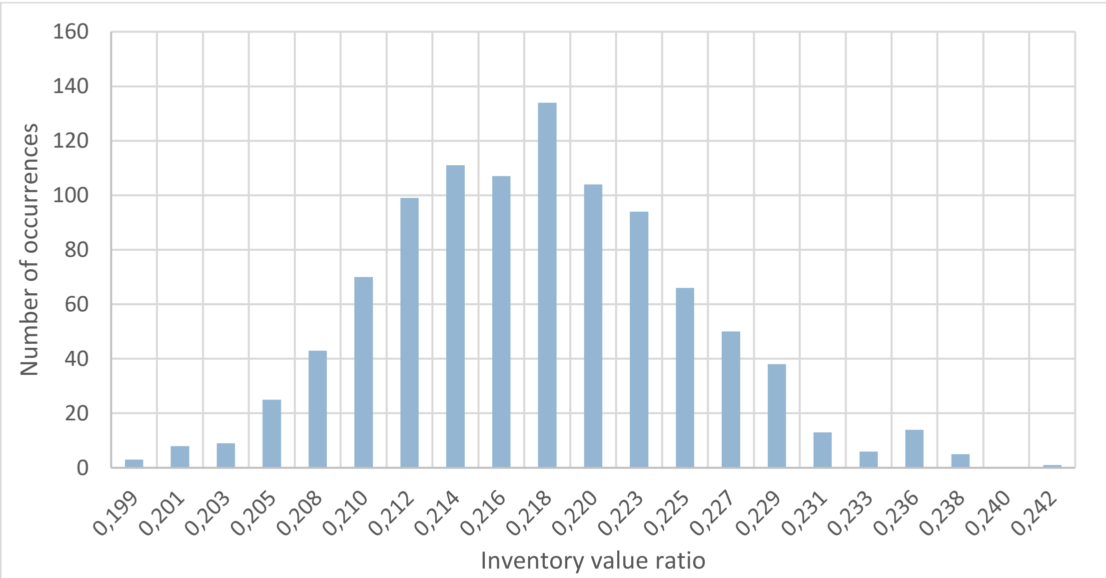


Figure 25: *Distribution of average inventory values in the MTO consistency test*

While the distribution is more spread out, five closest categories to the total average still represent over half of the total sample size. The larger spread is likely explained by the less consistent buffer inventories present in the MTO process, which makes the inventory values fluctuate more compared to the rather stable variation shown for example in figure 24. Overall the results were shown to be quite consistent as even the variation between the highest and lowest average inventory values is quite small when compared to the total level of inventory.

5.7.4 Stability

The stability of the process is important as it ensures that the processes work consistently and there are no extreme situations or breakdowns which might tarnish the results of the simulation.

To evaluate the stability, the simulation was run 3 times for 500 weeks in both CTO and MTO. The example scenario was the same as the one used in section 5.7.2, so the demand was 8, capacity modifier 0,9 and supply lead time modifier 1. A comparison run of the normal 50 weeks was implemented with the same starting variables. Some overall inventory variables from the CTO test runs are shown in table 14. The values shown are the averaged from the 3 runs.

Table 14: *Inventory variables from the CTO stability test*

Weeks	Average inventory	Standard deviation	Min inventory	Max inventory
50	0,79	0,06	0,68	0,92
500	0,80	0,05	0,64	0,95
Difference	-0,03 %	6,56 %	5,62 %	-3,82 %

In table 14 it can straightaway be seen that the average inventory value is reasonably similar. The difference between minimum and maximum inventory values is larger in the 500-week run, which is natural as more weeks will more likely include more extreme circumstances. The standard deviation is lower in the 500-week run however, which means that the simulation stayed stable even when run over a longer period.

Table 15*: Inventory variables from the MTO stability test*

Weeks	Average inventory	Standard deviation	Min inventory	Max inventory
50	0,22	0,04	0,14	0,31
500	0,22	0,04	0,12	0,33
Difference	0,94 %	4,35 %	11,69 %	-8,07 %

The MTO test results in table 15 are comparable to the CTO test results. The average inventory differs a bit more but still the difference is very small. The standard deviation difference is even lower than in CTO, which means that the MTO process also works in a very stable way. The minimum and maximum inventory levels also don't stray too far away from the 50-week run levels.

5.7.5 Module demand approximation

This section describes how the distribution for a module's weekly demand can be derived analytically and how the analytical and numerical approaches compare to each other.

As mentioned in section 5.4.1, the total probability distribution for a module's weekly demand can be calculated as a sum of binomial distributions. The binomial distribution is basically a series of yes/no questions with the same probability. The probability for each number of yesses, also called successes, can be calculated with

$$P(yes) = \binom{n}{k} p^k (1-p)^{n-k}, \tag{4}$$

where n is the number of experiments, k is the number of yes-answers and p is the probability of the answer being yes. To get the distribution, formula 4 is then repeated for each k from 0 to n. (Florescu & Tudor 2013, pp. 92)

For calculating the module demand, the binomial distributions for each system demand level are needed. For this example, they were calculated for a scenario where the maximum demand is 8. The binomial distributions are shown in table 16 where each column depicts a system demand level and each row marks the number of successes. In this case a success is a system with one robot while a failure is a system with two robots.

Table 16: *Binomial distributions for each system demand level*

K \ N	0	1	2	3	4	5	6	7	8
0	100 %	11,3 %	1,28 %	0,14 %	0,016 %	0,002 %	0,0002 %	0,00002 %	0,000003 %
1		88,7 %	20,05 %	3,34 %	0,512 %	0,072 %	0,0098 %	0,00129 %	0,000166 %
2			78,68 %	26,67 %	6,028 %	1,135 %	0,1924 %	0,03044 %	0,004586 %
3				69,78 %	31,54 %	8,911 %	2,0138 %	0,39824 %	0,072003 %
4					61,90 %	34,97 %	11,856 %	3,12606 %	0,706490 %
5						54,90 %	37,226 %	14,7229 %	4,436509 %
6							48,701 %	38,5228 %	17,412319 %
7								43,1981 %	39,051144 %
8									38,316775 %

The data in table 16 shows the probabilities for each individual scenario and therefore the values in each column add up to 100%. To get the actual probability, each value is divided by 9 to account for the probability of each scenario. Table 17 shows the adjusted values.

***Table 17:** Distributions from table 10 adjusted to consider the probability of the system demand level*

K \ N	0	1	2	3	4	5	6	7	8
0	11,11 %	1,26 %	0,14 %	0,016 %	0,0018 %	0,0002 %	0,00002 %	0,000003 %	0,0000003 %
1		9,86 %	2,23 %	0,38 %	0,057 %	0,008 %	0,0011 %	0,00014 %	0,000019 %
2			8,74 %	2,96 %	0,67 %	0,13 %	0,021 %	0,0034 %	0,00051 %
3				7,75 %	3,50 %	0,99 %	0,22 %	0,044 %	0,008 %
4					6,88 %	3,89 %	1,32 %	0,35 %	0,078 %
5						6,10 %	4,14 %	1,64 %	0,49 %
6							5,41 %	4,28 %	1,93 %
7								4,80 %	4,34 %
8									4,26 %

In table 17, all the values add up to 100%, which shows that it details the total distribution for the weekly module demand. Because for the robots a yes-answer represents a system with 1 robot and a no-answer a system with 2, the actual number of robots each cell represents can be calculated with multiplying the number of failures with 2 and adding the successes and failures together. A formula for this is therefore

$$N(Shuttles) = (n - k) * 2 + k. \tag{5}$$

For example, the highlighted cell in table 17 shows that there is a 0,22% chance of a week with 6 systems requiring (6-3) * 2 + 3 = 9 robots (formula 5). Table 18 shows the numbers of robots each cell represents. The cell which was just discussed is highlighted in table 18 too.

***Table 18:** The level of robot demand each cell represents*

K \ N	0	1	2	3	4	5	6	7	8
0	0	2	4	6	8	10	12	14	16
1		1	3	5	7	9	11	13	15
2			2	4	6	8	10	12	14
3				3	5	7	9	11	13
4					4	6	8	10	12
5						5	7	9	11
6							6	8	10
7								7	9
8									8

The total probabilities of each robot demand level can be derived from the data from tables 17 and 18. The table in appendix G shows how the total values are calculated. The same cell is highlighted as in the previous two tables for reference.

The distribution of total values is also shown in figure 26, where the different colors indicate different weekly system demands. As stated earlier if the maximum system demand is 8, the possible demand for robots ranges from 0 to 16, but the higher the demand, the lower the possibility.

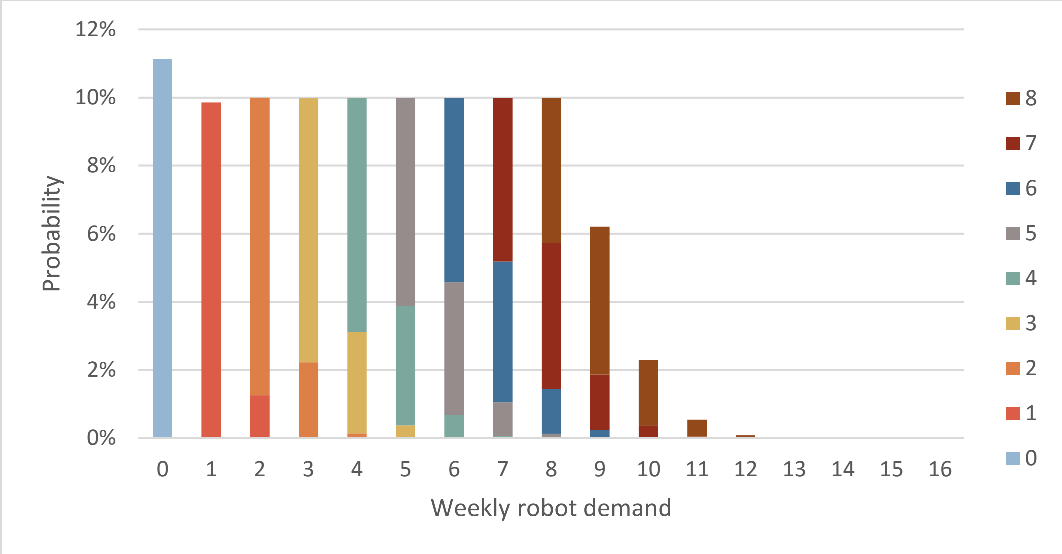


Figure 26: *Theoretical robot demand distribution with the different system demand levels differentiated by colors*

Figure 26 shows quite well why the probability of having 0 robot demand is higher than the rest. If the weekly demand is 1 (bright red color), some of the systems have 2 robots, and therefore the weeks which have only 1 system with 1 robot is lower than 0. Table 19 details the comparison between the numerical and analytical methods. The values for the numerical and analytical methods are from tables 10 and appendix G respectively.

Table 19: *Comparison between the numerical and theoretical distributions*

Demand	Numerical	Analytical	Difference
0	11,11 %	11,11 %	0 %
1	9,9 %	9,85 %	0,02 %
2	9,9 %	9,99 %	0,02 %
3	10,1 %	9,98 %	0,02 %
4	9,9 %	9,98 %	0,008 %
5	10,0 %	9,98 %	0,04 %
6	9,9 %	9,98 %	0,02 %
7	10,0 %	9,98 %	0,01 %
8	10,0 %	9,98 %	0,03 %
9	6,2 %	6,21 %	0,01 %
10	2,3 %	2,30 %	0,006 %
11	0,6 %	0,54 %	0,002 %
12	0,1 %	0,081%	0,0007 %
13	0,01 %	0,008 %	0,0002 %
14	0 %	0,0005 %	0,0005 %
15	0 %	0,00002 %	0,00002 %
16	0 %	0,0000003 %	0,0000003 %

The difference between the methods is miniscule, fractions of a percentage at most. Therefore, it is justifiable to utilize the numerical method, especially when considering the more general applicability and ease of use.

5.8 Implementation

This chapter explains the implementation of the simulations. The first section showcases the scenarios used and the number of runs conducted. The simulation sets consist of multiple runs of each scenario to increase the accuracy of the results. Especially in operations planning a number of runs is needed to discern the range of possible outcomes for each scenario as there is a lot of possibilities for the output variables (Robinson 1994, pp. 169).

A single simulation run consists of a warmup period of 16 weeks and then the actual simulation of 50 weeks. The warmup period eliminates the effects of the starting stock levels and therefore gives a better understanding of the actual performance of the system (Robinson 1994, pp. 159; Miclo et al. 2019) The 50 week run was decided on because it represents roughly the amount of time the factory is operational in a year.

5.8.1 Main simulation

The main simulation included multiple different scenarios to gain insight into the impact of the different input variables. There is one main and two secondary input variables which make up the scenarios. The main input variable is the level of demand. The two secondary variables are modifiers to capacity and supply lead time.

Table 20 shows the different demand levels the scenarios will include. These levels mimic possible current and future demand levels for the Agilon factory. The demand level is determined as the max weekly level. The minimum weekly demand is 0.

Table 20: *Demand scenarios*

Max/week	Average/week	Average/year
2	1	50
3	1,5	75
4	2	100
6	3	150
8	4	200
10	5	250
16	8	400

The demand level is the main variable because it determines the level of uncertainty in the simulation. The higher the demand is, the more uncertainty there is.

As mentioned above, the secondary input variables are modifiers to capacity and supply lead time. This means that e.g. the supply lead times of all the components are multiplied with the modifier.

Capacity modifier is used to adjust the number of modules which can be produced in a week. It affects the production of robots and access points, as the capacity to pick the frames was determined to be infinite for the purposes of the simulation. Table 21 shows the limitations of secondary input variables.

Table 21: *The limitations of secondary input variables*

	Supply lead time	**Capacity**
Lower limit	0,75	0,6
Upper limit	1,25	1

The secondary input variables will act as a type of sensitivity analysis to discern the effects of the corresponding factors. To achieve this, the simulation creates 100 pairs of capacity and lead time modifiers which are then assigned to each of the 7 demand levels. That comes up to a total of 700 simulation runs for both strategies. The pairs are randomly generated between the limits. However, to get consistent ranges for the simulations one of the pairs always consists of the lower limits and another of the upper limits of the secondary variables.

Lastly, while the demand patterns and input modifiers differ from each other, it should be noted that both processes were run with the same parameters. This means that both processes were run exactly the same number of times with exactly the same input variables and demand patterns. This is an important consideration as to reinforce the credibility and usability of the results which the simulation produces.

6. RESULTS

This chapter goes through the results of the simulation runs. First, the results are analyzed from the tied-up capital point of view and afterwards in terms of delivery reliability. Lastly, a summary of the results is gathered to give a concise overview of the findings.

The main variables used to analyze the results are the ones determined in the secondary research question: capacity, supply lead time and demand uncertainty. The capacity and supply lead time are measured with the modifiers each scenario included. The demand uncertainty is measured by the percentual deviation from the average demand. This measure does not encompass the demand uncertainty fully, but it is the best approximation available due to more detailed demand pattern analyzing being impossible with the data available. The inventory data in this chapter is substituted by ratios similarly the previous chapter.

6.1 Tied-up capital

This section gives an overview of the inventory investment results in the main simulation. As stated in section 5.8.1, each demand level was run 100 times with different input variables and demand patterns. First, the section strives to give a rough understanding of the levels of tied-up capital needed for the processes. The next section tries to explain the influence of the different input variables more closely by conducting regression analyses of the results in the different demand scenarios. Lastly, a summary of the inventory investment results is given.

Table 22 shows the results for the CTO process. The minimum and maximum inventory values are the lowest and highest average inventory values for each demand level. The average value and the module inventory percentage was calculated by calculating the averages for all 100 runs of each level.

Table 22: *Tied-up capital results for the CTO process*

Scenario	Min	Avg	Max	%-of module inventory
2	0,23	0,29	0,35	22,30 %
3	0,31	0,39	0,48	23,83 %
4	0,36	0,46	0,57	26,36 %
6	0,51	0,63	0,75	28,79 %
8	0,64	0,82	0,99	28,58 %
10	0,82	1,00	1,20	28,97 %
16	1,24	1,56	1,87	29,46 %

The average inventory value of the CTO process constantly grows by roughly 0,08 units when the demand increases by 1. Interestingly the difference between minimum and maximum values gets bigger when the demand increases. This is due to the impact of the input variables get bigger when the demand increases. For example, the supply lead times change equally with the different modifiers, but the number of components needed for a time period is higher. Also, the difference between module inventory levels, which is dictated by the capacity modifier, is bigger in the higher-demand scenarios.

The module inventory percentage generally gets higher when demand rises. However, the increase is relatively modest as it stays between 20% and 30%. This is most likely since the component buffers get comparatively smaller as meeting the theoretical maximum demand gets less and less likely when the demand increases. At the same time the module buffers increase quite linearly. due to more work in progress (WIP) being created as some modules are waiting for components when some have already been assembled.

Table 23 shows that in the MTO process, the increase of average inventory per demand is a bit more sporadic, but on average it is roughly 0,016 units. This is roughly one fifth of the increase in the CTO process. The difference is naturally due to less components being buffered, and no module inventory being kept. Figure 27 illustrates this discrepancy well.

The difference between going from 2 to 3 and 3 to 4 is explained by how the maximum demands are calculated. In the 2-demand scenarios the maximum demand for frame-modules and retail AP component buffering is 1, but in 3 and 4 it is 2. Therefore, the buffer is set at similar levels in 3 and 4, but 3 has less demand so the average inventory stays higher. 4-demand scenarios still have higher average inventories due to more buffering for robots and industrial APs. This shows the importance of accurate forecasting.

Table 23: *Tied-up capital results for the MTO process*

Scenario	Min	Avg	Max	%-of module inventory
2	0,08	0,12	0,15	9,46 %
3	0,09	0,14	0,19	10,58 %
4	0,10	0,15	0,19	13,30 %
6	0,13	0,17	0,22	16,53 %
8	0,16	0,21	0,28	16,86 %
10	0,18	0,25	0,32	17,45 %
16	0,25	0,37	0,46	17,53 %

Like in the CTO process, the module inventory gets higher as the demand increases in the MTO process. This is probably due to more WIP accumulating as some modules are being assembled for a system while other modules are still waiting for components to arrive. The percentage and especially the actual value are lower than in the CTO process as module inventory is not held regularly.

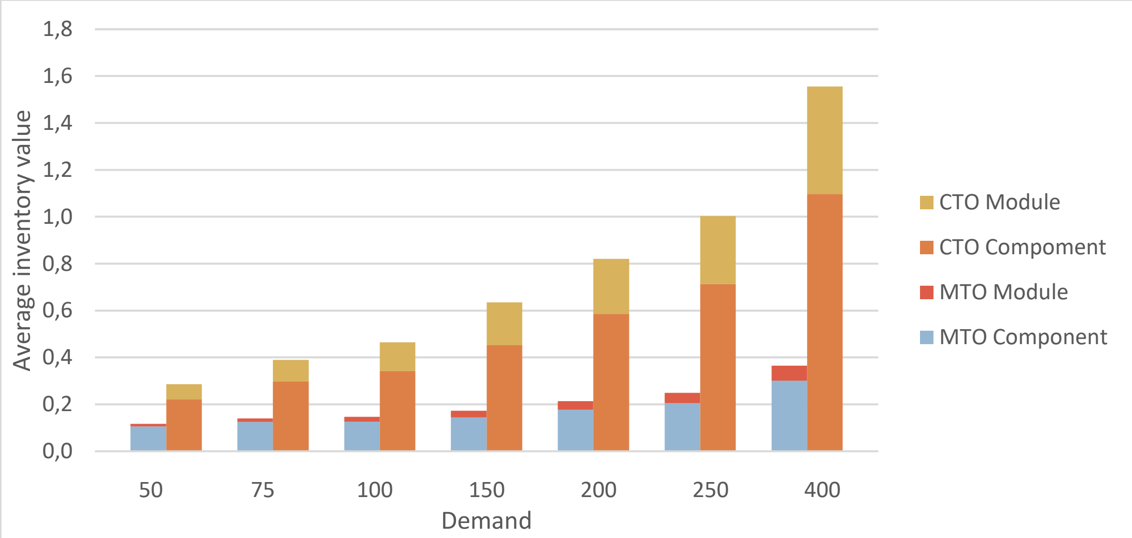


Figure 27: *Average inventory investment comparison*

As seen in from tables 22 and 23, starting from demand level 6, the module inventory value of the CTO process is higher than the total value of the MTO process inventory. Also, when maximum weekly demand is 2, the CTO tied-up capital is a bit over two times the MTO equivalent while at the 16 level it's almost four times more. the difference is due to the bigger increments in the CTO process.

Table 24 shows the average inventory investment needed for one average weekly system demand. For example, when the max demand is 2, the average weekly demand is 1 so the average inventory is wholly spent to on average produce 1 system, so the value is the same as the average inventory. Similarly, in the case of maximum demand of 16, the average demand is 8, so the average inventory was divided by 8 to get the inventory investment needed for one system.

Table 24: *Average inventory investment needed for one system*

Scenario	CTO	MTO
2	0,29	0,08
3	0,26	0,06
4	0,23	0,05
6	0,21	0,04
8	0,21	0,04
10	0,20	0,04
16	0,19	0,03

The investment needed for one system decreases in both processes when the demand increases. This is likely again due to relatively less buffering needed in the higher-demand scenarios as the probability of the maximum demand being realized is significantly lower. The investment drops roughly 30% in the CTO process, while in the MTO process it drops about 60%. This is likely due to the buffers increasing less per demand level, so the impact of the division is higher.

6.1.1 Regression analysis method

Linear regression analysis was used to discern the actual effects of the input variables on the tied-up capital. The analysis was conducted for each different demand level separately. This created an individual regression formula for the processes in each demand scenario. The equation strives to describe the relative impact of the different input variables. (Montgomery et al. 2012, pp. 1)

This section briefly goes over the procedure used and gives one example of the analysis process. Lastly, the section outlines a synthesis of the results and lists all the resulting formulas. IBM SPSS Statistics-software was used to conduct the analysis.

The independent variables, also called regressors, considered in the regression analysis were the capacity, supply lead time and the percentual deviation from the average demand in the scenario. The dependent variable or response is in this case the average inventory value of the scenario. The goal was not to create the most predictive model but rather to create an equation that describes how the input variables influence the simulation to create the observed values.

To get a good impression of the effects all three variables have, the regression model creation loosely followed the backward elimination method. In the backwards elimination all the regressors are initially entered to the model and then their relative importance is evaluated. If some variables don't meet the criteria of an important regressors, they are discarded, and a new model was created with the remaining variables. (Elliott & Woodward 2007, pp. 101; Montgomery et al. 2012, pp. 346-347) The criteria used is that the statistical significance of a regressor needs to be over 0,1 in the model. This is default criteria in SPSS and the same which Montgomery et al. (2012, pp. 347) use as an example. However, this makes it possible that if all variables are deemed statistically significant, they are all accepted into the model.

After an eligible model has been found, the model was then critically evaluated. This was done as Elliot and Woodward (2007, pp. 102) instruct by assessing the R-squared values and using residual analysis. However, the model selection process was not carried out in an overly strict manner, as the aim was not finding a model which perfectly depicts the impact of the variables but rather get a grasp of how the input variables might act in comparison to each other. All in all, the goal was to find a simple regression equation which describes the effects of the input variables on the inventory investment. Table 25 shows an example model created by SPSS with the specifications which were discussed earlier. The standard error of the estimate was modified like the other inventory variables.

Table 25: *Example regression model summary*

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.989[a]	0,979	0,978	0,012

a. Predictors: (Constant), cap, deviation, dt
b. Dependent Variable: average total value

The relationship between the regressors and the response should also be confirmed with a statistical test (Elliott & Woodward 2007, pp. 100). SPSS automatically does this by conducting an analysis of variance (Anova) test to study the significance of regression. The test came out statistically significant for the candidate model. This means that there is at least some linear relationship between the variables and therefore the model can likely be used to predict the value. (Montgomery et al. 2012, pp. 84-85) Table 26 shows the important characteristics of the model in more detail. The unstandardized coefficients and the confidence intervals have been modified to the same ratio scale as the other inventory values in this thesis.

Table 26: *Example regression model coefficients*

Variable	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95,0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
(Constant)	0,23193	0,01163		19,94	6,53E-36	0,2088	0,2550
Deviation	-0,00087	0,00013	-0,098	-6,57	2,61E-09	-0,0011	-0,0006
Supply lead time	0,00508	0,00008	0,962	64,25	1,07E-80	0,0049	0,0052
Capacity	0,00096	0,00011	0,132	8,88	3,88E-14	0,0007	0,0012

a. Dependent Variable: Average total value (8)

As shown in table 26, all the variables have a very low significance (Sig.) value, which conversely means they all are significant, and therefore they all are eligible for the model. They all also have a high t-score which also indicates significance. This significance is likely due to these variables being the main differences between the creation of the data-points and therefore they naturally have an impact on the output of the model.

Table 26 also shows the error and the confidence intervals and the standard error estimates. The coefficients in table 26 show the influence of the variables. The constant is the X-intercept of the regression formula. The standardized beta shows the relative impact of the variables and therefore it can be used to compare the models against each other. (Montgomery et al. 2012, pp. 67-68)

The unstandardized B, as in beta, shows the actual multiplier for the regression formula for the constant and the variables. (Montgomery et al. 2012, pp. 67-68) They can be interpreted as when the delivery time modifier increases by 1%, the average inventory kept increases on average by a bit over 0,005 units. The formula for this scenario is therefore as follows:

$$\text{Average inventory} = -0{,}0009 * \text{Deviation} - \% + 0{,}005 * \text{SLT Modifier} + 0{,}001 * \text{Capacity modifier} + 0{,}23$$

After the model was created, residual analysis was conducted to identify faults or deviations in the model. The summary variables such as the r-squared don't always tell the full truth of the validity of the model. Residuals are the deviations of the observed points to the regression line. The conditions to show a valid model which the residuals can prove are as follows:

- The error term has zero mean.
- The errors have a constant variance.
- The errors can't be correlated
- The errors need to be normally distributed. (Montgomery et al. 2012, pp. 129)

The mean of 0 and the normality of the residuals is confirmed by the distribution in appendix H. While not perfect, the shape does resemble a normal distribution. Montgomery et al. (2012, pp. 136) also state that small deviations from the normality clause don't affect the validity of the model in a big way.

The constant variance and correlation can be investigated by plotting them in a scatter plot with the predicted Y values. The scatter plot should not show any recognizable patterns, such as curves and variance should be roughly even throughout the spread. The figure in appendix H shows a reasonably even spread, which therefore reinforces the assumptions second and third assumption on the residuals. (Montgomery et al. 2012, pp. 131)

As none the needed tests and assumptions were failed, it can be concluded that the model is valid and can be used to estimate the impact of supply lead time on the inventory level and to conclude that the deviation and capacity don't contribute significantly in the result. Figure 28 show a graph where the model is fitted with the observed values. The predicted values are not linear due to the graph being squashed into just 2 dimensions.

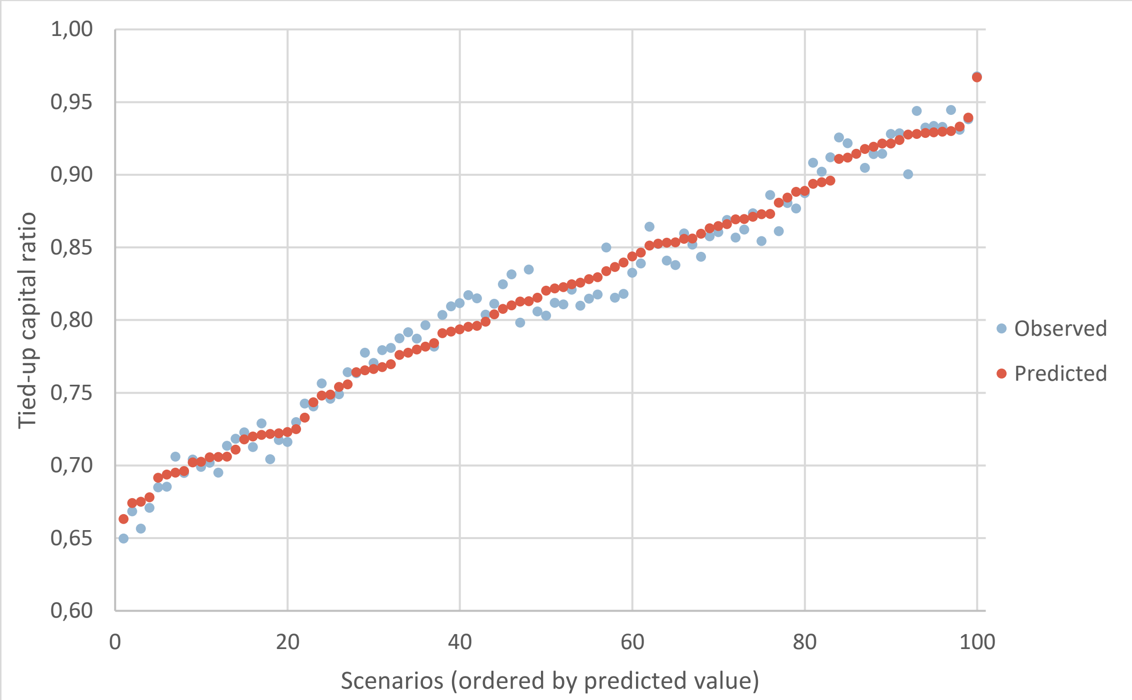


Figure 28: *Example regression model compared to the observed values*

While the fit is very good, some values do differ from the model a bit. However, if the difference would have been significant enough for the data points to be deemed to be outliers, they would have been detected in the residual analysis (Montgomery et al. 2012, pp. 130-131). Also, even the biggest deviations are slightly less than 0,04 units which conforms to the confidence intervals and is quite irrelevant when considering the overall accuracy of this study.

6.1.2 Regression analysis results

This section goes through the regression models. The regression models are displayed in two different tables for both processes, one table with some overall characteristics and the standardized coefficients and one table with the unstandardized coefficients. These tables are 27 and 29 for the CTO and 28 and 30 for the MTO process respectively. The results of the regression turned out to be so that all models included all the variables. As discussed earlier, this is likely due to them being essentially the only variables which can be tracked, and which individualize the scenarios.

The overall characteristics shown are the R squared values which represent how well the model captures the variation in the dataset and standard errors which indicate the size of the error, which shows the size of an average error in the model.

It should be noted that there were in total 14 datapoints which were considered as outliers by SPSS. These were discarded as failures of the measurement tool as they might skew the model results (Montgomery et al. 2012, pp.152). This was not deemed to be a problem as the total dataset includes 1400 results, so the discard rate was only just 1%.

The standardized coefficients are useful in comparing the models to each other because they have been standardized to extract the magnitude of the coefficient and to just display the relative effects of the coefficients. However, one should be cautious when analyzing these kinds of partial regression coefficients as they display the effects of the variables only when the other variables are constant. However, they can give useful insights to the model. (Montgomery et al. 2012, pp. 68, 114-115)

Table 27: *Model characteristics and standardized coefficients for the CTO regression models*

CTO	R Square	Std. Error	Cap	DT	Dev
2	0,990	0,003	0,076	0,987	-0,065
3	0,987	0,005	0,086	0,987	-0,072
4	0,983	0,006	0,098	0,981	-0,080
6	0,985	0,008	0,114	0,990	-0,107
8	0,979	0,012	0,132	0,962	-0,098
10	0,980	0,014	0,149	0,952	-0,111
16	0,979	0,022	0,159	0,973	-0,111

The R-squared values in table 27 and 28 indicate that in both processes the models explain the variance in the datasets very well. In the MTO process the R-squared values decrease as the demand increases which might be an indication that the importance of factors which are not explained well by the input variables, such as demand patterns, increases as demand increases. The standard errors increase as the demand increases which is natural as the variance of the data also increases as seen in tables 22 and 23.

The standardized coefficients show that the delivery lead time modifier is by far the most influential of the three variables. However, it can also be seen that the influence of the other two also increases as the demand increases. The capacity modifiers' importance doubles when going from the lowest to the highest demand.

The MTO process shows similar trends which can be seen in table 28. The supply lead time is the most important but not as dominant especially in the high-demand scenarios. The relative importance of the capacity modifier and the deviation from average demand show a big increase as the demand rises.

***Table 28:** Model characteristics and standardized coefficients for the MTO regression models*

MTO	R Square	Std. Error	Cap	DT	Dev
2	0,945	0,005	-0,099	0,949	-0,131
3	0,955	0,006	-0,093	0,956	-0,130
4	0,947	0,006	-0,163	0,939	-0,137
6	0,906	0,008	-0,215	0,906	-0,292
8	0,920	0,009	-0,184	0,875	-0,231
10	0,875	0,014	-0,214	0,815	-0,273
16	0,853	0,019	-0,247	0,825	-0,331

Interestingly the capacity modifier has an inverse effect when compared to the CTO process. This is likely due to the capacity modifier enabling more efficient deliveries but increasing the module inventory in the CTO process. The efficient assembly capacity means that the deliveries can be fulfilled quickly as the components arrive and therefore less inventories are being kept both at the component and the module level.

The unstandardized coefficients indicate the strength of the variables and considers the magnitude of their effects. The data in table 29 shows that for example 1% increase in supply lead time increases the average inventory by 0,002 units in the 2 demand scenarios and almost 0,0096 units in the 16 demand scenarios. These are of course average values and the error margins of the models create uncertainty in the actual results. However, they can give good insights. Interestingly as seen by the relative increase in table 27, the magnitude of the capacity and deviation increases to roughly 10 times the value from demands 2 to 16, while the supply lead times effect increases only 5 times in the same range.

***Table 29:** Unstandardized coefficients for the CTO regression models*

Demand	Constant	Cap	DT	Dev
2	0,0657	0,0002	0,0020	-0,0002
3	0,0943	0,0003	0,0027	-0,0003
4	0,1227	0,0004	0,0031	-0,0004
6	0,1784	0,0006	0,0040	-0,0007
8	0,2319	0,0010	0,0051	-0,0009
10	0,2672	0,0014	0,0063	-0,0012
16	0,4228	0,0022	0,0095	-0,0018

The results for the MTO process show similar patterns. The values are just quite a lot lower, even to the point of all but the supply lead time being less than zero. This is possible as the supply lead time value is so much higher, and it gets values from 75% to 125%. The absolute values are consistently lower than in the CTO process except for deviation, which seems to indicate that the deviation from average demand affects both processes in a similar way, but the effect is stronger for the MTO process because the overall inventory investment is lower.

Table 30: *Unstandardized coefficients for the MTO regression models*

Demand	Constant	Cap	DT	Dev
2	-0,0017	-0,0002	0,0013	-0,0003
3	-0,0153	-0,0002	0,0017	-0,0003
4	0,0162	-0,0004	0,0016	-0,0004
6	0,0631	-0,0005	0,0015	-0,0008
8	0,0679	-0,0005	0,0018	-0,0008
10	0,1006	-0,0007	0,0021	-0,0011
16	0,1753	-0,0011	0,0027	-0,0019

In the higher demand scenarios, the increasing impact of the capacity and deviation can be seen well. The values grow almost to the same levels as the supply lead time. This is a bit misleading however, as in the unstandardized coefficients the range of the input values should be considered when the coefficients are analyzed. In this case the average input value is highest for the supply lead time, so the coefficient is multiplied on average more than the other values.

6.1.3 Inventory investment result summary

Overall the results show that in the CTO process, the inventory investment is a lot higher than in the MTO process. This is quite an obvious conclusion as the CTO process simply keeps more buffer inventories. The magnitude of the difference is the interesting factor in this matter. The difference starts by CTO inventories being just over 2 times more in the 2 demand scenarios to almost 4 times more in the high-demand scenarios.

The supply lead time is the dominant factor by which the tied-up capital is dependent on. This is quite a natural conclusion as the supply lead time modifier is the main component when calculating the buffer inventories for each component in the simulation. The absolute effect shown by the unstandardized coefficients is much more powerful in the CTO process as there are more components which have buffers dictated by the supply lead time.

Also, the effects of the deviation are consistent between the processes. When the demand is higher than the average demand by which the buffers are determined, the average inventories naturally lower. As mentioned in the previous section, this effect is rather similar for both processes, as the absolute effect quite consistently stays at the same levels. Yet, the relative effect is higher for the MTO process due to the overall lower inventories.

As mentioned earlier, the capacity has an inverse effect between the processes. It is likely that in both processes the higher capacity enhances the delivery process performance which makes for better circulation of parts and less modules waiting for deliveries. However, it seems that in the CTO process this effect is outdone by the fact that the module buffers increase as the capacity increases as the maximum module inventories are tied to the weekly module assembly capacity.

6.2 Delivery reliability

The delivery reliability was measured by calculating the percentage of late deliveries. After an initial analysis of the main simulation results, it was noted that a supplementary simulation would complement the main simulation in regard to the delivery reliability. Therefore, a supplementary simulation was conducted to gain a broader understanding of the impact of the capacity modifier on the CTO process.

This section first goes through the main simulation results. After this, the main simulation results are combined with the supplementary simulation and the combined results are analyzed. Lastly a summary of the delivery reliability results concisely goes through the insights gathered during the analysis.

6.2.1 Main simulation

The results of the main simulation in table 31 show that on each demand level the late percentages for both processes are very low. On average 1,5% of deliveries in the MTO process are late while virtually no deliveries are late in the CTO side with the input variables in the main simulation. This seems to indicate that in similar conditions the MTO process as it is set up in this simulation tends to perform worse than the CTO process when it comes to delivery reliability.

Table 31: *Delivery reliability results from the main simulation grouped by demand level*

Demand	Late % CTO	Late % MTO
2	0,08 %	1,74 %
3	0,00 %	2,28 %
4	0,02 %	0,71 %
6	0,19 %	0,99 %
8	0,01 %	1,73 %
10	0,00 %	1,33 %
16	0,00 %	1,10 %

However, it should be noted that as is the case in the inventory investment, the different input variables have differing impacts on the delivery reliability. This is shown very well for the MTO process in appendix E. While the results are rather inconclusive for the CTO process, it does show that under these conditions, the CTO process manages to fulfill almost every order perfectly.

In the MTO process, the deviation from average demand seems to increase the amount of late deliveries, which is quite a natural progression, as when there is less deliveries, it is easier to fulfill them. The supply lead time modifier and shows an opposite interaction, where the delivery reliability gets better when the modifier increases. This is likely due to there simply being more buffering for the different components, which helps in responding to the differences in demand.

However, the neither of those input variables show as strong an interaction as the capacity. Most of the late deliveries come from the scenarios which have the lowest capacity modifiers. This reinforces the natural notion of capacity being of the utmost importance in deciding the delivery reliability of an order fulfillment process.

Table 32 shows the distribution of delivery times for the late deliveries. It should be noted that the absolute values for the late deliveries in the processes is different as there is a lot more late deliveries in the MTO process as shown by table 31. Therefore, the values in table 32 for the CTO and MTO process represent the roughly 0,1% and 1,8% of late deliveries respectively.

Table 32: *Delivery times for the late deliveries in the main simulation*

Delivery time (weeks)	CTO	MTO
5	97,6 %	82,7 %
6	2,4 %	15,5 %
7	0,0 %	1,8 %

Table 32 shows that while there was a couple of deliveries which were 2 weeks late in the CTO process, almost all of them were just one week late. In the MTO process there is a lot more deliveries which were 2 weeks late and even a couple which were late by 3 weeks.

6.2.2 Supplementary simulation

The results of the main simulation were a bit deficient for the CTO process due to the lack of late deliveries. Therefore, a supplementary simulation run was needed. Due to the highest impact of the capacity input variable, the range of the capacity modifier was increased for the second simulation. The simulation was specified otherwise similarly to the main one, but with the capacity modifier going from 0.3 to 0.6.

The rather drastic decrease in the capacity modifier predictably brought with it a drastic increase in late deliveries. This doesn't mean that the processes don't perform well but rather reveal their limitations. Table 33 shows the results when both simulations are considered. As expected, the lower capacity modifiers bring about a lot more late deliveries, and actually seem to form a trend where the delivery reliability gets worse when the demand increases. This is likely explained by the increase in demand uncertainty and the decrease in the absolute capacity values.

Table 33: *Delivery reliability results grouped by demand level for the extended simulation*

Demand	Late % CTO	Late % MTO
2	10,67 %	20,51 %
3	10,03 %	19,86 %
4	13,13 %	23,58 %
6	14,71 %	25,07 %
8	14,77 %	25,17 %
10	17,05 %	26,91 %
16	17,26 %	27,17 %

In the CTO Process the percentage of late deliveries increases from around 10% to a bit over 17%, while the results for the MTO process are roughly 10% higher for each demand scenario. It should be noted that the difference between the values in tables 31 and 33 come solely from the supplementary simulation, so it can already be seen that the capacity reduction has a very big impact on the delivery reliability.

Figure 29 showcases the impact of capacity well. Up to the 0.6 capacity modifier the reliability is almost perfect for both processes, but when it drops below 0.6 the MTO process's percentage starts to grow almost exponentially. A similar development can be seen for the CTO process but with the spike starting from the 0.5 margin. It should be noted that the 0.5 modifier marks the capacity which corresponds to the average demand in the process.

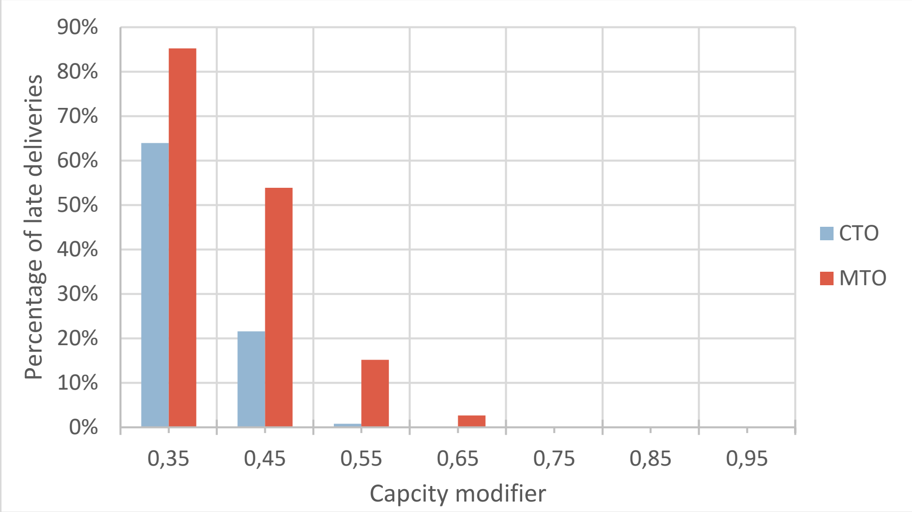


Figure 29: *Extended delivery reliability results grouped by capacity*

The deviation shows similar behavior as in the main simulation. Figure 30 shows that the increase can be seen in both processes. It should be noted that the scale is not even as the deviation scale is normally distributed, the data has more representation on values close to 0. The number of data points was evened out by increasing the range of the more far out values. Even then, the amount of representation is larger in the middle of the figure, which might skew the +10% category to be the highest. However, overall the delivery reliability decreases when the system demand increases.

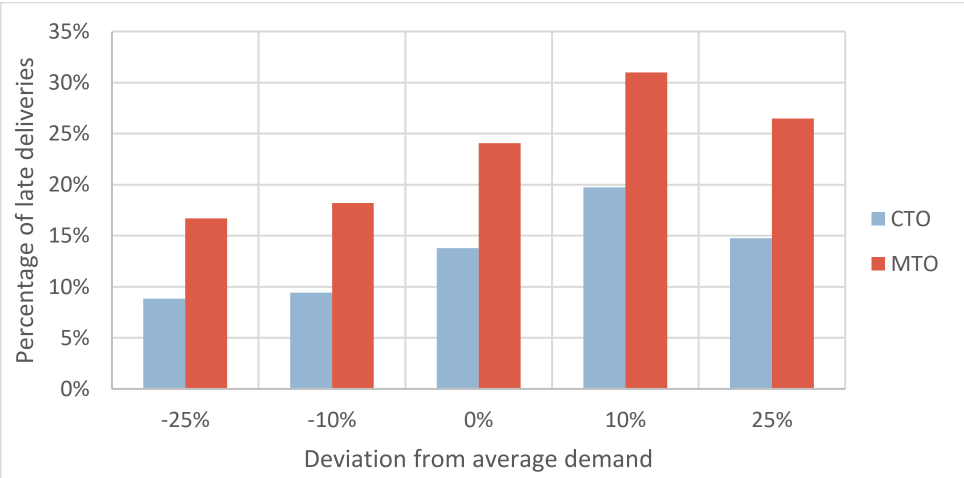


Figure 30: *Impact of demand deviation on delivery reliability*

While both processes responded quite similarly to the two input variables discussed above, the supply lead time doesn't give quite as obvious results, as seen in figure 31. In the CTO process the general trend seems to be that there is an increase in late deliveries when the supply lead time increases. The lowest point is at the 1.1 mark, however. MTO doesn't quite follow the same logic as there is a big increase in the 0.9 category, and there are no obvious trends. The lowest point is at the 1.1 modifier as in the CTO process.

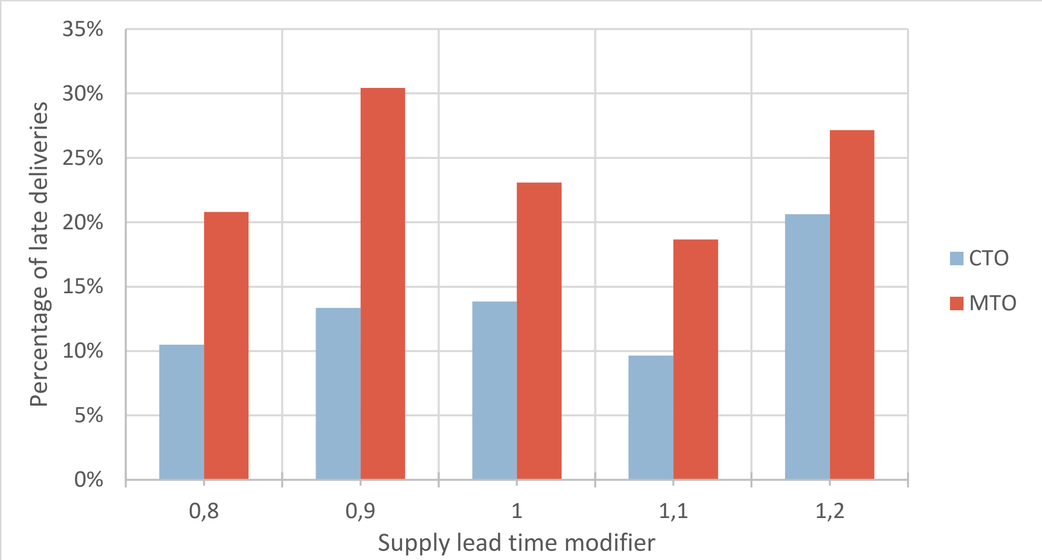


Figure 31: *Impact of the supply lead time modifier on delivery reliability*

The sporadic impact of supply delivery time modifier is likely due to the fundamental role it plays when determining the buffer inventories. In the MTO process the decision on keeping buffer inventories is dictated by whether the supply delivery lead time of a component is longer than the customer delivery time. Therefore, if there are big groups of components with similar supply characteristics, they behave in the same way when the supply lead time goes over certain boundaries. This makes it so that at certain intervals even small changes to the supply lead time can change the characteristics of a simulation run in a big way.

Table 34 shows the delivery time statistics for the extended simulation. Again, the values represent the roughly 13% and 24% of deliveries for CTO and MTO respectively, as the rest of the deliveries were delivered in the set 4 weeks. Right away it can be seen that the extension provides a lot more variability in the delivery reliability, as there are deliveries which were over 10 weeks late. This might not be quite a realistic statistic in the real world as maybe the contract would be canceled if it ran that late. However, for the purposes of the simulation it was deemed interesting to see how long the delivery times would stretch.

Table 34: *Delivery times for the late deliveries in the extended simulation*

Delivery time (weeks)	CTO	MTO
5	22,1 %	13,4 %
6	14,2 %	11,7 %
7	10,6 %	10,2 %
8	9,1 %	8,9 %
9	7,2 %	8,1 %
10	6,3 %	6,8 %
11	5,0 %	6,0 %
12	4,1 %	5,4 %
13	3,6 %	4,4 %
More	17,8 %	25,1 %

It can be seen in table 34 that while some deliveries are extremely late, the majority is still just a couple of weeks late. The percentage decreases gradually the longer the delivery time. In the CTO process the percentage of deliveries which were just 1 or 2 weeks late is higher than in the MTO process, but after that the percentages are quite similar until the more-category. Also, it should be noted that almost all the late deliveries were created from the supplementary simulation.

6.2.3 Capacity utilization

As the delivery reliability was most clearly influenced by the capacity modifier, analyzing the differences in capacity utilization might give some insights on the differences between the processes.

Two sets of scenarios were picked to represent the capacity utilization data, one with ample capacity and one with more constrained capacity. Table 35 shows key values from the example scenarios. The sets consisted of roughly 30 scenarios chosen and grouped by the capacity.

Table 35: *Key values from the capacity utilization examples*

	Demand	Avg capacity modifier	Maximum capacity	Capacity utilization		Late percentage	
				CTO	MTO	CTO	MTO
High	8	0,75	7	55,88 %	55,88 %	0,01 %	0,34 %
Low	8	0,50	5	77,72 %	77,38 %	1,95 %	22,16 %

The overall capacity utilization is very close between the CTO and MTO processes in both examples due to the weekly capacity and the demand patterns being identical between the processes. However, the delivery reliability shows very drastic differences especially in the low capacity example. Clues to why this occurs can be found by investigating the weekly differences in the capacity utilization. Figure 32 shows the distribution of robot production numbers in weeks in the high capacity example. For example, there was on average roughly 8 weeks on which neither of the processes produced any robots and roughly 4 weeks when they both produced 1 robot.

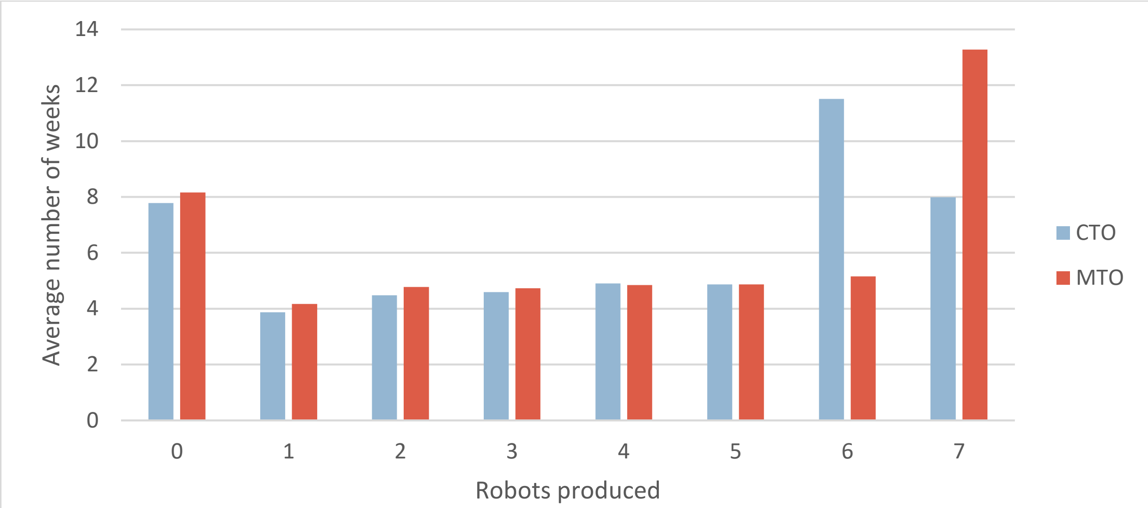


Figure 32: *Capacity utilization in the high capacity scenarios example*

While the graph shows very similar behavior between the processes in the lower end of the production numbers, the interesting part is the two highest production numbers. The MTO process had 13 weeks of max capacity production while the CTO process only had 8. Also, the second highest number is higher for the CTO process, so they had more weeks where they had floated excess capacity. Figure 33 shows the same statistics for the lower capacity example.

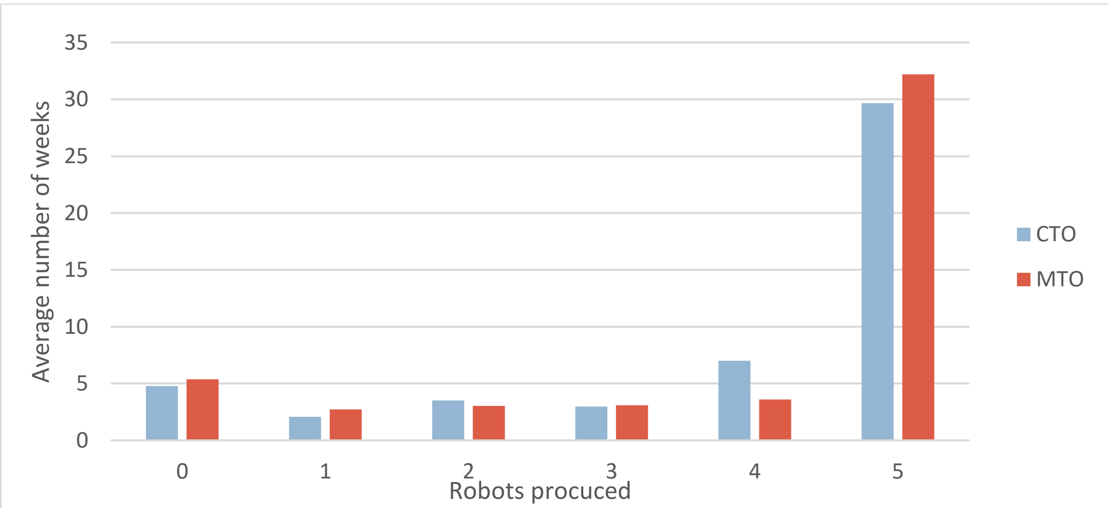


Figure 33: *Capacity utilization in the low capacity scenarios example*

The differences between the processes are much smaller in the small capacity example. The MTO process has roughly 5 more maximum capacity days. Due to the lo modifier both processes had to work over half of the year on full capacity.

However, the delivery reliability statistics shown in table 35 shows a big difference between the processes. There are two possible reasons for this. The first is that the inventory handling in the MTO process performs worse. This is not likely as there is no evidence for such behavior in the high-capacity scenarios of the main simulation.

The other possibility is that the fact that some of the CTO processes full capacity weeks are possibly ones in which the demand was not quite at the maximum level and the modules were placed into the module inventory. This ability of spreading out the production of the critical modules would explain the much better delivery reliability and less capacity-dependent performance of the CTO process.

6.2.4 Summary

The simulations show that the CTO process seems to be more reliable with deliveries both in situations where there is ample capacity and in more tight capacity scenarios too. The CTO process managed the main simulation with virtually no late deliveries, but the delivery reliability for the MTO process was also quite good at on average about 98% of deliveries were delivered on time.

When analyzing the effects of the different input variables, the capacity modifier proved to be the most influential. The other variables show some command over the trends in the distribution of late deliveries, but when the capacity gets closer to the maximum demand, the delivery reliability naturally increases close to perfect, especially when there is no manufacturing or supply uncertainty.

As expected, the supplementary simulation showed a lot more variation in the delivery reliability in both processes. The impact of the capacity was reinforced even more, as the trends which the other variables seemed to bring about were outweighed even more by capacity.

Deviation from average demand naturally seemed to decrease the delivery reliability when there was more demand. There were seemingly no clear trends created by supply delivery time, which is likely due to its sporadic effects on the BOM and how the BOM then inherently influences the inventory buffers of the MTO components.

The customer delivery times stayed quite reasonable in the main simulation with the MTO process having a couple of deliveries be three weeks late while the other late deliveries were mostly late by just a week. In the supplementary simulation the delivery times ranged all the way to over 10 weeks late. The really late deliveries were probably in the scenarios where the overall delivery reliability was close to 0 as the capacity was simply too low to keep up with the demand even at the beginning.

The most interesting takeaway in the delivery reliability front was the differing capacity thresholds for the processes where the delivery reliability really started to suffer. For the MTO process the threshold was close to the average demand as in when the capacity to produce dips below the average demand, the delivery reliability starts to drop exponentially. For the CTO process similar behavior can be seen at the 40% mark so the CTO process seems to be able to withstand the capacity deficiency better.

The fact that the MTO process produced late deliveries even when there was ample capacity seems to indicate that either the inventory buffers help the delivery reliability a lot or that the MTO process is worse at handling and recovering from extreme demand patterns.

The analysis of the capacity utilization shows that while the overall capacity utilization is very similar, the CTO process has a lot more weeks which there would be capacity left while the MTO process has more weeks where either 0 or the maximum amount was produced. This shows the rather obvious fact that the ability to spread the production out on more days would help a lot with delivery reliability.

7. DISCUSSION

This chapter first reflects on the results of this study and compares them to the theory background on the subject. Afterwards some sources of error in this study are identified. In the next section, some managerial implications are gathered to provide a concise set of recommendations and things to consider when making decisions on matters within the scope of the research. The last section suggests some topics for further research based on the findings and limitations of this study.

7.1 Conclusions

This section evaluates how well the research managed to find answers to the research questions and compares the results to the literature. Overall the study succeeded in finding reasonable answers for both research questions and providing valuable insights into the problems at hand. Also, the results of the study are quite comprehensively in line with the literature on the subject. The main theoretical contribution of this study is reinforcing many interesting findings from previous studies on the subject, as well as providing some additional context to some theories.

The first research question was targeted to identify and measure differences between the processes in terms of inventory investment and delivery reliability. This was very successful in the scope of the research. The inventory investment results show that the CTO process requires a lot more inventories. This larger inventory investment requirements in CTO compared to MTO are well recognized in literature (e.g. Vollmann et al. 2005, pp. 23; Willner et al. 2014). In addition to this, the results give interesting insights on the magnitude of this effect when processes similar to the ones modelled in this study.

For the delivery reliability results, the CTO process was found to perform better under these circumstances. An interesting thing to note is that the CTO process should almost by definition be able to deliver products faster than a pure MTO process, as some production steps are done in advance (e.g. Vollmann et al. 2005, pp. 22). For example, Su et al. (2010) mark this as one of the main benefits of a CTO process. This shows that the MTO process was by-design at a disadvantage in the delivery reliability portion of this research, which manifested itself in more constrained production capabilities and worse delivery reliability. However, as the CTO process was in the same vein likely forced to keep some of the already assembled products in inventory until the delivery date, the inventory investment results could have been worse than what they might have been.

The study's findings provide ample evidence for the impact of the different input variables, which was the focus of the second research question. The notion that capacity is very tightly linked with delivery reliability is a natural conclusion and for example Harrison and Skipworth (2007) remark that the lack of sufficient excess capacity can cause serious delivery reliability complications. Closs et al. (2010) and Nyaga et al. (2007) find similar results, and this is also reinforced by the results of this study.

Both Closs et al. (2010) and Nyaga et al. (2007) also come to the conclusion that in a CTO environment increasing capacity starts to provide diminishing returns when a certain threshold is passed. They both set the threshold at around 50% over the average capacity (Nyaga et al. 2007; Closs et al. 2010). This is very clearly reinforced by this study as seen in figure 29. After the capacity modifier goes over 0.5, which is the average capacity in this case, the delivery reliability is already almost perfect, so increasing the modifier doesn't enhance the delivery performance that much anymore.

Nyaga et al. (2007) state that the optimum capacity for a CTO process at 125% of the average capacity and if the capacity goes under the average demand, delivery reliability starts to suffer greatly. This is partly reinforced as the exponential spike in late deliveries starts roughly at the 50% mark for the CTO process. However, while this study did not strive to discern an optimum value for the capacity modifier, it can be seen that after that this process manages an almost perfect delivery record almost all the way to the average demand, which might indicate that the optimum value for this process is lower than that which Nyaga et al. (2007) studied.

However, in their research on CTO processes Closs et al. (2010) found not relationship between inventory levels and order fill rate, as long as the inventory levels are sufficient to meet the demand. Similar findings could be theorized from the results of this study. The absence of a relationship between delivery reliability and supply lead time modifier combined with the very strong relationship between supply lead time modifier and inventory investment seem to reinforce this theory.

Another consideration in the delivery process is the OPP positioning, which dictates the production orders and inventories. This type of effect is stated by for example Olhager (2003) who contemplates the role of the OPP in the compromise between delivery lead time and reduction of inventories. As stated before, a similar trade-off is seen in this study between the delivery reliability and inventory investment. By enabling part of the manufacturing to happen even when a customer order is not yet placed, the capacity utilization of the CTO process is much more relaxed, but the more continuous production require more inventory buffering.

Another thing enabling the earlier production is the modular product structure, which reinforces the notion that a modular product structure and delayed differentiation enable quicker deliveries and in this case more reliable production as there is more room to spread the module assemblies out time-wise (Harrison & Skipworth 2007).

It could also be noted that, on a more theoretical level, the inclusion of more uncertainty management strategies in the form of increased safety buffers and delaying differentiation of the products enabled the CTO process to perform better in the presence of demand uncertainty. Christopher & Holweg (2011) state that variability decreases performance by causing poor capacity utilization and increased inventory buffers. Both can be detected in the simulation especially in the scenarios which performed well.

7.2 Error evaluation

As is the case with all simulations, the one used in this study has its shortcomings. This section elaborates on the things that should be considered when evaluating the results of the simulation. Much of the inconsistencies in the operations stems from the scope-complexity dilemma which plagues all simulations. The more realistic and precise the simulation is, more work goes into creating and especially validating the model. However, overall the simulation provided consistent and reasonable results, but there naturally was a couple of disparities in its theoretical basis and operation.

There are some very important things to consider when analyzing the results of the simulation. The fact that the only form of uncertainty implemented in the simulation is the demand uncertainty is one of the biggest considerations from a theoretical standpoint. While the scope of the study was to analyze the effects of demand uncertainty, it should be noted that including supply and process uncertainties into the model would change the results a lot. The buffer inventories would likely need to be adjusted and some type of safety lead times would need to be included in the weekly module assemblies.

The power of the threshold values which arise from the intermittent supply delivery time spread of the components created some irregularities in the data, which likely increased the error values in the regression analysis and probably also muddied the results of the delivery reliability. This was touched on in section 6.2.2 when the seemingly sporadic effect of supply lead time on delivery reliability was discussed. The phenomenon arises when big groups of components go from non-buffered to buffered at certain supply lead time boundaries which causes the inventory investment and the delivery reliability benefits of buffers to increase by large leaps rather than incrementally.

Also, similar threshold values were found to occur when the adjusted maximum demand values were calculated. This resulted in some inconsistencies in the inventory investment-section, as for example 2 and 3 demand scenarios had similar maximum dependent demands for modules and then there was a big gap to the values in the scenarios which had demand of 4, as some modules and their components were determined to need more buffering.

The effects of the aforementioned threshold-values were likely not too harmful to the overall precision of the simulation. More intricate decision-making could have diminished the influence of these thresholds but also made the simulation more complex which might have caused more problems of the same nature.

The fact that the simulation did not resolve very late deliveries as lost transactions probably caused some degree of overestimation of the delivery reliability statistics towards the lower end of the capacity spectrum. Including a mechanism to degree a delivery lost and to free the already created modules back for usage in later deliveries would certainly change the statistics and maybe provide some deeper insights to delivery reliability. A mechanism was actually implemented for this purpose at an early stage of the simulation model creation, but it proved quite complex to implement and it was scrapped after some discussion to maintain the simplified structure of the simulation.

It could also be argued whether the regression analysis was the most apt way of describing the inventory investment results. However, while the precision and statistical characteristics of the analysis were on point, the aim of the regression analysis was to describe the dataset and not to predict the values the precision is not as important but rather the simplicity and clarity of the way the regression equations can explain the relationships in the dataset. Also, it should be noted that the results of the regression analysis should never be extrapolated as the regression provides no guarantees that the equations work outside the scope of the input variables (Elliott & Woodward 2007, pp. 96).

7.3 Managerial implications

The essential managerial implication of this study is the valuable insights it provides into the differences of the investigated processes. While some of the results may seem obvious, the magnitude of the differences and the different behavior of the processes in different environments can contribute into creating a better understanding of the possibilities a company has in its operations.

In the inventory investment the inventory values provided can be used to help in decision making as rough guidelines when planning future processes. Nevertheless, the fact that changes to the product structure or even the supply profile of some of the more impactful components can deviate the results quite a lot should not be overlooked.

It can be seen from the results that if customer delivery time is an essential source of competitive advantage, the CTO process is probably the better option of the two, even at the cost of larger inventories. Also, if demand uncertainty is seen as a big factor, then the flexibility provided by the CTO process can prove itself invaluable.

It should be noted that given ample capacity, the MTO process is able to perform at the same level of delivery reliability as the CTO process with much less inventories. However, as the simulation did not include e.g. supply or process uncertainty, the performance of a company's processes and suppliers should be critically evaluated case by case before taking these results as granted.

7.4 Further research topics

This study provided ample evidence of possible interesting avenues for future research. Many suggestions can be drawn from the limits of this research's scope. A natural direction for research could be to implement different types of uncertainty into the study. This could be done for example by swapping out demand uncertainty and adding supply uncertainty or adding another type of uncertainty into a model which already imitates demand uncertainty. This could create a more realistic behavior for the model and maybe enable setting the inventory and lead time buffers at more precise levels.

A longitudinal time-horizon on a research on this subject might provide some interesting insights into the study. A longer time-horizon would enable better demand distribution creation and a real-life comparison for the model validation. In addition, it could make modelling the aforementioned uncertainty patterns possible if the study could utilize longitudinal data on assembly and supply performance.

As discussed in section 7.1, comparing CTO and MTO processes with the same delivery time requirements may not be the most insightful approach. Therefore, it could prove very interesting to compare the performance with a freer delivery window which could promote the studying of the discussed quicker delivery performance of the CTO process and how that would affect the inventory levels and delivery reliability.

Implementing more intricate ordering policies could improve the precision of the study. Taking a cost perspective instead of the straight-up tied-up capital could give more valuable insights into the optimal ordering policies and create a better understanding of the actual costs incurred by the different processes. This might also give a better view into the optimal inventory policies for the inventory buffers in both processes if the ordering costs would overweigh the inventory holding costs which might increase the optimal inventory investment for either process.

REFERENCES

Angkiriwang, R., Pujawan, I.N. & Santosa, B. (2014). Managing uncertainty through supply chain flexibility: reactive vs. proactive approaches, Production & Manufacturing Research, Vol. 2(1), pp. 50-70.

Appelqvist, P. & Gubi, E. (2005). Postponed variety creation: case study in consumer electronics retail, International Journal of Retail & Distribution Management, Vol. 33(10), pp. 734-748.

Atan, Z., Ahmadi, T., Stegehuis, C., Kok, T.d. & Adan, I. (2017). Assemble-to-order systems: A review, European Journal of Operational Research, Vol. 261(3), pp. 866-879.

Baldwin, C.Y. & Clark, K.B. (1997). Managing in an age of modularity, Harvard business review, Vol. 75(5), pp. 84-93.

Chen-Ritzo, C., Ervolina, T., Harrison, T.P. & Gupta, B. (2010). Sales and operations planning in systems with order configuration uncertainty, European Journal of Operational Research, Vol. 205(3), pp. 604-614.

Cheng, F., Ettl, M., Lin, G. & Yao, D.D. (2002). Inventory-Service Optimization in Configure-to-Order Systems, Manufacturing & Service Operations Management, Vol. 4(2), pp. 114-132.

Christopher, M. & Holweg, M. (2011). "Supply Chain 2.0": managing supply chains in the era of turbulence, International Journal of Physical Distribution & Logistics Management, Vol. 41(1), pp. 63-82.

Closs, D.J., Nyaga, G.N. & Voss, M.D. (2010). The differential impact of product complexity, inventory level, and configuration capacity on unit and order fill rate performance, Journal of Operations Management, Vol. 28(1), pp. 47-57.

Davis, T. (1993). Effective supply chain management, Sloan management review, Vol. 34(4), pp. 35-46.

Duray, R., Ward, P.T., Milligan, G.W. & Berry, W.L. (2000). Approaches to mass customization: configurations and empirical validation, Journal of Operations Management, Vol. 18(6), pp. 605-625.

Elliott, A.C. & Woodward, W.A. (2007). Statistical Analysis Quick Reference Guidebook, SAGE Publications, Inc, Thousand Oaks.

Florescu, I. & Tudor, C.A. (2013). Handbook of probability, John Wiley & Sons, Incorporated, Hoboken.

Fogliatto, F.S., da Silveira, Giovani J. C. & Borenstein, D. (2012). The mass customization decade: An updated review of the literature, International Journal of Production Economics, Vol. 138(1), pp. 14-25.

Gosling, J. & Naim, M.M. (2009). Engineer-to-order supply chain management: A literature review and research agenda, International Journal of Production Economics, Vol. 122(2), pp. 741-754.

Guo, S., Choi, T., Shen, B. & Jung, S. (2019). Inventory Management in Mass Customization Operations: A Review, IEEE Transactions on Engineering Management, Vol. 66(3), pp. 412-428.

Harrison, A. & Skipworth, H. (2007). Implications of form postponement to manufacturing: a cross case comparison, International Journal of Production Research, Vol. 46(1), pp. 173-195.

Haug, A., Ladeby, K. & Edwards, K. (2009). From engineer-to-order to mass customization, Management Research News, Vol. 32(7), pp. 633-644.

Heikkilä, J., Karjalainen, T., Martio, A. & Niininen, P. (2002). Products and modularity: Managing competitive Product Portfolios through Holistic Platform Thinking. TAI Research Centre, Helsinki University of Technology, Espoo.

Heikkilä, J. & Ketokivi, M. (2005). Tuotanto murroksessa: strategisen johtamisen uusi haaste, Talentum, Helsinki.

Hopp, W.J. & Spearman, M.L. (2000). Factory physics: foundations of manufacturing management, 2nd ed. Irwin McGraw-Hill, New York; Singapore.

Jacobs, M., Vickery, S.K. & Droge, C. (2007). The effects of product modularity on competitive performance, International Journal of Operations & Production Management, Vol. 27(10), pp. 1046-1068.

Klee, H. & Allen, R. (2011). Simulation of dynamic systems with MATLAB and Simulink, 2nd ed. CRC Press, Boca Raton, FL, 795 pages.

Laguna, M. & Marklund, J. (2013). Business process modeling, simulation, and design, 2nd ed. Taylor & Francis, Boca Raton, xiii, 509 pages.

MacCarthy, B.L. (2013). An analysis of order fulfilment approaches for delivering variety and customisation, International Journal of Production Research, Vol. 51(23-24), pp. 7329-7344.

Miclo, R., Lauras, M., Fontanili, F., Lamothe, J. & Melnyk, S.A. (2019). Demand Driven MRP: assessment of a new approach to materials management, International Journal of Production Research, Vol. 57(1), pp. 166-181.

Montgomery, D.C., Peck, E.A. & Vining, G.G. (2012). Introduction to linear regression analysis, 5th ed. Wiley, Hoboken.

Nyaga, G.N., Closs, D.J., Rodrigues, A. & Calantone, R.J. (2007). The impact of demand uncertainty and configuration capacity on customer service performance in a configure-to-order environment, Journal of Business Logistics, Vol. 28(2), pp. 83-104.

Olhager, J. (2003). Strategic positioning of the order penetration point, International Journal of Production Economics, Vol. 85(3), pp. 319-329.

Pagh, J.D. & Cooper, M.C. (1998). Supply chain postponement and speculation strategies: How to choose the right strategy, Journal of Business Logistics, Vol. 19(2), pp. 13.

Partanen, J. & Haapasalo, H. (2004). Fast production for order fulfillment: Implementing mass customization in electronics industry, International Journal of Production Economics, Vol. 90(2), pp. 213-222.

Robinson, S. (1994). Successful simulation: a practical approach to simulation projects, McGraw-Hill, London, 246 pages.

Saeed, M., Loya, S. & Loya, A. (2016). A Review on Strategic Positioning of Bottleneck around the Customer Order Decoupling Point and Issues on Production Planning in Supply Chain, Journal for Manufacturing Science and Production, Vol. 16(4), pp. 243.

Salvador, F., Forza, C. & Rungtusanatham, M. (2002). Modularity, product variety, production volume, and component sourcing: theorizing beyond generic prescriptions, Journal of Operations Management, Vol. 20(5), pp. 549-575.

Sanchez, R. & Mahoney, J.T. (1996). Modularity, Flexibility, and Knowledge Management in Product and Organization Design, Strategic Management Journal, Vol. 17(S2), pp. 63-76.

Saunders, M., Lewis, P. & Thornhill, A. (2009). Research methods for business students, 5th ed. Prentice Hall, Harlow.

Simangunsong, E., Hendry, L.C. & Stevenson, M. (2011). Supply-chain uncertainty: a review and theoretical foundation for future research, International Journal of Production Research, Vol. 50(16), pp. 4493-4523.

Sterman, J.D. (2000). Business dynamics: systems thinking and modeling for a complex world, McGraw-Hill, Boston (Mass.).

Stevenson, M., Hendry, L.C. & Kingsman †, B.G. (2005). A review of production planning and control: the applicability of key concepts to the make-to-order industry, International Journal of Production Research, Vol. 43(5), pp. 869-898.

Stevenson, W.J. (2011). Operations management, 11th ed. McGraw-Hill/Irwin, New York.

Su, J.C.P., Chang, Y. & Ferguson, M. (2005). Evaluation of postponement structures to accommodate mass customization, Journal of Operations Management, Vol. 23(3), pp. 305-318.

Su, J.C.P., Chang, Y., Ferguson, M. & Ho, J.C. (2010). The impact of delayed differentiation in make-to-order environments, International Journal of Production Research, Vol. 48(19), pp. 5809-5829.

Ulrich, K. & Eppinger, S.D. (2012). Product design and development, 5th ed. McGraw-Hill Irwin, New York, NY.

Ulrich, K. (1995). The role of product architecture in the manufacturing firm, Research Policy, Vol. 24(3), pp. 419-440.

Vollmann, T.E., Berry, W.L., Whybark, D.C. & Jacobs, F.R. (2005). Manufacturing planning and control for supply chain management, 5th ed. McGraw-Hill, Boston.

Wikner, J. & Johansson, E. (2015). Inventory classification based on decoupling points, Production & Manufacturing Research, Vol. 3(1), pp. 218-235.

Willner, O., Powell, D., Duchi, A. & Schönsleben, P. (2014). Globally Distributed Engineering Processes: Making the Distinction between Engineer-to-order and Make-to-order, Procedia CIRP, Vol. 17 pp. 663-668.

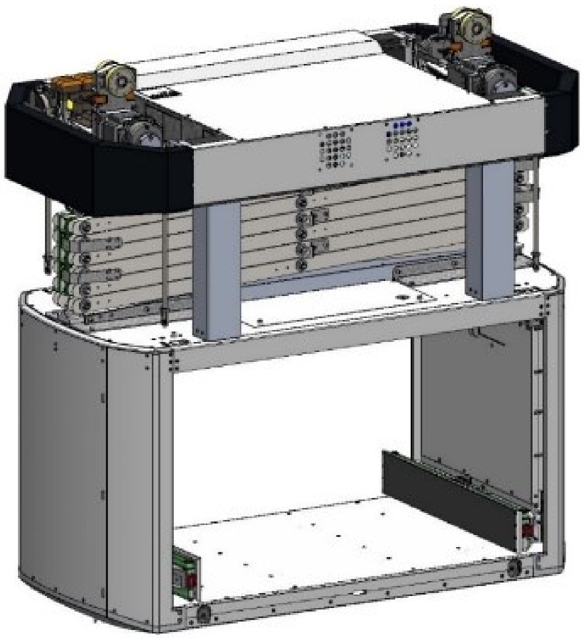
Yang, B., Burns, N.D. & Backhouse, C.J. (2004). Management of uncertainty through postponement, International Journal of Production Research, Vol. 42(6), pp. 1049-1064.

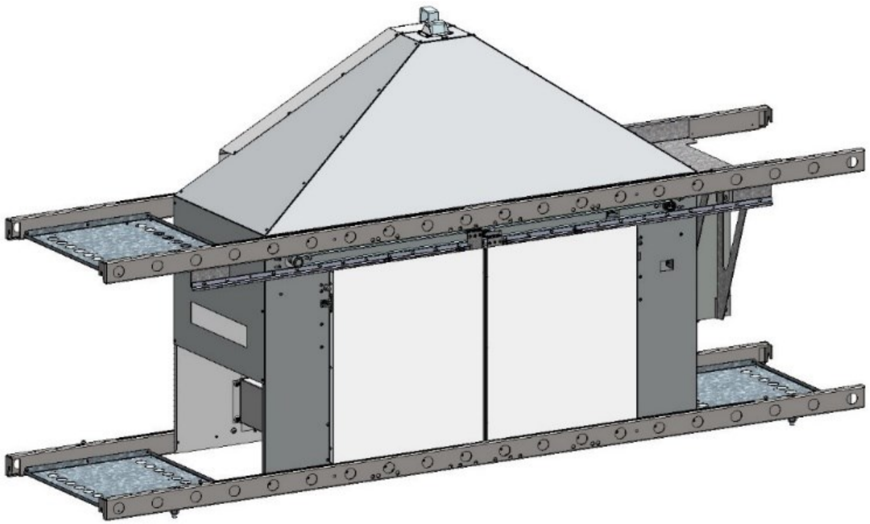
Yang, B. & Burns, N. (2003). Implications of postponement for the supply chain, International Journal of Production Research, Vol. 41(9), pp. 2075-2090.

Zäpfel, G. (1998). Customer-order-driven production: An economical concept for responding to demand uncertainty? International Journal of Production Economics, Vol. 56(1), pp. 699-709.

APPENDIX A: THE AGILON MODULES

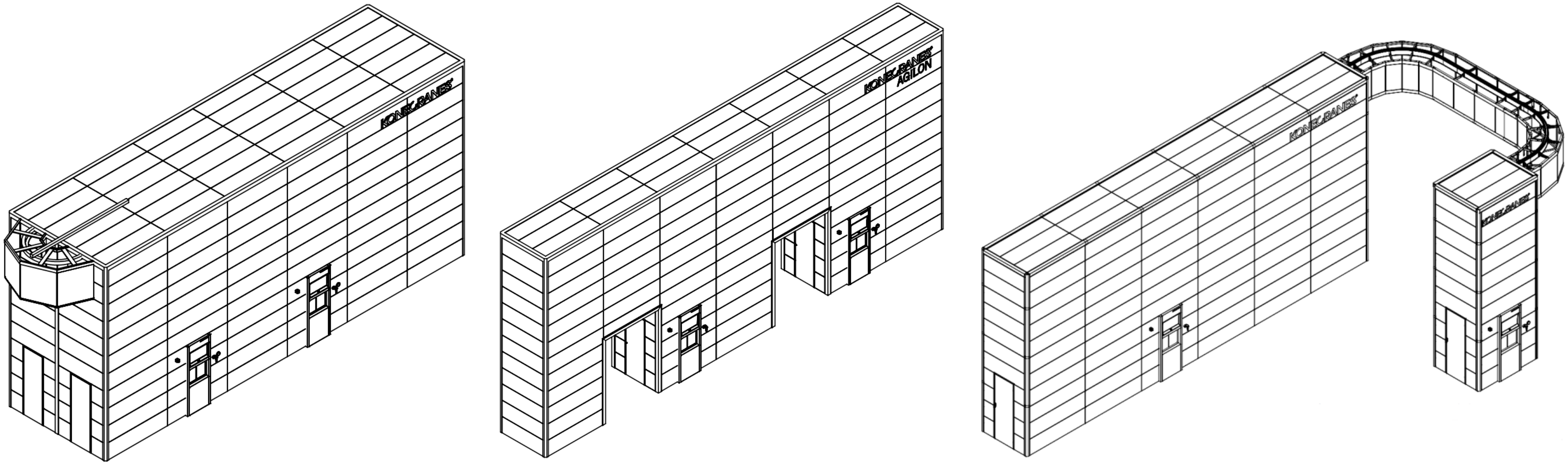
Left to right: Robot, Access point, Industrial mask.



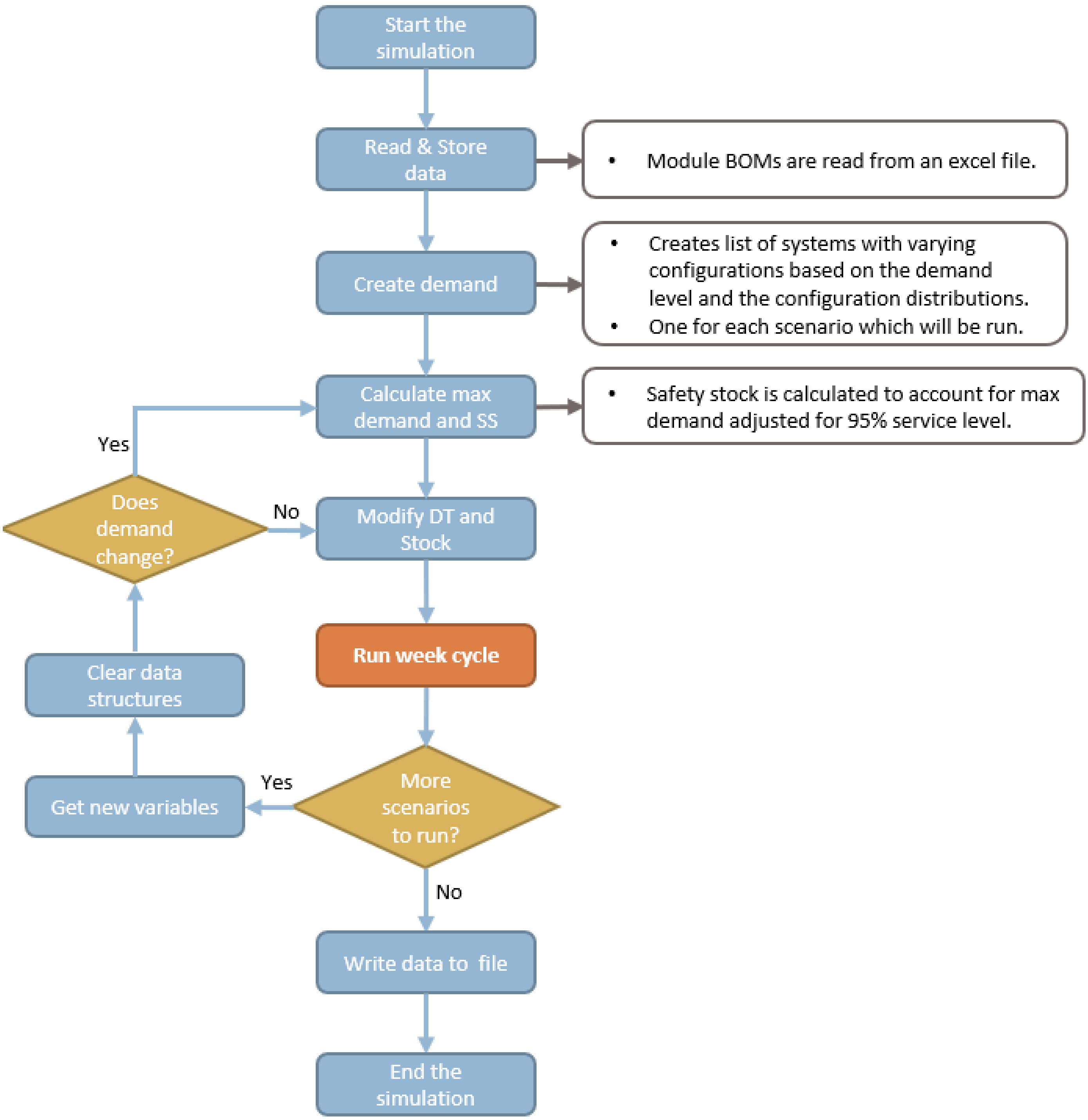


APPENDIX B: AGILON SYSTEM EXAMPLES

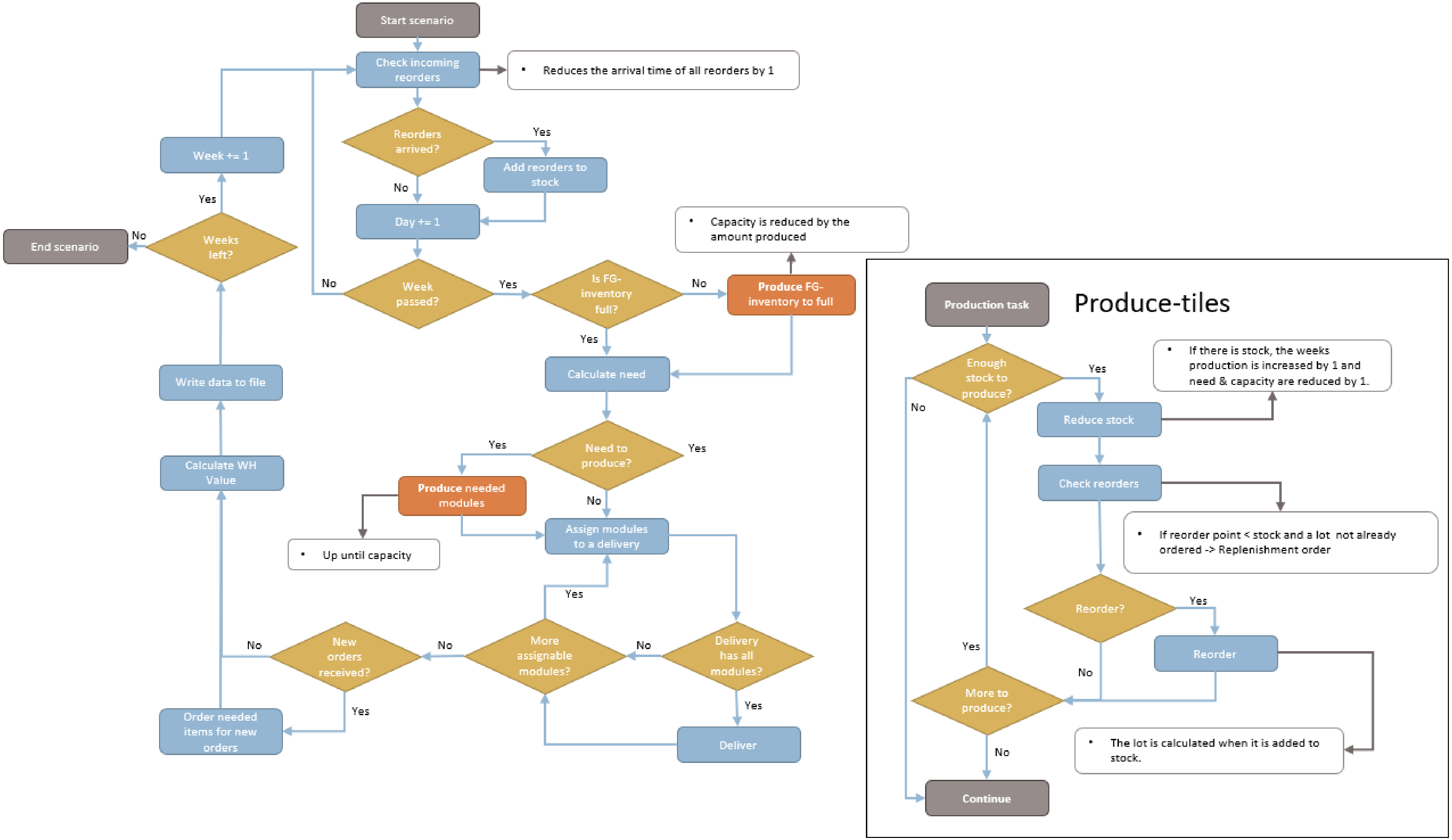
Left to right: A stacked Agilon, An Agilon with two underpasses and an Agilon with a connection tube connecting two units.



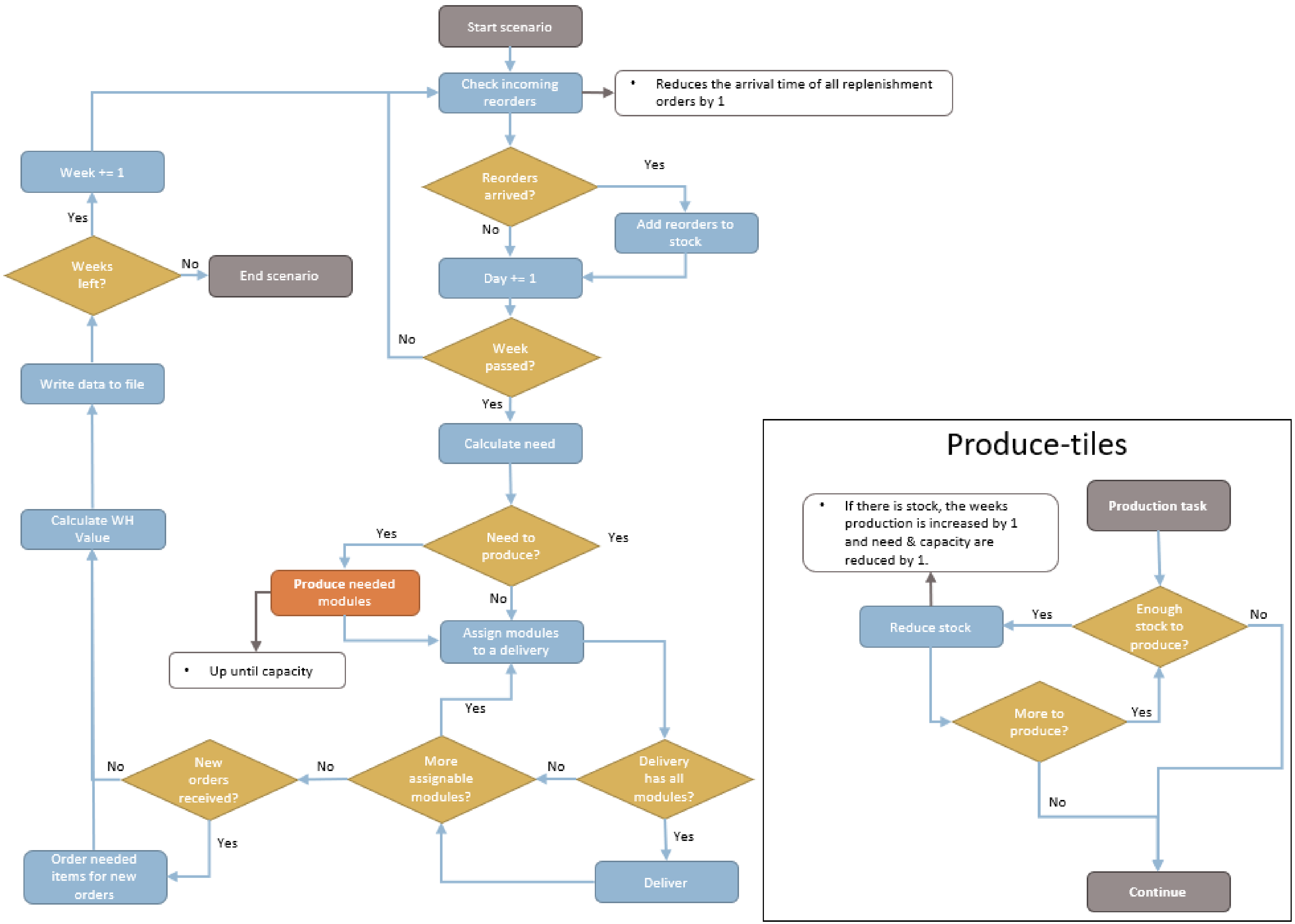
APPENDIX C: INITIALIZATION AND SCENARIO CYCLE FLOW CHART



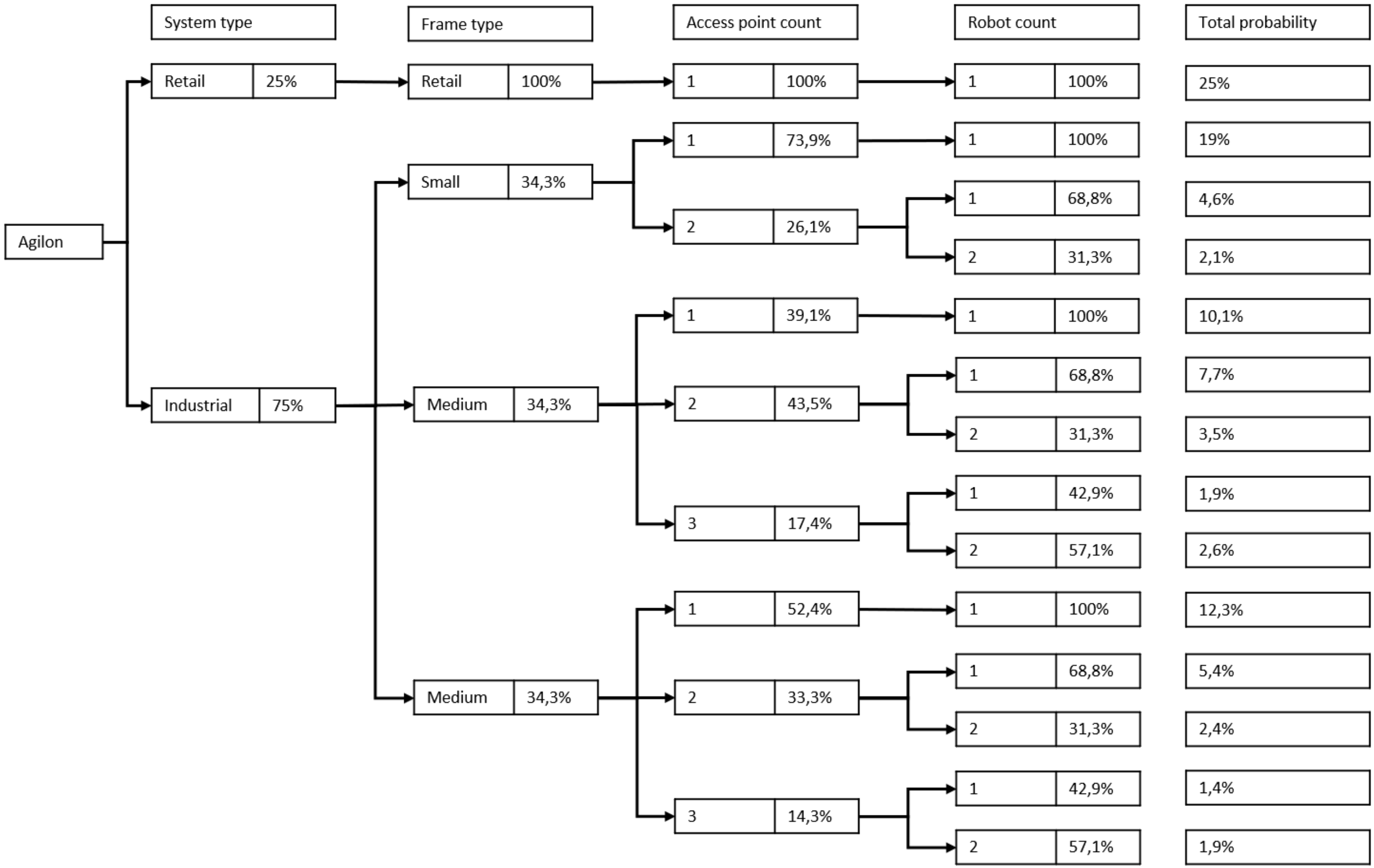
APPENDIX D: COMPLETE CTO FLOW CHART



APPENDIX E: COMPLETE MTO FLOW CHART



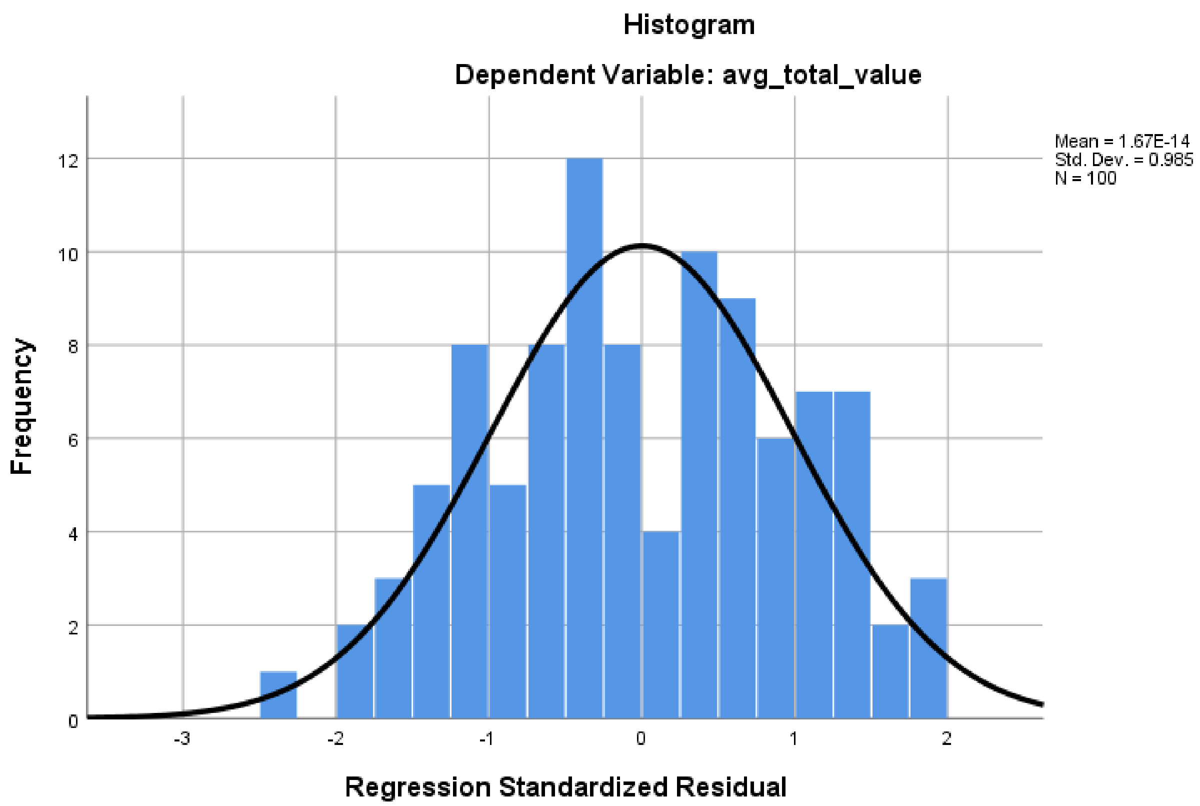
APPENDIX F: SYSTEM CONFIGURATIONS AND PROBABILITIES

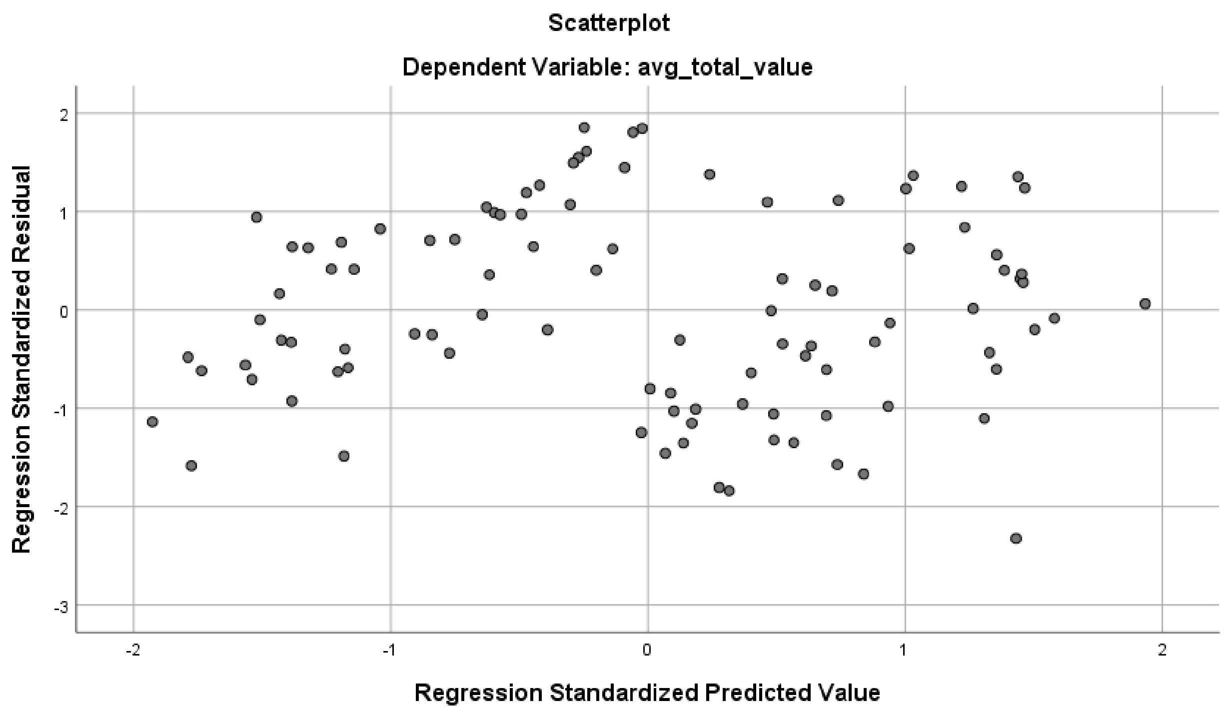


APPENDIX G: TABLE OF THE CORRESPONDENCE BETWEEN SYSTEM AND ROBOT DEMAND AND THE TOTAL PROBABILITIES FOR ROBOT DEMAND

Robot \ *System*	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Total**
0	11,11 %									11,1111 %
1		9,86 %								9,8556 %
2		1,26 %	8,74 %							9,9974 %
3			2,23 %	7,75 %						9,9814 %
4			0,14 %	2,96 %	6,88 %					9,9832 %
5				0,38 %	3,50 %	6,10 %				9,9830 %
6				0,02 %	0,67 %	3,89 %	5,41 %			9,9830 %
7					0,057 %	0,99 %	4,14 %	4,80 %		9,9830 %
8					0,002 %	0,13 %	1,32 %	4,28 %	4,26 %	9,9830 %
9						0,008 %	0,22 %	1,64 %	4,34 %	6,206 %
10						0,0002 %	0,021 %	0,35 %	1,93 %	2,303 %
11							0,0011 %	0,044 %	0,49 %	0,5383 %
12							0,00002 %	0,0034 %	0,078 %	0,0819 %
13								0,00014 %	0,008 %	0,0081 %
14								0,000003 %	0,00051 %	0,0005 %
15									0,000019 %	0,00002 %
16									0,0000003 %	0,0000003 %

APPENDIX H: EXAMPLE REGRESSION RESIDUAL ANALYSIS RESULTS





APPENDIX I: MAIN SIMULATION DELIVERY RELIABILITY RESULTS GROUPED BY THE INPUT VARIABLES