MASTER

Minimizing storage capacity based on reorder levels for an online retailer
developing a model to translate historical sales into storage capacity

Vissers, L.

Award date:
2017

Utrecht, January 2017

Minimizing storage capacity based on reorder levels for an online retailer

Developing a model to translate historical sales into storage capacity

Industrial Engineering & Innovation Sciences
Eindhoven University of Technology

By
L. (Lieke) Vissers

BSc Industrial Engineering & Management Science – TU/e 2014
Student identity number 0758482

In partial fulfillment of the requirements for the degree of
Master of Science
In Operations Management and Logistics

Supervisors TU/e:	Dr. ir. R.A.C.M. Broekmeulen, OPAC
	Dr. A.E. Akçay, OPAC
Supervisor bol.com:	Ing. E. Kempkes, Team Lead Supply Chain

TUE - School of Industrial Engineering

Series Master Thesis Operations, Management and Logistics

Subjects: sales forecasting, trend & seasonality, reorder level determination, capacity planning, re-warehousing/healing

Abstract

This study includes the development of a model in order to translate historical sales data into the required storage capacity at the distribution center. It minimizes the required capacity by reducing the required reorder levels considering the target fill rates. The initiated problem is the lack of an automatic procedure on how to determine the required storage capacity. Currently, a monthly sales forecast results in a lot of manual work that has to be done in order to estimate more sufficient forecasts. More importantly, it results in being unable to timely communicate the required forecasts to the distribution center. Different classification methods regarding trend and seasonality, forecasting methods and reorder level determination methods have been compared and discussed. Finally, the developed model that represents the best practice strategy regarding forecasting the required storage capacity based on minimizing the reorder levels is provided. The redesigned model visualizes the trade-off between performance indicators in the retail industry. This research provides new insights into classification systems, reorder level determinations and capacity planning in order to improve future decision-making.

Management summary

This report represents the results of a master thesis project, which is conducted at bol.com, a Dutch online retailer. In 2015, problems regarding inbound, storage and outbound occurred. Commonly, storage capacity would have been the main issue. However, if an accurate capacity forecast was timely communicated, enough storage capacity would have been available. Due to an inefficient process of forecasting the required storage capacity, bol.com was not able to timely communicate an accurate capacity forecast. Therefore, the distribution center was not prepared for the large amounts of inbound, storage and outbound streams. Consequently, the issue in this research is how to automatically determine the required storage capacity. In this way, the distribution center can be timely prepared and capacity would not be an issue.

Existing literature lacks of information regarding this matter. Consequently, this research contributes to the literature on the following topics: e-commerce retailing with regard to capacity planning, the classification of stock keeping units regarding trend and seasonality, determination of reorder levels and translating historical sales data into the required storage capacity.

Project definition

Due to a dedicated storage policy, the required storage capacity is equal to the capacity required for the maximum inventory on hand. Moreover, the maximum inventory on hand is based on the determined reorder levels and fixed case pack sizes. Therefore, the reorder level per stock keeping unit is the decision variable in this research. It consist of cycle stock and safety stock, with the target fill rates as input variable. The fill rate is defined as the long-term fraction of the demand that is immediately delivered from stock. Consequently, the objective of the research is *to minimize the storage capacity based on the required reorder levels per stock keeping unit.*

The mentioned problems led to the following research question:

What is the required storage capacity at Docdata in order to achieve the desired fill rates?

Docdata is defined as the distribution center in scope. In order to answer the research questions, the following steps have been taken: a classification procedure regarding trend and seasonality, forecasting the sales, estimating the reorder levels and finally, defining the forecast of storage capacity. The sales forecast consists of the expected units of sales per stock keeping unit per week. The stock forecast contains the required maximum inventory levels based on the reorder levels per stock keeping unit per week. Finally, the forecast of storage capacity includes the required storage capacity per stock keeping unit in cubic meters per week.

Conceptual design

During the development of the model, different challenges have to be addressed. Firstly, it should be mentioned that lost sales apply. The lost sales are defined as the fraction of demand that is not immediately delivered from stock (1 – fill rate). Furthermore, many different stock keeping units with their own characteristics require a classification method in order to choose a suitable forecasting method. The classification method proposed by Gardner Jr & McKenzie (1988) (GaMcK) is compared with the method that selects the most appropriate forecasting method based on the smallest forecast error (MSE), which is mentioned by Hyndman, Koehler, Snyder, & Grose (2002) as well. For the determination of the sales forecast, different exponential smoothing models are used, which take level, trend and seasonality components into account. Furthermore, determining the reorder levels is a challenge. Initially, the mean and variance resulting from the forecasting methods are used. Four methods are evaluated in order to determine the reorder levels:

1. If forecasted relationship between mean and variance does not exist according to Adan, van Eenige, & Resing (1995), fit to Poisson distribution. Use Adan, van Eenige, & Resing (1995) to create probability mass function. (MMM)
2. If forecasted relationship between mean and variance does not exist or no sufficient data is available, use aggregated variance over SKU's acquired by Taylor's Power law (Taylor, 1961). Use Adan, van Eenige, & Resing (1995) to create probability mass function. (MMM-PL)
3. If forecasted relationship between mean and variance does not exist, fit to Poisson distribution. Use method of van Eenige (1996) to create probability mass function. (MMM-vE)
4. If forecasted relationship between mean and variance does not exist or no sufficient data is available, use aggregated variance over SKU's acquired by Taylor's Power law (Taylor, 1961). Use method of van Eenige (1996) to create probability mass function.(MMM-PL-vE)

Looking at the general relationship between the mean and variance, the aggregated variance is acquired with use of a power law. The power law provides the relationship between the mean and the variance of the population through time or through space. This method is used as a statistical description of aggregation for the formulation of transformations that stabilize the variance (Taylor, 1961). The method of van Eenige (1996) uses the maximum demand observation as input variable in order to create a probability mass function with limited support. The reorder levels are determined using the probability mass function while assuming a lost sales inventory system, positive lead times and target fill rates using a periodic review policy (van Donselaar & Broekmeulen, 2013). The maximum inventory on hand is determined based on the reorder levels and the case pack size. The required capacity is estimated by multiplying the maximum inventory on hand by the volume per SKU.

The order-up-policy proposed by Eaves & Kingsman (2004) is used as a benchmark model to compare and evaluate the models with corresponding results. Using this model, the minimal reorder levels per stock keeping unit that would have resulted in achieving the target fill rate are determined. These levels are equal to the minimal reorder levels required for achieving the target fill rates. It is not reasonable to assume that the proposed methods can provide the same results. However, it is used as an indicator for comparison and to determine which of the proposed methods fits best.

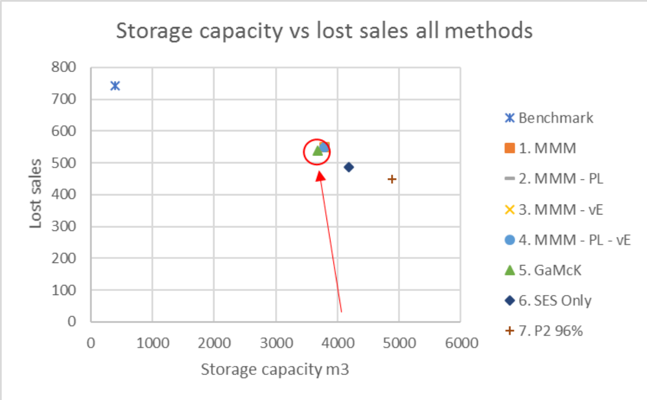
Furthermore, three scenarios are discussed in order to determine the most appropriate model and to determine which factors are the most important drives of the model. The first scenario is using the classification method proposed by Gardner Jr & McKenzie (1988) to identify the most appropriate forecasting model (GaMcK). The second scenario is excluding the classification methods and is using simple exponential smoothing to forecast the mean and variance (SES Only). The last scenario is setting all target fill rates to 96% in order to look at the impact of the target fill rate (P2 96%). It can be concluded that these scenario's did not result in improvements of the model.

Formal model

The objective is translated into a mathematical model that minimizes the storage capacity per stock keeping unit by reducing the required reorder levels with a service level constraint. Based on the amount of lost sales, and the required storage capacity, the most appropriate model is selected which should be used for all SKU's.

Results

The results of the four different methods and three scenarios are provided in the Figure below.



As can be derived from the Figure, the method proposed by Gardner Jr & McKenzie (1988) is selected as most appropriate model in order to estimate the forecast of required storage capacity. Comparing to the other methods and scenarios, this method resulted in the least required storage capacity and low levels of lost sales. Taylor's Power law (Taylor, 1961) or the method of van Eenige (1996) did not results in any improvements.

Conclusion & Recommendations

This study developed a model in order to translate historical sales data into the required storage capacity. If desired, the sales and stock forecast can be derived as well. The following recommendations for bol.com are provided:

1. *Implement the automatic forecasting model that deals with the dynamic environment of e-commerce retailing.*
2. *Update model frequently and provide a m-step ahead forecast to Docdata*.
3. *Determine if broad assortment with many slow movers is advantageous.*
4. *Use fill rate as service level measurement*.
5. *Determine if the ABC- classification is sufficient*.

This research contributes to the literature on e-commerce retailing with regard to capacity planning based on reorder levels. E-commerce retailing differs significantly from traditional retailing on several areas. This research clarifies these differences and provides recommendations to deal with e-commerce retailing with regard to capacity planning. Moreover, classification methods of SKUs regarding trend and seasonality are discussed and evaluated, which resulted in interesting outcomes. Additionally, it contributes to the literature with regard to fluctuating order patterns and reorder level determinations. These issues make forecasting and developing a capacity plan more difficult. Finally, this research contributed to the aspect of inventory management. Several methods of forecasting the mean and variance for calculating the reorder levels have been addressed.

Preface

This report is the result of a master thesis project, which is conducted at bol.com in Utrecht. It is written in partial fulfillment of the requirements for the degree of Master of Science in Operations, Management & Logistics at Eindhoven University of Technology. Finishing this research and completing this thesis brings an end to the journey as a student.

I hope you will enjoy reading this thesis as much as I have enjoyed doing research at and being part of a positive, open, enthusiastic and ambitious company as bol.com. I was part of a great team, which helped me when needed, provided me the necessary distraction and made sure I had a great time. At the beginning, my mentor at bol.com Erwin said to me: “I am going to turn you into a professional!” He supported me, guided me and made sure I learned from every step I made. Thank you, Erwin. I have learned along the way, met a lot of people from different departments, gained a lot of knowledge which helped me with my research and beyond. Particularly, I want to thank team Supply Chain Lifestyle and Sales & Operations Planning for sharing their skills and knowledge and enriching my time at bol.com.

Moreover, I particularly would like my first supervisor, Rob Broekmeulen for always being enthusiastic, optimistic, and positive. A great mentor with feedback where I could work with. He made time to explain me everything carefully and shared his experience and knowledge that provided me additional insights into many issues. These characteristics of our collaboration resulted in a pleasant cooperation. Furthermore, I would like to thank my second supervisor, Alp Akçay. His feedback helped me to provide a structured report.

Finally yet importantly, I would like to thank my parents, my brother and sister, my friends and my boyfriend who supported me in every step I took in order to achieve the goals I have chosen. Without them, finishing this project would have been “mission impossible”, thanks!

I wish you all the best!

Lieke Vissers

List of Figures

Figure 1: Revenues of bol.com from 2002 until 2015 1
Figure 2: Main processes of bol.com 2
Figure 3: Comparison of items and SKU's stored at Docdata, peak versus regular weeks 3
Figure 4: Research model by Mitroff et al. (1974) (left) and the regulative cycle of van Strien (1986) (right). 4
Figure 5: Followed steps in order to answer the research question 10
Figure 6: Stock forecast accuracies vs volume unit Toys 16
Figure 7: Sales forecast accuracies vs volume unit Toys 16
Figure 8: Time series divided into training and holdout period 18
Figure 9: Outliers per week product category Toys - Initial method 20
Figure 10: Outliers per week product category Toys - Adapted method 20
Figure 11: Disaggregated variance to aggregated variance 26
Figure 12: PMF Adan, Eenige van , & Resing (1995) and van Eenige (1996) 27
Figure 13: Process of translating historical data into the required capacity 29
Figure 14: Demand patterns per classification group 30
Figure 15: Estimated forecast vs actuals 31
Figure 16: Power law of different phases (historical (left), training period (middle), holdout period (right)) 32
Figure 17: Results 4 methods on training period 34
Figure 18: Results all methods and scenarios 37

List of Tables

Table 1: Division of sales over the last month of 2015 7
Table 2: Target fill rates and target discrete ready rate per class according to ABC-classification 8
Table 3: Results benchmark model 11
Table 4: Bias stock forecast over 2015 17
Table 5: Bias sales forecast over 2015 17
Table 6: Product categories per unit 18
Table 7: Percentage SKU's and items sold per unit of 2015 19
Table 8: Functional requirements and design parameters per phase 20
Table 9: Fixed and variable design parameters 21
Table 10: Different possible forecasting models by GaMcK and MMM 24
Table 11: Formulated hypotheses 28
Table 12: Number of SKU's classified per model by both classification methods 30
Table 13: Second model chosen by MMM 31
Table 14: Bias of forecast of developed model 32
Table 15: Sales forecast accuracy of developed model 32
Table 16: Division SKU's ABC-classification 33
Table 17:Percentage target fill rate not achieved and percentage lost sales of different methods and scenarios 37
Table 18: Results selected model (GaMcK method) on holdout 38

List of definitions and abbreviations

Item	Total number of products.
SKU	Stock keeping unit: a distinct type of item for sale
Unit	The four main areas of bol.com (Toys, Entertainment, Lifestyle and Home & Living). The units consist of different product categories.
Product category	Bol.com has assigned their SKU's to 18 product different categories. Several product categories together form a unit.
LvB	Logistics via Bol.com
Base demand	Number of items sold per week without promotions
Adjusted base demand	Base demand after substitution of outliers
LF	Lift factor. The lift in volume in times of promotions
IOQ	Incremental order quantity. Is equal to the CPS
Q	Case pack size. Ordering should be done in multiples of the case pack size
MOQ	Minimal order quantity. The minimal amount that should be ordered of a SKU. In this research, the minimal amount is equal to the case pack size.
Review period (R)	Time between two moments when the inventory levels are reviewed
Lead time (L)	Time between a replenishment order is placed and when the order is received
Inventory on hand (IOH_t)	The actual inventory levels at Docdata on a specific time t
OOS	Out Of Stock. When the inventory on hand of an SKU is zero, that particular product is out of stock. In this case, the sales cannot be back ordered and are lost.
GaMcK	This abbreviation corresponds to article of Gardner & McKenzie (Gardner Jr & McKenzie, 1988)
MMM	Minimizing MSE method
P2	Fill rate, the long-term fraction of the demand that is immediately delivered from stock
$P3^{discrete}$	Discrete ready rate, probability of positive inventory on hand just before a potential delivery moment at specific discrete moments in time

Contents

Abstract i
Management summary ii
Preface v
List of Figures vi
List of Tables vii
List of definitions and abbreviations viii
1. Introduction 1
1.1 Introduction to bol.com 1
1.2 Research objective 2
1.3 Methodology 4
1.4 Research outline 4
2. Project definition 5
2.1 Analysis of current processes regarding forecasting and capacity planning 5
2.2 Research question and sub questions. 8
2.3 Benchmark method 11
2.4 Scope 11
2.5 Deliverables 12
2.6 Summary of project definition 12
3. Literature review 13
3.1 Forecasting model 13
3.1.1 Classification models 13
3.2 Determining reorder levels 14
3.3 Gaps in literature 14
4. Diagnosis of current performance 15
4.1 Forecast accuracy 15
4.2 Forecast accuracy vs volume 16
4.3 Under and over forecasting 16
4.4 Obsolete stock 17
4.5 Summary of diagnosis 17
5. Plan for redesign 18
5.1 Time series data 18
5.1.1 Analysis on sales data 19
5.2 Conceptual design 20
5.2.1 Dependent & independent variables 21
5.2.2 Key performance indicators 22

5.2.3 Assumptions ... 22
5.2.4 Mathematical model ... 23
5.3 Proposed plan per phase ... 23
5.3.1 Classifying the SKU's ... 23
5.3.2 Forecasting ... 25
5.3.3 Reorder level determination ... 27
5.4 Proposed hypotheses ... 28
5.5 Summary of the plan for redesign ... 28
6. Redesigned model ... 30
6.1 Results of classifying the SKU's ... 30
6.2 Results of forecasting ... 31
6.3 Results of reorder level determination ... 32
6.4 Results required capacity ... 33
7 Scenario analysis ... 35
7.1 Analysis initial scenario ... 35
7.2 Scenario A: GaMcK as classification method ... 36
7.3 Scenario B: Simple exponential smoothing ... 36
7.4 Scenario C: Change of Fill rates ... 36
7.5 Selected model ... 37
7.5 Sensitivity analysis ... 38
7.6 Evaluation of hypotheses ... 39
8 Conclusion & Recommendations ... 41
8.1 Answers to the research question ... 41
8.2 Scientific contributions ... 42
8.3 Generalizability ... 43
8.4 Recommendations & Implementation ... 43
9 Discussion & Limitations ... 45
9.1 Discussion ... 45
9.2 Limitations ... 45
9.3 Future research ... 46
Bibliography ... 48
Appendix ... 53

1. Introduction

The e-commerce market is rapidly growing, “it grew like a virus” is how Kendall (2000) describes it in his research. According to ITU (2009) and OECD (2009), the Netherlands has 87.1% internet use penetration and 0.97% secure server penetration. Compared to the rest of the world, only the US scores higher on secure server penetration than the Netherlands. Secure server penetration and internet user penetration represent the indicators for respectively the supply side and the demand side of an e-commerce market. When a country is ranking high on both internet user penetration and secure server penetration, which is the case in the Netherlands, it indicates that e-commerce activities have the potential to be in growth mode. Bol.com is one of those rapidly growing e-commerce companies in the Netherlands and Belgium. As can be seen in Figure 1, the revenues of bol.com showed substantial growth over the years and are even doubled since 2012. However, not only a growth in revenue is remarkable, bol.com itself has grown as well. From only 300 employees in 2010 to almost 1100 employees in 2016 and this number is still increasing. Furthermore, bol.com realized a growth in net consumer sales of 35.3% in 2015. The Christmas period counts for 30% of this growth (Bol.com, 2016). Compared to traditional retailing, there is a higher probability of significant fluctuation in customer demand. Therefore, e-commerce retailers experience more valleys and more peaks. Furthermore, minimal stock levels cannot be held due to demand uncertainties, order volumes and seasonal peaks (Tarn, Razi, Wen, & Perez Jr, 2003). How to deal with all those fast changes? How to automatically forecast the required storage capacity for an e-commerce retailer in order to timely prepare the distribution center? These issues will be addressed in this research.

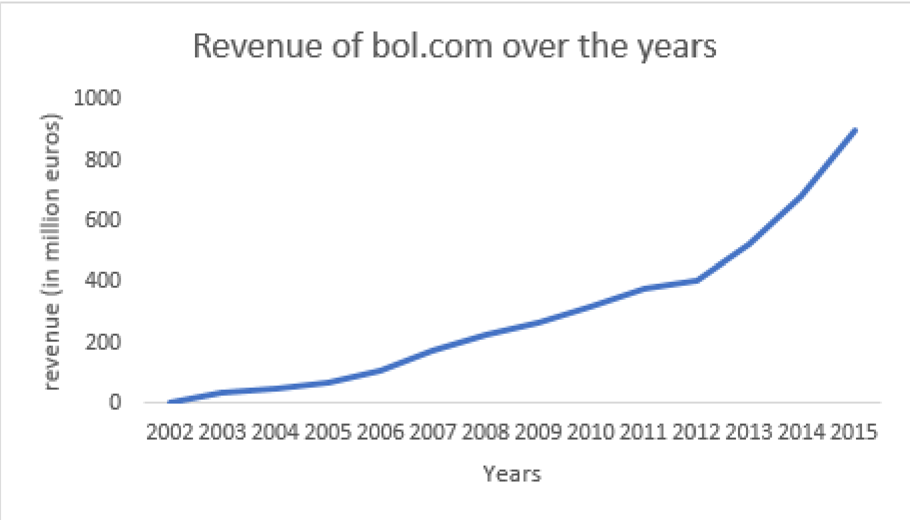


Figure 1: Revenues of bol.com from 2002 until 2015

In the following sections, an introduction to the company is provided, the research objectives are explained and the methodology used in this research is shown. Finally, the outline of the thesis is provided.

1.1 Introduction to bol.com

As mentioned before, this research is conducted at bol.com. In this section, a short introduction to the company is given. The German media group Bertelsmann AG (Bertelsmann on-line) established bol.com in 1999 as an online bookstore. In 2012, bol.com has been taken over by Ahold. The company developed into an online store in the Netherlands and Belgium. Three main activities are provided via their website. Firstly, their own assortment is sold online to customers (B2C). Secondly, the provision of second-hand products by individuals (C2C) and lastly, the concept of ‘Plaza’ (B2C). Brand owners and other retailers, can provide and sell their products through this platform. They can choose to either regulate the further handling of the orders themselves or use the service provided by bol.com, called LvB (Logistics via bol.com). LvB service facilitates packaging of ordered products, sending, reverse logistics and customer service. Docdata, the warehouse in scope, is divided into several areas where stock keeping units (SKU's) from one product category are stored per SKU together in an area. The different types of distribution of inbound and the division per product category are shown respectively in Figures 19 and 20 of Appendix A. This research is conducted at the department of Supply chain Management, which is responsible for the operational and commercial activities to go as efficient and effective as possible, in collaboration with the department of Sales & Operations Planning. In Appendix B, the organization chart regarding the scope of this project is shown. Based on this information, the research objective can be clearly defined.

The main processes of bol.com, with regard to the processes in scope, are shown in Figure 2. Veerweg is the location of Docdata, the distribution center in scope.

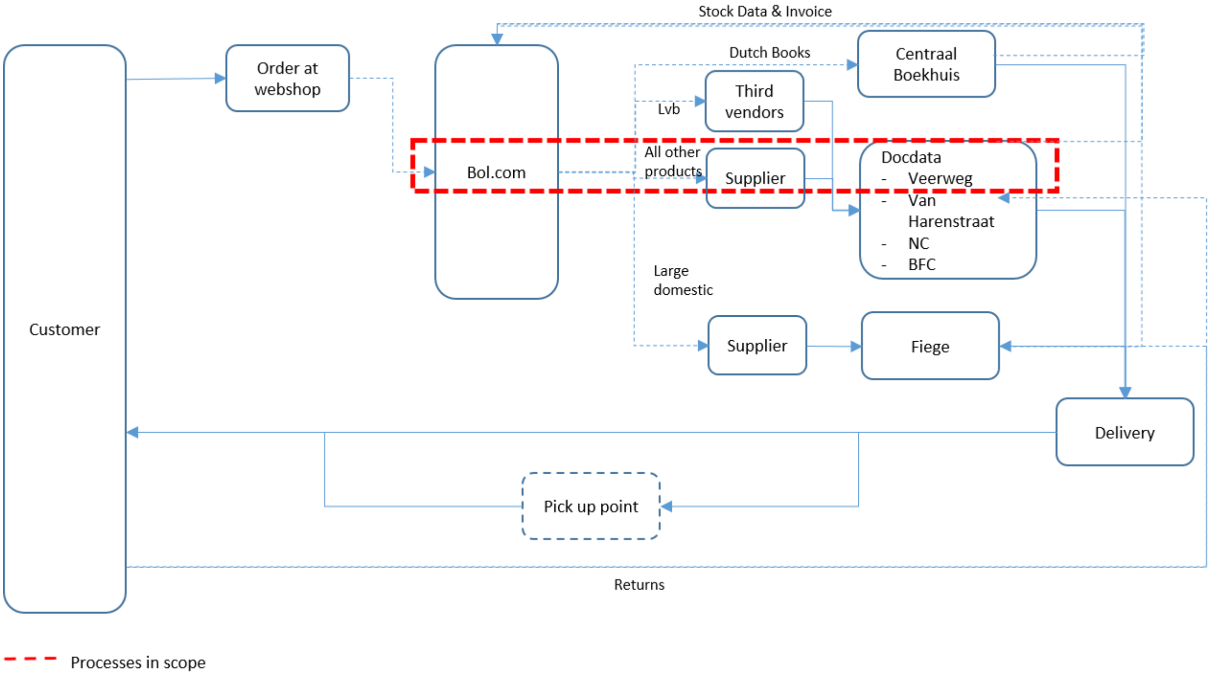


Figure 2: Main processes of bol.com

1.2 Research objective

The research objective is separated in two parts. Firstly, the research objective according to what is to be found in existing literature is described. Secondly, the research objective provided by stakeholders of bol.com is explained.

Research objective regarding existing literature

According to Tarn, Razi, Wen, & Perez Jr (2003), E-commerce retailing differs from traditional retailing. Fulfilling orders placed through the internet is different from fulfilling orders placed via other channels. Reasons such as a larger customer base, more order transactions and smaller order sizes in comparison with traditional retailing are mentioned. Moreover, fulfilling the demand of individual customers is more time consuming than fulfilling the same order for just one business. In addition to that, a higher probability of significant fluctuation in customer demand is present. Seasonality applies more frequently and is less predictable. With e-commerce warehousing, minimal stock levels cannot be held due to demand uncertainties, order volumes and seasonal peaks. However, articles about how to deal with these issues and to determine the capacity required in order to store all products, is not frequently found in literature. Therefore, this research will contribute to the literature as it focusses on e-commerce retailing with regard to reorder level determination and capacity planning.

Furthermore, the operations of a warehouse contribute to a significant portion of performance, costs and success of a Supply Chain (Koster de, Le-Duc, & Roodbergen, 2007). This industry has grown many folds with the growth of demand for several products. This resulted in problems in areas as slot design, layout, slotting, re-slotting and inventory management due to mass customization (Ackerman, 1990). In addition to that, there is a lack of literature about the topic of translating sales history into a capacity planning. Previous researchers made it clear that forecasting of demand is the key to a capacity plan. Literature about the process of the development of a method to translate sales into the required storage capacity is rarely found. Which methods could be used in order to determine the required storage capacity? How to deal with issues as fluctuating order patterns and forecasting the variance? Hence, how to translate the historical sales into the capacity required at a distribution center? These questions have not yet been addressed often in literature. In this way, the distribution center can be timely prepared for inbound, storage and outbound streams in order to avoid capacity problems?

Research objective regarding bol.com

According to stakeholders, bol.com was facing inbound, storage and outbound problems in 2015. An overflow location has even been used in order to replenish Docdata. This resulted in increasing costs and the risk to be unable to deliver on time. Commonly, storage capacity would have been the main issue. However, if an accurate capacity forecast was timely communicated, enough storage capacity would have been available. The mentioned problem is not a capacity problem. Due to an inefficient process of forecasting the required storage capacity, bol.com was not able to timely communicate an accurate capacity forecast. Therefore, the distribution center was not prepared for the large unexpected amounts of inbound, storage and outbound streams. When Docdata receives accurate forecasts some period in advance, enough storage capacity and employment capacity can be reserved. Facing the growth of bol.com, this problem could be even larger in coming years. It would be even more important to provide accurate forecasts that are timely communicated.

In order to give an indication of the relevance of the problem, a clear overview of the items and SKU'S stored at Docdata is shown in Figure 3. The peak period of 2015, where problems arose, a regular period of 2016 and a forecast of 2016's peak period are provided. Items are defined as the total number of products and SKU's as a distinct type of item for sale. Remarkable is that during last year's peak period, less SKU's are stored comparing to this year's non-peak period. This indicates that the assortment of bol.com is growing rapidly. Consequently, the amount of SKU's and items during this year's peak period will rapidly grow as well, which is noticeable from the Figure. Therefore, higher inbound and outbound streams will occur. This requires an accurate forecast of inbound, storage and outbound, which was lacking in 2015. Concluding, it is necessary to focus on these problems in order to fulfill all customer orders as efficient and effective as possible and achieve the target fill rates. Hence, this research is relevant and interesting to conduct.

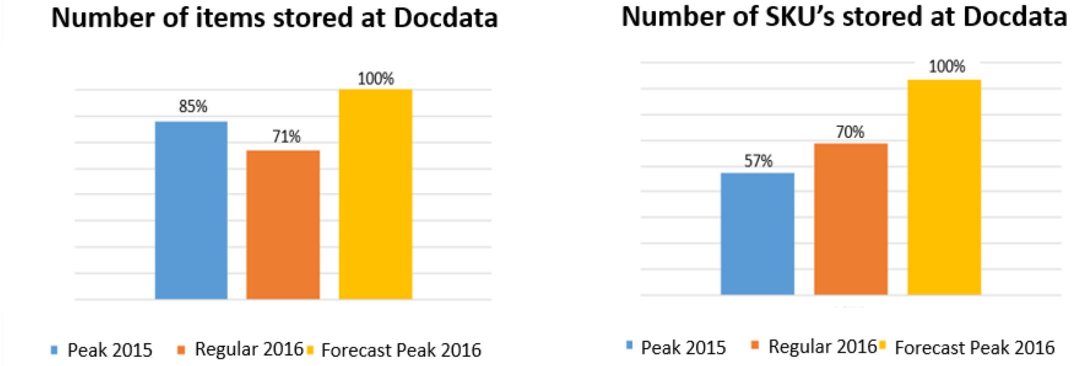


Figure 3: Comparison of items and SKU's stored at Docdata, peak versus regular weeks

In conclusion, the problem of bol.com is an inefficient process of forecasting the required storage capacity. Due to that, a lot of manual work is required and bol.com is not able to timely communicate an accurate forecast. Therefore, Docdata is not prepared well for the Christmas peak period for example. Furthermore, this research will offer three main contributions. Firstly, it contributes to the literature on e-commerce retailing with regard to capacity planning. Additionally, it will contribute to the literature with regard to the classification of SKU's regarding trend and seasonality. Moreover, different input scenarios to determine the reorder levels are evaluated. Finally, insights in the impact of different methods for determining the storage capacity on KPI's are provided. In section 3.4, these and other gaps in other gaps in literature are discussed.

1.3 Methodology

Quantitative modeling in operational research includes the need to develop explanatory and predictive theory regarding operational processes and operational management has become apparent (Bertrand, Will, & Fransoo, 2002). The methodology used for this research is described with use of the model of Mitroff, Betz, Pondy, & Sagasti (1974) and van Strien (1986), shown in Figure 4. The methodological model of Mitroff et al (1974) provides an approach in helping to answer the already formulated research question and corresponding sub-questions. When operational research emerged as a field, initial approaches as conceptualization, modeling, model solving and implementation were developed. Creating a model fit between the model developed and the observations for a reality is the primary concern in empirical research. The regulative cycle of van Strien (1986) is a combination of the analysis of the problem, a design process that results in an object design and the implementation and evaluation of the object design.

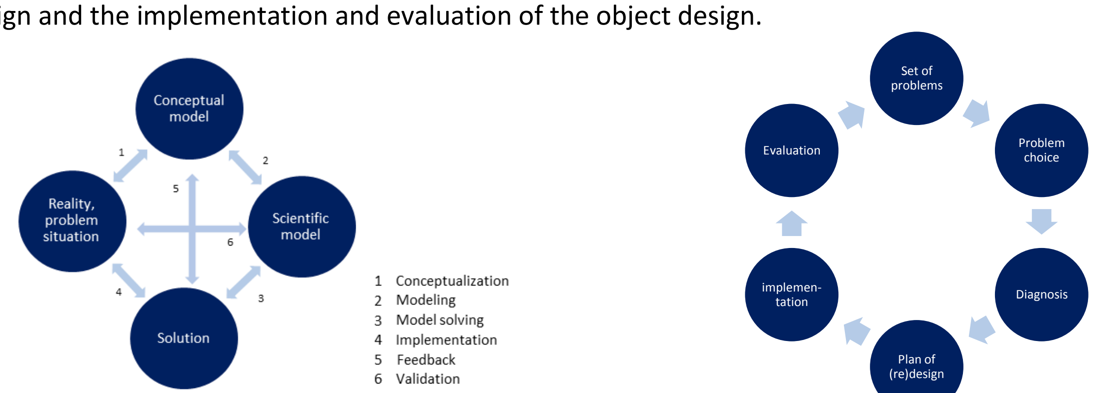


Figure 4: Research model by Mitroff et al. (1974) (left) and the regulative cycle of van Strien (1986) (right).

Microsoft Acces in combination with VBA programming is used in order to build the developed model and provide results. With use of the DoBr-tool, the reorder levels are determined (Broekmeulen & van Donselaar, 2015). In Appendix L, an introduction into the DoBr-tool is provided.

1.4 Research outline

The next Chapter provides a definition of the problem. An analysis of the current processes is conducted and research questions are provided. Moreover, the benchmark model is explained. Furthermore, the scope is provided and deliverables shown. A concise literature review on relevant topics is compiled in Chapter 3. Subsequently, a diagnosis of the current performance is provided in Chapter 4. Thereafter, in Chapter 5, the plan for redesign is described including an analysis on the available data and the conceptual model. In Chapter 6, the redesigned model is clarified with associated results. Different scenarios are applied to the designed model in order to conduct a sensitivity analysis in Chapter 7. Furthermore, conclusions are drawn and the most relevant recommendations provided in Chapter 8. Finally, a discussion is conducted, limitations of the research are listed and interesting topics for future research are shown in Chapter 9.

2. Project definition

In this Chapter, the current processes regarding forecasting are explained. The first two steps of the regulative cycle of Figure 4 are discussed: set of problems and problem choice. This corresponds to the step ‘reality, problem, situation’ of the model by Mitroff et al. (1974) as well. Based on this information and the issues addressed in Chapter 1, the research question and associated sub-questions are provided. Afterwards, the benchmark method is described. Furthermore, the scope of the research is described which provides additional insights in what can be expected from this research. Finally, the deliverables are given.

2.1 Analysis of current processes regarding forecasting and capacity planning

A clear process of forecasting the required storage capacity is currently lacking, as mentioned in Chapter 1. In order to clarify the current situation and provide the problems on which the research questions are based, the processes of forecasting are discussed. Furthermore, the concepts of healing and re-warehousing are provided in order to understand the issues of timely communicating the forecasts.

Process of forecasting

Three processes occur in order to develop the sales, stock and finally the storage capacity forecast in items and SKU’s. These obtained predictions are monitored, revised and adapted periodically. The sales forecast is defined as the expected amount of products sold per period. The stock forecast is defined as the amount of products expected to be stored at the distribution center. Lastly, the storage capacity forecast is defined as the expected cubic meters of storage space required to store the products. The composition of the three forecasts are discussed separately:

- Sales forecast: Slim4, the current inventory replenishment system, is used in order to determine the sales forecast on which decisions are made to order products at suppliers. The Finance department composes a sales forecast as well. This forecasts is used in order to determine if the sales goals are achieved.
- Stock forecast: parallel to the sales forecast process runs the stock forecast process. SCM develops a stock forecast per product category for items stored at Docdata, based on historical sales and experience. There is no one way of working to develop this stock forecast.
- Capacity forecast: S&OP, in collaboration with Docdata, translates the stock forecast into capacity planning with regard to inbound, outbound and storage in order to achieve reliable fulfillment and realize efficiency at Docdata.

In the case of over forecasting, too much capacity is reserved at Docdata in terms of storage and employment. Consequently, holding costs are increasing and unrequired employees are sent home with the risk that they will not come back. In the case of under forecasting, not enough storage capacity is reserved and not enough employees are scheduled. This could result in under capacity at the inbound, storage and outbound capacity. These issues indicate problems that occur when an insufficient process results in inaccurate forecasts or the distribution center does not receive an accurate forecast on time. The time period the distribution center needs to prepare for the inbound, storage and outbound streams depends on the time required for re-warehousing and healing, discussed in the next section.

Re-warehousing and healing

According to Kofler M. , Beham, Wagner, & Affenzeller (2015), due to modifications to the picking line, variations in the order mix, demand fluctuations and infrastructure changes the optimal storage assignment could become outdated. Because of seasonal fluctuations in the demand, item placement in practice need to be reviewed periodically and the warehouse should be re-organized in order to keep it operating efficiently. If the distribution center does not receive the forecasts on time, it cannot prepare timely for the inbound, storage and outbound streams and capacity problems occur. However, if changes in the forecasts are timely communicated, the distribution center can adapt timely. The smaller the changes, the shorter the time period required for preparing.

When only a few SKU's are moved per day iteratively over a longer period, it is referred to as healing. To ensure that the quality of a determined storage assignment does not deteriorate, which could be the case at companies with fluctuating demands and a strong seasonality, re-location of products might be a solution. It is often not necessary or feasible to implement a completely new slotting in each period. This is only worthwhile if the re-location efforts are lower than the efficiency gains (Kofler M. , Beham, Wagner, & Affenzeller, 2015). Re-ordering the existing layout of the allocation of SKU's is called re-warehousing and requires the movement of thousands of items. Therefore, for an extensive period, personnel and material handling equipment should be reserved. In practice, this is conducted infrequently (Kofler M. , Beham, Wagner, & Affenzeller, 2015).

Three measures are included to calculate a combined robustness rank of each move in order to determine how often re-warehousing or healing should occur, namely: stability, urgency and importance. Re-warehousing or healing is desirable when a SKU is stable, which means that not many targets are suggested during a period otherwise, the new place may rapidly become outdated again. Urgency is about how many periods a product has already been in the wrong place. Lastly, importance means that target storage that change significantly between two periods are preferred over target changes which are neighbor places for example (Pierre, Nieuwenhuyse van, Dominanta, & Dessel van, 2003).

Based on the re-warehousing and healing frequency of Docdata, the frequency of providing Docdata with an accurate forecast can be determined. However, it should be mentioned that the lower the frequency, the lower the accurateness of the forecast. A reason for this is that when Docdata needs the forecast six months ahead for example, a six months ahead forecast should be provided which is less accurate then a one month ahead forecast.

Current inventory replenishment system

Slim4 is the inventory replenishment system developed by Slimstock B.V., currently used by bol.com. It is a fill rate driven forecasting and inventory management system. The fill rate is defined as the fraction of demand that is immediately delivered from stock (van Donselaar & Broekmeulen, 2014). According to stakeholders, there are some disadvantages to the inventory replenishment system used. The first and most relevant one in this research is that monthly sales forecasts are estimated. Since the sales are fluctuating drastically during a month, a monthly forecast is not sufficient. For example, the division of the sales over the last 4 weeks of 2015 of a product category are shown in Table 1. The difference between two consecutive weeks, week 153 and 154, is indicating that the sales are fluctuating extremely. Furthermore, Docdata requires a weekly storage forecast in order to prepare the distribution center for all incoming and outgoing products. Moreover, Slim4 uses exponential smoothing methods with a fixed alpha for all SKU's to estimate the sales forecast, which might not be sufficient for all SKU's. To correct these issues and to provide a sufficient sales forecast, employees are handling these issues by conducting a lot of manual work every day. Furthermore, Docdata (and the other distribution centers) receive many shipments of small quantities. This could result in smaller inventory levels, however; it also results in higher handling costs. Concluding, weekly forecasts are sufficient and desirable.

Week nr.	% sales per week
153	36%
154	19%
155	29%
156	17%

Table 1: Division of sales over the last month of 2015

Storage policy

Generally, three different methods exists in literature to assign SKU's to storage locations in picking areas. The first one is the dedicated storage policy, which assigns a unique area to each SKU. This means that other SKU's cannot be located at that specific location. Secondly, the SKU's can be assigned randomly which means that incoming items can be stored anywhere in the entire picking area. This includes the sharing of areas over multiple SKU's. Lastly, the SKU's can be assigned to a location following a class-based assignment policy. A space in the picking area is reserved which can be shared by a set of predetermined product groups. Therefore, SKU's are restricted to a specific area. However, within this area, the location of SKU's can change over time and are randomly allocated. (Trevino, Wutthisirisart, Noble, & Chang, 2009; Larson, March, & Kusiak, 1997). With a dedicated storage policy, more storage space and less material handling costs are required. On the contrary, with a randomized storage policy, the material handling costs are greater and less storage space is required. The class-based storage policy is a compromise between those two (Larson, March, & Kusiak, 1997).

For example, Amazon is using the randomized storage policy, including a limit to the amount of different SKU's in an aisle. Bol.com makes use of the dedicated storage policy in combination with family grouping where similar products are located in the same area of the warehouse (Koster de, Le-Duc, & Roodbergen, 2007).

Inventory policy

The (R,s,nQ) inventory policy applies at bol.com, where the review period (R) is fixed and an order is placed when the inventory position is less than the reorder level (s). The ordering amount is equal to a plurality of the case pack size (Q) (van Donselaar & Broekmeulen, 2013). When Q is equal to 1, the (R,S) inventory policy applies. The review period and lead-time (L) are fixed, depending on the specific supplier.

Inventory control systems are usually based on the inventory position instead of on the inventory on hand (van Donselaar & Broekmeulen, 2014). The inventory position is defined as the sum of the inventory on hand added to the not yet delivered items minus backorders. The inventory on hand is defined as the actual inventory stored. An important remark in this research is that demand that is not met is lost. When the inventory on hand is zero, the offer on the website automatically changes into an offer of a Plaza partner if available. In this way, the sale is substituted to a Plaza partner. When no other offer is available, the sale is lost. Therefore, with regard to the inventory located at Docdata, back ordering does not apply. For that reason and because a dedicated storage policy applies, the storage space that should be reserved at Docdata is equal to the storage capacity required for the maximal inventory on hand. Therefore, the maximum inventory on hand is the concept in scope. Therefore, for both inventory policies, the maximum inventory on hand is equal: $Max(IOH) = s_i + Q - 1$. Where $s_i = S_i$ for the (R,S) inventory policy.

Key performance indicators

In this section, the key performance indicators (KPI's) used at bol.com are defined. Because backordering does not apply, the unobserved demand is hard to identify. Therefore, the following KPI's are used:

- Availability. This availability is equal to the discrete ready rate, which is defined as the probability that inventory on hand $(IOH) > 0$ just before a potential delivery moment. It is measured at specific discrete moments in time (van Donselaar & Broekmeulen, 2014). Bol.com is measuring the rates once per week on a fixed day.
- Percentage of obsolescence. This concept is discussed in section 4.3.4. It is defined as the percentage of stock that is classified as stock that is expected to be stored at the distribution center for more than 30 days.
- Sales from stock. One of the objectives is to deliver as many items as possible within one day to the customer. This is only possible with certainty when the items are delivered from bol.com's own stock. Therefore, as many items as possible should be delivered from one of the distribution centers instead from cross-dock.

The difference between the discrete ready rate and the ready rate, which is defined as the fraction of time the inventory on hand is positive, is that the ready rate continuously measures the product availability where the discrete ready rate measures it at fixed times (van Donselaar & Broekmeulen, 2014). Notice that it is remarkable that bol.com is using the discrete ready rate to measure their performance instead of fill rate, which is commonly the preferred method (van Donselaar & Broekmeulen, 2014).

One would not expect that the discrete ready rate is the most appropriate KPI to be totally customer oriented. This measurement does not estimate the service level towards the customer. This is due to the just mentioned issue of lacking to notice the unobserved demand. When an historical sales observation of a SKU at a certain week is zero, it indicates no sales during that particular time period. Causes for this to happen are when there is simply no demand or when an Out Of Stock (OOS) occurred. When the latter is the case, the demand of the SKU that is OOS is unobserved and lost. Using only the sales history can therefore lead to an underestimated mean and variance. Especially for low demand products with a higher frequency of observations with zero sales, this may result in poor performance when these are not taken into consideration. However, Tan & Karabati (2004) showed that when lost sales and demand are unobserved, the desired fill rate could still be achieved. Because of these reasons, the KPI used in this research in order to determine the most appropriate method is the fill rate.

Classes

SKU's are categorized according to the A, B or C – classification, which is based on the Pareto-analysis. This classification method is widely used to streamline the management of inventories that consist of large numbers of SKU's (Teunter, Babai, & Syntetos, 2010). Classification of SKU's is necessary in order to generalize the chosen methods. Looking at the first 3 months of the sales forecast, the products that will be responsible for the first 50% of the actual sales are classified as A items. The following 30% is classified as B-items and the last 20% as C-items. The targets for the fill rate and discrete ready rates also depend on the ABC-category. In Table 2, the target levels per class are shown.

ABC	Target Fill rate	Target Discrete ready rate
A	96,0%	94,5%
B	94,5%	94,5%
C	85,0%	89,5%

Table 2: Target fill rates and target discrete ready rate per class according to ABC-classification

2.2 Research question and sub questions.

Based on the research objectives mentioned in Chapter 1 and on the analysis of the current processes described in this chapter, the research question is formulated as follows:

What is the required storage capacity at Docdata in order to achieve the desired fill rates?

The objective proposed by bol.com is to acquire the forecast of storage capacity and when required, the sales and stock forecast as well. Some possible methods to estimate the reorder levels are to directly forecast the reorder levels, to set the reorder level equal to the demand of a certain cover period, equal to a part of the case pack size or equal to a part of the reorder quantity. However, to consider demand variability, lead times and the trend and seasonality components, it would be more appropriate to take these concepts into account. Therefore, multiple steps are taken in order to determine the reorder levels. Moreover, the results can be more easily compared to the current inventory replenishment system. In this way, a more flexible model that gains more insight into the impact of the characteristics of the SKU's and concepts as seasonality, trend and the composition of reorder levels can be created.

A trade-off between required capacity in cubic meters and units of lost sales have to be made. Lost sales are indicating the impact of lower achieved fill rates. Higher fill rates probably comes with higher required storage capacity. It is not desirable to store too much or too little, since additional costs as costs of lost sales and holding costs are involved. With the recommendations provided by answering this research question, a process for determining the weekly forecast of storage capacity can be developed. Using this, bol.com could communicate the forecasts timely and Docdata will be better prepared during the year with regard to storage capacity, inbound streams and outbound streams. The formulated research question is translated into the following research assignment: *Develop a model that translates historical sales data into a forecast of the required minimal storage capacity based on the required reorder levels, while achieving the target fill rates*. Therefore, minimizing the required reorder levels per SKU to minimize the required capacity is the main focus of this research.

To support this research question, the following sub-questions need to be answered:

1. What is the performance of the current situation regarding the forecasting of sales, stock and capacity?

Even though a clear process to translate historical data into a capacity plan is not present, it still needs to be clear how those weekly forecasts are established in order to fully understand the strengths and weaknesses of these processes. How well is the process currently performing, what are the achieved KPI's and current degree of obsolescence and how accurate were the sales and the stock forecast?

2. How can the SKU's be classified regarding trend and seasonality?

The developed model should deal with a large amount of SKU's. According to Gardner Jr & McKenzie (1988) and Hyndman, Koehler, Snyder, & Grose (2002), a classification should be made in order to deal with these issues and to determine the most suitable forecasting model per SKU. Which SKU's are heavily affected by seasonality or trend for example and therefore require different approaches?

3. How can the sales forecast per SKU per week be determined?

In order to determine the required stock levels to deal with the demand, forecasting is an appropriate method. The aim of forecasting is to reduce the uncertainty that confounds decisions in the future (Hyndman & Athanasopoulos, 2014) and future activity can better be predicted (Tersine, 1994). Currently, bol.com uses Slim4 to forecast the sales in order to determine how much they should order. A disadvantage of Slim4 is that it forecasts in months and uses exponential smoothing with fixed parameters for all SKU's. Bol.com wants to forecast their sales and capacity at least in weeks. Therefore, in this research, an aggregation level of one week will be used which will make the forecast more adequate. Based on the classification methods, the most appropriate forecasting method will be used to provide a weekly sales forecast per SKU.

4. How to determine the associated inventory levels per SKU per week?

As mentioned before, the inventory for which capacity should be reserved are equal to the storage capacity required for the maximum inventory on hand due to the dedicated storage policy. The maximum inventory on hand is derived from the reorder levels. Reorder levels are consisting of cycle and safety stock. SKU's with higher means should have relatively less safety stocks in comparison with SKU's with lower means. Based on the forecasted mean, forecasted variance and the specifications per supplier (such as lead time and case pack size) the desired reorder levels per SKU can be calculated taking the target fill rates into account. However, several issues are addressed as variance for example. How to forecast not only the mean but also the variance properly?

5. What is the required storage capacity per product category per week at Docdata?

The determined maximum inventory on hand could be translated in the required storage capacity in order to store all the forecasted products. Because bol.com does not have insights in the size of the storage locations and the required storage locations per SKU, only the volume of the SKU is taken into account. Docdata is storing the products per product category. Therefore, how much storage capacity per product category is required?

Concluding, 5 steps are taken in order to acquire the required storage capacity, shown in Figure 5.

Figure 5: Followed steps in order to answer the research question

Current problems & challenges

Based on the processes described above the following problems arose: the processes of forecasting the sales, stock and determining the required capacity at Docdata are currently different processes. Furthermore, Slim4 is using fixed input variables and delivering monthly sales forecasts, which is causing inaccurate weekly sales forecasts and therefore a lot of manual work. Furthermore, the developed model has to deal with the following challenging characteristics:

- Big data. Using such a large dataset may not only be time-consuming, many different issues should be handled.
- Different characteristics of SKU's. Every SKU has different characteristics, as trend and seasonality but also the dimensions for storage. Therefore, different approaches apply.
- Dynamic environment. E-commerce retailing has to adapt quickly to changes internally or externally. This also includes dynamic stock.
- The variance of the time series. Because the demand is non-stationary, not only the mean has to be forecasted but the variance as well.
- Historical sales data instead of historical demand data. Because lost sales apply, historical demand is not easily derived.
- Increasing amounts of SKU's and items.

2.3 Benchmark method

Multiple models are proposed and evaluated which are discussed in Chapter 5. In order to decide which model is best performing, a benchmark model is used. The order-up-policy proposed by Eaves & Kingsman (2004) is a simple and less time-consuming method that determines the minimal reorder levels that would have resulted in the target fill rate. This means that the results from this benchmark model are equal to the minimal number of items per SKU that should have been stored in order to achieve at least the target fill rate. Therefore, it is not reasonable to assume that one of the proposed methods can provide the same results. However, it is an indicator for comparison and to determine which method fits best. The results of the benchmark method are shown in Table 3. The mean and variances differ over time. Therefore, in order to compare the benchmark model with the other methods, the benchmark method is measured at a fixed moment at the end of the training period, which will be explained in section 5.1. The results of the other methods will be estimated at the same moment in time. The capacity is estimated as will be explained in section 5.2.5. The lost sales are defined as the fraction of demand that is not immediately delivered from stock, 1 – fill rate. Assuming that the reorder level of a SKU is at least one, the sum of the max inventory on hand for the subset is 27525 items and the total required capacity for the subset equal to 402.70 m3. The average amount of lost sales is equal to 741.16. In Appendix C, the results are shown per product category.

ABC	Maximum inventory on hand	Min fill rate	Avg fill rate	Avg discrete ready rate	Capacity in m3	Lost sales
A	8641	0.96	0.97	0.77	110.99	391.82
B	10093	0.95	0.99	0.90	170.37	167.04
C	8791	0.85	0.98	0.90	121.34	182.30
Totaal	**27525**		**0.98**	**0.86**	**402.70**	**741.16**

Table 3: Results benchmark model

2.4 Scope

As mentioned before, Docdata is the warehouse in scope. This distribution center fulfills 70% of the daily sales and at bol.com. Furthermore, regarding the capacity of Docdata, only stocked products of bol.com's own assortment are taken into account. The assortment of LvB partners is located at Docdata as well. However, because this concept is rather new, not enough historical data is available. Therefore, this concept will not be taken into account when determining the storage capacity. Furthermore, some product categories are stored at CB or Fiege. Those categories are Dutch books, International books and Domestic Appliances. Some product categories are out of scope as well because they are new to bol.com or new to Docdata and no historical sales are available. The composition of some product categories has changed as well. Beauty & Care, for example, is now divided into Beauty & Care and Health & Intimacy. In this research, the aggregate is used. However, this model could be used for the separate product categories in the future.

Furthermore, because capacity is not the issue in this research, the redesigned model assumes endless capacity. When the ordering of products at suppliers takes place is out of scope as well as the delivery of products. Concluding, the actual lead-time and review period are out of scope since no actual performance is available. Since routing is not considered in this research, the storage location within a bulk location, forward location or fast-mover area make no difference.

Additionally, promotions are out of scope as well as returns. Moreover, the incremental order quantity (IOQ), from now on denoted as Q, is known and taken into account. The minimal order quantity (MOQ) is equal to Q. Minimal order values depending on the particular supplier are out of scope. The characteristics of outdating are not taken into account. Products with an expiring date, as medicines and food for pets, are treated as unperishable.

Finally, the main focus is on SKU's that require most storage capacity. Those SKU's have the largest amount of sales and are probably most profitable. This research is restricted to the accessible data. Therefore, only historical sales data will be used.

2.5 Deliverables

The following research methods are used in this research to gather data and relevant information. Academic literature is consulted and cited when used. The bibliography of the used literature is shown in Chapter 'Bibliography'. Furthermore, interviews are held with different stakeholders of bol.com. Moreover, data is gathered in order to form and to support the conclusions to the research questions.

At the end of this research, a model is developed that determines the required storage capacity at Docdata in order to better plan capacity during the year. The following deliverables are formulated:

1. Literature review on the process of translating historical sales data into a storage capacity plan.
2. An analysis and a method to support the decision-making process
3. Analysis of the link between the variance of the time series and the required storage capacity.
4. Dynamic plan of storage capacity. This includes which products in which amount should be located at Docdata each period considering the target fill rate.

2.6 Summary of project definition

Firstly, the current three different forecasting processes that should be combined into one automatic process are discussed and the concepts of healing and re-warehousing explained. Furthermore, an aggregation level of one week is used in order to cope with fluctuations in demand patterns. Slim4 is currently providing monthly sales forecasts and using fixed parameters. Moreover, the discrete ready rate is used at bol.com as key performance indicator instead of the more commonly used fill rate. This is due to lost sales in the case of out of stocks, which results in unobserved demand. However, the fill rate is a more accurate criterion to measure the service towards the customer. Therefore, this is used as the service level constraint in this research. These issues resulted in the proposed research questions, the steps that should be taken in order to answer them and the deliverables. In order to determine the required storage capacity the maximum inventory on hand should be determined because of a dedicated storage policy. The maximum inventory on hand is derived from the reorder levels and since bol.com can manage the reorder levels, this is the decision variable in this research.

3. Literature review

The gap between existing knowledge and analyzed potential solution directions should be bridged with the existence of scientific research in the academic field. Existing literature on the topics in scope is provided in this Chapter, followed by the gaps found in literature.

3.1 Forecasting model

Forecasting could play a very important role in many different areas of a company. Therefore, it should be an integral part of the decision-making process. The aim of forecasting is to reduce the uncertainty that confounds decisions in the future. A short-term forecast is required for the scheduling of transportation, production and personnel (Hyndman & Athanasopoulos, 2014). There are two major types of forecasting, namely qualitative forecasting and quantitative forecasting. Qualitative forecasting relies on the judgmental and expert opinion and quantitative forecasting is based on the use of statistical methods. When historical data of demand is available, quantitative models, which use historical sales, are better methods to predict future activity (Tersine, 1994). This model can be applied when numerical information of the past is available and it is reasonable to assume that patterns of the past will continue in the future (Hyndman & Athanasopoulos, 2014). Mainly, there are two types of quantitative forecasting, causal forecasting and time series. Time series, which uses the historical past and prior experience to predict future attributions, is the type that is used in this research. Three components could be interacting. The first one is the level component, also known as raw-data, which is defined as the central tendency of the time series at a given time. The second component is the trend. This concept includes the continuing pattern that exhibits either a decline of an incline in the growth rate. The last component is the seasonal factors. These factors correspond to fluctuations, which are repeatedly returning every season (Anusha, Alok, & Shaik, 2014).

Methods used in this research are univariate methods based on time series techniques. The exponential smoothing techniques are examples of such techniques (Gardner, 2006). Other techniques are the ARIMA models for example. De Gooijer & Hyndman (2006) showed in their research that simple exponential smoothing methods are performing better than ARIMA models, especially when data is not normally distributed. A reason for this is that exponential smoothing models are not subject to model selection problems when the parameters are optimized. Exponential smoothing is developed in order to manage and compose forecasts on routine for many SKU's. Little computing time and data storage is necessary. Furthermore, this approach is responsive to changes. It is widely used for SKU's with similar properties, as is the case in sales forecasting and inventory control (Chatfield, Koehler, Ord, & Snyder, 2001). Another advantage is that the implementation of the exponential smoothing methods is relatively simple (Bartolomei & Sweet, 1989).

3.1.1 Classification models

In practice, model identification is often ignored and the same models are used to every time series. However, SKU's that are affected by seasonality, by trend or by both, require different approaches and different forecasting models (Gardner Jr & McKenzie, 1988). However, classical references as Brown (1983) for example, do not offer guidance on the identification of models. They only suggest to visually inspect the plot of the data. Relying on a subjective analysis could be misleading (Gardner Jr & McKenzie, 1988; Box & Jenkins, 1976). Therefore, classification models are appropriate. Gerrodette (1987) proposed a power analysis for detecting trend. Furthermore, Hyndman, Koehler, Snyder, & Grose (2002) compared several estimation methods for fitting exponential smoothing models on the data. Minimizing the MSE is found to result in forecasts that are more accurate. However, according to Chatfield (1995), empirical modelling is often prone to overfitting. Another classifiaction is suggested by Gardner Jr & McKenzie (1988), they suggested a simple and efficient robust procedure.

3.2 Determining reorder levels

The reorder levels consist of cycle stock and safety stock. Cycle stock is defined as stock during lead-time and review period and can be calculated by calculating the average demand during lead time and review period (Yamazaki, Shida, & Kanazawa, 2016). Safety stock is defined as the part of the inventory that absorbs differences between supply and demand. These differences could arise resulting from differences between average and actual demand, from changes in the timing of the purchases of customers, from machine breakdowns, fluctuations in demand, product defects, employee absenteeism, material shortages and more. Safety stocks are largely contributing to the required storage capacity at a distribution center (Bartholdi & Hackman, 2014). Van Donselaar & Broekmeulen, (2013) developed a method to determine the reorder levels taking lost sales into account. With the Forecast System method, the required level of safety stock depends on the variability of the demand forecast errors. This system requires 15% less safety stock compared to the Demand System method, where the level of safety stock depends on the variability of demand and still provides the same level of customer service (Zinn & Marmorstein, 1990). If the demand process experiences trend or seasonality in addition to random fluctuations, the demand process is non-stationary. When demand is non-stationary, higher forecasting errors apply and higher safety stock levels are required (Bhatnagar & Teo, 2009). Therefore, the higher the forecast accuracy, the lower the safety stock and the lower the required capacity. This indicates that the derived variance together with the derived mean are important factors for calculating the safety stocks. Different methods are providing how to forecast the variance properly. Nahmias & Olsen (2015) are proposing using the forecast error as input variable for determining the safety stocks. Taylor (1961) developed Taylor's Power Law (TPL), a method to determine the aggregated variance. Van Eenige (1996) proposed a method to estimate the probability mass function with limited support, which can be used in order to determine the safety stocks.

3.3 Gaps in literature

Taking the main contributions explained in Chapter 1 into account as well, the main contributions that this research offers are provided. Firstly, it contributes to the literature on e-commerce retailing with regard to capacity planning based on reorder levels. E-commerce retailing differs significantly from traditional retailing on several areas. This research clarifies these differences and provides recommendations to deal with e-commerce retailing with regard to capacity planning. Fluctuating order patterns, as trend and seasonality, makes forecasting and developing a capacity plan more difficult. Classification methods of SKUs regarding trend and seasonality are evaluated and discussed. Additionally, it contributes to the literature with regard to forecasting of variance. Furthermore, in this research the reorder levels are determined using several methods. This could result in lower reorder levels while the desired fill rates are still achieved. This is rarely found in literature. Finally and most importantly, this research contributes on the translation of historical demand into a capacity plan. From this point of view, structural ways of translating historical data into a storage capacity plan are lacking in existing literature.

4. Diagnosis of current performance

The next step of the regulative cycle of van Strien (1986) is the diagnosis. This Chapter encompasses a diagnosis on the current situation. The diagnosis consists of the current performance of the sales and stock forecasting processes, of the stored inventory and of the key performance indicators in order to develop a plan for redesign.

In this research, the current performance is the performance over 2015 because complete data is available and therefore used. The process of sales forecasting occurs parallel to the process of stock forecasting. Therefore, the current performance of the two forecasts is explained separately as well.

4.1 Forecast accuracy

Commonly, the objective of forecasting is to predict the demand accurately. In this research, the objective is to predict not only the mean accurately, but also the variance. The deviation between the forecast and the actual demand should be as small as possible. The MSE, a forecasting error measurement method, can be directly tied to variability of the forecast errors. Because each residual is squared, the larger the forecast error, the heavier the penalizing of the error. This could be beneficial in situations when the costs are increasing with an increasing error. This is mostly the case in production planning or inventory control. It is a proper error measure in situations when larger errors result in greater costs. This forecast error measurement is also used to forecast the expected sales, explained in section 5.3.2. The following formula shows how the MSE is calculated (Sanders, 1997):

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(\widehat{Y_t} - Y_t)^2$$

$$Where$$

$$\widehat{Y_t} = Forecast\ demand\ value\ at\ time\ t$$

$$Y_t = actual\ demand\ value\ at\ time\ t$$

$$n = number\ of\ time\ periods$$

RMSE is the square root of the MSE. Just like the MSE, the RMSE penalizes errors according to their magnitude. However, the RMSE is easier to understand and results can be compared because it is provided in the same unit as the data. Currently, bol.com is not measuring their sales forecast accuracy. Therefore, the accuracy of their sales forecast will be measured with use of the Mean Absolute Percentage Error (MAPE) and the Root Mean Square Error (RMSE) (Anusha, Alok, & Shaik, 2014). With the MAPE measurement, the average relative magnitude of forecast errors is provided in comparison to the actual forecast error. It is used due to its ease of understanding and the possibility to easily compare time series (Sanders, 1997).

$$MAPE = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{Y_t - \widehat{Y_t}}{Y_t}\right|$$

Because stock forecast data is only available from week 14 of 2014, the forecast accuracy is measured from week 14 until the end of the year. The forecast accuracy and RMSE of the different product categories for 2015 are shown in Table 20 of Appendix D. The forecast accuracy is equal to 100% - MAPE.

These forecast accuracy measurements, which are used on historical sales, are probably providing better results than a forecast accuracy measurement that takes the unobserved sales into account.

This is due to the concept of lost or unobserved sales, which is mentioned before and also discussed in the research of Tan & Karabati (2004). Furthermore, a lower forecast accuracy could be due to the experience of the Supply Chain Specialist and the historical data available per category. Figures of forecast accuracies per units are shown in Appendix D. Overall, it can be concluded that a protocol or method for developing a capacity plan is a relevant issue for bol.com. These processes are currently two separate processes and should be translated into one process.

4.2 Forecast accuracy vs volume

In this section, the current performance of the forecast accuracy versus the volume is explained. Not only the forecast accuracy is important, but the volume that corresponds to it is essential as well. A higher the volume of a certain product category, acquires a higher forecast accuracy in order to avoid waste or out of stocks. The MSE measurement does already consider this. The sales forecast accuracy versus the volume (left) and the stock forecast accuracy versus volume (right) for the unit Toys are shown in respectively Figure 10 and 11. The forecast accuracy is again defined as 100%-MAPE.

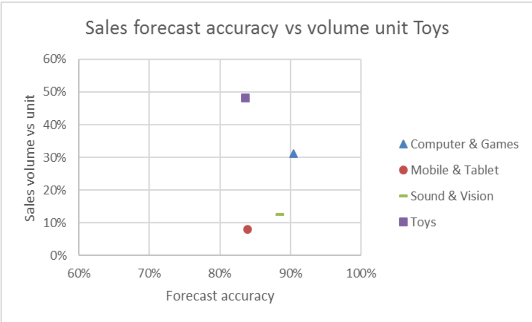


Figure 7: Sales forecast accuracies vs volume unit Toys

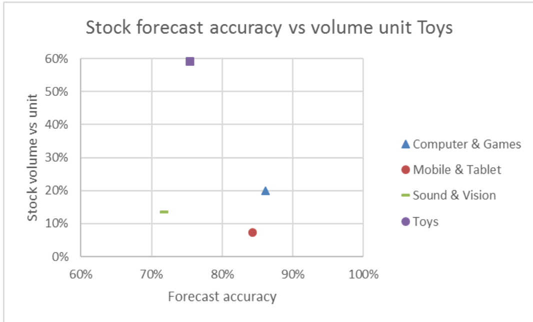


Figure 6: Stock forecast accuracies vs volume unit Toys

As can be seen from these Figures, product category Toys is the product category with the highest volume. However, compared to the other categories in the unit Toys, the lowest sales forecast accuracy. This could have a great influence on the sales forecast and capacity planning. Therefore, for those product categories, it is even more important to review the forecasting process and develop a model to determine the required storage capacity. In Appendix E, the other product categories are reviewed as well.

4.3 Under and over forecasting

The MAPE and other forecasting error measurement methods do not take over or under forecasting into account. When over forecasting of stock applies, the surplus will result in higher holding costs, higher throughput times, and employees will be send home with the risk they will not come back when they are needed. What is even worse is under forecasting of stock. This could result in more out of stocks could occur, not enough employees and lower fill rates and discrete ready rates. Bol.com is measuring their stock forecast accuracy in percentage deviation per product category per week. This is almost equivalent to the MAPE measurement, but the under and under forecasting is taken into account. The following formula is used:

$$FCacc_t = 100\% * (1 - \frac{(\widehat{Y_t} - Y_t)}{Y_t})$$

If $FCacc_{pc,t} > 0$, the forecast is higher than the actual stock, over forecasting occurs. When $FCacc_{pc,t} < 0$, the forecast is lower than the actuals, under forecasting occurs.

Bias of the current situation

In Appendix F, the over and under forecasting percentages for the stock are exposed for the different product categories for 2015 are shown. The MPE measurement is used, which measures the bias of the forecast (Chockalingam, 2010). The MSE already incorporates the bias and the RMSE is the standard deviation for an unbiased estimator. A forecasting bias could occur when a general tendency exists of consistent differences between actual demand and forecasted demand, also known as under- and over forecasting.

$$MPE = 100\% * \frac{\sum_1^t \widehat{Y_{pc,t}} - Y_{pc,t}}{\sum_1^t Y_{pc,t}}$$

The results of this estimation per product category of the sales forecast and stock forecast are shown in respectively Table 5 and 6. It should be mentioned that for the product category Baby, a structural over forecasting of sales applies. Jewelry, Watches & Accessories often over forecasts the sales even as the stock. Domestic Appliances and Beauty & Care are often under-forecasting. This could be due to promotions. In general, no structural over or under forecast occurs.

Sales forecast	
Product Category	Bias
Entertainment	-3,24%
Ereader	-6,87%
Cooking, Dining & Houseware	5,80%
Domestic Appliances	-5,46%
Home Furnishing	1,49%
Home Improvement & Gardening	7,18%
Baby	58,18%
Beauty & Care	-12,73%
Jewellery, watches & Accessories	20,23%
Pet	-7,59%
Sport & Leisure	3,71%
Computer & Games	-4,15%
Mobile & Tablet	-14,09%
Sound & Vision	-9,12%
Toys	-12,04%

Table 5: Bias sales forecast over 2015

Stock forecast	
Product Category	Bias
Entertainment	12,80%
Ereader	-16,03%
Cooking, Dining & Houseware	12,09%
Domestic Appliances	-25,76%
Home Furnishing	-3,05%
Home Improvement & Gardening	28,56%
Baby	2,02%
Beauty & Care	-25,76%
Jewellery, watches & Accessories	20,49%
Pet	21,16%
Sport & Leisure	-15,76%
Computer & Games	9,26%
Mobile & Tablet	-6,47%
Sound & Vision	-4,79%
Toys	8,61%

Table 4: Bias stock forecast over 2015

4.4 Obsolete stock

Overall, the discrete ready rates were achieved. However, on the other hand, the numbers of obsolete stock are interesting as well. It could be the case that the achieved discrete ready rates were quite high. However, this could result in high levels of obsolescence. To determine the amount of healthy stock in 2015, the percentages of obsolete stock have to be evaluated. All SKU's are subject to obsolescence. However, when demand is not obsolete, demand is healthy (Song & Zipkin, 1996). The numbers of obsolescence are based on the sales forecast in Slim4. The products of the current inventory on hand that are expected to be sold during the first 30 days is classified as green stock. The products that are expected to be sold within three months are classified as orange stock and if the products are not expected to be sold before that time, they are classified as red stock, meaning obsolete stock. For example, imagine that a SKU has 1000 items on stock. The sales forecast for the next couple of months is 100 items per month. 100 items will be classified as green stock, 200 items as orange stock and 700 items as red stock. The achieved discrete ready rates over 2015 with corresponding percentages of obsolescence are shown In Appendix G. It can be seen that high levels of obsolescence occurred. Items that are obsolete either go into the outlet, bought up by a buyer or are destroyed. However, different departments are involved when making these decisions. Therefore, it is not easy to remove obsolete stock and higher the average throughput times.

4.5 Summary of diagnosis

Concluding from the current performance, the forecast accuracies point out that there is still room for improvement. Furthermore, a higher volume does not necessarily indicate a higher forecast accuracy. Moreover, the percentages of bias indicated that no structural over or under forecasting applies. Finally, the target discrete ready rates of 2015 are achieved. However, high levels of obsolete stock are showing that unnecessary storage capacity was required.

5. Plan for redesign

This Chapter encompasses the plan for redesign in which the used time series data is analyzed and the conceptual model is provided. This includes the input variables, the KPI's, the mathematical model and the assumptions made. This all represents the modelling phase of the model by Mitroff et al. (1974). The methods for detecting seasonality and trend are explained and forecasting methods are clarified. Thereby, the methods for estimating the inventory levels are provided. Different methods that could be used are already mentioned in Chapter 3. Using these methods, the final required capacity can be determined.

5.1 Time series data

This section includes information about the gathering of the required data of historical sales. In order to use a representative dataset and to draw significant conclusions, weekly observations per SKU of the last three years are taken. Data from the first week of January 2013 until the last week of December 2015 is provided, 156 weeks in total. The data is split up into 2 periods, the training period and the holdout period. The training period is usually about 20% of the total sample, although the length depends on the sample size according to Hyndman (2014). Therefore, the first two years of the data (2013 and 2014) are used as the training period as shown in Figure 8. Using data of these two years, initial values are estimated, parameters are optimized and the KPI's per model are determined. Since it is a dynamic model and mean and variances change constantly, the models are evaluated at one moment in time. These moments are indicated with the red arrows in the Figure. At the end of the training period, the reorder levels are estimated. Looking back at a certain time period, depending on the target fill rates, the achieved service levels are estimated and evaluated. When the target fill rate is equal to 96%, it means that for 1 of the 25 ordered products, it is allowed to not be able to fulfil the demand. Therefore, the model is evaluated looking back at the last 25 ordered products from that moment. The models can be compared to the benchmark method, clarified in Section 2.3. Based on this information, the most appropriate model can be selected.

The third year (2015) is used as the holdout period. The sales of the holdout period are forecasted using the selected model. The determined forecasting methods and optimized parameters determined at the end of the training period are used, which is discussed in section 5.3.2. At the end of the holdout period, the reorder levels are measured and evaluated looking back at the certain time period, depending on the target fill rates as mentioned before. The acquired required storage capacity and realized lost sales are used in order to measure how likely the selected model is to forecast well on new data.

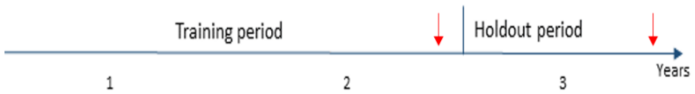


Figure 8: Time series divided into training and holdout period

The data provided consists of SKU's of different product categories, shown in Table 6, ordered by unit.

Unit	Toys	Entertainment	Lifestyle	Home&Living
Product categories	Computer & Games	Entertainment	Baby	Cooking, Dining & Houseware
	Mobile & Tablets	E-reader	Beauty & Care	Domestic Appliances
	Sound & Vision		Jewelry, Watches & Accessories	Home Furnishing
	Toys		Pet	Home Improvement & Gardening
			Sport & Leisure	

Table 6: Product categories per unit

According to stakeholders, the product category Toys causes the most challenges regarding storage capacity, especially during the Christmas peak period. During this period, the largest part of the inventory located at Docdata belongs to Toys. The unit Entertainment has the most SKU's per product category where the Lifestyle unit has more but smaller product categories. In total, the unit Lifestyle has the most SKU's as can be seen in Table 7. The forecasting of and dealing with promotions is a characteristic of the product category Beauty & care. Additionally, the product category Baby is also different from the others since it includes fashion products like baby clothes. Moreover, Jewelry, Watches & Accessories has the most slow moving SKU's. It should be noticed that the unit Toys has the least SKU's but has sold the most items in 2015.

	Entertainment	Home&Living	Toys	Lifestyle	Total
% SKU dataset	27%	21%	19%	33%	100%
% Items sold 2015	23%	17%	37%	23%	100%

Table 7: Percentage SKU's and items sold per unit of 2015

Selected subset

For measuring the KPI's, daily historical data observations are required. These are not obtainable for all SKU's of the dataset. Therefore, a subset is taken in order to evaluate and to discuss all methods. As will be discussed in section 5.3.3, the method of Van Eenige (1996) only provides results when the maximum demand is larger than 1. Therefore, SKU's where the maximum demand of the training period is larger than 1 are selected. The sample consists of 6880 SKU's that are randomly chosen from every unit. According to Green (1991), this sample size is definitely more than enough because a minimum of 50+8k is required. The variable K is equal to the predictors of the model. In conclusion, for the evaluation of the proposed models, the subset is taken to make a fair comparison. However, the evaluation of the different classification methods are based on all SKU's (410.020 SKU's), in order to have more support.

5.1.1 Analysis on sales data

An analysis should be done on the data beforehand in order to select appropriate methods for the model. It is shown that the time series has a discrete distribution and the used outlier detection method is provided.

Probability distribution

Demand is stochastic when the demand is unknown until it is received. The stochastic nature may results in stock-outs and lost sales or backorders are possible (Sobel & Zhang, 2001). However, in this research only the actual sales are known. This means that mixtures of the Geometric, the Binomial, the negative Binomial and the Poisson distributions could apply to historical sales. To determine the distribution of the data, the mean-variance relationship could be used (Molenberghs, Verbeke, & Demetrio, 2007). When the variance is equal to the variance ($\mu = \sigma$), the Poisson distribution applies. When the data is over-dispersed, $\mu < \sigma$, the data is negative binomial distributed and when the data is under dispersed, $\mu > \sigma$, the data has a binomial distribution (Molenberghs, Verbeke, & Demetrio, 2007). Adan, van Eenige, & Resing (1995) developed an effective and simple procedure based on this relationship to fit a theoretical discrete probability distribution on the mean and variance where the four discrete distributions are considered. Van Donselaar & Broekmeulen (2013) used this method as well. In Appendix H, the method and results are shown in more detail. As can be seen, some relationships between the mean and the variance does not exists in discrete distributions, marked by the shaded areas in Figure 25. The results of the proposed method show that the dataset assumes a combination of negative binomial and binomial distributions. However, using the method of Van Eenige (1996) where the maximum demand value is used as input value for determining probability mass function, the binomial distribution applies. This method is discussed in section 5.3.3.

Outlier detection

According to Ledolter (1989), outliers could have significant effects on the forecasts. In this research, only univariate outliers are taken into account. Prediction intervals could become severely misleading and they could inflate the estimated variance. Outliers do affect the autocorrelation structure, which is used in order to identify the presence of seasonality and trend. The estimated autocorrelation, extended autocorrelation functions and partial autocorrelation could be biased (Nare, Maposa, & Lesaoana, 2012). The impact of promotions is not accounted by univariate models as exponential smoothing methods (Trapero, Kourentzes, & Fildes, 2015). Therefore, it is necessary to substitute outliers before the seasonality and trend is identified. Promotions, defined as temporary reductions in price or marketing advertisements for specific products for a specific period, are out of scope and should be substituted as well. The selected method for detecting outliers is the base-times-lift method (Cooper, Baron, Levy, Swisher, & Gogos , 1999; Huang, Fildes, & Soopramanien, 2014). The results of the initial outlier detecting method are shown in Figure 9. As can be seen in this Figure, the most outliers are detected during the Christmas peak period. However, bol.com wants to include this period as well. Therefore, in accordance with the stakeholders of bol.com, the Christmas peak period is excluded from the outlier detection method. This may result in worse forecasting accuracies. On the other hand, the forecast model will react quicker to the demand pattern. The results for the product category Toys of the adapted model are shown in Figure 10. The detected outliers are replaced with the baseline sales. More details about the outlier detection methods are shown in Appendix I.

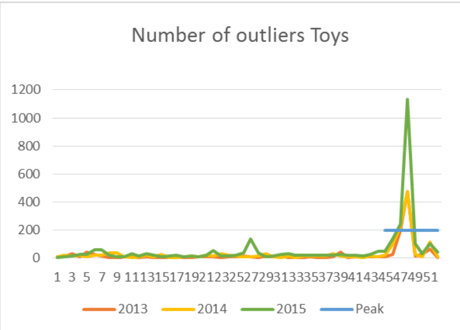


Figure 9: Outliers per week product category Toys - Initial method

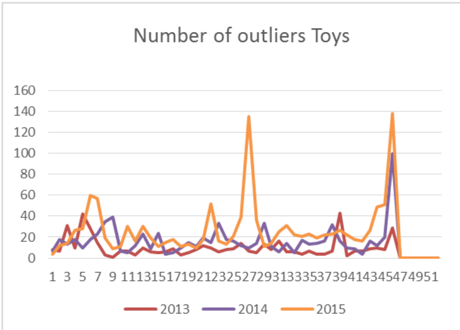


Figure 10: Outliers per week product category Toys - Adapted method

5.2 Conceptual design

The conceptual model of Mitroff, Betz, Pondy, & Sagasti (1974) is addressed in this section. According to Suh (1990), a design should start with functional requirements, design parameters and assumptions. Design is about what to achieve and how to achieve it. The functional requirements state what the design should be able to do or achieve. The design parameters represent characteristics of the design and the assumptions represent limitations and constraints of the design. In Table 8, the functional requirements with corresponding design parameters per main phase are provided.

Phase	Functional requirements	Design parameters
Forecasting	- Weekly sales forecast in items per SKU - High forecast accuracy	- Sufficient data available - Level component - Trend component - Seasonal component - Initial values
Determine Reorder levels	- Reorder levels in items per product category per week - Target fill rates achieved	- Variance - Mean - Maximal demand observation - Case pack size
Determining Capacity	- Storage capacity plan in m3 per product category per week	- Volume of SKU

Table 8: Functional requirements and design parameters per phase

As mentioned before, the sales are fluctuating significantly during a month. Therefore, the aggregation level in this research is one week. Different methods are evaluated in order to forecast the mean and variance properly and to determine the reorder levels. These and the design parameters are discussed in the next sections.

5.2.1 Dependent & independent variables

The choice of the independent and dependent variables have an important effect on the model and the depth of the model. The dependent variable is *the sum of the reorder levels.* This amount is estimated based on the forecasted mean and variance. The objective is to minimize the required reorder levels in order to minimize the required storage capacity. The fixed and variable independent variables are shown in Table 9. These parameters are all discussed in this Chapter.

Parameter	Description	Value(s)
Fixed		
W	Number of weeks in a year	52
$\boldsymbol{T_{tot}}$	Length historical data	3*W
$\boldsymbol{T_T}$	Length training period	2*W
$\boldsymbol{T_H}$	Length Holdout period	1*W
BaseMA	Length of Base Moving Average.	6
LF	Lift Factor	3
MO	Minimal Baseline sales for detecting outlier	10
M	m-step ahead forecast	1
$\boldsymbol{L_{0,SES}}$	Initial value level component of Simple exponential smoothing	$\bar{D} = \frac{\sum_{i=1}^{BaseMA} D_i}{BaseMA}$
$\boldsymbol{L_{0,H}}$	Initial value level component of Holt's exponential smoothing	$OLS_{Level,BaseMA}$
$\boldsymbol{L_{0,HW}}$	Initial value level component of Holt Winter's exponential smoothing	$OLS_{Level,BaseMA}$
$\boldsymbol{T_{0,H}}$	Initial value trend component of Holt's exponential smoothing	$OLS_{Slope,BaseMA}$
$\boldsymbol{T_{0,HW}}$	Initial value trend component of Holt Winter's exponential smoothing	$OLS_{Slope,BaseMA}$
S	Number of seasons	52
$\boldsymbol{C_t}$	Seasonal indices for seasonal demand patterns	$C_t = \frac{(D_t + D_{t+S})}{2 * \bar{D}}$
L	Lead-time	1
R	Review period	1
Variable		
t	Peak week	0, 1
$\boldsymbol{ABC_i}$	ABC-classification	A,B,C
$\boldsymbol{P2_T}$	Target Fill rate	0.96, 0.945, 0.85
$\boldsymbol{P3^{discrete}{}_T}$	Target discrete ready rate	0.945, 0.895
$\boldsymbol{M_{GaMcK}}$	Model by GaMcK	1,2,3,4
$\boldsymbol{M_{MMM}}$	Model by MMM	1,2,3,4,5,6
$\boldsymbol{\alpha_i}$	Smoothing constant for level parameter	$0 < \alpha < 1$
$\boldsymbol{\beta_i}$	Smoothing constant for trend parameter	$0 < \beta < 1$
$\boldsymbol{\theta_i}$	Damped or autoregressive parameter	$0 < \theta < 1$
$\boldsymbol{Q_i}$	Case Pack size	$1 \leq Q \leq 720$
$\boldsymbol{V_i}$	Volume of SKU in square meters	$1.9 * 10^{-7}m^3 \leq V_i \leq 1.4m^3$
$\boldsymbol{\sigma_i^2}$	Variance	$\sigma_{PLFC}^2, \sigma_{FC}^2$

Table 9: Fixed and variable design parameters

A fixed amount of 52 seasonal periods (S) is considered, which corresponds to weekly time periods. This is equal to the number of weeks per year (L). For each period, the seasonal indices are determined if a SKU is classified as seasonal. Moreover, the total length of the historical data (T) is defined as 3 times L. The holdout period is taken as L, so 2 times L remains for the estimation of parameters for the forecasting method. The BaseMA is equal to 6 weeks and is used for estimating the initial values. These are discussed in section 5.3.2. The minimum required amount of history is 9 weeks in the training period and 1 week in the holdout period. In this way, the parameters can be optimized properly and only SKU's are forecasted that still are included in the assortment in the hold out. The minimum amount of weeks required to calculate the seasonal indices is equal to S.

5.2.2 Key performance indicators

The following key performance indicators are measured per proposed model. Based on these values, the possible scenarios are compared and the one that fits best is evaluated and discussed. The key performance indicators are:

- Achieved fill rate of the SKU, which can be translated into the lost sales per SKU
- Required storage capacity per product category in cubic meters

The discrete ready rate levels are also measured in order to compare the different service level measurement methods, the proposed service level (fill rate) and the service level bol.com is currently using (discrete ready rate).

5.2.3 Assumptions

In order to define the scope of the model more clearly and to provide the principles for design, the following assumptions have been made:

- Future values are linearly related to past observations. Quantitative models can be applied when numerical information of the past is available and it is reasonable to assume that patterns of the past will continue in the future. Future sales will follow the trend and/or seasonality or have a similar pattern of historical sales. (Hyndman & Athanasopoulos, 2014).
- The lead-time and review period are constant, deterministic and both equal to one. A reason for this is the assumption that items are always available. Therefore, items can always be delivered to Docdata in one day.
- Demand that is not met is not back ordered but is lost. When the inventory on hand is zero, the offer on the website automatically changes into an offer of a Plaza partner if available. In this way, the sale is substituted to a Plaza partner. When no other offer is available, the sale is lost. Therefore, with regard to the inventory located at Docdata, back ordering is not possible.
- Seasonal indices can be calculated using the same weeks of each year.
- The amount of products introduced to the assortment (phasing in) is equal to the products excluded from the assortment (phasing out) due to the life cycles of products. Divide the product life cycle into stages results in the need for different strategies regarding production planning and control, for example forecasting with trend models. The sales are assumed equally as well.
- The reorder levels are at least one for each SKU per period and therefore, the maximum inventory on hand is at least equal to one as well.
- A historical sales point of zero is defined as no demand and the unobserved sales are assumed equal to zero.
- Just-in-time strategy for the capacity planning.
- Capacity is infinite.

5.2.4 Mathematical model

The mathematical model of this research, taking the objective of the research with the restrictions into account, is shown below:

Objective:

$$C = \min\left\{\sum_{i}^{j} V_i * (s_i + Q_i - 1)\right\} \tag{1}$$

Subject to

$$\forall_i: \sum_{s} \boldsymbol{P2_{i,s} * X_{i,s} \geq P2^*} \tag{2}$$

$$\boldsymbol{\frac{1}{N}\sum_{s} P3^{discrete}{}_{i,s} * X_{i,s} \geq P3^{discrete^*}} \tag{3}$$

$$\forall_i: \sum_{s} \boldsymbol{X_{i,s} = 1} \tag{4}$$

The objective is to minimize the total storage capacity. This is calculated by taking the aggregated storage capacity required over SKU's. The storage capacity per SKU is determined by multiplying the volume with the maximum inventory on hand. The maximum inventory on hand is equal to $s_i + Q_i - 1$ (1). Note that for the (R,S) inventory policy, $s_i = S_i$. This is subject to the fill rate of the determined required reorder levels per SKU should be larger than or equal to the optimal fill rate (P2) (2). Even as the average discrete ready rate ($P3^{discrete}$) (3). The decision variable for estimating the reorder levels, $X_{i,s}$, is equal to 1 (4).

5.3 Proposed plan per phase

Completing the different steps in order to answer the research question, as described in section 2.2, various theories and methods are used as already shortly discussed in Chapter 3. Per phase, the different methods are shown, restrictions are given and important remarks are provided in order to fully understand the different steps, processes and the final model. Firstly, the methods of classifying the SKU's in order to determine the most appropriate forecasting method are provided. Thereafter, the corresponding forecasting are explained. Subsequently, the method of determining the reorder levels, including appropriate forecasting of the variance, is shown. Using these numbers, the maximum inventory on hand and finally the required storage capacity can be determined.

5.3.1 Classifying the SKU's

Commonly, the linear trend model of Holt is used for forecasting. However, for data sets consisting of many SKU's, is more efficient and accurate to identify different characteristics of SKU's. SKU's that are affected by seasonality, by trend or by both, require different approaches. Demand patterns including a linear or damped trend factor require more complex methods than SKU's with a constant level. Therefore, a classification should be made in order to determine the most appropriate forecasting method for each SKU. Trend is defined as an increase or a decrease during a longer time period. Seasonal factors are affecting the time series when the sales depend on the time of the year for example (Hyndman & Athanasopoulos, 2014). The classification methods discussed in Chapter 3 are explained in the following sections. As already mentioned, empirical modelling is often prone to overfitting according to Chatfield (1995). Therefore, multiple classification models are evaluated in this research.

Classification method of Gardner & McKenzie

Gardner Jr & McKenzie (1988) proposed an efficient and simple procedure to identify the most suitable exponential smoothing model regarding seasonality and trend. From now on, this method is abbreviated to GaMcK. GaMcK proposed 6 models: no seasonality with constant level, no seasonality with damped trend, no seasonality with linear trend, seasonality with constant level, seasonality with damped trend and seasonality with linear trend. The procedure of identification is based on comparing the variances of relevant differences between the data and the order of differencing yielding minimum variance. Three types of variances are calculated. The variance that corresponds to the variance of the original data, the variance that is differenced once and the variance that is differenced twice. If the first variance is the smallest one, the SKU is classified to a model with constant level. If the second mentioned variance is the smallest one, the SKU is classified to a damped trend model. If the third variance represents the smallest one, linear trend applies. Suggested advantages are that the models give a better fit to the data and that this method yields frequently more parsimonious models.

Classification method by minimal forecast error

Moreover, Hyndman, Koehler, Snyder, & Grose (2002) compared several estimation methods for fitting exponential smoothing models on the data. Minimizing the MSE is found to result in forecasts that are most accurate. Gardner Jr & McKenzie (1988) are mentioning this as well, the process of estimating parameters that provide the minimum MSE is a way to identify the model for the data. Additionally, according to Sanders (1997), in environments where large errors are costlier than small errors and when variability is important, like it is in safety stock calculations and inventory control, MSE can be an excellent error measure. Therefore, to evaluate the GaMcK classification method and to determine the most appropriate forecasting method, all forecasting methods shown in Table 13 are applied to all SKU's. The forecasting are discussed in more details in section 5.2.2. The forecasting method that results in the smallest MSE is selected. For each evaluated forecasting method, the parameters (α, β, θ) are optimized since it is dangerous to guess them (Gardner, 2006). From now on, this classification method is abbreviated to the Minimal MSE Method (MMM). For the SKU's classified by GaMcK to non-seasonal (models 1, 2 and 3), the MSE is calculated for the three forecasting methods without seasonal indices, the first three methods. For the SKU's classified as seasonal (model 4), the MSE is calculated for all forecasting methods shown in Table 13.

Model GaMcK	Series yielding a minimum variance	Selected model	Model MMM	Forecasting method
1	X_t	No seasonality, Constant level	1	Simple exponential smoothing
2	(1-B) X_t	No seasonality, Damped trend	2	Holt's exponential smoothing
3	$(1-B)^2 X_t$	No seasoanilty, Linear trend	3	Holt's exponential smoothing
4	$(1-B^p) X_t$	Seasonal	4	Holt Winters exponential smoothing constant trend
			5	Holt Winters exponential smoothing damped trend
			6	Holt Winters exponential smoothing linear trend

Table 10: Different possible forecasting models by GaMcK and MMM

Restrictions to classification models

When insufficient data is available, meaning that a SKU does not have historical observations in the holdout period or has less than 9 weeks of observations in the training period, the SKU is classified to model 0 and no classification method is applied. The objective of forecasting is to forecast the mean and variance as accurate as possible. Therefore, a small MSE is desirable.

If the two methods result in different outcomes, it could indicate that either the GaMcK method or the MMM- method is more suitable for this SKU to identify the most appropriate model. Therefore, both methods are evaluated. In the following chapters, the MMM-method is used as classification model because the minimal MSE is found to result in forecasts that are more accurate (Hyndman, Koehler, Snyder, & Grose, 2002). In Chapter 7, the GaMcK-method is evaluated and discussed as well.

5.3.2 Forecasting

As already mentioned, six forecasting methods are proposed. With forecasting, historical sales are analyzed in order to extract demand patterns that count for demand patterns in the future (Ord & Fildes, 2012). De Gooijer & Hyndman (2006) showed in their research that simple exponential smoothing methods are performing better than ARIMA models, especially when data is not normally distributed. A reason for this is that exponential smoothing models are not that subject to model selection problems when the parameters are optimized. This is reasonable, because model 3 is equal to model 2 with the parameter phi equal to 1 for example. In Appendix J, the methods of forecasting are shown in more detail. The underlying demand distribution will not remain constant indefinitely. Therefore, the alpha for exponential smoothing is introduced in order to balance responsiveness and stability. Exponential smoothing lags behind a trend if there exists one (Nahmias & Olsen, 2015). Holt's method is introduced to deal with the trend and Holt Winter's method is used in order to deal with seasonality. Since the parameters are optimized in this research, the exponential smoothing models are appropriate models for this research and are for that reason used in order to determine the required mean and variance for the determination of the safety stock levels. As mentioned before, the parameters (α, β, θ) are optimized for each method, since it is dangerous to guess them (Gardner, 2006).

Restrictions to forecasting models

It is not possible to forecast negative sales. Therefore, the restriction is made that the determined level component and the corresponding estimated forecast can not drop below 0.01. In this way, a negative forecast is ruled out. Without this restricction, a forecast that would drop below zero during the simulation would not be able to easily become positive again. SKU's within model 0 are forecasted using Simple exponential smoothing with α = 0.3, which is widely used in practice (Gardner, 2006). The average of estimated optimal alpha for SKU's that are classified to group 1, by both the GaMcK classification method as the MMM classification method, is on average 0.3 as well. Examples of the different forecasting methods applied to a SKU that is classified to group 4 by both classification methods, are shown in Appendix J.

Initial values

Before the forecasting procedure or the MMM classification methods starts, the initial values have to be estimated. Notice that the initial values for the level and slope (trend) components loose its importance when n is large and therefore, does not have a significant effect on the forecast made by an exponential smoothing method. When n < 20, the initial value becomes more important (Capar, 2015). The average of the BaseMA is taken as the initial values for the simple exponential smoothing method. For the other methods, fitting the ordinary least squares method (OLS) to the BaseMA is used in order to include the fixed drift or trend component (Gardner & Mckenzie, 2011). This analysis fits a straight line to a set of data and determines the level and slope component. The rest of the training period is used for optimizing the parameters. For the seasonal models, the initial values are determined according to the seasonal factors for stationary series (Nahmias & Olsen, 2015). These initialization methods are explained in more detail in Appendix K.

Forecasting the variance

The variance is an input variable for determining the reorder levels. However, the question is: how to forecast the variance properly? Nahmias & Olsen (2015) are mentioning that the obvious variance to use is the historical variance, however, the correct variance to use is the variance of the forecast error.

This is due to the introduction of a sampling error into the process of estimation. This sampling error accounts for the forecast error variance. Therefore, initially, the forecast error (MSE) represents the variance used for estimating the required reorder levels. However, what is the variance when the SKU did not have sufficient historical data to find the optimal forecasting method with corresponding parameters? Moreover, which variance should be taken if the forecasted relationship between mean and variance does not exist when the demand distribution is discrete according to Adan, van Eenige, & Resing (1995) as discussed in section 5.1.1? Is the determined variance by forecasting sufficient enough or should different methods be used?

Initially, the variance of the forecast error is used in order to determine the reorder levels. However, when SKU's did not have sufficient historical data in order to optimize the parameters, simple exponential smoothing is used as mentioned before. However, this acquired variance could also be insufficient because of a lack of data. Additionally, looking at the mean and variance acquired by the forecast methods, almost 20% of the SKU's are classified into the impossible areas from the theory of Adan, Eenige van , & Resing (1995). These issuses point out, that the determined variance should be revised. For these SKU's, the method of Van Donselaar & Broekmeulen (2013) fits the Poisson distribution to the SKU.

In order to deal with these discrepancies, the power law may be a more appropriate method than fitting the Poisson distribution. Taylor's Power Law (TPL) provides the relationship between the mean and the variance of the population through time or through space. This method is used as a statistical description of aggregation for the formulation of transformations that stabilize the variance (Taylor, 1961). The log-transformed variance of the power law is plotted as a function of the mean, which is log-transformed as well. This function exhibits a characteristic slope (Taylor, 1977). $\log(variance) = \log a + b * \log(mean)$. On an arithmetic scale, this becomes a power law in the following form: $variance = a * (mean)^b$. How these variances are transformed is shown in Figure 11. In this way, the variance is stabilized and this leads to less dispersion of the variance. This relationship indicates the aggregated relationship between variance and mean. Therefore, by estimating the power law of mean and variance for the SKUs that are not affected by seasonality or trend and for the SKUs that have relationships between mean and variance that exists, the aggregated variance could be acquried. Using this variance for SKU's with relationship that did not exists of for SKU's with insufficient observations, the variance seems more reasonable because the SKU is not classified to the impossible areas anymore.

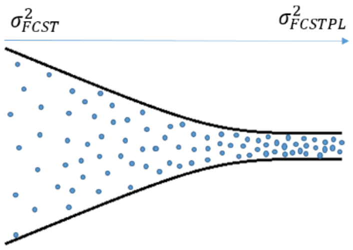


Figure 11: Disaggregated variance to aggregated variance

In conclusion, the following steps are taken during the classification and forecasting phase:

1. Determine proposed forecasting model GaMcK
2. Determine initial values per model
3. Determine proposed model with the MMM method by measuring the MSE and optimizing the parameters (α, β, θ)
4. Forecast the expected sales with the associated forecasting method on holdout

5. Analysis of relationship between forecasted mean and variance acquired by the exponential smoothing models

5.3.3 Reorder level determination

After the mean and variance are forecasted, the reorder levels can be determined. The reorder levels consist of cycle stock and safety stock. Cycle stock is defined as stock during lead-time and review period and can be calculated by calculating the average demand during lead time and review period. The cycle stock is the minimal amount that is should be reserved at Docdata and the safety stock is used to deal with uncertainties. However, high safety stocks result in a decrease of the profitability of the Supply Chain (Bhatnagar & Teo, 2009). The formula used is: $s = \mu * (L + R) + SS$, where $\mu * (L + R)$ is equal to the cycle stock and SS denoting the safety stock (Yamazaki, Shida, & Kanazawa, 2016; van Donselaar & Broekmeulen, 2013). The cycle stock can be calculated using the forecasted mean, the lead time and review period are both equal to one.

An important remark is no historical demand data is available and because of lost sales not easy to identify. For that reason, the mean and variance of the demand could be underestimated (Nahmias & Olsen, 2015). Therefore, the reorder levels are determined according to the proposed method by van Donselaar & Broekmeulen (2013), which takes this issue into account. They proposed a method to determine the required reorder levels in a lost sales inventory system with a periodic review system and a target fill rate, which is the case at bol.com. The target fill rate depends on the ABC-classification. Therefore, after the forecasting phase is finished, the SKU's are all classified to either the A, B or C class looking at a three months ahead forecast. The target fill rates of 96%, 94.5% and 85% apply for SKU's of respectively the A, B and C class.

Initially, the method of Adan, Eenige van, & Resing (1995) is used in order to create the probability mass function for discrete distributions. However, another method that is interesting to evaluate and has not been addressed much in literature so far is the method of van Eenige (1996). This method includes the maximum observed demand as input variable and creates a probability mass function (PMF) with the same mean and variance but with limited support. Due to this, reorder levels are estimated to deal with the largest demand observations. The target fill rate levels are taken into account. All data can then be assumed to be binomially distributed. Therefore, especially for data that is negative binomial distributed, this method has a large effect. In Figure 12 the consequences of the this method are shown in comparison with the method of Adan, Eenige van, & Resing (1995). The blue line defines the PMF acquired with the method of Adan, Eenige van , & Resing (1995) and the orange line indicates the PMF acquired with the method of van Eenige (1996). This is an example of a SKU with mean 16, stDev of 5 and a maximum observation of 99. Resulting from this, less safety stock could be required in order to achieve the target fill rates.

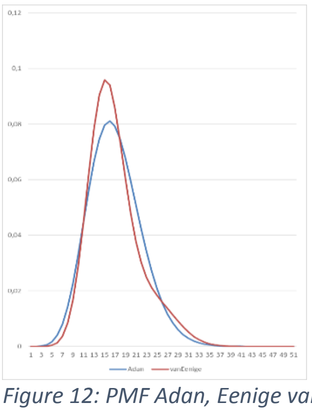


Figure 12: PMF Adan, Eenige van , & Resing (1995) and van Eenige (1996)

Concluding, the reorder levels can be estimated, using the mean and variance acquired by the forecasting methods, by using the following methods:

1. If relationship between mean and variance does not exist according to Adan, van Eenige, & Resing (1995), fit to Poisson distribution. Use Adan, van Eenige, & Resing (1995) to create PMF.
2. If relationship between mean and variance does not exist or no sufficient data is available, use aggregated variance acquired by TPL. Use Adan, van Eenige, & Resing (1995) to create PMF.
3. If relationship between mean and variance does not exist, fit to Poisson distribution. Use method of van Eenige (1996) to create PMF.
4. If relationship between mean and variance does not exist or no sufficient data is available, use aggregated variance acquired by TPL. Use method of van Eenige (1996) to create PMF.

The reorder levels can be translated into the maximum inventory on hand for which the storage capacity should be reserved. Concluding, the following steps are taken in order to determine the reorder levels:

1. Determine target fill rate by the ABC-classification method.
2. Determine reorder levels for the four different models per SKU in units with method proposed by van Donselaar & Broekmeulen (2013)
3. Determine maximum inventory on hand considering the case pack size in units per SKU per period.

The required capacity can be determined by multiplying the required maximum inventory on hand by the volume per SKU. The aggregated volume over SKU's per product category is determined.

5.4 Proposed hypotheses

Based on the proposed conceptual model and the described plan for redesign, the formulated hypotheses are shown in Table 11.

Number	Hypothesis
1	The MMM method is a more appropriate classification method than the GaMcK classification method
2	The required storage capacity is dependent on the classification methods.
3	Determining the safety stocks with input of method van Van Eenige (1996) performs better than using the method of Adan, van Eenige, & Resing (1995).
4	Using the aggregated relationship between mean and variance, acquired by the power law, for SKU's with a relationship between mean and variance that does not exists or for SKU's with insufficient data, performs better than fitting the Poisson distribution.
5	The fill rate is a more appropriate service level measurement than the discrete ready rate

Table 11: Formulated hypotheses

5.5 Summary of the plan for redesign

In this chapter, the used data is shown and the selected subset provided. It is shown that the data is binomial and negative binomial distributed. Furthermore, outliers are removed. The whole process of translating historical sales data into the required storage capacity is shown in Figure 13. The first step to do is analyzing the data, which involves substituting outliers, analyzing demand patterns and defining the probability distribution of the dataset. Thereafter, the most suitable model per SKU is determined by the GaMcK method and the MMM method. Afterwards, the acquired best-fitting model with associated optimal parameters are used in order to forecast the sales. Based on the forecasted sales, the SKU's are classified following the ABC- classification. The target fill rates, the maximal observed observation and forecasted mean and variance using the four different methods are used in order to determine the reorder levels. The reorder levels are translated into the maximum inventory on hand, taking the case pack size into account. These levels indicate the products for which storage capacity should be reserved.

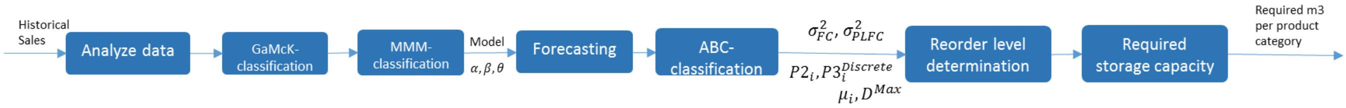


Figure 13: Process of translating historical data into the required capacity

In the following chapter, the results are provided per phase. Subsequently, different scenarios are provided, elaborated and evaluated. Using this information, the most appropriate model can be selected and the hypotheses can either be supported or be rejected.

6. Redesigned model

In this Chapter, results of the plan of redesign are shown. Firstly, the results of the classification of the SKU's are evaluated. Secondly, the results of forecasting are discussed. Furthermore, validation and intermediate results are shown. Finally, the final required storage capacity per method is provided. In Chapter 7, other scenarios are discussed and the best performing model is selected.

6.1 Results of classifying the SKU's

Examples from the historical sales from different models where both classification models selected the same model are shown in Figure 14. As can be seen, clear demand patterns arose.

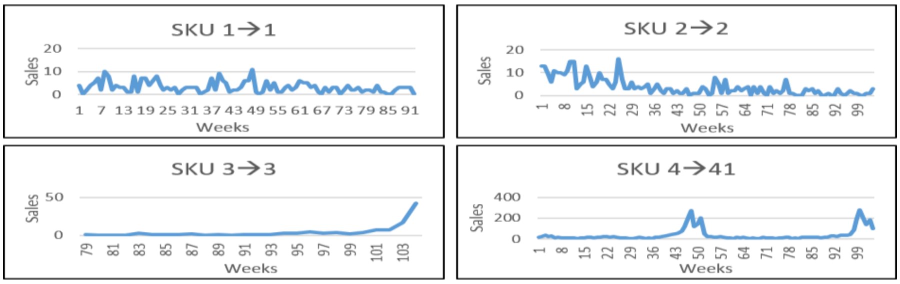


Figure 14: Demand patterns per classification group

Comparing method of Gardner Jr & McKenzie (1988) with minimal MSE method

In Table 12, the results of the two classification methods are shown. The model chosen by the MMM method is denoted as 'the first choice' of the MMM-method. Almost half of all SKU's did not have sufficient historical data and are classified to model 0. A reason for this could be that the product is relatively new and does not have sufficient history in the training period, also referred to as phasing in. Another reason could be the phasing out concept. Furthermore, it is remarkable that in (100%-63.62% =) 36.38% of the cases, the model chosen by the GaMcK method is not the model chosen by the MMM-method. These numbers are determined leaving the SKU's of model 0 out of scope. Furthermore, because the GaMcK method does not differ between the different seasonality models, all SKU's classified to a seasonality model by the GaMcK as well as by the MMM method are considered as classified to the same model.The division between seasonality and no seasonality is almost done equally for the both methods. Not even one percent of the SKU's assigned to a model with seasonality performed better with a model without seasonality. Moreover, models with damped trend are often preferred over models with a linear trend. These models are more likely to occur in practice and they are more robust (Gardner Jr & McKenzie, 1988). This is reasonable, because the model with linear trend is equal to the model with damped trend with the parameter phi equal to 1.

Model by GaMcK	Model by MMM	# SKU's	% of total without 0 classified to the same model
0 (no sufficient data)	No seasonality, constant level	199243	48.59%
No seasonality, constant level	No seasonality, constant level	117833	55.90%
No seasonality, constant level	No seasonality, damped trend	43940	
No seasonality, constant level	No seasonality, linear trend	16951	
No seasonality, damped trend	No seasonality, constant level	13957	
No seasonality, damped trend	No seasonality, damped trend	8479	4.02%
No seasonality, damped trend	No seasonality, linear trend	1584	
No seasonality, linear trend	No seasonality, constant level	174	
No seasonality, linear trend	No seasonality, damped trend	21	
No seasonality, linear trend	No seasonality, linear trend	51	0.02%
Seasonality, constant level	No seasonality, constant level	13	
Seasonality, constant level	No seasonality, damped trend	11	
Seasonality, constant level	No seasonality, linear trend	21	
Seasonality, constant level	Seasonality, constant level	3206	3.67%
Seasonality, constant level	Seasonality, damped trend	3797	
Seasonality, constant level	Seasonality, linear trend	739	
Total # SKU's		**410020**	
Total # SKU's without 0		**210777**	**63.62%**

Table 12: Number of SKU's classified per model by both classification methods

Where the model with the second smallest MSE determined by MMM is equal to the model determined by GaMcK, is shown in Table 13, denoted as 'the second choice' of the MMM-method. The average improvement of the MSE comparing the first choice with the second choice is shown as well. As can be seen, especially for the model with linear trend, large improvements have been made going for the first choice. After the second choice, another 24% of the SKU's is classified to the same model. However, still almost 15% of the SKU's is not classified properly, looking at the two first choices of the MMM-method. This could indicate that either the MMM classification method or the GaMcK-classification method is more appropriate model for the time series.

Model by GaMcK	2nd Model MMM	# SKU's	% 2nd classified to the same model	Avg % MSE first choice improved
No seasonality, constant level	No seasonality, constant level	35666	19.96%	2.85%
No seasonality, damped trend	No seasonality, damped trend	12397	51.61%	6.19%
No seasonality, linear trend	No seasonality, linear trend	181	73.58%	76.79%
Seasonality, constant level	Seasonality, constant level	2674	34.34%	13.28%
Total		**50918**	**24.16%**	**24.78%**

Table 13: Second model chosen by MMM

6.2 Results of forecasting

In Figure 15, the acquired forecast and the actuals are shown for product category Toys. As can be concluded visually, the forecasting method is working properly since the forecast is following the actuals quite well. During the peak period, the forecast lays behind. This is reasonable because a smaller alpha could result in a model that reacts later to the fluctuations in demand, as the peak period. In Appendix N, the forecasting results for the other product categories are shown as well. Concluding, the forecasting methods seem to work properly.

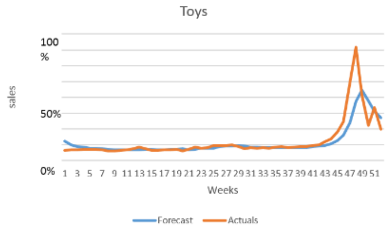


Figure 15: Estimated forecast vs actuals

Forecast accuracy of the developed model

The sales forecast accuracy acquired by the model looking at the holdout period is shown in Table 15 with use of the Forecast accuracy and RMSE. Because the peak weeks are excluded from the outlier detection method, a lower forecast accuracy is automatically expected. However, in order to evaluate the accuracy of the model, the forecast accuracy (100% - MAPE) is shown for regular weeks. Comparing these results with the results from Appendix D, only the products categories Entertainment, Cooking, Dining & Houseware, Domestic Appliances and Computer & Games are slightly less accurate. Especially large improvements of the category Baby have been made. Ereader is improved but still relatively medium accurate.

Bias of the developed model

Furthermore, as mentioned in section 4.3, structural over and under forecasting could bias the forecast. The results, based on the MPE, which determines the bias (Chockalingam, 2010), are shown in Table 14. As can be seen, a structural under forecasting applies. This is due to fluctuating order patterns, the lag of the forecast and the large impact of peak periods, where obviously under forecasting applies. In this case, higher safety stocks could deal with this.

Sales forecast accuracy		
Product Category	FA	RMSE
Entertainment	87,74%	24386,65
Ereader	72,22%	1369,06
Cooking, Dining & Houseware	89,23%	6034,58
Domestic Appliances	89,65%	4809,24
Home Furnishing	85,49%	2717,90
Home Improvement & Gardening	89,16%	2483,72
Baby	82,98%	3097,73
Beauty & Care	88,11%	10933,03
Jewellery, watches & Accessories	82,13%	1288,38
Pet	89,89%	891,78
Sport & Leisure	87,86%	3134,65
Computer & Games	88,02%	8895,92
Mobile & Tablet	88,77%	2249,04
Sound & Vision	90,37%	3614,72
Toys	87,16%	33424,33

Table 15: Sales forecast accuracy of developed model

Product Category	Bias
Entertainment	-5,53%
Ereader	-14,30%
Cooking, Dining & Houseware	-5,42%
Domestic Appliances	-3,94%
Home Furnishing	-10,92%
Home Improvement & Gardening	-9,55%
Baby	-18,76%
Beauty & Care	-7,86%
Jewellery, watches & Accessories	-15,56%
Pet	-11,30%
Sport & Leisure	-6,02%
Computer & Games	-6,28%
Mobile & Tablet	-11,17%
Sound & Vision	-5,48%
Toys	-7,38%

Table 14: Bias of forecast of developed model

Results of powerlaw

In Figure 16, the relationships between the mean and variance after the three different phases are shown for SKUs from model 1 (no seasonality, constant level). Only SKU's that have a feasible relationship between mean and variance according to Adan, Eenige van, & Resing (1995) are taken into account. Visually, the models seem to have a good fit. Moreover, as can be seen in the upper right corners of the figures, the R-squared and the formula that indicates the relationship between the mean and the variance are provided. The R-squared ranges from 0 to 1 and shows the proportion of the total variability that is explained by the designed model. The closer to one, the better the fit of the model. The three phases are tested. As can be seen from the Figure, these numbers indicate a good fit. Note that the variance of GaMcK is corrected with the found alpha corresponding to Nahmias & Olsen (2015), where the variance of the forecast error is equal to $\sigma^2(\frac{2}{2-\alpha})$.

For the aggregated variance acquired by the power law, the used relationship is the relationship after the training period: $\log(\sigma_\varepsilon^2) = 0.4303 * \log(\mu) + 0.0511$. An example is a SKU with a forecasted mean of 0.03157056 and variance of 0.01836082816. This SKU is classified to the areas that do not exists within the discrete distributions according to of Adan, Eenige van, & Resing (1995). Using the power law, the forecasted variance is equal to $\log(\sigma_\varepsilon^2) =$ log(0.03157056)*0.4303+0.0511. After this adaptation of the variance, the relationship between mean and variance does exists according to Adan, Eenige van, & Resing (1995).

Figure 16: Power law of different phases (historical (left), training period (middle), holdout period (right))

6.3 Results of reorder level determination

After the forecasting phase, the target fill rates are acquired using the ABC-classes. This information can be used in order to estimate the reorder levels using the four different methods.

ABC classes

Firstly, the target fill rates have to be estimated using the ABC-classification. The results are shown in Table 16. As can be seen, 10% of the total amount of SKU's are classified to class A, which count for 50% of the sales. Note that the SKU's that did not have sufficient historical sales in order to forecast the expected sales, no sales in the holdout period for example, are also taken into account. These are probably phasing out and are therefore classified as C-products. These classes indicate the target fill rates that are used as input variable when determining the safety stocks.

ABC	# SKU's	% SKU's
A	688	10,00%
B	2041	29,67%
C	4151	60,33%
Total	6880	100,00%

Table 16: Division SKU's ABC-classification

The minimal reorder levels for which capacity should be reserved are equal to the demand during review period and lead time, also known as the cycle stock. These minimum reorder levels per product category per period are provided in Appendix O. However, in order to deal with uncertainties safety stock has to be included. Using the target fill rates corresponding to the ABC-classes as input variables, together with the maximum demand observations and the forecasted mean and variance, the reorder levels are determined per method.

The obtained reorder levels are measured at the end of the training period, as mentioned in section 5.1. In Appendix P, the obtained reorder levels per method are provided. Only SKU's are included with a maximum demand larger than one as mentioned in section 5.1. Furthermore, for eight SKU's the ratio between variance and mean was too large to include in the DoBr-tool (Variance to mean larger than 10000). However, all these SKU's are phasing out. Therefore, it is not advisable to store those products at Docdata. For that reason, these SKU's are excluded from the model and a reorder level of 1 is assigned. The acquired reorder levels are translated in the maximum inventory on hand using the following formula: $\max(IOH) = s_i + Q - 1$.

6.4 Results required capacity

Figure 17 shows the results, the required storage capacity versus the corresponding lost sales, visually. As can be seen, the method of van Eenige (1996) did not end up in any improvements. The results where exactly the same for methods 1 and 3 as well as for method 2 and 4. Therefore, method MMM and MMM-vE as well as MMM-PL and MMM-PL-vE are positioned on exactly the same place in the Figure. Furthermore, the difference between the powerlaw and MMM method are also almost not noticable. The required capacity are for method 1 and 3 equal to 3787.55 m3 and for method 2 and 4 equal to 3791.09 m3. The lost sales are equal to respectively 548.51 and 548.50 for methods 1 and 3 and methods 2 and 4. As can be seen in the Appendix, for each class the target fill rates and discrete ready rates are on average achieved. For 15% of the SKUs, the target fill rates are not achieved. This resulted in lost sales of 2,2% of the total sales.

It was expected that using the aggregated variance instead of fitting the SKU's to the Poisson distribution or using the method of van Eenige (1996) would result in a better performance. However, as can be seen in the described results, no significant improvements have been made. Decisive reasons for this remain unclear. Only 8% of the SKU's has a mean larger than 10. This indicates that the assortment of bol.com consists of many slow moving products. Furthermore, the outlier detection method could obviously substitute the largest demand observations. Hence, the model of van Eenige (1996) would not provide large improvements. These reasons also indicate that the power law could not provide large improvements. Looking at the models where aggregated variance is used, the reorder levels were slightly higher than for the models without aggregation of the variance. The mean of most of the SKU's, for which the power law is used in order to determine the variances, is smaller than one. Therefore, the power law was not able to provide significant results.

Comparing the results to the minimal required storage capacity determined with the benchmark method, some interesting issues arise. The required capacity increased drastically. The required capacity is more than 9 times as much as the required capacity determined with the benchmark method. This is not a surprise because the benchmark shows the reorder levels required for just achieving the target fill rates. However, the lost sales also decreased. This is due to higher fill rates. The reorder levels were increased for SKU's that have an influence on the total required capacity and on the total lost sales.

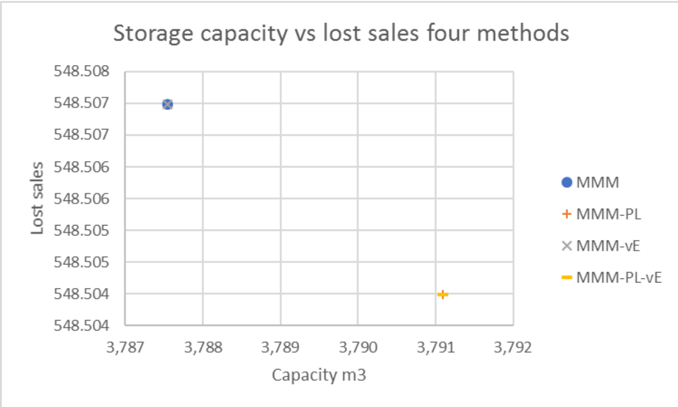


Figure 17: Results 4 methods on training period

7 Scenario analysis

In this Chapter, different scenarios are analyzed in order to conduct a sensitivity analysis and to select the most appropriate method or scenario. Firstly, the results of the initial scenario, which includes the four methods evaluated in chapter 6, are discussed and motivations for the scenarios are provided. Thereafter, the different scenarios are evaluated based on the amount of lost sales and required storage capacity. Finally, a sensitivity analysis is provided.

7.1 Analysis initial scenario

The results of the four different methods for determining the reorder levels and therefore the required capacity are discussed in Chapter 6. The method of Van Eenige (1996) did not provide any improvements. The required capacity are for method 1 and 3 equal to 3787.55 m3 and for method 2 and 4 equal to 3791.09 m3. The aggregated variance even resulted in a worse performance. The lost sales are equal to respectively 548.51 and 548.50 for methods 1 and 3 and methods 2 and 4. For all methods applies that for 15% of the SKUs, the target fill rates are not achieved. This resulted in lost sales of 2,2% of the total sales. Because the method of Van Eenige (1996) or the power law did not provide any improvements to the developed model, it has not been taken into account when evaluating the scenarios.

The initial model considers several issues. For example: classification regarding trend and seasonality, most appropriate forecasting method with optimized parameters, ABC – classification and forecasting of variance. However, how would the model react and how would the results be if these issues were not addressed but assumptions were made. These scenarios could simplify the model, the computation time would be reduced.

Introduction to scenarios

Firstly, it can be seen that the GaMcK method resulted in 40% of the SKU's in a different proposed forecasting method than the MMM-method. Therefore, it is interesting to check how the model reacts when the forecasting model is solely depending on the GaMCK classification method instead on the MMM-method. Which method provides the best results, meaning the smallest required capacity and the lowest lost sales. This issue is addressed in Scenario A.

Secondly, Scenario B involves evaluating whether a classification method is sufficient. The classification models classify the SKU's according to level, trend and seasonal components and the forecasting model that fits the SKU best is chosen. However, are these comprehensive models necessary, or would simple exponential smoothing give sufficient results as well? This scenario is discussed in Section 7.3.

Thirdly, the SKU's are all classified to either the A, B or C category, which provides them with corresponding service levels constraints. A lower target fill rate will probably result in lower reorder levels and therefore a lower required capacity. However, what is the required capacity when all the target fill rates are increased and set equal to 96%? This scenario could be worthwhile considering as well for bol.com. It could be interesting to revise the target fill rates of all products if the target fill rates are not hugely influencing the required storage capacity.

Evaluating the scenarios

Just as the four methods discussed in chapter 6, the results of the scenarios are estimated with use of the data observations of the training period as well. The results are evaluated using the same amount of data as for the evaluation of the four methods depending on the target fill rates, explained in section 5.1. Concluding, in order to evaluate the different models and scenarios, the acquired reorder levels at the end of the training period are evaluated looking back from a certain data point at the end of the training period.

7.2 Scenario A: GaMcK as classification method

In this Scenario, the classification method of GaMcK is used as classification method instead of the MMM-method. Consequently, the forecasting method proposed by GaMcK is used to forecast the sales. The parameters are still optimized and the initial values estimated. The results for this Scenario are shown in Appendix R. In Figure 18, the results for the benchmark, all methods and all scenarios are shown visually.

As can be seen, the capacity required determined by the benchmark method did result in achieving the target fill rates. Scenario A resulted in more than 9 times as much required capacity. However, the lost sales are less than the lost sales of the benchmark as well. This is due to higher achieved fill rates. In Table 17, the percentage of SKU's that did not achieve the target fill rates and the percentage of lost sales are shown per method or scenario. As can be seen 14,7% of the SKU's did not achieve the target fill rates which resulted in 2.17% lost sales. In the Appendix is shown more clearly that this Scenario results in the lowest required capacity and the lowest lost sales. This scenario performs better than the four methods discussed in Chapter 6. This could indicate that the MMM-method is overfitting and therefore, does not result in the best outcomes. This is also mentioned by Chatfield (1995).

7.3 Scenario B: Simple exponential smoothing

This scenario only includes the simple exponential smoothing method as forecasting method. This means that no identification model is required. The results for simple exponential smoothing are shown in Appendix R. An alpha of 0.3 is widely used in practice (Gardner, 2006). Therefore, this alpha is used when forecasting the sales. Looking at Figure 18, it can be seen that this scenario is performing worse than all methods or scenarios described above. The lost sales did decrease. However, the required capacity is increasing as well. The required capacity is almost 10,5 times as much as the required capacity estimated with the benchmark method. Because the achieved lost sales are below target, the additional required storage capacity is not necessary. This can be due to higher forecast error resulting in higher reorder levels, which is unnecessary. Because the assortment consists of many slow movers, the impact of improved fill rates is not that large compared to the impact of increasing the reorder levels. Furthermore, 13.08% of the SKU's did not achieve the target fill rates, which resulted in 1.91% lost sales.

These results indicate that classification of SKU's and the optimization of parameters is necessary. Simple exponential smoothing is not the most appropriate forecasting method for all SKU's. In addition to that, because trend and seasonal components are often present, a fixed alpha is not sufficient.

7.4 Scenario C: Change of Fill rates

Currently, the ABC classification method is used to rank the SKU's and provide them with different target fill rates. This scenario evaluates if changing the fill rates for all SKU's to 96% provides sufficient outcomes as well. The sales forecast is acquired by the forecasting methods per SKU, proposed by the MMM-method. The parameters are still optimized and the initial values estimated. Only the target fill rates are changed to 96% for all SKU's. The results for this scenario are provided in Appendix R as well. As can be seen, the target fill rates do have a significant influence on the model. The required capacity is increased drastically. For this scenario, the required capacity increased with more than 1200%, compared with the benchmark method. However, the lost sales did not decrease with the same degree. Moreover, 10.38% of the SKU's did not achieve the target fill rates, which resulted in 1.80% lost sales.

The results of scenario C are indicating that changing the target fill rates has a large impact on the required storage capacity. It shows that for increasing the fill rates, the required storage capacity has to grow exponentially. Therefore, setting the fill rates to an appropriate level is of great importance.

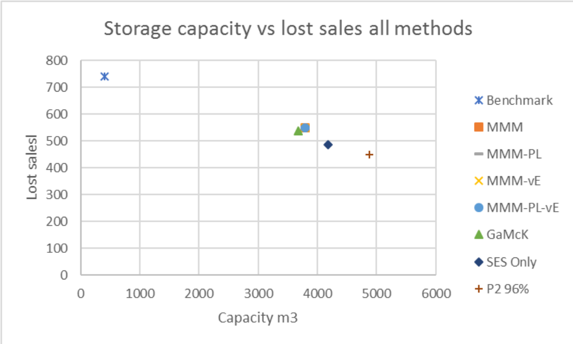


Figure 18: Results all methods and scenarios

In Table 17, the percentage of SKU's for which the target fill rates are not achieved with the evaluated method are shown in the second column. In the third column the percentage of lost sales are shown. These numbers indicate the consequences of when the target fill rates are not achieved. These lost sales are determined by taking the ratio of lost sales and mean. The lost sales are determined by multiplying the mean with 1 minus the achieved fill rate. This table shows that the percentage of SKU's for which the target fill rates are not achieved and the percentage of lost sales does not differ much between the different methods. This indicates that a large part of the SKU's are slow movers and intermittent demand patterns occur frequently.

Method	% SKU's for which achieved fill rate < target fill rate	% Lost sales
MMM	14.95%	2.20%
MMM-PL	14.95%	2.20%
MMM-vE	14.95%	2.20%
MMM-PL-vE	14.95%	2.20%
GaMcK	14.70%	2.17%
SES only	13.08%	1.19%
P2 96%	10.38%	1.80%

Table 17:Percentage target fill rate not achieved and percentage lost sales of different methods and scenarios

7.5 Selected model

With use of the results provided in Chapter 6 and 7, the most appropriate model is selected and evaluated. Based on the training period, the reorder levels per method are estimated and the required capacity determined. Using these levels, the models are evaluated looking back at a certain part of training period as explained in section 5.1. Looking at Figure 18, the MMM-method and the GaMcK are closest by the benchmark method, which is desirable. The benchmark methods indicates the minimal amount of storage capacity required to achieve the target fill rates. However, the GaMcK method resulted in less required capacity and less lost sales, comparing to the MMM-method. Therefore, comparing to the other methods and scenarios, the GaMcK method (Scenario A) is selected as best performing model. The MMM-method seems to be overfitting. This is not unusual in empirical modelling and it can be due to including many variables into the model (Chatfield, 1995).

Performance of selected model

In Table 18, it is shown how the selected model is performing on the holdout. Note that the parameters were optimized over the training period. The found parameters were used to forecast the sales for the holdout period, without intermediately updating them. The same applies for the found forecasting model. This indicates a pessimistic scenario. Still, one week ahead forecasts are provided, which is a more optimistic scenario. Furthermore, due to outphasing SKU's, less SKU's were taking into account. This resulted in lower reorder levels, corresponding maximum inventory on hand and capacity. In Appendix R the results are shown per product category.

ABC	Reorder level	Maximum inventory on hand	Min fill rate	Avg fill rate	Avg discrete ready rate	Capacity in m3	Lost sales
A	3286	8509	0.32	0.82	0.74	102.16	302.16
B	11066	14765	0.34	0.91	0.88	372.98	192.53
C	9143	11313	0.36	0.94	0.93	155.05	97.35
Totaal	23495	34587		0.89	0.85	630.18	592.04

Table 18: Results selected model (GaMcK method) on holdout

As can be seen, the target fill rates are not achieved for the A and B class. This could be due to the just mentioned lack of updating the most appropriate forecasting model with corresponding parameters. Especially with the fast changing assortment of bol.com, is it likely that products are phasing out (or phasing in). The most appropriate forecast model could change from a seasonal model to a trend model for example. In this case, the model should be updated again in order to use the most appropriate model and parameters.

Results of fill rate versus discrete ready rates

In order to determine the correctness of the use of the discrete ready rate, the two service levels are compared. In Appendix R is shown that the differences between the achieved fill rates and the discrete ready rates could be large. The impact of an out of stock is measured with the fill rate measurement. For the discrete ready rate, the real impact is not defined. However, when demand patterns are not fluctuating drastically, these measurements could be more or less equal. On the other hand, when demand patterns are fluctuating, the fill rate will be relatively lower. Therefore, it is important to measure the fill rate because it measures the impact of an out of stock and thus the impact it has on the customers. How many customers are disappointed and what is the impact when it is chosen to let a SKU go out of stock for example are questions that could be answered.

7.5 Sensitivity analysis

It is expected that models that are more complicated yield better results, however this does not have to be the case. The bias could be reduced. However, other parameters, like variance, could be increased because more parameters are involved. Due to including many variables into a model, the method could be overfitting (Chatfield, 1995). To evaluate if the developed model is robust against changes into the model, a sensitivity analysis is performed. Looking at the results provided in previous chapters, it could be concluded that the different methods and scenarios do have a significant effect on the results.

As can be seen from Figure 18, the target fill rates are the main drivers of the model. Increasing the fill rates results in an exponentially increasing required capacity. Furthermore, excluding the classification methods, and solely using simple exponential smoothing, is an important driver as well. Moreover, optimizing the parameters seems necessary looking at scenario B. Which one of the last two has to largest influence remains unclear. Additionally, changing the classification method to the method of GaMcK, provided better results. Relying on the classification method of MMM will result in a model that is overfitting. GaMcK assumes a corrected historical variance, which is a more sufficient variance than the variance acquired by the MMM method. Other drivers, of which the extent to which they will change the model are not investigated, are the lift factor, the minimum observation for detecting outliers, the initial values of the forecasting models, the lead time and review period. More importantly, the impact of an m-step ahead forecast is not considered as well. This driver could also have an essential impact on the proposed model. Currently, a one-step ahead forecast is determined instead of a more ideally m-step ahead forecast in order to timely communicate a stock forecast. In this way, Docdata can prepare for inbound, storage and outbound streams through the concepts of healing and re-warehousing discussed earlier in this research. Lastly, the frequency the model is updated could be an important driver. This probably depends on the frequency of healing and re-warehousing as well.

7.6 Evaluation of hypotheses

Due to the results described in Chapter 6 and the section above, the hypotheses can be either supported or rejected. In the following areas, the six hypotheses are discussed.

1. **The MMM- method is a more appropriate classification method than the GaMcK method**

Looking at the results in section 6.1, it is shown that in almost 40% of the SKU's the GaMcK provided different results than the MMM-method. This could indicate that one of the two models is more appropriate for the classification of SKU's. However, the seasonality of products were identified well by both methods. Taking the results of Chapter 7 into account, the GaMcK classification method resulted in a better performance than the MMM method. Less storage capacity is required and less lost sales occurred. Therefore, the hypothesis is rejected. The MMM method is not a more appropriate classification method than the GaMcK method. It should be mentioned that the opposite applies, the GaMcK classification method seems to be a more appropriate classification method than the MMM method. The MMM method seems to be overfitting.

2. **The required storage capacity is dependent on the classification methods**

In section 7.3, the scenario of having no classification method is discussed, where only simple exponential smoothing is used as forecasting method. The required storage capacity increases exponentially with the use of only simple exponential smoothing and a fixed alpha. Concluding, the required storage capacity is dependent on the classification methods. This is due to large forecast errors and therefore large variances as input variables for the determination of reorder levels. The models including a classification method provided better results, GaMcK and the MMM method resulted in less required capacity. On the other hand, the lost sales were less than with the classification methods. However, the target fill rates were the guidelines to achieve and therefore more storage capacity is not desirable. Concluding, the hypothesis is supported.

3. **Determining the safety stocks with input of method van Van Eenige (1996) performs better than using the method of Adan, van Eenige, & Resing (1995).**

As mentioned in section 6.3, methods 3 and 4 did not provide any improvements to the model. The proposed model did not even in any differences at all. Decisive reasons for this remain unclear. Only 8% has a mean larger than 10. This indicates that the assortment of bol.com consists of many slow moving products. Furthermore, the outlier detection method substituted the largest demand observations. Hence, the method of van Eenige (1996) does not provide improvements to the developed model for the SKU's of the subset. Therefore, the hypothesis is rejected in this research.

4. **Using the aggregated relationship between mean and variance, acquired by the power law, for SKU's with a relationship between mean and variance that does not exists or for SKU's with insufficient data, performs better than fitting the Poisson distribution**.

Against the expectations, the power law did not show any improvements as well. Apparently, fitting the Poisson distribution for SKU's where the relationship between mean and variance was classified as unfeasible by the method of Adan, Eenige van , & Resing (1995) is sufficient as well. The revised SKU's have low means. This could be due to intermittent or slow moving demand patterns. Therefore, no significant results could be made. Therefore, the hypothesis is rejected for the time series in this research.

5. **The fill rate is a more appropriate service level measurement than the discrete ready rate**

As mentioned in this section, the discrete ready rate and fill rates differ significantly. For some product categories, the achieved discrete ready rates are lower than the achieved fill rates. For other product categories, the achieved discrete ready rates are higher than the achieved fill rates. This is due to the impact of an out of stock. When the impact is high (and it resulted in many lost sales), the fill rate will be lower than the discrete ready rate. When the impact is low, it is the other way around. Therefore, the fill rate is an important measurement that always measures the service level towards the customer. It measures the unobserved demand. Therefore, the hypothesis is supported. However, in the case of lower impacts of an out of stock, the discrete ready rate may be sufficient as well.

8 Conclusion & Recommendations

In this chapter, conclusions are drawn and recommendations are provided, referred to the implementation part of the regulative cycle. Firstly, the general conclusions are given. The selected model is explained and results are provided. Secondly, the research and subquestions are answered. Furthermore, how other retailers could use the model developed in this research is discussed before the scientific contributions are explained. Finally, the recommendations and steps for implementation are presented.

8.1 Answers to the research question

In this section, the research question is answered taking the answers on the subquestions into account. The research question of this research is: *What is the required storage capacity at Docdata in order to achieve the desired fill rates?* The developed model includes the most appropriate classification model (GaMcK), the most appropriate exponential smoothing forecasting methods per SKU and most appropriate variance determination, which leads to the required forecasting capacity. In order to answer the research question properly, the subquestions are answered.

6. How are the processes at bol.com designed regarding forecasting of sales and capacity and what is the current performance?

In section 1.1 and 2.1, the different processes of bol.com regarding sales, stock and capacity forecasting are explained. It has been made clear that those forecasting procedures are independent processes developed by different departments of the organization. Bol.com has the desire to change this into one automatic, consecutive process. Furthermore, the distribution centers receive many shipments of small quantities that results in higher transportation and handling costs. In addition to that, Slim4 provides monthly sales forecasts that causes al lot of operational work. Moreover, the higher the volume does not necessarily indicate a higher forecast accuracy. Besides, the percentages of bias indicated that no structural over or under forecasting applies. Finally, the target discrete ready rates of 2015 are achieved. However, high levels of obsolete stock are showing that unnecessary storage capacity was required.

7. How can the SKU's be classified regarding trend and seasonality?

For the own assortment of bol.com, different exponential smoothing forecasting methods are introduced in order to forecast SKU's with diverse characteristics. These techniques are well suited to companies that manage many SKU's and produce their forecasts semi-automatically. Little computing time and data storage are necessary. Furthermore, it is widely used for SKU's with similar properties, as is the case in sales forecasting and inventory control (Chatfield, Koehler, Ord, & Snyder, 2001). However, the most suitable forecasting method depends on the trend and seasonality component. Two different classification methods are discussed. The GaMcK- method (Gardner Jr & McKenzie, 1988) is selected as most appropriate in order to classify the products to select the most appropriate forecasting method. The method is discussed in section 6.1. In sections 7.2 and 7.5, the results are provided.

8. How can the sales forecast per SKU per week be determined?

When determining the sales forecast using the forecasting method proposed by GaMck, the parameters (α_i, β_i and θ_i) are optimized. In section 5.3.2, the different exponential smoothing forecasting methods are discussed. These are explained in more detail in Appendix J. In section 6.2 the results are shown. As can be seen, the sales are properly forecasted per period. However, structural under forecasting applies with could be caused by peak periods. Higher reorder levels are set in order to deal with these forecasting errors. The mean and variance are forecasted with the forecasting methods. If insufficient data is available or the relationship between mean and variance does is not feasible, the SKU is fitted to the Poisson distribution.

9. How to determine the associated inventory levels per SKU per week?

After the forecasting phase, the ABC classification procedure have been conducted in order to determine the target fill rates. The reorder levels are estimated with use of the method proposed by van Donselaar & Broekmeulen (2013) discussed in section 5.3.3. With use of Adan, van Eenige, & Resing, (1995), the probability mass function is created for determining the reorder levels. Based on the reorder levels and case pack size per SKU, the maximum inventory on hand is estimated. Storage capacity should be reserved for these levels. The results are shown in section 6.3.

10. What is the required capacity per period at Docdata?

The last step is to determine the required capacity at Docdata based on the inventory on hand. The capacity is dependent on the volume of the SKU. The selected method (GaMcK) estimated the variances that resulted in the smallest reorder levels and required capacity. The results are 630.18 m3 storage capacity and 592.04 lost sales for the selected subset at the end of the holdout period. However, the target fill rates are not achieved. This is probably due to the fast changing assortment of bol.com. Consequently, demand patterns are changing and therefore different forecasting methods and optimized parameters than estimated over the training period may apply.

8.2 Scientific contributions

This research provides several contributions to the literature. Additional knowledge and new angles to existing literature in the academic field is developed. The academic relevance is discussed by identifying lacking issues of existing literature. Different matters are elaborated. Afterwards, the practical contributions are given as well.

As mentioned by Tarn, Razi, Wen, & Perez Jr (2003), e-commerce retailing differs significantly from traditional retailing regarding reorder level and capacity determination. However, literature is lacking. There is a higher probability of significant fluctuation in customer demand and seasonality applies more frequently and is less predictable. E-commerce retailers experience more valleys and more peaks compared to traditional retailers. Therefore, with e-commerce warehousing, minimal stock levels cannot be held due to demand uncertainties, order volumes and seasonal peaks. In this research, a model is developed in which these issues are addressed. This contributes to the literature as it focused on e-commerce retailing with regard to capacity planning.

Furthermore, it contributes to the literature regarding the classification of SKU’s regarding trend and seasonality. The classification procedure of Gardner Jr & McKenzie (1988) and the minimial MSE procedure proposed by Hyndman, Koehler, Snyder, & Grose (2002) are discussed and evaluated. Looking at the results, the conclusion can be drawn that the MMM method is not sufficient for the time series. It seems to be overfitting, which is also addressed by Chatfield (1995).

The retailing industry has grown many folds with the growth of demand for several products. Due to mass customization, problems in areas as layout, slotting, re-slotting and inventory management occurred (Ackerman, 1990). Previous researchers made it very clear that forecasting of demand is the key to a capacity plan. This research contributed to the aspect of inventory management. Several methods of forecasting the variance for calculating the reorder levels have been addressed. The method of van Eenige (1996) has not been addressed much in literature. It could provide significant improvements to reorder levels or safety stock determinations.

Finally, this research contributes on the translation of historical sales into a capacity plan, which is lacking in the literature. Several aspects have been taken into account as the classification of SKU's regarding trend and seasonality, forecasting methods, forecasting of variance and determination of reorder levels.

8.3 Generalizability

Are the developed model, the resulting outcomes and the contributions to science applicable to a broader perspective? The developed model and corresponding results are based on characteristics of bol.com and the SKU's involved. However, since such a large dataset is used, many different issues have been addressed which developed in a model with different possibilities and opportunities. The final selected method to determine the reorder levels is the most appropriate in this research. However, for other companies with different composition of SKU's, other methods could be more profitable. Furthermore, when generalizing the models, it would be an improvement to take other drivers into account as well. For example, the costs aspects have not been addressed in this research. Additionally, capacity constraints could have changed the model as well. In this research, the problem is not a capacity problem. The main focus was on SKU's with high means because they require more storage capacity. This is an aspect of the problem proposed by bol.com. Still, the acquired results are generalizable to many other categories, since the model includes widely accepted classification models and logic. Therefore, the model could be used in many other environments.

8.4 Recommendations & Implementation

Based on the diagnosis of the current performance, the evaluated models and the results that came with it, the following recommendations are made. For a company as bol.com, that has storage issues because of a lack of a clear procedure of accurately forecasting the required storage capacity, the selected method is important. As can be seen from the results, the GaMcK method is selected as most appropriate method.

Furthermore, the following recommendations for implementation are provided:

1. *Implement the automatic forecasting model that deals with the dynamic environment of e-commerce retailing.* Monthly sales forecast does not provide sufficient information in order to deal with the fluctuating demand patterns and manual work still has to be done to deal with this deficiency. The developed model automatically translates historical sales into the required capacity. This forecast can be used as a baseline model for the required storage forecast. Therefore, the model should be used in parallel with experience of the Supply Chain Specialist in order to extend the model to the expected capacity including promotions and current stock positions. If desired, a sales forecast or stock forecast can be derived from the model as well. Furthermore, Docdata can be better prepared during the year and conduct activities as healing and re-warehousing (Kofler M. , Beham, Wagner, & Affenzeller, 2015). If the current inventory position is higher than the determined inventory position, the surplus should be reallocated, put into the outlet of be destroyed. This will avoid having inbound, storage and outbound issues.

2. *Update model frequently and provide a m-step ahead forecast to Docdata*. Because of the changing assortment and therefore demand patterns, the models will provide better results when the models are updated frequently. This includes selecting the most appropriate forecasting models with corresponding optimal parameters per SKU. Furthermore, the reorder levels should be revised frequently in order to deal with the dynamic environment. Using the updated forecasting models with corresponding updated parameters, weekly forecasts can be determined. For example, when Docdata requires the stock forecasts one month ahead in order to timely adapt to the changes, the a month ahead forecast should be provided and the model should be updated just before the month ahead forecast is created.
3. *Determine if broad assortment including many slow movers is advantageous.* Slow-moving items are occupying a relatively large part of the available capacity. This results in higher storage costs. It could be the case that this is not advantageous. Therefore, it is recommended in order to determine the costs aspects of this issue and decide if it is profitable and interesting to hold such a broad assortment.
4. *Use fill rate as service level measurement*. The fill rate measures the service level towards the customer. This might be a more appropriate measurement than the discrete ready rate, which is currently used. However, the demand history is required in order to accurately define the fill rate. Therefore, it is recommended to start estimating historical demand in addition to historical sales. By click analyses for example, lost sales can be estimated and expected demand can be determined. By doing this, the achieved fill rate can be determined and a more accurate estimation of the required stock and associated storage capacity can be estimated.
5. *Determine if the ABC- classification is sufficient*. As can be seen from the results, increasing the target fill rate levels is increasing the required capacity drastically. Bol.com should be determine if products that have large variances should have high target fill rate as well, since a lot of storage capacity is required.

In the next Chapter the proposed methodology is discussed and limitations are provided. These limitations lead to directions for further research.

9 Discussion & Limitations

The last step of the regulative cycle of Figure 4 is evaluation. In this Chapter, the methodology of this research is discussed and limitations to it are provided. Using these issues, the purposes of the research and the model are clarified and interesting topics for future research are provided.

9.1 Discussion

The designed model translates historical sales data into the required capacity. Furthermore, interim results, as demand forecast and stock forecast, could also be acquired from the model. Even though the research and methodology have several limitations, it could be used in many areas of study. Furthermore, many topics for further research would be interesting looking at the acquired results and the limitations that are proposed in section 9.2.

The outlier detection method used is only detecting outliers at the upper bound. Demand observations that were relatively large, indicating promotions for example, were substituted. However, it could be desirable to detect outliers at the lower bound as well, substituting demand observations that did not have enough demand. The issue of unobserved demand could be solved. Furthermore, the outlier detection method itself should be evaluated as well. The method of van Eenige (1996) did not provide any improvements, this could be due to the outlier detection method.

In section 6.2 the results of the forecasting phase are provided. As can be seen in Table 17, under-forecasting often applies. Especially for categories Baby and Jewellery, Watches & Accessories, the forecasts are biased. This is largely due to forecasting the peak period and excluding the Christmas peak from the outlier detection method. Moreover, the MSE forecast error that is used in order to determine the required reorder levels already incorporates this, higher reorder levels apply. However, since structural under forecasting applies, it is probably desirable to take more action.

The costs aspects are not incorporated in this research. The driver of the model are the reorder levels with corresponding required capacity. However, storage costs increase when new floors or halls have to be reserved at Docdata. Therefore, it could be desirable to allow lower service levels in order to decrease storage costs. On the other hand, it could be desirable to increase the reorder levels of profitable products when a floor is already reserved.

The selected model, the GaMcK method, did provide the best results on the training period. However, looking at the holdout period, the target fill rates were not achieved. Before implementing the method, it should be proved that the proposed model with updating processes results in achieving the target fill rates. Moreover, the results are based on one week ahead forecasts. Since Docdata should be prepared a period in advance, it should be determined if the same results would apply when a m-step ahead forecast would be evaluated. Finally, Scenario B is testing two drivers: forecasting method and fixed or variable parameters. In order to determine the impact of both drivers, the two should be evaluated separately.

9.2 Limitations

In this section, the limitations of the executed research are provided and, consequently, the directions for future research shown.

- According to Walter (1977), the probability of detecting trend and seasonality is dependent on the sample size and the proportional amplitude of variation from the average of the series to the maximum. Even a sample size of 250 and an alpha of 0.20, will only yield in a 50% chance of detecting the trend. A reasonable sample size for detecting trend is 15 (Gerrodette, 1987). The probability of a type I error, a significant result when no seasonal trend is present, is small.

However, the probability of a type II error, where a seasonal trend that is present is not detected, is large.

- The Christmas peak is not included in the outlier detection method. Therefore, it would be appropriate to exclude those weeks from the MSE calculation as well.
- The lead time and review period are set equal to one. This in the ideal situation. However, in reality, this is often not the case and larger numbers apply. In this case, higher reorder levels would be necessary in order to deal with the demand uncertainty.
- In this research, the sales and stock forecast are composed conforming a 1-step ahead forecast. In this way, the model can adapt to changes in the historical data and transform into a better fit. However, taking healing and re-warehousing into account, it is desirable for Docdata to receive a m-step ahead forecast in order to timely prepare for the inbound, storage and outbound streams.
- The training period is used in order to optimize the parameters. These parameters are then fixed for the holdout period. However, these parameters could still be optimized frequently taking healing and re-warehousing into account. In the case of bol.com, it would be ideal to optimize the parameters. This applies for the safety stocks as well.
- Slow movers were not the focus of this research. However, the trend or seasonality in demand cannot be assumed due to a lack of evidence. This is due to a dataset with many data points of zeros (Teunter & Duncan, 2009). These SKU's probably require different approaches and the model would be improved by adding those approaches to the model.
- Phasing in and phasing out could have a large impact on the results. It is stipulated that all SKU's have at least one product stored per period. Bol.com is changing and enlarging their assortment during the year, especially towards the peak period. As product life cycles are getting shorter, the unpredictability and randomness of the demand processes have become even greater (Graves, 1999)
- When the product was not available, due to Out Of Stocks (OOS), the historical sales were zero. Therefore, the demand of that customer is unobserved and could be lost or substituted with other products. Especially for low demand products with a higher frequency of observations with zero sales, this may result in poor performance when these are not taken into consideration. Therefore, the true mean and variance of the demand could be underestimated.
- An assumption is made that seasonal indices are calculated using the same weeks each year. However, Carnival occurs every year on different weeks. In the optimal form, the weeks of carnival, independent of the number of the weeks, are used for calculating the seasonal indices.
- The storage capacity plan is based on a just-in-time strategy. Items arrive just in time and preparation for the peak period is excluded from the model.
- Docdata is the warehouse in scope and focus is laid on the Veerweg, one location of Docdata. However, the productions could also be located at other distributions centers. This division is not taken into account.
- A large subset is taken from the whole dataset in order to evaluate and discuss the results. However, there is a probability that the subset is not representative for the whole time series. This could be due to the randomly selected SKU's. It could be the case that the majority of the selected SKU's is phasing out or are slow moving for example.

9.3 Future research

Based on the generalizability, recommendations, implementation and limitations of the research, interesting topics for future research could include the following issues.

Firstly, the proposed method of van Eenige (1996) could improve the model. However, for now, it is only tested on a subset of the data. It could be interesting to test if the method is profitable for the entire dataset of bol.com. This would probably include more products with higher demands. In this way, simple improvements could be made regarding the required capacity. Furthermore, it would be interesting to investigate if the outlier procedure is most suitable to conduct before any of the analyses is done. It is expected that the method of van Eenige (1996) would provide improvements to the developed model when the outliers where treated differently. Another outlier detection procedure could be interesting or to investigate if the outlier detection procedure might be more appropriate at another moment in time. Thirdly, the method proposed by Gardner Jr & McKenzie (1988) is selected as most appropriate. This classification method identifies the most appropriate forecasting method according to the determined minimal variance. The MMM method is evaluated in combination with the power law (Taylor, 1961) or the method of van Eenige (1996). However, it could be interesting to investigate if one of those methods, or both, would be advanteous in combination with the method proposed by Gardner Jr & McKenzie (1988).

Moreover, trend or seasonality is hard to detect when SKU's are intermittent. Furthermore, the used forecasting methods as simple exponential smoothing are unsuitable in intermittent demand scenarios. These methods will lead to stocking decisions that are not optimal. These SKU's were not the focus of the research. However, forecasting methods as simple exponential smoothing are unsuitable in intermittent demand scenarios. These methods will lead to stocking decisions that are not optimal (Teunter & Duncan, 2009). The model can be improved by focusing on these products and find appropriate ways to forecast the required capacity. A method to deal with intermittent demand patterns is Croston's method (Teunter R. H., Forecasting Intermittent Demand, 2006). Furthermore, an aggregation level of one week and per SKU is used. It would be worthwhile to investigate if these levels are properly chosen for slow moving SKU's. Other options could be aggregated on month level or even on quarter level for these products. Since the total required storage capacity per product category is the desired output, aggregation level per product category or per sub-product category or other lower levels, could also be an improvement. Moreover, another option for slow moving or intermittent demand patterns is to directly forecast the reorder levels. Therefore, future research including this topic would be interesting.

Additionally, it would be interesting to investigate the identification of products that are out phasing in order to exclude them from the model. In this way, the required storage capacity could be reduced because many bins are occupied by just one SKU. Moreover, many product categories have to deal with many promotions during the year that causes difficulties. To model could be extended in order to deal with promotions as well and forecast them properly. The same applies for forecasting during peak periods. The 'last like' rule is commonly used, which assumes the promotional quantity as the previous similar promotion (Trusov, Bodapati, & Cooper, 2006) or using the Lift Factor model, which is equal to the ratio of promotional sales and baseline sales (Cooper, Baron, Levy, Swisher, & Gogos , 1999).

Furthermore, it would be interesting to investigate the impact of fixed parameters and the impact of only one forecasting method separately. Lastly, the selected forecast model with corresponding optimized parameters are fixed based on the training period. It would be worthwhile to investigate if the model would be improved and the target fill rates would be achieved when these variables are frequently recalculated. Lastly, the evaluated results are based on one step ahead forecasts, it should be determined if m-step ahead forecasts would provide the same results.

Bibliography

Ackerman, K. B. (1990). *Practical Handbook of Warehousing.* New York: Van Nostrand Reinhold.

Adan, I., Eenige van , M., & Resing, J. (1995). Fitting discrete distributions on the first two moments. *Probability in the Engineering and Informational Sciences, 9*(4), 623-632.

Adan, I., van Eenige, M., & Resing, J. (1995). Fitting discrete distributions on the first two moments. *Probability in the Engineering and Informational Sciences, 9*(4), 623-632.

Anusha, S. L., Alok, S., & Shaik, A. (2014). Demand Forecasting for the Indian Pharmaceutical Retail: A Case Study. *Journal of Supply Chain Management Systems, 3*(2).

Bartholdi, J. J., & Hackman, S. T. (2014). *Warehouse & Distribution science.* Atlanta: Georgia Institute of Technology.

Bartolomei, S. M., & Sweet, A. L. (1989). A note on a comparison of exponential smoothing methods for forecasting seasonal series. *International Journal of Forecasting*, 111-116.

Bertrand, M., Will, J., & Fransoo, J. C. (2002). Operationals management research methodologies using quantitative modeling. *International Journal of Operations & Production Management, 22*(2), 241-264.

Bhatnagar, R., & Teo, C.-C. (2009). Role of logistics in enhancing competitive advantage. *International Journal of Physical Distribution & Logistics Management, 39*(3), 202-226.

Bol.com. (2016). *Intranet*. Retrieved 05 23, 2016, from http://intranet.local.nl.bol.com/

Box, G. E., & Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control.* San Francisco, Calif: Holden-Day.

Boylan, J. E., Syntetos, A. A., & Karakostas, G. C. (2009). Classification for forecasting and stock control: a case study. *Journal of the Operational Research Society, 59*, 473-481.

Broekmeulen, R., & van Donselaar, K. (2015). *DoBr v3.*

Brown, R. G. (1983). *Smoothing, Forecasting and Prediction of Discrete Time Series.* Englewood Cliffs, NJ: Prentice-Hall.

Capar, S. (2015). Importance of initial value in exponential smoothing methods. *Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 17*(3), 291-302.

Chase, C. W. (1994). Customer demand forecasting. *The Journal of Business Forecasting Methods & Systems, 13*(3), 26-28.

Chatfield, C. (1995). Model uncertainty, data mining and statistical inference. *Journal of Royal Statistical Society, 38*, 161-178.

Chatfield, C., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2001). A nwe look at models for exponential smoothing. *The Statistician, 50*(2), 147-159.

Chockalingam, M. (2010). Response to MAPE and MPE calculations. *USA*.

Coffman, E. G., Courcoubetis, C., Garey, M. R., Johnson, D. S., McGeoch, L. A., Shor, P. W., . . . Yannakakis, M. (1991). Fundamental Discrepancies Between Average-Case Analyses Under

Discrete and Continuous Distributions: A Bin Packing Case Study. *Proceedings of the twenty-third annual ACM symposium on Theory of computing*, 230-240.

Cooper, L. G., Baron, P., Levy, W., Swisher, M., & Gogos , P. (1999). Promocast: a new forecasting method for promotional planning. *Marketing Science, 18*(3), 301-316.

de Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. *International Journal of Forecasting, 22*(3), 443-473.

Eaves, A. H., & Kingsman, B. G. (2004). Forecasting for the ordering and stockholding of spare parts. *Journal of the Operational Research Society, 55*, 431-437.

Gardner. (2006). Exponential smoothing: the stage of art, Part II. *International Journal of Forecasting, 4*(1), 1-28.

Gardner Jr, E. S., & McKenzie, E. (1988). Model Identification in Exponential Smoothing. *The Journal of the Operational Research Society, 39*(9), 863-867.

Gardner, E. S., & McKenzie, E. (1985). Forecasting trends in time series. *Management of Science, 31*, 1237-1246.

Gardner, E. S., & McKenzie, E. (1989). Seasonal exponential smoothing with damped trends. *Management science*(35), 372-376.

Gardner, E. S., & Mckenzie, E. (2011). Why the damped trend works. *The Journal of the Operational Research Society, 62*(6), 1177-1180.

Gerrodette, T. (1987). A Power Analysis for Detecting Trends. *Ecology, 68*(5), 1364-1372.

Graves, S. C. (1999). A Single-Item Inventory Model for a Non-Stationary Demand Process. *Manufacturing and Service Operations Management, 1*, 50-61.

Green, S. B. (1991). How many subject does it take to do a regression analysis. *Multivariate Behavioural Research*, 499-510.

Hall, N. G., Ghosh, S., Kankey, R. D., Narasimhan, S., & Rhee, W. T. (1988). Bin Packing Problems in One Dimension: Heuristic Solutions and Confidence Intervals. *Computers & Operations Research, 15*(2), 171-177.

Huang, T., Fildes, R., & Soopramanien, D. (2014). The value of competitive information in forecasting FMCG retail product sales and the variable selection problem. *European Journal of Operational Research, 237*(2), 738-748.

Hyndman, R. J. (2014). Measuring forecast accuracy. 1-8.

Hyndman, R. J., & Athanasopoulos, G. (2014). *Forecasting: Principles and Practice.* Otexts.com.

Hyndman, R. J., Koehler, A. B., Snyder, R. D., & Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting, 18*(3), 439-454.

Keaton, M. (1995). Using the gamma distribution to model demand when lead time is random. *Journal of Business Logistics, 16*(1), 107-131.

Kendall, K. (2000). E-commerce aesthetics: E-commerce for E-commerce's sake. *Information Resources Management Journal, 13*(3), 3-4.

Kofler, M., Beham, A., Wagner, S., & Affenzeller, M. (2015). Robust storage assignment in warehouse with correlated demand. *Studies in Computational Intelligence, 595*, 415-428.

Kofler, M., Beham, A., Wagner, S., Affenzeller, M., & Achleitner, W. (2011). Re-Warehousing vs. Healing: Strategies for Warehouse Storage Location Assignment. *International Symposium on Logistics and Industrial Informatics*, 77-82.

Koster de, R., Le-Duc, T., & Roodbergen, K. J. (2007). Design and control of warehouse order picking: a literature review. *European Journal of Operational Research, 182*(2), 481-501.

Larson, N. T., March, H., & Kusiak, A. (1997). A heuristic approach to warehouse layout with class-based storage. *IIE Transactions, 29*, 337-348.

Ledolter, J. (1989). The effect of additive outliers on the forecasts from ARIMA models. *International Journal of Forecasting*(5), 231-240.

Levitt, T. (1965). Exploit the product life cycle. *Harvard Business Review, 43*(6), 81-94.

Li, Y., & Lu, J. (2011). Model and algorithm for periodic storage allocation based on correlations. *Journal of Mechanical Engineering, 20*, 75-80.

Manzini, R., Accorsi, R., Gamberi, M., & Penazzi, S. (2015). Modeling class-based storage assignment over life cycle picking patterns. *International Journal of Production Economics, 170*, 790-800.

Miller, M. K., Childers, A. K., & Taaffe, K. M. (2009). Improving Reorder Quantities and Forecasting Methodologies through Simulation. *IIE Annual Conference. Proceedings*, 1664-1669.

Mitroff, I. I., Betz, F., Pondy, L. R., & Sagasti, F. (1974). On managing science in the systems age: two schemas for the study of science as a whole systems phenomenon. *Interfaces, 4*(3), 46-58.

Molenberghs, G., Verbeke, G., & Demetrio, C. G. (2007). An extended random-effects approach to modeling repeated, overdispersed count data. *Lifetime Data Analysis, 13*(4), 513-531.

Nahmias, S., & Olsen, T. L. (2015). *Production and Operations Analysis.* Waveland Press. Inc.

Nare, H., Maposa, D., & Lesaoana, M. (2012). A method for detection and correction of outliers in time series data. *Afircan Journal of Business Management, 6*(22), 6631-6639.

Nu.nl. (2016, 03 24). *NU.nl*. Retrieved 05 23, 2016, from http://www.nu.nl/internet/4236418/bolcom-opnieuw-verkozen-beste-webwinkel-van-nederland.html

OECD, (. o. (2009). *Communications outlook: information and communications technologies.* Zurich, Switzerland.

Ord, J. K., & Fildes, R. (2012). *Principles of Business Forecasting.* Mason, OH: South-Western Cengage Learning.

Pierre, B., Nieuwenhuyse van, B., Dominanta, D., & Dessel van, H. (2003). Dynamic ABC storage policy in erratic demand environments. *Jurnal Teknik Industri, 5*(1), 1-12.

Ravinder, H. V. (2013). Forecasting with Exponential Smoothing - What's The Right Smoothing Constant. *Review of Business Information Systems, 17*(3), 117-126.

Rhee, W. T. (1991). Stochastic Analysis of a Modified First Fit Decreasing Packing. *Mathematics of Operations Research, 16*(1), 162-175.

Rosenblatt, M. J., & Roll, Y. (1984). Warehouse design with storage policy considerations. *International Journal of Production Research, 22*(5), 809-821.

Sanders, N. R. (1997). Measuring forecast accurcay : Some practical suggestions. *Production and Inventory Management Journal, 38*(1), 43-46.

Sharma, S., & Shah, B. (2015). A proposed hybrid storage assignment framework: a case study. *International Journal of Productivity and Performance Management, 64*(6), 870-892.

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

Sobel, M. J., & Zhang, R. Q. (2001). Inventory policies for systems with stochastic and deterministic demand. *Operations Research, 49*(1), 157-162.

Song, J.-S., & Zipkin, P. H. (1996). Managing inventory with the prospect of obsolescence. *Operations Research, 44*(1), 215-221.

Suh, N. P. (1990). *The principles of Design.* Oxford: Oxford University Press.

Syntetos, A. A., Boylan, J. E., & Croston, J. D. (2005). On the categorization of demand patterns. *Journal of the Operational Research Society, 56*, 495-503.

Tan, B., & Karabati, S. (2004). Can the desired service level be achieved when the demand and lost sales are unobserved? *IIE Transactions, 36*, 345-358.

Tarn, M. J., Razi, M. A., Wen, H. J., & Perez Jr, A. A. (2003). E-fulfilment: the strategy and operational requirements. *Logistics Information Management, 16*(5), 350-362.

Taylor, L. R. (1961). Aggregation, variance and the mean. *Nature 189*, 732-735.

Taylor, L. R. (1977). Aggregation, migration and population mechanics. *Nature 265*, 415-421.

Tersine, R. J. (1994). *Principles of inventory and materails management* (4th ed.). EngleWood Cliffs NJ: Prentice Hall.

Teunter, R. H. (2006). Forecasting Intermittent Demand. 3-24.

Teunter, R. H., & Duncan, L. (2009). Forecasting intermittent demand: a comparative study. *Journal of the Operational Research Society, 60*, 321-329.

Teunter, R. H., Babai, M. Z., & Syntetos, A. A. (2010). ABC classification: Service Levels and Inventory Costs. *Production and Operations Management, 19*(3), 343-352.

Trapero, J. R., Kourentzes, N., & Fildes, R. (2015). On the identification of sales forecasting models in the presence of promotions. *Journal of the Operational Research Society, 66*, 299-307.

Trevino, S., Wutthisirisart, P., Noble, J., & Chang, A. (2009). An approach for order-based warehouse slotting. *IIE Annual Conference. Proceedings*, 569-573.

Trusov, M., Bodapati, A., & Cooper, L. (2006). Retailer Promotion Planning. *Journal of Interactive Marketing, 20*, 71-81.

Union, I. (. (n.d.). Retrieved 05 19, 2016, from www.itu.int/library

van Aken, J. (2005). Management Research as a Design Science: articulating the Research Products of Mode 2 Knowledge Production in Management. *British Journal of Management*, 19-36.

van Donselaar, K. H., & Broekmeulen, R. A. (2013). Determination of safety stocks in a lost sales inventory system with periodic review, positive lead-time, lot-sizing and a target fill rate. *International Journal or Production Economics, 143*(2), 440-448.

van Donselaar, K. H., & Broekmeulen, R. A. (2014). Stochastic inventory models for a single item at a single location. *Research School for Operations Management and Logistic*, 2-56.

van Eenige, M. J. (1996). *Queueing Systems with Periodic Service.* Eindhoven: Technische Universiteit Eindhoven.

van Strien, P. J. (1986). *Praktijk als wetenschap.* Assen: van Gorcum.

Walter, S. D. (1977). The power of a test for seasonality. *British Journal of Preventive and Social Medicine, 31*, 137-140.

Widiarta, H., Viswanathan, S., & Piplani , R. (2004). *Evaluation of top-down versus bottom-up forecasting strategies for substitutable products.* Centre for Supply Chain Management. Singapore: Nanyang Technological University.

Williams, T. M. (1984). Stock Control with Sporadic and Slow-Moving Demand. *Journal of the Operational Research Society, 35*(10), 939-948.

Yamazaki, T., Shida, K., & Kanazawa, T. (2016). An approach to establishing a method for calculating inventory. *International Journal of Production Research, 54*(8), 2320-2331.

Zhang, X. M., Hou, J. C., & Ren, H. Z. (2014). Study on dynamic slotting optimization in storage system. *Advanced Materials Research, 933*, 260-264.

Zinn, W., & Marmorstein, H. (1990). Comparing two alternative methods of determining safety stock levels: the demand and the forecast systems. *Journal of Business Logistics, 11*(1), 95-110.

Appendix