HEMA

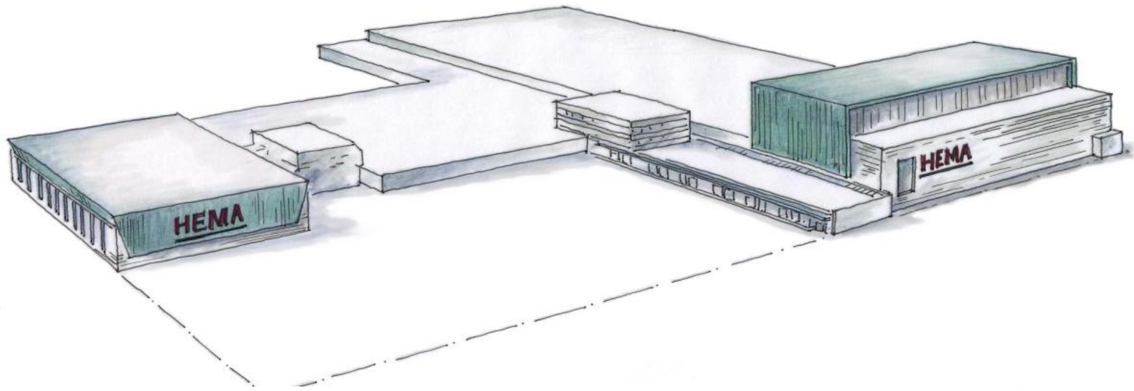
UNIVERSITY OF TWENTE.

Forecasting the required external storage space for HEMA

An approach based on historic stock levels

Master Thesis

Public Version



Author B.J. Holland

Date December, 2012

“If a man gives no thought about what is distant,
he will find sorrow near at hand.”

Confucius (551-479 BC)

"Forecasting the required external storage space for HEMA"

Bart Johan Holland
December, 2012

Master Thesis

In order to obtain the Master's degree of the study:

Industrial Engineering & Management at the University of Twente
Track: Production & Logistics Management

HEMA

UNIVERSITY OF TWENTE.

Ing. J.J. (Jordy) Harte
Manager Outbound DC

Dr. P.C. (Peter) Schuur
School of Management and Governance

Dr. ir. L.L.M. (Leo) van der Wegen
School of Management and Governance

SUMMARY

The research in this thesis was conducted at the company HEMA in the context of completing the study Industrial Engineering and Management at the University of Twente. We first present an introduction to the organization and the motivation for this research. Subsequently, the approach and results are outlined. Finally, the conclusions and recommendations are given.

Introduction

HEMA is a large retail organization which delivers to several countries in Europe. Its roots however are established in the Netherlands. Headquarters is located in Amsterdam and its distribution center is located in Utrecht.

Confidential

The following research objective is the central theme of this thesis:

"The main goal is to develop a practical model that is able to forecast the required internal and external storage capacity needed up to 12 months ahead which will allow HEMA to decide whether additional storage locations are required and if the contract with Beusichem can be cancelled."

Distribution center processes

The part of the HEMA assortment that is handled by the DC consists of a large variety of products. To enhance its manageability, it can be split into divisions (three) or categories (fourteen). Supplied products are transported on pallets in transport packages. The latter consists of one or more picking units. The smallest component is the sales unit. Products are stored on pallets or in blue totes. Storage is possible in the halls, in one of the four order picking systems or at one of the three external locations. Picked items are transported to the stores by truck.

Data collection

Because the research objective is to forecast the stock level of the distribution center, we analyze the current planning process to collect data. It turns out that there are many problems when we use the data to forecast the stock level. These are related to the product dimensions, lead time variability, the lack of standard measuring units and the unreliable forecasts of the demand chain management system. Because the correlation between the historic stock level and the historic demand is very high (ρ=0.84), it's possible to use the former to estimate future stock levels. Therefore, we can consider the stock level data as a time series. This allows the use of demand forecast techniques to generate medium- and long-term forecasts for the stock level. We identify four distinguishable patterns (i.e., seasonal, trend, cyclical and irregular) in the historic data. The medium-term forecast model should be able to capture the trend and seasonal component. The long-term forecast model should be able to capture the trend and cyclical components. We omit the irregular component because its influence is small.

Medium-term forecast

We select the mean squared error (MSE) for estimating the parameters of the available medium-term forecast models. To be able to compare between the medium-term forecast models, we select the mean absolute percentage error (MAPE) and mean absolute scaled error (MASE). Literature research concludes that *standard exponential smoothing* models are the most suitable to use for HEMA. Three methods are selected: 1) simple exponential smoothing; 2) Holt's damped method and 3) Holt-Winters' seasonal method. We test all three methods on four validation intervals with three different horizons. Parameter estimation of the models is performed by using a grid search technique. It turns out that the Holt-Winters model is the most suitable to forecast the aggregate stock level.

Forecast improvement

To improve the medium-term forecast model, we test whether forecasts made on a disaggregated level can improve the accuracy. It appears that the accumulated estimates made for the stock levels of the three divisions (i.e., fashion, hardware and various) provide better aggregate forecasts. We refer to this new model as the *combined forecast model*. Furthermore, a 95% prediction interval is added which indicates the accuracy of the estimates. Besides, a tracking signal is added which indicates whenever the forecast errors contain significant bias. A call for human intervention is given. Finally, the influence due to Easter on the stock level of the hardware division is estimated by decomposing the seasonal component. Easter is the only yearly recurring event with changing start dates. We update the seasonal factors which improves the MAPE accuracy in the weeks before.

Long-term forecast

Confidential

Determining capacity

Now we arrive at the main target of this research: to forecast the required external storage space. To be able to use the medium- and long-term forecast model, it is necessary to express the actual storage space in sales units.

Confidential

The LP-model is based on three storage properties. These are the storage period, storage costs and height classes at the locations. The model requires the number of sales units for each category as input variable. Consequently, it can also be used to make a storage division based on forecasted sales unit values. An example of a possible storage division (percentage of the total stock per category that should be stored at a certain location) for the amount of stock at week 1 of 2012 based on an empty warehouse problem is given in the table below:

Nr.	Category	DPS/ VPS	Halls/ HBW	eDC	Veem	Beus-ichem
1	Ladies' and men's apparel	x%	x%	x%	x%	x%
2	Babies and children's apparel	x%	x%	x%	x%	x%
3	Lingerie, underwear and nightwear	x%	x%	x%	x%	x%
4	Hosiery	x%	x%	x%	x%	x%
5	Leather goods, accessories, swimwear	x%	x%	x%	x%	x%
6	Interior textiles and home accessories	x%	x%	x%	x%	x%
7	House wares	x%	x%	x%	x%	x%
8	Do-it-yourself and maintenance	x%	x%	x%	x%	x%
9	Stationery, toys and Christmas	x%	x%	x%	x%	x%
10	Personal care	x%	x%	x%	x%	x%
11	Services	x%	x%	x%	x%	x%
15	Materials	x%	x%	x%	x%	x%
17	New markets	x%	x%	x%	x%	x%
30	New services	x%	x%	x%	x%	x%

SUITABLE PRODUCT DISTRIBUTION ACCORDING THE LP-MODEL FOR THE HEMA CATEGORIES IN WEEK 1 OF 2012

The storage division shows that the depot in Beusichem still is valuable for product storage although the transport costs are high compared with the eDC and the Veem.

Conclusion

Confidential

ACKNOWLEDGEMENTS

This thesis is the completion of my Master Industrial Engineering and Management at the University of Twente. Moreover, it is also a result of my research done at HEMA, the *Hollandse Eenheidsprijzen Maatschappij Amsterdam*. I really appreciate the help that so many people have given me.

First of all, I would like to thank my HEMA supervisor Jordy Harte. The many discussions I had with him considering the project helped me to gain insight in the problem I was actually solving. He provided me with the necessary resources and data needed for analysis and challenged me to stay focused and 'just ask' people. Although his agenda was very busy sometimes, he always made time to see me and help me when necessary.

Secondly, I owe my gratitude to my colleagues at the office, Mimoun Bekkaoui, Stijn van der Putte and Dick Smit. They were always available to answer one of the many questions I had. I will sincerely miss their humor and jokes. Furthermore, Bart van Harten and Jimmy Zandbergen helped me tremendously during my research. Only they knew the answer to some specific questions. It was always possible for me to make an appointment with them to discuss the topic.

Also, I thank my supervisors from the University of Twente, Peter Schuur and Leo van der Wegen for the time they spent in evaluating my reports, visiting the Distribution Center in Utrecht and supporting me in my ideas throughout the research.

Last but not least, without the help and continuous support of my parents, my girlfriend Fleur, my friends and family I would have never been able to accomplish this project. Sometimes when things looked bleak they pulled me through.

Today I finished my report and I will continue to challenge myself in the future with what I learned, it's not the end but just the start.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ANN	Artificial neural network
ARIMA	autoregressive integrated moving average
AS/RS	Automatic storage and retrieval system
CPS	Car picking system
CTU	Container terminal Utrecht
DC	Distribution center
DCM	Demand chain management
DPS	Dynamic picking system
eDC	Electronic distribution center
ETA	Estimated time of arrival
GMRAE	Geometric mean relative absolute error
HEMA	Hollandse Eenheidsprijzen Maatschappij Amsterdam
HQ	Headquarters
HBW	High bay warehouse
IT&TS	Information technology & technical services
KPI	Key performance indicator
LP	Linear programming
MA	Moving average
MAD	Mean absolute deviation
MAPE	Mean absolute percentage error
MASE	Mean absolute scaled error
MdAPE	Median absolute percentage error
MdRAE	Median relative absolute error
MRAE	Mean relative absolute error
MSE	Mean squared error
OCB	Order consolidation buffer
PU	Picking unit
RF	Radio frequency
RMSE	Root mean squared error
RMSPE	Root mean Square percentage error
SES	Simple exponential Smoothing
SKU	Stock keeping unit
SLIM	Selling less is more
SOQ	Sales order quantity
SU	Sales unit
TPP	Transport packaging
VDDC	Distribution floor
VPS	Voice picking system
WMS	Warehouse management system

1 INTRODUCTION

This report is the result of six months of research conducted at the HEMA, the *Hollandse Eenheidsprijzen Maatschappij Amsterdam.* In this thesis, a model is developed to forecast the stock level of the distribution center (DC), located in Utrecht. It can support decisions related to the total storage capacity that the DC requires. This research is a Master Thesis which finalizes the study *Industrial Engineering and Management* at the *University of Twente* with specialization *Production and Logistics Management*.

In this first chapter we indicate the central problem and present a design for solving it. Section 1.1 gives an introduction to the organization and states the research background. Subsequently, we state the latter more formally by defining the research goal in Section 1.2. The research questions that are posed to split the problem into manageable parts can be found in Section 1.3. To make sure that project stays within bounds, the research scope is made clear in Section 1.4. The methods and tools that are used to answer the research questions are presented in Section 1.5. Furthermore, the practical and theoretical significance of this research is discussed in Section 1.6. Finally, in Section 1.7 we present the structure of this thesis and introduce the following chapters.

1.1 RESEARCH BACKGROUND

HEMA is a Dutch dimestore chain. It was part of the Maxeda Company until June 2007, when it was bought by Lion Capital LLP. The chain is characterized by relative low pricing of generic house wares, which are mostly made by and for the chain itself. This is often combined with original design. While the global headquarters (HQ) is located in Amsterdam, the *non-food* products intended for the retail stores are supplied from the DC in Utrecht. Within the DC, approximately x% of the products are stored on *pallets* while the remaining x% is stored in *blue totes.* Over the past years, the organization has become much larger, more complex and diverse. Hence, in 2008 HEMA invested in a large automatic storage and retrieval system (AS/RS) to improve material handling. This high bay warehouse (HBW) provided additional storage capacity of x pallet locations. Currently, HEMA can store approximately x pallets and x totes. A detailed overview is given in Table 1-1.

	Location	# Pallets	# Totes
Int.	Halls	x	
	HBW	x	
	DPS		x
	VPS		x
Ext.	eDC	x	
	Veem	x	
	Beusichem	x	
	Total	**x**	**x**

TABLE 1-1 NUMBER OF INTERNAL AND EXTERNAL STORAGE LOCATIONS

The various internal and external locations available for storage are explained into more detail in Section 2.2.2 and Section 2.2.4.

HEMA currently supplies more than 600 retail stores in The Netherlands, Germany, Belgium, Luxembourg and France and this amount is still growing.

Confidential

In the summer of 2010, HEMA introduced the *selling less is more* (SLIM) concept. It meant that for a large number of products the prices were lowered to increase sales and assure fast market penetration (McCarthy, 2001). However, HEMA more or less became victim of the *bullwhip effect.* It occurs when order variability is amplified as products move up the supply chain (Lee, Padmanabhan, & Whang, 1997). The SLIM concept complicated forecasting and HEMA's demand chain management (DCM) system (see Section 3.2.4) overestimated expected demand significantly. It resulted in a large amount of excess stock at the DC. This is shown graphically in Figure 1-1, in which the inventory level expressed in *sales units* (see Section 2.2.2.1 for more detail) per week over the past years is plotted. The stock level was quite stable from 2006 to 2009. In week 35 of 2010 however the stock level started to increase due to the introduction of the SLIM concept.

Confidential

In 2011, HEMA was forced to rent additional storage locations in the municipalities of Beusichem and Almere. The aforementioned increase of stock continued until week 41 of 2011. HEMA's stock reduction measures started to have effect and the level started to decline again. In February 2012, the contract with the storage depot in Almere was discontinued because the lessor went bankrupt. It meant that x pallet locations were scrapped.

Confidential

Summarizing, the DC currently has three external locations (i.e., Veem, eDC and Beusichem) available for pallet storage. Considering the rent and transport costs it's questionable whether or not these are still beneficial and DC's management is unsure whether the depot in Beusichem should be kept. The contract is signed for one year and by the end of 2012 it has to be reviewed. It takes more than a year before the HBW2 is finished. Meanwhile, the storage capacity of the DC is decreased.

HEMA wants to have an indication of the DC's future stock level for the *medium-term* (i.e., 1-12 weeks) and *long-term* (i.e., 4-12 months). This indication should support the decision whether the contract with Beusichem should be discontinued or renewed. This depends on both the costs to store items at the depot and the estimated future stock levels.

1.2 RESEARCH GOAL

The purpose of this study is to gather insight in the current way of forecasting the expected stock levels at the DC and identifying possibilities for improvement. This will allow HEMA to make a better assessment of the required storage space for its inventory for the medium and long term. The following research objective is the central theme of this thesis:

"The main goal is to develop a practical model that is able to forecast the required internal and external storage capacity needed up to 12 months ahead which will allow HEMA to decide whether additional storage locations are required and if the contract with Beusichem can be cancelled."

1.3 RESEARCH QUESTIONS

To divide the research goal into manageable components, the following research questions are formulated:

1. *How is the distribution of goods across the HEMA supply chain currently organized?*
    - HEMA has a wide assortment and many products rely on the DC for distribution. Knowledge about this flow is essential and will serve as a basis for the forecast.

2. *Which data are required to make medium- and long-term forecasts?*
    - Forecast quality is always based on the accuracy of the input data. Therefore, it's important to specify which data can be used and which should be omitted.

3. *Which model is appropriate to make medium-term forecasts?*
    - Scientific literature describes several medium-term forecast methods. The most practical for HEMA with a sufficiently high quality level will be chosen.

4. *How can the medium-term forecast model be improved?*
    - Medium-term forecast are very important for HEMA. Important capacity decisions are made several weeks or months in advance. Consequently, medium-term forecasts should be easy interpretable and reliable and we will try to improve the selected model.

5. *How can HEMA management make long-term forecasts?*
    - It's necessary to make long-term forecasts because storage capacity decisions can require much time. The most suitable long-term forecast model for HEMA will be chosen based on scientific literature.

6. *How should HEMA use the medium- and long-term forecasts to determine the total required internal and external capacity and improve the product allocation?*
    - HEMA should be able to use the medium- and long-term forecast model without any problems. Furthermore, to know whether the depot in

Beusichem is still useful, we require a suitable division of the products over the locations and need to determine the part of items that's stored there.

1.4 RESEARCH SCOPE

First of all, the scope of this research is limited due to time constraints. Only six months are available to complete the research. Tradeoffs have to be made in the number of instruments used, the width and depth of the theoretical model and the number of factors considered in the study (Cooper & Schindler, 2008). In this thesis, only *non-food* articles will be taken into account. The food products flow across different supply chain processes and are therefore excluded. Consequently, there is no need to apply a deterioration rate to the items (Pakkala & Achary, 1992).

Not all locations that are available for pallet storage in the DC will be part of the research. For example, a considerable number of pallet locations are reserved for returned goods from the retail stores. The locations that will be part of this research are the halls, the HBW, the dynamic picking system (DPS), the voice picking system (VPS) and the external depots. These are explained into more detail in Section 2.2.2 and Section 2.2.4.

HQ is responsible for the purchase of goods and promotion campaigns. Only they can influence the height of the stock level and the DC is simply required to cope with the inevitable fluctuations. Consequently, we assume that the future stock level of the DC can only be forecasted.

1.5 RESEARCH DESIGN

In this section, we split the research questions posed in Section 1.3 in manageable sub questions. This will facilitate the conducted research and accomplishing the research goal stated in Section 1.2. We use bullet points to describe the methodology.

1. How is the distribution of goods across the HEMA supply chain currently organized?
   a. *Which products are part of the HEMA assortment?*
      - We check the internal databases.
      - We make a graphical overview of the assortment to show the different product levels.
   b. *How has HEMA organized the flow of these products across its supply chain?*
      - We check the internal databases and other internal sources.
      - We conduct semi-structured interviews with the heads of the different departments.
      - We describe the flow of the products part of the current assortment across the supply chain.

2. Which data are required to make medium- and long-term forecasts?
   a. *Which data are used in the current planning process?*
      - We conduct semi-structured interviews with the stock planners and warehouse managers.

  - We perform a case study concerning four different products to gain more insight in the planning process.
  - We make a description of the current planning process to indicate which data is used by the planners.
b. *Which problems arise when this data is used for forecasting principles?*
  - We analyze the data to assess its usefulness to forecast the DC stock level.
  - We identify and explain the problems occurred during the case study.
c. *How can the data complexity be reduced to overcome the identified problems?*
  - We reduce the data complexity by checking it for significant correlations.
  - We combine data or discard other data when we find a significant relation.
d. *Which patterns can be recognized in the historic stock data?*
  - We analyze the historic data to identify recurring patterns.
  - We conduct a semi-structured interview with a stock planner form headquarters and discuss the effect of the identified patterns.

3. Which model is appropriate to make medium term forecasts?

a. *What is a good measure for forecast quality?*
  - We perform a literature study to identify suitable quality measurement methods.
  - We select the methods that best fit the research goal.
b. *Which medium-term forecast models are available according to literature?*
  - We perform a literature study to identify suitable forecast models.
  - We make a selection of the forecast models that best fit the research goal.
c. *Which criteria are used to test the preselected forecast models?*
  - We test the forecast models on different validation intervals.
  - We test the forecast models on different time horizons.

d. *How can the required parameters for the selected forecasting methods be found?*
  - We perform a literature study to determine common methods for parameter estimation.
  - We perform a grid search to estimate the parameters that yield the best result for our data set.
e. *Which forecast model should be used to make medium-term forecasts for the aggregated stock level?*
  - We do a benchmark between the preselected forecast models to compare their performance on the aggregated data.
  - We select the model which yields the highest accuracy according the quality measurement methods for the proposed validation intervals and time horizons.

4. How can the medium-term forecast model be improved?
    a. *Does a disaggregated forecast improve the results?*
        - We repeat the forecast procedure with the same test settings for the disaggregated level.
        - We compare the disaggregated results with the aggregated results and adjust the forecast model accordingly.
    b. *How can the assessment of the forecasts made by the model be improved?*
        - We add a 95% prediction interval to the point forecast to improve managerial assessment.
    c. *Which techniques are available to identify inaccurate forecasts?*
        - We add a tracking signal to the forecast model which can notify management whenever human intervention is required.
    d. *How should the identified patterns be incorporated in the medium-term forecast model?*
        - We decompose the historic data and incorporate the parameters in the model when required.

5. How can HEMA management make long-term forecasts?
    a. *How should the identified patterns be incorporated in the long-term forecast model?*
        - We perform a literature study to identify suitable long-term forecast models.
        - We decompose the historic data and incorporate the parameters in the model when required.
    b. *Which aspects should be estimated by HEMA management?*
        - We propose a framework for making long-term forecasts by adding human judgment.

6. How should HEMA use the medium- and long-term forecasts to determine the total required internal and external capacity and improve the product allocation?
    a. *How many actual storage locations are available?*
        - We determine the fill grates of the storage locations and calculate the actual storage space.
        - We express the storage space in a singular unit.
    b. *What are the future stock levels that can be expected?*
        - We identify the bottleneck periods.
        - We use the forecast models to estimate the medium and long term stock levels.
    c. *Which storage properties determine whether a storage location is suitable for a product?*
        - We attend weekly group meetings with the decision makers.
        - We make a list of properties and discuss them.
    d. *Which model is adequate to support HEMA in this product allocation?*
        - We construct a practical model based on the identified properties that can be used by HEMA to allocate stored and supplied goods.

1.6 RESEARCH SIGNIFICANCE

According to Geurts (1999), the outcome of a research study will have *practical* or *theoretical* significance. This thesis is a practice-oriented research. Therefore, its main objective is to contribute to the knowledge of HEMA. This doesn't necessarily exclude the contribution it gives to the existing theories.

Contribution to practice: This research will focus on the inventory levels of HEMA within the next year. The organization currently lacks the knowledge about the required storage space it needs. Consequently, HEMA is unable to determine whether it requires extra storage locations. It implies that a practical forecast model will be given to HEMA to determine its required storage space for the coming months.

Contribution to theory: The theoretical framework in Chapter 4 will discuss what's already known about forecasting. The medium- and long-term forecast model will be based on these existing theories. Therefore, it can be argued that the forecast model could be generalized beyond the bounds of this study and could also be of use for other organizations with a comparable warehouse (Cooper & Schindler, 2008).

1.7 THESIS STRUCTURE

The structure of this thesis is outlined in Figure 1-2. It follows the proposed structure of the research questions. First of all, the research background, research questions and the research design are given in Chapter 1. Subsequently, the distribution of goods across the HEMA supply chain is discussed in Chapter 2. It provides a general overview of the processes in the DC which eventually lead to successful store delivery. The medium-term forecast model is solely based on the historic stock level. This is justified and explained in Chapter 3. Subsequently, several market aspects are identified which should be incorporated in the medium- and long-term forecast models.

The medium-term forecast model that is appropriate according to literature to estimate the future stock levels for the DC is selected and discussed in Chapter 4. Thereafter, in Chapter 5 we increase the accuracy of the initial model and add features that can improve managerial judgment. In Chapter 6, we develop a long-term forecast model based on historic data and expert knowledge. Thereafter, in Chapter 7 both the medium- and long-term forecast models are used to determine the number of storage locations HEMA needs. We construct a practical storage model which can support management in allocating the products. The conclusion and recommendations that result from the conducted research are presented in Chapter 8.

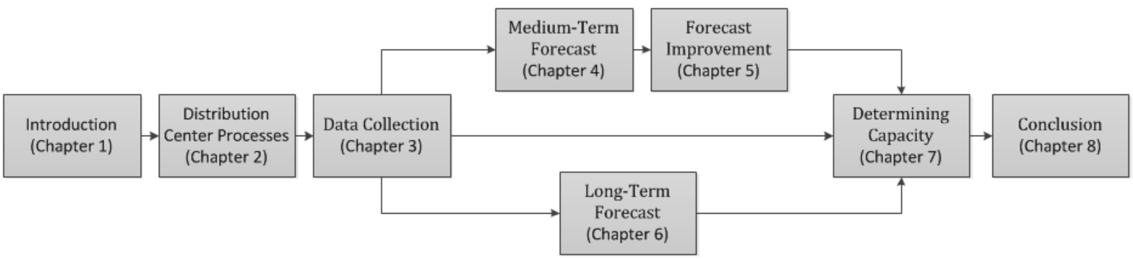


FIGURE 1-1 THESIS STRUCTURE

2 DISTRIBUTION CENTER PROCESSES

This chapter gives an overview of the type of products that HEMA offers and how they are processed by the DC. It's essential to understand this flow of goods because it defines which data is required in the following chapters. Consequently, we try to accomplish this by answering the first research question: *"How is the distribution of goods across the HEMA supply chain currently organized?"*

In Section 2.1 the items that are processed by the DC are introduced. The flow of these goods across the supply chain is discussed in Section 2.2. We explain the various storage types and methods plus the order picking and slotting strategies. Finally, we recapitulate on the main points of this chapter in Section 2.3 and answer the first research question.

2.1 PRODUCT ASSORTMENT

First of all, some definitions are needed for clarity. A *product* is defined as a type of good, while the individual stock keeping units (SKUs) are called *items* (Rouwenhorst et al., 2000). At HEMA, the customer is always put first. This statement is expressed in a wide assortment of products, low prices and a high degree of self-service in the stores. The organization distinguishes itself from other retailers by selling its own brand. HQ develops all new products but the production itself is outsourced under private label to a large number of global suppliers (HEMA, 2008). The structure of the assortment HEMA offers to its customers is shown graphically in Figure 2-1.

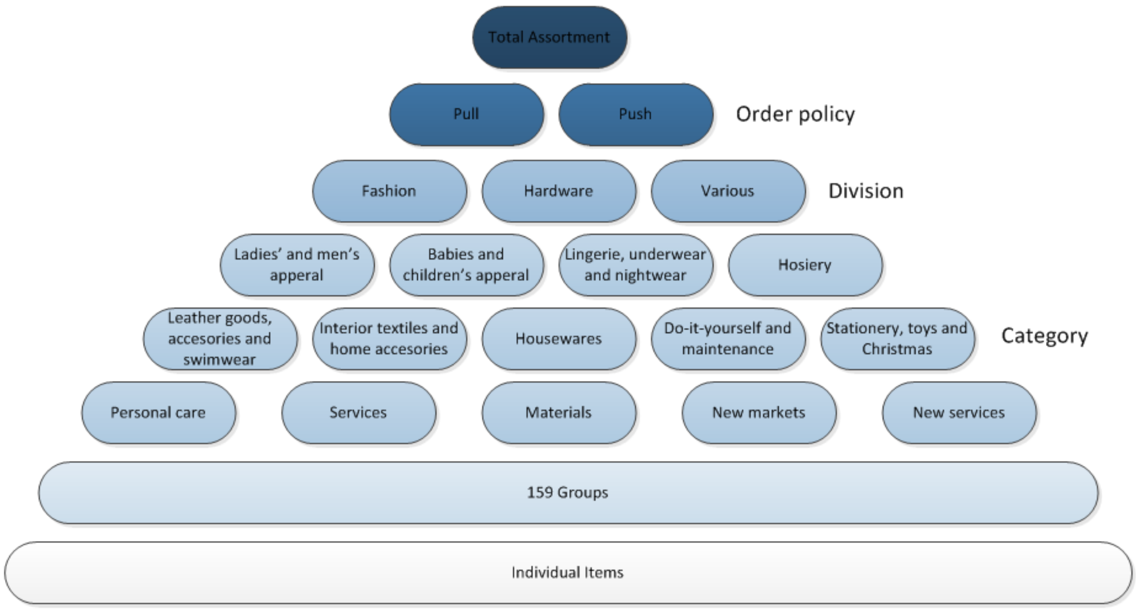


FIGURE 2-1 THE HEMA ASSORTMENT OF 2012, SPLIT INTO THREE DIFFERENT LEVELS

Generally spoken, items that are supposed to remain in store for a short time are labeled as *push*, the opposite are labeled as *pull*. They are discussed in Section 2.1.1 and Section 2.1.2. Furthermore, the assortment can be split into three *divisions*: fashion, hardware and various. However, the food division is omitted because these products are not supplied from the DC. Each division represents a certain number of *categories*: 14 of them are dependent on the DC for storage and distribution. Thereupon, these categories are split up into 159 *groups*. Each group consists of a selection of products that serve a particular consumer category or market (De Groot, 2011). We present a more detailed overview per category in Appendix I.

2.1.1 PUSH PRODUCTS

The push flow comprises seasonal and promotion products. HQ makes a forecast based on expected demand for the seasonal articles and divides the total volume over the stores based on the assortment grades and size. The push orders are more or less known in advance and can be scheduled for processing based on the required counter week (HEMA, 2008). Push orders usually involve all the stores but a limited number of products each week. HEMA has four discount promotions per year in which a rise in the push flow can be noted. Furthermore, holidays also tend to provoke an increase. We elaborate on this in Section 3.4.1

2.1.2 PULL PRODUCTS

Contrary to the planned push flow, the pull flow is driven by the cash desks at the retail stores. The daily sales are registered such that a calculation can be made based on minimum stock levels to see whether the shelf needs to be replenished. Depending on their size and assortment, each store is supplied several times per week (with a maximum of 4). A predefined order calendar assures that this is performed at the agreed time. Consequently, the outlets do not need to create any stock aside from the items on the shelf. This has a positive influence on the holding costs (HEMA, 2008).

2.2 DISTRIBUTION PROCESS

Approximately 40% of the HEMA suppliers are located in the Far East (e.g., Cambodia, China). When the products are ordered by HQ, an *estimated time of arrival* (ETA) at the DC is calculated based on the lead time and made agreements. The goods are packed into containers with the *snakepack* technique. It means that they are packed together according their article number which speeds up the container unloading process (Wegert, Roth, & Kraus, 2012). Depending on the required speed of delivery and attached costs, goods can either be transported by sea, by air or a combination of these two (i.e., sea-air transport). Sea containers arrive at the harbor of Rotterdam from where they are shipped by barge to the container terminal in Utrecht (CTU). Containers can usually stay there costless for a few days. The containers are transported to the DC by truck. Air containers arrive at Schiphol in Amsterdam from where they are shipped to the DC by truck.

The remaining 60% of the HEMA suppliers are located in Europe (e.g., Italy, Turkey). In accordance, an ETA is calculated based on lead times and made agreements. All goods are transported to the DC by truck. Products that are delivered by truck are stacked on pallets.

The DC in Utrecht makes sure the products are received and stored until they have to be transported to the stores. The departments involved in the distribution process are human resource, finance, operations, IT & TS and the operations office. An organizational chart is presented in Appendix II. The goods flow through the different departments depending on their dimensions and weight. The different paths an item can take within this process are shown schematically in the floor plan in Figure 2-2.

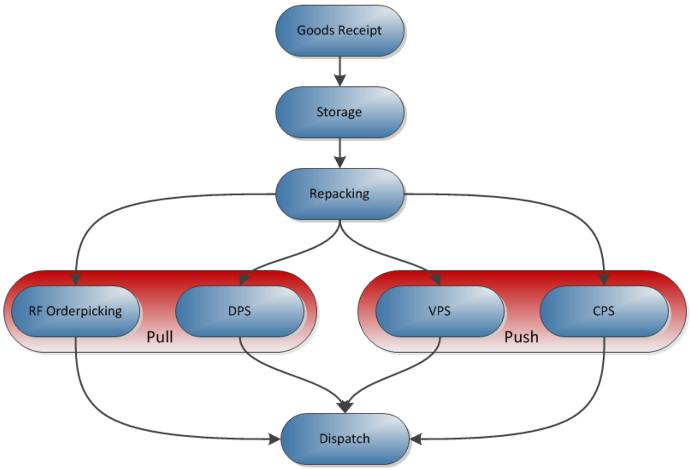


FIGURE 2-2 SCHEMATIC FLOOR PLAN OF THE DISTRIBUTION CENTER

In Section 2.2.1 we clarify the receiving process of the products. Thereafter, they inevitable have to be stored for a certain period. The various storage modes are explained in Section 2.2.2. Some products are repacked from pallets to totes (discussed in Section 2.2.3), before they are picked in the DPS or VPS (discussed in Section 2.2.4). Others products are picked straight away from pallets through radio frequency or in the CPS (also discussed in Section 2.2.4). Finally, in Section 2.2.5 we explain how the products are loaded in the trucks and distributed to the stores.

2.2.1 GOODS RECEIPT

The goods receipt department is responsible for the receiving process. Products can be delivered in three ways: 1) on pallets; 2) snakepacked in a sea container; or 3) in blue totes. Only a few suppliers (<1%) use the third option. Furthermore, products that are snakepacked in a sea container are always manually packed on pallets. The operations office estimates the number of pallets a single container will generate. Products that are already supplied on pallets only require quality and quantity checks. To simplify subsequent calculations, we will assume that *all* supplied goods to the DC arrive on pallets.

Goods receipt is responsible for storing the products at one of the pallet locations (see Table 1-1). The pallet racks have different heights however. The classes and the number of available pallet locations are shown in Table 2-1.

Class	2	3	4	5	6
Height (m.)	x	x	x	x	x
# Pallet locations	x	x	x	x	x

TABLE 2-1 HEIGHT CLASSES FOR PALLET STACKING

The height of the classes varies between 0.58 and 2.18 meter. As can be seen, the small classes are underrepresented. The height of stacking of the goods is determined based on the number of free locations per class. After a suitable location for a pallet is chosen, the transport is done using a forklift or a conveyor belt.

2.2.2 STORAGE

The storage of goods can be divided in three layers. In Section 2.2.2.1, the different types of item packing is discussed. Furthermore, pallets and totes are used to distribute the packed items within the DC or to one of the external depots. We explain them in Section 2.2.2.2. Finally, we discuss the various storage locations for pallets and totes in Section 2.2.2.3.

2.2.2.1 PACKAGING

According to Hellström and Saghir (2007), the functions that packaging must perform are manifold and complex: it has to protect, contain, preserve and communicate the product. They argue that packaging can be classified in three different levels: 1) primary packaging, the material that in direct contact with the product; 2) secondary packaging, this is designed to contain several primary packages; and 3) tertiary packaging, an assembly of a number of primary or secondary packages. At HEMA, the packaging for items has a maximum of three layers (HEMA, 2011).

First of all, the *sales unit* (SU) is comparable with primary packaging. It's in direct contact with the product and is packed in a way such that it's commercially presentable. It refers to the units purchased by the customer in the stores. To increase the stability and speed of handling the sales unit is bundled into a *picking unit* (PU). For example, if a retail order consists of ten SUs, the picking can be done ten times as fast. Consequently, stores are only allowed by HEMA to order PUs. The PU is expressed in terms of the number of SUs that it contains. Finally, the *transport packaging* (TPP) is a collection of PUs and is expressed in terms of the number of SUs it contains. It is the outermost packing layer and usually consists of a carton box. An example of a packaging structure is given in Figure 2-3.

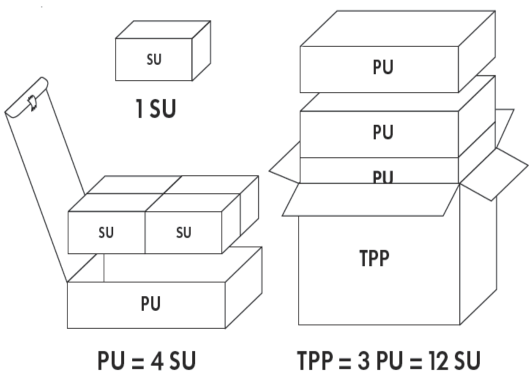


FIGURE 2-3 PACKAGING STRUCTURE EXAMPLE

The example shows one carton box (i.e., the transport packaging) wherein three picking units can be packed. Every picking unit consists of four sales units. Therefore, this TPP is composed of twelve SUs.

In the remainder of this thesis, the *selling unit* will be used to perform calculations regarding the stock level. It's the smallest available packaging structure and can give early notifications about changes in the level. Furthermore, it's easy accessible in the HEMA databases and simple to define for each product type.

2.2.2.2 TRANSPORT

HEMA uses standardized wooden EURO pallets (shown in Figure 2-4) to store the TPPs. The pallets can be transported with forklifts to their preferred location in the warehouse or depot. Bartholdi and Hackmann (2011) recommend using pallet storage whenever possible to avoid high handling and labor costs.

FIGURE 2-4 EURO PALLET

FIGURE 2-5 COLLAPSIBLE BLUE TOTE

Besides pallets, the collapsible tote (shown in Figure 2-5) is used for storage of PUs in the DPS and VPS. Both order picking systems and are explained in Section 2.2.4. When the tote is the preferred storage method for an item, the TPP has to be broken up through repacking (see Section 2.2.3) such that the separate PUs can be stored. Determining the preferred storage method for an item (i.e., pallet or tote) is discussed in Section 3.1.3.

2.2.2.3 LOCATIONS

As was shown in Table 1-1, the DC has four physical locations for internal storage (i.e., storage at the DC). The DPS and VPS are order picking systems with storage space for blue totes. We describe them in Section 2.2.4. Pallets can be stored in the high bay warehouse (HBW) and in racks in the halls. The HBW is an automatic storage and retrieval system and provides space for x pallets. The four halls (i.e., Hall1, Hall2, Hall3 and Hall4) initially provided space for x pallets. However, as discussed in Section 1.1, in summer 2012 the constructing of a new automatic storage and retrieval system (HBW2) started and Hall4 had to be demolished. This decreased the total storage space of the halls to x pallet locations.

The three rented depots are considered as external storage modes and are mainly used for two types of products: 1) those that are not in the current assortment because their selling season has ended; and 2) those that already have a (too) high coverage ratio at the distribution center. Also, products that are completely out of the HEMA assortment are usually transported to one of the external locations to remain stacked on pallets until their final use is known. Eventually, all products will return to the DC for further processing.

HEMA rents the *Veem* for a long time already. It is located in Utrecht near to the DC and provides storage space for x pallets. The Veem is mostly used for products which with a high coverage ratio. The second depot is the *eDC*. It's operational since February 2012 and also located in Utrecht. It has x pallet locations available. It often provides temporary storage for products which will return soon. Finally, HEMA rents a depot in the municipality of *Beusichem* where x pallet locations are available. Considering the larger travel distance, Beusichem is most suitable for products that require a longer storage duration (discussed in Section 7.3.1 more formally).

In summer 2011 HEMA rented a large amount of locations in the municipality of Almere. Approximately x pallets could be stored. HEMA had negotiated that for the first x pallets a fixed price and for the remaining x pallets a variable price would be paid. However, in February 2012 the lessor went bankrupt and the building was cleared.

2.2.3 REPACKING

When picking units are small enough to fit in a tote, they are sometimes repacked. Whether goods are repacked or not depends on certain criteria, these are discussed in Section 3.1.3. The products arrive from one of the halls or from the goods receipt department via the conveyor belt. The TPPs stacked on pallets are manually repacked to PUs and stored in totes. They are transported to either the DPS or the VPS (see Section 2.2.4). The warehouse management system (WMS) calculates the number of products that should be repacked based on historical figures, the number of order lines received and a forecast for the next few days (HEMA, 2008).

2.2.4 ORDER PICKING

Order picking is done using one the four following systems or techniques: 1) the dynamic picking system (DPS); 2) the volume picking system (VPS); 3) radio frequency (RF) controlled order picking; and 4) the car picking system (CPS). The preferred order picking system for a product is based on whether it fits in a tote and its order policy. It will be explained into more detail in Section 3.1.3. Push orders are picked using the VPS or CPS while pull orders are picked using the DPS or RF-controlled picking. The preferred system for an item can be found by using Table 2-2.

	Pull	**Push**
Totable	Dynamic Picking System	Volume Picking System
Non-totable	RF-Controlled Picking	Car Picking System

TABLE 2-2 PREFERRED ORDER PICKING SYSTEM FOR ITEMS BASED ON THEIR SIZE AND ORDER POLICY (SCHOLTUS, 2009)

The next sections elaborate on the four order picking systems. The DPS is discussed in Section 2.2.4.1, the VPS in Section 2.2.4.2, voice picking through radio frequency in Section 2.2.4.3 and the CPS in Section 2.2.4.4.

2.2.4.1 DYNAMIC PICKING SYSTEM

In the dynamic picking system, totes arrive on a conveyer belt from the repacking department and are stored in the tote warehouse. When required, the totes from the storage location flow to the picking area. This front is divided into 20 aisles, which are likewise sub-divided into 70 pick zones. The movement of the totes is completely automated with the use of conveyors, lifts and cranes. Order picking is done manually and the monitor at the picking station indicates which locations are to be picked from. Once the order picker has confirmed that all the picks have been completed, the tote is automatically carried away (HEMA, 2008).

2.2.4.2 VOLUME PICKING SYSTEM

The voice picking system is designed to handle push items that fit in a tote. The order picking process is similar to the DPS and makes use of the 'pick-to-light' principle. Furthermore, products that do not fit into totes but are stored in the HBW on pallets can directly picked in the VPS. The pick aisle is equipped with a carton discharge conveyer belt which carries the waste transport packs off to a carton press (HEMA, 2008).

2.2.4.3 RF-CONTROLLED PICKING

A small part of all pull PUs (7% of all picks) is not suitable to fit in a tote. They are too large, cause a poor fill rate or are too hazardous or vulnerable. Consequently, an alternative order picking method has been developed: the radio frequency controlled order picking. In this area, which is located in Hall1, the order picker doesn't use a truck but walks through the aisles with a roll container and picks the required articles. This process includes the use of barcodes, scanners and radio frequency controlled terminals.

2.2.4.4 CAR PICKING SYSTEM

Approximately 25% of the push PUs do not fit in a standard tote and can't be processed in the VPS. Therefore, they are picked in the CPS located in Hall3. In this area, the orders are picked manually by employees driving small trucks. A truck can contain four roll containers in which the PUs can be stacked. Order picking is enhanced with the help of radio frequency controlled terminals and barcode scanners. Also, a clever slotting technique is used to guarantee the best possible fit for the roll containers (HEMA, 2008). The easy stackable products are picked first and the 'ugly' ones are performed last.

2.2.5 DISPATCH

Dispatch is responsible for the loading of goods. After the totes are picked in the DPS or VPS they are transported to the *order consolidation buffer* (OCB). The totes are automatically stacked on dollies and sorted based on store number and store layout. Products arriving from the RF department or the CPS are stored in roll containers and are already sorted. After the dollies and roll containers are collected and loaded in the trucks they await further transportation, regulated by expedition.

2.3 CONCLUSION

The HEMA assortment consists of a large variety of products. Therefore, it is split up in four manageable categories. The first distinguishable level is the product order policy, which can be either push or pull. Push products usually remain in the assortment shorter than pull products. Furthermore, the push flow is planned beforehand while the pull flow is driven by sales data of the stores. The second level is the product division. Products can be of type fashion, hardware or various. The third level is the product category. The distribution center is responsible for fourteen types of these. Finally, the current assortment can be split in 159 groups.

Suppliers either deliver in containers or by truck. Goods should arrive on a beforehand estimated time (ETA). The goods receipt department is responsible for their unloading, quality and quantity check and storing them in the warehouse. For simplicity, we assume that all products enter the DC packed on pallets. This is justified because almost all products arrive on a pallet or are packed on a pallet by goods receipt. HEMA distinguishes between three different packaging structures: 1) the transport packaging (TPP); 2) the picking unit (PU); and 3) the sales unit (SU). Because the sales unit is easy accessible in the database and easy to use for calculations, we use it in this thesis to express the stock level. For transport of the PUs and TPPs, HEMA uses pallets and totes. Sometimes it's beneficial to repack products from pallets to totes to simplify the order picking. Totes with pull products are stored in the dynamic picking system (DPS) while totes with push products are stored in the voice picking system (VPS). Pallets with pull products are stored in Hall1 and picked with radio frequency (RF- order picking) while pallets with push products are stored in Hall3 and picked with the car picking system (CPS). Besides, they can also be stored internally in the HBW or Hall2 or externally at the Veem, the eDC or at the depot in Beusichem. Dispatch is responsible for loading of trucks with totes (stacked on dollies) from the DPS or VPS and the roll containers with items (from the CPS or picked through radio frequency). Expedition finally makes sure the goods are transported to the stores.

The assortment of HEMA products that depends on the DC for distribution is explained. Furthermore, their flow across the HEMA supply chain is made clear. Now it's possible to use this information and collect the required data required to forecast the stock level expressed in sales units. We do this in Chapter 3 by discussing the planning process performed by the operation office.

3 DATA COLLECTION

Most forecasting models are used to estimate future demands as a function of past data (Winston, 1994). In this chapter we will analyze and filter the data used by HEMA management in order to keep only the relevant part. We try to accomplish this by answering the second research question: "*Which data are required to make medium- and long-term forecasts?*"

To know which data are used by the operations office, the planning process is described in Section 3.1. Secondly, the difficulties encountered when we would use this data to forecast the stock level are made clear in Section 3.2. In Section 3.3 we argue that only the data of the historic stock level should be used to forecast the future levels. Subsequently, we discuss certain market aspects which should be forecasted correctly because they influence the stock level on the medium- and long-term. These could already be noted in Figure 1-1 and we elaborate on them in Section 3.4. Finally, we summarize this chapter and answer the posed research question in Section 3.5.

3.1 PLANNING PROCESS

In this section, the used methods and techniques performed by the operations office for the planning of the HEMA products across the supply chain is explained. We do this in order to acquire relevant data which could be used to forecast the stock level. The planning starts with the initial forecast (Section 3.1.1) of the supplied goods, after which the products are scheduled (Section 3.1.2) before they are picked (Section 3.1.3) and distributed to the stores (Section 3.1.4).

3.1.1 FORECASTING

The amount of supplied pallets to the DC each week mainly depends on purchase decisions taken at HQ. Purchasers generate supplier orders and enter these in the joint database. Operation office planners can extract and use this data. A supplier order consists of various products in differing quantities which are packed together into a truck or container. As stated in Section 2.2, an estimated time of arrival at the DC is calculated. However, during this lead time the actual arrival date can change due to various (unforeseen) reasons. Eventually, an unloading date is set which usually is quite reliable.

3.1.2 SCHEDULING

[*Confidential*]

3.1.3 ORDER PICKING

Planners from the operations office use as guideline that when maximal two picking units of the same product fit in a tote the transport packaging shouldn't be repacked (see Section 2.2.3). Another guideline is that when a product is supplied in large quantities it's not preferable to store them in totes. The guidelines are stated to prevent poor fill rates and a too large amount of totes in the DPS or VPS which could lead to an overload of the picking systems.

For the initial picking schedule, the planners consider the total volume of order totes (i.e., a tote with picked items intended for a store) as a fluid model. Bartholdi and Hackmann (2011) describe the flow to a single store with the following equation:

$$flow\ (\#\ totes/store/week) = \left(\frac{\#\ items/store/week}{volume/tote}\right) volume/item \qquad \textbf{3.1}$$

The actual estimation used for the picking schedule is generated by the WMS. This is a complicated calculation because every product has its own sort group. Only products which are part of the same group can be stored together in an order tote because it speeds up the unpacking at the retailers. Furthermore, differing product dimensions (see Section 3.2.1) may also prohibit some items to be stored together.

3.1.4 DISPATCHING

Orders in the DPS and the VPS are picked in order totes and stacked on dollies. Orders picked in the CPS and through radio frequency are packed in roll containers. Both are transported to dispatch and loaded in the trucks. The pull flow is clearly larger in terms of sales units than the push flow. A calculation performed with data over the past four years show the former to comprise 73% of the total volume and the latter only 27%. Therefore, pull goods are planned for delivery up to four times a week to certain (large) retailers while the push goods are delivered only once per week.

3.2 DATA COMPLEXITY

Optimally, all data for the planning process should be combined and used for the forecasting model. However, several problems concerning the product dimensions, availability of the products, the standard calculation methods and the demand chain management system impede this.

It’s hard to estimate the amount of weekly supplied products to the distribution center. We discuss this in Section 3.2.1. In Section 3.2.2 we note that the product dimensions are incorrect in many cases. This makes it difficult to use the weekly estimated volume of order totes to forecast the stock level. Subsequently, the DC consists of various departments (as explained in Section 2.2*)* and they all use their own terminology. In Section 3.2.3 we argue that it’s difficult to unify the data from the described planning process because there is no standard calculation unit. Finally, two years ago HEMA introduced their new DCM system which forecasts pull demand. However, in Section 3.2.4 we show that these forecasts aren’t always reliable. Consequently, we don’t use the generated demand forecasts by the demand chain management system in this thesis.

3.2.1 DELIVERY PROBLEMS

Not all suppliers are reliable. The ones located in the Far East usually have longer lead times compared with the ones in Europe. Furthermore, delivery data shows that the lead times of the Far East suppliers are also quite variable. Unforeseen circumstances can lead to problems with the delivery and the estimated time of arrival has to be extended.

3.2.2 PRODUCT DIMENSIONS

Headquarters is responsible that the specified product dimensions are correctly inserted in the joint database when a supplier order is given. This allows the operations office to perform calculations for the number of picks, required workforce and the transportation volume. However, errors in the stated dimensions are not uncommon. The number of times the dimensions are incorrect can be 20% for certain weeks. Though these error calculations sometimes offset (i.e., 50% of the time the dimensions of a product is underestimated and 50% of the time it's overestimated), they can seriously disturb the planned outgoing volume of order totes.

3.2.3 STANDARD UNIT

The goods receipt department doesn't consider sales units. Their *key performance indicator* (KPI) is the number of stored pallets per week. Repacking however uses the number of processed transport packagings per hour as KPI. Calculations for the picking areas are performed in picking units. This gives a better comparison between the order pickers. On the contrary, dispatch and expedition use the number of loaded and delivered dollies and roll containers per day as KPI. Consequently, there is no standard unit for making calculations and measuring performance.

3.2.4 DCM SYSTEM

Confidential

3.3 COMPLEXITY REDUCTION

Considering the problems mentioned in Section 3.2, it's very hard to combine and filter the data from the planning process and use them to forecast the stock level. However, the inbound (supplied goods) and outbound (distributed goods) flows are very similar. On average, all supplied items are distributed to the stores within three weeks. When we compare the average amount of received sales units per week in 2009 to 2012 with the average demand (shifted three weeks to the left) from the same period, a very resembling pattern is obtained. This is shown graphically in Figure 3-1. The resemblance is caused by the fact that the DC simply serves as a transfer point for most products.

Confidential

Most scientific forecasting techniques are specifically designed to forecast customer demand. We assume that the outbound flow of products distributed to the stores equals customer demand. When we compare the weekly demand with the weekly stock level two weeks before (averaged with data from 2009 to 2012), it is clear that both are highly correlated. A detailed calculation shown in Appendix III acknowledges this (i.e., ρ=0.84). In periods with low customer demand the stock level is significantly lower and vice versa.

Now we can reduce the problem complexity by using techniques designed to forecast demand. We only need the data from the historic stock level. This way, the problems mentioned in Section 3.2 are bypassed. Instead of analyzing two variable flows (i.e. inbound and outbound), only one *time series* (i.e., the historic stock level) is required to make medium- and long-term stock level forecasts. In this thesis, we consider a time series to be a collection of observations of well-defined data points obtained through repeated measurements over time (Yaffee & McGee, 2000).

3.4 DATA DECOMPOSITION

The time series of the historic stock level is subject to certain market influences. The medium- and long term forecasting model should be able to capture these variations. Therefore, it's necessary to decompose the time series. We analyze each individual component such that the time series can be projected into the future. Hanke and Reitsch (1998) state that for series measured in time periods less than a year (e.g., weekly data) each original value can be considered to be the multiplicative product of four components:

$$Y = TSCI \qquad \textbf{3.2}$$

with Y is the actual value, T is the trend component, S is the seasonal component, C is the cyclical component and I is the irregular component. We will discuss the four components in the next section and try to identify the visible patterns. We use all available historic stock level data (i.e., week 1 of 2006 to week 35 of 2012).

First of all, seasonal variations that recur every year are explained in Section 3.4.1. In Section 3.4.2 we argue that the historic stock level data exhibit a distinguishable upward trend. Furthermore, a cyclic pattern can be identified in the historic stock level. This is shown in Section 3.4.3. We omit irregular patterns and state the reason in Section 3.4.4.

3.4.1 SEASONAL DECOMPOSITION

A significant part of the HEMA assortment consists of products that incur *seasonality* each calendar year. Seasonality is defined to be the tendency of data to exhibit behavior that repeats itself each period (Yaffee & McGee, 2000). For HEMA this means that peaks in the stock level can be expected in certain weeks. These peaks will repeat themselves every year. Figure 3-2 shows the seasonal events that are of importance for HEMA.

Confidential

The seasonality percentages in Figure 3-2 are estimated by using the m*oving average* (MA) method (Hamburg & Yough, 1994). The n-period moving average, as of the end of period t, is given by:

$$\bar{y}_{t,n} = \frac{1}{n}(y_t + y_{t-1} + y_{t-2} \cdots y_{t-n+1}) \quad \textbf{3.3}$$

Where the y's are the actual observed data in the corresponding period t. We follow the notation proposed by Makridakis et al. (1998) and do a "2 x 52-MA" procedure. It means that the initial 52-period moving average is averaged through a 2-period moving average. This is necessary to make the results symmetric. The 100% line in the figure represents the yearly average stock level.

A part of HEMA's selling strategy is to do four discount promotions per year (shown in purple squares in Figure 3-2). They last for two weeks and are named 'HEMA op Hol' and 'Stapelgoed'. These are usually performed in week X, week X, week X and week X. As can be seen from the figure, the demand peaks are just before the start week because the stores must have the products on shelf the week before.

Confidential

The start of Easter varies between week 12 and 16. Therefore, it's shown in the gray arrow in Figure 3-2 to emphasize its changing start date. It has a similar influence on the peak demand as the above mentioned holidays. Table 3-1 shows the start weeks of all events that influence the stock level.

Event	Week	Event	Week
Promotion I	9-10	**Promotion III**	35-36
Easter	12-16	**Promotion IV**	43-44
Queen's Day	18	**Santa Claus**	49
Mother's Day	19	**Christmas**	51
Promotion II	21-22	**New Year**	52
Father's Day	24		

TABLE 3-1 EVENTS THROUGHOUT THE YEAR WHICH INFLUENCE THE STOCK LEVEL OF THE DISTRIBUTION CENTER

In Chapter 4 and Chapter 5 we use the seasonal data shown in Figure 3-2 and Table 3-1 to make medium-term forecasts.

3.4.2 TREND DECOMPOSITION

A *trend* exists when there is a long time increase or decrease in the data which doesn't have to be linear (Makridakis, Wheelwright, & Hyndman, 1998). It's no surprise that the historic stock level from Figure 1-1 shows an upward trend. In 2006, HEMA supplied a little over 300 stores. In 2012 this number has increased up to 600. It means that on average each week a new store was opened. Of course, it needs to be supplied.
In trend analysis the independent variable is time. There are many different trendlines that can be drawn (e.g., linear, logarithmic, exponential) through a series of data points.

Because of its widespread use and ease of understanding, we only consider linear trends. Consequently, we use linear regression and apply the *least squares* method to calculate the parameters of the equation (Hanke & Reitsch, 1998). This approach computes the line that best fits a group of points by minimizing the squared errors made. The linear trend equation is:

$$\hat{Y} = b_0 + bX \qquad \textbf{3.4}$$

with $\hat{Y}$ is the predicted trend value for Y in period X, b_0 is the value of the trend when $X = 0$ and b is the average increase in $\hat{Y}$ for each increase of one unit of X.

A trendline is most reliable when the *coefficient of determination* R^2 is at or near 1 (Hanke & Reitsch, 1998). The R^2 is a number from 0 to 1 that reveals how closely the estimated values for the trendline correspond to the actual data. The coefficient is defined as follows:

$$R^2 = 1 - \frac{\sum_t (y_t - f_t)^2}{\sum_t (y_t - \bar{y})^2} = 1 - \frac{SS_{err}}{SS_{tot}} \qquad \textbf{3.5}$$

with y_i is the actual value in period t, $\bar{y}$ is the mean of all observed value and f_t is the forecasted value with the trend equation. The residual sum of squares SS_{err} and the total sum of squares SS_{tot} are the formal notations of the numerator and denominator. According to Ozer (1985) there's no standard cut-off value for a 'low' versus 'high' R^2, but instead is at the discretion of the researcher. Therefore, we assume that the linear regression model can give predictions if $R^2 \geq 0.6$. We use the trend equation in Chapter 4 and Chapter 5 to make medium-term forecasts and in Chapter 6 to make long-term forecasts.

3.4.3 CYCLICAL DECOMPOSITION

Another effect seen in the historic data is a *cyclical* pattern. It exists when data exhibit rises and falls that are not of fixed and known period (Makridakis, Wheelwright, & Hyndman, 1998). The SLIM concept introduced in 2010 caused an increase in the stock level. This upward cycle continued until the end of 2011. At that time, the stock reduction policy started to have effect and the level turned downwards again. The cyclical component is defined as follows:

$$C = \frac{Y}{TSI} \qquad \textbf{3.6}$$

with all variables are as defined before. To identify the values, we need to remove the other components. We already stated in the introduction of this section that we omit irregular patterns (see also Section 3.4.4). Furthermore, the seasonal component in the data can be removed by applying the "2 x 52-MA" procedure (see Equation 3.3). The trend component can be eliminated by averaging out its effects. This method is referred to as the *residual method* (Hanke & Reitsch, 1998). We perform the above mentioned steps and identify the cyclical component in the data to make long-term forecasts in Chapter 6.

3.4.4 IRREGULAR DECOMPOSITION

Irregular patterns in the historic data are hard to identify. Most of the irregular component is made up of random variability. For instance, the amount of sales of some HEMA products exhibits a particularly high correlation with the weather. Umbrellas are a perfect example for this. When a rainy period is expected their demand is usually much higher than during a dry season. The same applies for sun cream during sunny days. These irregular effects have only minor influence on the stock level. Therefore, decomposition of the irregular component of the time series is omitted.

3.5 CONCLUSION

We need to collect data which can serve as basis for the medium-and long-term forecast model. Therefore, we analyze the planning process performed at the operation office. First of all, the amount of supplied pallets is forecasted. With this data, the order picking schedule is generated. After simulation of this schedule, the actual order picking is performed and the products are dispatched and expedited to the stores.

Confidential

A lot of techniques are designed to forecast demand. We show that the correlation between the historic stock level and the historic demand is very high (ρ=0.84). Therefore, we can reduce the data complexity by using demand forecasting techniques to forecast the stock level when we consider the historic level as a time series.

We decompose the time series to clarify the patterns seen in the data from Figure 1-1. First of all, a seasonal component is identified. HEMA carries out four discount promotions every year. Prior to these, a significant increase in the stock level can be seen due to a larger push volume. The same applies for the weeks prior to the Dutch holidays. A rise in stock level is seen due to the increased push and pull demand. Secondly, a linear trend is seen in the time series due to HEMA's expansion ambitions. We also identify a cyclical component, caused by the SLIM concept and the subsequent stock reduction. We omit the irregular component because its effect is only minor.

The large amount of data is reduced to a manageable form. We only need to consider the historic level as a time series to make estimates about the future level. In Chapter 4 we will choose between several quality measurement techniques and perform a literature study to identify suitable medium-term forecast models. The forecast models should be able to capture trend and seasonality. We test the models based on several settings. After estimation of the optimal parameters we will select the most accurate forecast model for making forecasts up to 12 weeks ahead. In Chapter 6 we will explain long-term forecasting and incorporate the cyclical component.

4 MEDIUM-TERM FORECAST

Forecasting is concerned with predicting the future. According to Makridakis, Wheelwright and McGee (1983), a distinction can be made between *qualitative* and *quantitative* forecasting methods. The former consists of subjective techniques based on the opinion and judgment of consumers and experts. They are appropriate form making long-term forecasts and will be discussed in Chapter 6. The latter are used to forecast demand as a function of past data and can support medium-term decision-making. This chapter focuses on quantitative techniques and we will try to answer the third research question: *"Which model is appropriate to make medium-term forecasts?"*

First of all, we need an appropriate measure of forecast quality related to error measurement (Hoover, 2009). We do this in Section 4.1, in which we discuss several measures and choose three of them. Subsequently, Silver, Pyke and Peterson (1998) reckon three distinct steps involved in statistically forecasting a time series, these are:

1. Select an appropriate underlying model of the time series pattern through time
2. Select the values for the parameters inherent in the model
3. Use the model and the parameter values to forecast future values

The sections in this chapter follow these steps. In Section 4.2 we present a literature review and select three forecast models for further analysis. Section 4.3 defines the various settings used to select the most suitable model. In Section 4.4, the optimal parameters for the proposed models are estimated according the data set. In Section 4.5 we test the models with their optimal settings and choose the most suitable model. Finally, we give the conclusion and answer the third research question in Section 4.6.

4.1 ERROR MEASUREMENT

Forecasting quality is measured by applying error-metrics. The forecast model that generates the smallest error terms is the most adequate in explaining the variability in the observations (Gardner, 1985). This is known as the *model fit*. According to Hyndman (2006), there are four types of forecast-error metrics: 1) scale-dependent metrics; 2) percentage-error metrics; 3) relative-error metrics; and 4) scale-free error metrics.

We discuss the first type of metric in Section 4.1.1. These are the oldest error measurement techniques but are still widely used. In Section 4.1.2 we explain the easier interpretable percentage-error metrics. Subsequently, relative-error metrics use a benchmark method but have wide applicability. They are presented in Section 4.1.3. The fourth technique according the taxonomy of Hyndman (2006) (scale-free error metrics) is discussed in Section 4.1.4. We select the most suitable method for measuring medium-term forecast quality in Section 4.1.5.

4.1.1 SCALE-DEPENDENT METRICS

Scale-dependent metrics are useful when comparing different forecasting methods applied to the same set of data. According to Winston (1994), the measure of variability that is often used in fitting errors to historical data is the mean squared error (MSE). It's defined as follows:

$$MSE = \frac{1}{n}\sum_{t=1}^{n}\left(y_t - f_{t-1,t}\right)^2 \qquad \textbf{4.1}$$

with $f_{t-1,t}$ is the forecast made in period $t-1$ of the expected value in period t and y_t equals the actual observed value each of n unit time periods. Although its widespread use and simplicity, the MSE isn't meaningful for assessing a method's accuracy across multiple series with different scales. Other well-known but less used methods are the mean absolute deviation (MAD) and the root mean squared error (RMSE).

4.1.2 PERCENTAGE-ERROR METRICS

Percentage errors have the advantage of being scale-independent so they are frequently used to compare forecast performance across different data sets (Hyndman & Koehler, 2006). The most commonly used metric is the mean absolute percentage error (MAPE). Likewise, it was the primary error measure in the M2 forecasting competition (Makridakis, et al., 1993). Since it's very intuitive and not affected by the magnitude of the observed value it's a commonly used measure for forecast quality. The MAPE is defined as follows:

$$MAPE = \left[\frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - f_{t-1,t}}{y_t}\right|\right] \cdot 100\% \qquad \textbf{4.2}$$

As can be seen, the MAPE is scale-independent since its unit free. It can be used to assess and compare accuracy across a range of items. Unfortunately, the MAPE has a few drawbacks. For instance, the return value will be zero when a perfect forecast is given but in regard to its upper level the MAPE has no restriction. Furthermore, according to Armstrong and Collopy (1992) the method is only relevant for ratio-scaled data (i.e., data with a meaningful zero) but isn't very appropriate for series with intermittent values. Other widely used methods, like the median absolute percentage error (MdAPE) and the root mean square percentage error (RMSPE) possess the same issues (Hyndman, 2006)

4.1.3 RELATIVE-ERROR METRICS

Relative-error metrics are also scale independent. They are characterized by the fact that the obtained error is divided by the error obtained using some benchmark forecast method. Usually, the benchmark is a *naive* method where the new forecast is simply equal to the last observation. Fildes (1992) recommends the use of the geometric mean relative absolute error (GMRAE) because it can be used to assess forecast accuracy across multiple series. It is given by:

$$GMRAE = \left[\prod_{t=1}^{n}\frac{\left|f_{t-1,t} - y_t\right|}{\left|f_{rw,t-1,t} - y_t\right|}\right]^{1/n} \qquad \textbf{4.3}$$

with $f_{rw,t-1,t}$ is the generated forecast made in period $t-1$ for y_t with a random walk as naive method. Other methods are the mean relative absolute error (MRAE) and the median relative absolute error (MdRAE). Unfortunately, relative-error metrics are flawed by the fact that they can't be used for intermittent series with small errors (Hyndman & Koehler, 2006). Therefore, use of the naive method as benchmark is no longer possible because it involves division by zero.

4.1.4 SCALE-FREE ERROR METRICS

Scale-free error metrics can be used to compare different forecast methods on a single series and to compare forecast accuracy between series. Hyndman and Koehler (2006) propose to use the mean absolute scaled error (MASE). It's defined as follows:

$$MASE = \frac{1}{n}\sum_{t=1}^{n}\left(\left|\frac{y_t - f_{t-1,t}}{\frac{1}{n-1}\sum_{i=2}^{n}|y_i - y_{i-1}|}\right|\right) \quad \textbf{4.4}$$

where the numerator is the forecast error in period t, defined as the actual value y_t minus the forecast value $f_{t-1,t}$, and the denominator is the average forecast error of the one-step naive method. A naive forecast method will realize $MASE = 1$. Forecast methods that score below this value are considered to be of higher quality and vice versa.

The MASE is well suited for intermittent series. It doesn't provide infinite or undefined values except in the irrelevant case where all historical data are equal. A closely related method is the MAD/Mean ratio proposed by Kolassa and Schütz (2007). It weights individual products and is a generalization of the MAPE. It scales the errors by the in-sample mean of the series instead of the in-sample mean absolute error. However, the main advantage of the MASE over the MAD/mean ratio is that it's more widely applicable. Furthermore, the MAD/mean ratio assumes that the mean is stationary over time, which isn't true for data which show trend or seasonality (Hyndman, 2006).

4.1.5 CHOICE OF METRIC

All error measuring metrics have their pros and cons. Consequently, it's impossible to find a single best method that suits every situation (Silver, Pyke, & Peterson, 1998). The best forecast method may vary according to the metric used (Makridakis & Hibon, 2000). Therefore, we will use three error metrics to determine forecast quality.

We will use the MSE for parameter estimation (see Section 4.4) because of its simplicity and suitability for assessing accuracy on a single series. For assessing quality *between* series, we will use both a percentage-error metric and a scale-free error metric. According to Hyndman (2006), the former is preferred because of its scale-independency and intuitiveness while the latter gives good comparisons for multistep forecasts. Consequently, we use the MAPE and the MASE to select between forecast models because they were considered best in the specified categories.

4.2 METHOD SELECTION

Box and Jenkins (1970) developed the autoregressive integrated moving average (ARIMA) models. These are fitted to time data in cases where data show evidence of non-stationary. An initial differencing step (corresponding to the integrated part of the model) can be applied to remove the non-stationary.

A methodology that has recently been proposed as a tool for forecasting is the use of artificial neural networks (ANN). A neural network is a mathematical model consists of an interconnected group of artificial neurons and can change its structure during a learning phase. While they provide a great deal of promise, they also embody much uncertainty. Researchers to date are still not certain about the effect of key factors on forecasting performance of ANNs (Zhang, Patuwo, & Hu, 1998). Consequently, the implementation of these non-intuitive methods is not very easy. This makes it harder to get management support.

A widely used forecasting technique is *standard exponential smoothing*. It can be applied to time series and was first suggested by Brown (1956). The method assigns exponentially decreasing weights to past observations over time. Gardner (1985) showed that the traditionally used ARIMA models, despite being more complicated and harder to implement, are simply a subset of standard exponential smoothing. Furthermore, the M3-competition of Makridakis (2000) featuring 3003 data sets showed that ANNs do not outperform standard exponential smoothing. Therefore, we only consider standard exponential smoothing methods in this thesis to make medium-term forecasts. They are easy to understand and implement and generate forecasts with sufficient quality.

The taxonomy first described by Pegels (1969) and Gardner (1985) and extended later by Hyndman et al. and Taylor (2003) is helpful in describing the class of standard exponential smoothing methods. This framework with the naming convention according to Gardner is presented in Table 4-1.

	Seasonality		
Trend	**N (None)**	**A (Additive)**	**M (Multiplicative)**
N (None)	N-N	N-A	N-M
A (Additive)	A-N	A-A	A-M
DA (Damped Additive)	DA-N	DA-A	DA-M
M (Multiplicative)	M-N	M-A	M-M
DM (Damped Multiplicative)	DM-N	DM-A	DM-M

TABLE 4-1 STANDARD EXPONENTIAL SMOOTHING MODELS WITH INCORPORATED TREND AND SEASONALITY COMPONENT ACCORDING TO THE NAMING CONVENTION OF GARDNER (1985)

Each method is denoted by one or two letters for the trend and one letter for seasonality. Method N-N denotes simple exponential smoothing (SES). It incorporates no trend and no seasonality (Brown R. , 1959). The other non-seasonal methods are Holt's (1957) additive trend (A-N), Gardner and McKenzie's (1985) damped additive trend (DA-N), Pegels' (1969) multiplicative trend (M-N) and Taylor's (2003) damped multiplicative trend (DM-N). All seasonal methods are extensions of Winters' (1960) method.

Additive seasonal components are preferred when the variations are constant across the series, while the multiplicative method is preferred when the variations are changing proportional to the level of the series (Makridakis, Wheelwright, & Hyndman, 1998). Kalekar (2004) used both additive and multiplicative models and he shows that the latter are superior in all cases.

We will explore three exponential smoothing methods and try to identify which is most suitable to forecast the stock level. Because of its ease of use, simple exponential smoothing (N-N) is the first selected model. Furthermore, Fildes (2001) acknowledges that in aggregate selection, it is difficult to beat the damped-trend version of standard exponential smoothing. Consequently, the second selected model is Holt's method with a damped additive trend (DA-N). In addition, we showed in Section 3.4.1 that the historic data is subject to seasonality. Therefore, the well-known Holt-Winters multiplicative seasonal method (A-M) is the third selected model.

4.3 FORECAST SETTINGS

To choose between the three forecast models and select the most suitable for HEMA we need to review them. We use an estimation interval of four years (i.e., historic data from 2006 to 2009) to estimate the optimal parameters of all models because the Holt-Winters requires at least this time for initialization (Silver, Pyke, & Peterson, 1998). Optimization is done by minimizing the MSE in the estimation interval.

We select four periods for validation of the forecasts. In order to have a reliable sample, each validation period consists of 16 weeks (four months). The first validation interval is week 1 to week 16 of 2010. To determine the predicted loss of forecast accuracy due to the SLIM concept, week 36 to week 51 of 2010 is the second validation interval. Because the Christmas season always shows a peak in the stock level, we choose week 36 to week 51 of 2011 as the third validation interval to see whether the models can cope with this effect. The fourth validation interval is week 20 to week 35 of 2012 because these are the last weeks for which historic data is available. We use the MAPE and MASE techniques to compare the results between the models during the validation interval. This is shown in Section 4.5.

The accuracy of the different methods may depend upon the length of the forecasting horizon (Makridakis & Hibon, 2000). Consequently, we choose three different forecast horizons. When selecting between forecast models, its common practice to use a one week horizon ($\tau = 1$). Furthermore, we use an eight week horizon ($\tau = 8$) as intermediate value. Finally, we use a twelve week horizon ($\tau = 12$) because Armstrong (2001) reckons this to be the maximum number of periods for which accurate forecasts can be made. It's also in accordance with the desired horizon set by HEMA's management for medium-term forecasts.

We make forecasts until the horizon value exceeds the last week of the validation period and compare the average MAPE and MASE values of the three models for each horizon. The error metrics, forecast models and test aspects are summarized in Table 4-2.

Validation intervals	Forecast methods	Error metrics	Horizons
Week 1 – 16 of 2010	SES	MSE	1 week
Week 36 – 51 of 2010	Holt's damped	MAPE	8 weeks
Week 36 – 51 of 2011	Holt-Winters	MASE	12 weeks
Week 20 – 35 of 2012			

TABLE 4-2 SUMMARY OF THE USED INTERVALS, ERROR METRICS AND HORIZONS FOR TESTING THE THREE FORECAST METHODS

We give an example to explain the validation method by making forecasts in the first validation interval with a 12 week horizon. For all three forecasts methods, the optimal parameters are estimated based upon the historic data from week 1 of 2006 to week 52 of 2009 by minimizing the MSE. We use the parameter settings to make a forecast for the next 12 weeks and calculate the MAPE and MASE values. Afterwards, the actual stock level value of week 1 of 2010 is added and the parameters are updated accordingly. New forecasts are made and the MAPE and MASE values are calculated again. The actual value of week 2 is added and the parameters are updated again. This procedure is repeated until week 5 of 2010 because one week later the current week plus the horizon value exceeds the validation period (i.e., week 16). We calculate the average MAPE and MASE value in the validation period for each model and compare the results.

4.4 PARAMETER ESTIMATION

Parameters are estimated by performing a *grid search* (Gardner, 2006). With this technique, we try all possible parameter combinations and select the set which results in the lowest MSE. The optimal values in the different weeks of the validation interval parameters may vary. In the next subsections, only the optimal parameter values for the first week of the validation interval are shown in the tables.

In Section 4.4.1 we explain simple exponential smoothing and estimate the optimal parameter values. Holt's damped method is the subject of Section 4.4.2. Finally, in Section 4.4.3 we discuss the Holt-Winters method and estimate the parameters.

4.4.1 SIMPLE EXPONENTIAL SMOOTHING

Simple exponential smoothing is one of the most widely used statistical methods for short- and medium-term forecasting. Recent observations are given relatively more weight in forecasting than the older observations which makes it an affective procedure when the underlying demand model is composed of level and random components (Winston, 1994). On the downside however, if the data contains a trend this method will always lag behind and outliers can have significant impact (Silver, Pyke, & Peterson, 1998). However, simple exponential smoothing is capable at handling these trends to a certain degree because it keeps reference to the old data. The underlying model for the expected stock level y_t in period t is as follows:

$$y_t = a + \varepsilon_t \qquad \textbf{4.5}$$

with a is the level estimate and ε_t is an independent, normally distributed variable which causes random errors with a mean of 0 and a constant variance σ^2. Updating of the model is done using the following procedure:

$$\hat{a}_t = \alpha x_t + (1-\alpha)\hat{a}_{t-1} \qquad \textbf{4.6}$$

where α is a smoothing constant that satisfies $0 < \alpha < 1$ and $\hat{a}_t$ is the new estimate of the level in period t. If stable predictions with smoothed random variation are required then a small value of α is required. On the other hand, if a rapid response to a real change in the pattern of observations is desired, a large value of α is appropriate (Yaffee & McGee, 2000). Initialization of the method is done by using the stock level of week 1 in 2006 as the first estimate for a. The forecast $f_{t,t+\tau}$, made at the end of period t for any future period $t+\tau$ is:

$$f_{t,t+\tau} = \hat{a}_t \qquad \textbf{4.7}$$

When simple exponential smoothing is applied to the historic data, (very) high values for parameter α are found when we minimize the MSE. A summary of the results is shown in Table 4-3.

Interval start	1 week horizon	8 week horizon	12 week horizon
Week 1, 2010	$\alpha = 1.00$	$\alpha = 0.95$	$\alpha = 0.81$
Week 36, 2010	$\alpha = 1.00$	$\alpha = 0.98$	$\alpha = 0.87$
Week 36, 2011	$\alpha = 1.00$	$\alpha = 1.00$	$\alpha = 1.00$
Week 20, 2012	$\alpha = 1.00$	$\alpha = 1.00$	$\alpha = 0.93$

TABLE 4-3 PARAMETER ESTIMATION FOR SIMPLE EXPONENTIAL SMOOTHING

According to Gardner (1985), a value of α between 0.01 and 0.30 is quite reasonable. The α value in this case is well above the reckoned upper limit. It implies that simple exponential smoothing probably isn't suitable for this time series.

4.4.2 HOLT'S DAMPED

When time series contains a trend, Holt's method often yields good forecasts (Winston, 1994). It's an extension of simple exponential smoothing by adding a growth factor to adjust for the trend. The basic underlying demand model can be described with the following formula:

$$y_t = a + bt + \varepsilon_t \qquad \textbf{4.8}$$

where the level a and error term ε_t are defined as before and the factor b refers to the estimated trend t periods ahead. We use the formulas suggested by Brown (1963) to obtain the initial values for a and b named $\hat{a}_0$ and $\hat{b}_0$:

$$\hat{a}_0 = \frac{6}{n(n+1)}\sum_t tx_t + \frac{2(2n-1)}{n(n+1)}\sum_t y \qquad \textbf{4.9}$$

and:

$$\hat{b}_0 = \frac{12}{n(n^2+1)}\sum_t tx_t + \frac{6}{n(n+1)}\sum_t y_t \qquad \textbf{4.10}$$

We use week 1 to week 4 of 2006 (i.e. one month) to find the initial base level and trend. However, the trend is rather erratic and Holt's (1957) original method with its linear forecast function is criticized for tending to overshoot the data beyond the short term (Taylor, 2003). Therefore, Gardner and McKenzie (1985) address this problem by including an extra parameter ϕ in Holt's original model to dampen the projected trend. The updating formulas are:

$$\hat{a}_t = \alpha y_t + (1-\alpha)\left(\hat{a}_{t-1} + \phi \hat{b}_{t-1}\right) \quad \textbf{4.11}$$

$$\hat{b}_t = \beta(\hat{a}_t - \hat{a}_{t-1}) + (1-\beta)\phi \hat{b}_{t-1} \quad \textbf{4.12}$$

where both α and β are smoothing constants for determining the updated trend and level. Consequently, large weights will result in more rapid changes in the components and vice versa. If $0 < \phi < 1$, the trend is damped and if $\phi = 1$, the trend is linear. If $\phi > 1$, the trend is exponential and can potentially provide erroneous results (Gardner & McKenzie, 1985). Forecasting is done using the following formula:

$$f_{t,t+\tau} = \hat{a}_t + \sum_{j=1}^{\tau} \phi^i \hat{b}_t \quad \textbf{4.13}$$

When we apply Equation 4.9 and 4.10, the following results are found: $\hat{a}_0 = X$ and $\hat{b}_0 = X$. We perform a simultaneous grid search again to find the optimal parameters. The results are shown in Table 4-4.

Interval start	1 week horizon	8 week horizon	12 week horizon
Week 1, 2010	$\alpha = 1.00, \beta = 1.00, \phi = 0.10$	$\alpha = 0.95, \beta = 0.09, \phi = 0.07$	$\alpha = 0.12, \beta = 0.74, \phi = 0.15$
Week 36, 2010	$\alpha = 1.00, \beta = 1.00, \phi = 0.11$	$\alpha = 0.98, \beta = 0.03$ $\phi = 0.19$	$\alpha = 0.86, \beta = 0.31, \phi = 0.21$
Week 36, 2011	$\alpha = 1.00, \beta = 1.00, \phi = 0.11$	$\alpha = 0.98, \beta = 0.13, \phi = 0.21$	$\alpha = 1.00, \beta = 0.21, \phi = 0.09$
Week 20, 2012	$\alpha = 1.00, \beta = 0.94, \phi = 0.07$	$\alpha = 0.95, \beta = 0.07, \phi = 0.62$	$\alpha = 1.00, \beta = 0.06, \phi = 0.42$

TABLE 4-4 PARAMETER ESTIMATION FOR HOLT'S DAMPED METHOD

Again, very high values of α are found for every setting. The dampening factor ϕ is also high for some settings. But because the trend factor β is low for all 8 and 12 week settings, the dampening won't have much effect. It implies that Holt's damped model probably won't provide satisfactory results.

4.4.3 HOLT-WINTERS METHOD

Many time series exhibit a certain form of seasonality. As shown in Section 3.4.1, the DC stock level is subject to seasonal influences. Therefore, to accommodate for both trend and seasonality, Winters (1960) method adds a seasonal parameter to Holt's original model. This Holt-Winters model is suitable for forecasting univariate time series in presence of outliers. It attempts to best fit a smoothing constant, trend constant and seasonal constant to past data (Silver, Pyke, & Peterson, 1998). The underlying demand model is as follows:

$$y_t = (a + bt)S_t + \varepsilon_t \qquad \textbf{4.14}$$

where S_t is the seasonal coefficient for period t while the other variables are as defined before. The season is of length P periods. Because we assume the effects to be yearly, we use $P = 52$ for a weekly forecast interval. Like in Section 3.4.1, we perform a "2 x 52-MA" (see Equation 3.3) to estimate the initial seasonal coefficients. The parameters are updated according the following three equations:

$$\hat{a}_t = \alpha(y_t/\hat{S}_{t-P}) + (1-\alpha)\big(\hat{a}_{t-1} + \hat{b}_{t-1}\big) \qquad \textbf{4.15}$$

$$\hat{b}_t = \beta(\hat{a}_t - \hat{a}_{t-1}) + (1-\beta)\hat{b}_{t-1} \qquad \textbf{4.16}$$

$$\hat{S}_t = \gamma(x_t/\hat{a}_t) + (1-\gamma)\hat{S}_{t-P} \qquad \textbf{4.17}$$

where α, β and γ are three smoothing constants that all lie between 0 and 1. If $f_{t,t+\tau}$ is the forecast made at the end of period t for the demand in period $t + \tau$ and $\hat{S}_{t+\tau-P}$ is the most recent estimate of the seasonal index for period $t + \tau$, then forecasting is done with:

$$f_{t,t+\tau} = (\hat{a}_t + \tau\hat{b}_t)\hat{S}_{t+\tau-P} \qquad \textbf{4.18}$$

We use Equation 4.9 and 4.10 of Brown (1963) again to estimate the initial base level and trend and find: $\hat{a}_0 = X$ and $\hat{b}_0 = X$. As before, we perform a simultaneous grid search to find the optimal values for the smoothing parameters. We always use the last two years of data of the estimation interval to determine the optimal updating parameters (i.e., α, β and γ). The years before are used to estimate the weekly seasonal coefficients (i.e., S_t)

Because a full grid search is rather time consuming for the Holt-Winters method, Silver et al. (1998) suggest a range of values in which the optimal values can be found (i.e., $0.02 < \alpha < 0.50$, $0.005 < \beta < 0.176$ and $0.05 < \gamma < 0.50$). They are used as lower and upper limit of the grid search. We add another constraint for stability purposes because the value of α should be kept at least twice as high as the value of β (McClain & Thomas, 1973). The grid search is programmed in VBA for Excel and the code is shown Appendix IV. The results of the grid search are shown in Table 4-5.

Interval start	1 week horizon	8 week horizon	12 week horizon
Week 1, 2010	$\alpha = 0.50, \beta = 0.01, \gamma = 0.50$	$\alpha = 0.13, \beta = 0.05, \gamma = 0.50$	$\alpha = 0.11, \beta = 0.05, \gamma = 0.50$
Week 36, 2010	$\alpha = 0.50, \beta = 0.08, \gamma = 0.50$	$\alpha = 0.50, \beta = 0.02, \gamma = 0.50$	$\alpha = 0.50, \beta = 0.02, \gamma = 0.50$
Week 36, 2011	$\alpha = 0.50, \beta = 0.12, \gamma = 0.50$	$\alpha = 0.50, \beta = 0.03, \gamma = 0.48$	$\alpha = 0.43, \beta = 0.04, \gamma = 0.29$
Week 20, 2012	$\alpha = 0.50, \beta = 0.09, \gamma = 0.40$	$\alpha = 0.47, \beta = 0.08, \gamma = 0.46$	$\alpha = 0.41, \beta = 0.08, \gamma = 0.26$

TABLE 4-5 PARAMETER ESTIMATION HOLT-WINTERS METHOD

As can be seen, the optimal α, β and γ vary for each setting. The α value is equal to the upper bound set by Silver et al. (1998) for all 1 week forecasts. For larger horizons it situates between the bounds. The γ value is equal to the upper bound for the older forecasts because less data is available. When the estimation interval increases, more seasonal factors become available and γ decreases.

4.5 FORECAST RESULTS

Now that the optimal parameters for the methods are selected, we can make forecasts in the validation intervals. As stated in Section 4.3, we measure the forecast performance of the three models by their averaged MAPE and MASE scores. The results are shown in Table 4-6.

Method	Horizon	Error	Week 1, 2010	Week 36, 2010	Week 36, 2011	Week 20, 2012
SES	1 week	MAPE	2.88%	2.72%	1.91%	3.02%
		MASE	1.00	1.00	1.00	1.00
	8 weeks	MAPE	3.67%	6.05%	5.51%	4.99%
		MASE	0.99	1.00	1.00	1.00
	12 weeks	MAPE	4.54%	9.62%	**5.12%**	5.14%
		MASE	0.95	1.05	**1.00**	0.95
Holt's damped	1 week	MAPE	2.93%	**2.68%**	1,81%	**2.97%**
		MASE	1.02	**0.99**	0.95	**0.98**
	8 weeks	MAPE	3.67%	**6.04%**	5.49%	5.05%
		MASE	0.99	**1.00**	1.00	1.01
	12 weeks	MAPE	3.65%	**8.63%**	5.15%	5.12%
		MASE	0.75	**0.94**	1.01	0.94
Holt-Winters	1 week	MAPE	**2.66%**	4.49%	**1.65%**	3.11%
		MASE	**0.93**	1.64	**0.87**	1.02
	8 weeks	MAPE	**2.55%**	7.92%	**4.23%**	**4.40%**
		MASE	**0.70**	1.31	**0.79**	**0.88**
	12 weeks	MAPE	**2.73%**	10.76%	5.72%	**5.01%**
		MASE	**0.57**	1.17	1.13	**0.91**

TABLE 4-6 FORECAST RESULTS FOR THE N-N, DA-N AND A-M METHOD BASED ON THEIR MAPE AND MASE SCORES FORE THREE TIME HORIZONS AND FOUR VALIDATION PERIODS

The forecast with the lowest MAPE and MASE scores for every combination is shown in blue. The Holt-Winters method scores the best on overall performance. However, the method has difficulties to forecast the cyclical influences due to the SLIM concept (see the second validation period, starting in week 36 of 2010). This applies to Holt's method in a lesser extent, but none of the three forecast models is able to capture this cyclical component satisfactorily. We show this graphically in Figure 4-1, in which we compare three forecasts with a horizon of 12 weeks made at the start of week 36 in 2010 with the actual values.

Confidential

The sudden difference with the actual stock level is very noticeable. The Holt-Winters method is able to capture the seasonality quite well, but not the cyclical influence. Holt's method provides a slightly better forecast because it estimates a small upward trend, but it's also lagging behind. Obviously, simple exponential smoothing just forecasts a straight line.

4.6 CONCLUSION

Perfect forecasts are nonexistent. Therefore, a measure for the deviation is required. Four types of error measurement are available: 1) scale-dependent metrics; 2) percentage-error metrics; 3) relative-error metrics; and 4) scale-free error metrics. All have their pros and cons. We use the mean squared error (MSE) for parameter estimation because of its widespread use and simplicity. We use the mean absolute percentage error (MAPE) and mean absolute scaled error (MASE) to compare between forecast models because both are scale independent and easy to comprehend.

A lot of research is done in the field of forecasting techniques. Examples include autoregressive integrated moving average (ARIMA) models and artificial neural networks (ANN). However, these don't outperform the much older standard exponential smoothing methods. Therefore, we prefer the latter in this research due to their simplicity and capability of providing accurate forecasts. According to the taxonomy of Gardner, fifteen standard exponential smoothing methods are available. We select three methods for further research: 1) simple exponential smoothing (N-N); 2) Holt's damped method (DA-N); and 3) Holt-Winters seasonal multiplicative method (A-M).

We use four validation intervals (i.e., week 1 to 16 of 2010, week 36 to 51 of 2010, week 36 to 51 of 2011 and week 20 to 35 of 2012) and three horizons (i.e., 1 week, 8 weeks and 12 weeks) to test the accuracy of the three models. All previous historic data is used for the estimation interval.

We optimize the parameters by performing a grid search by minimizing the MSE. The α value for simple exponential smoothing and Holt's damped method are quite high. It indicates that both methods probably won't provide satisfying results. The parameters for Holt-Winter's method are constrained to decrease the estimation time.

We use the three models to make forecasts with the three horizons for the four validation intervals. Their average MAPE and MASE scores are calculated and compared for every combination. The Holt-Winters model clearly outperforms the other models, except for the second validation interval. All three models have difficulties to make accurate forecasts in this interval due to the introduction of the SLIM concept.

The Holt-Winters model shows the best overall performance and is the most suitable to make medium-term forecasts for the aggregated stock level. However, some inaccuracies are still present because the cyclical component is difficult to forecast. In Chapter 5 we try to improve the quality of the forecasts generated by the Holt-Winters model. We test whether a forecast made on a more disaggregate level could provide better results. A confidence interval is added which improves managerial judgment. We add a tracking signal is to warn management when the forecasts start to deviate and human intervention is required. Finally, we make seasonal adjustments to the forecasts in the Easter period.

5 FORECAST IMPROVEMENT

The three models tested in the previous chapter make forecasts based on aggregated historic data. Because the SLIM concept caused an upward cycle, which was proven hard to forecast, we try to improve the accuracy of the Holt-Winters model in this chapter. This is done by answering the fourth research question: "*How can the medium-term forecast model be improved?*"

It might be possible that a more accurate forecast can be made by disaggregating the stock data. We test this hypothesis in Section 5.1 by disaggregating the data (analogous to Figure 2-1) to the division level and the category level and comparing the outcomes with Table 4-6. In Section 5.2 we try to improve managerial judgment by adding a 95% prediction interval to the point forecast. Furthermore, we add a tracking signal to the forecast model in Section 5.3 to prevent bias. This will allow HEMA management to intervene in time when forecasts start to deviate due to a cyclical effect. In Section 3.4.1 we argued that the stock level in the Easter period lack accuracy because it has shifting start weeks. We adjust for this in Section 5.4 Finally, in Section 5.5 we summarize the preceding sections and introduce the next chapter.

5.1 DISAGGREGATED FORECASTING

Armstrong (2001) argues that disaggregated data are noisier than the aggregates constructed from them and appear harder to forecast. Hugos (2011) reckons that aggregate forecasts are more accurate than for segments. The variance is smaller because the extremes cancel out. Kahn (1998) acknowledges the fact that aggregated data at the highest level is smoother than lower level data. This corresponds to lower percentage differences about the mean compared with disaggregated time series. However, he also notes that top-level data exhibits more randomness than lower level data. Makridakis (1986) adds to this that managers should concentrate their efforts on forecasting aggregated series and delegate the task of predicting the disaggregated series to their subordinates.

On the other hand, many researchers do believe that disaggregation can lead to better forecasts. To be able to fully comprehend and predict the fluctuations in the stock level caused by unknown variability, every stock keeping unit should be analyzed separately in terms of forecasting (Zotteri, Kalchschmidt, & Caniato, 2005). Therefore, the stock level over the past years should be completely disaggregated and reviewed for each HEMA product. This hypothesis is supported by Martin and Witt (1989), who acknowledge that separating and disaggregation of the data could produce better forecasts.

The total assortment of HEMA however is far too large to estimate all the significant parameters. Hendry and Hubrich (2005) suggest in these cases to include disaggregate information or variables in the aggregated model opposed to using only lagged aggregate information. Dekker et al. (2004) propose to make forecasts at an in-between level. It means that products with similar seasonal patterns should be combined to increase the accuracy.

We use the model of Zotteri et al. (2005), shown in Figure 5-1, to test whether disaggregating the stock level can improve the forecast quality. It suggests that the optimal forecast level is a tradeoff between the ability to capture variability (which is high at the disaggregated level) and the ability to manage variability (which is high at the aggregated level).

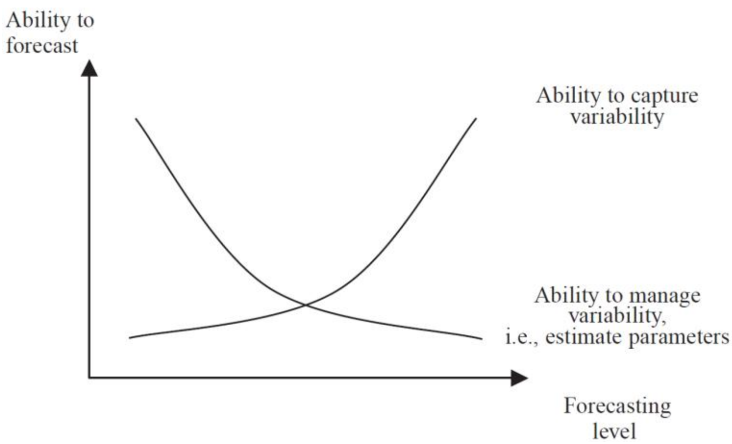


FIGURE 5-1 RELATION BETWEEN THE FORECAST ABILITY AND FORECAST LEVEL (ZOTTERI, KALCHSCHMIDT, & CANIATO, 2005)

When we consider the HEMA assortment as shown in Figure 2-1, it seems that forecasts made for the *category* level and *division* level will give promising results. They are near the trade-off point in the center of Figure 5-1. We test this hypothesis for category stock disaggregation in Section 5.1.1 and for division stock disaggregation in Section 5.1.2. We omit disaggregation to the push and pull level. In accordance with Zotteri et al. (2005), we assume that it would be too difficult to capture all variability. Likewise, disaggregation to the product group or individual level is omitted because it's too hard to estimate all relevant parameters.

5.1.1 CATEGORY LEVEL DISAGGREGATION

HEMA's assortment contains fourteen categories that are dependent on the DC for distribution. They are shown in Table 5-1 for convenience. A detailed overview of the HEMA assortment is presented in Appendix I.

Nr.	Category	Nr.	Category
1	Ladies' and men's apparel	**8**	Do-it-yourself and maintenance
2	Babies and children's apparel	**9**	Stationery, toys and Christmas
3	Lingerie, underwear and nightwear	**10**	Personal care
4	Hosiery	**11**	Services
5	Leather goods, accessories and swimwear	**15**	Materials
6	Interior textiles and home accessories	**17**	New markets
7	House wares	**30**	New services

TABLE 5-1 PRODUCT CATEGORIES IN THE HEMA ASSORTMENT THAT ARE STORED AT THE DISTRIBUTION CENTER

The historic stock of the categories from 2006 to 2012 is graphically shown in Figure 5-2. To improve the clarity of the figure, the stock levels are deseasonalized with a "2 x 52-MA" method (see Equation 3.3).

Confidential

The cyclic effect due to the SLIM concept is clearly visible in the figure. Furthermore, we note a distinct upward trend in fashion. The next section will elaborate on this.

We only use the Holt-Winters model to forecast the disaggregated stock levels. It was shown in Section 4.5 that the model produces reliable forecasts on the aggregated level. Therefore, we assume that it also will provide good results for the disaggregated category data.

The test settings are as defined in Section 4-3. We estimate the optimal parameters for the Holt-Winters method again by performing a grid search. The fourteen disaggregated forecasts are summed and compared with the actual aggregate values. The average MAPE and MASE scores for each combination are shown in Table 5-2.

Horizon	Error	Week 1, 2010	Week 36, 2010	Week 36, 2011	Week 20, 2012
1 Week	MAPE	4.12%	6.01%	1.99%	3.08%
	MASE	1.18	1.56	1.12	1.01
8 Weeks	MAPE	2.78%	13.12%	5.82%	5.12%
	MASE	0.76	2.28	1.09	1.03
12 Weeks	MAPE	5.09%	15.28%	3.41%	6.23%
	MASE	1.12	1.62	0.89	1.23

TABLE 5-2 DISAGGREGATED FORECAST AT THE CATEGORY LEVEL WITH MAPE AND MASE ERRORS FOR TWO TIME HORIZONS AND FOUR DIFFERENT START WEEKS

When we compare the MAPE and MASE errors with the values from Table 4-6, it can be seen that disaggregating the data to the category level doesn't improve the forecast accuracy. Therefore, we do not conduct any further research at this level.

5.1.2 DIVISION LEVEL DISAGGREGATION

HEMA's assortment can be divided in three divisions that are all dependent on the DC for distribution. These are fashion, hardware and various. The historic stock of the divisions from 2006 to 2012 is graphically shown in Figure 5-3.

Confidential

As can be seen from the figure, fashion is more subject to seasonal influences than the other two. Hardware division only experiences strong seasonality during the Christmas season. The fashion stock shows an upward trend over the past six years. Various has neither trend nor seasonality patterns, it simply renders around a base level. Even the introduction of the SLIM concept had no effect for the stock level of these products.

As stated in the above paragraph, Figure 5-3 shows that the fashion stock shows strong seasonality influences. Therefore, we assume that the Holt-Winters model will provide the most accurate forecasts. Similarly, the various stock isn't subject to seasonal influences and has no trend. Therefore, we assume that simple exponential smoothing will provide the most accurate forecasts.

It's unclear however which of the three models provides the best hardware stock forecasts. Therefore, we test which model is preferred by using the settings from Section 4.3. For convenience, we only use a 12 week horizon. A grid search is used again to estimate the optimal parameters. The results are shown in Table 5-3.

Method	Horizon	Error	Week 1, 2010	Week 36, 2010	Week 36, 2011	Week 20, 2012
SES	12 weeks	MAPE	5.12%	9.61%	7.64%	4.80%
		MASE	1.02	1.03	1.01	0.98
Holt's damped	12 weeks	MAPE	4.89%	9.50%	**7.44%**	4.68%
		MASE	0.88	1.01	**0.96**	0.93
Holt-Winters	12 weeks	MAPE	**3.64%**	**5.03%**	8.67%	**4.63%**
		MASE	**0.65**	**0.49**	1.29	**0.89**

TABLE 5-3 FORECAST RESULTS FOR THE N-N, DA-N AND A-M METHOD BASED ON THEIR MAPE AND MASE SCORES FORE A 12 WEEK TIME HORIZON AND FOUR VALIDATION PERIODS

The forecast with the lowest MAPE and MASE scores for every validation interval is shown in blue. As can be seen, the Holt-Winters model clearly outperforms the other two. Therefore, we use it to make forecasts for the hardware division. For clarity, we summarize the used models to forecast the division stock in Table 5-4. We refer to this new model comprised of the three methods to generate forecasts as the *combined forecast model.*

Division	Forecast Model
Fashion	Holt-Winters
Hardware	Holt-Winters
Various	Simple exponential smoothing

TABLE 5-4 COMPONENTS OF THE COMBINED FORECAST MODEL

We use the settings as defined in Section 4-3 to test the combined forecast model. The optimal parameters are estimated using a grid search. We sum the three disaggregated forecasts on the four validation periods for the specified horizons and calculate the average MAPE and MASE values. The results are shown in Table 5-5.

Horizon	Error	Week 1, 2010	Week 36, 2010	Week 36, 2011	Week 20, 2012
1 Week	MAPE	2.80%	4.13%	**1.17%**	3.19%
	MASE	0.98	1.51	**0.61**	1.04
8 Weeks	MAPE	**2.41%**	8.82%	**4.03%**	**4.28%**
	MASE	**0.68**	1.46	**0.76**	**0.86**
12 Weeks	MAPE	**2.71%**	13.69%	6.50%	**3.87%**
	MASE	**0.56**	1.46	1.24	**0.71**

TABLE 5-5 DISAGGREGATED FORECAST AT DIVISION LEVEL WITH MAPE AND MASE ERRORS FOR TWO TIME HORIZONS AND FOUR DIFFERENT START WEEKS

When the MAPE and MASE values are compared with the scores from Table 4-6, it can be seen that disaggregating the data to divisions has a positive effect on the forecast accuracy for the 8 and 12 week horizons. It has a negative effect on forecasts made with a 1 week horizon. The blue numbers indicate the forecasts for which the accuracy has improved compared with the aggregated level.

In the remainder of this thesis, forecasts are made with a combination of simple exponential smoothing and the Holt-Winters model (i.e., the combined forecast model). The new model requires four input variables (e.g., the historic stock levels of the three divisions) and a specified horizon. The made forecasts are summed to obtain the aggregated stock level forecast.

5.2 PREDICTION INTERVALS

Until now, all made forecasts are in fact estimates of the middle of a range of possible values that the actual stock level could be. However, it's possible to accompany the point-forecast with a *prediction interval*. This gives a range of values which the actual stock level could take with relative high probability. It shows how much uncertainty is associated which each individual forecast. We assume the forecast errors to be uncorrelated and normally distributed. Then Makridakis et al. (1998) reckon the 95% prediction interval to be:

$$f_t \pm 1.96\hat{\sigma} \tag{5.1}$$

with f_t is the forecast made for period t and $\hat{\sigma}$ is an estimate of the standard deviation of the forecast distribution. The root mean squared error (RMSE, see Section 4.1.1) which is the square root of the MSE, can be used as an estimate for the standard deviation σ.

We use a forecast made at the start of week 1 of 2012 as an example to demonstrate the 95% prediction interval. The RMSE is calculated for all t-period forecasts made in the estimation interval (i.e., week 1 of 2006 to week 52 of 2011), with t set equal to the horizon value. In this example the horizon is set to 12 weeks. Table 5-6 shows the output.

Week	Lower Bound	Point Forecast	Upper Bound	Actual Level
1	x	x	x	x
2	x	x	x	x
3	x	x	x	x
4	x	x	x	x
5	x	x	x	x
6	x	x	x	x
7	x	x	x	x
8	x	x	x	x
9	x	x	x	x
10	x	x	x	x
11	x	x	x	x
12	x	x	x	x

TABLE 5-6 EXAMPLE OF A 12 WEEK HORIZON POINT-FORECAST WITH 95% CONFIDENCE BOUNDS MADE AT THE START OF WEEK 1 OF 2012

In this example all the actual values comfortably lie between the confidence bounds. The added 95% prediction interval to the made point-forecasts improves HEMA's assessment of the forecasted medium-term stock level. Because the confidence bounds are relatively wide, it's clear that in certain scenarios the aggregated level could increase or decrease quickly due to cyclical effects.

5.3 ADDING TRACKING SIGNAL

When forecasts are substantially above or below their actual values for a long period, they are subject to certain *bias* (Silver, Pyke, & Peterson, 1998). In general, the forecast error e_t should fluctuate around zero. A graph can indicate whether they show positive or negative trend. In Figure 5-4 we illustrate the bias of the combined forecast model in the first validation interval (i.e., week 1 to week 16 of 2010) with a horizon of 1 week. The figure shows that the made forecasts by the combined model do not possess any bias during this interval. The weekly forecast errors fluctuate around zero.

[*Confidential*]

When the combined forecast model is used to make forecasts for the second validation interval (i.e., week 36 to week 51 of 2010), the errors do show significant bias. Due to the cyclical effect, the estimates from week 36 to week 45 are consistently too low. This is shown in Figure 5-5.

[*Confidential*]

Harrison and Davies (1964) suggest the use of a cumulative sum technique to monitor the bias in a forecasting procedure. The cumulative sum of errors U_t is computed recursively using:

$$U_t = \frac{C_{t-1} + y_t - f_{t-1,t}}{\sqrt{MSE_t}} \quad \textbf{5.2}$$

where C_{t-1} is the cumulative sum of forecast errors as at the end of period $t-1$, $y_t - f_{t-1,t}$ is the one-period-ahead forecast error and MSE_t is the MSE at the end of period t. If the forecasting procedure is unbiased, U_t should fluctuate around zero.

Silver et al. (1998) distinguish between two types of corrective actions when bias is present. The first is to increase the values of the smoothing constants in the hope that the form of the underlying model is still correct. The second type of corrective action is more substantive. It involves human intervention to make significant changes in the model itself. We opt to use the former because it's flexible and reliable. Silver et al. (1998) use the following formula to define sufficient bias:

$$|C_t| > k\sqrt{MSE} \quad \textbf{5.3}$$

where k is some threshold value. Usually, $k = 4$ is considered to be plausible. When we use this tracking signal, a call for intervention would already be given in week 38 of 2010.

We add this tracking signal to the combined forecast model by keeping track of previous one-week forecasts. When sufficient bias is present, a signal is given and HEMA management is informed. If the effect is temporary, management could override the forecasts for a few periods. On the other hand, a more permanent effect can be handled by manually adjusting the smoothing constants.

5.4 SEASONAL ADJUSTMENT

The weekly seasonality factors are easy to predict because the holidays and discount promotions are in the same week(s) every year. However, Easter is the exception to the rule. We try to improve the forecasts in the Easter period by using the historic data from 2006 to 2012. The start weeks of Easter are given in Table 5-7.

2006	2007	2008	2009	2010	2011	2012
Week 15	Week 14	Week 12	Week 15	Week 13	Week 16	Week 14

TABLE 5-7 SHIFTING START WEEK OF EASTER FROM 2006 TO 2012

It turns out the stock level of the hardware stock is significantly higher for three consecutive weeks. This peak lasts until three weeks before the start of Easter. For instance, when Easter is in week 15, the peak is expected in week 10 to week 12. We call this the *Easter season.* Easter is always in the period from week 12 to week 16. Therefore, we need to improve the seasonality factors for hardware in week 7 to week 13.

We use Easter 2012 as an example. Thus, the Easter season is week 9 to week 11. A forecast with a 6 week horizon, made at the start of week 9, gives a MAPE value of 3.20%. To improve the historic seasonality factors, we decompose the seasonality factors from week 7 to week 13 of 2006 to 2011. We take the average of all seasonality factors from week 9 to week 11 when they are part of the Easter season in that specific year.

Consequently, we take the average of all seasonality factors from week 7, week 8, week 12 and week 13 when they are not part of the Easter season in that specific year. We use the new averaged seasonality factors to calculate the new stock levels of hardware. The results are given in Table 5-8.

Week	Old factor	Old estimate	New factor	New estimate	Actual level
9	x	x	x	x	x
10	x	x	x	x	x
11	x	x	x	x	x
12	x	x	x	x	x
13	x	x	x	x	x
14	x	x	x	x	x

TABLE 5-8 UPDATED ESTIMATES FOR THE EASTER SEASON

The MAPE value decreases to 2.45%. To improve short and medium term forecasts for the coming years, the Easter dates of 2013 to 2020 are implemented in the model. The new factors will be estimated automatically based on the historic data.

5.5 CONCLUSION

Aggregation of time series data is a common way to improve the forecast accuracy. This enables the forecast model to capture all variabilities. Various authors however propose to disaggregate the data. Therefore, we test whether a disaggregate forecast for the DC stock level is more accurate than the aggregate forecasts made in Chapter 4. We disaggregate to the division level and category level. The test settings are kept the same as in Section 4.2. It turns out that disaggregation of the data to the category level has a negative effect on the forecast accuracy. However, disaggregation to the division level has a positive effect on the forecast accuracy. Nearly all MAPE and MASE values are improved. Fashion and hardware stock is forecasted with the Holt-Winters method, various stock is forecasted with simple exponential smoothing. The three forecasts are summed to obtain the aggregate forecast. This combined forecast model will be used in the remainder of this thesis to make medium-term forecasts.

To increase managerial judgment of the made forecasts, a 95% prediction interval is added. The confidence bounds are relatively wide because it takes previous cyclical effects into account.t

When forecasts are substantially above or below the actual values, some bias is present. We use a cumulative sum technique to monitor the forecast bias. When we analyze the generated errors made by the combined forecast model in the second validation interval, we notice a large bias due to the SLIM concept. Future bias can be prevented by the use of the added tracking signal. Management is warned and they have the option to override the forecasts for a short time or alter the smoothing parameters.

Because the start of Easter varies throughout the year, its seasonal influence is hard to predict. Therefore, we decompose the seasonal components in the Easter season. Adjustments to the model results to a MAPE decrease from 3.20% to 2.45% in 2012.

Now that the model for medium-term forecasts is finished, the focus shifts to long-term forecasting. In Chapter 6 we develop a model that can support HEMA management in long-term forecasting. This enables decision making for storage capacity up to one year ahead.

6 LONG-TERM FORECAST

In the two preceding chapters we have made a forecast model for the medium-term (i.e., up to 12 weeks) and improved it by adding some features. However, HEMA is also interested in long-term forecasts (i.e., up to one year ahead). This supports decision making about the required capacity well in advance. We accomplish this goal by answering the fifth research question: "*How can HEMA management make long-term forecasts?*"

Armstrong (2001) argues that when long term forecasting is the goal, incorporation of managerial judgment usually is essential. Furthermore, he reckons that forecasts combined with expert input are never *less* accurate. In 30 empirical comparisons, the reduction in ex ante errors for equally weighted combined forecasts averaged about 12.5%. This hypothesis is supported by Fildes, Goodwin, Lawrence and Nikolopoulos (2009), who conclude from extensive research that adjusting forecasts by incorporating human judgment can improve the accuracy significantly.

According to Hanke and Reitsch (1998), long-term forecasts should be made by decomposition of the time series and examining the trend and cyclical component. We define the historic trend of the stock level in Section 6.1. We decompose the cyclical effect due to the SLIM concept in Section 6.2 and argue that human judgment is required to correct for future effects. Finally, Section 6.3 concludes the chapter.

6.1 TREND ESTIMATION

We make an estimate of the trend in the historic data from 2006 to 2012. Based on Equation 3.4, we obtain the following regression line for the aggregated stock level:

Confidential

As discussed in Section 3.4.2, we need to calculate the coefficient of determination to assess the credibility of the trend line. The residual sum of squares is 4.97×10^{15} and the total sum of squares is 1.48×10^{16}. Subsequently, Equation 3.5 gives the following:

$$R^2 = 1 - \frac{4.97 \times 10^{15}}{1.48 \times 10^{16}} = 0.66 \qquad \textbf{6.1}$$

According to classic regression analysis, it means that 66% of the change in the stock level is explained by a change in the time variable t. In other words, the progressing of time explains for 66% the increase in stock. We assume that the linear trend equation is a good predictor because $R^2 \geq 0.6$. We use the trend estimation in Chapter 7 to make long-term forecasts.

6.2 CYCLICAL ESTIMATION

We make an estimation of the cyclical variation in the stock level from 2006 to 2012. In total, 347 weeks of data are available. We deseasonalize the data first by applying the "2 x 52-MA" procedure (see Equation 3.3) and remove the seasonal component. Subsequently, we use Equation 3.6 to calculate the weekly cyclical component and express it as a fraction. This implies simply the division of the actual stock level for a certain week by the expected trend value, calculated with Equation 6.1. Figure 6-2 shows the cyclical chart that is developed by performing this method. The linear trend equation is represented as 1.00 (i.e., the base line).

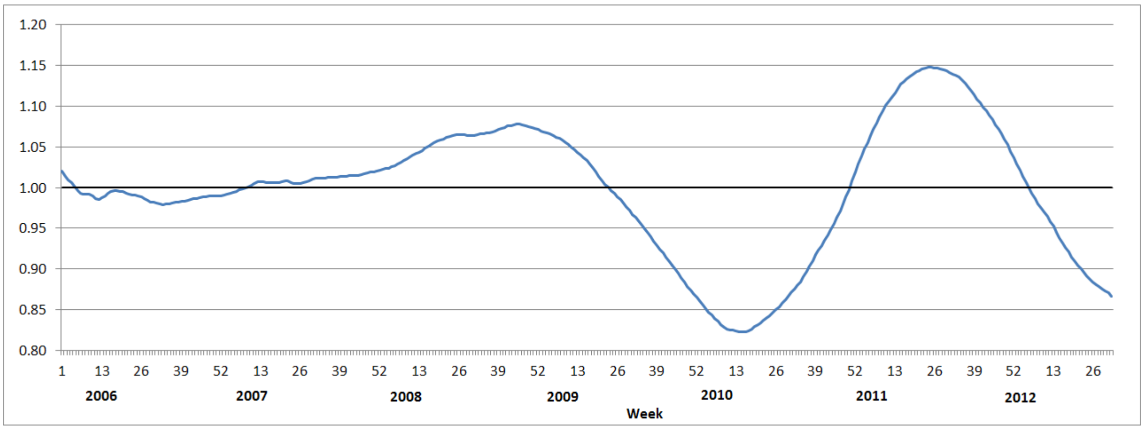


FIGURE 6-1 CYCLICAL CHART FOR 2006-2012

The figure shows that the time series cycles. The maximum and minimum cyclical factors are 1.15 and 0.85. The cycle follows the decisions taken by HQ (i.e., the introduction of the SLIM concept and the stock reduction policy). Up to 2009 however the maximum cyclical factor was only 1.07. In Chapter 7 we use these estimations to make long-term forecasts.

It's very hard to estimate the cyclical component up to a year into the future (Hanke & Reitsch, 1998). It can only be done by HEMA management because expert judgment is necessary. No forecast solely based on historic data is able to make predictions of this component.

6.3 CONCLUSION

Confidential

7 DETERMINING CAPACITY

In this chapter we intend to answer the final research question: *"How should HEMA use the medium- and long-term forecasts to determine the total required internal and external capacity and improve the product allocation?"* When this is clear, HEMA can make the decision whether the current storage capacity is sufficient.

In the previous chapters, the storage space is given in either pallets or totes. We unify these two in Section 7.1 by expressing the storage space in sales units. Subsequently, in Section 7.2 we use the medium- and long-term forecast models to forecast the expected stock levels for the bottleneck periods in the next year.

In Section 7.3, we state the properties for determining a suitable storage location for the products. As we will explain in this section, HEMA makes its storage decisions quite intuitively. We argue that a better product allocation can decrease the incurred storage costs. Therefore, we construct a model that can support HEMA in picking a suitable location in Section 7.4. Finally, we present the conclusion and introduce the last chapter in Section 7.5.

7.1 ACTUAL STORAGE SPACE

The actual storage space is less than stated in Section 1-1. Some of the locations are subject to fill rate constraints. We present these in Section 7.1.1 and recalculate the actual storage space. Subsequently, to compare the storage space with the output of the forecast models, we convert the pallet and tote locations to sales units in Section 7.1.2.

7.1.1 FILL RATES

As shown in Table 1-1, HEMA has approximately x pallet and x tote locations. However, the actual amount of pallets and totes that can be stored is less. The halls, the HBW, the DPS and the VPS require free locations to avoid jams and allow for movement of the goods. The halls shouldn't be filled more than 95%. The HBW, DPS and VPS have a fill rate of 90%. External locations can be filled up to 99%. Therefore, we multiply the numbers from Table 1-1 with these percentages. The results are presented in Table 7-1.

	Location	# Pallets	# Totes
Int.	Halls	x	x
	HBW	x	x
	DPS	x	x
	VPS	x	x
Ext.	eDC	x	x
	Veem	x	x
	Beusichem	x	x
	Total	x	x

TABLE 7-1 ACTUAL STORAGE SPACE OF THE INTERNAL AND EXTERNAL LOCATIONS

The actual number of available pallet locations equals x while the actual number of available tote locations equals x. Subsequently, these numbers should be converted to sales units for comparison with the medium- and long-term forecasts.

7.1.2 STORAGE SPACE IN SALES UNITS

Confidential

7.2 FUTURE STOCK LEVEL

As said before, the historic data shows that the maximum stock level in a given year usually occurs during the Christmas season. Hence, we only need to consider these weeks to determine the capacity bottlenecks. We use the medium-term forecast model to estimate the stock levels during the Christmas season of 2012 in Section 7.2.1. In Section 7.2.2 we follow the same procedure but use the long-term forecast model for the Christmas season of 2013. We discuss the implications of both forecasts in Section 7.2.3.

7.2.1 MEDIUM TERM FORECAST

Historic data is available until week 35 of 2012. We use the combined forecast model to forecast the expected stock levels in week 36 to week 47 of 2012. This is equivalent to a 12 week horizon forecast. The results are shown in Table 7-4.

Week	Various	Hardware	Fashion	Total (LB)	Total (Av.)	Total (UB)
36	x	x	x	x	x	x
37	x	x	x	x	x	x
38	x	x	x	x	x	x
39	x	x	x	x	x	x
40	x	x	x	x	x	x
41	x	x	x	x	x	x
42	x	x	x	x	x	x
43	x	x	x	x	x	x
44	x	x	x	x	x	x
45	x	x	x	x	x	x
46	x	x	x	x	x	x
47	x	x	x	x	x	x

TABLE 7-2 FORECAST MADE WITH THE COMBINED FORECAST MODEL FOR THE STOCK LEVEL IN SALES UNITS DURING THE CHRISTMAS SEASON OF 2012

Confidential

7.2.2 LONG TERM FORECAST

Confidential

7.2.3 IMPLICATIONS

Confidential

7.3 STORAGE PROPERTIES

In this section we discuss three storage properties that determine the most suitable location for a product from a given category. We assume that the warehouse is *empty* and all products require a storage location. First of all, we use the samples of Section 7.1.2 (i.e., the actual number of sales units stored on pallets and in totes as in week 16 of 2012) to determine the average number of sales units that's stored on a pallet or in a tote for each category. The results are shown in Table 7-8.

Nr.	Category	SUs/pallet	SUs/tote
1	Ladies' and men's apparel	x	x
2	Babies and children's apparel	x	x
3	Lingerie, underwear and nightwear	x	x
4	Hosiery	x	x
5	Leather goods, accessories and swimwear	x	x
6	Interior textiles and home accessories	x	x
7	House wares	x	x
8	Do-it-yourself and maintenance	x	x
9	Stationery, toys and Christmas	x	x
10	Personal care	x	x
11	Services	x	x
15	Materials	x	x
17	New markets	x	x
30	New services	x	x

TABLE 7-3 AVERAGE NUMBER OF SALES UNITS PER PALLET AND TOTE FOR EACH CATEGORY

These values play a role in determining the values of the three storage properties (explained in the following subsections). The average time period that the products stay in storage is discussed in Section 7.3.1. Subsequently, in Section 7.3.2 the different height classes for pallet storage are described. This is an important property because the number of height classes is not evenly distributed per location. Finally, in Section 7.3.3 we calculate the accumulated costs to store a sales unit based on the chosen location, height class and required storage time.

7.3.1 STORAGE PERIOD

The first property that determines the most suitable location is the expected period that a product stays in storage. We use the *turnover ratio* to get an indication of this value. This is the number of times that the inventory cycles through the warehouse in a year (Ballou, 2000). The turnover ratio can be calculated using the following formula:

$$Turnover = \frac{Yearly\ demand}{Average\ yearly\ inventory} \tag{7.1}$$

We calculate the 2011 inventory turnover ratio for the fourteen categories. The yearly demand is the sum of all weekly demand. We calculate the average yearly inventory by taking the average of all weekly inventory levels. A company wants its turnover ratios to be high, because that means that it has to carry fewer inventory and spend less money on inventory carrying costs (Stickney, Veil, & Schipper, 2009). We also estimate the number of weeks an item is on stock by dividing 52 (i.e., the number of weeks per year) by the turnover ratio. The results of the turnover calculations are shown in Table 7-9.

Nr.	Category	Average inventory	Yearly demand	Turnover ratio	Turnover weeks
1	Ladies' and men's apparel	x	x	x	x
2	Babies and children's apparel	x	x	x	x
3	Lingerie, underwear and nightwear	x	x	x	x
4	Hosiery	x	x	x	x
5	Leather goods, accessories and swimwear	x	x	x	x
6	Interior textiles and home accessories	x	x	x	x
7	House wares	x	x	x	x
8	Do-it-yourself and maintenance	x	x	x	x
9	Stationery, toys and Christmas	x	x	x	x
10	Personal care	x	x	x	x
11	Services	x	x	x	x
15	Materials	x	x	x	x
17	New markets	x	x	x	x
30	New services	x	x	x	x

TABLE 7-4 TURNOVER RATIOS FOR THE STOCK LEVEL OF THE FOURTEEN HEMA CATEGORIES

[*Confidential*]

In Section 7.3.3 we use the turnover ratios of the categories to estimate the costs to store a sales unit.

7.3.2 STORAGE HEIGHT

HEMA uses various height classes for pallet storage. These vary from 0.58 meter (class 2) to 2.18 meter (class 6). Because only the number of pallet locations is known, we assume a uniform distribution of the number of sales units that can be stored per centimeter of height. The height classes and the number of sales units that can be stored at each location are presented in Table 7-10.

Class Height	2 0.58m	3 0.98m	4 1.38m	5 1.78m	6 2.18m	Total
Halls/HBW	x	x	x	x	x	x
eDC	x	x	x	x	x	x
Veem	x	x	x	x	x	x
Beusich.	x	x	x	x	x	x
Total	x	x	x	x	x	x

TABLE 7-5 NUMBER OF SALES UNITS THAT CAN BE STORED AT EACH HEIGHT CLASS

Subsequently, we need to determine the average pallet height (required in Section 7.3.3 to determine the storage costs) by applying linear interpolation. We take the dot product of the two vectors that represent the height in meters and the total number of sales units per class. We divide by the total number of sales units and get the following result:

$$\frac{(0.58m, \cdots 2.18m) \cdot (121{,}401, \cdots 6{,}604{,}199)}{28{,}197{,}550} = 1.63m$$

Thus, the average pallet height is set to 1.63 meter.

Confidential

7.3.3 STORAGE COSTS

A few weeks prior to the estimated time of arrival of a new product, management from the operations office and goods receipt department discuss the most suitable storage location. This decision is made based on the product dimensions and the time it is required on stock. It’s a matter of human judgment without any cost calculations done beforehand. However, we argue that it’s possible to calculate the incurred costs to store the products.

We assume that a variable price is paid to store a single pallet on a single location for one week. For convenience, this price is a mixture of fixed and variable costs (e.g., rent, internal transport, electricity, service and cleaning). The calculated storage costs are given in the second column of Table 7-11.

Furthermore, the travel distances to the external locations are known. The Veem and the eDC are nearby (i.e., 3 kilometers and 2 kilometers), but the depot in Beusichem requires 35 kilometers of travel. We make the assumptions that a truck makes a round trip (starting and ending at the DC in Utrecht) to a single depot and that it drives on average 50 kilometers per hour, is fully loaded (i.e., 33 pallets) and requires one hour for loading and unloading. We include the gas, lease, tax insurance, maintenance and driver tariffs in the variable costs. The encountered transport costs (to and from the locations taken together) are shown in the third column of Table 7-11.

Location	Storage costs	Transport costs
DPS/VPS	€x.xx/tote/week	-
Halls/HBW	€ x.xx /pallet/week	-
eDC	€ x.xx /pallet/week	€ x.xx /pallet
Veem	€ x.xx /pallet/week	€ x.xx /pallet
Beusichem	€ x.xx /pallet/week	€ x.xx /pallet

TABLE 7-6 STORAGE AND TRANSPORT COSTS PER LOCATION FOR A PALLET OR TOTE

The cost calculation shows that Beusichem is the most suitable location to store pallets. However, it also has the highest transport costs. Storage at the DC requires no transport and is favored when the products have a short turnover time.

We use the average number of sales units per pallet and per tote from Table 7-8, the turnover in days from Table 7-9 and the storage and transport costs from Table 7-11 to determine the average costs to store a sales unit at a location during the estimated duration of stay. For the pallet locations we assume an average height of 1.63 meters. The results (multiplied by 1,000 for clarity) are shown in Table 7-12. Comparison can only be done horizontally since the turnover periods per category are different. The storage locations with the lowest costs are made (dark) blue.

Nr.	Category	DPS/ VPS	Halls/ HBW	eDC	Veem	Beus-ichem
1	Ladies' and men's apparel	€ x	€ x	€ x	€ x	€ x
2	Babies and children's apparel	€ x	€ x	€ x	€ x	€ x
3	Lingerie, underwear and nightwear	€ x	€ x	€ x	€ x	€ x
4	Hosiery	€ x	€ x	€ x	€ x	€ x
5	Leather goods, accessories and swimwear	€ x	€ x	€ x	€ x	€ x
6	Interior textiles and home accessories	€ x	€ x	€ x	€ x	€ x
7	House wares	€ x	€ x	€ x	€ x	€ x
8	Do-it-yourself and maintenance	€ x	€ x	€ x	€ x	€ x
9	Stationery, toys and Christmas	€ x	€ x	€ x	€ x	€ x
10	Personal care	€ x	€ x	€ x	€ x	€ x
11	Services	€ x	€ x	€ x	€ x	€ x
15	Materials	€ x	€ x	€ x	€ x	€ x
17	New markets	€ x	€ x	€ x	€ x	€ x
30	New services	€ x	€ x	€ x	€ x	€ x

TABLE 7-7 COSTS TO STORE A SALES UNIT DURING THE ESTIMATED TURNOVER PERIOD (RESULTS ARE MULTIPLIED BY 1,000) ON A 1.63 METER PALLET OR IN A TOTE

To illustrate the cost calculations we use stock from category 10 (i.e., personal care), stored at the Veem, as an example. According to Table 7-8, a pallet can store X sales units of this item. Furthermore, according to Table 7-9, the expected turnover period is 4 weeks. Finally, the storage costs per week for a pallet are €x.xx and the transport costs to the Veem are €x.xx per pallet. We get the following cost calculation:

Confidential

We use the cost calculations in Section 7.4 for the storage model.

7.4 STORAGE MODEL

Now that we formulated the properties that determine a suitable storage location and calculated the storage cots, it is still quite complex to make the actual allocation. One has to use Table 7-12 and the cost column of every location (except the DPS/VPS column) unfolds five times due to the five height classes. To support HEMA management in making the correct location choice we introduce a simple storage model in this section. This model is programmed in Excel as a *linear programming* (LP) model (Dantzig, 1998). Furthermore, it's an *empty warehouse model*. We assume that the warehouse is empty and that all products arrive at once and should be stored at a suitable location.

The input variable for the model is the number of sales units on stock per category. The **objective** of the model is to minimize the total costs incurred by storing the items. As stated in Section 7.3.1, we assume that all products are stored for the duration of their corresponding category turnover period. We use the following constraints:

- **Constraint 1:** Each sales unit should be allocated to a storage location (i.e., the DPS/VPS or a height class at a depot)
- **Constraint 2:** Each storage location can be filled up to maximal 100%
- **Constraint 3:** Each storage location should at least be filled with the indicated percentage part of stock for each category

The third constraint requires some explanation. The model allows HEMA to indicate that a fixed percentage of stock of a category must be stored at a certain location. For instance, during the Christmas season, stock of category 9 (i.e., stationery, toys and Christmas) requires a high availability and items should be stored at the distribution center. Management could decide that at least 40% of the products should be stored at the DPS/VPS and at least 40% at the Halls/HBW. The model then tries to minimize the total costs with this constraint in mind. We define the LP in the following *standard form*:

Indices:

c	$c = \{1 \cdots 14\}$ which represents the 14 categories
l	$l = \{1 \cdots 21\}$ which represents the 21 locations with $\{1\}$ = DPS/VPS, $\{2 \cdots 6\}$ = Halls/HBW class $2 \cdots 6$, $\{7 \cdots 11\}$ = eDC class $2 \cdots 6$, $\{12 \cdots 16\}$ = Veem class $2 \cdots 6$ and $\{17 \cdots 21\}$ = Beusichem class $2 \cdots 6$

Parameters:

$stock_c$	Stock in sales units per category c that has to be stored
$cost_{c,l}$	Cost to store all items of category c at location l

$capacity_l$	Capacity in sales units per location l
$fill\_grade_{c,l}$	Minimum fill grade per category c of sales units that has to be stored at location l

Variable:

$X_{c,l}$ Fraction of all items from category c that's stored at location l

Objective:

$$min \sum_c \sum_l cost_{c,l} \cdot X_{c,l}$$

Subject to:

$$\sum_l X_{c,l} = 1 \ (\forall c) \qquad \textbf{Constraint 1}$$

$$\sum_c X_{c,l} \cdot stock_c \leq capacity_l \ \ (\forall l) \qquad \textbf{Constraint 2}$$

$$X_{c,l} \geq fill\_grade_{c,l} \ (\forall c, l) \qquad \textbf{Constraint 3}$$

$$0 \leq X_{c,l} \leq 1 \ (\forall c, l) \qquad \textbf{Variable restriction}$$

To provide an example of the storage model, we use the stock levels per category from week 1 of 2012 as input. In total 36 million sales units require storage. We assume that at least 40% of category 3 should be stored at the DPC/VPS, at least 15% of category 30 should be stored at the halls/HBW, at least 35% of category 2 should be stored at the eDC, at least 25% of category 17 should be stored at the Veem and at least 45% of category 7 should be stored at the Veem. When we run the model, the objective gives a value of €574,558. The division percentages (i.e., the model output) are given in Table 7-13.

Nr.	Category	DPS/ VPS	Halls/ HBW	eDC	Veem	Beus-ichem
1	Ladies' and men's apparel	x%	x%	x%	x%	x%
2	Babies and children's apparel	x%	x%	x%	x%	x%
3	Lingerie, underwear and nightwear	x%	x%	x%	x%	x%
4	Hosiery	x%	x%	x%	x%	x%
5	Leather goods, accessories, swimwear	x%	x%	x%	x%	x%
6	Interior textiles and home accessories	x%	x%	x%	x%	x%
7	House wares	x%	x%	x%	x%	x%
8	Do-it-yourself and maintenance	x%	x%	x%	x%	x%
9	Stationery, toys and Christmas	x%	x%	x%	x%	x%
10	Personal care	x%	x%	x%	x%	x%
11	Services	x%	x%	x%	x%	x%
15	Materials	x%	x%	x%	x%	x%
17	New markets	x%	x%	x%	x%	x%
30	New services	x%	x%	x%	x%	x%

TABLE 7-8 SUITABLE PRODUCT DISTRIBUTION ACCORDING THE LP-MODEL FOR THE HEMA CATEGORIES IN WEEK 1 OF 2012

The model shows clear differences between the storage locations. It's beneficial based from a cost perspective to store expensive goods with a high turnover rate at the distribution center. However, products that require a longer storage period can be held at an external depot.

[*Confidential*]

As stated in the introduction of this section, the model is adequate for allocating stock in an empty warehouse. Because this is just a theoretic situation for HEMA, we recommend using the empty warehouse storage model occasionally (e.g., once every three months) and compare the actual situation with the preferred situation and making gradual changes.

Furthermore, it's possible to use the LP model as an *incremental stock model*. For instance, the current stock per category can be set as *fixed* (by changing the values of constraint 3). New arrivals are then used as input. By running the model a new stock division is calculated. The category forecast model explained in Section 5.1.1 can be used to estimate this input factor.

Finally, HEMA has the option to alter the associated costs, height or turnover ratio. The values used in this thesis can be recalculated to improve the model's accuracy.

7.5 CONCLUSION

[*Confidential*]

Based on the three storage properties, we constructed a linear programming (LP) storage model appropriate for an empty warehouse. HEMA can use the storage model to determine a suitable location for its inventory. The number of sales units on stock for each category is required as input variable. Its objective is to minimize the total storage costs. The model is subject to three constraints which prevent a capacity overflow. HEMA has the option to change the constraints and can indicate for example that a certain part of stock should be stored at a certain location. The model's output is a suitable storage division expressed in percentages. Furthermore, it's possible to use the model for incremental changes. The current stock is set as fixed and the LP model can calculate a suitable location for new products. The category stock forecast model from Section 5.1.1 can be used to estimate future stock levels.

When HEMA wants to make forecasts up to 12 weeks ahead, they can use the improved medium-term forecast model. Parameters for the base level, trend and seasonal influences are estimated based on historic data. When HEMA management wants to make forecasts up to one year ahead, they can use the long-term forecast model. The trend is estimated based on historic data. Cyclical influences can only be estimated by using human judgment. HEMA can compare the forecast outcome with the approximate capacity available to determine whether additional storage capacity is required. Once it's certain that the current capacity is sufficient, HEMA can use the storage model to determine a suitable location for the products.

Because all posed research questions of Section 1.3 are answered, we can now reflect on the research goal of Section 1.2. In Chapter 8 we give the conclusion and recommendations that are the result of the conducted research.

8 CONCLUSION

The aim of this research (as stated in Section 1.2) is to develop a practical model that is able to forecast the required internal and external storage capacity needed up to 12 months ahead which will allow HEMA to decide whether additional storage locations are required and if the contract with Beusichem can be cancelled. We evaluate the research goal and present the outcome of the conducted research in this final chapter.

In Section 8.1 we recapitulate the previous chapters and present the research conclusions. Subsequently, in Section 8.2 we give the recommendations for HEMA relating to the research aim. Finally, in Section 8.3 we give suggestions for further research.

8.1 CONCLUSIONS

Analysis shows that the future stock level of the distribution center can be forecasted by considering the historic data as a time series. When medium-term forecasts (i.e. up to 12 weeks ahead) are required, the mean average percentage error (MAPE) and mean absolute scaled error (MASE) are good measures to assess the accuracy. Literature shows that the class of standard exponential smoothing methods is most suitable for HEMA to use. We test three exponential smoothing methods to identify which one gives the best results. These are: 1) simple exponential smoothing; 2) Holt's damped trend model; and 3) Holt-Winters' seasonal method. The various parameters of the three models are estimated by minimizing the mean squared error (MSE). It turns out that forecasts made for the aggregated level should be done by using the Holt-Winters model. It provides accurate forecasts for every selected horizon and almost all validation intervals. Only the second interval, during the introduction of the selling less is more (SLIM) concept, the Holt-Winters model underperforms.

We hypothesize that a disaggregated forecast can improve the results. This hypothesis is falsified for the accumulated forecasts made for the category level. However, accumulated forecasts made for the division level supports our hypothesis. This combined forecast model outperforms the initial Holt-Winters model. Stock levels for fashion and hardware are forecasted with Holt-Winters, stock levels for various are forecasted with simple exponential smoothing and the outcomes are added to gain the aggregated forecast. Managerial judgment is improved by adding a 95% prediction interval. The same applies for adding a tracking signal to the forecasts. It calls for action when forecasts show significant bias and human intervention is required.

[*Confidential*]

8.2 RECOMMENDATIONS

The recommendations that directly follow from the conclusions are given in Section 8.2.1. Furthermore, we give some general recommendations in Section 8.2.2 that became clear during the conducted research at HEMA.

8.2.1 RESEARCH SPECIFIC RECOMMENDATIONS

1. Use the combined forecast model to make medium-term forecasts with a maximum horizon of 12 weeks. Update the stock levels of the three divisions weekly and take action when the model gives a tracking signal which indicates sufficient bias.

2. Apply linear regression and ascertain that the constraint $R^2 \geq 0.6$ is satisfied. Project the historic trend into the future to make long-term forecasts up to one year ahead.

3. Make use of a project team to estimate the cyclical effect on the long-term. Test different scenarios by changing the cyclical factor and check its influence on the predicted trend.

Confidential

8.2.2 GENERAL RECOMMENDATIONS

Confidential

8.3 FURTHER RESEARCH

Further research is possible in several ways. First of all, it would be interesting to test whether other forecasts methods can outperform the combined forecast model. This could be accomplished by using a more accurate model with just the aggregated stock as input variable. But also by disaggregating the data to a lower level and capturing more variability. HEMA's demand chain management system could be helpful for this if the input data is more reliable.

Furthermore, it would be interesting to know whether the LP storage model can be improved by adding more properties. For instance, a seasonal factor could be added to distinguish between products that are in the current assortment and products which are not. The product groups should be used as input variable instead of the product categories. This would require more data input, but maybe this process can be automated.

The used variables and parameters for the LP storage model require further explanation. Simple calculations are used to determine the storage costs for a sales unit. These could be improved. Subsequently, the turnover period is just an average for all the products from a category. A more specified expected storage period can increase the accuracy of the results. The storage model can be used by HEMA but it depends on the quality of the input.

REFERENCES

Armstrong, S. J. (2001). *Principles of Forecasting: A Handbook for Researchers and Practitioners.* New York, NY: Springer.

Armstrong, S., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: Emperical comparisons. *International Journal of Forecasting, 8*(1), 69-80.

Ballou, R. (2000). Evaluating inventory management performance using a turnover curve. *International Journal of Physical Distribution & Logistics, 30*(1), 72-85.

Bartholdi, J. J., & Hackmann, S. T. (2011). *Warehouse & Distribution Science.* Atlanta, GA: Georgia Institute of Technology.

Box, G. E., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control.* San Fransisco: Holden-Day.

Brown, R. (1956). *Exponential Smoothing for Predicting Demand.* Cambridge, MA: Arthur D. Little Inc.

Brown, R. (1959). *Statistical forecasting for inventory control.* New York: McGraw-Hill.

Brown, R. G. (1963). *Smoothing, Forecasting and Prediction.* Englewood Cliffs, NJ: Prentice-Hall.

Cooper, D. R., & Schindler, P. S. (2008). *Business Research Methods* (10th ed.). Boston: McGraw-Hill.

De Groot, S. (2011). *The impact of local interventions on the performance of the automatic store replenishment system: A case study at HEMA.* Amsterdam: Erasmus University Rotterdam.

DeHoratius, N., & Raman, A. (2008). Inventory Record Inaccuracy: An Empirical Analysis. *Management Science, 54*(4), 627-641.

Dekker, M., Van Donselaar, K., & Ouwehand, P. (2004). How to use aggregation and combined forecasting to improve seasonal demand forecasts. *International Journal of Production Economics, 90*(2), 151-167.

Fildes, R. (1992). The evaluation of extrapolative forecasting methods. *International Journal of Forecasting, 8*(1), 81-98.

Fildes, R. (2001). Beyond forecasting competitions. *International Journal of Forecasting, 11*(2), 47-56.

Fildes, R., Goodwin, P., Lawrence, M., & Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting, 25*(1), 3-23.

Gardner, E. S. (1985). Exponential Smoothing: The State of the Art. *Jorunal of Forecasting, 4*(1), 1-28.

Gardner, E. S. (2006). Exponential Smoothing: The State of the Art - Part II. *International Journal of Forecasting, 22*(4), 637-666.

Gardner, E. S., & McKenzie, E. (1985). Forecasting Trends in Time Series. *Management Science, 31*(10), 1237 - 1246.

Geurts, P. (1999). *Van probleem naar onderzoek.* Bussum: Coutinho.

Hamburg, M., & Yough, P. (1994). *Statistical Analysis for Decision Making.* Belmont, CA: Wadsworth Publishing Company.

Hanke, J., & Reitsch, A. (1998). *Business Forecasting* (6th ed.). Upper Saddle River, NJ: Prentice-Hall.

Harrison, P. J., & Davies, O. (1964). The Use of Cumulative Sum (CUSUM) Techniques for the Control of Product Demand. *Operations Research, 12*(2), 325-333.

Hellström, D., & Saghir, M. (2007). Packaging and Logistics Interactions in Retail Supply Chains. *Packaging Technology and Science, 20*(3), 197-216.

HEMA. (2008). *Distributiecentrum Utrecht: Het logische vervolg.* Culemborg: Present Press.

HEMA. (2011). *Packaging instructions HEMA logistics.*

Hendry, D. F., & Hubrich, K. (2005). *Forecasting Aggregates by Disaggregates.* Oxford: Department of Economics, Oxford University and Research Department, European Central Bank.

Holt, C. C. (1957). *Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages.* Pittsburgh, PA: Carnegie Institute of Technology.

Hoover, J. (2009). How to track forecast accuracy to guide forecast process improvement. *Foresight, 2009*(14), 17-23.

Hugos, M. H. (2011). *Essentials of Supply Chain Management* (3th ed.). Hoboken NJ: Wiley.

Hyndman, R. (2006). Another look at forecast-accuracy metrics for intermittent demand. *Foresight, 2006*(4), 43-46.

Hyndman, R., & Koehler, A. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting, 22*(4), 679-688.

Kahn, K. B. (1998). Revisiting top-down versus bottum-up forecasting. *The Journal of Business Forecasting, 17*(2), 14-19.

Kalekar, P. S. (2004). Time series Forecasting using Holt-Winters Exponential Smoothing. *Kanwal Rehki School of Information Technology*, 1-13.

Kolassa, S., & Schütz, W. (2007). Advantages of the MAD/Mean Ratio over the MAPE. *Foresight: The International Journal of Applied Forecasting, 6*, 40-43.

Lee, H., Padmanabhan, V., & Whang, S. (1997). The Bullwhip Effect in Supply Chains. *Sloan Management Review, 38*(3), 93-102.

Makridakis, S. (1986). The art and science of forecasting: An assessment and future directions. *International Journal of Forecasting, 2*(1), 15-39.

Makridakis, S. G., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and Applications* (3th ed.). New York, NY: Wiley.

Makridakis, S. G., Wheelwright, S. C., & McGee, V. E. (1983). *Forecasting: Methods and Applications* (2nd ed.). New York, NY: Wiley.

Makridakis, S., & Hibon, M. (2000). The M3-Competition: Results, conclusions and implications. *International Journal of Forecasting, 16*(4), 451-476.

Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., Mills, T., Ord, K., et al. (1993). The M2-Competition: A real-time judgmentally based forecasting study. *INternational Journal of Forecasting, 9*(1), 5-22.

Martin, C. A., & Witt, S. F. (1989). Forecasting tourism demand: A comparison of the accuracy of several quantitative methods. *International Journal of Forecasting, 5*(1), 7-19.

McCarthy, E. J. (2001). *Basic Marketing: A Managerial Approach* (13th ed.). Illinois: Homewood.

McClain, J. O., & Thomas, L. J. (1973). Response-Variance Tradeoffs in Adaptive Forecasting. *Operations Research, 21*(2), 554-568.

Ozer, D. (1985). Correlation and the Coefficient of Determination. *Psychological Bulletin, 97*(2), 307-345.

Pakkala, T., & Achary, K. (1992). A deterministic inventory model for deteriorating items with two warehouses and finite replenishment rate. *European Journal of Operation Research, 57*(1), 71-76.

Pegels, C. (1969). Exponential Forecasting: Some New Variations. *Management Science, 15*(5), 311-315.

Scholtus, L. (2009). *Cost Transparancy at HEMA: The development of a supply chain cost model.* Enschede: University of Twente.

Silver, E. A., Pyke, D. F., & Peterson, R. (1998). *Inventory Management and Production Planning and Scheduling* (3th ed.). Hoboken, NY: John Wiley & Sons.

Stickney, C., Veil, R., & Schipper, K. (2009). *Financial Accounting: An Introduction to Concepts, Methods and Uses* (13th ed.). Mason, OH: South-Western.

Taylor, J. W. (2003). Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting, 19*(4), 715-725.

Van Dillen, D. (2011). Uitbreiding DC Lage Weide. Amsterdam.

Wegert, E., Roth, O., & Kraus, D. (2012). On Beurling's boundary value problem in circle packing. *Complex Variables and Elliptic Equations: An International Journal, 57*(2-4), 397-410.

Winston, W. L. (1994). *Operations Research: Applications and Algorithms* (3th ed.). Belmont CA: International Thomson Publishing.

Winters, P. R. (1960). Forecasting Sales by Exponentially Weighted Moving Averages. *Management Science, 6*(3), 324-342.

Yaffee, R. A., & McGee, M. (2000). *Introduction to Time Series Analysis and Forecasting.* New York, NY: Academic Press.

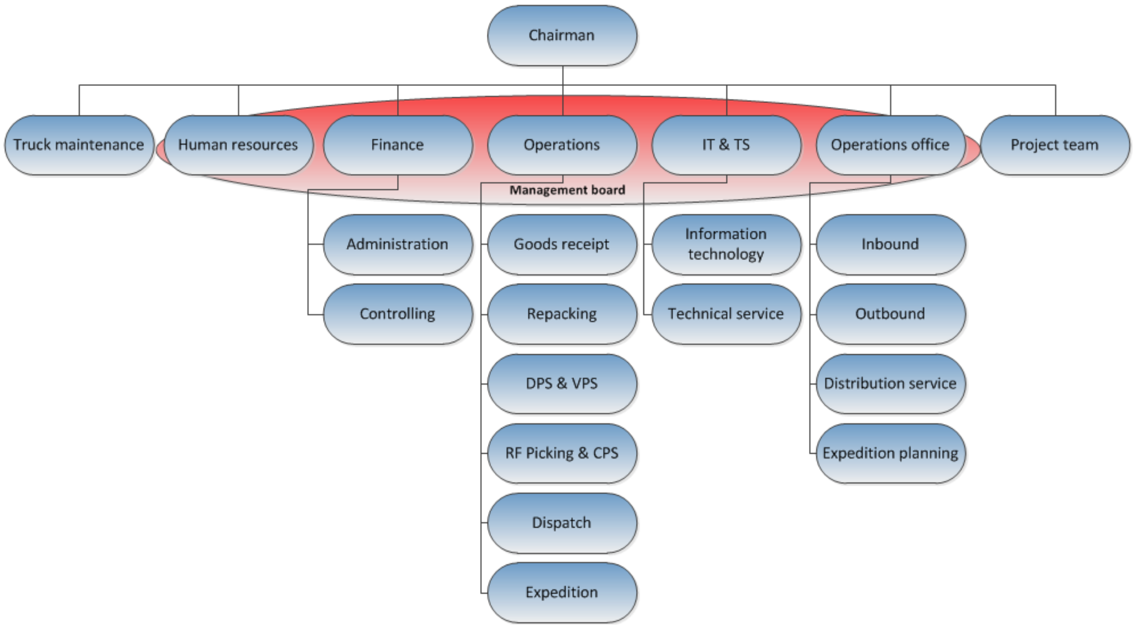
Zhang, G., Patuwo, B., & Hu, M. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting, 14*(1), 35-62.

Zotteri, G., Kalchschmidt, M., & Caniato, F. (2005). The impact of aggregation level on forecasting performance. *International Journal of Production Economics, 93-94*(1), 479-491.

APPENDIX I: HEMA ASSORTMENT

Confidential

APPENDIX II: ORGANIZATIONAL CHART



APPENDIX III: CORRELATION INVENTORY AND DEMAND

Confidential

APPENDIX IV: GRID SEARCH HOLT-WINTERS

```
Private Sub grid_search(nr_rows As Integer, alpha As Double, beta As Double, gamma As
Double, level As Double, trend As Double, stock_level() As Double, seasonalities() As
Double, forecast_grid_search() As Double, forecast() As Double, SSE() As Double)

  Dim i, current_row As Integer, nr_SSE As Integer
  Dim previous_level As Double, stock_level_value As Double, seasonality_factor As Double
  Dim MSE As Double, min_MSE As Double
  Dim best_alpha As Double, best_beta As Double, best_gamma As Double
  Dim best_level As   Double, best_trend As Double
  Dim initial_level As Double, initial_trend As Double
  Dim initial_seasonalities(1 To YEAR) As Double

  min_MSE = MaxInt

  initial_level = level
  initial_trend = trend
  Call copy_seasonalities(seasonalities, initial_seasonalities)

  For alpha = 0.02 To 0.50 Step 0.01
    For beta = 0.005 To 0.176 Step 0.01
      For gamma = 0.05 To 0.50 Step 0.01

        current_row = nr_rows - YEAR - YEAR
        MSE = 0
        nr_SSE = 0

        level = initial_level
        trend = initial_trend
        Call copy_seasonalities(initial_seasonalities, seasonalities)

        For i = 1 To YEAR + YEAR

          Call make_forecast(current_row, level, trend, seasonalities,
forecast_grid_search())
          Call calculate_SSE(nr_rows, current_row, stock_level, forecast_grid_search, SSE,
nr_SSE, MSE)

          current_row = current_row + 1
          previous_level = level
          stock_level_value = stock_level(current_row)
          seasonality_factor = seasonalities(get_seasonality_factor(current_row))

          level = update_level(alpha, level, trend, stock_level_value, seasonality_factor)
          trend = update_trend(beta, previous_level, level, trend)
          seasonalities(get_seasonality_factor(current_row)) =
update_seasonality(gamma, stock_level_value, level, seasonality_factor)
          Call normalize_seasonalities(seasonalities())
```

```
        Next i

        MSE = MSE / nr_SSE

        If MSE < min_MSE And alpha > 2 * beta Then
          min_MSE = MSE
          best_alpha = alpha
          best_beta = beta
          best_gamma = gamma

          Call make_forecast(nr_rows, level, trend, seasonalities, forecast())

        End If

      Next gamma
    Next beta
  Next alpha

End Sub
```