TAMPERE UNIVERSITY OF TECHNOLOGY

SASU MÄNTYNIEMI

DEVELOPMENT OF DEMAND FORECASTING PROCESS

Master of Science Thesis

Prof. Miia Martinsuo and lect. Ilkka Kouri have been appointed as the examiners at the Council Meeting of the Faculty of Business and Technology Management on April 4th, 2012.

ABSTRACT

TAMPERE UNIVERSITY OF TECHNOLOGY
Master’s Degree Programme in Industrial Engineering and Management

MÄNTYNIEMI, SASU: Development of Demand Forecasting Process
Master of Science Thesis, 107 pages, 14 appendices (23 pages)
May 2012
Major: Industrial Management
Examiners: Professor Miia Martinsuo, lecturer Ilkka Kouri
Keywords: Demand, demand forecasting, demand management, demand planning, demand forecasting process, development

The purpose of demand forecasting is to predict the future demand of products or items, and thus, to ensure that right amount of products or items is available when needed. Because future events cannot always be known beforehand, forecasts are usually incorrect. For this reason, companies need to make contingency plans on account of the inaccuracy, resulting in more costs. By improving different aspects of demand forecasting, more accurate forecasts can be made, leading to decreases in costs and increases in service level. The demand forecasting process combines different aspects of demand forecasting into a multi-step process, which can be used as a framework for how companies should handle their demand forecasting. However, there are several interpretations of how the demand forecasting process should function.

The case company of this study is a Finnish paints and coatings manufacturer, which operates in both industrial and consumer markets. The purpose of this study is to use the concept of Demand Forecasting Process to evaluate and improve demand forecasting in the case company in order to provide the company with more accurate forecasts. This is done by evaluating how different phases of the demand forecasting process are handled in the case company. Afterwards possible alternate approaches are suggested and their effects are further estimated or tested. The company’s use of a specific forecasting software as the main tool with demand forecasting limits some of the recommendations and alternatives that are presented in this study. The data that is used in this study is mostly the sales data of different products, which is provided by the case company.

The results of this study indicate that there are some steps in the demand forecasting process of the case company which could be improved. This means that some recommendations can be made on how the demand forecasting process should work in the case company. Because of the external approach of this study, which lead to the lack of proper information in some cases, and the limitations that the forecasting software as part of the demand forecasting process created, some of the findings of this study are not necessarily applicable in other studies and some of the solutions that were presented are only the best possible from the ones that are available for the case company.

TIIVISTELMÄ

TAMPEREEN TEKNILLINEN YLIOPISTO
Tuotantotalouden koulutusohjelma

MÄNTYNIEMI. SASU: Kysynnän ennustamisprosessin kehittäminen
Diplomityö, 107 sivua, 14 liitettä (23 sivua)
Toukokuu 2012
Pääaine: Teollisuustalous
Tarkastajat: professori Miia Martinsuo, lehtori Ilkka Kouri
Avainsanat: Kysyntä, kysynnän ennustaminen, kysynnän hallinta, kysynnän ennustamisprosessi, ennusteprosessin kehittäminen

Kysynnän ennustamisen tarkoitus on laskea tai arvioida jonkin tuotteen tulevaa kysyntää ja näin ollen varmistaa, että oikea määrä kyseistä tuotetta on saatavilla tarvittaessa. Koska tulevaisuuden ennustaminen on hankalaa, ovat ennusteet usein pielessä, mikä tarkoittaa, että yritysten tarvitsee tehdä suunnitelmia ennustevirheiden varalle. Tämä aiheuttaa yleensä lisäkustannuksia yrityksille. Ennusteiden tarkkuutta voidaan parantaa kehittämällä ennustamisen osa-alueita, mikä taasen johtaa kustannusten laskuun ja palvelutason paranemiseen. Kysynnän ennusteprosessi yhdistää kysynnän ennustamisen osa-alueet yhdeksi monivaiheiseksi prosessiksi, mitä voidaan käyttää viitekehyksenä mietittäessä, miten kysyntä ennustamista voidaan parantaa. Ennusteprosessin etenemisestä on kuitenkin olemassa useita erilaisia tulkintoja.

Tutkimuksen kohdeyritys on suomalainen maalien ja pinnoitteiden valmistaja, jonka asiakkaita ovat sekä eri teollisuudenalat että kuluttajat. Tutkimuksen tarkoitus on käyttää kysynnän ennusteprosessi -konseptia arvioimaan ja parantamaan kysynnän ennustamista ja ennusteiden tarkkuutta kohdeyrityksessä arvioimalla, miten eri kysynnän ennusteprosessin vaiheet suoritetaan kohdeyrityksessä ja tarjoamalla vaihtoehtoisia ratkaisuja, ja arvioimalla näiden ratkaisujen vaikutusta ennusteprosessin laatuun. Tutkimusta rajoittaa ennusteohjelmiston käyttö, mikä tarkoittaa, että jotkut ratkaisuvaihtoehdot ja jäävät tutkimuksen ulkopuolelle. Data, jota tutkimuksessa käytetään, koostuu suurimmaksi osaksi eri tuotteiden historiallisesta myyntidatasta.

Tutkimuksen tulosten perusteella voidaan sanoa, että kysynnän ennusteprosessin eri osa-alueita voidaan parantaa yrityksessä. Tämä tarkoittaa, että erilaisia ratkaisuja ja toimenpide-ehdotuksia, miten prosessin tulisi vastaisuudessa toimia, pystytään tarjoamaan kohdeyritykselle. Samalla niiden vaikutusta pystytään osittain arvioimaan. Tutkimuksen ulkopuolisen näkökulman johdosta, mikä johti osittain tarvittavan tiedon puuttumiseen, ja ohjelmiston käytön aiheuttamien rajoitteiden vuoksi jotkut ratkaisut eivät välttämättä ole verrattavissa muihin tutkimuksiin asiasta. Tämän lisäksi jotkut tässä tutkimuksessa esitetyt ratkaisut ovat ainoastaan parhaat niistä vaihtoehdoista, joita kohdeyritykselle voidaan tarjota ennusteohjelmistossa.

PREFACE

First and foremost I would like to thank Teknos Oy and especially Jaakko Koskenpää for giving me this great and interesting opportunity for my Master's Thesis.

Big thanks goes out to the two examiners of this thesis, professor Miia Martinsuo and lecturer Ilkka Kouri, whose support, guidance and feedback has helped me over the course of my writing and study process.

Additionally, I would like to thank my friends and family for their support over the years, which has helped me progress to this phase in my studies in the first place.

Last but not least, I would like to thank Irina Pravet for the proofreading of this thesis.

Tampere 22.5.2012

Sasu Mäntyniemi

sasu.mantyniemi@tut.fi

TABLE OF CONTENTS

ABBREVIATIONS

AE	Absolute Error
AC	Architectural Coatings (product segment of the case company)
COV	Coefficient of Variation
ERP	Enterprise Resource Planning
FVA	Forecast Value Added
GI	General Industry & Heavy Duty (product segment of the case company)
IM	Road Marking and Floor Coatings (product segment of the case company)
IW	Industrial Wood (product segment of the case company)
MAD	Mean Absolute Deviation
MAE	Mean Absolute Error (Same as MAD, but this term is used in the software)
MAPE	Mean Absolute Percentage Error
MdAPE	Median Absolute Percentage Error
ME	Mean Error
MSE	Mean Squared Error
PC	Powder Coatings (product segment of the case company)

1. INTRODUCTION

This study is done for Teknos Oy, a Finnish manufacturer of paints and coatings. The aim of this thesis is to focus on the demand forecasting process in the case company. The reason why forecasting can be seen as a somewhat new entity in the case company is the implementation of a new ERP-software (IFS Demand Planning) including the forecasting software, which is the heart of the case company's forecasting process. The software was introduced in the case company in the beginning of the year 2010, and the case company itself has not had the resources to evaluate the different attributes of the software, its use in forecasting and the overall performance of the company's forecasting process.

The products of the case company, manufactured for both consumer and industrial markets, include those manufactured based on customer orders and those being kept in the stock continuously. In the latter case, demand forecasting is needed if, for example, the acceptable delivery lead time of a product is shorter than the production or replenishment lead time. This is because the company has to keep a certain safety stock level at all times in order to ensure that it can deliver its products to customers when needed. Therefore the demand forecast has a direct impact on the safety stock levels, which again affects the company's ability to ensure a continuous flow of products to its customers.

1.1. The purpose and scope of the study

The purpose of this study is to develop and improve the demand forecasting process of the case company in order to provide the case company with more accurate forecasts. The theoretical background will provide a framework for the concept of the demand forecasting process. The demand forecasting process presented in the theory section will act as a benchmark that the actual demand forecasting process of the case company will be compared to. Based on a thorough literature review the following question will be answered:

1. Which actions and procedures, related to forecasting, should a company implement in order to ensure an effective demand forecasting process?

In other words: the first phase of this study is to use theory and concepts of forecasting to define a multi-step model, which is the demand forecasting process. In addition to the theoretical review, the second phase of the study will include the analysis of the case company and its current demand forecasting process. The aforementioned will include analysis of the company, its customers, its products, demand for those products, forecasting practices and possible specificities affecting forecasting practices in the case company. Based on the theoretical framework, the analysis of the case company and the current state of its demand forecasting process, the main research problem of this study will be answered:

2. How to develop and improve that process and thus provide more accurate forecasts for the case company?

To help answer the main research problem, the study will include an empirical section, where different aspects of the demand forecasting process are analyzed and possible alternatives tested.

At this point it should also be mentioned that at the heart of the case company's demand forecasting process lies a specific demand forecasting software. This is why this study's approach to demand forecasting is limited to the use of this software. Analysis and improvements of the demand forecasting principles in the case company will thus focus only on solutions which utilize the aforementioned software. Therefore, some of the alternatives which would normally be suitable may be discarded if they do not belong to the alternatives provided by the software. For example, when the accuracies of the statistical models were tested, only the models available in the software were included. This means that the best possible option that is suggested in this study is not absolutely the best possible option, but it is only the best possible available option for the case company.

There are some steps of the demand forecasting process that are not discussed in this study. These are: planning of dependent demand and data gathering. The former was left out because the main concepts related to it were seen as parts of planning rather than forecasting. The latter was also excluded because it was not seen as a direct part of demand forecasting procedures of the case company. Another factor that influenced the scope was the external perspective from which this study was conducted. This meant that the needed information was not always available to help the analysis or to find areas of improvement. Therefore, some assumptions had to be made based only on the demand data and the information that was available.

1.2. The structure of the study

This study is divided into three main sections. The first section, the literature review, will provide a framework to which the latter sections can be compared. The second section, analysis, includes the overall analysis of the case company and its demand forecasting process. After which the third section presents some possible improvements for the process are presented and the effects of changing some of the procedures and parameters are tested. Based on the results of the tests, some guidelines and suggestions for improvements are presented.

The theoretical section of this thesis is discussed in chapters two and three. The purpose of the literature review section is to provide the reader a proper comprehension of the subject, and additionally to provide a framework for the analysis and results that are later addressed in this thesis. The literature review consists of collecting, choosing and combining the theoretical material used in this study. The theoretical material includes books related to supply chain management and operations management, academic journals and articles relating to previous studies around different aspects of forecasting practices.

The purpose of using information from books related to the research subject is to provide readers with a general understanding of forecasting theory and the best practices described in literature. However, the prevailing weakness of the literature is that it is mainly limited to theory and practices of forecasting in consumer markets. Although the forecasting theory of consumer markets is partially applicable to industrial markets as well, there are certain practices that should be dealt with differently depending on the type of market. That is why not all of the best practices presented in the books are applicable to forecasting principles in industrial markets, one of the areas that this thesis focuses on.

The aforementioned problem was dealt with by collecting theory from academic journals and articles regarding forecasting. Even though most of these articles and previous studies are somewhat focused on the same principles as the books, they are able to provide a broader understanding to the subject. Additionally, in them the distinction between the practices involving forecasting in industrial and consumer markets is much better in comparison to books. In short, the general theory and concepts of forecasting that is applicable in both industrial and consumer markets is usually derived from the books, whereas the theory about differences of forecasting practices between the two markets is derived from journals, articles and other publications dealing with the research subject.

In addition to the basic forecasting practices, the theoretical section will introduce the reader to the concept of the demand forecasting process. To fully understand the meaning of the aforementioned concept is important, because it is the basis of this whole thesis. The demand forecasting process has been addressed in the literature and some other studies involving forecasting. However, its meaning has often varied depending on the author, the context or the study. That is why in this study the concept is defined based on the characteristics of this particular study. In other words: the mission is not to create a new way of studying the concept, but rather to explain what the concept includes in this study.

After the theory section, the case company and its demand forecasting process are presented in chapter four. This includes an analysis of the case company itself, its customers and end products. After the case company analysis, the demand forecasting process of the case company is presented. This is done by describing how each step of the process is being done in the case company. In the very center of the forecasting process in the case company is the use of specific software, with which all of the different steps of the process are made. Therefore, this study will focus on how those steps are handled with the use of the software. This means that the data that is used in the analysis is the data provided by the forecasting software and also the additional information about the guidelines and rules regarding the usage of the software.

In the third phase of the study some possible areas of improvement and alternative approaches in the demand forecasting process are identified and their effects on the quality of the forecasting process are further tested. To measure the quality of the demand forecasting process, this study uses the output of the process, which is accuracy of the forecast, as a measure to evaluate whether or not an alternative approach could improve the process. In some cases the effect of the change on forecast accuracy cannot be directly tested, which means that in those cases the study merely estimates if a change could improve the demand forecasting process or not. The material that is used in the second and the third section is discussed further in subchapter 1.3.

The third phase is presented in chapters five and six. Chapter five consists of testing or estimating the possible alternatives for different steps of the process and presents the results, whereas chapter six gathers all the findings presented in chapter five and presents, based on the results, some possible modifications or recommendations and suggests some courses of action that could be taken to improve the demand forecasting process. Chapter seven consists of conclusions made about the entire study, the usage of its results, as well as possibilities for further research.

1.3. Material and methodology of the study

As previously mentioned, the case company uses a forecasting software in its current demand forecasting process. The basic information about the use of the software is available in the software manuals and specific guide books of the case company, which are partially used as an analysis tool for the current demand forecasting practices. However, to gain a deeper understanding of how the software is actually used as a forecasting tool and which of its specific features are being used on a day-to-day basis, meetings were held in the case company. The attendees included the company’s production director, who provided instruction on how the forecasting software is used and how the software’s data can be accessed and modified. These meeting were always informal. However, some notes were taken and used as the basis for some of the analysis of the current demand forecasting process.

The forecasting software's data, also used in this study, is the sales data of the case company's products, hereon in referred to as demand data. The reason why a distinction between the two terms has to be made is because of different possibilities to define demand; sales and demand do not always mean the same thing. However, in this study, when referring to the demand data of the products or demand data in the software, this study actually talks about the sales data. When the data was used, it was always in the form presented in appendix 1. However, it could be organized in a number of different ways, depending on what was searched.

The aforementioned means that in some cases the data could be limited to include only certain products or forecast groups or only certain values (e.g. different error measures). This was useful, for example, when testing the effects of some modifications, which could be done by changing certain settings in the software. However, in some cases (in order to compare the original settings and the modified ones), the data was copied to Excel in order to make further calculations about the effects of the changes. This had to be done because the calculations could not be done in the software itself. In some cases the settings of the software were not changed but the data was organized in different ways in the software in order to identify certain situations, where changes to the existing practices would be applicable. Appendix 1 shows how the data is presented in the software and in which ways it can be organized and how the effects of changing some settings impact the data.

When using the demand data of different products, some general limitations are made because of the abundance of different products that the case company manufactures. Hence, only some of those products are taken into account for the analysis in this study. First, products of certain inventory classes are excluded. The study will only include Make-to-Stock products, whereas other inventory classes, which are Make-to-Order, Make-to-Lot and Deleted products (classification of the case company) are excluded because the demand of these products is not forecasted. Second, only four out of the five product segments are included. These segments are: architectural coatings (AC), general industry and heavy duty (GI), powder coatings (PC) and industrial wood (IW). The segment Road Marking and Floor Coatings (IM) was excluded because of its specific characteristics and the relatively low importance based on sales of Make-to-Stock products (1 %).

At this point it should also be mentioned that this study is conducted mostly from an external perspective. This means that, for example, the actual behaviour of people involved in the demand forecasting process of the case company was not observed and all in all, the communication with the case company was relatively limited, apart from the meetings in the company. Because of this, the assumptions about the daily use of the forecasting software are based on the suggested practices and guidelines of the case company, which means that in this study it is not absolutely clear whether or not the people involved in the process are actually using the software according to the aforementioned guidelines.

The reason why this sort of approach is taken is because the case company requested an external perspective about the use of the forecasting software in the demand forecasting process. The benefit of this approach is that a completely external perspective can focus efforts on certain areas that do not necessarily come as a suggestion from the company. However, disadvantages include a lack of information about the state of the actual forecasting practices and the fact that some of the suggestions have to be made on a more abstract level because of this.

2. DEMAND FORECASTING

The theoretical part of this study is divided into two sections. The first section, which is this chapter, focuses on general practices and theory related to demand forecasting. The aforementioned includes theory about forecasting principles, methods and special characteristics depending on the markets and customers. The purpose of it is to introduce the reader to the fundamental aspects of forecasting.

In the third chapter of this study, which is the second theory chapter, forecasting is viewed as a process within a company. Therefore, the third chapter will focus on describing that process and its parts. Another purpose of it is to create a theoretical framework of the process, suitable for the specific requirements of this study. The materials in the theory chapters are collected from operations management and supply chain management literature and from related academic journals and articles.

2.1. Demand and its special characteristics

Demand is usually defined as customers' willingness to purchase some specific product, which can be either a commodity or a service. However, demand should not be limited to the purchasing operation between a company and its customers but rather, considered to be a versatile movement of products between two or more parties. (Kiely 1999) According to Chambers et al. (2004, pp. 327–330) demand can be divided into two categories: independent and dependent demand. Independent demand is a type of demand that cannot be known beforehand with utmost certainty, whereas dependent demand is derived from a known factor.

An example of dependent demand is the demand of components or raw materials that are needed to manufacture a certain product. In such a case the number of components can be calculated from the number of products being manufactured. However, even though the demand of components and raw materials is dependent, the demand of the product being manufactured can be, and in most cases is, independent. It is because of the independent demand that companies need demand forecasting and planning. (Chambers et al. 2004, pp. 327–330)

Kiely (1999) states that demand is usually measured by the number of units of a certain product sold in a specific time period. When all different demands in their respective time periods are taken into account, the development of demand over time can be depicted as a demand curve or a time series of a demand. Based on a time series, it is possible to analyze, among other things, the historical patterns of demand and use it to estimate the future development. Buffa (1983, p. 59-60) identifies five different components or patterns of demand: average levels, trend, seasonal, cycle and random variation. Average level means an average demand for any particular period of time, which is more of a component that can be used in forecasts (Buffa 1983, p. 60).

Trend refers to a long-term upward or downward movement in the data (demand in this case) which can be either linear or exponential. Linear trend refers to a trend where the demand of a certain product increases or decreases regularly, whereas exponential trend refers to a trend where the demand increases or decreases in amount of a specific percentage every time period (Holt 2004). For example, a decrease in the price of the product might account for increased sales which could cause increasing trend (Armstrong and Collopy 1993).

Seasonal variations often refer to fairly regular variations which usually occur during a year. Good examples of seasonal products are winter or summer car tires. However, depending on the branch of the business or product itself seasonal variation can occur in a much shorter period of time such as one month, a week or even one day. (Chambers et. al 2004, pp. 363–364) The aforementioned short-term variations are more common amongst businesses that provide services (Radas & Shugan 2008). Cycles are similar to seasonal variations. The difference being that cycles are a case of a more long-term type of variation. The duration of cycles is usually one year or more and they are often related to, for example, economic or political conditions (Stevenson 2007, p. 72).

In addition to the first four patterns of demand, the time series normally includes random variability and possibly some irregular variations. Irregular variations are due to unusual, unpredictable circumstances such as natural disasters, political changes or a major change in a product itself. It is very important that once these kinds of variations are identified, they are removed from the data because they do not reflect typical behaviour, thus including them in the series (and later on in the forecast) will most likely distort the overall picture. Random variability is categorized as residual component that is left remaining – unless the demand is constant, which is unlikely – after all other patterns and variations have been accounted for. The change in demand between certain limits is categorized as random variability. (Stevenson 2007, pp. 72–73)

In addition to the five patterns, there is one special case that cannot be neglected: sporadic demand. A time series can be called sporadic (or intermittent), if no demand is observed in several periods. An example of this is C-class items, for which demand can often be sporadic. (Stadler & Kilger, 2008, p. 156) The difference between sporadic demand and irregular variation is that irregular variations are usually due to unusual circumstances and they do not happen very often, whereas sporadic demand happens more frequently even if the occurrence of it can be relatively random. These demand patterns can be seen in figure 1.1.

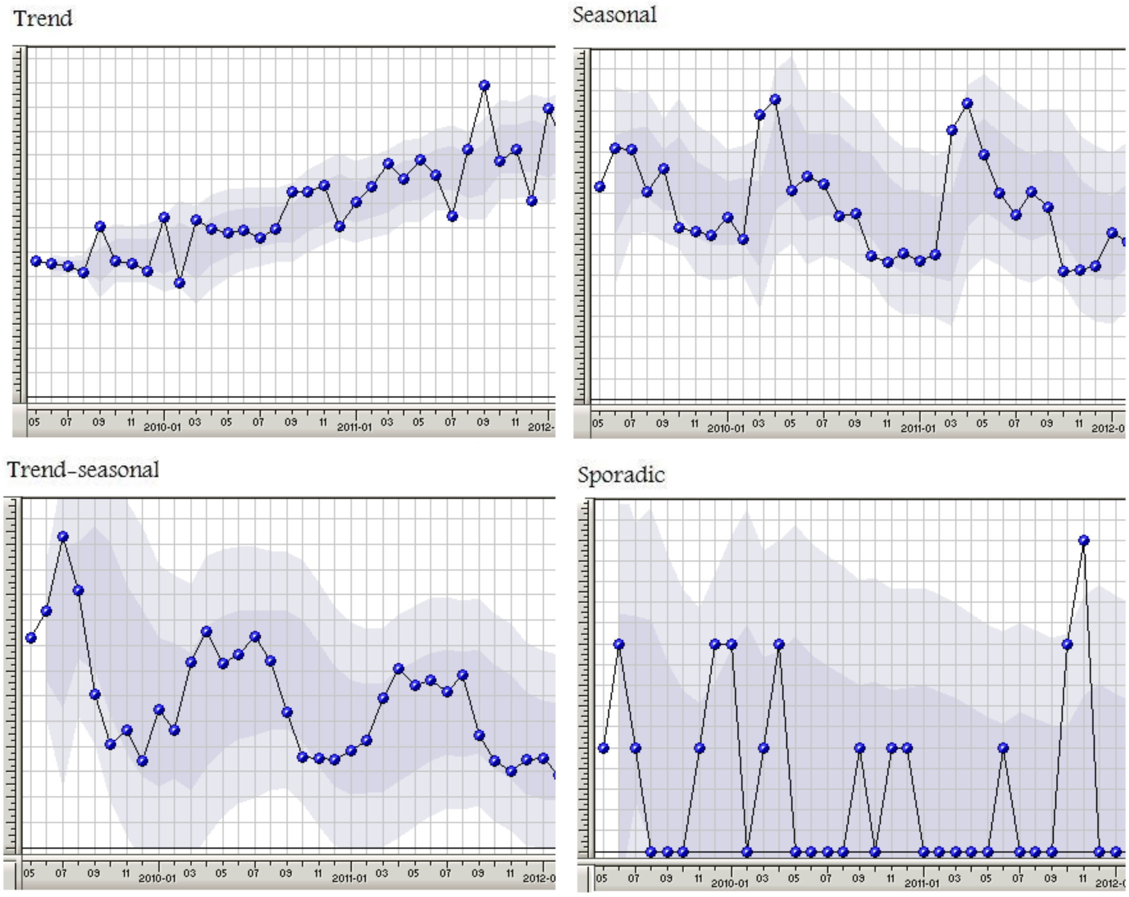


Figure 1.1. Different characteristics or patterns of a time series.

As depicted in figure 1.1, it is possible that a time series consists of a combination of two or more individual patterns. The declining trend and seasonality in the third case is an example of such a case. Random variation around a level demand can be seen in the top left graph of figure 1.1, within the first eight periods before the increasing trend. Additionally, the seasonal variation in the second case could be interpreted as a cycle if the time span during which it occurs would be two to three years instead of the 6-7 months seen in the second case of the figure 1.1.

2.2. General aspects of demand forecasting

The following subchapters will introduce some general concepts of forecasts and forecasting needs. At this point, it should be emphasized that the terms forecasting and demand forecasting are being used interchangeably throughout this study, because of their interchangeability in the different materials on which the literature review is based. In other words, in the material on which the literature review is based as well as this study, both of the terms mean the same thing.

2.2.1. Characteristics of a forecast

Merriam-Webster Dictionary (2012) defines forecast as a calculation or a prediction of some future event or condition, which is usually a result of a study and an analysis of available pertinent data. Another way of defining what forecast really is, is to look into the characteristics of forecasts. Stevenson (2007, p. 69) lists four characteristics that are said to hold true regardless of the forecasting model being used. One of these four characteristics is the statement: forecasts are almost always incorrect since they are merely estimates or predictions. The other three are: forecasts usually assume that the future resembles the past, aggregate or combined forecasts are more accurate than individual ones and the longer the time horizon of the forecast, the less accurate the forecast will be.

A number of empirical studies have shown that the fourth characteristic mentioned by Stevenson is true. For example Lawrence et al. (1985), Brown et al. (1987), Lawrence and Madrikakis (1989) and Hopwood and McKeown (1990) have all come to the conclusion that forecasts with shorter time horizon have proven to be more accurate and less volatile than forecasts with long time horizon. (O'Connor and Webby 1996) An important consequence to this is that the more flexible organizations, which are quicker to respond to changes in demand, and therefore able to make short term forecasts. Hence, they benefit from more accurate forecasts. The reason why aggregate forecasts are generally more accurate is because the random variations of individual demands usually overrule one another. (Stevenson 2007, p. 69)

Madrikakis et al. (1998) state that forecasts should not exclude known information (Case company material [1]). This is backed up by Buffa (1983, p. 57) who states that the planning and control of operations depends on the combination of intelligence about what is actually happening to demand and what is expected to happen. It should also be stated that because demand can be defined as a planned or issued quantity of a product on a desired, promised, planned or issued date from customer orders and return material authorizations, it is sometimes difficult to ascertain what real demand is (Kiely 1999). Because forecasts are derived from demand, it is important for organizations to use a definition for demand which is comparable to the real demand based on which the forecast is made. (Case company material [1]). Stadler & Kilger (2008, p. 156) mention stock-outs as an example in which the case of real demand might cause a problem. According to them a frequent occurrence of stock-outs, which eventually leads to no sales, might imply that the time series is sporadic and therefore the real demand might be underestimated.

2.2.2. The need for forecasting

Madrikakis et al. (1998) state that if there is a time lag between the need to know about an event in order to plan for it, and the occurrence of that event, there is a need for forecasting (Case company material [1]). Buffa (1983, pp. 57-59) argues that a forecast is the single most useful and important data base for operation management decisions and is needed for different planning horizons. These are: short-range, medium-range and long-range. Even though it is generally easier to predict what will happen in the near future, long-range forecasts should not be ruled out. According to Buffa (1983, p. 59) long-range forecasts are needed in plans for capacity and location decisions, changing product and service mix, and the exploitation of new products and services. This is backed up by Stevenson (2007, p. 68) long-range forecasts can prove to be valuable in evaluation of future trends.

Hogarth and Madrikakis (1981) state that medium-range forecasts are usually derived from long-range ones. Medium-range plans include capacities of personnel, materials and equipment for the upcoming one to 12 months (Buffa 1983, p. 59). According to Hogarth and Madrikakis (1981) short-range forecasts are made in accordance to operational planning and managing of production. Short-range plans are needed to plan for current operations and the immediate future. Hence, short-range forecasts are a prerequisite to scheduling the production, stock decisions, distribution, allocation and procurement of resources and managing the supply chain (Stadler & Kilger 2008, pp. 133-134). It is said that short term forecasts are the only part of the forecasting process that can repeatedly create actual benefits and cost saving opportunities (Hogarth and Madrikakis 1981).

In addition to the abovementioned time dimension, there are also two other dimensions along which forecasts can be structured: product and geography (Stadler & Kilger 2008, pp. 135-139). Structuring the forecasts based on product dimension means making forecasts not only for individual final product, but also for different product groups. Forecasting on a group level usually results in a more aggregated forecast. Product groupings can be made in numerous ways based on size, color, packaging, among others and depending on the industry. Another way is to make the aggregation based on geography. In such a case, customers can be grouped by different sales regions or distribution centers. The aforesaid can help determine the key customers (or customer groups) or aid in determining the need of certain raw materials for a specific kind of products. This claim is supported by Mentzer and Moon (2005) who maintain that forecasting should be focused only on the most important customers and products (Kerkkänen et al. 2008). The reason why these kinds of decisions are beneficial on a more aggregate level is because aggregated forecasts are generally more accurate than forecasts made for individual products. (Stadler & Kilger 2008, pp. 135-139)

Kerkkänen et al. (2008) state that because there are many potential sources of information and forecasting requires the combined information from those sources, the number of ways to distribute the responsibility of forecasting grows. Kerkkänen et al. add that insufficiently clear organizational responsibilities are a threat. This would imply that it is not always clear within the company, who should make the forecasts.

Stevenson (2007, p. 69) claims that forecasting is usually the responsibility of the sales and marketing department, since they have access to the best demand information. However, this is only one way of allocating forecasting responsibility. The important factor worth emphasizing here is that the departments, whether sales or others, who should be involved in the forecasting process, are the same ones which have access to the relevant sources of information needed to make the forecast (Kerkkänen et al. 2008). Croxton et al. (2002) are in accordance with the previous statement and emphasize the importance of information sharing between the different functions and departments within the company as an important basis to create more accurate forecasts.

2.3. Forecasting methods

Forecasting techniques are commonly divided into two different categories: quantitative, also known as objective ones and qualitative, also known as subjective ones (Chambers et. al 2004, p. 196). Quantitative techniques involve either the attempt to forecast the future from the historical data or the development of associative models that try to utilize causal (explanatory) factors in order to make a forecast. Quantitative techniques rely on hard data and avoid personal biases, whereas qualitative techniques are subjective and include so-called soft information, such as human factors, personal opinions or intuitions. (Stevenson 2007, p. 70)

Some academics, such as Buffa (1983, pp. 57-58) divide the forecasting techniques to predictive techniques and actual forecasting techniques. According to Buffa the difference between predicting and forecasting is that predicting means integrating subjective and objective information to form an estimate of the future. Predictive methods are used when there is little experience on which to base the future estimates. Forecasting, on the other hand, uses statistical techniques in order to project the historical data into the future. These methods require historical data to be able to describe the record in future terms. Even though the terms are slightly different, the categorization made by Buffa is analogical to the categorization of Chambers et al.

Kerkkänen (2010, p. 26) points out, however, that all forecasting involves human judgment in one way or another. According to Kerkkänen, human judgment can occur either in making the forecast, formulating a forecasting model or selecting the forecasting technique. Additionally, even the most sophisticated models rely at least a bit on human judgment, for example, in the model identification phase or in the selection of the independent variables. The two categories and some of the most common forecasting techniques are summarized in the table 2.1.

Table 2.1. Categorization of most common forecast methods (adapted from Buffa 1983 and Stevenson 2007).

	Quantitative (Forecasting)	**Qualitative (Predicting)**
Objectivity	Objective	Subjective
Methods	- Time series analysis: • Naive • Exponential smoothing • Moving average • Fourier series least squares fit - Causal models: • Simple linear regression • Multiple regression • Econometric models	- Internal expert opinions: • Managers • Sales staff • Delphi - External expert opinions: • Consumer surveys • Industrial surveys - Historical analogy and life cycle analysis

The forecasting techniques summarized in table 2.1 are discussed more thoroughly in the next two subchapters.

2.3.1. Quantitative techniques

As previously mentioned, quantitative techniques include both time series methods and causal models. Time series forecasts use past values of the demand to project the future values. In other words, historical demand data is used under the assumption that the future is like the past and that the time series has some sort of time-related regularity. However, this assumption, which is the basis of the time series methods, is also considered to be the main weakness of said methods. This is because they do not account for other factors (e.g. causality) that have an effect on the demand but merely assume that things are the way they are because they were so before. Time series forecasts are nevertheless quite popular because the ideas behind them are relatively simple and nowadays the calculations can be done very quickly by computers and different statistical softwares. (Chambers et al. 2004, pp. 197–198) Time series techniques are best used when random variability is low (Croxton et al. 2002).

As reported by Stevenson (2007, p. 71) there are a number of different time series techniques. Some of them attempt to smooth out random variations in historical data, whereas others attempt to identify certain patterns such as trend and seasonality and then project these patterns into the future. The simplest of the time series methods is the naïve method, where the forecast of the next time period is the same as the actual demand in the current period. This method is commonly used as a benchmark for other methods: if a forecast of a certain technique is less accurate than that of a naïve one, this technique should be abandoned. (Stevenson 2007, pp. 71–78) Two other very common techniques are moving average and exponential smoothing. Moving average takes the previous *n* periods' demand, calculates their average and then uses this average as a forecast for next period. (Chambers et al. 2004, pp. 197–198) Exponential smoothing forecasts next period's demand by taking into account the actual demand of the current period and the forecast made for the current period. It does so with a smoothing constant that gives more weight to recent periods. (Winters 1960)

Different patterns of times series can be included in different time series methods. The most common of these is the inclusion of either trend, seasonality or both of them. For example with exponential smoothing, the original forecast can be adjusted with the addition of a trend estimate, which is calculated as a difference of the demands of two previous periods. Seasonality can be included by adjusting the forecasts with a seasonal index. To calculate the seasonal index, data from at least the previous twelve months is needed. Calculation of the seasonal indices for each month is done by dividing the monthly demand by the annual average. When both trend and seasonality are included the trend adjusted forecast is further adjusted with a seasonal index. (Buffa 1983, pp. 64-69) The mathematical formulas of the aforementioned methods can be seen in the appendix 2.

While time series techniques try to project the future from past values, the causal models attempt to identify related variables that can be used to predict the values of the variable of interest. The essence of these techniques is to develop an equation that can summarize the effects of the predictor variable (used to predict values of the variable of interest). The most common method is regression. (Stevenson 2007, p. 88) Regression can be a simple linear regression or a more complex multiple regression. Simple linear regression tries to determine the relationship between two variables, whereas more complex models comprise many variables and relationships each with their own set of assumptions and limitations. (Chambers et al. 2004, pp. 200–201) Armstrong and Green (2006) state that in addition to forecasting, causal models can be used to examine the effects of marketing activity, such as price reduction and therefore they provide information for contingency planning.

In addition to the previous there are other more complex econometric forecasting methods. They are an extension of regression analysis. However, they include a system of simultaneous regression equations of several variables. Furthermore, interdependence between the variables usually exists. (Buffa 1983, p. 78) One additional method deserving mention is a method presented by Buffa (1983, p. 58) but not included in the categorization made by Stevenson (2007, p. 68) or Chambers et al. (2004, p. 196) and is the Fourier series least squares fit, which fits a finite Fourier series equation to empirical data, projecting trend and seasonal values. It is used a short-range forecast. However, the Fourier series least squares fit requires at least two years of historical data. These methods do not belong within the scope of this study so they will not be addressed further.

2.3.2. Qualitative techniques

Qualitative techniques are usually used in situations where no historical data is available or management must have a forecast quickly and there is no time to gather and analyze quantitative data. These instances might occur, for example, when launching a new product or when conditions, such as economic or political change and the available historical data may consequently become irrelevant or obsolete. (Stevenson 2007, p. 71) According to Kerkkänen (2010, p. 26) there is a wide range of qualitative methods available. Kerkkänen also adds that they are very difficult to categorize, while the simplest of them are based fully on intuition, whereas some of them are iterative methods or require some team work. This claim is supported by Armstrong and Green (2006) who also list a number of different qualitative methods. Some examples of qualitative methods, as listed by Armstrong and Green (2006) and Stevenson (2007, p. 71), are executive opinions, consumer surveys, opinions of the sales staff and opinions of experts.

Executive opinion forecasts are made by small group of upper-level managers who meet collectively in order to make a forecast. This kind of approach is often used in situations where a new product is being developed. Forecasts made by sales staff is usually considered a good source of information because of the direct contact which sales people have with customers, especially in industrial markets. The drawbacks of these approaches are, however, that the sales people may sometimes have difficulties distinguishing what customers would like to do and what they are actually going to do in addition to personal biases. (Stevenson 2007, p. 71) There are also some empirical studies, such as Winklhofer et al. (1996) which have shown that forecasts made by sales people are notoriously inaccurate. However, Lawrence et al. (2006) present different studies which have concluded that even though the biases' of forecasters can be irrational and lead to suboptimal performance, there are also contradicting findings that show that there are also cases when biases may be rational as well.

One commonly used approach is the Delphi method, which uses a panel of experts (both inside and outside of the company) to answer a series of questionnaires. After the first questionnaire the answers are summarized and made available to the panel to aid in answering the next questionnaire. This process is repeated for several rounds until a convergence of results is obtained. In addition to expert opinions there are consumer surveys or the analysis of consumer behavior, which can be used as extremely valuable input to predict the future market demand. The aforementioned surveys can also be supplemented by referencing the performance of previous comparable kinds of products or product families. This is a case of historical analogy and life-cycle analysis. (Buffa 1983, pp. 79-81)

2.3.3. Integrating different forecasting methods

Even though there are different strategies for selecting a method, there is no technique which consistently outperforms others in varying situations (Chambers et al. 2004, p. 202). However, there are a number of studies, such as a comprehensive review made by Hogarth and Madrikakis (1981) or Mahmoud (1982), which have proven that certain techniques perform better under specific circumstances. For example, time series analysis is usually proven to be a good method in short-term forecasting, whereas causal methods are better suitable for long-range forecasting. However, Armstrong & Yokum (1995) point out that in addition to accuracy criterion, there are also other factors, such as cost and ease of use that should be taken into account when comparing and choosing different forecasting methods.

Lee (2002) presents an important factor that should also be taken into account when making the choice between different forecasting methods. In addition to the choice of method or forecasting approach, the characteristics of the product should also be taken into account. Products with stable demand and long life-cycle (so-called functional products, such as basic household items) should be treated differently than products with highly volatile demand and short life-cycle (so-called innovative products, such as fashion or electronics). Lastly, a choice of method may also be derived from the market (industrial or consumer) in which the company operates (Mentzer & Kahn 1995). and will be discussed further in chapter 2.4.2.

A possibility is also to use both quantitative and qualitative techniques since a combination of these is also possible and recommended in many cases. Some previous studies have shown that the best results in forecasting are achieved by combining two or more forecasting techniques. Mahmoud (1982) concludes in a broad summary of empirical investigations concerning forecasting accuracy that integrating techniques indeed improves forecast accuracy. This is backed up by O'Connor and Webby (1996) who also state forecasts are generally improved when using integrated forecasting techniques.

One of the reasons for improved forecasts is the combined benefit from multiple methods. An example is the integration of unbiased mathematical methods with the information that the mathematical methods do not have available, such as promotional activities or customer feedback. (Stadler & Kilger 2008, p. 142) The previous is in accordance with Armstrong and Collopy (1993), who state that even though statistical methods can make better use of the historical data, the experts might see a lot more in the data than is warranted.

There are different ways of integrating quantitative (objective) and qualitative (subjective) forecasting techniques. According to O'Connor and Webby (1996) the approaches that are most commonly used are combination of two or more different methods and adjustment of statistical forecast with human judgment and sometimes even the other way around. Armstrong and Collopy (1992) state that even though different approaches differ in ease-of-use, credibility and costs, they all have the ability to regularly increase forecasting accuracy. However, it is important to emphasize that when integrating statistical and judgmental methods, the presence of contextual information is of the utmost importance especially when wanting to increase forecasting accuracy (O'Connor & Webby 1996). In other words, there is no point in adjusting the statistical forecast with manual human judgment if there is no additional information available.

One case where combination is proved to produce especially good results is the case of sporadic demand. In this case, the use of the common statistical methods would not make any sense because of the random occurrence of periods with zero demand. Additional judgmental forecasting would probably not increase the quality either. For these items it is recommended to get forecasts with low costs and low time effort for human planners. Hence, there are different procedures for automatic calculation of forecasts for sporadic demand. The purpose of these methods is usually the forecasting of two components, the occurrence of a period with positive demand and quantity of demand, separately. It is proven that these methods are able to significantly reduce the forecast error, if the sporadic demand process has no specific influence on the demand pattern. (Stadler & Kilger 2008, pp. 155-156)

2.4. Forecasting in an industrial context

At this point it should be emphasized that most of the theoretical materials used in this study do not define the differences of forecasting in consumer or industrial markets. When forecasting or demand forecasting is mentioned, especially in operations management or supply chain management literature, it usually implies forecasting procedures in consumer markets. Although there are a lot of similarities, between the consumer and industrial markets in terms of general characteristics of a forecast, forecasting needs and forecasting methods, there are also some differences.

In the previous chapters the aim was to provide the reader with general knowledge of forecasting theory. Even though the theory in those chapters was mainly adapted from literature and journals that did not distinguish the differences between the two different markets, the concepts mentioned in those chapters are still applicable to industrial markets. In other words, the purpose of this chapter is not to dismiss the previous chapters of the literature review but to supplement them and introduce some of the differences and specialties of the industrial markets to the reader, while explaining implications they have to the aforementioned forecasting practices.

Even though the studies used in this chapter, such as Mentzer and Kahn (1995) and Herbig et al. (1993), make a distinction between forecasting practices in consumer and industrial markets, they do not clarify, which type of industry or operational environment the focus is on. Instead, they tend to generalize and merely talk about industrial markets. Therefore, some of the findings presented in this chapter are not necessary applicable to all types of industries. However, they are used in this chapter because of the lack of more specific research on the subject.

2.4.1. The differences between industrial and consumer markets

According to Mentzer and Kahn (1995) industrial markets consist of organizations that acquire goods and/or services in order to use them in the production or offering of other products or services. Alternatively, consumer markets include individual consumers and households who buy goods or services for personal consumption. Mentzer and Kahn (1995) define three special characteristics of industrial markets which differ from consumer markets:

1) Industrial markets have fewer customers
2) Closer relationships between customer and seller is more common
3) The demand for products in industrial markets can be derived from the end-customers’ demand

Since there are fewer customers in industrial markets, the importance of a single customer is far greater than in consumer markets, which makes the demand more volatile (Kerkkänen 2010, p. 18). Kerkkänen (2010, p. 18) points out another factor which increases demand volatility in industrial markets, namely the fact that the demand of those markets can and usually is derived from the end-customers’ demand. This is backed up by Mentzer & Kahn (1995) who state that in the short run, the demand in industrial markets is inelastic, but in the long run it can fluctuate dramatically because of slight changes in the end-customer demand.

Closer relationships with customers could have implications on the availability of demand information. If the relationships are closer, it is possible that demand information is not only available in the form the previous sales data, but also for example in contracts, inquiries, preliminary orders, customers’ inventory levels and production plans, customers’ own forecasts and estimates about the future demand. (Kerkkänen 2010, p. 21)

Stadler & Kilger (2008, p. 156) also present one more specificity present in industrial markets but not in the consumer markets: the case of back-orders. Industrial customers are likely to accept back-orders, if the product is not available. However, this is not the case in consumer markets: if the product is not available, consumers are very likely to take their business elsewhere instead of waiting for the product to arrive, which means lost sales for the company (Chambers et al. 2005, p. 415).

The case of back-orders is closely related to the problem of real demand mentioned in chapter 2.2.1. There it was already mentioned that in case of stock-outs and lost sales the underestimation of real demand is possible. This would also imply that if stock-outs did occur, it would be easier for companies operating in industrial rather than consumer markets to estimate the real demand of their products. However, this obviously assumes that back-orders are possible in industrial markets.

2.4.2. Forecasting practices in industrial markets

In general, the characteristics of a forecast and the forecasting needs of a company are similar regardless of whether the company is operating in consumer or industrial markets. However, the differences in the industrial and consumer markets do require somewhat different business practices. In terms of forecasting, this usually implies the use of different kind of forecasting methods in industrial versus consumer markets. (Mentzer & Kahn 1995)

As previously mentioned in subchapter 2.3.3, there are several studies which handle different forecasting methods, approaches and their popularity. Kerkkänen (2010, p. 41) points out, however, that the major shortcoming of most of the studies is that they rarely distinguish between industrial and consumer companies and are conducted with surveys. Another problem is that these surveys provide information on which methods are being used but not why or how. However, Mentzer & Kahn (1995) state, based on their study, that in industrial markets the preference is usually that forecasts be made by the sales force.

The aforementioned can be justified with the claim that a closer relationship between sales people of the company and the customers encourages companies operating in industrial markets to use their sales force for forecasting. (Mentzer & Kahn 1995) The previous statements are backed up by Kerkkänen et al. (2008) who state that in an environment where demand patterns are more volatile, human judgment plays a more important role in the forecasting than predicting the future demand based on the historical demand. The situation is reversed in consumer markets, where lack of direct customer information forces companies to identify other factors which affect sales or try to extrapolate sales history in order to predict future values (Mentzer & Kahn 1995).

Because of the inability to distinguish between consumer and industrial companies in most of the surveys, as reported by Kerkkänen et al. (2010, p. 41), it is difficult to say that forecasting in one market is regularly easier than in the other or that the accuracies of the forecasts are regularly better in the other. This is true even though the demand is usually more volatile in industrial markets, which makes forecasting a bit more difficult, at least theoretically. However, this is not always the case. In a paper by Mentzer & Kahn (1995), the forecasting accuracy in the two markets was studied and no major differences between industrial and consumer markets in terms of accuracy.

However, there are contradictory findings such as the one made by Herbig et al. (1993). In their study they found that consumer market companies thought that their forecasting processes were more accurate, whereas industrial market companies felt their forecasting processes being less accurate (Mentzer & Kahn 1995). This can be partially explained by Kerkkänen et al. (2008), who state that because most of the forecasting methods have been developed for and are applied in consumer markets, their accuracy targets are also higher in the consumer markets. It must be remembered though that the findings presented in this chapter deal with industrial forecasting in general, and therefore differences between the forecasting practices may exist, depending on the operational environment.

2.5. Forecast errors

One of the mean characteristics of a forecast, as mentioned in subchapter 2.2.1, is that forecasts are most often likely to be incorrect. That is why it is normal to use certain limits of forecast accuracy between which the forecast should remain (Stevenson 2007, p. 69). In order for this to work, forecasting accuracy should be measured and calculated ongoingly. Chopra and Meindl (2001) state that measuring forecasting accuracy serves two main purposes. First, managers can use the error analysis to determine whether the current forecasting method predicts the systematic component of demand accurately. Second, managers are able to estimate forecasting error because a contingency plan should account for such an error. (Kerkkänen 2010, p. 32)

Stadler and Kilger (2008, pp. 149) are in agreement with ongoing forecast accuracy measuring as they state that forecast error is an important building block in the forecasting process, because it can be used to check the performances of both statistical and additional judgmental input. Calculation of forecast errors is important also because safety stock calculations are usually based on forecast error. This is highly important because safety stock is the key factor which affects the service level of the supply chain. (Stadler & Kilger 2008, pp. 149). Mentzer and Moon (2005) affirm that it is also important to use metrics which relate the forecast accuracy to performance measurement of the company, such as costs or customer service (Kerkkänen 2010, p. 34).

2.5.1. Error measures

Measuring the actual accuracy of the forecasts can be done in a number of ways. Mentzer and Moon (2005) presents a categorization of error measures. According to them there are three categories. The most common categories are absolute and relative measures, but there is also a third category which relates the forecasting technique to another technique. (Kerkkänen 2010, p. 33) Absolute measures are all based on calculating the difference between actual sales and forecasts in different ways. It is worth emphasizing, however, that the basis of all measures, not only the actual measures, is the simple forecast error which is the difference between the actual and the forecasted quantity (Winters 1960). Based on that error, more sophisticated calculations can be made and then used to evaluate or compare the accuracy of the forecast. The most common absolute error measures are mean error (ME), mean absolute deviation (MAD) and mean squared error (MSE) (Buffa 1983, p. 64). These can be used for different purposes. For example, the mean error shows whether the forecast is continuously too high (positive ME) or too low (negative ME), meaning whether or not there is bias in the forecast (Buffa 1983, p. 64).

The most common relative measure is the mean absolute percentage error (MAPE), which shows how much, measured by percentage, an individual forecast or forecasts on average deteriorate from the actual demand. It can also be used to compare the quality of the forecasts by comparing them: the lower the MAPE, the more accurate the forecast. (Stevenson 2007, pp. 93-94) According to some studies (e.g. Mentzer & Cox 1984 or Mentzer & Kahn 2004) MAPE is one of the most popular error measures. According Mentzer and Moon (2005) an example of a method that compares the forecasting technique into another technique is Theil’s U, which calculates the ratio of the accuracy of the technique to the naive forecast. If Theil’s U is less than 1, the method being used is better than the naïve method. However, if Theil’s U is more than or equal to 1, the naïve method is as good as or better than the forecast model chosen and should therefore be used. (Kerkkänen 2010, p. 33) The mathematical formulas of the error measures are presented in the appendix 3.

The shortcoming of some methods could be, for example MAD and MSE, that they are absolute quantities, and thus they cannot be benchmarked against or compared to other products with higher or lower average demand (Stadler & Kilger 2008, p. 151). Additionally, not all methods are suited in all environments, for example, MAPE cannot be used if the demand is intermittent (Kerkkänen et al. 2008). Hyndman & Koehler (2005) are in accordance with this and add that the other problem of MAPE is that it is constantly larger than its corresponding median average percentage error (MdAPE), of which use could be more applicable than the use of MAPE. In general, it is difficult to say, which error measure is the best one since there are a number of different opinions, depending on the researcher.

For further discussion of comparison and the problems of error measures, Hyndman and Koehler (2005) present a critical view towards most of the traditional forecasting measures, including some of the ones that were discussed in this chapter. Additionally, Hyndman and Koehler (2005) also present some modifications for the popular MAPE and their own point of view on the subject and also some of the previous conclusions made by other researchers for the best accuracy measure. However, since the forecasting software does only include most of the traditional values (e.g. MAD, ME, MAPE) some of the more complex values, such as the ones of Hyndman and Koehler (2005) are not addressed further.

In addition to the calculation of forecast accuracy, there are some other values which can be used to investigate forecasts. Milliken 2006 introduces an additional value, which is the Coefficient of Variation (COV). According to Milliken, the COV is a value which can be used to determine whether or not an item is forecastable. It is calculated as the ratio of Standard Deviation of Demand and Average Period Demand: the more volatile the demand, the higher the COV, which also makes forecasting more difficult. When calculating the COV, a minimum of 12 months of demand data should be used. According to Milliken, if the COV is equal to or less than 0.8, the item is forecastable. However if the COV is more than 0.8, the item is not forecastable, which means that it should be managed some alternative way. Milliken adds that there are some exceptions, such as seasonal products, which may have a high COV value but can be still be forecasted successfully.

2.5.2. Importance of forecast accuracy and the costs of forecast errors

Perhaps the greatest disadvantage of forecast error measures in general is that even though accuracy plays a huge role in selecting forecasting methods and evaluating their performances, there is no universally accepted measure of accuracy. The problem is as well that there are no certain error limits between which the forecast should always be. Instead, the degree of accuracy that is sufficient for their purposes is usually defined by the company itself. For example a 10 % MAPE may be sufficient for some companies whereas an MAPE of more than 5 % can be a disaster for other companies. (Mahmoud 1984)

Bunn and Taylor (2001) state that another problem with accuracy measurement is that it is often difficult to receive 100 % valid information about the actual demand. This is due to the fact that demand is often manipulated with, for example, price discounts and delivery dates. Kerkkänen (2010, p. 33) Mahmoud (1984) concurs with the aforementioned and adds that the problem arises from the fact that in real life forecasting situations, the forecaster must always start with the data available for certain forecasting problems and this data is not necessarily the data that should be used, at least based on theory.

Even though there is no universally accepted measure of accuracy or value of a good accuracy, there are studies (e.g. Mentzer and Cox 1984, Mentzer and Kahn 1995) which conclude that accuracy is the most important criterion when evaluating forecasting performance. If the forecast is accurate it can be offset through poor planning, whereas an inaccurate forecast can ruin the best of plans (Mahmoud 1984). As it was already mentioned that one of the purposes of measuring forecasting accuracy is that managers are able to estimate the effect of the forecast error and make a contingency plan accordingly. It is also suggested that the forecast accuracy and forecast errors should be linked to business performance (Kerkkänen et al. 2008)

A way to link forecast error to business performance is to link the forecast accuracy to costs. According to Mentzer and Moon (2005) the forecasting costs can be divided into two categories: costs of making the forecasts and costs resulting from forecasting errors. The cost of making forecasts includes the forecasting software, personnel, training and time taken from other activities in the company (Kerkkänen 2010, p. 34), which is not always taken into consideration when assessing the costs. For example, a study conducted by Mentzer and Kahn (2005) in which the forecast accuracy was the most important criterion for the respondents, concluded also that the cost of forecasting or return on investment were not seen as important criterions when evaluating the performance of forecasting, implying that making the forecasts is not necessarily evaluated based on financial measures like the effect of forecast errors.

Forslund and Jonsson (2007) state that supply chain performance is typically related to metrics reflecting cost, tied-up capital and customer service while suppliers may need to use internal actions in order to compensate poor customer service, which is usually the source of the costs. This is observably related to costs of forecast errors, since the purpose of forecasts is to predict the customer demand as accurately as possible and therefore, in the case of inaccurate forecasts, some internal actions may be needed. Forslund and Jonsson (2007) divide the internal actions to corrective or preventive actions. Corrective actions may include rush orders and overtime, whereas safety stocks and extra capacity are measures of preventive actions.

All of these aforementioned actions create costs in their own way. For example, safety stock or extra capacity and excess inventory will obviously increase inventory holding costs and can also lead to obsolescence or deterioration with age (Chambers et al. 2004, p. 415). Another source of costs in inaccurate forecasting includes costs that occur when the actual demand is higher than forecasts. In addition to corrective actions, there are the costs of lost sales or stock-outs. As mentioned in subchapter 2.4.1, the possible case of stock-outs and lost sales is more present in the consumer than in the industrial markets, because of the fact that industrial customers are sometimes ready to accept backorders, unlike consumers who usually take their business elsewhere when the product they want is not available.

Even though some studies based on, for example, surveys responses emphasize the importance of forecast accuracy, they do not mention in which operational environment the accuracy criterion is especially important. There are certain specific products or industries in which the costs of forecast errors are especially high. Lee (2002) focuses on this by making a distinction between functional and innovative products. The accuracy of forecasts is especially important in the case of innovative products. Examples of innovative products are fashion or electronics and other industries with highly volatile demand. In these cases lost sales or excess inventories – which might lead to selling at mark-down prices or selling at a loss because of the short age of product could ,in the worst cases, exceed the total costs of manufacturing.

As previously mentioned, there are different reasons for costs caused by forecast error, which can make it difficult to ascertain the actual costs. A measure that is used later, in chapter 5 of this study, is Value of MAD (or Value of MAE in the software), which is calculated (for single product) by multiplying the mean average deviation by inventory value of the product (appendix 3). This is a fairly simple way of calculating the cost of an error, or a multiple errors, because it only takes the inventory value into account. In the case where an error is negative, which means that the forecast is larger than the actual demand, a proper estimate about the inventory holding costs is obtained and could be suitable for products which do not lose their value in the inventory. Additionally, in the case where the forecast error is positive, which means that the actual demand is larger than the forecast, the cost of error is not determined by the inventory holding costs, but for example corrective actions or lost sales. In those cases the Value of MAD does not necessarily give a proper estimate about the actual costs.

3. DEMAND FORECASTING PROCESS

This chapter is the second section of the theoretical part of this study. The forecasting practices and theory, which were presented in the first section of theory (chapter 2) of this thesis, are now taken into more thorough inspection. In this chapter forecasting, or demand forecasting, procedures are analyzed as part of a larger entity, which in this case is called the demand forecasting process. The problem is, however, that there are different definitions of said process depending on the source.

The purpose of this chapter is to create a framework of the demand forecasting process which is suitable for the special requirements, derived from the attributes of the case company of this study. In other words, the aim of this chapter is not to create new knowledge about the topic but to merge different previous studies surrounding the area to suit the specificities and structure the content of this study.

3.1. Defining the concept of demand forecasting process

When trying to define what the demand forecasting process is, there are two main problems that have to be dealt with. The first problem is that demand forecasting, as a process, is not often discussed in the operations management or supply chain management literature or previous studies, to which the forecasting theory also belongs. This is backed up by Kerkkänen (2010, p. 39) who also found that there is no single established way to describe the demand forecasting process. It is very often the case that the literature or previous studies around the area focus only on specific individual areas of demand forecasting, such as using proper and valid demand data, describing different methods and comparing their accuracy or debating which accuracy measure is the most suitable one for error measurement. In other words, it is usual that only the individual parts of demand forecasting process are discussed, not the whole entity.

The aforementioned statement is backed up by Kerkkänen (2010, p. 40) who states that the forecasting process is often divided into smaller tasks, after which suggestions are provided on how to perform these tasks. However, the thing that is common for all individual tasks of forecasting is that the main purpose of them is to produce more accurate forecasts. In other words they can be characterized as specific actions with a one common goal, which is exactly what a process is by definition (Merriam-Webster Dictionary). Davenport (1993) also defines a business process as a structured set of activities designed to produce a specific output (Bititci & Muir 1997). Thus, when all the different actions and tasks of demand forecasting are linked and combined, it can be said that when talking about demand forecasting, we are talking about a business process that serves a specific purpose and produces an output, which in this case is the final forecast.

Before it was mentioned that there are two problems related to definition of the demand forecasting process and one of them was that the term is not often seen in operations management or supply chain management literature. The second problem is partially linked with the first one. When the term is actually discussed, it can sometimes be mixed with two other similar terms: demand management and demand planning. When one or more of these terms are used, the distinction between them can be relatively unclear from time-to-time since they are at least closely related but sometimes even used interchangeably depending on the source or the context.

For example, Croxton et al. (2002) define demand management as a supply chain management process which balances the customers' requirements with the capabilities of the supply chain. The process is however, not only limited to forecasting. It includes several sub-processes which attempt to synchronize supply and demand, increase flexibility and reduce variability. (Croxton et al. 2002) Another definition is by Chambers et al. (2005, p. 487) who define demand management as the management of customer orders and sales forecasts, including several processes such as sales order entry, demand forecasting, order promising, customer service and physical distribution. Figure 3.1. shows how Croxton et al. (2002) link demand forecasting to companies' other processes and to supply chain management.

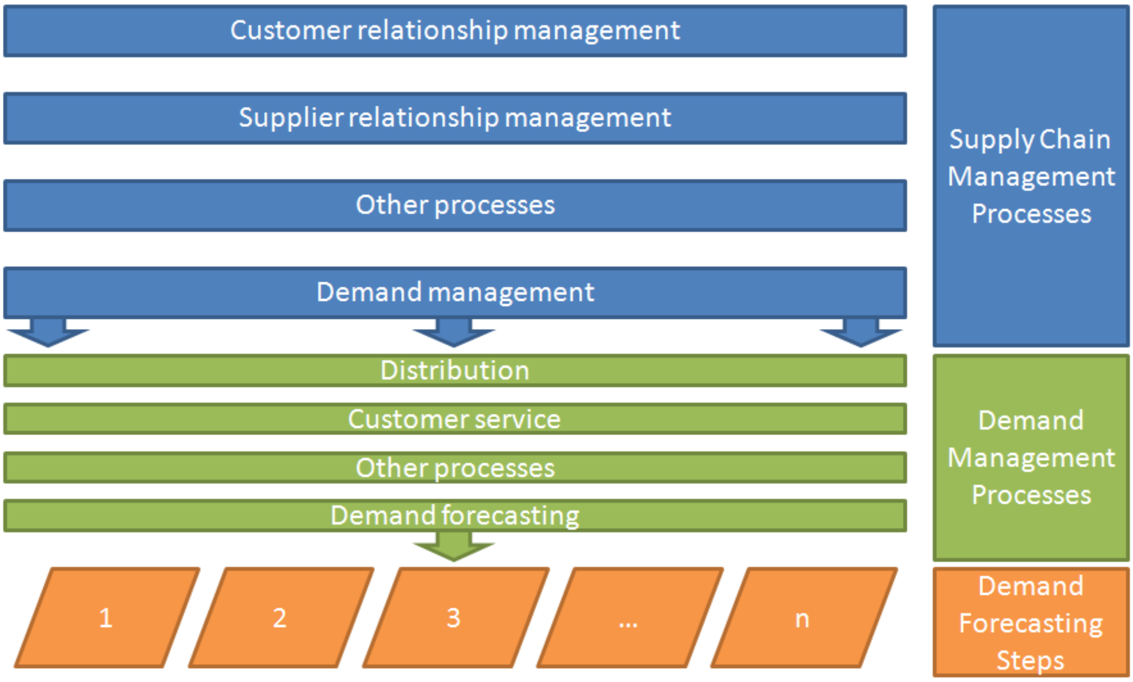


Figure 3.1. Demand forecasting as a part of demand management (adapted from Croxton et al. 2002).

In other words, both Croxton et al. (2002) and Chambers et al. (2004) see demand forecasting process as a sub-process of demand management process. Sometimes demand forecasting is also discussed in relation to planning as, for example, in the case of sales and operations planning (S&OP).

Muzumdar and Fontanella (2006) define sales and operations planning as a set of processes which allow an enterprise to respond effectively to demand and supply variability. This is backed up by Olhager et al. (2001) who state that sales and operations planning is often referred to as a fundamental which maintains a balance between aggregate supply and aggregate demand. Olhager et al. (2001) see sales and operations planning as a long-term planning of production and sales relative to forecasted demand and the supply of capacity. Therefore, they also divide S&OP roughly into a sales plan, which is made based on forecasted demand and a production plan, which affects inventory and capacity requirements.

The aforementioned is very close to the concept of forecasting and planning (F&P) described by Hogarth & Madrikakis (1981). They state that planning requires the existence of values and goals, alternative courses of action, the assessment of those alternatives and the implementation of alternative selected, whereas forecasting is used in the generation and assessment of the previously mentioned alternatives. A similar definition is used by Stadler and Kilger (2008, pp. 139-157), who discuss the demand planning concept.

Stadler and Kilger (2008, pp. 139-157) state that the purpose of demand planning is to determine the planned sales and forecasts and to improve decisions which affect demand accuracy. For forecasting this would mean predicting future sales and for planning, for example, safety stock calculations in order to reach a predefined service level. Demand planning can be linked to demand management and again further to supply chain management is the following way: a demand plan is the basis for the performances of each supply chain entities, which are the basis of the overall performance of the entire supply chain (Stadler & Kilger 2008, pp. 139-157).

As can be seen, the second problem regarding the differences between the three terms is somewhat complicated. Even though there might be some differences in the concepts depending on the author, or the case study in which they are discussed, what is similar in them is that they separate forecasting and planning. Additionally, demand forecasting is often seen as a part of demand planning even though it might not fall under the category of planning. This is discussed by Stadler and Kilger (2008, pp. 139-157) who state that forecasting is not a real planning or decision-making process as it only aims to predict the future as accurately as possible without influencing the demand. Therefore, the forecasting process of demand planning lays the foundation on which the planning process can be based.

Even though Stadler and Kilger (2008, pp. 139-157) describe a multi-step demand planning process, most of the phases of the process actually have more to do with forecasting than planning. Bearing that in mind that the premise of this study is to focus simply on the demand forecasting process, we can apply certain related concepts presented by Stadler and Kilger (2008, pp. 139-157) when describing process suitable for this particular case study. This is important because, as previously mentioned, even though there are a number of different studies which describe different procedures of forecasting, the existence of previous frameworks or process descriptions is very limited. Therefore, the process described by Stadler and Kilger (2008, pp. 139-157) can be used as a basis for demand forecasting process discussed in the next chapter.

3.2. Description of the demand forecasting process

In their description of the demand planning process, which is later referred to as demand forecasting process, Stadler and Kilger (2008, pp. 140-160) divide the process into six different phases. The process of Stadler and Kilger consists of following steps:

1. Preparation of demand planning structures and historic data
2. Computation of statistical forecast
3. Judgmental forecasting
4. Consensus forecasting
5. Planning of dependent demand
6. Release of the forecast

As it was already mentioned in the previous chapter, that even though the process described by Stadler and Kilger (2008, pp. 140-160) is called the demand planning process, most of the steps have more to do with forecasting than planning. The one exception is the fifth step of the process, which is planning of dependent demand, which does not fit to the scope of this study and thus is not discussed in detail. In short, it is done by computing the demands of different components and raw materials needed to make or assemble the final product. The dependent demand is computed based on the forecast of the demand of final product. Planning of dependent demand is of course determined also during master planning and materials requirements planning. (Stadler & Kilger 2008, pp. 142-144)

Another part of the process that does not entirely fit to the scope of this study is the first step, preparation of demand planning structures and historical data. Later on in this study when the analysis of the demand forecasting procedures of the case company is analyzed the focus is not on the gathering of the demand data but more on the analysis and the proper interpretation of that data. Because of this, chapter 3.2.1, where the matter is discussed, focuses also on the principles of the analysis and updating of the data, not on how it is or should be gathered.

The shortcoming of the process description by Stadler and Kilger (2008, pp. 139-157) is that it lacks an important part and that is the measurement stage of the process.

According to Kerkkänen et al. (2008) the lack of performance measurement is a common shortcoming in the forecasting studies since the previous studies and literature emphasizes producing the forecasts instead of their use in decision making. Kerkkänen et al. (2008) also state that based on empirical research, producing the forecasts is managed much better than is the evaluation of its impacts and for example, in many companies the forecast errors are measured, but the errors' impacts are not assessed equally well. That's why the inclusion of the measurement stage can be seen as an important addition to the process by Stadler and Kilger. Measurement, as well as other phases, will be discussed further in the next subchapters.

3.2.1. Preparation of demand data and computation of statistical forecast

As it was already mentioned the process starts with the preparation of demand data, which consists of gathering and updating the historical data regarding demand. In addition to the gathering of sales figures, the updating of data includes for example possible changes in product groups, including new products and deactivating products that will not be sold anymore and, hence will not be forecasted either. It is of utmost importance that the forecast data does not have any inconsistencies, such as different quantities between the product levels (e. g. subgroup, product and packaging levels). (Stadler and Kilger, 2008, pp. 140-142) Croxton et al. (2002) state that an important aspect of updating the data is to take into account, not only the sales figures, but also the possible returns. For example in a situation where a lot of returns have occurred, taking into account only the original sales the overall numbers of the demand will be inflated and hence the forecast might be unnecessarily overestimated.

Recognizing the demand pattern is important part in the second phase because at that phase a proper statistical technique is chosen based on the demand pattern of the product of which future demand it tries to predict (Stadler & Kilger 2008, pp. 140-142). However, this is also important in the preparation phase because at that phase different kinds of unusual variations, such as promotional activities or introduction of new products should be identified and cleaned from the demand data for their inclusion in the time series might distort the overall picture (Stevenson 2007, pp. 72–73). In addition to the demand history, there are also other possible sources of information such as, contracts, inquiries, preliminary orders and customers' future plans depending on the operational environment. The aforementioned other sources are more available in industrial markets than on consumer markets as it was mentioned in chapter 2.4.1.

The second step of the process includes computation of the statistical forecast based on the updated demand. Nowadays, this is normally done with the help of software designed for this purpose. As mentioned in the chapter 2.3.1 there are different ways and techniques to do the statistical forecast. The selection – whether it is automatically done by the software or manually by the user – of the proper technique should always be based on the demand data and the pattern of the time series. For example a time series with seasonality should be the basis of a forecast technique that has a seasonal adjustment. The importance of the availability of enough historical data that is needed to recognize the patterns of the time series and get statistically significant results cannot be emphasized enough. For example in order to recognize seasonality at least two or three seasonal cycles (usually a year) are needed. (Stadler & Kilger 2008, pp. 141-148)

Armstrong and Green (2006) list a number of statistical methods that can be used and emphasize the importance of the regular comparison of different forecasting methods and their accuracy. The comparison can be done with many different accuracy methods, which were presented in chapter 2.5.1. Stadler and Kilger (2008) are in accordance with the aforementioned and state that it is important that the selection of the statistical forecasting model and the estimation of its necessary parameters should be done more or less frequently. Nowadays this can be done very easily with statistical software for they usually include a so-called Pick-the-Best or Best Fit -option. (Stadler & Kilger 2008, pp. 156-157)

The advantage of the Best Fit -option is that the system searches all available statistical forecasting techniques and different parameter combinations and then selects the one which is able to produce the best accuracy in the specified time-segment that is defined by the user. Hence, the user does not have to check if a given model is suitable with the time series under consideration. However, Best-Fit should not be always used as an only forecasting tool but more of a guide in search for the appropriate procedure or technique. This is because historical data is not always available enough, especially in the beginning of the process. The second reason is the one already mentioned in chapter 2.3.3 that the forecast accuracy should not always be used as a sole criterion in determination of the appropriate forecast method. (Stadler & Kilger 2008, pp. 156-157)

3.2.2. Judgmental input, consensus forecast and release of the final forecast

In the third phase of the process judgmental forecasts are made. This includes forecasts made by different departments of the company, usually sales, product management and marketing (Stevenson 2007, p. 69). Croxton et. al (2002) emphasize that at this point the importance of knowledge sharing between the different departments of the company is pivotal. For example, it is important because certain campaigns or price reductions, that marketing department is aware of, can cause momentary changes in the demand and this should be acknowledged to the production department by marketing. In that case production department is able to understand that the possible change is merely momentary and they can adjust their operational decisions based on that.

According to Armstrong (2006) the choice between judgmental forecasting techniques depends on the life cycle of the product and in which phase of that cycle the product is currently. This is because certain methods provide more accurate forecasts at different phases of the product life cycle. It is different trying to forecast the demand of a new product in comparison to one that has already been in the market for some time. Armstrong (2006) states as well that nowadays the movement has been from purely judgmental approaches to quantitative models. Stadler & Kilger (2008, pp. 142-143) support this by stating that integrating statistical and judgmental forecast is reasonable only, if information is not double counted in both forecasts. An example would be the already mentioned unusual variations that are removed from demand data before statistical forecast.

Making both statistical and judgmental forecasts is not always necessary and not even possible. For example forecasting a demand of a new product without any demand data cannot be done with statistical methods, in which case judgmental methods would be used. Based on literature and previous studies, there has been a lot of controversy which one of these methods should be preferred in forecasting in general. There are studies (e.g. Makridakis 1988) that claim that if there are statistical techniques available they should be the number one priority and that the judgmental forecasts should only be used in special circumstances (Bunn & Wright 1991). However, there are also contradicting studies that show a well structured judgmental process outperforming various statistical measures (Lawrence et al. 2006).

Stadler & Kilger (2008, pp. 142-143) argue however that the integration of forecast methods is a key factor that determines the efficiency of the demand forecasting process. This is backed up by some of the previous studies on forecasting accuracy (e.g. Mahmoud 1982, Lawrence 1986, O'Connor & Webby 1996). As it was mentioned in chapter 2.3.3 there are a number of ways to integrate forecasting methods. However, the most commonly used are combination and judgmental adjustment (O'Connor & Webby 1996).

Combination of forecasting methods means using two or more different forecasting methods and making the forecast based on the forecasts of these methods. The methods themselves can vary substantially: it is possible to use different statistical methods, different judgmental methods or a mixture of both statistical and judgmental methods for example. When combining methods it is important to have formal procedures for the combination. This means that a different weight is to be given to different methods, for example giving equal amount of weight to each method. However, it is important to point out that it is ideal to give the weights mechanically. Thus, no bias occurs when giving weights to certain methods, while people might have a tendency to favor one method above others, even though there is no ground for it. Weights can be given for example based on historical accuracy of methods being used. (Armstrong 2001)

According to Armstrong and Green (2006) judgmental adjustment means making revisions based on a judgment of an expert (or a person who is responsible) to a statistical forecast. However, this should only be done based on predefined triggers

(Stadler & Kilger 2008, p. 142) such as promotions or campaigns. The reason for this is the one already mentioned in chapter 2.3.2 that sometimes judgmental forecasts involve hidden agendas of the people making them. O'Connor & webby (1996) support this and state that judgmental adjustment is effective when it is based on contextual information and add, that if the adjustment is made based on some other information or incentive, the outcome could be contradictory.

Armstrong and Green (2006) introduce another way of adjustment as well. In this case a judgmental forecast is created based on relevant data. After that a forecast is made using statistical methods. Then, the original judgmental forecast can be revised based on the statistical forecast. It is argued that these methods should be used because the judgmental forecasts should be used as inputs to statistical forecasts, not just to adjust the outputs. (Armstrong and Green 2006) Stadler & Kilger (2008, p. 142) state the advantage is as well that this method leaves more control of the process to the human planner. Regardless of the way it is done, forecast resulting from the integration process is the basis of a consensus forecast, which is the fourth phase of the forecasting process.

The purpose of the consensus forecast is to settle possible open issues, such as influence of promotions or campaigns. This can include for example a what-if-analysis that enables the user to view the consequences of different scenarios and actions, which then allows the user to plan promotions, new product launches and other things that have or might have an effect on the demand even further. If there are multiple departments of people involved, making the final consensus forecast can be done by weighing the forecasts of different departments or experts based on the accuracy that is achieved in the past. An important factor here is the feedback that needs to be given to the people involved in forecasting. (Stadler & Kilger 2008, pp. 142-143) Importance of feedback is discussed further in the next subchapter. The last step of the process described by Stadler & Kilger (2008, p. 144) is the release of the demand forecast. At this step the demand forecast is formally approved and made available for other processes.

3.2.3. Measurement of forecasting process

In the beginning of subchapter 3.1 it was mentioned that demand forecasting is a process and the output of that process is the forecast. In order to have information about the quality of the process it must be measured somehow: "When you can measure what you are speaking about, and express it in numbers, you know something about it" (Lord Kelvin 1824-1907). Importance of measurement of performance has been recognized by academics and practitioners for a long time (Neely et al. 2005). Behn (2003) argues that measurement of performance can be used to evaluate, control, budget, motivate, promote, celebrate, learn and improve. Neilimo and Uusi-Rauva (2005, pp. 300-301) concur with this and list seven different purposes of measurement. According to them measurement, at its best, motivates, emphasizes the value of the thing that is measured, guides to focus on right things, sets clearer targets, helps communication, causes rivalry or competition and helps with compensation.

Neely et al. (2005) define performance measurement as quantifying an action. In the case of the demand forecasting process and its measurement the focus would be on the output of the process, which is the final forecast. In order to do this, there should be a proper performance measure, which is a metric that quantifies the process (Neely et al. 2005). In the case of forecasting, it could be the accuracy of the forecast. Stadler and Kilger (2008, p. 149) state that one of the reasons why measuring the forecasting process is important is because the other processes in the company which use forecasts as foundations for their decisions, such as pre-production or procurement require some sort of quality measure in order to comprehend the forecast accuracy and the dimension of the possible deviations of the forecast from the actual demand. Therefore, they need to be sure of the quality of the forecast when making their own decisions.

Stadler and Kilger (2008, p. 149) call the measurement part of the forecasting process the controlling phase and state that its purpose is to control the quality of the forecast and additionally the quality of the process itself. Stadler and Kilger (2008, p. 149) add that the other purpose of forecast controlling is to provide information of the forecast quality and offer feedback based on the quality to the contributors in the forecasting process. This is in accordance with the previous statements made by Behn (2003) and Neilimo and Uusi-Rauva (2005), which were regarding the purposes of performance measurement. To summarize, the performance measurement of the demand forecasting process includes three larger entities: measuring the output (calculating the accuracy of the forecast), modifying the process based on the results of the measurement and giving feedback to people involved in the process.

Some authors claim that measuring forecast errors improves forecast accuracy (e.g. Wacker and Sprague 1995 or Mentzer and Moon 2005), but the mere measurement of errors does not provide information that is sufficient for setting targets for forecast accuracy and finding development areas in the demand forecasting process. (Kerkkänen et al. 2008) This is backed up by Behn (2003), who states that what people measure is not what they always want done and that even though performance measures shape behavior, they may shape it in both desirable and undesirable ways. For this reason, Behn (2003) emphasizes the importance of selecting proper performance measures, mentioned previously in this chapter. In the case of forecasting, the measure that is normally used is the accuracy of forecasts. However, Kerkkänen et al. (2008) state that even though the performance measurement includes other metrics than accuracy, accuracy measurement is nevertheless a part of performance measurement of forecasting.

However, as was also mentioned in the subchapter 2.5, the most common problem is that there is no universally accepted measure for accuracy and therefore it is up to the company itself to decide what measure they use and what values they consider to be accurate or inaccurate. Additionally, not all of the accuracy measures are applicable in all environments. Kerkkänen et al. (2008) state that accuracy targets should be set and in addition roots of the forecast errors should be found to make corrective actions. Gardner (1983) agrees with this by stating that forecast methods should be regularly monitored in order to ensure that forecasting system remains in control and the correct forecasting method or technique is being used. Moreover there should be a reaction to possible unacceptable forecast errors and model parameters or the forecasting technique altogether should be changed if necessary. This can be seen as a part of the process modification based on the results of measurement.

Croxton et al. (2002) are in accordance with the aforementioned and state that with monitoring of forecasting accuracy it is easy to notice the possible presence of systematic error made by the forecasting model. To help monitor the forecast accuracy and its development, the forecast accuracy should be calculated over time (i.e. calculating the forecast accuracy each of past time periods), which is the most common method seen in forecasting literature and studies (Stadler & Kilger, 2008 p. 141). To help monitor forecast accuracy Gardner (1983) suggests the use of a control card, which can be seen in the figure 3.2. The control card in figure has been taken from the forecasting software used by the case company of this study and it shows error values, which have been cleared from the figure because of the confidentiality agreement, of a certain product of the case company.

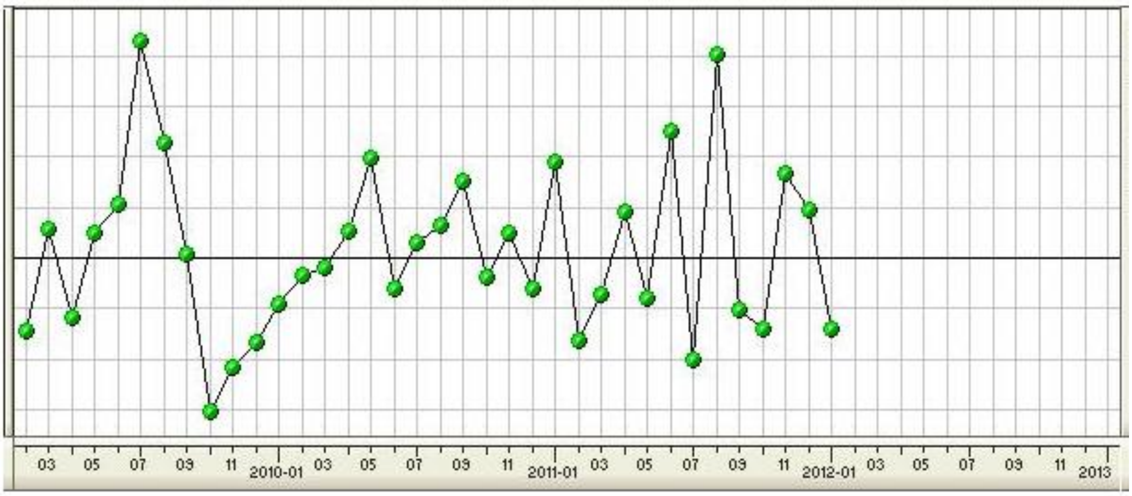


Figure 3.2. Control card (adapted from the Forecasting Software used by the case company).

Stevenson (2007, p. 95) supports the use of "control card" by stating that it is an excellent visual aid to help notice systematic or unusual errors and if the forecast is in control. Examples of these sorts of errors would be that there are relatively more errors on one side of the mean error (ideally zero) or that the errors show a trend or some other pattern. This is in accordance with Gardner (1985), who states that in order for the forecasting system to remain in control, the cumulative sum of forecast errors should fluctuate around zero.

If multiple departments are involved in the forecasting process it could be useful to calculate the Forecast Value Added (FVA), which measures if certain steps or phases in the overall forecasting process are paying off. An example would be to compare the statistical forecast to a naïve one. The purpose is that every successive step of the forecast process is adding value to the process and therefore, the statistical forecast should be more accurate than a naïve one. After that a possible revised judgmental forecast made by a certain department or a person can be compared to the statistical forecast. The accuracy measures presented in subchapter 2.5.1 can be used when making the comparison. (Stadler & Kilger, 2008 p. 152)

In addition to quality measurement and possible modification or improvement of the process, it should be noted that communicating and feedback of the results of the measurement is also important. This is supported by Croxton et al. (2002), according to whom an important part of learning process of forecasting is that after analyzing the forecast errors it is important to fine-tune the forecasting methods and give feedback to the people involved in the process. Stadler and Kilger (2008, p. 149) support this by stating that the quality of the forecast should be used as a feedback mechanism for the people involved in order to them to receive information about the quality of their contributions. These statements are in accordance with O'Connor and Webby (1996), who affirm that feedback has been found to have a beneficial effect on task performance.

O'Connor and Webby (1996) point out, however, that the effect depends also on the type of feedback. For example, presenting a summative error or progress of the process may be more efficient than presenting a simple value. Neilimo and Uusi-Rauva (2005, p. 304) agree with this and state that the in order to take advantage of measurement information in decision making, it is important that the information is: reliable, up-to-date, meaningful and that it is presented in a proper and clear way. The importance of feedback type is also discussed in other studies. For example, Brehmer (1980) argues that presenting a simple value or an outcome does not facilitate performance because outcomes are probabilistic and hence, it is not possible to learn from them. Another reason is that the end result is always affected by a number of different relations and thus, representation of merely the end result does not provide any information how the relations actually behave (Todd and Hammond 1965). In addition to this, providing information on errors of individual forecasts can improve the ability to combine statistical and judgmental forecasts. (Fischer and Harvey 1999)

3.3. Summary of the demand forecasting process

This chapter summarizes the individual aspects of the demand forecasting process described in chapter 3.2 and depicted in figure 3.3. It should be emphasized that the process described in this section of the study is based on the process description made by Stadler and Kilger (2008, pp. 140-160) and a compilation of different preceding studies and research regarding different forecasting practices and approaches. It must be remembered that most of the studies, which have been used as a basis for this, build upon the specificities of those prior studies, not as general frameworks about the demand forecasting processes of companies.

However, as mentioned in the beginning of this subchapter, the purpose was not to create a universally applicable demand forecasting process description but to structure the contents of this thesis and to create a framework for this particular case study. Additionally, as mentioned in subchapter 2.4, most of the forecasting research focuses on consumer markets and there is a general lack in distinguishing the difference between forecasting practices in consumer or industrial markets. Therefore it can be suggested that the process is perhaps more applicable in consumer than in industrial markets.

Another shortcoming of prior research is the lack of knowledge about the operational environment or the specific industry or branch, in which a company is engaged, and its effects on the forecasting practices. However, because of the lack of previously created frameworks or process descriptions for specific operational environments, the demand forecasting process described in this chapter is used later throughout this particular case study, whether in the context of consumer or industrial markets.

Another reason why this particular framework is used in this thesis is because it includes all the different parts or aspects of forecasting that fit to the characteristics of this study. It must also be remembered that certain characteristics and limitations of this particular case study, like the exclusion of planning of dependent demand, have had influence on the demand forecasting process that was described earlier. Hence, the demand forecasting process seen in figure 3.3 is not necessarily applicable to later case studies about the subject.

To summarize, the demand forecasting process starts with the stage of making the forecast, which consists of following steps:

1) Preparation of demand planning structures and historic data
2) Computation of statistical forecast
3) Judgmental forecasting
4) Consensus forecasting and release of forecast

After the initial forecast has been made comes the stage of performance measurement of the process, which has three additional steps:

5) Calculation of forecasting errors
6) Modification of parameters
7) Performance feedback

Taking both of the aforementioned stages (forecast & performance measurement) into account, we have a demand forecasting process consisting of seven different stages, which is summarized in figure 3.3.

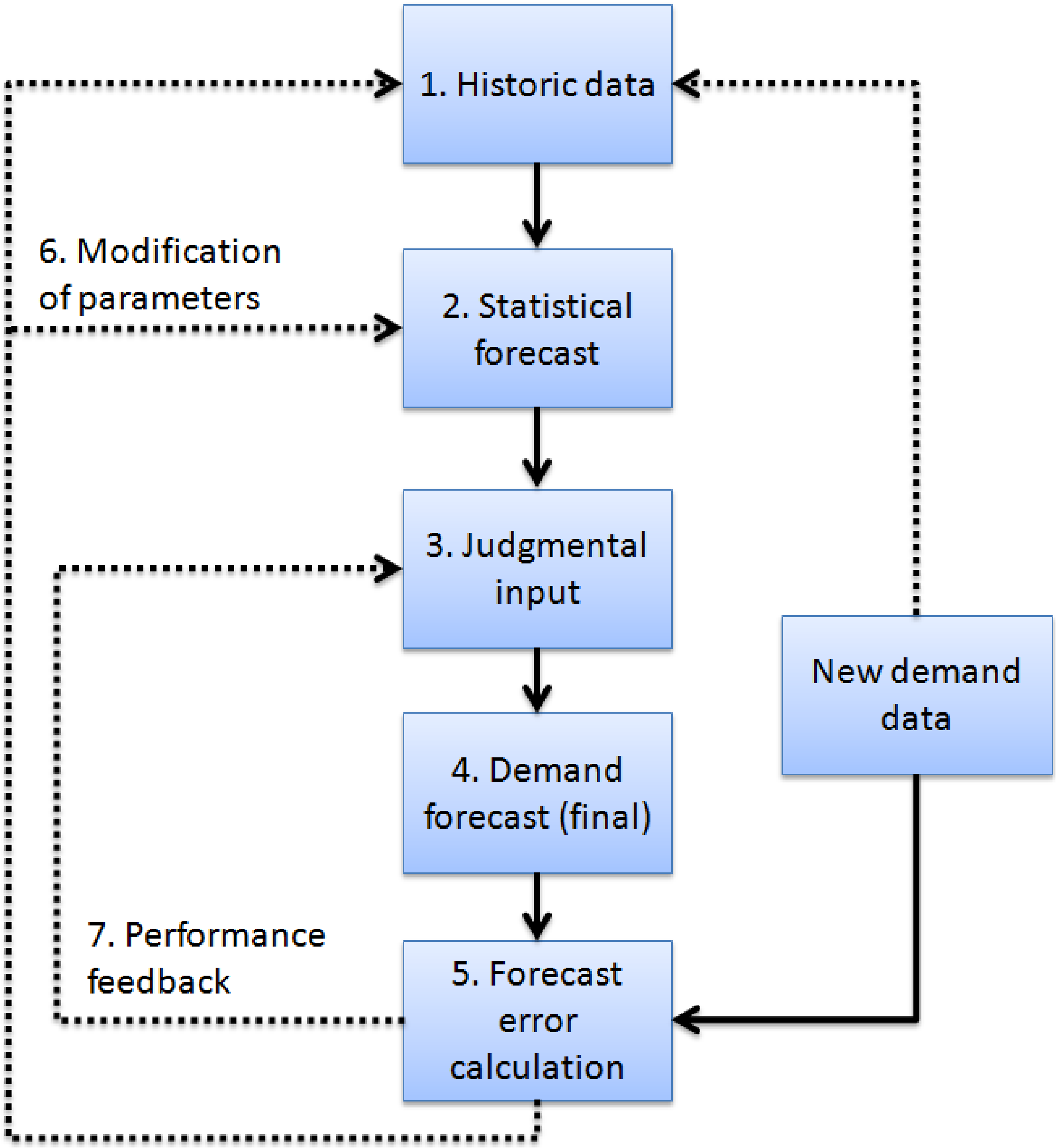


Figure 3.3. The demand forecasting process.

An important aspect that was not dealt with earlier is that even though the different steps of the demand forecasting process are numbered in an ascending order from one to seven, the process does not necessarily progress in a sequential order. Instead, some of the steps can be done simultaneously. In addition, the demand forecasting process is an iterative process, in which inclusion of new demand data, calculation of forecast errors and possible modification that follows restart the whole process. This is one of the reasons why the performance measurement part is of utmost importance. It is based on the performance measurement that the possible modifications are made, which ideally would improve the forecast accuracy.

There are also some exceptions to the process seen in the figure 3.3. In some situations, some of the steps of the process can be left out completely. If there is no historical data, for example, when a new product is introduced, the first two phases cannot be done. In that case the process starts by making only a judgmental forecast for that particular product. However, it is possible, as it was mentioned in subchapter 2.3.2, that the judgmental forecast is done by using historical data of another similar product. Another example would be that the demand is relatively predictable, which would mean that the statistical models are likely to be quite accurate, in which case, a judgmental adjustment is not necessarily needed.

4. CASE COMPANY ANALYSIS

This chapter consists of the analysis of the case company, its current forecasting practices and the use of the forecasting software in the demand forecasting process. The material used in this section includes the sales data of company's products and additional information gained in the meetings held in the case company. The sales data used in this section can be accessed in the forecasting software used by the case company. Appendix 1 shows how the forecast software (and more specifically the forecast client, which is the entity where forecasts are made and modified) looks like and what kind of data and attributes it includes.

Even though the data used in this and the next section is the sales data of different products available in the forecasting software, it is sometimes referred to as demand data instead of sales data. However, in this and following chapters, these mean the same thing unless stated otherwise. It should be stressed that even though these terms are used interchangeably in the following chapters, they do not necessarily always mean the same thing, as it was mentioned in chapter 2.1.

4.1. The case company

The case company of this study is Teknos Oy (referred to simply as Teknos), which is part of Teknos Group, one of Europe's leading suppliers of industrial coatings and a major participant in the retail and architectural paint markets. The Teknos Group was founded in 1948 in Finland under the name Teknos-Tehtaat Oy. Over the course of 60+ years, Teknos has grown from a local firm into a group of companies with an international presence. (Case company material [2])

Nowadays the company is called Teknos Group and it runs its subsidiaries in Finland (Teknos Oy), Sweden, Norway, Denmark, Germany, the UK (England, Scotland & Northern Ireland), Poland, Slovenia, Ukraine, Belarus, Russia and China. In addition to this, it has a network of representatives in about twenty other European countries. (Case company material [2])

Teknos Group is one of the biggest family firms in Finland, the ownership now belonging to the fourth generation of the founding family. Teknos Group employs a staff of around 1000 (in 2010 number of people employed was 993) people and had an annual turnover of over 200 million Euros in 2010. Table 4.1 shows the key figures of Teknos Group from the previous five years (2007-2011). (Case company material [2])

Table 4.1 Key figures of Teknos Group (adapted from case company material [2])

	2007	*2008*	*2009*	*2010*	*2011*
NET SALES (million EUR)	247	228	197	215	247
EBITD (million EUR)	27	25	22	28	31
Personnel	987	922	891	870	993

The case company in this thesis, Teknos Oy is the Finnish subsidiary of Teknos Group. Teknos Oy's portion of the net sales of the Group in 2011 was 128 million euros, which accounts for approximately 52 % of the entire Group's sales. Teknos Oy employs 500 people, which account for about a half of the personnel of Teknos Group. (Case company material [2])

4.2. Products and markets

As previously mentioned, Teknos operates in both industrial and consumer markets by selling its products to both industrial companies and end customers. Teknos' products can be categorized into three main segments:

1) Architectural coatings
2) Metal and mineral coatings, powder coatings
3) Industrial wood

The first segment, architectural coatings serves consumer markets, whereas other segments are meant for industrial markets.

In the architectural coatings (AC) segment Teknos' goal is to serve both professional and do-it-yourself painters. These products serve the purposes of end-customers and are therefore normally sold to them, for example via retail shops. Products include various indoor and outdoor paints, as well as some specialty products. (Case company material [2]) Based on the sales data of AC-products, which was available in the forecasting software of the case company, it can be said that they are subject to some level of seasonality (appendix 4). The peak of their demand has so far been during March and April. At this point it should be emphasized that not all products are subject to seasonality: there are hundreds of different products with different demand patterns. For example, indoor paints are not subject to same amount of seasonality as the outdoor paints. However, on average the demand curve of AC-segments has seasonal variation.

The second segment includes products for various industrial applications. It can be divided into three product groups which are general industry & heavy duty, powder coatings and road marking & floor coatings. General industry (GI) includes products for different types of machines, components and process equipment made from metal, plastic or composite materials, whereas heavy duty paints are intended for larger entities or constructions such as cranes, oil refineries, power plants and bridges. (Case company material [2]) Based on the sales data general industry is by far the largest sub-segment of industrial products with more than a thousand different products and the demand curve (appendix 4) of GI-products does not include seasonal or any other pattern but it includes a lot of random variation. This is in line with prior research presented in this thesis, according to which the demand for industrial products is usually more volatile than the demand for consumer products.

The second product group, powder coatings (PC), is a product group, in which Teknos is a market leader in the Nordic countries. These products include various domestic and household appliances such as steel furniture, garden furniture, bicycles and car parts. Powder coatings are environmentally friendly, which makes them a good alternative for paint shops that thrive on environmental friendliness. Nowadays the importance of powder coatings is rising and in order to meet the growing demand Teknos has built a totally new powder coating factory in Rajamäki, Finland in 2007. (Case company material [2]) The growing demand of powder coatings can also be seen in the sales data of PC-products, which shows an increasing trend (appendix X).

The third product group of the second segment – road marking and floor coatings (IM) – consists of a variety of special marking products to different kinds of customer applications. Road marking applications include markings used in highways and urban streets, courtyards as well as parking and restricted areas. Floor coatings comprise a wide range of products used for coating concrete floors or other surfaces. In addition to the aforementioned, Teknos also has products for certain niche markets such as coatings for grass pitches and anti-graffiti glazes. (Case company material [2]) IM-products are the smallest sub-segment of industrial products with only a handful of products and, as mentioned in subchapter 1.3, they are not analyzed in this study.

The third segment, industrial wood (IW) is comprised of paints and coatings for the wood and joinery industry. The range of products covers a wide variety of surface, both outdoor and indoor, and treatment requirements. Outdoor surfaces include for, example solid wood structures (beams, bridges), doors and windows, cladding or garden furniture, whereas the indoor surfaces consist of panels and moldings, furniture or internal doors and windows. In addition to these, there are also products for special damp spaces, such as solid wood structures in cold stores, swimming halls or ice rinks. (Case company material [2]) Like the powder coatings group, the sales data shows that there is also an increasing trend in the demand curve over the previous years (appendix 4).

4.3. Forecasting practices in the case company

This subchapter will include the forecasting practices which are applied in the case company at the moment. The aspects will be discussed step-by-step in consistency with the demand forecasting process, which was introduced in subchapter 3.3. The case company uses demand forecasting software, where the forecasts are made. This is why it should be emphasized that some of the procedures and aspects that are presented are dependent on this software and its applications. That is why, for example, only those statistical methods or error measures that are included in the software are introduced here.

Helping the later analysis products of each of the four product segments (AC, GI, PC and IW), were further divided based on the demand pattern of products. This helped limit the number of products per group. Another reason why this particular option was chosen was because of the importance of applying certain forecasting practices to products with specific historical demand. The reason was also discussed in the literature review of this study, where it was mentioned that recognizing the demand pattern is important for the correct choice of a statistical forecasting technique, which serves as a basis for the final demand forecast.

The division of the products based on their demand pattern is done automatically in the forecasting software by the software itself by recognizing the demand pattern of a product and classifying it into one of the ten demand groups of the software. The ten different groups of the forecasting software, to which products can belong, are:

1) Intermittent: products with intermittent (sporadic) demand
2) Level: products with constant (on average) demand, with random variations
3) Level/Season: in addition to the group two, these products have some seasonal variations
4) Level/Negative trend: products with a declining trend in their demand
5) Level/Negative trend/Season: products with a declining trend and seasonal variations
6) Level/Positive trend: products with an increasing trend in the demand
7) Level/Positive trend/Season: products with an increasing trend and seasonal variations
8) New Parts: new products with less than 24 months of data
9) Season: products with seasonal variation
10) Terminated: products with several consecutive months of zero demand

The reason why the division of products into demand groups is beneficial is because it divides the hundreds, or in some cases thousands of products, into smaller groups, which can then be analyzed more closely. The significance of each group is summarized in table 4.2, where the importance is estimated based on the groups' sales values in relation to the sales value of the product segment to which the group belongs.

Table 4.2. Importance of each demand group based on the sum of sales values of products of each group the group.

	AC	***GI***	***PC***	***IW***
Intermittent	2 %	16 %	9 %	12 %
Level	25 %	28 %	27 %	21 %
Level/Season	35 %	6 %	9 %	10 %
Level/Trend(-)	1 %	2 %	1 %	2 %
Level/Trend(-)/ Season	1 %	0 %	0 %	0 %
Level/Trend(+)	11 %	33 %	47 %	31 %
Level/Trend(+)/ Season	7 %	12 %	3 %	21 %
New Parts	15 %	2 %	2 %	3 %
Season	4 %	1 %	1 %	0 %
Value of sales of the group in relation to sales in the segment	More than 25 %	10-25 %	5-10 %	Less than 5 %

Table 4.2 is later used in chapter 5 and 6 as a reference when estimating the importance of the findings.

4.3.1. Preparation of demand data and statistical forecast

The first step of the process is the preparation of demand data, which includes data gathering, analysis and possible modifications or updating of the data. Gathering of data does not fall within the scope of this study, therefore practices on how it should be done will not be discussed in this thesis. However, it is worth mentioning that the demand data used, is the sales data of different products, which is automatically collected and uploaded into the forecasting software, and presented in appendix 1.

Even though the analysis and possible modification of demand data is the first step of the demand forecasting process, described in subchapters 3.2 and 3.3, it is not addressed here further. Instead, it is discussed in relation to the other steps of the process (e.g. choosing statistical models based on data, or analyzing data to find where judgmental adjustments are needed).

Statistical forecast is made automatically every month (meaning that the time period of a forecast is also one month) by the forecasting software. In the forecasting software it is possible to make forecasts for different product groups, based on different variables, or even entire product segments. The norm is, however, that the software will calculate and update its systematic forecasts for each individual product. There are twelve different statistical models, which can be used to make the statistical forecast. The models are:

1) **Manual:** Gives an average demand per period (month) from a fixed yearly demand that is decided manually.
2) **Naïve:** The forecast for the next future period is the same as the demand for the current period.
3) **Moving average:** The forecast for the future period is the average of the demands of previous *N* periods. As time passes, each new average is calculated by dropping the oldest observation and including the next one.
4) **Exponential smoothing (EWMA):** Applies an unequal set of weights to past data. Model calculates a weighted average of past observations using weights that decrease exponentially. For example, when the weighting/smoothing parameter is set equal to 0.20, the most recent period is weighted 20 %, the next recent period is weighted 16 % (= 0.20 * (1 - 0.20)), and the previous period is weighted 12.8 % and so on. This is done until the oldest period is reached.
5) **Exponential smoothing with trend:** Trend estimate is added to the original smoothing model by calculating the trend as a weighted average of previous demands and an earlier trend using a smoothing parameter. Note: the smoothing parameter to weigh trend is not (however it can be) the same as the smoothing parameter of the original model.
6) **Adaptive exponential smoothing (AEWMA):** Similar to exponential smoothing, difference being that this model allows the smoothing parameter to be modified, in a controlled manner, as changes in the pattern of data occur.
7) **Brown's smoothing with trend:** Similar to exponential smoothing, difference being that the original smoothing parameters are calculated differently.
8) **Regression/least squares:** This forecast model is a straight-line fitting of the demands of historical periods in accordance with the least squares fitting rule.
9) **Multiple regression:** Includes one value to be predicted (demand) but two or more explanation variables that can be used to calculate the prediction variable.
10) **Bayesian:** An average of four other forecast modes, which are all given a weight of ¼ in the final forecast. The models included are: moving average, adaptive exponential smoothing, least squares, and Brown's smoothing with trend.
11) **Best fit:** Runs different models in competition on the last known historical demand. The model with the best result is chosen as a forecast model for this part. Models are run with different parameters, so that the parameters are also optimized. The models included are: manual, moving average, exponential smoothing, exponential smoothing with trend, naive, adaptive exponential smoothing, least squares, Brown's smoothing with trend.
12) **Croston's intermittent:** A special model to be used on parts with intermittent demand (slow movers). It calculates the occurrence of demand and the size of demand separately.

(Case company material [3]) Mathematical formulas of each models are presented in the appendix 2.

The case company has chosen to divide the use of statistical forecasts into four different situations. These situations are:

1) Standard situation; historical data from at least the previous two years.
2) New product without a known seasonality pattern or historical data.
3) New product with a known seasonality pattern but no historical data.
4) Less than 2-year-old product with known seasonality and historical data.

In the case of a standard situation, the case company uses an automatic statistical forecast. In order to notice possible trends or seasonal variations, this requires at least two years of demand history. This means that in standard situations the forecast can be adjusted with an inclusion of a seasonal profile, which in this case is an automatic profile determined by the software based on the data from the previous two years (24 months). In the case of a standard situation, the Bayesian forecast model is being used. In the second case, a new product without a known seasonal pattern or historical data, the forecast is made with a combination of manual forecasts and the Bayesian method. Firstly, a manual forecast is given for approximately from two to six months. After that the Bayesian method is being utilized in the same way as in a standard situation. (Case company material [4] & [5])

In the case of new product with a known seasonal pattern but no historical data of its own, the situation is slightly different. In this case, the forecast model being used in the beginning is manual and after one year of demand data it will change to a moving average. At first, when there is no historical data, the user has to enter the cumulative demand manually for the forecasting horizon (usually one year) and the seasonal profile. After this, the software calculates the demand for each time period based on the given inputs. An example of this kind of situation would be the addition of a new product to a specific product family, where the other, older products have a known seasonality pattern. (Case company material [4] & [5])

The seasonal profile that can be included in the third case is a predetermined profile called FI-OUTDOOR. This is used only in the AC-products' segment, specifically for outdoor paints because they are often influenced by seasonal variations. In the case of new products of industrial segments (GI, PC and IW), there is no predetermined profile at all. Even though the inclusion of seasonality varies between new products of consumer and industrial segments, other forecasting procedures are the same for all both of the aforementioned segments. After one year the forecast model changes to a moving average. This is the fourth case, when a product is less than two years old with a known seasonal pattern and historical data. Afterwards when there is data from the two first years, the product follows the principles of the standard situation. (Case company material [4] & [5])

4.3.2. Judgmental input, consensus forecast and release of the forecast

As previously mentioned, the software will create and update the statistical forecast for each individual product every month. However, the statistical forecast can be modified by the users of the forecasting software (in this case the time period is also one month). This should be done if the people involved have, not only the information provided by the software, but also some potential hidden information (e.g. campaigns). It is worth emphasizing that if the forecasts are modified in the software, the new modified forecast is always the final forecast for that particular month. Thus, the steps of consensus forecast and release of the forecast belong basically to the step of judgmental input made in the case company. However, if there is no judgmental adjustment, i.e. forecasts are not modified then the original forecast calculated by a statistical model will remain in effect. (Case company material [4] & [5])

There are number of people in the case company who can manually modify the forecasts and the responsibility of making the modifications varies. Theoretically, everyone with access to the software and the proper rights of the software can manually change the forecast. However, there are obviously certain rules and guidelines on who should make the adjustments (judgmental input) and how they should be done. Usually people in certain departments, such as people of sales department, who have closer contact with some customers, are responsible for products of those customers. However, it is unclear what level of co-operation exists with the customers and how adjustments actually occur. If the changes are made, there are certain norms which should always be followed. (Case company material [5])

First and foremost, if adjustments are made, they should always be made only for a few, usually from one to three, future periods (months) at a time. The reason for this is that if the manual forecast is not updated regularly, it is likely that the old manual forecast (made for example to six months in the future) is less accurate than the systematical (statistical) forecast. A case where a forecast should be adjusted downwards compared to the statistical forecast is when the demand history has some sort of anomalies, such as exceptionally high demand because of a certain project. However, in this case in addition to the forecast the demand should also be adjusted so that the historical data would represent the actual “real demand” in the future periods. (Case company material [4])

If the forecast is adjusted upwards compared to the statistical forecast, the case company has made some rules which should be followed, namely: if the product only has a little amount of demand occurrences per period (e.g. less than 10 per month) the accuracy is likely to be low. In this case, the adjusted forecast should be at least 50 % larger than the systematic (statistical forecast) for the adjustment to be justified. If the product has a lot of occurrences per month, then the accuracy is usually higher and then even smaller adjustments (approximately 20-30 %) can be made. The norms were meant to be applied mostly in the cases where there is only one product, for which demand is forecasted. In addition to individual products forecasts can be manually adjusted for whole product groups as well. (Case company material [4] & [5])

However, when adjusting the forecast of an entire group, the products in that group have to have the same seasonal profile and the forecast model. If an adjustment is made for an entire product group, then the group's products have a manual forecast for the next 12 months and forecasts of the group's products are not systematically adjusted based on historical demand anymore. When the adjustment is made for entire product groups, it is divided by each product based on the ratios of historical demand of the products. Changes of entire product groups should only be made for such groups, where all the products have historical data from at least the previous 12 months. Products with less data should always be adjusted individually. (Case company material [4] & [5])

4.3.3. Measurement of the forecasting process

In addition to the calculation of statistical forecast for each product, the software calculates different kinds of forecast errors after each period, which serves as the basis for the performance measurement of the forecasting process. The calculated errors can be seen in two different parts of the display (appendix 1). In the first part the errors are seen automatically as a cumulative sum of each value. The values seen here are: mean error (ME), mean absolute error (MAE), Value of MAE (which is the inventory value multiplied by MAE), mean absolute percentage error (MAPE), percentage variation explained (PVE), tracking signal and Theil's U-statistic. In the second part, the control card shows the development of forecast accuracy. Here the possibilities for error measures are: mean absolute error (MAE), absolute error, mean absolute percentage error (MAPE), absolute percentage error, mean error (ME) and the simple error. (Case company material [3])

Some of the aforementioned methods were already presented in subchapter 2.5.1. Measures not mentioned in chapter 2.5.1 were the tracking signal and percent variation explained. Tracking signal is used when wanting to find out whether the forecast model includes some sort of bias, whereas percent variation explained shows whether the forecast is improving over time or not. The mathematical formulas of all of the accuracy measures can be seen in the appendix 2. Additionally, the difference between error measures and mean error measures (e.g. difference between absolute percentage error and mean absolute percentage error) is that the mean is always the average of the forecast error occurrences. (Case company material [3])

So far the case company has not created any guidelines on which accuracy measure or measures should be used in different situations. The only target the case company has, is an MAPE value of approximately 20 % or less, which is seen as acceptable accuracy. Additionally there has not been a specific focus on the concept of performance measurement: the whole has been on the level where the forecasting software automatically calculates the errors and they are merely accepted as the unavoidable disadvantage of the forecasting process. This has been explained with the abundance of products and their data and lack of time resources of the people involved in the process. (Case company material [5])

Because the performance measurement has not received much of a focus, steps 6 and 7 (figure 3.3) have also been neglected. Thus far, there has not been any change of parameters of statistical models or even forecast models themselves. Additionally the feedback that the people involved are provided with is, in most cases, only the information provided by the software, which includes the aforementioned error calculations and the control card. Afterwards, it is up to the people involved to find the relative data that could be beneficial in their estimation of the quality of the forecasts and the whole forecasting process. (Case company material [5]) Because of the external approach of this study, it cannot be said with utmost certainty that there are no other sources of performance feedback. However, based on the fact that the whole demand forecasting process of the case company is centered on the forecasting software, it can be suggested that it is also the main conduit for feedback.

4.4. Summary of the demand forecasting process of the case company

This chapter summarizes the analysis of the previous subchapters and additionally uses the literature review of chapters 2 and 3 as a benchmark to ascertain the problems of current forecasting practices. It must be remembered though, that prior studies and research focus usually on forecast practices applied in consumer markets, and the case company is not only engaged in consumer markets but also in industrial markets, within which some variation exists. In addition, there are some other shortcomings of previous studies, which are discussed in this chapter.

The characteristics of the case company also have an effect on the forecasting practices, and the suitability of prior research to this particular situation. For example, the heterogeneous customer base also means that there are number of sources of information available and not just the historical demand data, which is usually the focus in the prior studies. This is true especially in the industrial markets, but in this case also in consumer markets. For example, the demand for outdoor paints can be depended on weather conditions.

It should also be noted that because of a vast product mix of the case company, it is not always possible to incorporate all of the demand forecasting procedures to all of the different products. Therefore, some of the forecasting practices presented in the literature review are not always possible because they do not always distinguish the difference between an ideal and a real-life situation. The difference between those two is that in real-life resources are limited, and as it is in this case, there are number of different individual products in the mix which require at least partially different approaches. In addition, prior studies are often merely generalizations about the best practices, which means that they are not necessary applicable to all of the situations. Even though there are some shortcomings in the literature review, it is still applicable, when comparing the demand forecasting process of the case company (figure 4.1) to the one described in chapter 3.2 (figure 3.3).

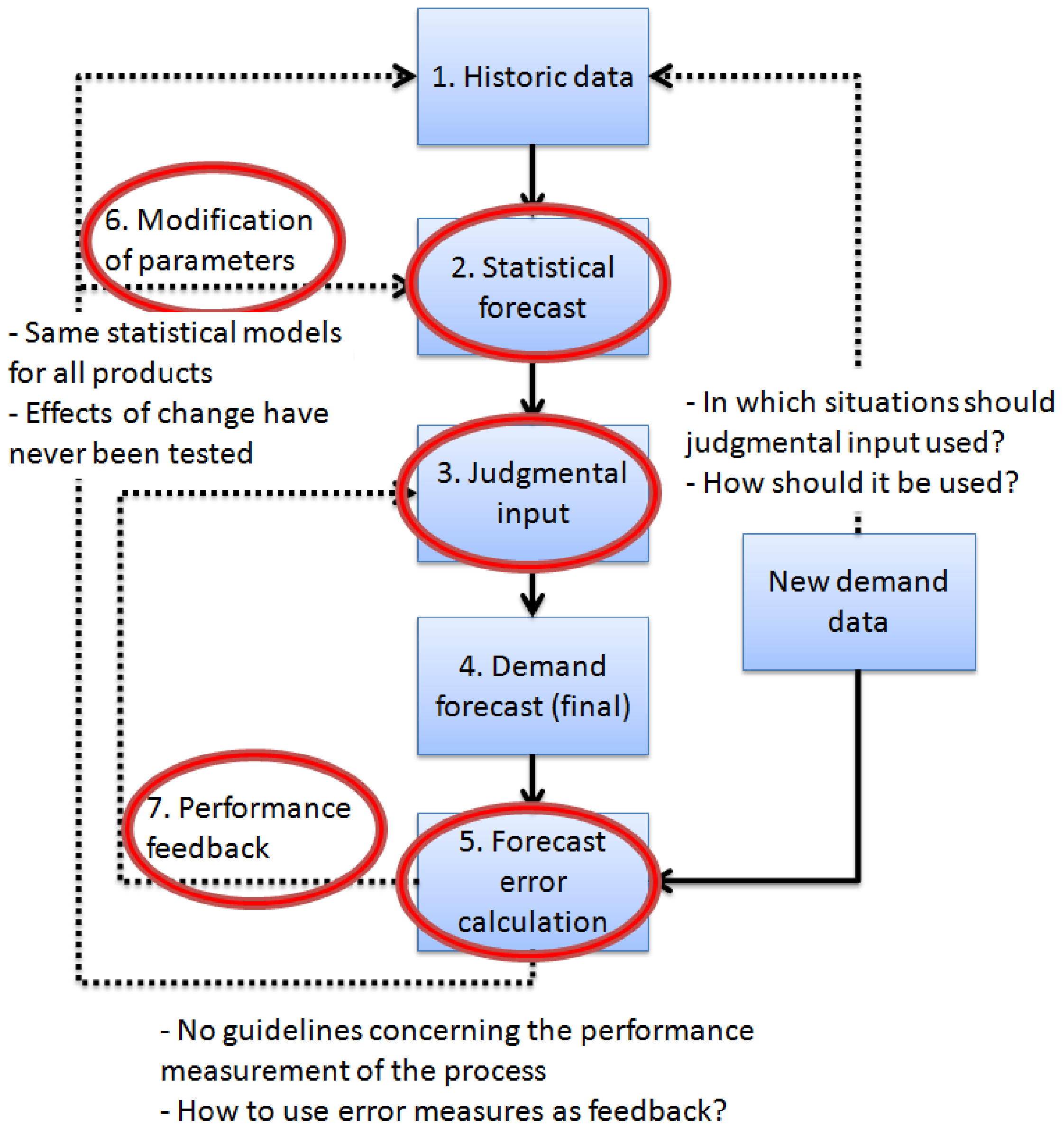


Figure 4.1. The demand forecasting process of the case company.

When comparing the actual situation of the demand forecasting process of the case company to the ideal one presented in subchapter 3.3, some general differences can be noticed. The different statistical models that have been used have always been the same, as well as their parameter combinations. There combinations are used by the forecasting software to calculate the systematical forecast and are used for all of the product segments. Even though there are twelve different models available, including specific models for certain kind of demand patterns, only two of the available models are used. These two models are used for all of the products regardless of their product segments or demand patterns.

Another problem of the case company is the incorporation of judgmental input to the forecasting process: the rules and previous guidelines relating to adjustments focused only on how they should be done in the forecasting software, not in which situations or for which products. However, this problem is also a shortcoming in some of the previous studies: even though they present different ways on how judgmental adjustment of forecasts can be done, they do not define properly, when it is needed and in when not. An additional problem related to judgmental input is that it is not always possible to do because, based on the guidelines of the case company, it should be done for individual or at least a small number of products, which obviously takes considerable amount of time because of the vast product mix.

The problem related to the performance measurement is the fact that even though the calculation of forecast errors is done automatically every period by the software, this information has not been properly used by the people involved in forecasting. The same goes for performance feedback. This problem was also mentioned in literature review, where it was stated that too often the errors are calculated but their impacts on performance are not measured in any way. One reason for this is the fact that the entire demand forecasting process of the company works around forecasting software, which means that the so-called feedback is in the raw data calculated by the software. This could also mean that, even though the implementation of the forecasting software helps in most of the forecasting procedures, it can also have a negative effect because it handles everything automatically, which can lead to the fact that the users of the software rely too much on it.

5. DEVELOPING AND TESTING OF ALTERNATIVE DEMAND FORECASTING PRINCIPLES

This chapter consists of alternative solutions to the current demand forecasting process. They are derived based on the problems in the demand forecasting process of the case company, which were discussed in subchapter 4.4, and they are divided into three main sections: alternatives of statistical forecast (steps 2 and 6), incorporation of judgmental input (step 3) and performance measurement of the process (steps 5, 6 and 7). The purpose of this chapter is to present the findings; they are discussed in more detail in chapter 6, where recommended actions are also suggested based on the findings presented here and also on the previous studies presented in the literature review of this study.

The first section of this chapter includes the testing of different statistical models, which is done in order to find out if the statistical model for each product group is the best possible one out of the alternatives provided by the forecasting software. An additional purpose of the testing of statistical models is to find out whether the future demand of products can be properly forecasted using any of the available statistical measures. The second section focuses on the incorporation of judgmental input. However, because of the external approach of this study it is difficult to define some of the problems of judgmental input: for example accuracy comparisons are not possible because it is not clear in which situations the adjustment of the statistical forecast has been made. That is why this study only attempts to identify the situations where the judgmental input would be applicable.

The third section this chapter includes the problems related to the performance measurement part of the forecast. Some of the problems related to this are also difficult to assess because of the external approach of this study. However, knowing that the accuracy calculation, which is the basis of performance measurement is done automatically by the forecasting software and that there is a general lack in using this data, this study focuses on improving the latter one. Therefore, the focus will be on the use of this data and how it can also function as a performance feedback to the people involved in the demand forecasting process.

The data used in this chapter is the sales data of different products, which is available in the forecasting software. Therefore, it must also be emphasized that because of the external approach of this study and lack of other applicable data, some of the assumptions made in this chapter are based solely on the sales data and its interpretation. It must also be remembered that some of the solutions that are presented in this chapter are based on the possibilities of the forecasting software. This means that there are some other possible alternatives as well. However, because they are not available in the software, they will not be analyzed here. To help the analysis of this chapter the product grouping of table 4.2 presented in subchapter 4.3 is used throughout this subchapter.

5.1. Testing the accuracy of different statistical models

As it was mentioned in subchapter 4.3.1 the statistical model that has been used so far, for most of the products is the Bayesian model. Because the case company has been content with the accuracy provided by the Bayesian model, the effects on the accuracy when using other models has never been tested (Case company material [5]). Bearing in mind the literature review of subchapter 3.2, which states that in the demand forecasting process the statistical methods should be tested frequently, for example every few months. In addition to the Bayesian model, there are cases (new products with seasonality) where a moving average is used, the use which was also tested.

In this study the comparison of different statistical models is done by creating a test sample of products and then testing the accuracy of different statistical models for the created sample. It should be highlighted that the statistical models tested included only the ones available in the forecasting software. This is because those are the models currently available for the case company and as it was mentioned in subchapter 1.2, that at some points the options and best choices are limited by the options available in the forecasting software. The results of the comparison are presented in this subchapter, whereas the implications and recommended courses of action based on the results are discussed in detail in chapter 6.

5.1.1. Creation of the test sample

Because of the vast number of products, a test sample of products was chosen to assess the accuracy of different methods. First, the products were divided based on their segment into four groups: AC, GI, PC and IW. After which the products were divided, automatically by the software, into the ten demand categories introduced in subchapter 4.3. The reason why this particular categorization was chosen was because of the importance of historical demand as a basis for the choice between forecasting methods, especially in the case of statistical methods. Another reason for the aforementioned categorization was that it helped limit the hundreds or thousands of products to tens of products per specific group.

After the categorization, a number of one to four products out of each demand group were chosen to be a part of the test sample. The products were chosen randomly. The number of products varied based on the demand group's characteristics and size. For example, groups with seasonal demand had products with different kind of seasonality, and one product representing each type of seasonality was chosen. Another example would be groups with level demand, in which the random variation (standard deviation) was either relatively small or mediocre, approximately 100 % or less of the average demand, or relatively large, more than 100 % of the average demand. In those cases two products, one with relatively small and one with relatively large random variation, were chosen. The characteristics of each product can be seen in the appendix 5. The number of test products out of each demand group is summarized in table 5.1.

Table 5.1. The test sample.

	AC	***GI***	***PC***	***IW***	***Sum***
Intermittent	3 / 139	3 / 490	3 / 174	3 / 76	12 / 879
Level	2 / 194	2 / 246	2 / 144	2 / 40	8 / 624
Level/Season	3 / 273	2 / 92	1 / 26	2 / 14	8 / 405
Level/Trend(-)	2 / 13	2 / 27	2 / 9	1 / 4	7 / 53
Level/Trend(-) /Season	2 / 16	0 / 3	1 / 1	0 / 0	3 / 20
Level/Trend(+)	2 / 43	2 / 121	2 / 91	2 / 36	8 / 291
Level/Trend(+) /Season	2 / 18	2 / 25	1 / 8	2 / 16	7 / 67
New Parts	4 / 146	2 / 73	2 / 18	2 / 21	10 / 258
Season	3 / 139	2 / 50	2 / 17	1 / 4	8 / 210
Sum	23 / 983	18 / 1134	16 / 491	15 / 213	71 / 2807

As it can be seen from table 5.1 there are two cases when no products were chosen at all. In the one case there were no products in the demand group and in the other case the demand pattern of the products was the same as in the other groups (e.g. level/season) but the software had categorized the products to the other group. In case only one product was chosen to represent the group, it was because the group was relatively small and all of the products had a very similar demand pattern. Additionally the group Terminated products was not included in this particular test because the products of that group were not forecasted after their categorization to be terminated. In the group New Parts of AC-products four products were chosen because it was of the special nature of the group; in that particular group the procedures are a bit different than in the other groups (chapter 4.4.1). In the New Parts of AC-products the standard statistical model is moving average and the product has a known seasonality pattern.

5.1.2. Conducting the test

Statistical forecasting models that were chosen varied between the demand groups because of the different demand patterns. In every case four different statistical models were selected: original model, The Best fit -option, and two additional models. Original model was in most cases the Bayesian, with the exception of New Parts of the AC-products. The Best fit -option meant that the forecasting software chose the model itself based on the demand pattern. The two additional models varied between the demand groups depending on the demand pattern. For example, products with intermittent demand were tested with the models designed specifically for them, whereas products with trend were tested with models that took the trend factor into account. The statistical models that were chosen for different demand groups can be seen in the appendixes 6-10, where their accuracies are also presented.

In addition to the demand pattern of a test product, an initial evaluation of accuracy was done to check that the model chosen could actually be competitive with the original one. This was done with the help of the forecasting software: when the change of methods is tested the accuracy of forecast can also be seen in the forecast graph (shadow in the graph, appendix 1). Therefore, if it was sometimes absolutely clear that some methods decreased the accuracy and because of that they were not included in the test at all. In some cases however, this was not that clear. Examples of this would be that the forecast accuracy changed a lot over time or that its pattern was altogether very different compared to the original one. Even though the initial estimation of accuracy could not be as thoroughly as the proper calculation of accuracies of different models, it had to be done in order to limit the possible options in the analysis.

The reason why the initial assessment of accuracy was needed in the forecasting software was because the comparison of accuracy of different models could not be done in the software itself. Instead the data had to be copied first, after which the accuracies were able to be calculated. Therefore, the initial limitation of some statistical models helped to limit the amount of data that had to be copied from the forecasting software. This was also the reason why different parameter combinations for some models were not tested; testing different parameters for each model would have taken considerable amount of time and therefore, only the standard parameters were used. Main reason why the accuracy calculations could not be done in the software itself was the fact that the software does not provide the tools necessary for the accuracy measurement that was used in this particular test. The accuracy measures and criterions which were used in this test are presented later on in this chapter.

In addition to comparing different statistical models, another test was made only for products that contained seasonal variation. These included the test products from the groups: Level/season, Negative trend/season, Positive trend/season and Season. The test was made because of the existence of the option of including the Automatic Seasonality -profile and FI-OUTDOOR -profile for New Parts of the AC-segment. So far the option of Automatic Seasonality has been used for all products that the software categorizes as part of the seasonality groups and FI-OUTDOOR has been used for New Parts of the AC-segment. Therefore, in this study it was tested whether or not:

1) The inclusion of Automatic Seasonality improves the forecast accuracy in the groups, where seasonal variation is present.
2) The inclusion of FI-OUTDOOR improves the forecast accuracy in the product group New Parts of the AC-segment.

After all the data was gathered the criterions for accuracy measurement were chosen. As stated in chapter 2.5.1, there are plenty of error measures that can be used to measure accuracy of a forecast. Out of the different accuracy measures, mean average percentage error (MAPE) and median average percentage error (MdAPE) were chosen. Even though MAPE is sometimes criticized (chapter 2.5.1), it was used here because of its simplicity and understandability as a benchmarking method. Another reason for its use was the fact, that MAPE is one of the error measures that the forecasting software itself uses, albeit its use here was slightly different than in the forecasting software. MdAPE was used to supplement the use of MAPE because MAPE is subject to be somewhat misleading in the presence of one or more significantly large values. However, the primary criterion was MAPE and in the case where MAPE's of two or more different models were the same or very close (within 1-2 %) to each other, MdAPE was used a secondary criterion.

The only group, where the uses of MAPE or MdAPE were not used as the primary and secondary criterion was the demand group Intermittent. This was because the presence of a lot of periods of zero-demands, which does not suit to the use of MAPE or MdAPE for that matter. Therefore they had to be replaced with some other appropriate measure. In the case of intermittent demand, mean absolute deviation (MAD) was used to compare the accuracies. Because MAD is an absolute measure its use had to be modified in order to make comparisons of products with different level of demand. Therefore, the ratio of average MAD and average demand was used as the criterion to evaluate the accuracy: the smaller the ratio, the better the forecast. In addition to the calculation of forecast accuracy the Coefficient of Variation (COV) was also calculated for the test products in order to evaluate whether or not the products are forecastable or not forecastable. However, these values are not presented in this chapter but in chapter 6, where the results are discussed further.

Apart from New Parts and in some cases products with intermittent demand, the historical data was always available from the previous three years (36 periods). It should be emphasized however, that the forecasts of those previous periods, which were used as the data in this test, were forecasts that the statistical model would have used as the forecasts of previous periods while making the forecast for the future period (expl. forecast, appendix 1). In other words, the forecasts that were used as a data were the statistical models' forecasts of the past, which are more accurate than the forecasts of the future. Therefore, the accuracies that were achieved in the tests were slightly better than the accuracies that could be achieved in reality.

However, the forecasts are still comparable because different models are based on different formulas, and therefore the forecasts are calculated differently. Additionally, because of the limited time resources and the length of the period being one month, this method was able to provide more data (36 periods) than for example following the future forecasts for a few months. In this case the effects of seasonal variations, trend or other recurring demand characteristics could also be taken into account. After all the data of previous forecasts of different models were copied, errors and percentage errors were calculated for each period of each product. After this the so-called final accuracies, which were MAPE, MdAPE and Average MAD/Average Demand in the case of intermittent demand, as well as the COV value were calculated. The results of the accuracies of different models are discussed in the next subchapter. However the COV values of test products are only presented in chapter 6, where the results and their significance and possible modifications they lead to are discussed further.

5.1.3. The results of the test

This chapter includes the results of the comparison of different statistical forecasting models. The results are discussed in more detail in the subchapters. However, the original accuracies of each demand group is presented in the table 5.2 below.

Table 5.2. Summary of the accuracies of the Bayesian model (exc. New Parts, where Moving Average is used) adapted from appendices 6-10.

	AC		***GI***		***PC***		***IW***	
	MAPE	MdAPE	MAPE	MdAPE	MAPE	MdAPE	MAPE	MdAPE
Intermittent								
Level	51 %	31 %	75 %	42 %	127 %	53 %	270 %	71 %
Level/Season	60 %	36 %	127 %	51 %	76 %	44 %	68 %	30 %
Level/Trend(-)	46 %	32 %	66 %	37 %	155 %	45 %	52 %	36 %
Level/Trend(-) /Season	59 %	35 %			50 %	39 %		
Level/Trend(+)	62 %	30 %	65 %	31 %	99 %	48 %	298 %	45 %
Level/Trend(+) /Season	34 %	24 %	23 %	18 %	82 %	28 %	68 %	33 %
New Parts	329 %	90 %	41 %	43 %	47 %	42 %	47 %	48 %
Season	76 %	48 %	70 %	55 %	211 %	44 %	36 %	34 %

In table 5.2 the accuracies are presented as average accuracies – using mean and median average percentage error – of demand groups with 1-4 test products using the original model. The only exception was the group Intermittent, where the accuracies were estimated using other measures. The colour-codes of the groups are the same as in table 4.2 (subchapter 4.3). As can be seen from table 5.12, there is room for improvement in the accuracies. As mentioned in subchapter 4.3 that the categorization of the groups' importance was done in order to guide the focus of further analysis in to the most important groups. Therefore, in the emphasis will be especially on the demand groups Level, Level/Season, Level/Positive Trend and Level/Positive Trend/Season.

The detailed results will be discussed in five different sections. First, Level and Level/Season are discussed. After that products with positive trend, including the groups Level/Positive Trend and Level/Positive Trend/Season, are discussed. After that the demand group New Parts is presented, which is followed by products with intermittent demand. Lastly, the three least important demand groups, Level/Negative Trend, Level/Negative Trend/Season and Season are presented.

Products with level demand

The products of group Level can be divided into different groups based on their random variation. Because this group contained hundreds of products (in total of 624, table 5.1.) it is safe to say that there are also tens, if not hundreds of different sort of patterns for random variation. However, in this test the products were divided crudely into two categories: one with relatively low or mediocre random variation (less than 100 % of the average demand) and one with relatively strong random variation (more than 100 % of the average). By making this distinction, it was possible to analyze the possible effect that random variation has on the forecasting accuracy.

The results of this particular test support hypothesis that was mentioned in the subchapter 2.3.1: time series methods are effective when the random variation is relatively low. Looking at the detailed results, it is clear that the accuracy of forecasts of products with mediocre random variation is significantly better in comparison to the accuracy of products with strong random variation. The average MAPE of products of group Level with mediocre random variation is 40 %, whereas it is 167 % in the case of products with strong random variation. When calculating the averages of MdAPEs, the results are better: with mediocre random variation the average is 19 %, whereas with the products of strong random variation it is 62 %.

The differences between MAPE and MdaPE can be explained, by the likely existence of some individual large errors that distort the MAPEs, especially in the segments PC and IW. That’s why the median (MdAPE) gives a better image of how the forecast accuracy actually is, or at least has the potential to be. However, even though MAPE might be distorted, it gives a good indication of what kind of effect individual large errors can have on the overall accuracy. At this point, the reason for such errors is not clear, which means that it cannot be said how they should be dealt with. This is discussed more in subchapter 5.2.

The results also indicate that the existence of seasonal variations does not have a strong effect on the forecasting accuracy, provided that the automatic seasonal profile is used. As it can be seen from the results (appendices 6-10) the inclusion of the automatic profile increases the accuracy of the forecasts drastically. With the automatic seasonal profile the accuracies within the group Level/Season are similar to the accuracies of the Level group. The averages of MAPEs and MdAPEs in the group Level/Season are 63 % and 42 %. However, unlike in the Level group, there were not any large deviations in the results. The results are summarized in table 5.3.

Table 5.3. Summary of the accuracies of most accurate models for the demand groups Level and Level/Season (adapted from appendices 6-10).

AC	***Characteristic***	***Best forecasting method***	***Best possible accuracy***	
			MAPE	MdAPE
Level 1	Mediocre random variation	Bayesian	31	15
Level 2	Strong random variation	AEWMA	62	39
Level/Season 1 (aut. season)	Relatively strong seasonal variation	Bayesian/AEWMA	74	34
Level/Season 2 (aut. season)	Mediocre seasonal variation	EWMA	38	26
Level/Season 3 (aut. season)	Individual peaks	Bayesian/AEWMA	59	36
GI				
Level 1	Strong random variation	EWMA	72	55
Level 2	Mediocre random variation	Bayesian	31	24
Level/Season 1 (aut. season)	Relatively strong seasonal variation	EWMA	87	93
Level/Season 2 (aut. season)	Mediocre seasonal variation	AEWMA	65	27
PC				
Level 1	Strong random variation	Least squares	160	74
Level 2	Mediocre random variation	Moving average	55	24
Level/Season 1 (aut. season)	Individual peaks	Best fit (Moving average)	71	49
IW				
Level 1	Strong random variation	AEWMA	372	80
Level 2	Mediocre random variation	Best fit (Naive)	46	13
Level/Season 1 (aut. season)	Relatively strong seasonal variation	Best fit (Moving average)	70	49
Level/Season 2 (aut. season)	Relatively strong seasonal variation	Bayesian	36	21

In the group Level/Season the random variation was much harder to see because of the difficulty in separating it from the seasonal variation, which for some products lasted the entire 12 months. However, the thing that might have had a slight effect on the results was the type of seasonal variation, which in some cases was very strong (during the variation demand is five or even ten times higher than normally) and in some cases only mediocre. The averages of MAPEs and MdAPEs for the strong seasonal demand were 67 % and 50 %, whereas in the cases of mediocre demand and individual peaks the averages were 58 % and 35 %. This just proves that when the variation increases, the forecast accuracy usually decreases at the same time. What is remarkable and definitely worthy of emphasis is that the Best-fit option was able to provide the most accurate result for only one product out of eight in the case of level demand and for two out of eight in the case of level and seasonality, even though theoretically it should choose the best option based on the historical data.

When it comes to choosing the best forecasting model, there are several options from which to choose in the Level group. The existence of so many different models is partially explained by the fact that in the software there are a lot more models which have been designed for level demand instead of trend. As can be seen, there are six different models which have proved to be the most accurate for the eight test products of the Level group. However, only two of the models, AEWMA and EWMA, can be analyzed further, because they were the only models, which occur amongst the best models in both of the test products of most of the segments. Other models such as, moving average and least squares cannot be used because they are present only once, which means that in the other case their accuracy is worse than the ones in appendices 6-10, which makes their analysis difficult. Table 5.4 summarizes the further analysis of the three models.

Table 5.4. Comparison of the accuracies of Bayesian model, AEWMA and EWMA.

	Bayesian (original)		***AEWMA***		***EWMA***	
	average of MAPEs	average of MdAPEs	average of MAPEs	average of MdAPEs	average of MAPEs	average of MdAPEs
AC	51	31	***45***	***32***	50	33
GI	75	42			***53***	***42***
PC	***127***	***53***	144	53	138	60
IW	270	76	***220***	***50***		

Based on the comparison, the Bayesian model is only most accurate for PC-products. In other cases either the AEWMA (for AC- and IW-products) or the EWMA (for GI-segment) are more accurate. However, it should be pointed out that if the comparison is made based on random variation the results are contradictory: if the random variation is mediocre or small, the Bayesian model is the most accurate model (average of MAPEs 44 % and average of MdAPEs 20 %), whereas in the case of strong random variation

AEWMA is the most accurate model (average of MAPEs 194 % and average of MdAPEs 65 %). This should be remembered when thinking whether the original model is changed or not. For example in the case of AC-segment, where the random variation is most likely, at least according to the previous studies, not as strong as in other (industrial) segments the change of model is not necessarily needed, whereas in the PC-segment the change could be the better option, even though the Bayesian model is the more accurate in the comparison seen in table 5.4.

The situation is a bit different when seasonality is included: in the group Level/Season, there are four different statistical models which have provided the best results. The models are: the Bayesian model, AEWMA, EWMA and moving average. To have a better insight into the potential accuracies of these models, the averages of accuracies of forecasts made by each model for the test products were calculated. The results are seen in table 5.5 below. As was the case previously, some models were not compared because there was only data from one of the test products.

Table 5.5. Comparison of Bayesian model, AEWMA, EWMA and the Moving Average.

	Bayesian (original)		***AEWMA***		***EWMA***		***Moving Average***	
	average of MAPEs	average of MdAPEs	average of MAPEs	average of MdAPEs	average of MAPEs	average of MdAPEs	average of MAPEs	average of MdAPEs
AC	60	36	***58***	***33***			62	37
GI	127	51			***83***	***61***		
PC	76	44	82	44	78	43	***71***	***49***
IW	68	30	83	29			***55***	***41***

The comparison shows that different models are optimal depending on the segment. Based on the comparison, the most accurate model for the group Level/Season of AC-products would be AEWMA, whereas in GI-products, EWMA is more accurate than the original Bayesian model. Moving average is most accurate method in the segments PC and IW, albeit the average of MdAPE of the original Bayesian model is much lower compared to moving average but since MAPE was used as the primary criterion and its average is lower, moving average is said to be the more accurate model.

Products with positive trend

Products with a positive trend include two demand groups: Positive Trend and Positive Trend/Season. As it can be seen from table 5.5, it is clear that the effect of random variation that was apparent in the previous results is also apparent in the results of product group Positive Trend. The average MAPE of products with strong random variation is approximately 214 %, whereas the average MAPE of products with mediocre random variation is approximately 44 %. Albeit, these figures are distorted by the one relative large value, without it the average of MAPEs is 110 %. In addition to the average MAPEs the average MdAPEs support the aforementioned effect of random variation. Average MdAPE of products with strong random variation is 51 %, whereas products with mediocre random variation have an average MdAPE of 28 %. Results of the tests for these two product groups are presented in table 5.6.

Table 5.6. Summary of the accuracies of most accurate models for demand groups Positive Trend and Positive Trend/Season.

	Characteristic	***Best forecasting method***	***Best forecasting method***	
AC			MAPE	MdAPE
Trend(+) 1	Strong random variation	Least squares	85	55
Trend(+) 2	Mediocre random variation	Best fit (EWMA trend)	19	15
Trend(+)/Season 1 (aut. season)	Regular seasonal variation	EWMA trend/Brown	18	13
Trend(+)/Season 2 (aut. season)	Irregular seasonal variation	Bayesian	49	33
GI				
Trend(+) 1	Strong random variation	Bayesian	98	39
Trend(+) 2	Mediocre random variation	Bayesian/Best fit (EWMA trend)	32	23
Trend(+)/Season 1 (aut. season)	Regular seasonal variation	Bayesian	25	18
Trend(+)/Season 2 (aut. season)	Regular seasonal variation	Best fit (Brown)/EWMA trend	20	13
PC				
Trend(+) 1	Strong random variation	Bayesian	147	54
Trend(+) 2	Mediocre random variation	Best fit (Naive)	50	39
Trend(+)/Season 1 (aut. season)	Regular seasonal variation	Best fit (EWMA trend)	82	28
IW				
Trend(+) 1	Mediocre random variation	Bayesian/Best fit (EWMA trend)	69	35
Trend(+) 2	Strong random variation	Bayesian	527	54
Trend(+)/Season 1 (aut. season)	Regular seasonal variation	EWMA trend/Brown	40	16
Trend(+)/Season 2 (aut. season)	Regular seasonal variation	Bayesian	94	40

Similar to previous section, the seasonal variation does not seem to have a strong effect on the accuracies, when the automatic seasonal profile is being used for products that belong to the group Positive Trend/Season. Like before, the inclusion of the seasonal profile increases the forecasting accuracy drastically on all accounts, even though, on one occasion the seasonal variation is somewhat irregular. The average MAPE of products with seasonal variation is 47 %, whereas the average MdAPE of those products is 23 %. So in this case, the forecasts were more accurate when seasonality was involved. In cases where seasonality was involved, it was not possible to divide the products based on strong and mediocre random variation because it was impossible to distinguish random variation from seasonal variation.

In the case of a positive trend, the Best-fit option provided the best results on six occasions (out of fifteen), which is better than before. Out of those occasions four were for the demand group "positive trend" (out of possible eight) and two for the demand group "positive trend and seasonality" (out of possible seven). However, its performance is still somewhat substandard, for example compared to the original Bayesian model; average of MAPEs and MdAPES with the Best-fit option for the group "positive demand" was 143 % and 38 %, whereas the corresponding values were 131 % and 38 % with the Bayesian model.

There are three statistical models which provide the most accurate results, except on two occasions. This makes sense because there are only a few statistical models that were chosen here were the ones that take trend into account. The three statistical models are: the Bayesian model, EWMA with level and trend and Brown's smoothing with trend. In the cases where seasonality is involved, the Bayesian model and EWMA with level and trend are most accurate three times, whereas Brown's smoothing and trend is the most accurate four times. The three methods were examined closer in a comparison of their accuracy for all the test products, which is summarized in table 5.7.

Table 5.7. Comparison of Bayesian model, EWMA with level and trend and Brown's smoothing with trend for products with positive trend and seasonality.

	Bayesian (original)		***EWMA with level and trend***		***Brown's smoothing with trend***	
	average of MAPEs	average of MdAPEs	average of MAPEs	average of MdAPEs	average of MAPEs	average of MdAPEs
AC	***34***	***24***			42	22
GI	***23***	***18***	***24***	***17***	***24***	***16***
PC	82	28	***76***	***43***	85	23
IW	***68***	***33***	97	28	96	27

The comparison indicates that the original Bayesian model is most accurate in two cases (AC and IW). In the case of AC, the EWMA with level and trend was not included since it only had values for one product (in the case of the second product the values were not as good as with the four models that are presented in appendices 6-10. However, bearing that in mind, it can be said that it would not have been more accurate than the original model in that particular case. In the case of GI products, all of the three methods are equally accurate, with their values very similar to each other's. The only case where the choice is a bit unclear is the case of PC segment, where EWMA with level and trend is a bit more accurate measured with the average of MAPE's but much worse measured by the average of MdAPEs. However, since MAPE is the primary criterion, EWMA level and trend is said to be the most accurate model.

In the case where there only positive trend occurs, it is apparent in the appendices 6-10, which models are the most accurate for which segment. This is why there will be no similar comparison as seen in table 5.7 for example. It is clear that the Bayesian model is the best in three out of the four cases, the only exception being the AC-segment where all the three models are equally good based on MAPEs. However, EWMA level and trend (20 %) and Brown's smoothing with trend (22 %) have a better average of MdAPEs compared to the Bayesian model (30 %). Out of those, EWMA level and trend is slightly more accurate and that's why it is chosen.

New Parts

New Parts are products with demand data from less than two years, in most cases approximately one year. This demand group is very important for consumer products (AC-products), however for industrial products, which includes the other three segments the importance is quite low. The forecasting practices are a bit different depending on whether the product belongs to AC-segment or not. In the AC-segment all the products have the predetermined FI-OUTDOOR -seasonality profile and the first forecast is made manually for one year. After that the software divides the yearly forecast for each month based on the seasonal profile. After one year of demand data, the forecasts are made with a moving average model.

In other segments, there is no seasonality profile and the original forecast is manual for two to six months. After a few months of demand data, the software starts to use Bayesian model as a statistical model for the forecasts. In addition to the forecasting accuracies the effect of the predetermined seasonal profile on accuracy was tested. Because the AC-segment is the only one where the profile is being used, four test products were chosen to provide a better insight on the effect. Because of the limited amount of data (approximately year or less), there is not much that can be said about the demand patterns of New Parts.

As can be seen from the appendix 12, the inclusion of the predetermined seasonal profile increases the accuracy for the products of the AC-segment. Even though the accuracy is improved with the seasonality profile, the overall accuracy of the AC-segment is very low, whereas the accuracy in the other segments is close to the accuracies of other groups in this test. The accuracy of the test products of AC-segment is on average 195 % based on MAPE and 62 % based on MdAPE. In the other segments the accuracies are very similar, the averages of MAPEs and MdAPEs of other three segments being 40 % and 36 %. Results of the test are summarized can be seen in table 5.8.

Table 5.8. Summary of the accuracies of the most accurate models for product group New Parts.

	Characteristic	***Best forecasting method***	***Best possible accuracy***	
AC			MAPE	MdAPE
New part 1	with FI-outdoor	Best fit (Brown)	362	35
New part 2	with FI-outdoor	AEWMA	68	73
New part 3	with FI-outdoor	AEWMA	96	62
New part 4	with FI-outdoor	Best fit (Brown)	254	79
GI				
New part 1		EWMA	52	61
New part 2		Best fit (Brown)	20	13
PC				
New part 1		AEWMA	45	34
New part 2		Best fit (Brown)	43	31
IW				
New part 1		Bayesian	39	28
New part 2		Best fit (AEWMA)	43	49

Table 5.8 shows that there are two competing models for the forecasts of new parts, Brown's smoothing with trend and adaptive exponential smoothing (AEWMA). Additionally the Best-fit option worked much better compared to the previous results. The statistical models were compared further in order to find the most accurate model for each segment. The original models were taken into comparison so it would also be apparent, how much the accuracy could be increased. Table 5.9 shows the results of said comparison.

Table 5.9. Comparison of original model, Brown's smoothing with trend and AEWMA.

	Original (Moving average or Bayesian)		***Brown's smoothing with trend***		***AEWMA***	
	average of MAPEs	average of MdAPEs	average of MAPEs	average of MdAPEs	average of MAPEs	average of MdAPEs
AC						
GI	41	43	***41***	***35***		
PC	47	42	***46***	***35***	46	39
IW	***47***	***48***			71	52

As it can be seen in table 5.9, Brown's smoothing with trend is the best model for GI and PC segments. In the case of IW segment, the original model provided the best results. The good performance of Brown's model, which is originally designed for products with trend, can be explained by the fact that most of the industrial segments products have a declining trend. However, because there is less than two years of data the software does not categorize them into the group Negative Trend. The comparison could not be made for AC-segment because the data did not include accuracies of the models for all both of the products of each segment, which means that it is unclear which model would be the most accurate for the whole AC-segment.

Intermittent demand

The Intermittent group includes products with sporadic (intermittent) demand, which means that there are some months were demand is zero. Contrary to the New Parts group, intermittent is significant to industrial segments but not for the consumer segment. Because of the occurrence of zero months, the same sort of accuracy measures could not be used in this test, which meant that the accuracy was estimated based on the ratio between average deviation and average demand. Table 5.10 summarizes the findings.

Table 5.10. Summary of the accuracies of the most accurate models for the group Intermittent.

AC	***Characteristic***	***Best forecasting method***	***Best accuracy (av. Deviation/ av. Demand)***
Intermittent 1	Mediocre occurence of zero months	Bayesian	1,10
Intermittent 2	High occurence of zero months	Best fit (EWMA)	1,03
Intermittent 3	Low occurence of zero months	Bayesian/Croston's	0,94
GI			
Intermittent 1	Low occurence of zero months	Bayesian	0,51
Intermittent 2	Mediocre occurence of zero months	Bayesian	1,25
Intermittent 3	High occurence of zero months	Best fit (EWMA)	1,33
PC			
Intermittent 1	Mediocre occurence of zero months	Bayesian/Croston's	1,00
Intermittent 2	Low occurence of zero months	Best fit (EWMA trend)	0,66
Intermittent 3	High occurence of zero months	Naive/Croston's	1,17
IW			
Intermittent 1	Low occurence of zero months	Bayesian	0,81
Intermittent 2	Mediocre occurence of zero months	Best fit (Moving average)/Croston's	0,80
Intermittent 3	High occurence of zero months	Croston's	1,05

Results indicate that there is a connection between the occurrence of months with zero demand and the accuracy of the forecast: the fewer months with zero demand, the better the forecasting accuracy. The average ratio for products with low occurrence of zero months is 0.73, whereas for mediocre it is 1.04 and for high 1.15. There are no great differences between the segments: forecasts are altogether somewhat inaccurate, with the average ratio varying between 0.90 and 1.00.

The two models which provided the most accurate results were the Bayesian model and Croston's intermittent, which is designed for intermittent demand. What is perhaps somewhat strange is that Croston's model is five times the most accurate method and it is designed for intermittent demand but for some reason the software did not choose it as the Best-fit option. Altogether, the Best-fit option was able to provide the best result only four times out of twelve.

The overall performance of original Bayesian model and Croston's intermittent was compared further in order to find out which model is the best for each segment. The results of the aforementioned comparison are summarized in table 5.11.

Table 5.11. Comparison of the original Bayesian model and Croston's intermittent.

	Bayesian (original)	***Croston's intermittent***
	average of deviations/demand	average of deviations/demand
AC	1,10	***1,06***
GI	***1,09***	1,15
PC	***0,97***	***0,97***
IW	1,03	***0,91***

As it can be gathered, there are some differences between the accuracies of these models. Croston's intermittent was more accurate in the case of AC- and IW-segments, whereas the Bayesian model was more accurate for GI-segment. In the case of PC-segment, the accuracies were the same. When doing the same comparison based on the occurrence of zero months the accuracies are quite similar. Only difference is that the Croston's intermittent is more accurate (1.22 vs. 1.35) for products with high occurrence of zero months. This is good to bear in mind when considering whether to change the model or not.

Products with negative trend and the group Season

All three demand groups that have not been assessed thus far are products with little of importance (table 4.2), especially groups containing products with negative trend. Because of their relatively small importance they are discussed briefly here. Unlike in the case of products with a positive trend, the effect of random variation is not as clear when it comes to products with negative trend. The accuracies are on average alike: average of MAPEs and MdAPEs for products mediocre random variation was 74 % and 40 %, whereas their corresponding values were 76 % and 36 % for products with strong random variation. In the group Negative Trend/Season there were only three products in this test. Their averages of MAPEs and MdAPEs were 55 % and 35 %. In all of the three cases the inclusion of automatic profile improved the accuracy significantly. Table 5.12 summarizes the findings for products with negative trend.

Table 5.12. Summary of the accuracies of most accurate models for products with negative trend.

	Characteristic	***Best forecasting method***	***Best forecasting method***	***Best forecasting method***
AC			MAPE	MdAPE
Level/Trend(-) 1	Strong random variation	Bayesian	59	35
Level/Trend(-) 2	Mediocre random variation	Bayesian/Brown/ EWMA trend	32	29
Trend(-)/Season 1 (aut. season)	Individual peaks	Bayesian/Best fit (EWMA trend)	75	44
Trend(-)/Season 2 (aut. season)	Regular seasonal variation	Bayesian	43	26
GI				
Level/Trend(-) 1	Strong random variation	Best fit (EWMA trend)	57	38
Level/Trend(-) 2	Mediocre random variation	EWMA trend	61	43
PC				
Level/Trend(-) 1	Mediocre random variation	Least squares	155	47
Level/Trend(-) 2	Strong random variation	Brown	113	35
Trend(-)/Season 1 (aut. season)	Regular seasonal variation	AEWMA	45	36
IW				
Level/Trend(-) 1	Mediocre random variation	Bayesian/Best fit (EWMA trend)	49	39

Because of the small number of test products included in it and its insignificance, the group Negative Trend/Season will not be discussed further in this thesis. For products with a negative trend, a further comparison of three forecasting models that proved to be the most accurate ones was made. These models were the same as in the case of products with positive trend, which makes sense because they are products designed for products with a trend. The models that were compared further are the Bayesian model, EWMA with level and Trend and Brown's smoothing with trend. The results of the comparison are presented in table 5.13.

Table 5.13. Comparison of the Bayesian model, EWMA with level and trend and Brown's smoothing with trend.

	Bayesian (original)		***EWMA with level and trend***		***Brown's smoothing with trend***	
	average of MAPEs	average of MdAPES	average of MAPEs	average of MdAPES	average of MAPEs	average of MdAPES
AC	***46***	***32***	46	34	48	34
GI	66	37	***59***	***41***	65	45
PC	155	45	166	42	*150*	*39*
IW	52	36	***49***	***39***	52	43

Based on the comparison the original Bayesian model was most accurate for AC-segment, whereas EWMA with level and trend was most accurate model for GI-segment and Brown's smoothing was most accurate for PC-segment. In the IW-segment, both the original Bayesian and EWMA with level and trend work equally well.

The last demand group, Season, includes products with a seasonal pattern. The difference to the group Level/Season is that if the seasonal variation does not occur, there is very often no demand at all, which means that in this group there are also products with intermittent demand. This group is a bit more important for the case company than the other two but all-in-all its significance is still relatively small. The results of the test for the group Season are presented in the table 5.14.

Table 5.14. Summary of the accuracies of most accurate models for product group Season.

	Characteristic	***Best forecasting method***	***Accuracy***	
AC			MAPE	MdAPE
Season 1 (aut. Season)	Mediocre variation	EWMA/AEWMA	38	35
Season 2 (aut. Season)	Strong variation	EWMA	103	54
Season 3 (aut. Season)	Individual peaks	EWMA/AEWMA	120	52
GI				
Season 1 (aut. Season)	Strong variation	Best fit (EWMA level and trend)	60	44
Season 2 (aut. Season)	Individual peaks	Moving average	64	45
PC				
Season 1	Individual peaks	EWMA	206	83
Season 2 (aut. Season)	Strong variation	Brown	47	29
IW				
Season 1	Individual peaks	Bayesian	36	34

As it can be seen above, there is significant variation in the accuracies regardless of the pattern of the seasonality. An example is the individual peaks, where there is a huge deviation in the accuracies. What is remarkable as well is that in two cases out of seven the inclusion of the automatic seasonal profile did not improve the accuracy, but made it worse. All in all the accuracy that was achieved in this test was: an 84 % average of MAPEs and a 47 % average of MdAPEs. When choosing the best statistical model the most accurate models can be seen in the appendices 6-10. For the AC-segment the most accurate model is clearly EWMA, whereas for GI-segment the most accurate model is moving average, and for PC- and IW-segments the Bayesian model.

5.1.4. Summary of the comparison of different statistical models

The purpose of the test of statistical models was to find the models which best suit specific demand groups in each segment, and consequently produce the most accurate forecasts for those groups. A summary of the findings discussed earlier can be seen in table 5.15, which presents the accuracies of the most accurate statistical models for each product segment.

Table 5.15. Summary of the accuracies of most accurate statistical models for each segment.

	AC		***GI***		***PC***		***IW***	
	MAPE	MdAPE	MAPE	MdAPE	MAPE	MdAPE	MAPE	MdAPE
Intermittent								
Level	45 %	32 %	53 %	42 %	127 %	53 %	220 %	49 %
Level/Season	58 %	30 %	65 %	27 %	71 %	49 %	55 %	41 %
Level/Trend(-)	46 %	32 %	59 %	41 %	150 %	39 %	49 %	39 %
Level/Trend(-) /Season	59 %	35 %			50 %	39 %		
Level/Trend(+)	63 %	20 %	65 %	31 %	99 %	48 %	298 %	45 %
Level/Trend(+) /Season	18 %	13 %	23 %	18 %	76 %	43 %	68 %	33 %
New Parts	329 %	90 %	41 %	35 %	46 %	35 %	47 %	48 %
Season	71 %	44 %	64 %	50 %	211 %	44 %	36 %	34 %

In addition to the summary of the accuracies, the effect of the changing of the models was calculated by comparing the original accuracies with the accuracies of the best statistical models. Only the effect of group Intermittent was not taken into account because of the different accuracy measure used. The overall effects are summarized in table 5.16.

Table 5.16. Effect of the changing of the model on the accuracy.

	Original models	***Most accurate models***
Average of MAPEs	93 %	89 %
Average of MdAPEs	41 %	39 %
Median of MAPEs	67 %	51 %
Median of MdAPEs	41 %	39 %

As table 5.16 indicates, the accuracy can be improved with the changes of the models. The summary of the most accurate models is presented in table 5.17.

Table 5.17. Most accurate statistical models.

	AC	***GI***	***PC***	***IW***
Intermittent	Croston's	BAYESIAN	BAYESIAN	Croston's
Level	AEWMA	EWMA	BAYESIAN	AEWMA
Level/Season	EWMA	AEWMA	MA	MA
Level/Trend(-)	BAYESIAN	EWMA trend	Brown's	EWMA trend
Level/Trend(-) /Season	BAYESIAN		BAYESIAN	
Level/Trend(+)	EWMA trend	BAYESIAN	BAYESIAN	BAYESIAN
Level/Trend(+) /Season	EWMA trend	BAYESIAN	EWMA trend	BAYESIAN
New Parts	AEWMA	Brown's	Brown's	BAYESIAN
Season	EWMA	Moving average	AEWMA	BAYESIAN

Another aspect which was tested was the inclusion of the automatic seasonal setting and the company's own FI-OUTDOOR seasonal profile for the New Parts of the AC-segment. When included, both proved to be beneficial for the forecast accuracy (appendix 12).

Further discussion of the results of the comparison of different statistical models and some recommended actions are discussed more in chapter 6, where the results are also reflected on prior studies and the research problem of this particular study.

5.2. Incorporation of judgmental input to the forecasts

Unlike the accuracy of different statistical models, the effect of the inclusion of judgmental input was not tested in this study. One of the reasons for this was the outsider's perspective that this study had to the case company, which led to the lack of clarity on how the judgmental adjustment of forecast is actually done and in which cases. Additionally, because the forecasting software does not categorize adjusted forecasts, but always shows only the final forecast, the possible adjusted forecasts could not be found with the demand data, which was the main source of data in the analysis made in this study.

Instead of testing accuracy, the aim of this section in this study is the identification of situations, where judgmental input should occur. The reason why this is important is because there are no general guidelines in the company about in which situations judgmental adjustment of statistical forecasts should occur. The identification of such situations is done by analyzing some of the largest errors occurring in the previous months and trying to find causes for such errors. This subchapter includes only the presentation of the findings; the findings themselves are discussed in further detail in chapter 6, where some recommendations for courses of action are given as well.

Because of the abundance of data, only a number of products from each product segment could be analyzed. Each of the four product segment was analyzed separately. To help the analysis, the grouping based on the demand pattern (table 4.2) was used in this section as well. After the products were divided into demand groups, a calculation was made to find out which of the groups included the largest errors by cost. The measure of cost that was used Value of MAD. As it was mentioned in the subchapter 2.5.2, there are different kinds of costs forecast errors, and it can be often difficult to ascertain what the actual cost of a certain forecast error is.

The value of MAD is a fairly simple measure, where the absolute value of an error is multiplied by the products inventory value. The problem of it is that in the case of a positive error, demand being higher than the forecast, the costs do not arise from inventory holding costs, but from corrective measures such as overtime or extra capacity. Nevertheless, in the absence of other cost related measures in the forecasting software, which limits the inclusion of alternatives, Value of MAD is used in the case of positive errors as well. Even though it may not be precise, it can give an idea about the possible costs that inaccurate forecasts will produce. As it was previously mentioned, the importance of each demand group based on the costs of errors, measured by Value of MAD, was calculated. This calculation is analogical to the one presented in subchapter 4.3 where the importance of product groups was based on the sales values (table 4.2). The result of the calculation can be seen in table 5.18.

Table 5.18. Importance of the demand groups based on the cost of errors measured by Value of MAD.

	AC	***GI***	***PC***	***IW***
Intermittent	5 %	24 %	16 %	23 %
Level	21 %	26 %	19 %	21 %
Level/Season	29 %	5 %	9 %	11 %
Level/Trend(-)	1 %	2 %	1 %	2 %
Level/Trend(-) /Season	1 %	0 %	0 %	0 %
Level/Trend(+)	7 %	23 %	40 %	23 %
Level/Trend(+) /Season	4 %	7 %	2 %	12 %
New Parts	26 %	12 %	2 %	7 %
Season	7 %	2 %	2 %	1 %
Cost of errors of group in relation to whole cost of errors in the segment	More than 25 %	10-25 %	5-10%	Less than 5 %

When comparing the results of the importance of the groups to table 4.2, it is apparent that the two tables are very similar. This is logical since the groups with more products also have more errors. Additionally, more expensive products (sales value) also have more expensive inventory value, which partially determines the overall cost here. Table 5.18 can also be compared to table 5.15 seen in the previous section: doing this it can be seen that the groups, where the accuracies of statistical models are the worst, the cost of errors are also the highest.

After the calculation of the costs of each group, the three most significant groups of each segment were taken into a closer analysis. Bearing in mind the similarity of some of the groups (e.g. the random variation in some groups that affects the results), the analysis of only three groups can be said to be sufficient enough in this study. This is especially the case when looking at the percentage of the overall cost that they constitute (AC: 75 %, GI: 73 %, PC: 75 % and IW: 67 %). Furthermore, the three largest groups include the ones in which the accuracies of different statistical models were the lowest ones. Depending on the size (number of products) of the group, roughly ten to twenty products, with the largest Value of MAD were closely inspected in order to find out, if there was some recurrence or pattern in the errors and if unusually large individual errors existed. The reasons for errors are summarized in table 5.19.

Table 5.19. The reasons for recurring or unusually large errors.

	Demand patterns with the most significant errors		
	1.	2.	3.
AC	**Level/Season**	**New Parts**	**Level**
	Large errors in specific months, e.g. September	Constantly too large forecasts (year 2011)	Large errors in specific months, e.g. September
GI	**Level**	**Level/Positive trend**	**Intermittent**
	Very strong random variation, demand peaks and lows	Strong random variation. Errors have been increasing during previous 12 months	Large inventory values -> even a small error causes a large MAE value
IW	**Level/Positive trend**	**Intermittent**	**Level**
	Strong random variation around the increasing trend	Irregularity of the months with zero demand	Strong random variation
PC	**Level/Positive trend**	**Level**	**Intermittent**
	Strong random variation. Months with an unusually large demand (though partially regular)	Strong random variation	Irregularity of the months with zero demand. Large inventory values -> even a small error causes a large MAE value

The results show that there were indeed some recurring reasons why large errors (based on Value of MAE) existed. As can be seen from the table 5.19, the causes of errors vary between segments that serve consumer markets and segments that serve industrial markets. AC-products is the consumer market segment and it can be seen that the reasons for most significant errors are quite different than in the other segments. Other three segments bear a lot of resemblance to each other, and what they have in common is that they all serve industrial markets.

The most difficult to handle and the ever-present cause for forecast errors is uncertainty or random variation. The existence of random variation was one of the main causes for large errors in the case company as well. Its occurrence is aligned with the literature review, which states that random variation is more present in the forecasting of industrial products. And as the table 5.19 shows, three of the four products segments where strong random variation existed, and was the overall main cause for errors, are all product segments with industrial markets.

Another major reason for inaccurate forecasts in the industrial markets' segments has been the sporadic demand of some products. The accuracies achieved in the testing of statistical models were not particularly high and it is no surprise here that sporadic demand is also the reason for some large errors as well. When the demand is sporadic, the problem of random variation is usually linked with the random occurrence of the months without demand at all, even though the demand of those months when it occurs can be quite predictable. The situation was partially like this for the case company of this study. In addition to the random occurrence of months without demand, the high inventory values were another reason why the Value of MAD of errors was so high.

In the segment of AC-products, which serve consumer markets the causes of errors were different than the ones industrial markets: random variation does not play such strong a role here as it does in the industrial segment, which is aligning with the literature review. For the most important demand group, as well as for the third important group, the main reason was the occurrence of unusually large individual errors that have occurred during specific time (usually one or two months) of the year. Most of these unusually large errors of groups Level and Level/Season have occurred in September and October. What was happening, was that the demand for (some, not all) AC-products has been uncharacteristically high in September and after that relatively low in October.

The reason for this has usually been that there has been knowledge of a price increase for the retailers, which has caused the retailers to acquire these products more than they normally would require in September, since it is the last month to buy it at the previously lower price. And because they have been buying more than they normally would in September, the retailers have had to buy less than usual in October. (Case company material [5]) The effect of a peak on error and costs created by the error can be illustrated with the help of an example, presented in table 5.20 and figure 5.1.

Table 5.20. Example of the effect of peaks to costs.

Time span	2 years
Peak	One: Sept 2010
Demand of the peak/Average demand excluding peak	7,60
MAD, peaks	1018,21
MAD, incl. peaks	118,00
MAD, excl. peaks	83,26
Inventory value (€/unit)	15,54
Value of MAE, peak (€)	15822,98
Value of MAE, incl. peaks (€)	1833,72
Value of MAE, excl. peaks (€)	1293,86
Difference (excl./incl.)	71 %

Even though the save achieved in this example was almost 30 %, it must be remembered that there are different kinds of peaks so the effect on costs is different depending on the relative size and number of occurrences of the peaks. The overall effect of the peak of the example on MAD can be seen in the figure 5.1.

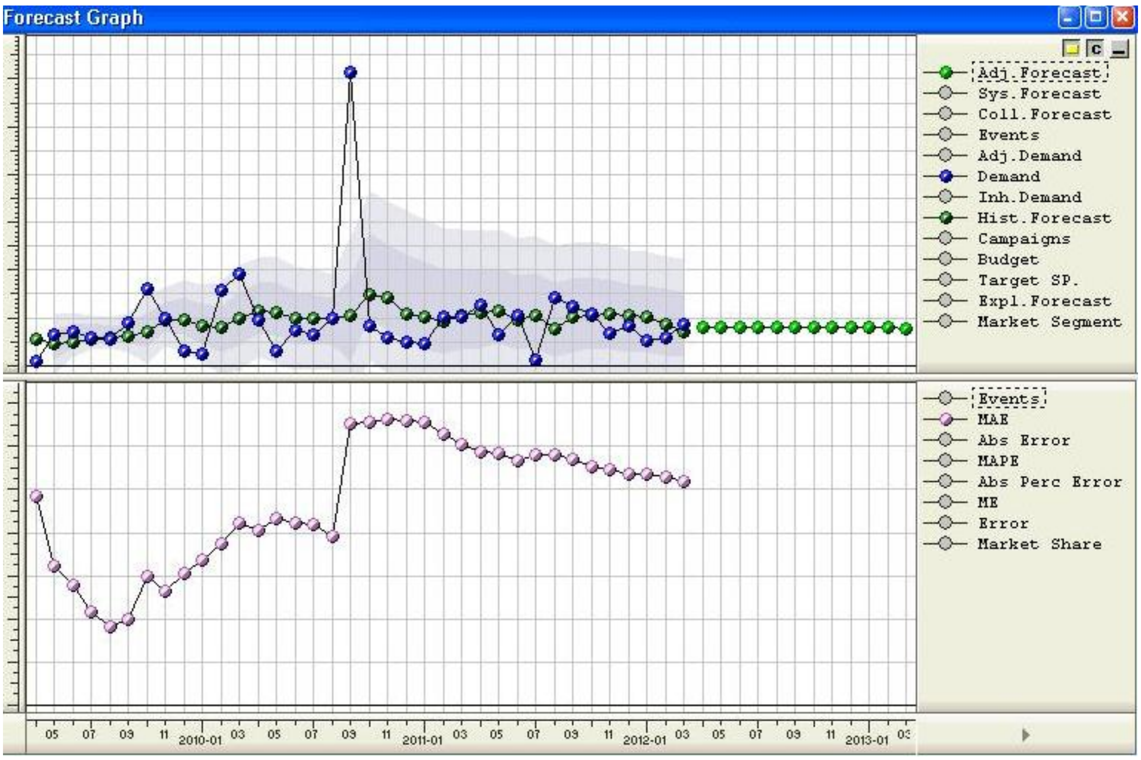


Figure 5.1. The effect of the peak on mean absolute deviation (adapted from the Forecasting Software).

Another product group in the AC-segment, in which some large errors have occurred, was the product group New Parts, which includes new products with no previous sales data and forecasts for these are made manually by giving the estimate of the demand for the whole year. After which the software divides it into months based on seasonal profile or other parameters that can be given. The regular cause for errors in this group has been that the forecasts are frequently too high.

There is another typical pattern that has been repeating in the group New Parts. Because all of the products of the group have the same seasonal index, there have been cases where the forecasts made for the beginning of the year, which is the peak according to the seasonal setting, have been clear overestimates about the demand. This has then affected to the forecasts of the few periods after, which have been much smaller than the demand. In short, the demand of all of the products has not changed exactly according to the seasonal index.

5.3. Performance measurement of the demand forecasting process

This subchapter focuses on the performance measurement part of the demand forecasting process. This includes error calculation and analysis, performance feedback and possible modification of parameters. The importance of the performance measurement of any process was emphasized in the literature review, because it helps to understand the quality of the process and additionally the development, the fact whether the results are improving over time or not, of the process.

As mentioned earlier, the output of the process is the final forecast and its quality can be measured by the forecast error, which is the difference between the final forecast and the actual demand. The forecasting software utilized by the case company has plenty of error measures that can be used to evaluate the progress of the forecasts. The problem in the case company thus far has been that even though there are different error measures that can be seen in the display of the forecasting software, it has been unclear how these measures could be used in order to increase the quality of process, i.e. improve the accuracy of the forecasts

Therefore, the purpose of this section is to go through the different measures offered by the forecasting software (appendix 1) and to provide examples on how the different measures can be used in the estimation of the quality of the forecasting process. Later on in the discussion subchapter, the aspects of the performance measurement are discussed further, with the inclusion of the previous findings presented in the literature review as well as other studies. Additionally, some recommendations are given in chapter 6 about how to handle the performance measurement in the future.

5.3.1. The average errors

As mentioned previously, the forecasting software includes several error measures that can be used to evaluate the accuracy of the forecast. There are average errors that can be used, for example, in comparison of different products or to organize products in an ascending order based on certain error measure. An example of organizing the products was the calculation of Value of MAD, presented in the previous chapter, which was helpful in finding the largest errors by cost. However, the same sort procedure can be done for products based on the percentage error or most of the other error measures.

The aforementioned procedure is very helpful when wanting to find out the average quality of the forecasts for certain products based on each value. Additionally the average values can be used to compare certain products or product groups. In this case the measure used in comparison would be the mean average percentage error. An important thing to remember is that the focus should not be only on one average value, because the values might provide information that completes the other.

For example, the mean average percentage error is a good way to evaluate the quality of the forecast model but when looking at the gravity of some errors it does not provide the user with all the needed details. To provide an understanding of the gravity of the forecast errors, the absolute error should be taken into account. It can be for example MAD, which shows the average deviation of the forecast (the average absolute error).

Again, an easy way to find products with largest absolute errors is to organize them as it was done in the previous subchapter. The best value to do this is MAD or Value of MAD, depending on if the largest errors, or by value largest errors are searched. In addition to finding the largest errors, the forecasting software can also be used to find the products of which average error deteriorates the most from the actual demand. By doing this (organizing based on ME) the cases where the forecasts are constantly higher or lower than the actual demand can be found.

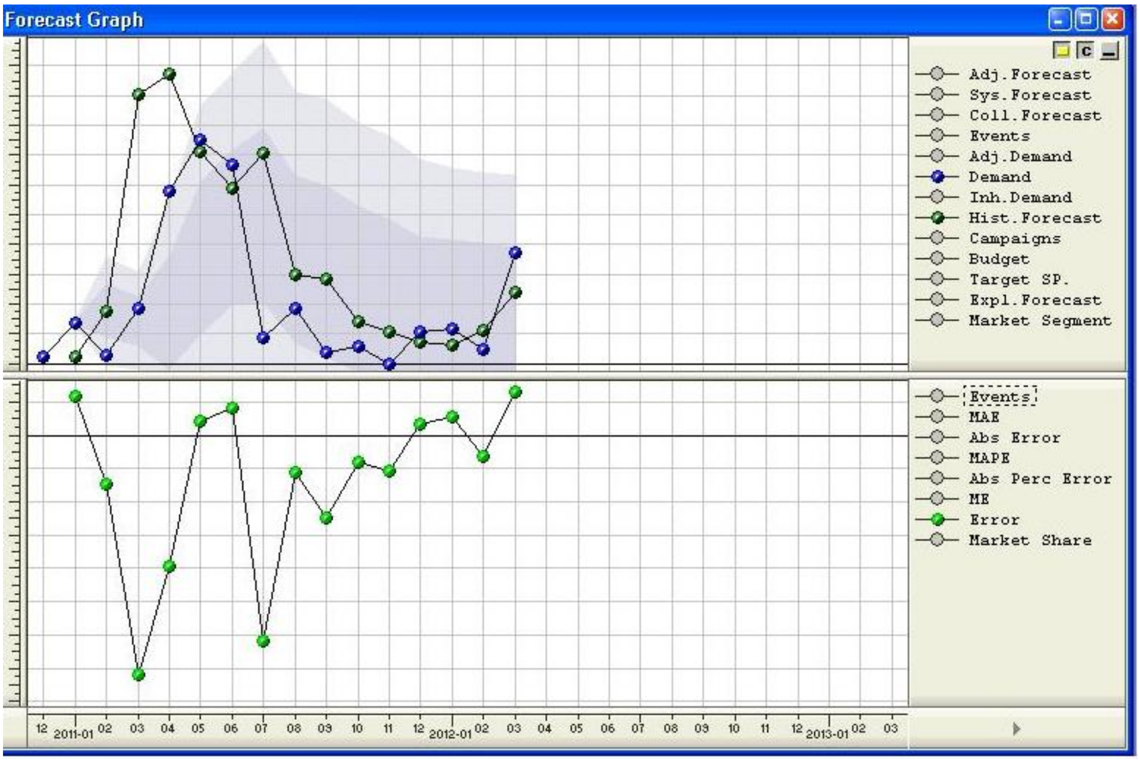
In addition to the aforementioned error measures, there are also three measures which provide information about the quality of the forecasting model. The quality can also be determined by the error measures above or based on the average percentage error. However, a quick way to find the cases, where the statistical model does not work is to categorize products based on the Theil’s U value. In the cases where the value of Theil’s U is larger than 1, the model that is used is inferior to the naïve model.

5.3.2. The use of control card

Another aspect of the error measurement is the use of a control card that can also be done in the forecasting software. However, in this case the errors are shown as individual occurrences in the control card, which shows the development of the forecast accuracy. Based on that, it can be, for example, estimated if the accuracy is getting better or worse over time. The use of control card was briefly discussed in the theory section of this study. While the theory section introduced briefly some of the reasons why the control card should be used, this chapter focuses on explaining the different examples it cannot be used for.

The examples that are presented here have been chosen based on typical error situation mentioned in subchapter 5.3 of this study. The specific control card presented here is the one available in the forecasting software. However, one of the most important things to remember when using the control card is that only one error measure should be chosen at once, because of the different scales and measures. For example, if percentage errors of tens and absolute errors of hundreds are shown simultaneously in the control card the scale is obviously determined by the larger errors, which means that it can be difficult to see the changes in the smaller (percentage) error values. This can deteriorate the overall image of the forecast and its errors.

The first example presents the situation where the forecast is regularly larger than the actual demand (AC-segment, New Parts). In this situation, most of the error values are below the zero line in the lower graph. The same can be seen in the upper graph, which shows the forecasted and the actual amounts, whereas the lower graph shows only errors. Figure 5.2 depicts the first situation.



Picture 5.2. Forecast is regularly larger than demand (adapted from the Forecasting Software).

The different error measures on the right column are the ones that can be seen in the control card and in the forecast graph as well. The first situation can best be seen when using error and mean error (also abs. percentage error). If the mean error was used in this situation it would be used the mean error (ME) line would be constantly below the zero, which implies that the forecast is uncontrollable (subchapter 3.2.3). In general the use of either of these measures, error or mean error, can help to find a possible bias in the forecasts.

The second situation, seen already in figure 5.1, shows how the control card can be used to find demand peaks or lows with the help of MAD. In this case, MAD is decreasing or relatively constant but after a demand peak or low, it increases drastically in one period and then starts to decline again or stay at a new constant relatively constant level. As figure 5.1 also shows, peaks or lows are easily detectable from the forecast graph as well and control card is not always needed to find them. However, a strong increase in MAD is a consequence of a relatively strong peak or low: if the peak or low is only strong by absolute value, the change in MAD is not always so apparent, as figure 5.3 shows.

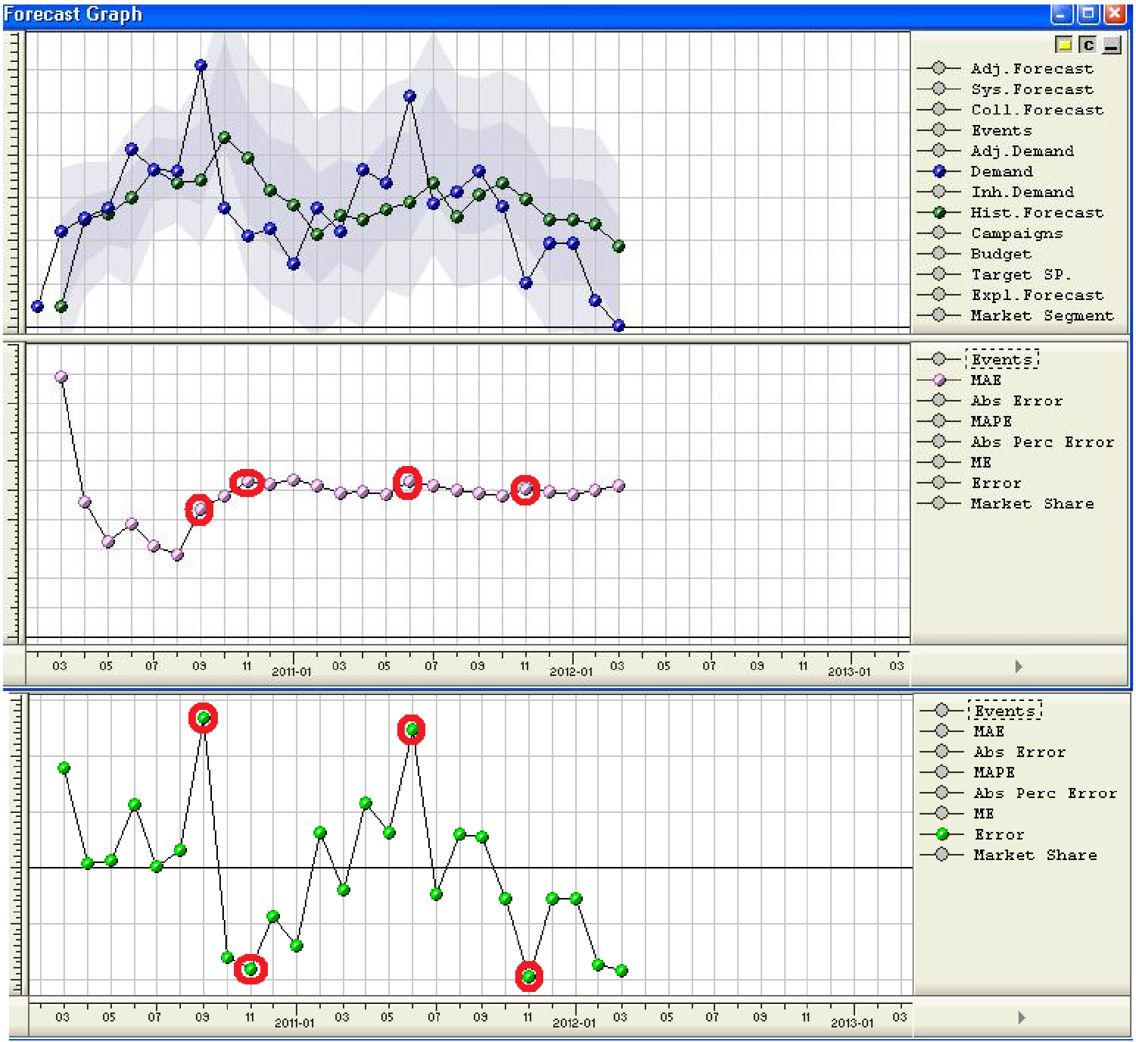


Figure 5.3. Demand peaks (adapted from the Forecasting Software).

As figure 5.3, in this case a proper measure to use is either error or absolute error. The aforementioned brings up an important with the use of control card and forecast graph, which is to be mindful of the scale of the graph and control card: sometimes the changes are relatively small but can be very significant in the absolute value. This is also one of the reasons why the use of both absolute values and control card complement each other.

In the third situation the forecast has been in control at first, with (absolute) error values close to zero, but after a while the errors have been increasing for some reason.

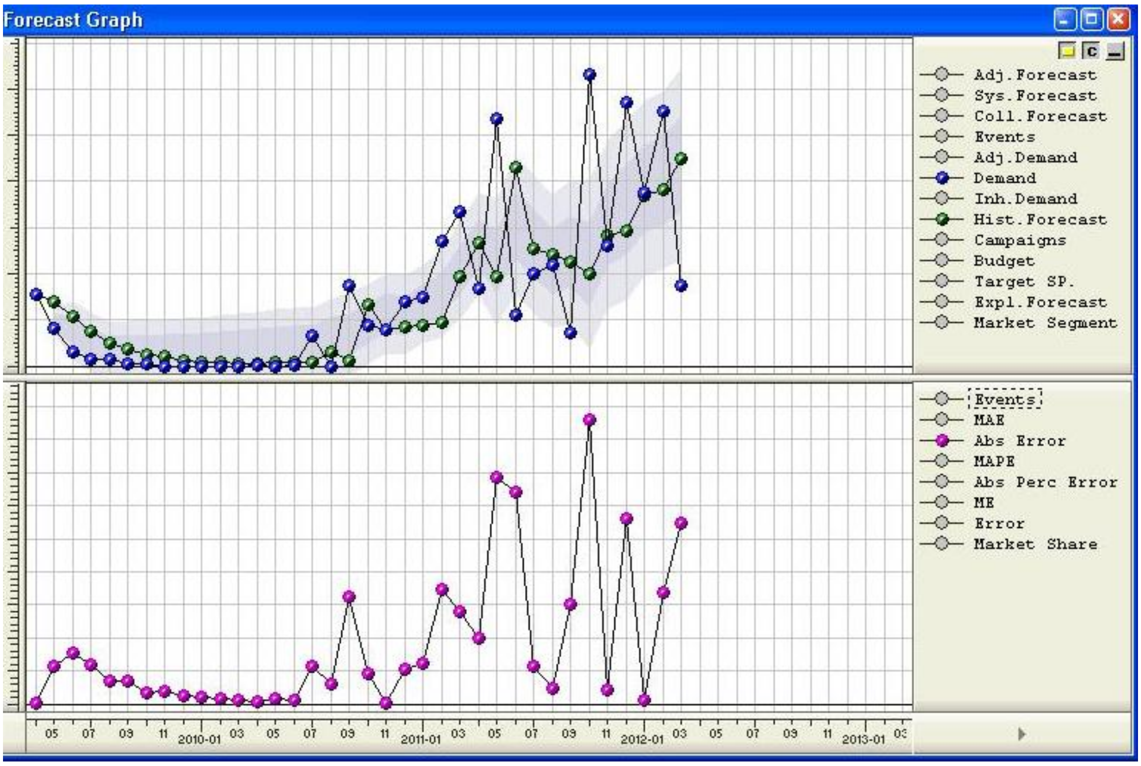


Figure 5.4. Forecast that is out of control (adapted from the Forecasting Software).

In the third situation the measure absolute error was used, because it shows the overall development of the accuracy. The same could have been noticed when using simple error: in that case the errors would have been close, in both sides of the zero in the beginning but after period 01/2011 they would have started to properly deteriorate on both sides. The error movement would have been similar to the one in the upper graph where the actual demand deteriorates to both sides of the forecast. After period 01/2011, the situation resembles a case where the random variation of demand for certain product is strong. The errors are distributed on the both sides of the zero line (when used the measure “error”) and are usually large on percentage value.

As previous examples showed, absolute error (or absolute percentage error) can be used when wanting to find out the progress of the forecasts; is it getting better (decreasing errors) or worse (increasing errors) over time. MAD and MAPE are also suitable for this since they show the average after every period; decreasing MAD or MAPE means forecasts are improving and vice versa. However, sometimes these measures can be slightly misleading, as was the case in figure 5.3. Additionally, MAD or MAPE should not be used alone, when estimating the progress of the forecast; especially in the case of strong random variation. Figure 5.5 shows why.

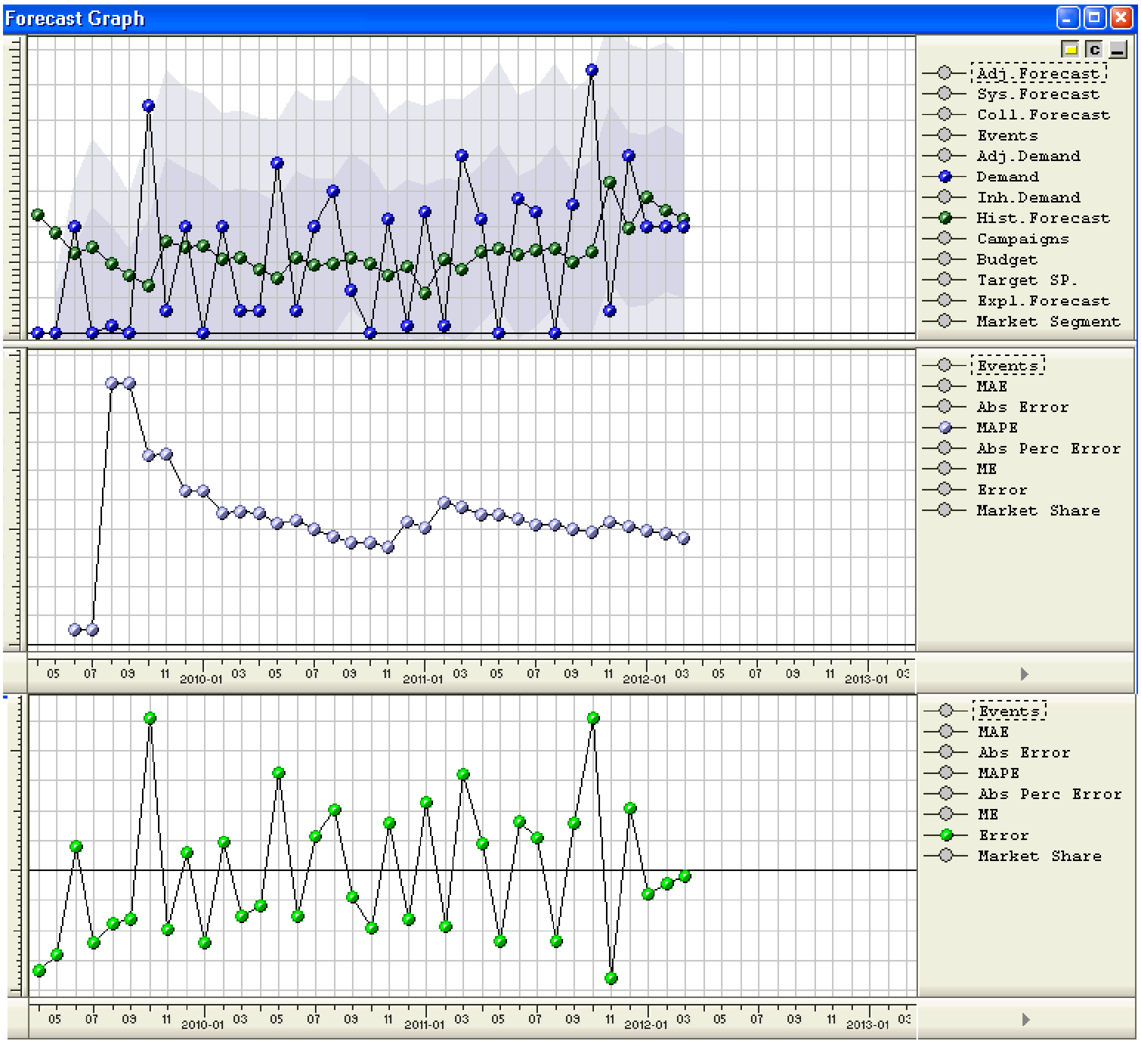


Figure 5.5. Misleading MAPE (adapted from the Forecasting Software).

As seen above, the development of MAPE would imply that the forecast is getting more and more accurate with the decreasing MAPE after period 09/2009. However, when looking at the individual errors it is quite difficult to suggest that the forecasting accuracy is actually improving. Surely there is some increase in accuracy in the beginning but the overall development is not quite as reliable.

The use of the aforementioned examples as feedback mechanisms are discussed further in chapter 6, which discusses the aspects presented in chapter 5 in more critical detail in relation to prior studies and their possible effect on the original research problem of this particular study.

6. DISCUSSION OF THE RESULTS

This chapter includes a summary of the results presented in the previous chapter and the further discussion of those results. In addition, some recommended courses of action are presented based on the results and prior studies, which were discussed in the literature review section of this study. This chapter will proceed analogically with chapter 5: first the results of changing the statistical model are discussed, after which the inclusion of judgmental input is discussed. After that, the performance measurement part of the process is discussed. Based on those sections a new approach to the current demand forecasting process is presented.

6.1. Test of changing the statistical models

As table 5.16 indicated, forecasting accuracy can be improved with the changes of the models. This means that the there are alternate approaches to the part of Statistical Forecast of the demand forecasting process, which could be applicable in the case company. However, there are several shortcomings in that part as well, as this test showed. First and foremost, even though the best models would be chosen, the overall accuracy is still not reliable enough in comparison to the expectations of the case company. When thinking about the original accuracy target that was set up by the case company, 20 %, it can be said that none of the models provide accuracies that would achieve this target on average. However, it should be emphasized that the accuracies seen in table 5.15 are averages of the sample products within each demand group and as it can be seen from the appendix 11, there are individual cases in which the accuracy target of 20 % or lower is achieved. However, there are only a few cases like this: out of the 59 test products (excluding intermittent demand) 4 products have a MAPE of 20 % or lower, however when measured with MdAPE there 8 products like this.

The problem is also that the average accuracy is the least reliable for the most important products, based on sales value, as seen in the table 5.15. Additionally, there was only a slight decrease in as the average of MAPEs decreased from 138 % to 134 % and the average of MdAPEs decreased from 47 % to 46 %. For the second-most important products the decrease was a bit better, from 98 % (MAPE) and 36 % (MdAPE) to 83 % and 35 %. However, the accuracies seemed to be the best for the third-most important products, where the error decreased from 79 % (MAPE) and 40 % (MdAPE) to 53 % to 28 %. The accuracy for products of least importance was decreased from 76 % (MAPE) and 41 % (MdAPE) to 73 % and 40 %. However, it must be remembered that the existence of some large individual MAPEs deteriorate the results especially in the case of most important products. Additionally, these groups were the ones, where some strong random variation (e.g. Level or Positive Trend) existed, which make the use of statistical models more difficult.

One interesting thing about the results is that the average of MAPEs and MdAPEs is best for GI-segment as table 6.1 indicates. The reason why this is interesting is because GI-products are manufactured for industrial markets, where the demand is usually more volatile, which means that the statistical models should be more accurate. However, as mentioned in chapter 2.4.2, there are not necessary any major differences in terms of accuracy, in the two, consumer and industrial, markets. This can be explained by the fact that the test sample included only a few products in comparison to the actual amount, and may not reflect the actual situation. Additionally, there are some large individual values that deteriorate the average of MAPEs in the AC-segment.

Table 6.1. Average accuracies of different segments.

	AC	***GI***	***PC***	***IW***
average of MAPEs	84 %	53 %	94 %	118 %
average of MdAPEs	40 %	38 %	43 %	38 %

Another, perhaps a bit surprising element is the overall bad performance of the Best-Fit option, which should choose the best possible model and parameters, based on the demand data. Based on this test, there is not a clear indication that it would be beneficial since it provided the most accurate option in total of 19 times out of the possible 71. This is perhaps a bit contradictory to literature review, according to which one of the advantages of Best-Fit option is that it chooses the best model or a parameter combination based on the demand data. However, based on this test, there is no indication that the Best-Fit option of the forecasting software would do so, which means that its use cannot be recommended here, even though its use is recommended based on the literature review

One result, which was quite clear in this test was the effect of the inclusion of seasonal profiles on accuracy, which can be seen in detail in appendix 12. The inclusion, whether it was the automatic seasonal profile for demand groups containing seasonality, based on the forecasting software’s categorization, or the predetermined profile for New Parts of the AC-segment, improved the forecast accuracy: including automatic seasonal profile decreased the average of MAPEs from 141 % to 72 % and the average of MdAPEs from 57 % to 44 %. The only exception was the one occasion in PC-segment in the group Season, where the accuracy was decreased with the inclusion of the profile. However, in that case the seasonal pattern of the product consisted individual peaks, which were categorized to seasonality by the software. Inclusion of FI-OUTDOOR decreased the average of MAPEs from 316 % to 195 % and the average of MdAPEs from 91 % to 62 %.

The inclusion of the seasonality in the statistical forecasts is aligning with the literature review (subchapter 3.2.1), which indicates that the use of statistical models should always be based on the demand pattern of the product. This means that if seasonality exists it should be accounted for in the forecasts. This is why it can be said that the inclusion of the automatic seasonal profile is the correct action and therefore there should not be any modifications for that. Also, it is recommended that if there still are some products with seasonal pattern and without an automatic profile as a setting, it should be rectified immediately. It should be point out, that even though the accuracy was increased in the group New Parts of the AC-segment, the accuracy was still quite poor. This can be explained by the fact that the profile is predetermined based on the whole demand of the group, which means that there are differences in the patterns of individual products. Additionally, because of the relatively low amount of data (compared to the other product groups), the statistical forecast can be assumed to be more inaccurate. Therefore, whether statistical models should be even used in the first place is debatable. The aforementioned is discussed later on in this chapter.

One case which was not included in the previous analysis was the group Intermittent, because of the different accuracy measures (ratio of mean absolute deviation and average demand, p. 55) used in that particular group. Results indicate that there is a connection between the occurrence of months with zero demand and the accuracy of the forecast: the fewer the months with zero demand, the better the forecasting accuracy. The average ratio (of mean absolute deviation and average demand) for products with low occurrence of zero months is 0.73, whereas for mediocre it is 1.04 and for high 1.15. There are no great differences between the segments: forecasts are altogether somewhat inaccurate, with the average ratio varying between 0.90 and 1.00. The problems relating to sporadic demand were discussed already in the literature review and therefore it is not a surprise that the accuracies are low in this test as well, even though there are models designed specifically for intermittent demand.

What is good about the results is the fact that the changes of statistical models were able to increase the forecast accuracy. Based on the results some recommendations can be given whether or not the statistical model should be changed. Because of the room for improvement within each segment and demand group and the ability of alternate models to provide more accurate forecasts, it can be said that the statistical models that were deemed to be the most accurate ones should be used. However, it should be remembered that this test has some limitations, which means that the results seen earlier cannot be generalized. Firstly, it should be noted that the tests included only a few test products per each demand group. The relatively small size of the sample group means that it is not clear how well the products actually reflect the situation within each group. Additionally, the individual characteristics of the test can have a strong impact on the results, which makes generalizations almost impossible.

For example, if a test sample with only products with relatively predictable demand would have been chosen, the results would have been different, in that case most likely more accurate, at least based on the literature review according to which statistical models are applicable, when the demand is fairly predictable. However, in this case the test sample contained more unforecastable than forecastable products, at least based on their COV value presented in subchapter 2.5.1. The effect of the COV value on forecast accuracy in this test can be seen in the appendix 13, which shows the accuracies of most accurate statistical models, for each product. In addition to this, it shows whether or not the product is forecastable or not. This can be later on used by the case company, when estimating whether certain product or a product group is forecastable or not.

The averages of MAPEs and MdAPEs for sample products, which are forecastable based on the COV are 73 % (average of MAPEs) and 31 % (average of MdAPEs), whereas the same values are 98 % (average of MAPEs) and 48 % (average of MdAPEs) for products that are unforecastable. In addition to this the correlation coefficients between the COV and MAPE (38 %) as well as COV and MdAPE (68 %) show that there is definitely some dependence between the two values (appendix 13). Based on the sample of this test and the averages of their COV values, which are seen in the appendix 14, it can be said that 16 out of the possible 34 demand groups, or 33 out of 59 products, are not forecastable. This means that the demands of the test sample's products were in general somewhat difficult to forecast. However, it should be emphasized that there are actually tens or in some cases hundreds of products within each demand group, which means that it cannot be said that the groups are not actually forecastable. Because of this it cannot be said that the average accuracies of statistical models in general are as poor as in the tests.

In addition to the aforementioned, when it came to the choosing of most accurate statistical models there was also some variation in the results within each demand group: for example in the case of group Level there were six different (for eight products) models that were most accurate. Even though the possible models could be limited when comparing them more carefully, it cannot be said with certainty that the models that were chosen, would be most accurate in most accurate when thinking about the entire demand group with tens or hundreds of products. The problem of choosing a different model was addressed earlier in chapter 2.3.3, where it was mentioned that there is no technique that outperforms others in every situation. It should also be remembered that the demand patterns of the products change all the time and basically when new demand data is included there is a change in the demand pattern, which can imply that the results change as well.

Because of the aforementioned uncertainties surrounding the results, it cannot be said with utmost certainty that a change of model would definitely increase the accuracy within the group. However, as the results of this test indicate, the Bayesian model is not the best model on all accounts. This is important because it at least proves that there is some room for improvement in the statistical forecast. Therefore, what is suggested is that in some cases the model is changed for certain product groups and the accuracies are monitored for few future periods (e.g. 3-6 months) and if an increase in accuracy is apparent (chapter 5.3.2 shows how it can be monitored), the new model is used afterwards. If the result is contradictory, the return to the old model can be done. This is in accordance with the literature review (chapter 3.2.1), according to which the use of statistical models should be monitored and updated frequently. This sort of procedure can be done first to only some of the groups or done only to certain product groups at a time if the case company does not want to risk the possible decrease in accuracy for the most important product groups. The statistical models that could be used in this can be the ones that were the most accurate based on the results of this test. Additionally, the performances of statistical models in different environments (e.g. low or high random variation) can be taken into account, when thinking about changing the models for some demand groups. Also the applicability of the statistical models for each product group can be estimated in relation to the respective COV values.

Another suggestion that can be made based on the inaccurate results is the use of other forecasting methods, other than statistical, in forecasting process. This means incorporating some judgmental input to the forecasting process. New Parts of AC-segment are one example in which all of the statistical models provided very inaccurate forecasts. It should be remembered though that when forecasts for that particular group is originally made, judgmental input is included in the process. In addition to the forecasts made for aforementioned group, forecasts for some other groups should be adjusted by the people involved in the forecasting process if the statistical model is not able to provide forecasts of sufficient accuracy or if the demand of a product is not forecastable. However, even though the results of this test sample were poor, they do not indicate with utmost certainty that this is the case also in general, which means that it cannot be said that recommended that statistical forecasts should be ignored and that judgmental forecast would then improve the accuracy. Additionally, there is no clear indication based on previous studies that the judgmental input increases the accuracy if it is not based on additional information other than the demand data available in the software. Judgmental input is further discussed in the next chapter.

6.2. Judgmental input

This chapter discusses the findings that were discussed in chapter 5.2. The purpose is to give guidelines on how the judgmental input can help the forecast accuracy in situations that were presented in chapter 5.2. The reason why judgmental input is suggested as a solution for these errors in the first place is because the errors have been so large, which indicates that the statistical model is not able to forecast the demand. At this point it should be mentioned however, that there are no guarantees that the judgmental input would necessarily increase the accuracy, and as mentioned earlier, its effect was on the forecast accuracy was not tested in this study. Therefore, most of the suggestions described here are based on the previous studies and literature review presented in this study as well as the can be interpreted from the demand data and information available about the case company's current forecasting process.

Main reason for large errors, especially in the industrial products' segments, was the large random variation. Random variation was discussed briefly in the previous chapter and its effect on the results of the test of the accuracy of the statistical models can be seen in the appendix 13, which shows the correlation between the COV and the forecast accuracy: the accuracy of a forecast accuracy made by a statistical model decreases when the COV, which is the ratio of standard deviation and the average demand, increases. In that section it was also suggested that since the statistical models are unable to improve the forecasting accuracy, judgment input could also be employed.

What makes the problem of random variation difficult is that there are no solutions that would help get rid of it entirely. Nevertheless, some previous studies provide suggest that additional information such as contracts, inquiries, preliminary orders, customers' inventory levels and production plans, customers' own forecasts and estimations about the future demand as mentioned in chapter 2.4.1. This is information that is usually more available in the case of industrial products, which is why it is often suggested, based on the previous studies, that because of the characteristics of industrial markets and the existence of aforementioned information, the forecasts could benefit from the additional judgmental input. However, it should be remembered that if there is no such information as described earlier, there is no indication that the judgmental input would prove the forecasting accuracy.

Problem might also be that because the entire demand forecasting process revolves around the forecasting software, and the statistical forecasts are calculated automatically, there is no specific incentive to incorporate judgmental input to forecasts by the people involved if it is not their primary function. In this study, it is not known how the relationship between the people making the forecasts and the industrial customers that the products are sold to is. In addition to giving straight solutions to coping with random variation, it is difficult to suggest how the case company could actually handle the use of other information sources. Therefore, it is possible to only identify certain situations, where the judgmental input should occur based on previous studies, not to give actual recommendations for in this specific case.

In consumer markets (AC-segment) the reasons for large errors were different compared to the ones in industrial markets. One of these was the occurrence of demand peaks and lows that was due to the price increase and its effect on demand. The reason why judgmental input should be used is because it is impossible for the statistical forecast to take this sort of anomaly into account, whereas people involved in the forecasting process have knowledge about it since they are the ones providing retailers with this information, which then leads to the situation, where the demand for last month of original prices is higher than normally, because they buy more products to be stored. After this the demand during first month of the new prices is lower than normally. Judgmental input can be used to diminish the effect of this situation on forecast errors.

This can be done by increasing the statistical forecast for the last month of the cheap prices and then decrease the forecast of the next month. When handled correctly it can be a good cost saving opportunity in the forecasting process as the example in table 5.20 showed. Even though the save that was achieved in the example was almost 30 %, it must be remembered that there are some limitations in the example. First, taking the effect into account is always easy afterwards, when the demand peak has already occurred and that the 30 % decrease was the absolute perfect that could be achieved because the effect was estimated after several periods of data after the event. In reality, the estimation of the full effects of the peaks is much more difficult, because it should be done prior to the event and therefore the decrease in overall error (in comparison to the situation where there is no intervention to the forecasts) is most likely much lower than the amount that was achieved in the example.

Additionally, the only cost that was used here was the inventory cost even though in the situation the demand was larger than the forecast, which means that the costs are due to increase of capacity, overtime and other costs, not inventory costs. However, inventory costs were used because it was the only cost available that helps to quantify the forecast error (MAD) in the forecasting software. Due to the aforementioned shortcomings in the example, it can be said that it does not quite fully reflect the cost situation in reality. However, even though the absolute costs are perhaps a bit different, the effect on the forecast error (MAD) is still the same. Additionally it can be seen that the relative cost of one occurrence can be very high in comparison to the average cost, in the example twelvefold.

Even if there are some difficulties in determining the overall effect of the price increase (or other cause for unusual changes in demand), the judgmental input should still be used because the statistical model does not take it into account unless it is a specific econometric model, where the price-demand dependence can be determined. However, these are not available in the software, which means that the only way to handle these situations is the use of judgmental input. When adjusting the forecast of the statistical model, effects of the previous price changes on demand can be analyzed in order to estimate the effect of the current price change. Even if the increase in accuracy is only relatively small, it can still be a good cost saving opportunity, for example if there are tens or hundreds of products where this sort of situation happens.

If the reason for unusually large errors is some other than the aforementioned change of prices or some other similar situations, effect of which can be estimated beforehand, then the proper solution is also a bit different. In that case it can be difficult to know before because in the consumer market segment the demand can be highly influenced by many different causes. According to the literature review, even if these causes are not known beforehand they should be taken into account afterwards, because the unusual large or low demand has implications to the future forecasts made by the statistical model because it uses the historical demand as a basis for the future forecasts. Therefore, if they do not reflect normal circumstances, they should usually be removed from the demand data so that they would not affect the future forecasts.

Another situation where certain regular errors occurred was the case of product group New Parts. The regular cause for errors in that particular group has been that the forecasts are frequently too high. Naturally it might make sense that slightly higher forecasts are made in the beginning, for example to make additional safety stock. However, as the case company has instructed, manual forecasts should be checked after a few months and updated if necessary. If the forecast is regularly too high the forecast should naturally be modified to lower than originally. In this case the problem is obviously the determination of the level of modification. The aforementioned depends on the magnitude of the error as well. If the percentage errors are large it makes sense to modify the forecast for example 20 % or even 30 % in the beginning and see what happens to the error. If the forecast is still larger than the forecast there can be additional modification. However, if the percentage error is small, even a 5 % or 10 % modification can be sufficient. The important thing is to follow the manual forecasts and monitor the development of the forecast error.

There is also another typical pattern that has been repeating in the New Parts group, which has been the fact that not all of the products are behaving according to the predetermined profile, which has lead to the fact that forecast errors have been quite large on some occasions. The solution to this can be related to the previous problem, where the modification of the forecasts is made every period. The recommendations that were suggested above are in accordance with the previous studies, chapter 3.2.2, which indicate that the judgmental input should occur based on predefined triggers such as campaigns or promotional activities (e.g. price changes in the example). Additionally if product is new and there is no demand data available, the forecasts should be using other measures than statistical. Additionally, judgmental input should only be used for a few future periods and its progress should constantly be followed and possible modifications should be made.

The obvious problem that arises from incorporating judgmental input in the demand forecasting process is the time and resources it takes. Since all of the products have a different sort of demand pattern and there are a huge number of different products, it is impossible to change the forecasts of all of the products individually (even though not all of the forecasts require changing). This is true, whether it is the case of reacting to random variation or individual unusual situations. However, an option that is highly recommended here is to choose the key products, which are the most important ones, where these sorts of errors are most costly or in some other way harmful for the case company. This is backed up by some of the previous studies, where it is suggested that forecasting should only include the most important products or customers.

It should be remembered that the aforementioned situations were only examples of the situations, where incorporation of judgmental input should be done. Additionally, judgmental input is much more complex entity than it is described here. However, because of the perspective of this study and the lack of available additional information about the procedures, how judgmental input is actually utilized on a day-to-day basis in the case company, making additional recommendations or suggesting other courses of action is rather difficult to do. Also, the effects of utilizing some of the recommendations presented in this chapter were not tested in the premise of this study, which means that their effects on the demand forecasting process of the case company is also difficult to properly evaluate. Therefore, the recommendations are merely based on the previous studies about the subject, which indicates that including judgmental input in the manner described can help to increase the forecasting accuracy.

6.3. Performance measurement of the demand forecasting process

Performance measurement procedures of the case company were partially discussed in chapter 5.3 of this study, whereas the findings of previous studies were presented mostly in chapter 2.5 and 3.2.3. The problem with the performance measurement in the demand forecasting process of the case company has been that it has not received the attention it deserves. This can be due to the fact that it has not been seen as a tool to improve the accuracy of the forecasts. Another reason is that because of the forecasting software and the abundance of data it exists, it is difficult to find the appropriate data when assessing the quality of the forecasts. The purpose of this chapter is to combine results of some of the previous studies presented in the literature review and the procedures presented in chapter 5.3 and to offer some guidelines how the performance measurement should be done in the case company. In addition to this, the ability of the forecasting software to function as a feedback mechanism is discussed.

The problem of data abundance can be helped with the same solution that was offered in the previous chapter, which is to focus only on most important products, which can be done by organizing the data of the forecasting software in different ways (the ways of organizing depend on the fact what are searched). Another problem is the existence of various error measures and their different uses. However, there should not be a use of only one error measure, because they can be used for different purposes. Mean error (ME) can be used to find bias (forecasts are constantly higher or lower than the actual demand) in the forecasts. If bias does occur (figure 5.2), the solution is, as mentioned in chapter 5.3: adjusting the forecasts to be larger (negative ME) or smaller (positive ME) than normally.

Error or mean absolute deviation (MAD) can be used to find some unusual individual errors like in figure 5.2 and 5.3. This was already discussed in previous chapter. Absolute error can be used to find out whether the forecast is in control or not (figure 5.4). If situations such as the one figure 5.4 are found it can be recommended that some other forecast model is used. After the model is changed the progress of absolute error can be further followed to find out, whether the accuracy is improved with the change of model or not. In general, the situation, as the one in the figure 5.4, requires further analysis since there could be a number of reasons why forecast accuracy is decreasing.

Another way to follow the progress of the forecast accuracy is to use MAPE or MAD. However, as figure 5.5 shows these can be sometimes slightly misleading and indicate that the forecast accuracy is improving, when it is necessarily not. According to the literature review, the mere measurement of the process does not improve its quality, but information that is achieved by the measurement has to be taken advantage of somehow. That is why it is emphasized that if situations, such as the examples in figures 5.2-5.5, occur, there should be an immediate reaction to these situations. This is one of the reasons why the performance measurement of the case company should not limit to the automatic calculations made by the forecasting software, but constant improvement should be pursued. One reason why the performance measurement is important as well is because other processes in the company use forecasts as foundations for their some of their decisions. Therefore, improving the quality of the demand forecasting process can also help to improve the quality of other processes of the company.

Therefore, it is recommended that the case company focuses some of the effort of their demand forecasting process to the future forecasts, but also concentrate a bit on the errors that have happened in the past. By doing this, it is much easier to avoid repeating and doing the same mistakes all over again. Additionally, because the forecasting software functions as the primary feedback mechanism to the people in charge of forecasting, understanding the errors and how they can be avoided later on and how the future forecasts can be modified based on the errors is of utmost importance. This is backed up by some of the previous studies according to which following the progress of the forecast accuracy is important.

Because the calculations of forecast accuracy are done in the software, it serves as a conduit for feedback about the process and its progress. This is somewhat problematic because the software presents only the values of the error or in the case of control card the development of the accuracy. However, for example the context (e.g. in which situations the largest errors occurred) are not presented and because of the abundance of data, the users of the software have to find the most significant errors themselves. Additionally, there are no warning-systems that would be triggered by certain accuracies or errors but the forecasters have to look for this sort of information themselves. Therefore, even though the automatic calculation of error values based on different accuracy measures helps the forecaster, it only provides the values which must be further interpreted by the people involved in the process.

However, it should be mentioned that even though previous studies indicate that the monitoring of forecast accuracy and the overall performance measurement of the forecasting is important, the studies also conclude that it is difficult to ascertain, what are the overall impacts on the forecast accuracy. It should also be remembered that there are other accuracy measures than the ones presented here that can be used in the performance measurement. Additionally, the performance measurement does not limit to the approaches discussed in this chapter but there other possible ways it can be done, depending on the context. Having said that, it should also be reminded that the purpose of this study was to focus on the approaches available for the case company. Hence, the focus was only on the accuracy measures available in the forecasting software and the use of that software as a feedback mechanism.

6.4. Overall summary of the modifications to the current demand forecasting process of the case company

The purpose of this chapter was to evaluate some possible modifications, analyze their effects and based on that, offer some recommendations on how the current demand forecasting process can be changed and possibly improved. The original demand forecasting process of the case company is depicted in figure 4.1 To improve the accuracy of the statistical forecast a change of models is recommended: based on the findings the Bayesian model is not necessarily most accurate one on all accounts, which shows that there is room for improvement in the phase of statistical forecast. However, because it cannot be said with utmost certainty that the competing models would necessarily lead to the improvement of the average accuracy, it is recommended that the models are changed first to only some of the demand groups at a time (or in the test platform of the software) and their performance is monitored for a few (e.g. 3-6 months)

After this, it can be further assessed whether the change actually improved the forecast accuracy or not and some courses of action are taken based on that result. If the results are improved with the change, the new model can be used in the future as well. However, if the results are not improved the old model can be taken into use again. Whichever the case is, the change can be done later for other product groups as well. The change can be based on the models of table 5.16 (or appendix 14) presented in chapter 5 for example, however other factors such as the existence of random variation within the demand group can also be taken into account when thinking about the changes of models. In some occasions, it was suggested that the statistical forecast should not be used at all because of its inability to cope with increasing random variation, as shown in the appendix 13.

The incorporation of judgmental input was found to be beneficial in certain situations: based on the literature review it should be used if information, other than the demand data, is available. The existence of some of these situations, such as random variation or promotional activities, was identified in the case company and the effects of judgmental adjustment of statistical forecast in these situations on the forecast accuracy and costs were estimated. Based on that estimation as well as the findings of the previous studies it was recommended that judgmental input should be included in the demand forecasting process at least in the existence of promotional activities. How judgmental adjustment can be done in those situations depends on the specific situation. Nevertheless, some examples were given to these situations. Additionally, it was recommended that in the presence of strong random variation, especially in the industrial products' segments, judgmental adjustment is used because of the inability of the statistical forecast to cope with the random variation. In that case alternate sources of information should be pursued to help adjust the statistical forecasts. If the forecast of a product is judgmentally adjusted, the forecast should be checked and possibly updated after each period and inclusion of new data. Additionally manual forecasts should not be made for too long in the future.

Because the lack of analysis of the accuracy measures as well as the performance measurement in the demand forecasting process of the case company, this study focused on offering guidelines on how the automatically calculated error values can be used to improve the quality of the forecasts and how they can be used as feedback mechanisms when estimating the progress of the forecast accuracy and the quality of the demand forecasting process. At least the following things should be checked with the help of the control card:

1) Is there any bias in the forecasts, i.e. are forecasts constantly too high or too low (mean error)? If yes, they should be modified based on bias. (New Parts)
2) Are there any peaks or lows in the demand curve and in the errors (error, MAD)? If yes, why and how could it be avoided in the future? (Level and Level/Season of the AC-segment)
3) Is the statistical model able to handle the random variation or should the statistical model be changed or additional judgmental input be included (absolute error, MAD, MAPE)? (GI-, PC- and IW-segments)
4) Is the change of model or incorporation of judgmental input increasing the accuracy of the forecast (MAPE and MAD)? (Situations where model is changed or judgmental input included)

Like the step of judgmental input, the overall impacts of the performance measurement procedures were not tested. Instead the previous studies and the literature review was used as a basis, when coming up with the recommendations suggested here. The aforementioned courses of actions are also summarized in the figure 6.1, which shows the demand forecasting process with the key recommendations for changes or improvements.

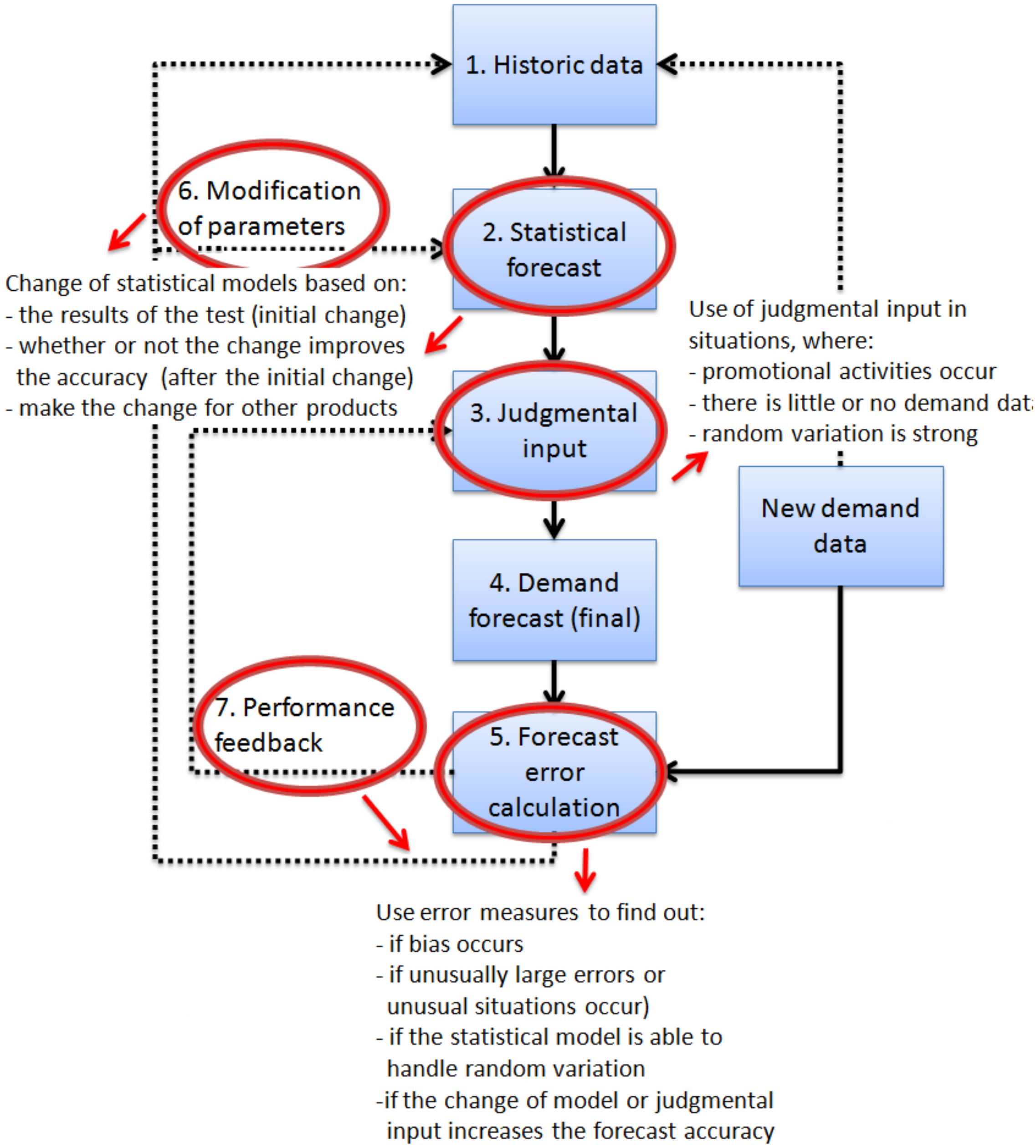


Figure 6.1. The demand forecasting process and recommendations.

It should be remembered that there are certain limitations that have to be included when thinking about suggesting the use of the aforementioned recommendations. Biggest one of these is that the case company has thousands of different products, which means that the forecasting practices that are presented here take considerable amount of time. However, some of these like the change of statistical models can be done for entire product groups (albeit it is always better when done only for individual products), which makes the situation a bit easier. However, judgmental input and the evaluation of the progress of the forecast accuracy cannot be done only for product groups, but for individual products as well because of the different demand patterns within the groups, which are needed to find errors that were presented earlier. A solution for this is to focus on only the most important products. Most important products can be found with the help of Value of MAD categorization presented earlier in this study.

Furthermore, what should be emphasized here is that there are some differences in different aspects of demand forecasting depending on the segment or product group, which is why the demand forecasting process in figure 6.1 is not necessarily accurate for all of the products. An example from situations, where some of the procedures might be different is the case of new products because there is no historical demand data available, which means that they cannot be forecasted with statistical models. This means that in this case the process starts from step 3 instead of step 1. Additionally, the inclusion of judgmental input can be focused on segments or product groups with more volatile demand, whereas in the case of forecastable products statistical models can relied on. It should also be remembered that the recommendations presented here are just the possibilities how the demand forecasting process can be modified. In the end, it is up to the people involved in the process to decide, which aspects of the aforementioned suggestions they seem suitable for which segments, product groups or individual products.

When conducting this study some general advantages and disadvantages in the use of the forecasting software as a tool in the demand forecasting were identified. Advantages are:

- The monthly automatic calculation of forecasts and the forecast errors for different products
- Different statistical models (and the possibility to easily change them) and accuracy measures
- The easiness of adjusting the statistical forecast if needed
- Data can be organized and limited to include what is searched or needed

Whereas the disadvantages are:

- Lack of documentation, e.g. if forecasts are adjusted by users, the adjusted forecast will remain in effect but it is not separated in any way, which makes the evaluation of its effect on the accuracy impossible
- Estimation of the effects of changes of models in the software is somewhat difficult and the proper calculation cannot be done
- Abundance of options, which might make it unclear to know, which statistical models or accuracy measures to use in which situations (for people involved in the demand forecasting process)
- Because everything is calculated automatically, the forecast might be taken for granted (for people involved in the demand forecasting process)

- The use of the data of the software as a feedback mechanism about the quality of the demand forecasting process (for people involved in the demand forecasting process)

It should be mentioned that there are some other advantages and disadvantages, depending on the perspective. The aforementioned list was only made in the perspective of this particular study.

7. CONCLUSIONS

This chapter presents the conclusions about the literature review, based on which the concept demand forecasting process was defined and the functionality of that process in the case company. The study is also reflected on the goals made in the beginning of the study. Lastly, the methodology of the study, its limitations, possible recommendations for usage purposes and the possibilities for further research or development of the study are discussed.

7.1. Reflection on the study, its purposes and original research problem

The purpose of this study was to develop and improve the demand forecasting process of the case company in order to provide the case company with more accurate forecasts. To improve the current demand forecasting process, a proper definition of what demand forecasting process actually contained, or at least should contain, was needed. This problem was assessed with the help literature review based on the operations management and supply chain management literature as well as previous studies regarding forecasting practices. In other words, the mission of the literature review was to answer the question, which actions and procedures, related to forecasting, should a company implement in order to ensure an effective demand forecasting process.

The problem that was encountered, when defining the concept of demand forecasting process, was the lack of process description in the literature or previous studies about the subject, which was backed up by Kerkkänen (2010). Additionally, there was no previous, universally accepted definition of demand forecasting process. Instead the prior research seemed to focus only on individual concepts of demand forecasting, not the entire process. In the presence of descriptions about the demand forecasting process, an additional problem was encountered: the interchangeable use of the terms demand management process, demand forecasting process and demand planning process (e.g. Hogarth & Makridakis 1981, Croxton et al. 2002, Chambers et al. 2004, Stadler & Kilger 2008). To at least partially solve the aforementioned problems, some aspects out of the different process descriptions and definitions of the three different terms were combined in order to form a definition for the demand forecasting process that would suit the specificities of this particular study.

Also, to gain a deeper understanding of the different aspects of the process, some additional concepts from prior research on individual aspects of demand forecasting were taken advantage of when defining what should be included in the demand forecasting process. Additionally, the suitability of some of the forecasting concepts to certain operational environments, such as different markets (consumer or industrial) was presented (Kerkkänen 2010, Mentzer & Kahn 1995). This was done because of the lack of distinction made in the forecasting literature between the applicable forecasting concepts between the two markets and because the case company to which the demand forecasting process was later applied was operating in both markets. However, the suitability of the concepts was made in a relatively abstract level because of the shortcoming of most of the previous studies to properly define to which situations (type of products, customers, demand etc.) they are most suitable for.

Despite some of the aforementioned problems, as a result of the literature review and the definition of the concept Demand Forecasting Process, a multi-step -model was able to be created. As a basis for this, the concept of Demand Planning Process made by Stadler and Kilger (2008) was used because of its suitability for the characteristics of this study. It should be emphasized that even though the process description by Stadler and Kilger was used as a basis, the demand forecasting process of this study combined aspects from other prior studies as well. Therefore, it can be said the demand forecasting process of this study is merely one interpretation of the concept and its use has some limitations. Nevertheless, the demand forecasting process of this study (chapter 3.3) included the following steps:

1) Preparation of demand planning structures and historic data
2) Computation of statistical forecast
3) Judgmental forecasting
4) Consensus forecasting and release of forecast
5) Calculation of forecasting errors
6) Modification of parameters
7) Performance feedback

Because the definition of the demand forecasting process was made mainly in accordance with the characteristics of this particular study some of the aspects of the demand forecasting process were neglected because of the scope and limitations of this study. Therefore, the process description that was made based on the literature review is not necessarily suitable for all situations and in some other studies some other aspects of forecasting might have been additionally included. It should also be remembered that even that the demand forecasting process is an iterative process and in some cases some of the steps can be done simultaneously or some of them can be completely excluded (e.g. new products).

Based on the multi-step –model that was created, an answer to the main research problem, *“how is the current demand forecasting process of the case company and how to develop and improve that process and thus provide more accurate forecasts for the case company?”* of this study was searched. To provide an answer to the research problem the current demand forecasting procedures of the case company were analyzed, after which they were compared to the ideal practices presented in literature. Based on those comparisons, three different targets of improvement were found. They were: computation of statistical forecast, judgmental input and the performance measurement of the process, which included steps 5, 6 and 7.

The problem with the computation of statistical forecast was that even though the forecasting software, in which the forecasts are made, includes several different statistical models designed for different products with different kind of demand patterns, only one of them, the Bayesian model, was used (except in the case on New Parts for AC-segment). This approach is contradicted by some of the previous studies, such as Armstrong and Green (2006) or Stadler and Kilger (2008), who emphasize the importance of the regular comparison of different forecasting methods and their accuracy. Because of this the accuracies of some other statistical models of the software were calculated for a group of test products, based on which it was recommended whether the statistical model that is used should be changed or not.

The results of the comparison showed that the Bayesian model was not the most accurate in all of the occasions, but there were other models with which a better accuracy was achieved. This indicated that there was some room for improvement in the accuracy of the statistical forecast, which meant that the changes of statistical models in some occasions could improve forecast accuracy and thus improve the quality of the demand forecasting process, which was the purpose of this study. However, because of the relatively small size of the test sample and the differences in the demand characteristics of the test products it could not be said with utmost certainty, in which cases the statistical model should definitely be changed. Instead, it was recommended that the change is first done for only of the product groups to test the performance of models that were most accurate in the test and only after a few months of monitoring the performance some additional changes could be made as well.

The problems related to the step of judgmental input were more difficult to assess because of the limitations of this study, which are discussed in more detail in chapter 7.1. Because of these limitations the effects of judgmental input on the forecasting accuracy could not be tested like the accuracies of different statistical models. One problem of judgmental input in the company, which was able to be studied further was, that there were no guidelines on which situations judgmental input should be included. These situations were searched by finding some of the largest error values for each product segment, because those were the cases where there was most room for improvement in terms of forecast accuracy. The situations where errors where relatively high where slightly different depending on the product segment: in industrial segments the large errors were due to the strong random variation, whereas in consumer markets they were mostly the cause of individual unusually large errors.

The fact that strong random variation played a role in the industrial segments was aligning with the previous studies (e.g. Kerkkänen 2010, Mentzer & Kahn 1995), which indicated that the demand is usually more volatile in the industrial market in comparison to consumer market, which also makes forecasting more difficult. The effect of increase of volatility on the forecast accuracy can also be seen in the appendix 13, where dependence of the two variables exists. In the case of strong random variation in the industrial markets some recommendations were based on the previous findings presented in the literature review. In the case of consumer segment, much more concrete recommendations were able to be given because of the different nature of the causes for large errors. There it was also shown (table 5.20) how the inclusion of judgmental input can improve the forecast accuracy. Even though the mere estimation of the impact on the accuracy was not always able to be made it can be said that identifying some of the cases were judgmental input could be applicable could be seen beneficial for the case company and the company's demand forecasting process because they are able to serve as guidelines in the future and therefore guide some of the aspects of demand forecasting to the parts, where improvements are possible based on the previous studies.

The problem related to the performance measurement was that it had received very little attention in the case company thus far, which is in contrast with the previous studies (e.g. Gardner 1983, Croxton et al. 2002, Stadler & Kilger 2008), according to whom performance measurement is an important part of the demand forecasting because it can be used for example to set targets, monitor the progress of the forecast and its quality. Because the forecast errors were automatically calculated every month by the forecasting software, this study focused on emphasizing the importance of the proper use of the different accuracy measures. The use of some of the measures was related to the phase of judgmental input, which meant that the step of judgmental input could be further improved with a proper analysis of different accuracy measures.

However, the impact of the use of different accuracy measures on forecast accuracy was not possible to assess here, which meant that the recommendations that were given were based mainly on previous findings about the subject. Having said that it should be noted, that based on those findings it can be said that the proper use of different accuracy measures and the overall improvement of the performance measurement can further improve the quality of the entire demand forecasting process. To summarize, the individual steps of the demand forecasting process can improve the quality of the process, which means that the aspects presented in the study and the recommendations they are based on, answer to the research problem of this study. Additionally, even though the summed up improvement of all the aspects is impossible to fully estimate, it can be said that the study has served its purpose, which was to improve the demand forecasting process of the case company. However, there are some limitations that are discussed in the next chapter.

7.2. Limitations of the study, usage purposes and further research opportunities

Even though it can be suggested that the research problem was answered and that the purpose of this study was filled, there are some limitations in this study, which mean that some of the solutions presented in this study are not absolutely the best ones. Because of the external perspective of this study on the forecasting practices, certain assumptions about the current demand forecasting process had to be done based only on the demand data available in the forecasting software. In other words, all of the information that was needed to analyze the current demand forecasting process was not necessarily available. For example, in situations such as the use of judgmental input or the performance measurement, it was not clear how some aspects of these two are actually handled on day-to-day basis within the company. This meant that some of the recommendations that were given in this study might already be applied in the case company. In addition to this, because of some of the modifications on the day-to-day practices could not actually be tested, in some cases the recommendations were mostly based on previous studies about the subject.

Another factor that limited some of the possible solutions was the specific forecasting software, which is the center of the demand forecasting process of the case company and in which all of the individual steps of the process are done. This lead to the fact, that the functionality of the aspects of the demand forecasting process, which was analyzed, was actually an analysis of the aspects of the particular forecasting software as a tool of demand forecasting. This meant that some of the alternatives or modifications for the process, which were studied, included only the ones that can be done with the forecasting software. For example, in the case of statistical models only the accuracies of the ones that were available in the forecasting software were tested. Therefore, the best possible accuracies that were achieved with some of the models were not necessarily absolutely the best possible accuracy that can be achieved. The same applies for the use of accuracy measures. In this study only the accuracy measures that were available in the software and their use was discussed, which means that other accuracy measures that can also be used to improve the quality of the process were excluded. Hence, it is unclear whether or not the quality of the demand forecasting process could be further improved and if so, how.

Third limitation of this study is the relatively low amount of products that were used in the different parts of the study. The reason why only a handful of products were chosen was because of the abundance of products and the limited resources of this study, which meant that there had to be some simplifications. However, this meant also that in some parts the generalization about the results is not possible. For example in the comparison of statistical models only a relatively small test sample was used, which made it impossible to recommend certain immediate changes for the computation of the statistical forecast, which meant that alternative actions had to be recommended.

Additionally, the situations for judgmental input were only identified for three out of ten product groups (albeit they were the most important ones), which meant that some of the product groups and possible some different situations where judgmental input could be used were excluded.

Because the aforementioned limitations have had an effect on the scope and areas of focus in this study, the recommendations are not necessarily applicable in other studies. Also, because there are only a few well-structured documentation of the demand forecasting process in the literature, the process described in chapter 3 was made for the specific research purposes of this particular study. Therefore, one of the possibilities of further research could be to ascertain the suitability of demand forecasting process in different forecasting situations, based on the characteristics of products, their demand, customers or the operational environment of the company. Additionally, the role of the specific forecasting software in the process of the case company meant that when some of the tasks were analyzed, the analysis focused more on how these tasks could be done with the software, which meant that they were not necessarily the absolute ideal solutions given in the literature. Therefore, some of the concrete recommendations (e.g. performance of statistical models) given in this study for the improvement of the demand forecasting process are mainly applicable only in this particular study.

However, some of the more abstract recommendations such as the role of judgmental input can be applied to other studies as well. The aforementioned is also an earlier where there is definitely room for additional research. Even though the role of the human judgment in demand forecasting has been recognized in the previous studies, the studies have failed to provide concrete recommendations, in which real-life situations is human judgment needed and in which situations it is not needed. Therefore, one possibility for additional research would be the further identification of specific forecasting situations, where human judgment should be included and additionally, how should it be included in the demand forecasting process. This is also related to one other problem of this study, which was that even though certain recommendations (e.g. use of accuracy measures as feedback mechanism) were seen as beneficial in improving the quality of the process, and therefore the forecast accuracy, their overall impact could not be quantified in terms of accuracy. Therefore, one other possibility for additional research could be to ascertain the impact of performance measurement of the demand forecasting process on the forecast accuracy.

BIBLIOGRAPHY

Armstrong, J. 1978. Forecasting with Econometric Methods: Folklore versus Fact. Journal of Business, Vol. 52, pp. 549-564.

Armstrong, J. 2001. Principles of Forecasting: A Handbook for Researchers and Practitioners, Norwell USA. Springer. 849 p.

Armstrong, J. & Collopy, F. 1993. Causal Forces: Structuring Knowledge for Time Series Extrapolation. Journal of Forecasting, Vol. 12, pp. 103-115.

Armstrong, J. & Collopy, F. 1992. Error Measures for Generalizing About Forecasting

Methods: Empirical Comparison. International Journal of Forecasting, Vol. 8, No. 1, pp. 69-80.

Armstrong, J. & Green, K. 2006. Demand Forecasting: Evidence-based Methods. Monash Econometrics and Business Statistics Working Papers 24/05. Monash University, Department of Econometrics and Business Statistics. 17p. http://ideas.repec.org/p/msh/ebswps/2005-24.html. Read 6.3.2012.

Armstrong, J. & Yokum, J. 1995. Beyond Accuracy: Comparison of Criteria Used to Select Forecasting Methods. International Journal of Forecasting, Vol. 11, pp. 591-597.

Ballou, R. 1999. Business Logistics/Supply Chain Management, Fifth Edition, New Jersey USA. Pearson Education. 789p.

Behn, R. 2003. Why measure performance? Public administration review. Harvard University. Vol. 63, No. 5, pp. 586-606.

Bititci, U. S. & Muir, D. 1997. Business Process Definition: A Bottom Up Approach. International Journal of Operations and Production Management. Vol 17, No. 4, pp. 365-374.

Brehmer, B. 1980. In One Word: Not from Experience. Acta Psychologica. Vol. 45, No. 1-3, pp. 223-241.

Brown, L., Hagerman, R. Griffin, P. & Zmijewski, M. 1987. Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings. Journal of Accounting and Economics, Vol. 9, pp. 61-87.

Buffa, E. 1983. Modern Production/Operations Management, Seventh Edition, USA. John Wiley & Sons Incorporated. 681p.

Bunn, D. & Taylor, J. 2001 Setting accuracy targets for short-term judgmental

sales forecasting. International Journal of Forecasting, Vol. 17, No. 2, pp. 159-169.

Bunn, D. & Wright, G. 1991. Interaction of Judgmental and Statistical Forecasting Methods: Issues & Analysis. Management Science. Vol. 37, No. 5, pp. 501-518.

Chambers, S., Johnston, R. & Slack, N. 2004. Operations Management, Fourth Edition, Harlow England, Pearson Education Limited. 793p.

Chopra, S. & Meindl, P. 2001. Supply Chain Management; Strategy, Planning and Operation, Upper Saddle River USA, Prentice Hall. 544p.

Croxton, K., Lambert, D. & García-Dastugue, S. 2002. The Demand Management Process. The International Journal of Logistics Management. Vol. 13, No. 2, pp. 51-66.

Davenport, T. 1993. Process Innovation. Boston USA. Harvard Business School Press. 352p.

Hogart, R. & Makridakis, S. 1981. Forecasting and Planning; an Evaluation. Management Science. Vol 27, No. 2, pp. 115-138.

Holt, C. 2004. Forecasting seasonals and trends by exponentially weighted moving averages. International Journal of Forecasting. Vol. 20, No. 1, pp. 5-10.

Hyndman, R. & Koehler, A. 2005. Another Look at Measures of Forecasting Accuracy. Working Paper 13/05. Monash University, Australia. Department of Econometrics and Business Statistics.

Fildes, R., Hibon, M., Makridakis, S. & Meade, N. 1998. Generalising about univariate forecasting methods: further empirical evidence. International Journal of Forecasting. Vol. 14, No. 3, pp. 339-358.

Fischer, I. & Harvey, N. 1999. Combining Forecasts: What Info Do the Judges Need to Outperform the Simple Average? International Journal of Forecasting. Vol. 15, pp. 227-246.

Forslund, H. & Jonsson, P. 2007. The Impact of Forecast Information Quality on Supply Chain Performance. International Journal of Operations and Production Management. Vol. 27, No. 1, pp. 90-107.

Gardner, E. 1983. Automatic Monitoring of Forecast Errors. Journal of Forecasting. Vol. 2, No. 1, pp. 1-21.

Gardner, E. 1985. Exponential Smoothing: The State of the Art. International Journal of Forecasting. Vol. 4, No. 1, pp. 1-28.

Herbig, P., Milewicz, J. & Golden, J. E. 1993. Forecasting: Who, What, When and How. Journal of Business Forecasting. pp. 16-21.

Hopwood, W. & McKeown, J. 1990. Evidence on surrogates for earnings expectations within a capital market context. Journal of Accounting, Auditing and Finance, Vol. 5, pp. 339-368.

Kerkkänen, A. 2010. Improving Demand Forecasting Practices in the Industrial Context. Thesis for the degree of Doctor of Science. Lappeenranta, Lappeenranta University of Technology, 71p.

Kerkkänen A, Huiskonen J, Korpela J. & Pirttilä T. 2008. Assessing demand forecasting practices in the B2B environment, 15th International Symposium on Inventories, Budapest, Hungary, 22.-26.8.2008.

Kiely, D. 1999. Synchronizing Supply Chain Operations with Consumer Demand Using Customer Data. The Journal of Business forecasting. Vol. 17, No. 4, pp. 3-9.

Lawrence, M., Edmundson, R. & O'Connor, M. 1985. An examination of the accuracy of judgemental extrapolation of time series. International Journal of Forecasting, Vol. 1, pp. 25-35.

Lawrence, M., Goodwin, P., O'Connor, M. & Önkal, D. 2006. Judgmental forecasting: A review of progress over the last 25 years. International Journal of Forecasting, Vol. 22 No. 3, pp. 493-518.

Lawrence, M. & Makridakis, S. 1989. Factors affecting judgemental forecasts and confidence intervals. Organisational Behaviour and Human Decision Processes, Vol. 42, pp. 172-189.

Lee, H. 2002. Aligning Supply Chain Strategies with Product Uncertainties. California Management Review. Vol. 44, No. 3, pp. 105-119.

Mahmoud, E. 1982. Short-term forecasting: matching techniques to tasks. An integrated framework and empirical investigation. Ph.D. Thesis. State University of New York at Buffalo.

Mahmoud, E. 1984. Accuracy in Forecasting: a Survey. Journal of Forecasting, Vol. 3, No. 2, pp. 139-159.

Makridakis, S. 1988. Metaforecasting. International Journal of Forecasting, Vol. 4, No. 3, pp. 467-491.

Makridakis, S., Andersen. A., Carbone. R., Fildes, R., Hibon, M., Lewandowski. R., Newton, J., Parzen, E. & Winkler, R. 1982. The Accuracy of Extrapolation (time series) Methods: Results of a Forecasting Competition. Journal of Forecasting, Vol. 1, pp. 111-153.

Makridakis, S., Wheelwright, S., & Hyndman, R. 1998. Forecasting Methods for Management, Third edition. New York USA. John Wiley. 642p.

Mentzer, J. & Cox, J. 1984. Familiarity, application, and performance of sales forecasting techniques. Journal of Forecasting, Vol. 3, No. 1, pp. 27-36.

Mentzer, J. & Kahn, K. 1995. Forecasting Technique Familiarity, Satisfaction, Usage, and Application. Journal of Forecastin, Vol. 14, No. 5, pp. 465-476.

Mentzer, J. & Kahn, K. 1995. Forecasting in Consumer and Industrial Markets. The Journal of Business Forecasting. Vol. 14, pp. 21-28.

Mentzer, J. & Moon, M. 2005. Sales Forecasting Management - A Demand Management Approach, Second Edition, Thousand Oaks Canada, Sage Publications. 368p.

Merriam-Webster Incorporated. 2012. Online dictionary. [http://www.merriam-webster.com/dictionary/]. Date of access 8.5.2012.

Millikan, A. 2006. Demand Planning: Managing the Unforecastables. The Journal of Business Forecasting. Vol. 25, No. 2, pp. 3-9.

Muzumdar, M. & Fontanella, J. 2006. The Secrets to S&OP Success. Supply Chain Management Review. http://alan.com.mx/wp-content/uploads/2010/04/SOP-must-read.doc. 8.5.2012.

Neely, A., Gregory, M. & Platts, K. Performance measurement system design. International Journal of Operations & Production Management. Vol. 25, No. 12, pp. 1228-1263.

O’Connor, M. & Webby, R. 1996. Judgemental and statistical time series forecasting: a review of the literature. International Journal of Forecasting. Vol, 12, No. 1 , pp. 91-118.

Olhager, J. Rudberg, M. & Wikner, J. 2001. Long-term capacity management: Linking the perspectives from manufacturing strategy and operations planning. International Journal of Production Economics, Vol. 69, pp. 215-225.

Stadtler, H. & Kilger, C. 2008. Supply Chain Management and Advanced Planning, Fourth Edition, Berlin, Springer. 556p.

Stevenson, W. 2007. Operations Management, Ninth Edition, New York USA, McGraw-Hill/Irwin. 903p.

Winkholfer, H. Diamantopoulos, A. & Witt, S. 1996. Forecasting Practice: A review of the Literature and an Agenda for Future Research. International Journal of Forecasting. Vol. 12, No. 2, pp. 193-221.

Winters, P. 1960. Forecasting Sales by Exponentially Weighted Averages. Management Science, Vol. 6, No. 3, pp. 324-342.

Wacker, J. & Sprague, L. 1995. The impact of institutional factors on forecast accuracy: Manufacturing executives perspective. International Journal of Productio Research, Vol. 33, No. 11, pp. 2945–2958.

OTHER REFERENCES

[1] Case company material: Notes of Demand Planning Seminar. Held on 23.-24.6.2010.

[2] Case company material: Teknos Oy. Homepage. [www.teknos.fi]. Date of access 3.4.2012.

[3] Case company material: IFS Demand Planning Software help. [http://dpprod/ifsdoc/documentation/en/Demand/CfrmDemandPlanClient.htm]. Date of access 3.4.2012.

[4] Case company material: IFS Demand Planning Software Manual. Made by Jaakko Koskenpää on 6.3.2011.

[5] Case company material: Notes of the meetings with Jaakko Koskenpää on 2.11.2011, 2.12.2011, 20.12.2011. 17.2.2012. Made by: Sasu Mäntyniemi.

APPENDICES (14 pieces)

Appendix 1: IFS Demand Planning software and its forecasting client.

Appendix 2: Mathematical formulas of different statistical models.

Appendix 3: Measures of forecast accuracy and their mathematical formulas.

Appendix 4: Demand curves (sales data in Finland) of the product segments of Make to Stock -products (beginning of the year is circulated). Data begins from 11/2009.

Appendix 5: Characteristics of the test products.

Appendix 6: Accuracies of the statistical models, AC-segment.

Appendix 7: Accuracies of the statistical models, GI-segment.

Appendix 8: Accuracies of the statistical models, PC-segment.

Appendix 9: Accuracies of the statistical models, IW-segment.

Appendix 10: Accuracies of the statistical models (MAD/av. DEMAND), Intermittent demand, all segments.

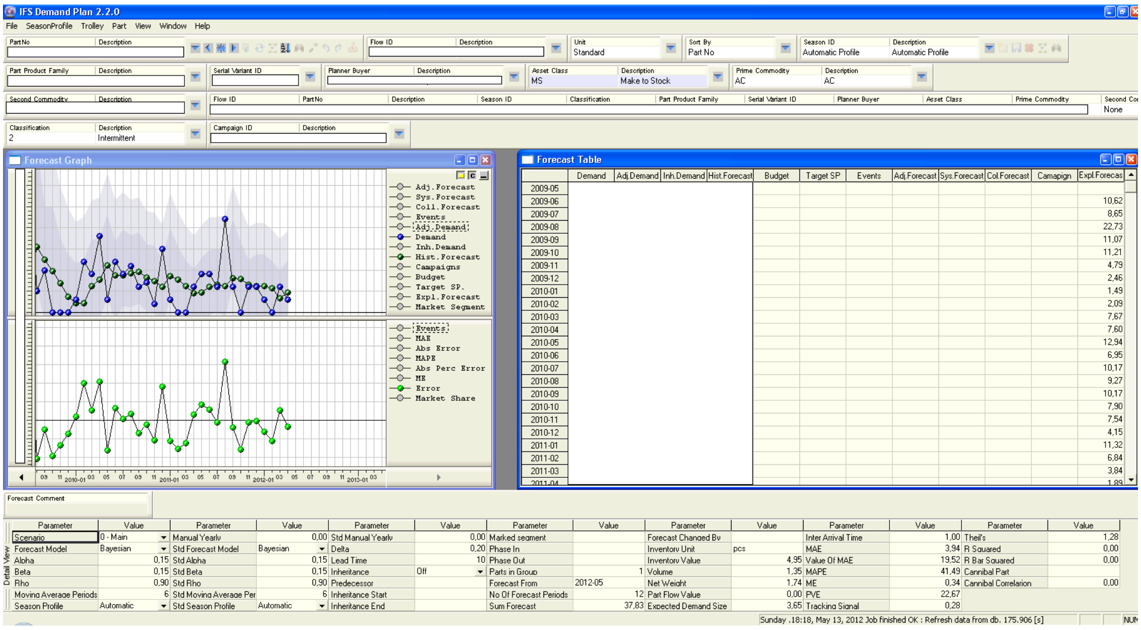
Appendix 11: Best accuracies with the use of different statistical models.

Appendix 12: Effects of the seasonal profiles on accuracy in the test of different statistical models.

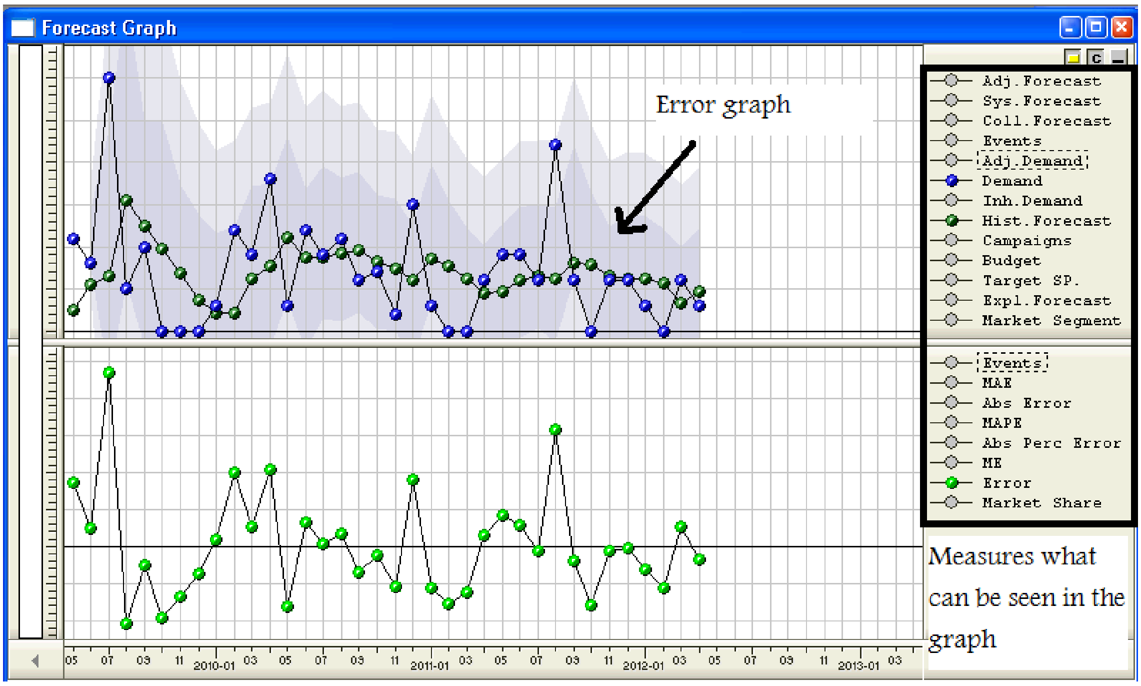
Appendix 13: Correlations between COV and MAPE and between COV and MdAPE.

Appendix 14: Forecastability of the test products based on COV value.

IFS Demand Planning software and its forecasting client.



Display of the forecast client with all of its features. Above the graph and the table are some options based on which the data can be organized, e.g. demand group, product segment, product number, different error measures and others similar.



The forecast graph (upper graph) and the control card (lower graph), in which what one wants to see can be determined by clicking the different measures on the right. When changing for example the error measure the graph in control card changes as well.

The shadow in the forecast graph is the error graph mentioned in page 54. When changing the statistical model the shadow also changes, which makes the initial assessment of forecast accuracy possible.

Forecast Table

	Demand	Adj.Demand	Inh.Demand	Hist.Forecast	Budget	Target SP	Events	Adj.Forecast	Sys.Forecast	Col.Forecast	Camapign	Expl.Forecas
2009-05	11,00	11,00		2,43								
2009-06	8,00	8,00		5,52								10,62
2009-07	30,00	30,00		6,50								8,65
2009-08	5,00	5,00		15,46								22,73
2009-09	10,00	10,00		12,49								11,07
2009-10	0,00	0,00		9,70								11,21
2009-11	0,00	0,00		6,84								4,79
2009-12	0,00	0,00		3,62								2,46
2010-01	3,00	3,00		2,13								1,49
2010-02	12,00	12,00		2,10								2,09
2010-03	9,00	9,00		6,28								7,67
2010-04	18,00	18,00		7,65								7,60
2010-05	3,00	3,00		11,13								12,94
2010-06	12,00	12,00		8,67								6,95
2010-07	9,00	9,00		8,72								10,17
2010-08	11,00	11,00		9,30								9,27
2010-09	6,00	6,00		9,51								10,17
2010-10	7,00	7,00		8,21								7,90
2010-11	2,00	2,00		7,33								7,54
2010-12	15,00	15,00		6,08								4,15
2011-01	3,00	3,00		8,54								11,32
2011-02	0,00	0,00		7,73								6,84
2011-03	0,00	0,00		6,19								3,84
2011-04	6,00	6,00		4,53								1,89

Table, in which the measures of the (upper) graph can be seen.

Expl. forecast includes the forecast of past periods (not actual ones) that the statistical model uses when making the forecast for future period.

Sys. demand shows the forecast of the statistical model for future period, whereas the adj. forecast shows the forecast that is adjusted (judgmental input) by the user.

Hist. forecast is the actual forecast of past period. It doesn't distinguish sys. and adj. forecasts afterwards.

Demand is the demand of each period; whereas adj. demand is the demand if it's adjusted somehow (e.g. projects, or unusual events)

Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value
Scenario	0 - Main	Manual Yearly	0,00	Std Manual Yearly	0,00	Marked segment	
Forecast Model	Bayesian	Std Forecast Model	Bayesian	Delta	0,20	Phase In	
Alpha	0,15	Std Alpha	0,15	Lead Time	10	Phase Out	
Beta	0,15	Std Beta	0,15	Inheritance	Off	Parts in Group	1
Rho	0,90	Std Rho	0,90	Predecessor		Forecast From	2012-05
Moving Average Periods	6	Std Moving Average Per	6	Inheritance Start		No Of Forecast Periods	12
Season Profile	Automatic	Std Season Profile	Automatic	Inheritance End		Sum Forecast	37,83

Parameter	Value	Parameter	Value	Parameter	Value
Forecast Changed By		Inter Arrival Time	1,00	Theil's	1,28
Inventory Unit	pcs	MAE	3,94	R Squared	0,00
Inventory Value	4,95	Value Of MAE	19,52	R Bar Squared	0,00
Volume	1,35	MAPE	41,49	Cannibal Part	
Net Weight	1,74	ME	0,34	Cannibal Correlarion	0,00
Part Flow Value	0,00	PVE	22,67		
Expected Demand Size	3,65	Tracking Signal	0,28		

Values that are shown below the forecast graph and the table, including the cumulative error values.

Mathematical formulas of different statistical models.

Manual:

The user sets the yearly demand which the software divides to different periods based on the predetermined seasonal profile (FI-OUTDOOR) (Case company material [3])

Naïve:

$F_{t+1} = D_t$, where

F_{t+1} = forecast for period t+1

D_t = demand in period t

Moving Average:

$F_{t+1} = \dfrac{D_{t-n} + D_{t-n+1} + ... + D_{t-2} + D_{t-1} + D_t}{n}$, where

n = number of periods

EWMA:

$F_{t+1} = \alpha D_t + (1-\alpha)F_t$, where

α = smoothing parameter between 0 and 1

EWMA with trend:

$F_{t+1} = S_{t+1} + T_{t+1}$, where

$S_{t+1} = \alpha A_t + (1-\alpha)(S_t + T_t)$ (forecast with trend of the previous period included)

$T_{t+1} = \beta(S_{t+1} - S_t) + (1-\beta)T_t$ (trend estimate for next period)

β = smoothing parameter between 0 and 1

AEWMA:

$F_{t+1} = \alpha_{t+1} D_t + (1-\alpha_{t+1}) F_t$, where

$$\alpha_{t+1} = \left| \frac{A_t}{M_t} \right|$$

$$A_t = \beta * e + (1-\beta) * A_{t-1}$$

$$M_t = \beta * |e| + (1-\beta) * M_{t-1}$$

$$e_t = D_t - F_t$$

β = smoothing parameter between 0 and 1

Brown's smoothing with trend:

$F_{t+1} = S_{t+1} + T_{t+1}$, where

$S_{t+1} = \alpha_{\mathrm{B}} D_t + (1-\alpha_{\mathrm{B}})(S_t + T_t)$ (forecast with trend of the previous period included)

$T_{t+1} = \beta_{\mathrm{B}} (S_{t+1} - S_t) + (1-\beta_{\mathrm{B}}) T_t$ (trend estimate for next period)

$$\alpha_{\mathrm{B}} = 1 - (1-\alpha\ )^2$$

$$\beta_{\mathrm{B}} = \frac{\alpha^2}{1-(1-\alpha)^2}$$

α = smoothing parameter between 0 and 1

β = smoothing parameter between 0 and 1

Least Squares:

$Y_t = L_{-1} + T_{-1}t$, where

$$L_{-1} = \overline{D} - \frac{T_{-1}(N+1)}{2}$$

$$T_{-1} = \frac{N\sum_{t=1}^{N} tD_t - \frac{N(N+1)}{2}\sum_{t=1}^{N} D_t}{\frac{N^2(N+1)(2N+1)}{6} - \frac{N^2(N+1)^2}{4}}$$

Multiple regression:

$Y_i = b_o + b_1 X_{1,i} + ... + ...b_k X_{k,1} e_1$, where

$Y_{1,}$, $X_{1,i}$, ... , $X_{k,i}$ = i:th observations of variables

b:s = fixed but unknown parameters

e = estimated error of the formula

Parameters are calculated based on least squares fitting rules.

Bayesian:

$F_{t+1} = \frac{1}{4}F1_{t+1} + \frac{1}{4}F2_{t+1} + \frac{1}{4}F3_{t+1} + \frac{1}{4}F4_{t+1}$, where

$F1$ = Moving average

$F2$ = AEWMA

$F3$ = Least squares

$F4$ = Brown’s smoothing with trend

Best Fit:

“This is a forecast model that runs every other model in competition on the last known historical demand. The model with the best result is chosen as a forecast model for this part based on which has the minimum Theil’s U.” (Case company material [3])

Croston’s Intermittent:

$Z_j = (1-\alpha)Z_{j-1} + \alpha Y^*$ and $P_j = (1-\alpha)P_{j-1} + \alpha Q^*$, where

Z = estimated demand size

P = inter-arrival time of the demand

α = smoothing parameter between 0 and 1

$Y_{n+h} = \frac{Z_l}{P_l}$, where

Formulas are adapted from Case company material [3], Buffa (1983) and Stevenson (2007)

Measures of forecast accuracy and their mathematical formulas.

Forecast error:

$e_t = D_t - F_t$, where

D_t = demand in period t

F_t = forecast in period t

Mean error:

$ME = (\frac{1}{n}) \times \sum_{t=1}^{n} e_t$, where

n = number of periods

Mean Squared error:

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

Mean Absolute Deviation:

$$MAD = (\frac{1}{n}) \times \sum_{t=1}^{n} |e_t|$$

Mean Absolute Percentage Error:

$$MAPE = \left[\frac{1}{n} \times \sum_{t=1}^{n} \left| \frac{e_t}{D_t} \right| \right] \times 100$$

Percent Variation Explained:

$PVE = \left[1 - \frac{MAPE}{MAPV} \right] \times 100$, where

$$MAPV = \left[\frac{1}{n} \times \sum_{t=1}^{n} \left| \frac{\overline{D} - D_t}{D_t} \right| \right] \times 100$$

$$\overline{D} = (\frac{1}{n}) \times \sum_{t=1}^{n} D_t$$

Theil's U-statistic:

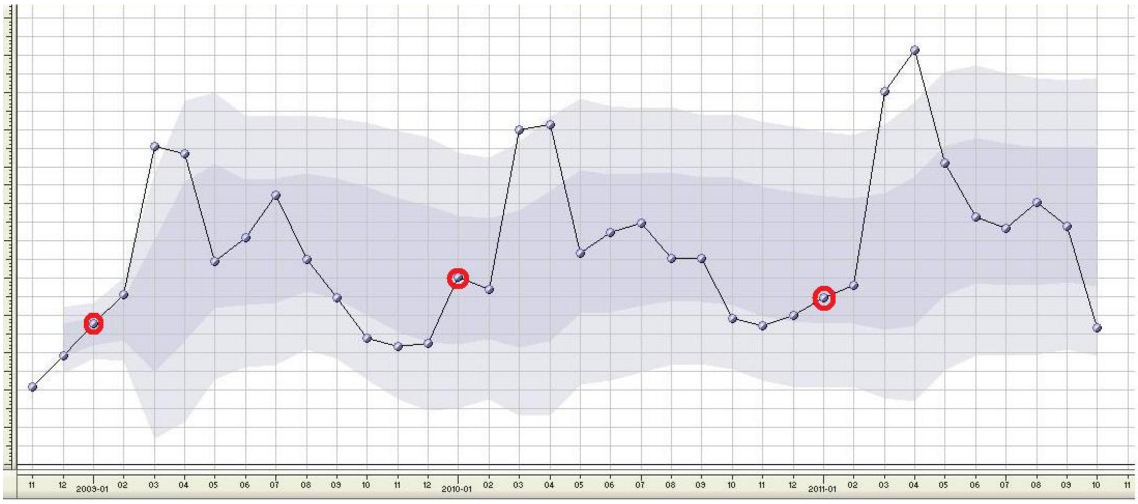
$$U = \sqrt{\frac{\sum_{t=1}^{n-1}(FPE_{t+1} - APE_{t+1})^2}{\sum_{t=1}^{n-1}(APE_{t+1})^2}}$$, where

$$FPE_{t+1} = \frac{F_{t+1} - D_t}{D_t}$$ (forecast relative change)

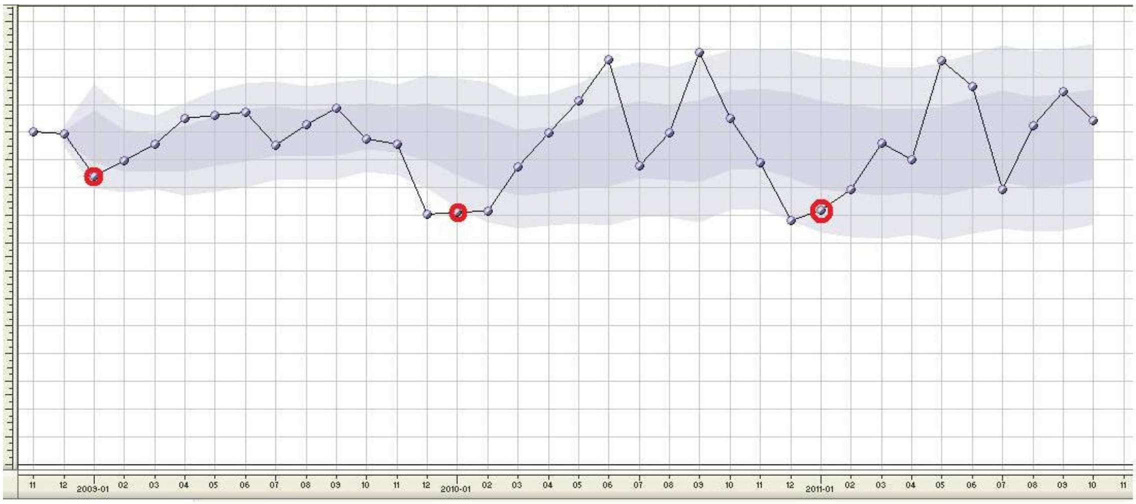
$$APE_{t+1} = \frac{D_{t+1} - D_t}{D_t}$$ (actual relative change)

Formulas are adapted from Case company material [3], Buffa (1983) and Stevenson (2007)

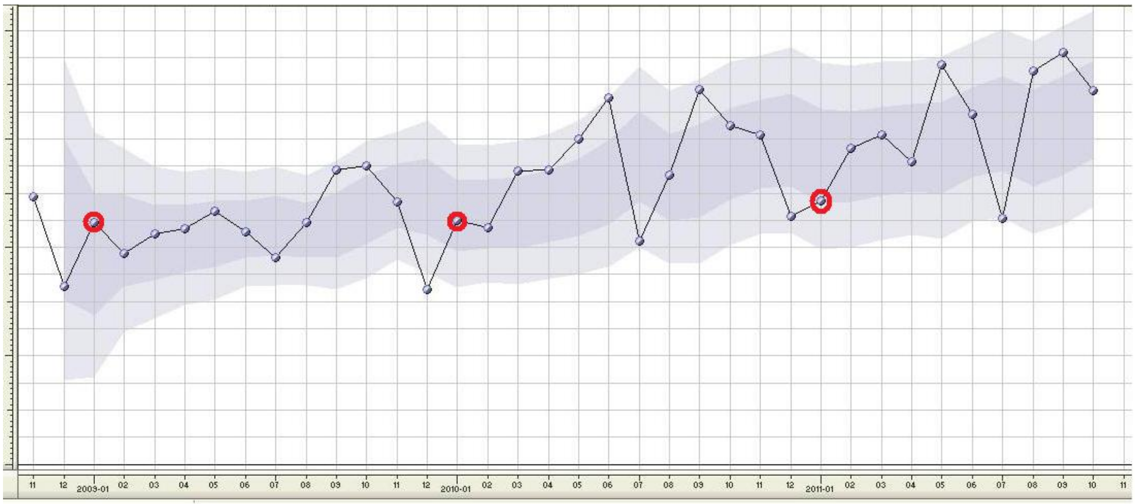
Demand curves (sales data in Finland) of the product segments of Make to Stock -products (beginning of the year is circulated). Data begins from 11/2009.



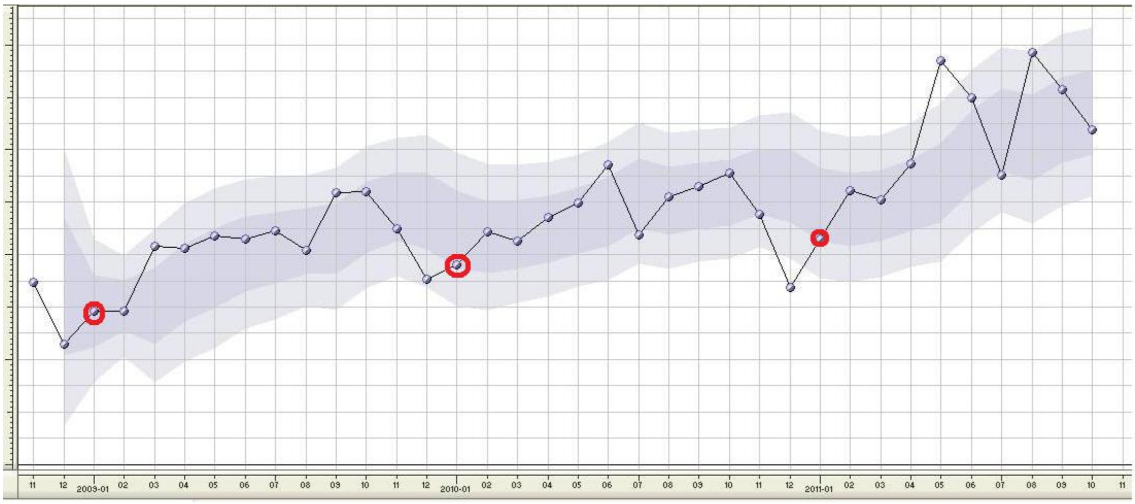
AC-products



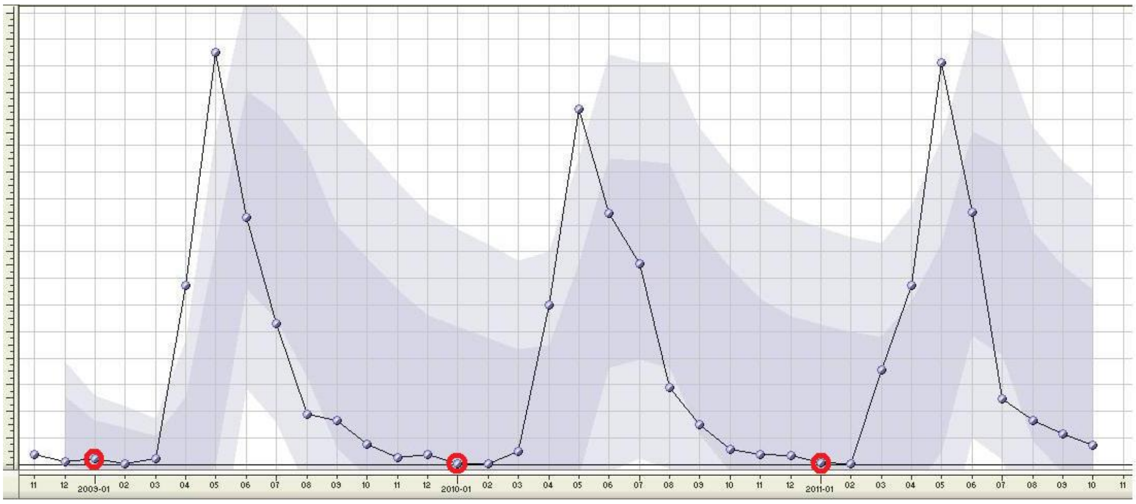
GI-products



PC-products



IW-products



IM-products

Characteristics of the test products.

	AC	***GI***	***PC***	***IW***
Intermittent 1	Mediocre occurence of zero months	Low occurence of zero months	Mediocre occurence of zero months	Low occurence of zero months
Intermittent 2	High occurence of zero months	Mediocre occurence of zero months	Low occurence of zero months	Mediocre occurence of zero months
Intermittent 3	Low occurence of zero months	High occurence of zero months	High occurence of zero months	High occurence of zero months
Level 1	Mediocre random variation	Strong random variation	Strong random variation	Strong random variation
Level 2	Strong random variation	Mediocre random variation	Mediocre random variation	Mediocre random variation
Level/Season 1	Strong seasonal variation	Strong seasonal variation	Individual peaks	Relatively strong seasonal variation
Level/Season 2	Mediocre seasonal variation	Mediocre seasonal variation		Relatively strong seasonal variation
Level/Season 3	Individual peaks			
Level/Trend(-) 1	Strong random variation	Strong random variation	Mediocre random variation	Mediocre random variation
Level/Trend(-) 2	Mediocre random variation	Mediocre random variation	Strong random variation	
Trend(-)/Season 1	Individual peaks		Regular seasonal variation	
Trend(-)/Season 2	Regular seasonal variation			
Trend(+) 1	Strong random variation	Strong random variation	Strong random variation	Mediocre random variation
Trend(+) 2	Mediocre random variation	Mediocre random variation	Mediocre random variation	Strong random variation
Trend(+)/Season 1	Regular seasonal variation	Regular seasonal variation	Regular seasonal variation	Regular seasonal variation
Trend(+)/Season 2	Irregular seasonal variation	Regular seasonal variation		Regular seasonal variation
New part 1	with FI-outdoor			
New part 2	with FI-outdoor			
New part 3	with FI-outdoor			
New part 4	with FI-outdoor			
Season 1	Mediocre variation	Strong variation	Individual peaks	Individual peaks
Season 2	Strong variation	Individual peaks	Strong variation	
Season 3	Individual peaks			

Accuracies of the statistical models, AC-segment.

AC-Products								
Product Group	***Model 1 (original)***		***Model 2 (Best Fit)***		***Model 3***		***Model 4***	
	MAPE	MdAPE	MAPE	MdAPE	MAPE	MdAPE	MAPE	MdAPE
Level	Bayesian		AEWMA		EWMA		Least squares	
1	31 %	15 %	28 %	25 %	30 %	22 %	38 %	28 %
	Bayesian		Moving average		EWMA		AEWMA	
2	71 %	46 %	86 %	52 %	70 %	43 %	62 %	39 %
Level/Season	Bayesian		Moving average		AEWMA		Least squares	
1	337 %	67 %	490 %	82 %	439 %	67 %	466 %	70 %
with seasonality	77 %	32 %	79 %	33 %	74 %	34 %	83 %	37 %
	Bayesian		Moving average		AEWMA		EWMA	
2	482 %	73 %	896 %	83 %	881 %	71 %	841 %	70 %
with seasonality	44 %	40 %	44 %	40 %	39 %	31 %	38 %	26 %
	Bayesian		EWMA		AEWMA		Moving average	
3	85 %	48 %	106 %	67 %	88 %	57 %	90 %	67 %
with seasonality	59 %	36 %	78 %	33 %	61 %	33 %	64 %	39 %
Trend(-)	Bayesian		Moving average		Brown's level and trend		EWMA level and trend	
1	59 %	35 %	68 %	39 %	62 %	40 %	60 %	37 %
	Bayesia n		Moving average		Brown's level and trend		EWMA level and trend	
2	32 %	29 %	34 %	27 %	33 %	27 %	32 %	30 %
Trend(-)/Season	Bayesian		EWMA level and trend		Moving average		AEWM A	
1	262 %	82 %	311 %	69 %	221 %	84 %	392 %	64 %
with seasonality	75 %	44 %	74 %	42 %	83 %	62 %	71 %	55 %
	Bayesian		Naive		Moving average		EWMA	
2	185 %	62 %	137 %	76 %	261 %	51 %	258 %	39 %
with seasonality	43 %	26 %	137 %	76 %	47 %	21 %	52 %	28 %
Trend(+)	Bayesian		Brown's level and trend		EWMA level and trend		Least squares	
1	105 %	42 %	105 %	26 %	106 %	24 %	85 %	55 %
	Bayesian		EWMA level and trend		Brown's level and trend		Moving average	
2	19 %	18 %	19 %	15 %	21 %	18 %	19 %	19 %

Trend(+)/Season	Bayesian		AEWMA		EWMA level and trend		Brown's level and trend	
1	26 %	20 %	32 %	25 %	30 %	24 %	30 %	23 %
with seasonality	19 %	14 %	20 %	17 %	18 %	13 %	18 %	12 %
	Bayesian		AEWMA		Brown's level and trend		Moving average	
2	134 %	52 %	124 %	63 %	155 %	51 %	152 %	50 %
with seasonality	49 %	33 %	49 %	43 %	65 %	32 %	60 %	29 %
Season	Bayesian		Moving average		EWMA		AEWMA	
1	130 %	83 %	185 %	100 %	111 %	81 %	163 %	86 %
with seasonality	46 %	38 %	63 %	55 %	39 %	33 %	38 %	35 %
	Bayesian		AEWMA		EWMA		Moving average	
2	141 %	53 %	161 %	76 %	147 %	69 %	188 %	67 %
with seasonality	106 %	57 %	96 %	55 %	103 %	54 %	104 %	58 %
	Bayesian		Moving average		AEWMA		EWMA	
3	324 %	76 %	256 %	77 %	355 %	61 %	320 %	63 %
with seasonality	129 %	47 %	128 %	53 %	120 %	52 %	122 %	54 %
New Parts	Moving average		Brown's level and trend		EWMA		Bayesian	
1	3187 %	96 %	1007 %	76 %	2764 %	275 %	691 %	116 %
FI-OUTDOOR	740 %	82 %	362 %	35 %	3187 %	96 %	366 %	108 %
	Moving average		EWMA		Bayesian		AEWMA	
2	244 %	89 %	252 %	77 %	102 %	66 %	108 %	41 %
FI-OUTDOOR	100 %	85 %	111 %	93 %	96 %	73 %	68 %	73 %
	Moving average		EWMA level and trend		Bayesian		AEWMA	
3	230 %	140 %	197 %	97 %	226 %	189 %	322 %	213 %
FI-OUTDOOR	94 %	77 %	126 %	78 %	111 %	77 %	96 %	62 %
	Moving average		Brown's level and trend		Bayesian		EWMA	
4	398 %	248 %	272 %	85 %	354 %	94 %	355 %	118 %
FI-OUTDOOR	383 %	114 %	254 %	79 %	286 %	83 %	313 %	88 %

Accuracies of the statistical models, GI-segment.

GI-Products								
Product Group	***Model 1 (original)***		***Model 2 (Best Fit)***		***Model 3***		***Model 4***	
	MAPE	MdAPE	MAPE	MdAPE	MAPE	MdAPE	MAPE	MdAPE
Level	Bayesian		Brown's level and trend		EWMA		AEWMA	
1	118 %	60 %	116 %	47 %	72 %	55 %	111 %	55 %
	Bayesian		Naive		MA		EWMA	
2	31 %	24 %	35 %	33 %	34 %	28 %	33 %	29 %
Level/Season	Bayesian		Naive		EWMA		MA	
1	358 %	78 %	259 %	94 %	476 %	62 %	460 %	85 %
with seasonality	150 %	69 %	259 %	94 %	87 %	93 %	121 %	64 %
	Bayesian		Moving average		AEWMA		EWMA	
2	192 %	60 %	174 %	60 %	107 %	64 %	160 %	63 %
with seasonality	103 %	32 %	79 %	32 %	65 %	27 %	78 %	29 %
Trend(-)	Bayesian		EWMA level and trend		Moving average		Brown's level and trend	
1	66 %	35 %	57 %	38 %	63 %	34 %	65 %	46 %
	Bayesian		Moving average		EWMA level and trend		Brown's level and trend	
2	65 %	38 %	67 %	38 %	61 %	43 %	64 %	44 %
Trend(+)	Bayesian		Brown's level and trend		EWMA level and trend		Moving average	
1	98 %	39 %	134 %	46 %	136 %	47 %	136 %	47 %
	Bayesian		EWMA level and trend		Brown's level and trend		Moving average	
2	32 %	23 %	34 %	20 %	34 %	20 %	33 %	25 %
Trend(+)/Season	Bayesian		Brown's level and trend		EWMA level and trend		Moving average	
1	54 %	33 %	53 %	28 %	59 %	29 %	52 %	28 %
with seasonality	25 %	18 %	28 %	19 %	28 %	21 %	27 %	20 %
	Bayesian		Brown's level and trend		EWMA level and trend		Moving average	
2	33 %	25 %	33 %	30 %	33 %	30 %	33 %	27 %
with seasonality	21 %	17 %	20 %	13 %	20 %	13 %	21 %	19 %

Season	Bayesian		EWMA level and trend		AEWMA		Moving average	
1	142 %	50 %	186 %	42 %	191 %	42 %	110 %	49 %
with seasonality	64 %	51 %	60 %	44 %	73 %	53 %	64 %	55 %
	Bayesian		EWMA		Moving average		AEWMA	
2	76 %	59 %	93 %	59 %	83 %	61 %	101 %	74 %
with seasonality	76 %	59 %	68 %	73 %	64 %	45 %	71 %	45 %
New Parts	Bayesian		Brown's level and trend		EWMA		Moving average	
1	56 %	59 %	61 %	56 %	52 %	61 %	58 %	63 %
	Bayesian		Brown's level and trend		EWMA		Moving average	
2	25 %	26 %	20 %	13 %	27 %	22 %	35 %	20 %

Accuracies of the statistical models, PC-segment.

PC-Products								
Product Group	***Model 1 (original)***		***Model 2 (Best Fit)***		***Model 3***		***Model 4***	
	MAPE	MdAPE	MAPE	MdAPE	MAPE	MdAPE	MAPE	MdAPE
Level	Bayesian		EWMA		AEWMA		Least squares	
1	195 %	80 %	219 %	96 %	230 %	85 %	160 %	74 %
	Bayesian		AEWMA		EWMA		Moving average	
2	59 %	26 %	58 %	20 %	56 %	24 %	55 %	24 %
Level/Season	Bayesian		Moving average		AEWMA		EWMA	
1	99 %	63 %	96 %	67 %	120 %	67 %	108 %	68 %
with seasonality	76 %	44 %	71 %	49 %	82 %	44 %	78 %	43 %
Trend(-)	Bayesian		Brown's level and trend		EWMA level and trend		Least squares	
1	177 %	42 %	186 %	43 %	190 %	40 %	155 %	47 %
	Bayesian		EWMA level and trend		Brown's level and trend		Moving average	
2	132 %	47 %	141 %	44 %	113 %	35 %	121 %	44 %
Trend(-)/Season	Bayesian		Naive		AEWMA		EWMA	
1	110 %	57 %	97 %	53 %	135 %	48 %	139 %	53 %
with seasonality	50 %	39 %	97 %	53 %	45 %	36 %	68 %	42 %
Trend(+)	Bayesian		Naive		Brown's level and trend		EWMA level and trend	
1	147 %	54 %	188 %	58 %	160 %	46 %	149 %	60 %
	Bayesian		Naive		Brown's level and trend		EWMA level and trend	
2	50 %	42 %	49 %	39 %	51 %	40 %	54 %	42 %
Trend(+)/Season	Bayesian		EWMA		Brown's level and trend		EWMA level and trend	
1	126 %	37 %	114 %	41 %	142 %	38 %	136 %	43 %
with seasonality	82 %	28 %	76 %	43 %	85 %	23 %	85 %	25 %

Season	Bayesian		Moving average		AEWMA		EWMA	
1	278 %	86 %	216 %	100 %	206 %	83 %	269 %	89 %
with seasonality	371 %	61 %	371 %	62 %	382 %	62 %	346 %	59 %
	Bayesian		Brown's level and trend		Least squares		AEWMA	
2	89 %	51 %	93 %	51 %	107 %	68 %	97 %	57 %
with seasonality	51 %	26 %	47 %	29 %	68 %	48 %	69 %	35 %
New Parts	Bayesian		Brown's level and trend		AEWMA		Moving average	
1	48 %	42 %	48 %	39 %	45 %	34 %	46 %	45 %
	Bayesian		Brown's level and trend		AEWMA		Moving average	
2	46 %	41 %	43 %	31 %	47 %	43 %	55 %	31 %

Accuracies of the statistical models, IW-segment.

IW-Products								
Product Group	***Model 1 (original)***		***Model 2 (Best Fit)***		***Model 3***		***Model 4***	
	MAPE	MdAPE	MAPE	MdAPE	MAPE	MdAPE	MAPE	MdAPE
Level	Bayesian		Moving average		AEWMA		EWMA	
1	485 %	136 %	444 %	95 %	372 %	80 %	418 %	114 %
	Bayesian		Naive		Least squares		AEWMA	
2	55 %	15 %	46 %	13 %	74 %	18 %	67 %	19 %
Level/Season	Bayesian		Moving average		AEWMA		EWMA	
1	318 %	58 %	680 %	73 %	456 %	49 %	507 %	62 %
with seasonality	99 %	38 %	70 %	49 %	128 %	34 %	112 %	34 %
	Bayesian		Moving average		Least squares		AEWMA	
2	96 %	40 %	81 %	44 %	102 %	45 %	99 %	44 %
with seasonality	36 %	21 %	40 %	33 %	42 %	32 %	37 %	24 %
Trend(-)	Bayesian		EWMA level trend		Brown's level and trend		Moving average	
1	52 %	36 %	49 %	39 %	52 %	43 %	56 %	32 %
Trend(+)	Bayesian		Moving average		EWMA level and trend		Brown's level and trend	
1	69 %	35 %	80 %	46 %	71 %	36 %	74 %	37 %
	Bayesian		Moving average		EWMA level and trend		Brown's level and trend	
2	527 %	54 %	533 %	57 %	561 %	45 %	588 %	49 %
Trend(+)/Season	Bayesian		EWMA		EWMA level and trend		Brown's level and trend	
1	90 %	40 %	71 %	39 %	94 %	39 %	95 %	40 %
with seasonality	41 %	25 %	51 %	49 %	40 %	16 %	41 %	15 %
	Bayesian		EWMA		EWMA level and trend		Brown's level and trend	
2	193 %	75 %	205 %	48 %	282 %	59 %	253 %	50 %
with seasonality	94 %	40 %	125 %	73 %	154 %	39 %	150 %	38 %
Season	Bayesian		EWMA level trend		Moving average		Least squares	
1	36 %	34 %	39 %	39 %	45 %	43 %	50 %	53 %
with seasonality	36 %	34 %	65 %	50 %	63 %	53 %	61 %	42 %
New Parts	Bayesian		AEWMA		EWMA		Moving average	
1	39 %	28 %	99 %	54 %	78 %	92 %	63 %	75 %
	Bayesian		AEWMA		EWMA		Moving average	
2	54 %	68 %	43 %	49 %	45 %	51 %	63 %	50 %

Accuracies of the statistical models (MAD/av. DEMAND), Intermittent demand, all segments.

Product Group	***Model 1 (original)***	***Model 2 (Best Fit)***	***Model 3***	***Model 4***
AC	Bayesian	Naive	Croston's intermittent	Moving average
Intermittent 1	110 %	139 %	116 %	122 %
	Bayesian	EWMA	Croston's intermittent	Naive
Intermittent2	127 %	103 %	107 %	166 %
	Bayesian	Moving average	Croston's intermittent	Naive
Intermittent 3	93 %	96 %	94 %	117 %
GI	Bayesian	Moving average	Naive	Moving average
Intermittent 1	51 %	57 %	68 %	59 %
	Bayesian	Naive	Moving average	Croston's intermittent
Intermittent2	125 %	130 %	129 %	127 %
	Bayesian	EWMA	Naive	Croston's intermittent
Intermittent 3	150 %	133 %	183 %	158 %
PC	Bayesian	Moving average	Naive	Croston's intermittent
Intermittent 1	100 %	105 %	101 %	99 %
	Bayesian	EWMA level and trend	Moving average	Croston's intermittent
Intermittent2	72 %	66 %	71 %	76 %
	Bayesian	Moving average	Naive	Croston's intermittent
Intermittent 3	120 %	120 %	117 %	117 %
IW	Bayesian	Moving average	Naive	Croston's intermittent
Intermittent 1	81 %	89 %	96 %	88 %
	Bayesian	Moving average	Naive	Croston's intermittent
Intermittent2	90 %	80 %	130 %	80 %
	Bayesian	Moving average	Naive	Croston's intermittent
Intermittent 3	138 %	126 %	174 %	105 %

Best accuracies with the use of different statistical models.

	AC		***GI***		***PC***		***IW***	
	MAPE	MdAPE	MAPE	MdAPE	MAPE	MdAPE	MAPE	MdAPE
Level 1	31 %	15 %	72 %	55 %	160 %	74 %	372 %	80 %
Level 2	62 %	39 %	31 %	24 %	55 %	24 %	46 %	13 %
Level/Season 1 (aut. season)	74 %	34 %	87 %	93 %	71 %	49 %	70 %	49 %
Level/Season 2 (aut. season)	38 %	26 %	65 %	27 %			36 %	21 %
Level/Season 3 (aut. season)	59 %	36 %						
Level/Trend(-) 1	59 %	35 %	57 %	38 %	155 %	47 %	49 %	39 %
Level/Trend(-) 2	32 %	29 %	61 %	43 %	113 %	35 %		
Trend(-)/Season 1 (aut. season)	75 %	44 %			45 %	36 %		
Trend(-)/Season 2 (aut. season)	43 %	26 %						
Trend(+) 1	85 %	55 %	98 %	39 %	147 %	54 %	69 %	35 %
Trend(+) 2	19 %	15 %	32 %	23 %	50 %	39 %	527 %	54 %
Trend(+)/Season 1 (aut. season)	18 %	13 %	25 %	18 %	82 %	28 %	40 %	16 %
Trend(+)/Season 2 (aut. season)	49 %	33 %	20 %	13 %			94 %	40 %
New part 1	362 %	35 %	52 %	61 %	45 %	34 %	39 %	28 %
New part 2	68 %	73 %	20 %	13 %	43 %	31 %	43 %	49 %
New part 3	96 %	62 %						
New part 4	254 %	79 %						
Season 1 (aut. Season)	38 %	35 %	60 %	44 %	206 %	83 %	36 %	34 %
Season 2 (aut. Season)	103 %	54 %	64 %	45 %	47 %	29 %		
Season 3 (aut. Season)	120 %	52 %						

Effects of the seasonal profiles on accuracy in the test of different statistical models.

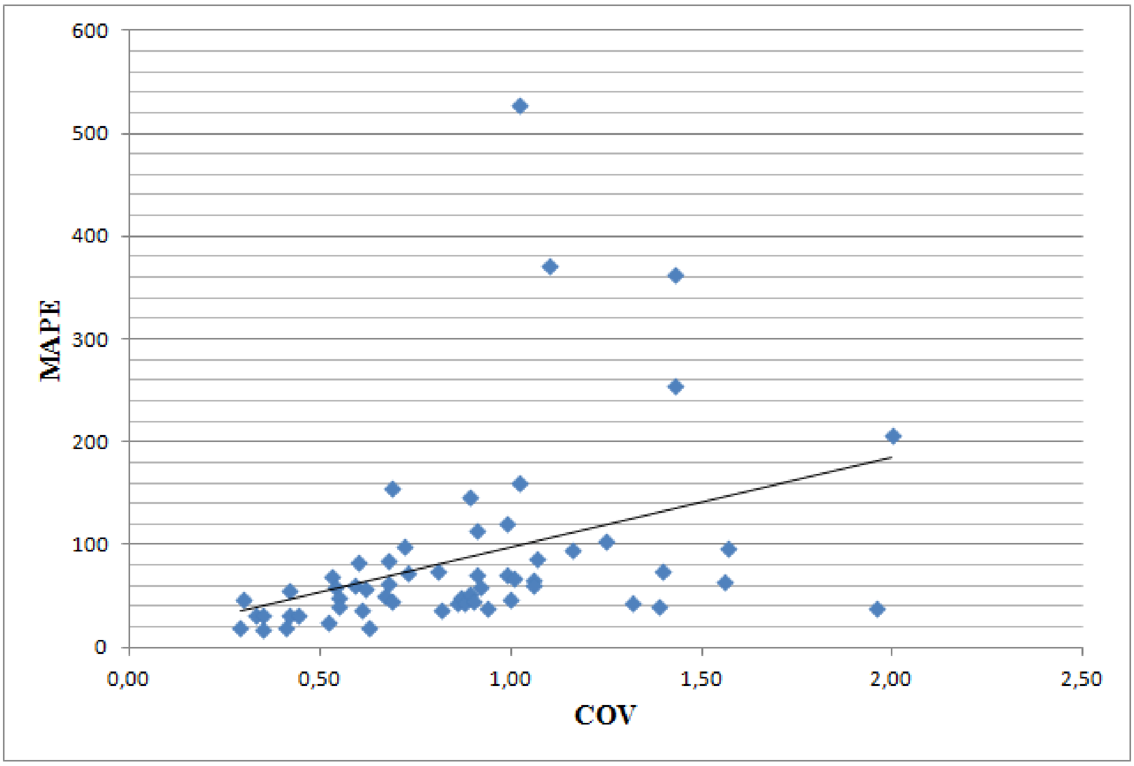
	Average without automatic seasonal profile		***Average with automatic seasonal profile***	
	MAPE	MdAPE	MAPE	MdAPE
AC	199 %	63 %	62 %	62 %
GI	106 %	54 %	53 %	40 %
PC	120 %	59 %	118 %	40 %
IW	140 %	50 %	55 %	32 %
average in total	141 %	57 %	72 %	44 %

Automatic seasonal profile

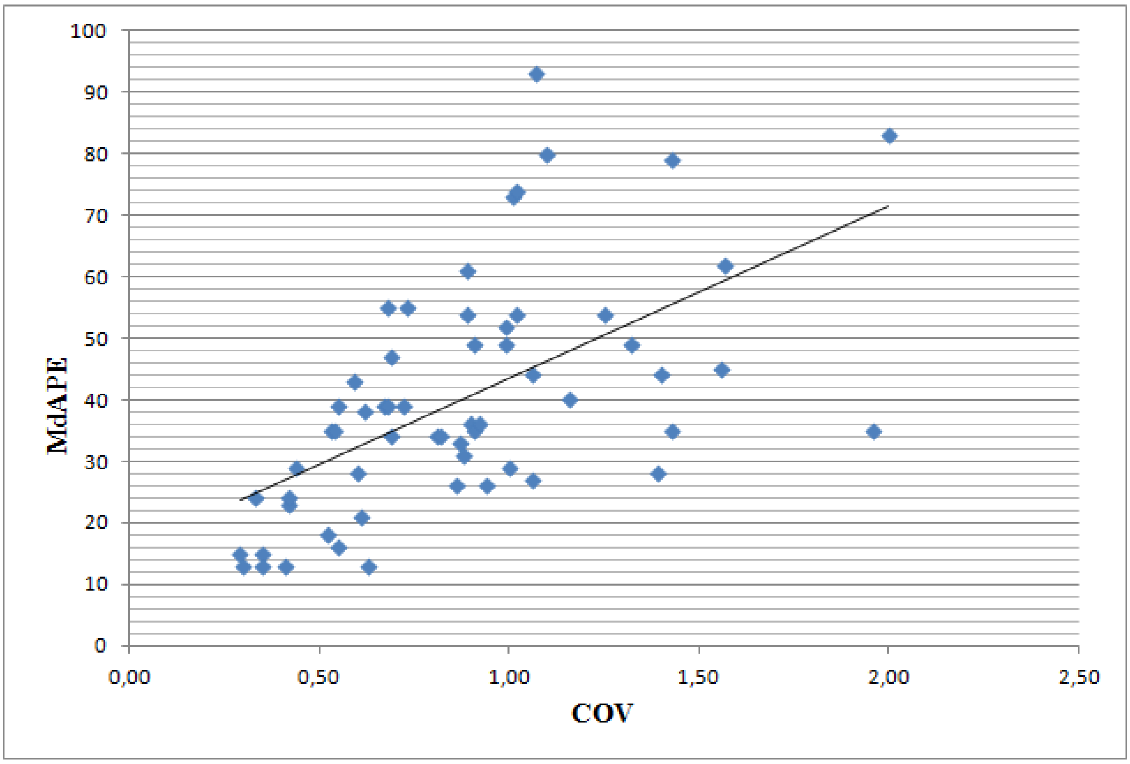
AC-PRODUCTS	***Accuracy without FI-OUTDOOR***		***Accuracy with FI-OUTDOOR***	
	MAPE	MdAPE	MAPE	MdAPE
Product 1	691 %	116 %	362 %	35 %
Product 2	102 %	66 %	68 %	73 %
Product 3	197 %	97 %	96 %	62 %
Product 4	272 %	85 %	254 %	79 %
average in total	316 %	91 %	195 %	62 %

FI-OUTDOOR seasonal profile

Correlations between COV and MAPE and between COV and MdAPE.



Correlation coefficient: 0,38.



Correlation coefficient: 0,68.

Forecastability of the test products based on COV value.

	AC	***GI***	***PC***	***IW***
Intermittent 1	1,31	0,76	1,41	0,89
Intermittent 2	1,85	1,69	0,87	0,87
Intermittent 3	1,01	2,19	1,81	1,70
Level 1	0,35	0,73	1,02	1,10
Level 2	0,68	0,33	0,42	0,30
Level/Season 1	0,81	1,07	0,99	0,91
Level/Season 2	0,94	1,06		0,61
Level/Season 3	0,92			
Level/Trend(-) 1	0,54	0,62	0,69	0,55
Level/Trend(-) 2	0,44	0,59	0,91	
Level/Trend(-) /Season 1	1,40		0,90	
Level/Trend(-) /Season 2	0,86			
Level/Trend(+) 1	0,68	0,72	0,89	0,53
Level/Trend(+) 2	0,29	0,42	0,67	1,02
Level/Trend(+) /Season 1	0,35	0,52	0,60	0,55
Level/Trend(+) /Season 2	0,87	0,41		1,16
New Parts 1	1,43	0,89	0,69	1,39
New Parts 2	1,01	0,63	0,88	1,32
New Parts 3	1,57			
New Parts 4	1,43			
Season 1	1,96	1,06	2,00	0,82
Season 2	1,25	1,56	1,00	
Season 3	0,99			

If COV is more than 1,40 item is not forecastable (marked with red colouring)