Lappeenranta University of Technology

School of Business and Management

Global Management of Innovation and Technology

Valentina Smelova

ACCURACY IMPROVEMENT OF A DEMAND FORECASTING MODEL

Examiners: Professor Leonid Chechurin

Associate Professor Ville Ojanen

Supervisor: Ville Ojanen

ABSTRACT

Lappeenranta University of Technology
School of Business and Management
Global Management of Innovation and Technology

Valentina Smelova

Accuracy improvement of a demand forecasting model
Master Thesis

93 pages, 12 figures, 8 tables, 4 appendices

Examiners: Professor Leonid Chechurin
Associate Professor Ville Ojanen

Keywords: Demand forecasting, forecasting model, inventory management, methods of forecasting, qualitative expert methods, time series analysis, moving average, error correction model, chain indices of seasonality, seasonal decomposition, forecasting error, forecasting accuracy measurement, seasonality, trend.

This research concerns different statistical methods that assist to increase the demand forecasting accuracy of company X’s forecasting model.

Current forecasting process was analyzed in details. As a result, graphical scheme of logical algorithm was developed. Based on the analysis of the algorithm and forecasting errors, all the potential directions for model future improvements in context of its accuracy were gathered into the complete list.

Three improvement directions were chosen for further practical research, on their basis, three test models were created and verified.

Novelty of this work lies in the methodological approach of the original analysis of the model, which identified its critical points, as well as the uniqueness of the developed test models. Results of the study formed the basis of the grant of the Government of St. Petersburg.

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor Professor Ville Ojanen for the opportunity to carry out master thesis work under his supportive guidance.

My supervisor from Saint Petersburg State Polytechnic University Professor Sergey Redko deserves special thanks for being a mentor to me from the very beginning of my research activities. It was he who introduced me to the world of mathematical modeling.

The reviewers have given insightful comments, which I have tried my best to take into account during the final stages of the process.

I would like to thank my employer supervisors Tomasz Blinski, Kejio Mäkela, Margus Kutsar and Cristian Kalve for giving me incredibly interesting research topic based on the real company performance, for their help, kind attention, friendly advice and support. It was a pleasure to deal with them during the entire work.

I also want to thank the Department of Industrial Management in LUT because it gave me an opportunity to study in modern, progressive European university and multiply my knowledge.

Finally, I would like to express my gratitude to my family and friends who have been there for me during this long working process.

Lappeenranta, January, 2015

Valentina Smelova

TABLE OF CONTENT

LIST OF SYMBOLS AND ABBREVIATIONS

ERP – Enterprise Resource Planning

SAMPO – Annual Regional Forecast

SBU – Strategic Business Unit

ARIMA – Autoregressive integrated moving average model

ANNs – Artificial Neural Networks

MAD – Mean Absolute Deviation

MSE – Mean Square Error

MAPE – Mean Average Percentage Error

MPE – Mean Percentage Error

sMAPE – Symmetric Mean Average Percentage Error

Median SAPE – Median Symmetric Absolute Percentage Error

Median RAE – Median Relative Absolute Error

PrIt – End Product hierarchy level

PrGr – Product Group hierarchy level

Pr2Gr – Product Secondary Group hierarchy level

PrLi – Product Line hierarchy level

SBU – Strategic Business Unit hierarchy level

Flat – Forecasting by Näive method

1 INTRODUCTION

Company X – is a family business, international company founded in Finland in 1958 specializing in the development, manufacturing and marketing of electrical systems and supplies for the distribution of electrical power as well as electrical applications. The core values of the company are long-term trusting relationships with customers and principles of environmental protection. Main customers are manufacturers of electrical and electronic products, distributors, wholesalers, construction and design organizations.

Operating in the energy field X creates innovative solutions that can significantly reduce energy costs. Company is separated into three divisions:

1. X Utility Networks (reliable and low loss power distribution solutions in the core of green thinking)
2. X Industrial Solutions (enclosing systems and industrial components for demanding environments)
3. X Building Technology (unique expertise and solutions in the field of integrated building technology)

X produces over 15 thousands different products. Main competitive advantage is high-qualitative products supply within the shortest lead time. This task can be completed when all products that have to be supplied are always in stock. In this case volumes of products must strictly comply with the current consumer demand, eliminating wasteful expenditures. This complicated management task can be solved through a system of effective inventory management.

1.1 Topicality

Previously, in order to assess the level of future demand objectively, the company directly asked their customers to send approximate monthly forecast. Despite its apparent simplicity, this approach has not been successfully implemented. In order to forecast the future demand, headquarter decided to take a look at the history of sales that has been continuously accumulated in the warehouse database and try to use this valuable information.

Statistical analysis of the data identified the recurring seasonal nature of the sales dynamic of products and product groups. After the conducted analysis, first version of the forecasting model was developed in MS Excel, model could forecast the demand for the next season with the certain accuracy rate, based on the previous two years of sales history. Currently, this model is exploited in the processes of production planning and procurement.

Comparative analysis of actual and forecasted sales data showed that forecasting model is not perfect. There are a number of problems which cannot be solved by this model. The author of the present study was requested to improve the accuracy of the current model. Wherein, all the appropriate approaches, methods and research scales were chosen independently. Thus, this study is done by the applicant individually.

1.2 Problem statement and research methods

The goal – is to develop an improved version of the forecasting model, which has a higher accuracy rate. Known statistical forecasting methods are main objects of this study. To achieve this goal it is necessary to solve the following tasks:

- Analyze the current process of demand forecasting.
- Make comparisons of actual sales data and forecast data of previous years.
- Explore all possible directions for future refinements of the model.
- Analyze known statistical methods for demand forecasting.
- Develop improved test versions of the forecasting model based on the selected statistical methods and conduct their approbation on the basis of available historical sales data.
- Analyze the possibility of practical application of an improved model of demand forecast in companies from other industries.

Basically, by the end of the study the weaknesses of the current model should be defined and then some practical solutions should be proposed and tested. The main question that has to get an answer is: how to improve the accuracy of the forecasting model of company X?

The research method used in this study is case study research. A case study is an empirical enquiry that investigates a phenomenon within its real life context where the boundaries of between the phenomenon and its context are not evident. An advantage of the case study method is that it allows the researcher to test theories in a real-life situation, providing deep insight into how the theory holds up in practice. The case study strategy helps to gain a rich understanding of the context of the research and the processes being enacted (Saunders, 2009).

In this study, the research method was also influenced by the ability of the author to take part in redesigning a demand forecasting process in company X. During the time of the writing, the author was hired by the company being a part of the project team that was tasked to study the forecasting model and the whole entire forecasting process implemented in the case company and to improve its accuracy. The project started in March 2013 and it keeps on going. This research project has some limitations. The basic idea, the forecast approach the model is based on should remain unchanged. Model could be clarified or supplemented in various ways. Also available forecasted and actual sales data used for the analysis was limited by the time period since May 2012 till April 2014.

1.3 Outline of the study

Thesis consists of eight chapters, including an introductory part. The second chapter provides a detailed description of the forecasting model, goals of the study, the basic idea of the forecasting method implemented in the current model, its main disadvantages. Also logic of the forecasting algorithm is presented there by the graphic scheme. Some of the known theoretical methods of demand forecasting, alternative approaches of the time series forecasting are presented in the third chapter. It contains a description of data sets required for demand modeling and various measures of the forecast accuracy also. In addition, results of the well-known experiment M3-competition of the time series models and important conclusions are presented in this chapter.

The fourth chapter provides an analysis of forecast errors and possible causes of the deviations of the forecasts from actual sales. The main outcome of the fourth chapter is the list of possible direction for the future refinements of the current model.

The fifth chapter describes three test models that were developed in the practical part of this study, based on the list of directions set out to improve the accuracy of current model forecast. All three models were tested on the past sales data. Accuracy of the forecasts compiled by the model versions was compared with the forecast accuracy of the original model.

Next sixth chapter provides a description and comparison of computation features in matrix-based software and Excel, presents a practical solution to implement the current forecasting algorithm in a matrix format; in the end it gives an assessment to the feasibility of such model.

Seventh chapter summarizes the obtained results and describes the possibilities of their practical application for companies from different industries.

All the results that have been achieved in this work and possible future steps for continuation of the research are described in the conclusion chapter of this study.

2 CURRENT FORECASTING MODEL ANALYSIS

This chapter gives the detailed description of the forecasting model that company X uses currently for production planning and procurement purposes. Here outlines the main objectives pursued by company in the implementation of forecasting activities, as well as a basic idea of the forecasting algorithm and the validity of chosen forecasting method. Sequence of the calculations was systematized and represented graphically through a step by step logical algorithm. Main disadvantages of the current model are highlighted in the end of this section. Mainly these disadvantages became the causes and triggers for this study.

2.1 Current forecasting process

Demand forecast figures that are generated by the original model are taking in consideration in the production planning of finished products and semi-finished products, procurement of raw materials on four different factories of company X based in Finland and Estonia. Demand forecasting process is based on prediction of the product volumes that will be sold during the next 3-month period. Company runs the model on monthly basis. Thereby, predicting seasonal changes in sales, forecasting model improves the efficiency of inventory management, reduces costs, and reduces the total time of product delivery and levels of safety stocks.

Forecast of the future sales volumes is built on the transaction data that is accumulated in two information systems ("IFS" и "Logisticar") within the previous three years, and also the results of the long-term regional annual forecast SAMPO. Figure 2.1 shows the scheme of interaction between two information systems mentioned above. ERP-system IFS is the main one, it plays the role of "muscles" in the logistic process. IFS database contains the transaction history about all the movements of goods from suppliers to end customers and back in case of a reverse supply.

On the other hand system Logisticar is a "brain" of the logistic process. This system is considered to be more robust in terms of ABC-categorization of products and other planning functions. In addition Logisticar does not interfere with the daily processes. It also stores the Life Cycle Code (LCC) information for all products. Changing one parameter in the system automatically updates all the required fields in the IFS system.

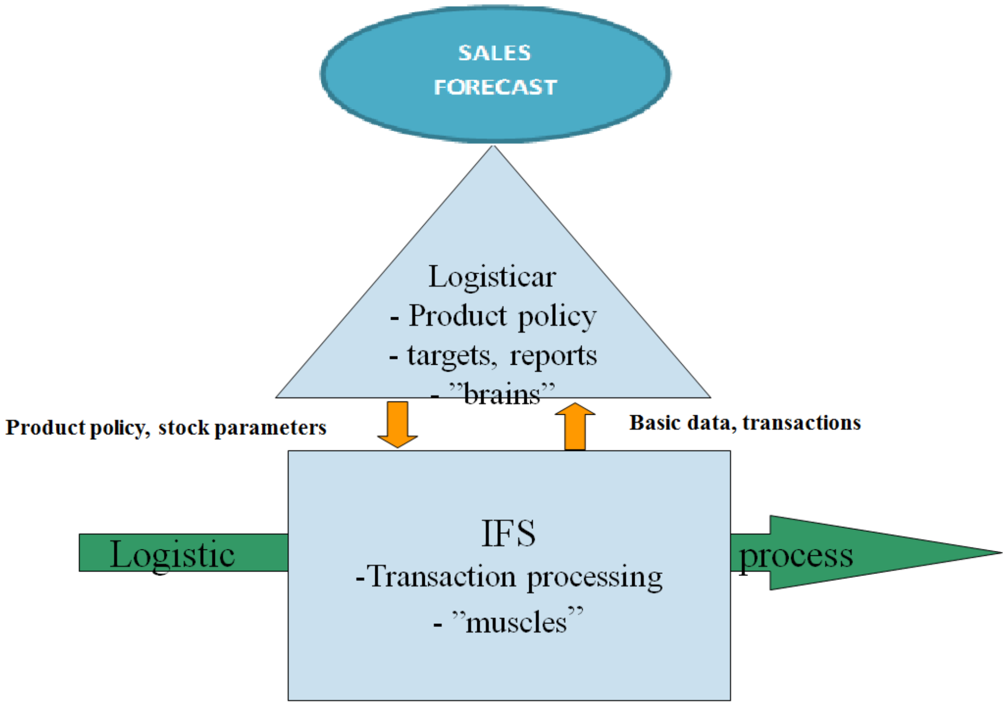


Figure 2.1. Interaction scheme for information systems and forecasting model

Statistical forecasting model is used for those products that are involved in external sales and meet the following requirements:

1. Product is not new, it is presented on the market more than one year;
2. Product LCC is defined;
3. Product is mature (stable ramp-up or stable ramp-down phases, Appendix 1);
4. Frequency of sales during the last year is over three months;

Further, Figure 2.2 presents three general steps of the forecasting process. On the first step sales departments in every country where X's products are presented on the market, update the annual forecast figures in SAMPO on the quarterly basis. In parallel, the logistic department provides support and updates the database, which contains information about the life cycle of products. This information influences the final results of the forecast. Directly, calculation stage of the forecasting process is divided into two parts. First, model calculates the forecast for all finished-products. Results are compared with the figures of the long-term annual forecast SAMPO for the product groups with the following manual corrections if necessary. Secondly, forecasted figures for the finished-products are exploded till the components level: materials, semi-finished products and finished products. After passing through these two stages calculation results can be corrected manually, if necessary based on the expert opinions. Finally, revised forecast results are loaded to the ERP-systems, where further systems implement an automatic capacity and material planning for machines and assemble cells, and other logistic parameters like reorder points, safety stock levels and batch sizes.

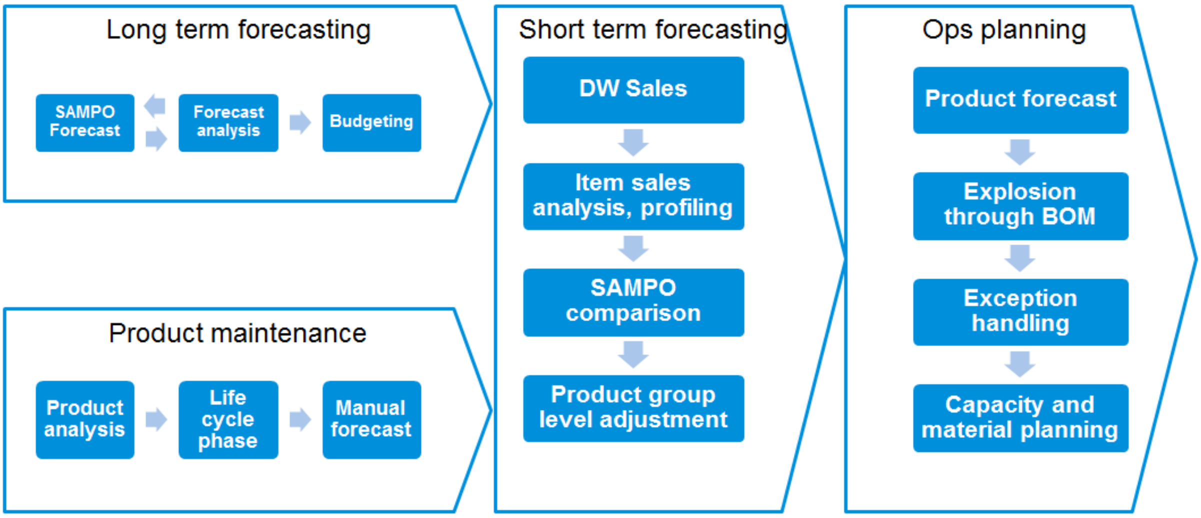


Figure 2.2. The general scheme of the forecasting process

2.2 Forecasting model

In order to describe the logic of current forecasting algorithm, it is necessary to take a look at the product hierarchy structure used in company X (Figure 2.3). Finished products range consists of approximately 15 000 items. Every product is related to one out of 250 product groups. At the same time product groups are united in 75 different secondary product groups. Secondary product groups form 25 product lines. On the top level of the hierarchy 25 lines are united into 3 strategic business units (SBU). Thus every single product is characterized by the certain set of 5 hierarchy levels. For instance, product X2ABP406020 belongs to the group 5907, secondary group 5MCX, and product line 5EC, which in turn refers to BUILDING TECHNOLOGY strategic business unit.

Current forecasting model was created by a group of developers for short-term demand forecasting purposes. It is based on the individual sales profile analysis for every finished product separately and identification of the seasonality in oscillations of its sales. However seasonal patterns are clearly seen not for all finished products. It is very important to mention, that seasonal behavior can be possibly identified on every hierarchy level.

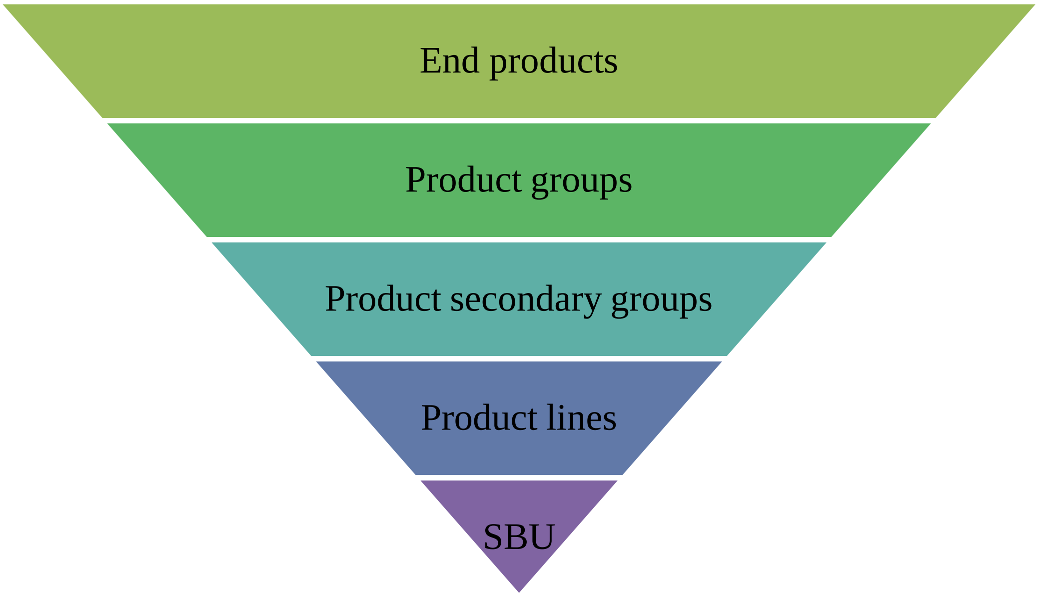


Figure 2.3. Product hierarchy structure

The literature describes a wide variety of statistical methods for the analysis of seasonal sales. Some of them will be presented in the next chapter. In turn, original forecasting model of company X helps to track and analyze seasonal behavior of the sales dynamic through individual seasonal chain indices. This mathematical approach in general successfully meets forecasting goals set by the company X management.

After loading the required data into MS Access Database it goes through a list of "cleaning" queries. These queries help to correct all the occurred mistakes and form the complete table layout for running the calculations. However, sometimes mistakes can happen while typing in information into the ERP databases. That is why all the following manual corrections of the wrong site of production and other issues still happen. As soon as all tables have been presented in the convenient layout and format for further work and all necessary information has been loaded into Excel-worksheets, it becomes possible to proceed with appropriate calculations of the short-term forecast figures.

Excel Pivot tables help to choose the correct data array for further calculations. Pivot table is a table, created in Excel, linked with the main table through window PowerPivot. This add-in in Excel is designed for efficient data analysis and sophisticated data models. Power Pivot helps to mash up large volumes of data from various sources, perform information analysis rapidly, and share insights easily, (Microsoft 2014). Each level of product hierarchy corresponds to a separate Excel worksheet. Individual seasonal indices K are calculated there in the worksheets for every month. Index *K* is calculated according to the formula (2.1).

$$Ki = \frac{Q3i}{Q4i}, \tag{2.1}$$

where *Q3i* – total number of products sold over the next three months (starting with the first day of the month *i*), *Q4i* – total number of products sold over the previous 4 months (starting with the first day of the month *i*, for which the calculation is performed).

For example, for particular product 1017300 index $K_{2012.02}$ for February 2012 is calculated as follows: number of product *1017300*, sold over the period February 2012 – April 2012 divided by number of the product *1017300* sold over the period October 2011 – January 2012. When calculating the *K* indices for the hierarchical level of the end products production volume *Q* is expressed in pieces. However, on higher hierarchy levels Q is expressed in monetary terms EUR. In case of any errors occurred during the calculation stage all indices of seasonality are set to 0.75 that basically means unchangeable demand in the next period. Coefficient 0.75 does not affect the final results. As a result of this stage every end product, as well as all higher levels of the hierarchy correspond to one-dimensional arrays, which consist entirely of individual indices of seasonality.

The next stage performed in Excel is related to correlation coefficients calculations for the pairs of annual time series, taken over the past two years, with respect to the month for which the forecast is calculated. Thereby five different correlation coefficients are calculated for every single end product on five hierarchical levels end product belongs to. Correlation coefficient is calculated for the time series of end products sales and aggregate levels time series end product belongs to. Correlation value can be insignificant on the end product level, but at the same time an appropriate product group can have a clear seasonal behavior of aggregated sales and vice versa. The most accurate short-term forecast can be made if end product is characterized by a high level of correlation between the time series of seasonal indices, indicating a clear seasonal nature of its sales.

On model development stage special boundary limit values for necessary and sufficient seasonal indices and correlation coefficients were calculated through experimental methods (Table 2.1).

Table 2.1 Sufficient boundary limits for seasonal indices and correlation coefficients

Correlation coefficient	Maximum seasonality index in the time series				
	Product	Group	Group2	Line	SBU
97 %	9	9	6	4	4
90 %	7	7	4,5	4	4
80 %	3,7	4	3	3	3
70 %	1,2	2	2	2	2
50 %	1	1,8	1,5	1,5	1,5
0% - 50%					

Calculated boundary limits set the restrictive conditions for maximum individual seasonal index in the time series and lower bounds for the correlation coefficient. Thus, each end product has five pairs of certain numbers: maximum index of seasonality in the time series and correlation coefficient on five hierarchy levels. Hierarchy level where seasonal index and correlation coefficient pair best meets the boundary limits of the Table 2.1, is considered to be a determinant level. Determinant level shows that significant seasonal behavior was identified on this particular level, and the entire future forecast will be based on its individual indices of seasonality.

If none of the pairs of coefficients for a certain end product does not fit the restrictive limits, then the forecast will be fully equal to the volume of sales in the previous period. The final value of the future individual seasonality index is calculated as the arithmetic mean between the indices of the corresponding months of the previous biennium.

For instance, Table 2.2 shows five pairs of coefficients calculated for the end product *13779*. Comparison of indices with boundary limit parameters showed that Product Group *5251* results meet all the restrictions and requirements. Thereby, future demand forecast of end product *13779* in March 2014 is calculated as arithmetic mean between two indices of seasonality, calculated for the product group *5251* in March 2013 and March 2012.

Table 2.2. Example of hierarchy level basis determination for the end product 13799

Hierarchy level	**Hierarchy level code**	**Correlation coefficient**	**Maximum index of seasonality**
End Product	13779	-0,48	2,94
Product Group	5251	0,37	1,20
Product Secondary Group	F5PES	0,17	1,12
Product Line	5PE	0,17	1,12
SBU	INDUSTRIAL SOLUTIONS	0,31	1,01

As a result, main Excel sheet with the forecast figures for the end products consists of the following data: number of products sold over the previous four-month period, calculated individual index of seasonality and the final forecast sales volumes for the next three-month period and year, expressed in units and euros. In addition, other useful descriptive information about the end product is presented in the final forecast sheet. At this stage, the short-term forecast step for end products step ends. After the prediction of end products sales is completed, the figures have to be exploded to the level of materials and components, based on the bill of materials (BOM). Bill of materials for all the variety of X's products is stored in the warehouse ERP system of the company. Final exploded forecast contains the next three-month sales figures for the materials, semi-finished products and end products. This model is exploited on monthly basis. Thereby, current model allows dynamically adjust future procurement plan and production, based on sales information for prior most recent periods.

End stage of the forecasting process concerns verification conducted by experts and making manual corrections if necessary. Usually, responsible expert goes through the whole list of product forecast figures, tracks specific codes that have got suspiciously high forecast figures. Logisticar software allows users to graphically compare dynamics and scale of last year's sales with sales this year. In practice, products with a long lead time are often problematic and require manual correction of the forecast, as well as the code of the product life cycle. Correction of the last is usually made in order to prevent future deviations of the model results from reality.

2.2.1 Logical algorithm of the demand forecasting

In the previous section the forecasting model was placed in the logistic process landscape, and milestones of the calculation routine were traced on the practical examples. This particular part of the study will attempt to summarize schematically logical principles of calculations underlying the model.

Method and forecasting model have different meanings. Forecasting method is a certain ordered set of simple techniques, aimed to calculate the forecast figures in general. In turn, the forecasting technique – is a specific form of theoretical or practical approach to the forecast creation, one or more mathematical or logical operations aimed to obtain a certain result in the process of forecast creation. Thus, forecasting model can be defined as an ordered set of methods and techniques designed to predict complex phenomena or processes, (Lapygin et al. 2009). In other words, the forecasting model is a functional representation that adequately describes the process under study. It becomes the basis for obtaining its future values and states. Certain studies have shown that none of the forecasting methods used apart from others may not provide a significant degree of accuracy. But certain combinations can be highly effective.

Demand forecasting for the whole range of products produced by a particular company is a complex and multi-faceted research. The analysis of the model structure and behavior was summarized in a graphical representation of the core logical algorithm of the current model (Figure 2.4). Graphic method of algorithm representation is more compact and intuitive compared to the verbal description. The sales data for each produced end product is analyzed through this algorithm. On its every step certain decision is taken that affects the accuracy of the final forecast of future sales volumes.

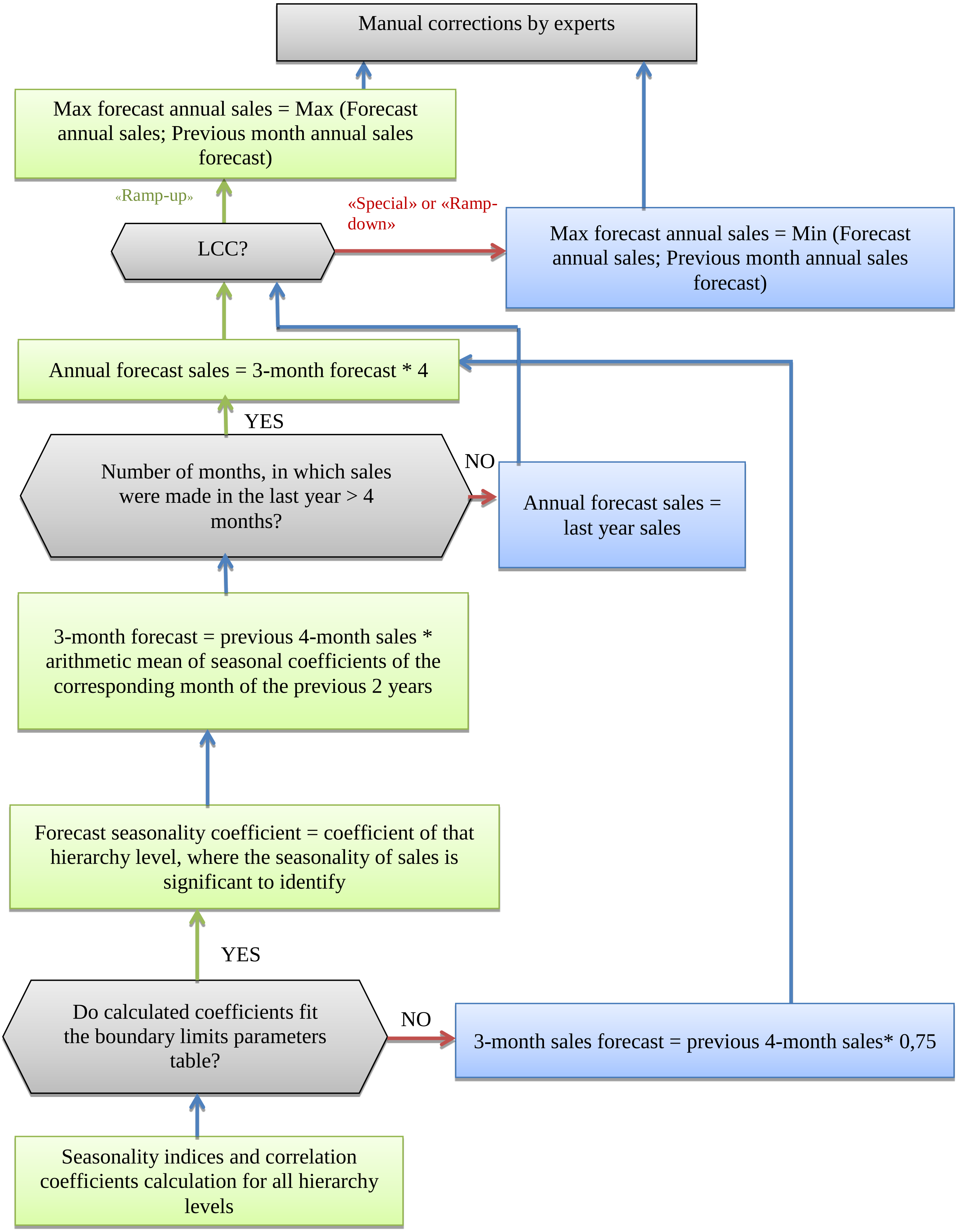


Figure 2.4. Logical algorithm of the demand forecasting

The earliest version of the demand forecasting model in company X was created three years ago. Over time of its usage, some shortcomings of the current model have been identified. Thereby, model gradually was supplemented and updated aiming to correct its defects, which can be identified only after some time of exploitation. For instance, early version of the current model did not take into account information about the products position on the product life cycle curve. Lack of this information decreases the accuracy rate. Forecasting model should be continuously improved and developed synchronously with the company and processes.

All calculation operations made in the model are shown as rectangles on the algorithm (Figure 2.4). Different conditions, the execution of a particular calculation will depend on, are depicted by hexagons. So, the whole figure represents a process of forecast formation underlying the current model. Figure 2.4 represents an algorithm of the forecasting logic that lays in the basis for the current model. Forecasting procedure can be divided into two parts - the calculation of model figures and their confirmation. Direct calculation of the sales forecast for the next period is not complicated. As described in more details earlier in the first chapter, it consists on preliminary calculation of the individual indices of seasonality for all months within the past two years and the correlation coefficients on each product hierarchy level. Final seasonal index is calculated as arithmetical mean of the indices that correspond to the certain month of the previous two years. Model will "trust" the forecasted index of seasonality of mature products that meet all the following conditions:

- Maximum index of seasonality and correlation coefficient meets the boundary limits requirements at least on one product hierarchy level.
- Sales of the product over the past year do not exceed 70% of sales in the year before.
- Sales of the product over the past year exceed sales, committed in the past four months, at least 25 times.
- During the last year, sales were carried out for at least four months.

If at least one of the above conditions is not met, the forecast value is calculated in accordance with the naive method of forecasting. In other words, the forecast will be equal to the volume of sales of the previous period. When calculation of the final demand forecast figures for the end products is completed, model estimates the maximum value of the forecast based on the product life cycle codes. In the end of the forecast procedure expert analyses all the model figures and introduces some manual corrections if necessary. In average, expert manually rechecks about one hundred values that cause him doubts. To make sure that significant forecast errors, that have been identified, will not appear in future, the expert usually manually changes the code of the product life cycle to the "special" or "ramp-down".

Product sales data is analyzed in accordance with the logical algorithm described above. The model can make an error in the forecast on any of its steps because the range of goods produced by the company is very diverse. Every product is unique and its sales behavior is unique too. Universal algorithm that underlies the model logic must be selected so that the total error of all forecasts was minimal. At the same time accuracy measures should be adjusted individually for each forecast system. The results of the forecast error analysis are described in the next chapter.

2.2.2 Disadvantages of the current forecasting model

Current model has been actively exploited by company X for several years. During this long time a few clarifications and changes have been introduced into the model to ensure a higher forecasting accuracy. However, the current model cannot be considered to be absolutely perfect. The biggest drawback of the forecasting model, according to the company's management opinion, is its insufficient accuracy. In particular, it refers to the product groups, which are characterized by long lead time, and unstable demand. Another problem of the current forecasting process is excessive amount of manual corrections introduced into the model figures. Employees, who use the final results of the model, do not trust its figures. These phenomena can be considered as a consequence of not sufficient accuracy and sporadic unexplained "outliers" for certain products. In some plants X production planning and procurement still occurs fully on the basis of old information about the past sales, as well as intuition of the decision makers. These disadvantages caused the content of this particular study. Theoretical part of the study contains the review of the existing methods of forecasting and assessment of their strengths and weaknesses.

3 FORECASTING APPROACHES REVIEW

This chapter is related to the theoretical demand forecasting methods. Different alternative methods of time series forecasting are described in more details. As a result of this review the most promising methods were selected, which were later used in the development of improved versions of the current model. In accordance to the forecasting object, section about the data required to model the future demand successfully, was also included in this chapter. As a determining criterion of the successful forecast for the company's management is its accuracy, different accuracy measures are also described in the end of this section. In addition, conclusions from M3-competition for time series forecasting models were also included in this chapter, because some established patterns of this experiment had influenced the course of the study.

3.1 Methods of forecasting

Prognostics development as a field of science now led to the invention of a variety of methods, procedures and techniques of forecasting (Chernysh et al. 2009). According to foreign and domestic researches, already more than 150 methods of forecasting are known, but just 20 out of them are widely used. A diverse array of methods and approaches, no doubt, requires classification. Mentzer and Moon (2005) suggested dividing forecasting methods into two large groups: quantitative and qualitative.

Qualitative approaches are based on expert judgment, intuition of experts. Such methods are subjective, that means interdependence between the forecasting results and the level of expert's competence, professionalism and intuition. Nowadays, methods from this group are widely used in the in marketing, economics and politics. In turn, most of the quantitative approaches are based on the certain mathematical dependences that allow the researcher to calculate the forecast values using the historical information about past states of the system. It is important to mention that any forecasting process involve indirect participation of the expert group. Human factor affects the forecast results when selecting certain forecasting method - basis of the model, as well as during model development stage. Even the most sophisticated statistical models are largely dependent on the decisions of experts (Wright & Goodwin 1998).

In practice qualitative methods are often used in the situations when there is a lack of accumulated information for analysis, as well as the unstable states of the system. One of the most popular qualitative methods is simple collecting of expert opinions and their subsequent generalization. Employees of the company may become experts (for example sales managers). Also experts can be invited from outside the company - third-party expert group (Armstrong 2002). More detailed classification of qualitative methods of forecasting is a challenging task. Methods can vary from the simplest approaches based entirely on intuition of one particular expert, to the whole group sessions, similar to the team work on decision making. Some methods look like the special forms of marketing research, and some are aimed at creating the most favorable conditions for the generation, collection and analysis of expert opinions.

Markidakis and Hibon (1979) named two main types of quantitative forecasting methods: causal forecasting methods, and time series forecasting. Causal methods are based on the idea of identification of the core factors that define the behavior of the system and its parameters. The process of finding these factors is a small part of the economic and mathematical modeling - construction of a model structure that explains the behavior of some specific entity, which takes into account the development of inter-related phenomena and processes occurring in the system. It should be noted that the application of multifactor forecasting requires solving a complex problem of finding and choosing certain significant factors that have some impact on the system behavior. This problem cannot be solved by purely statistical methods. It requires a very thoughtful study of the economic content of the phenomena or process (Bushueva 2004). Methods of time series forecasting are based on the extension of the trend component formed in the past up to the future situation. These mathematical forecasting models help to clarify the dependence between the future value and the past within the process itself and on this basis to calculate the forecast. The prevailing objective trends in the dynamics of economical parameters determine their future values to some extent. In reality, many market processes have some lag in their dynamic. Particularly it is clearly evident in the short term perspective. At the same time long time horizon forecast has to take into account the probability of changes in the conditions in which the market exists. These models are universal in terms of different application areas. Their overall structure does not depend on the nature of time series (Armstrong 2001). These methods in forecasting are the most suitable for a specific task set by X. Detailed description of some methods of time series analysis will be discussed in the next section. Additional information about the methods can be found in Appendix 2.

3.2 Data required for the demand modeling

A demand forecast is a central piece in a landscape of all operations of a modern firm. Very often companies use different time series methods or their combinations in order to get a sales forecast for the next time period. Time series techniques are based on the assumption that what happened in the past will happen in the future. Historical demand is projected into future in accordance with a mathematical formula (Arnold et al. 2008). This formula shows the mathematical relationship between the forecast and previous sales history. Thus, obviously, the key data required for demand modeling is consistent and full sales data. Time series forecasting methods that automatically capture seasonality typically require a minimum of two years of history and tend to work better with three or more. Some methods require even more years of accumulated information to analyze. In the context of the enterprise, it makes sense to forecast the future demand based on the sales data, expressed in monetary terms, because it is much more effective for a company to assess monetary loss. Certain products may be consumed in large quantities, but these costs are almost negligible for the company.

Additional information about the future from the outside environment can make the forecast more accurate, because it is very difficult to predict the future just looking into the past. The following are examples of data that could increase the forecast:

- Warehouse balances: if this information is available, potential out of stocks can be possibly cleaned from the sales history. This additional data estimates the lost sales and thus prepares better corrected past sales history to be used as a foundation for the calculated forecast.
- Campaigns and special events: if this information is available, such events should be replaced from the sales history because they are not systematic and happened rarely.
- Life Cycle Codes: forecasts may be influenced by a product's position in its life cycle. Product life cycle code contains information about the duration of fashion for a certain product, thereby enabling plans to be made to stockpile items in anticipation of a spurt in demand at the growth stage, for instance (Hollins, Shinkins 2006).

Forecasting model is a decision support tool which should consider different factors and justifies decisions. All sorts of additional information on the nature of sales, which will help the model to capture more clearly the possible patterns, promote more accurate forecasts. A demand forecast states the needed inventory that helps to overcome the fluctuations in demand. According to the information, a firm can start to plan its upcoming activities in a way that they can efficiently transform their inputs into outputs. Additionally, a demand forecast enables a corporation to provide its customer higher value. It distributes information including the needed products and stock keeping units (SKU), their quantities and the facilities required to fulfil the future needs. This way, the firm can gain better profit as forecast offers them a chance to lower their costs (Keath &Young 1996).

3.3 Methods of time series analysis

Dynamic processes occurring in economic systems are usually presented as a series of some economic indicator values arranged in a chronological order. Sequence of observations of a certain parameter, ordered according to increasing or decreasing values of another parameter, is called as dynamic series. If an ordering parameter is the time, this particular dynamic series is called as time series (Tatarenko 2008). Presenting the data as a time series is common in many spheres of life.

In case of economic sphere time series can represent the dynamics of daily stock prices, exchange rates, monthly sales, quarterly production volumes. Regardless of the origin of each time series, there are common challenges in the analysis of the original data set (Gardner 2006):

- description and graphical representation of the series characteristics;
- selection of the most suitable mathematical model to imitate the series;
- explanation of the series dynamic through other variables and evaluation of the degree of dependence between the observations;
- forecasting of the next elements in the time series, based on the previous observations;
- monitoring and control of the time series dynamic;

Investigation of regularities in the dynamics of any parameter over time is complex and time-consuming procedure. Every phenomena and process can be characterized by a wide range of factors acting in different ways and forms. These factors can be conventionally divided into two groups which differ by their impact on the time series. First group factors affect the trend component. At the same time, second group factors cause the random fluctuations. These factors shift the time series elements relative to the main trend in different directions. Thereby, time series can be decomposed into the following components (Chernysh et al. 2009):

- trend component - changes very slowly under the influence of long-term factors;
- cycle component – varies smoothly, represents the long periods of boom and bust;
- seasonal component – consists of certain seasonal sequence of repeating cycles;
- random component – remains after subtracting the regular component of the system, cannot be explained systematically;

Complete time series can be defined as a sum or multiplication of a certain set of components described above. Set content depends on the chosen forecasting model. Multiplicative models are used more often than additive models. The list of time series models is vast and diverse. Not every method involves the analysis of all four components of time series. The most commonly used theoretical forecasting models will be described in the next sections.

Naive methods of forecasting

Naive forecasting methods are based on the assumption that more recent observations predict future values of the time series with the higher accuracy rate than older observations. Usually the formula for forecast calculations here describes a very simple dependence between the forecast figures and past observations. Basic naïve model forecast value can be equal to the previous element in the time series for instance. Models of mean average are very popular, where forecast is calculated as a mean of several past observations. These models are more robust to different random fluctuations because all the non-systematic outliers are smoothed. Moving average models use in the calculations only recent data of a few last periods; old data is completely excluded from the forecast. Assumption that recent values of the time series adequately describe the future situation formed the basis for the class of weighted average models, where obsolete observations have smaller weights in the final calculations than the recent data (Svetunkov et al. 2009).

Exponential smoothing model

Nowadays, short-term forecasting became a very popular research topic. One of the well-known adaptive methods of short-term forecasting is Brown´s Method. Whereas in the simple moving average the past observations are weighted equally, exponential smoothing assigns weights that decrease exponentially over time. In this group of methods there is a special constant parameter α which determines the degree of dependence between the forecast and older observations. An influence of the past decreases exponentially as the data in the observations becomes older. Models of exponential smoothing are used successfully in the short-term forecasting of the main tendency for the next time period (Gardner 2006).

Brown's model has a very significant disadvantage. It is always "late" by one time period. In order to solve this problem several modifications of Brown's model were developed. These models assume an approximate trend in the time series model profile set a priori. In practice Holt's and Holt-Winters's models are used more often than other known modifications of the simple exponential smoothing models. In the middle of the previous century Holt developed an algorithm where level and trend components are exponentially smoothed. Moreover, the smoothing parameters are different. Extended Holt-Winters method uses three different smoothing parameters. It also takes into account seasonal component, (Svetunkov et al. 2009).

Main disadvantage of these models is caused by the assumption of the existence of the stable trend component in the time series that does not change over time significantly. In practice this assumption does not correspond to the reality correctly: smooth linear trends sometimes can be replaced by abrupt and highly non-linear dynamic, and the frequency of the cyclical component is not constant. This problem can cause a significant divergence of the forecast with actual figures.

Furthermore, usually it is a complicated challenge to find and choose appropriate smoothing coefficients, because their values define the forecasting capabilities of the model, but there is no universal approach that could systemize this task. As the result the researcher has to spend a lot of time selecting coefficients in order to prepare an adequate forecast (Lukashin 2003). Despite the fact that all approaches described above (naive methods and methods of exponential smoothing) are usually implemented in the field of business where the object of modeling is stable enough and is not that complicated, their accuracy rate is still not high enough. These approaches can be successfully used as supplementary and supportive methods in combination with others in different forecasting tasks, such as time series decomposition to the trend and seasonal components.

Error correction forecasting models

Forecasting models provided with an error correction mechanism take into account values of the past mistakes. Simultaneously these models predict two different parameters: the value of the main forecast *Yt* and its deviation from the actual value εt. The forecast figures can differ from the reality due to a few reasons. The first reason is incompleteness and inaccuracy of accumulated knowledge and data about the current state of the analyzed system. Secondly, structure of forecasting model and methods utilized in it predetermine the success of the forecast significantly. Certainly, the presence of random component directly affects the final accuracy of the forecast. As a result, error value can be decomposed into the system and random components also.

Clements and Henry (1999), as well as Entov and Nosko (2002) described four methods of accounting these errors. These techniques let to adjust forecasts in real time. The most frequent method is to adjust the forecast by a value of forecast error one step back, or by the value of the average error of all previous forecasts. However, corrective measure can be independent from the deviations of the actual values from the results of forecast calculations. Researcher can define this measure as a certain function or constant, if it is an expedient option. According to the results of Entov and Nosko research correction by the value of an average error can achieve significantly better results than most other methods. However both approaches are capable to remove the systematic error. Before researcher decides to implement one of the considered methods of error correction it is necessary to analyze the error data accumulated over a long period of time, in order to separate systematic errors from random variations.

Regression forecasting models

Class of regression forecasting models is based on certain function that describes the relationship between the quantitative characteristics of complex systems. Type of regression function is determined by fitting experimental data. The process of developing a regression model consists of two stages: the function type selection and calculation of function parameters. The first challenge does not have any systematic solution. Experience and intuition of the researcher usually help in this case. Another way to select a certain regression function is to go through a finite number of functions and select the best out of them. Most frequent are linear, quadratic functions, polynomials of higher order, logarithmic, exponential and power functions. Once function is chosen, the next step is to select its parameters so that the values of the function will be placed as close to the experimental values as possible. The parameters calculation is usually carried out by the method of least squares. When the function is completely defined, there is no difficulty to calculate the final forecast (Lapygin et al. 2009).

Decomposition forecasting models

For time series, which clearly contain certain trend and seasonal components, this information can be used in order to improve the accuracy of the forecast. In this case, the trend and seasonal component should be quantified. At the same time, taking out the trend and seasonality, it is possible to reduce significantly the amount of noise in the time series, and thereby to reduce the uncertainty of the future. The procedure of identification of non-random model components in the raw observations data is called decomposition (Arkhipov, 2008).

The classical procedure of time series decomposition begins with its smoothing in order to identify common trend. There are many smoothing methods, the most common are: consolidation of time slots, moving average, analytic alignment (replacement of time series by a certain smooth analytic function). Later on, as a result of arithmetic operation of division or subtracting the actual and smoothed time series (depending on model) we get another time series seasonal values, purified from the trend component. The resulting values of seasonal coefficients can be averaged in order to reduce the effects of noise. Thus, after extrapolation of the trend line for a few time periods in future the final forecast can be done through the multiplication or summation of the trend values and appropriate seasonal coefficients.

In reality, time series that have been cleaned from the seasonal component does not always fit well in a linear relationship because of the marked deviations. Consequently, in this case, the initial data should go through primary cleaning procedure in order to get rid of all sorts of random outliers. For more precise identification of seasonality by classical decomposition method it is highly recommended to have at least 4-5 full cycles of data. However, to isolate completely the effect of noise and to determine accurately the trend components of time series are very difficult tasks to solve. Another promising method of time series analysis, based on the effect of decomposition, is described in the literature. Greek statisticians Nikolopoulos and Assimakopoulos (2003) proposed a Θ-model, which is based on a modification of the curvature of the time series trend.

ARIMA models

In contrast to the time series forecasting method described above ARIMA methodology does not assume any strict model for forecasting. This forecasting methodology defines a general class or classes of models that will simulate time series. Then the algorithm like designer assembles the most appropriate forecasting model through internal parameters adjustment. Most of the work on studying this particular methodology was carried out by two statisticians G. E. P. Box and G. M. Jenkins (1976). These researchers developed a hierarchy of predictive models in this class. General model, which was developed in 1976, is a combination of autoregressive model and moving average model. Firstly, ARIMA determines the number of parameters in autoregressive and moving average forecasting methods to provide the most accurate results. After the assembly phase comes a phase of parameters estimation (Tebekin 2013). And finally, model with fully defined parameters can calculate the forecast values and confidence intervals (Statsoft 2013).

Artificial neural network models

Nowadays, models related to this class are considered to be very promising. They represent a set of elementary processors - artificial neurons, interconnected by synoptic connections. This network processes an information input and forms the outcome signals (Gorban et al. 1998). The apparatus of artificial neural networks (ANNs) is functioning as an independent component of the control decision-making system. Time series forecasting by neural networks is characterized by minimal analyst participation in the process of creating the model structure, as the neural network has the ability to learn. Learning algorithms adapt the weight coefficients in the models for a certain structure of the analyzed input. There are some significant difficulties in ANNs model development process. In particular, there is a real challenge to create a sample for network training purposes, which should meet the requirements of completeness: the sample should contain all valid examples of the test range, and consistency: sample must not contain contradictory examples (Chuchueva 2012). Thus, the mathematical model selection does not depend on the expert's choice. One of the main disadvantages of the neural network models is their opacity. After the automatic learning phase model works somehow, but the logic of decision-making is completely hidden from the expert.

Comparison of the forecasting models

As a result, the main advantages and disadvantages of all the forecasting approaches described above have been summarized in Table 3.1. In addition it should be noted that accuracy of the forecast was not evaluated for any considered groups of forecasting models. This was done due to the fact that the prediction accuracy of a given process depends not only on the model but also the experience of the researchers, the availability of data and hardware capacity and many other factors. Characteristics of the forecasting accuracy will be discussed in the next section.

Table 3.1. Advantages and disadvantages of the forecasting models comparison table (Chuchueva 2012)

Model name	**Advantages**	**Disadvantages**
Naive models	The simplicity of the design and results interpretation, transparency of the modeling; cheapness; often used for comparison with other models;	Inefficiency in case of the presence of a trend / seasonality; does not account any external influences; not suitable for long-term forecasting;
Regression models	Simplicity, flexibility, transparency of the modeling, uniformity of analysis and design;	The complexity of the identification of functional dependency and its coefficients;
Decomposition models	Separates seasonal component from the trend;	The applicability of these models is narrow;
ARIMA models	A powerful tool for generating short-term forecasts; flexibility; can describe a wide range of time-series;	Time-consuming and resource-intensive model identification; requires a large amount of raw data; complicated model adjustment process;
Exponential smoothing models	Simplicity of modeling; uniformity of analysis and design; easy correction of the smoothing parameter;	Lack of flexibility; narrow applicability of the models; lag effect;
Neural network models	Nonlinear models; scalability, high adaptability; uniformity of analysis and design; many examples of practical application;	Lack of transparency; the difficulty of choosing the structure of the model; strict requirements for the learning sample; the difficulty of choosing a learning algorithm; resource intensity of the learning process;

3.4 Combining forecasting methods

Usually, certain circumstances and insurmountable external factors strongly affect the success of the predictive model (Laurence et al. 2006). Very often, combination of the most appropriate quantitative methods of time series analysis and certain qualitative methods based on expert opinion could have a beneficial effect on the accuracy of the entire forecasting process. Demand forecasting accuracy for certain products produced by the company can be improved if the model end users do have some specific knowledge about products, processes, temporary projects or major customers, which cannot be depicted in the model (Webby et al. 2001).

Joanne Utley (et al. 2011), Bails and Peppers (1993) described a few feasible schemes of including forecast end users in the forecasting process:

- Adjustment: final forecast figures can be manually corrected by experts, if they do not trust the model results according to certain reasons.
- Combination: final forecast is formed as a simple average or weighted average of the model predictions and forecasts made by experts.
- Setting the structure and parameters of the model: experts may express their views on the structure of the model or its internal parameters.

3.5 Criteria for selecting a forecasting method

Well-known classes of time series analysis and forecasting have been described in the previous section. Each class is represented by a unique set of predictive models. There is a great variety of different models. Therefore, any company that decides to forecast future demand for its products undoubtedly faces a serious problem of choosing one or another method. In order to choose from the whole variety of known prognostic methods the most appropriate one, it should best meet the selection criteria.

Review of the relevant literature on the subject has shown that researchers have identified two most important criteria for forecasting models evaluation. Bails and Peppers (1993) and Yokum with Armstrong (1994) suggested to assess the predictive models by the criteria of their accuracy. On the other hand, Waddell and Sohal (1994), as well as Clifton and colleagues (1998) evaluated the success of the forecasting process by the criteria of matching the originally defined objectives of forecasting in the company. There are many other criteria for assessing the success of forecasting. For example, time horizons, size and completeness of the data available, costs, ease of use, as well as the significance and value of the results (Yokum & Armstrong 1994).

Taking into account the objectives that the company X pursued in this study, it was decided to evaluate and compare different prediction methods by the criterion of the forecast accuracy. Some measures of the forecast accuracy will be discussed later.

Accuracy of the forecast directly depends on the error value, occurred in the forecasting process. According to Chopra, S. and P. Meindl (2006) measurement of the forecasting accuracy has two main aims. First, forecast error analysis helps managers to determine how accurately the current method detects systemic trend component of demand. For instance, if the forecast is always accompanied by a positive error value the manager may conclude that the forecast method always overestimates the actual demand. Secondly, error analysis is an important component of the forecasting process, because any future plan must take into account possible deviations from the reality.

Forecasting accuracy – is the opposite concept of forecasting error. If the forecasting error is substantial, then the accuracy is low and vice versa. Menzer and Moon (2005) divided these measures into two groups: absolute and relative. All absolute measures are based on mathematical operation of subtracting the actual values out the forecast (Formula 3.1).

$$Error_t = E_t = Y_t - X_t \,, \qquad (3.1)$$

where Y_t, X_t *– the predicted and actual values, respectively, at time t*

One example of the absolute error evaluation is the average value of all errors, which is calculated by the following formula (3.2):

$$Mean\ Error_t = ME_t = \frac{\sum E}{N}, \qquad (3.2)$$

$N - number\ of\ time\ periods$

There are other ways to estimate absolute prediction error: mean absolute deviation (MAD), the root mean square error (MSE) and others. The second group of measures is based on the ratio of forecasted and actual values. Among the relative error measures average percentage error (MPE) and mean absolute percentage error (MAPE) are used more often.

MAPE fundamentally differs from another measure MPE because all MAPE values are positive due to the modulus function in the numerator. The formulas for calculating relative measures (3.3) and (3.4) are listed below.

$$MPE = \frac{X_t - Y_t}{X_t} * 100\%, \tag{3.3}$$

where X_t – actual value, Y_t - forecast

$$MAPE = \frac{|X_t - Y_t|}{X_t} * 100\% \tag{3.4}$$

Symmetrical mean absolute percentage error (sMAPE) is used as an alternative to MAPE error in case the actual value is equal to zero or close to zero. SMAPE measure is calculated by the following formula (3.5):

$$sMAPE = \frac{2}{N} \sum_{t=1}^{N} \frac{|X_t - Y_t|}{X_t + Y_t} * 100\% \tag{3.5}$$

One example of error estimated in relation to other prognostic techniques is «Theil's U». Typically, this value determines the accuracy of the forecast in relation to another forecast, calculated by the Naïve methods. Naive methods suggest using the actual value obtained in the previous time period as a future forecast without any adjustments. If the U-value is greater than 1, then the evaluated forecasting method is less accurate than naive approach. If U does not exceed 1, then the evaluated approach provides greater accuracy compared to naive methods. The same idea of the relative accuracy measurement can be implemented by dividing MAPE of the forecast model by MAPE of Naïve forecast (Mentzer & Moon 2005).

3.6 M3-competition for the forecasting models

Forecast accuracy directly affects the costs of the company, as well as the quality of service provided to customers. Therefore, one of the most important initial stages of designing the forecasting model for a certain company is a preliminary comparison of existing methods of forecasting. However, published descriptions of the advantages and disadvantages of forecasting techniques conducted in practice or in theory are not enough to form an objective and critical view about all the methods. Series of M-competitions for the forecasting models developed by Makridakis and Hibon (2000) is an interesting empirical research in the field of forecasting. As part of this competition some well-known methods of time series forecasting were compared by the criterion of their accuracy. Over one hundred different time series were analyzed by hundreds of experts from all around the world. The first series of M1-competition was held in 1979. All the forecast results generated by different models were compared with each other as well as with the results of the simple naive model and exponential smoothing model by the criterion of their accuracy.

Later in 2000 M3-competition was held where 3003 different time series were presented. To ensure maximum objectivity of the results time series were taken from the field of micro, macro economy, industrial, financial, and demographic areas. The time step in the time series varied as well. Over twenty different methods of time series analysis took part in the competition representing a group of naive models, exponential smoothing models, models of decomposition, ARIMA models, artificial neural networks and several expert models. The list of those methods of forecasting and a brief description are given in Appendix 3.

It is also important to mention that the comparison of the forecast results obtained by the participating in the competition methods was carried out according to the criteria sMAPE, Median sAPE, Median RAE error measures. Naive forecasting model and model of exponential smoothing were chosen as a benchmarking comparison basis. All the results of this interesting research and findings were published in the journal «International Journal of Forecasting». The difference in accuracy of six most accurate models and two selected samples is displayed in Table 3.2 and Table 3.3. Measure sMAPE was used to measure forecast accuracy. It can be seen that in both benchmarking analysis some of the forecasting methods outperformed in accuracy both naive and exponential smoothing methods. The absolute winner in both samples was the θ-method.

Table 3.2. The difference in the accuracy of the forecasts created by various methods of time series analysis and naïve forecasting (Makridakis & Hibon 2000)

	Forecasting horizon(s)				
	1	Average: 1–4	Average: 1–6	Average: 1–12	Average: 1–18
Theta	0.4%	0.6%	0.5%	0.5%	0.6%
ForecastPro	0.2%	0.4%	0.4%	0.3%	0.4%
ForecastX	0.1%	0.2%	0.3%	0.2%	0.1%
Comb S-H-D	−0.1%	0.0%	0.0%	0.0%	0.1%
RBF	−1.1%	−0.5%	−0.2%	−0.3%	−0.1%
ARARMA	−0.9%	−0.8%	−0.9%	−1.1%	−1.1%

Table 3.3. The difference in the accuracy of the forecasts created by various methods of time series analysis and exponential smoothing forecasting method (Makridakis & Hibon 2000)

	Forecasting horizon(s)				
	1	Average: 1–4	Average: 1–6	Average: 1–12	Average: 1–18
Theta	2.1%	2.2%	2.1%	2.3%	2.5%
ForecastPro	1.9%	2.0%	1.9%	2.1%	2.3%
ForecastX	1.8%	1.8%	1.8%	2.0%	2.0%
Comb S-H-D	1.6%	1.5%	1.5%	1.8%	2.0%
Dampen	1.7%	1.6%	1.5%	1.8%	1.8%
RBF	0.6%	1.1%	1.3%	1.5%	1.7%
ARARMA	0.8%	0.8%	0.7%	0.7%	0.7%

Analyzing the comparison results, researchers have formulated the following identified regularities:

1. Statistically sophisticated, complex methods usually do not outperform the simple methods in accuracy context.
2. Accuracy rate of the forecasting method depends on the selected error measures, as well as the nature of forecasted data.
3. Combination of different forecasting methods usually causes the decrease of the forecast accuracy.
4. The accuracy of the forecast depends on the time horizon of predictions.

Author of this study believe that all the results obtained in the experiment, can be considered an objective assistant for managers that helps select the most appropriate method of forecasting, which will have the greatest possible accuracy. Even a small advantage in predictive accuracy significantly affects the operation performance of the company, and thus helps to reduce inventory levels, to increase efficiency of the planning process, as well as to improve customer service and to streamline the company's costs (Makridakis 2000).

3.7 Summary

Second chapter provides a detailed description of the demand forecasting model exploited by the company X. Main goals of the forecasting process for company X are to improve the efficiency of inventory management, to reduce costs, reduce the total delivery time of products and levels of safety stocks, as well as to maintain a high level of customer service. This model belongs to the group of multiplicative decomposition models with individual chain indexes of seasonality. The choice of this particular method of forecasting was done due to the seasonal sales behavior of most of the products produced by company X. Main defect of the current model is its deficient forecasts accuracy and, as a result, an excessive amount of manual corrections introduced to the final forecast figures. Review of the existing theoretical methods of time series forecasting and measures for evaluation the accuracy rates of the forecasts helped to determine the strategy for further refinement of the current model. In the next chapter, main perspective directions of improving model accuracy will be discussed and formulated.

4 POSSIBLE DIRECTIONS FOR ACCURACY IMPROVEMENT OF THE MODEL

The fourth chapter provides the results of the analysis of forecast errors, compiled by the current model, suggests some causes of occurred deviations from the actual values. The outcome of this section is the list of possible directions of further improvement of the forecast accuracy rate.

4.1 Error analysis of the current model forecasts

One of the most important stages in the forecasting process is a verification of the results that means the evaluation of their accuracy and validity. Every forecast is characterized by considerable degree of uncertainty that has to be measured before taking the managerial decision to use the predictions about the future and apply the hypothesis about the prospects of the development of various systems, for instance. Developing the forecasts, experts are interested in improving their reliability. Detailed knowledge about the dynamic behavior of the forecast error helps to identify special products or product groups that are different from the majority of the forecasted products in context of their forecasting capabilities. It also helps to justify the choice of a particular forecasting method that forms the basis of the model. Comparative analysis of the forecasted figures and reality, as well as the creation of appropriate data bases with information about their deviations allow adjusting the forecasting methods parameters that are in use. Successful predictive models help to decrease average inventory levels of products in stock, as well as significantly improve the level of customer service (Kerkkänen 2010).

4.1.1 The methodology used in measuring the accuracy of the forecasting model

Objective assessment of the forecast accuracy is the result of a retrospective analysis of forecast errors. The accuracy of the particular forecasting model can be judged by the magnitude of the forecast errors - the difference between the predicted and the actual values of the variable. Such estimates can be obtained when the period of preemption ended and the actual value of the variable is already known. Such estimates are called a posteriori accuracy estimation of the forecast quality (Mentzer & Moon 2005). They include absolute and relative indicators that quantify the magnitude of the prediction error in units or as a percentage. Variety of accuracy measures have been previously described in the third chapter of thesis. Of course, each of these indicators contains unique information about the model performance. As has been said earlier in the section about the M3-competition of different forecasting methods, the final assessment of the accuracy of the method depends on the selected accuracy measures, as well as the nature of the forecasted data. It should also be noted that the accuracy of a single forecast has insufficient value for a researcher, because the formation of the phenomenon under investigation is influenced by many different factors, therefore, a complete coincidence, or significant divergence of the forecast and the reality may occur due to a particularly favorable or unfavorable circumstances. Single accurate prediction can be generated by bad model, and vice versa, thus the quality of model's predictions can be judged only by the multiple comparisons of the forecast figures with their actuals according to a variety of criteria.

The model used by X, predicts sales for all end products produced by five plants in the Northwest region. Sales volumes are expressed in units and cash equivalents. In order to assess the accuracy of the forecasts of future demand for company's products, all past forecasts created during the period from May 2012 to April 2014 were collected in one Excel table. Data arrays were cumbersome; each forecast consists of fifteen thousand lines on average. To speed up the calculations made in Excel evaluation of forecast accuracy was carried out only for products manufactured in a particular Estonian factory. All the products that were considered to be new were also excluded from the forecasts. This was done due to the fact that the model is used only for those products that are placed at the stage of maturity on the product life cycle curve. The accuracy of the forecasting model was measured using the mean absolute percentage error MAPE measure, as well as the coefficient of determination. Both measures were calculated for sales, expressed in euros.

Taking into consideration the specificity of the model under investigation, and completeness of the data available for analysis, MAPE measure was calculated for each month during the period from May 2012 to February 2014 for each product being sold in the following three-month period and the actual volumes of sales for this period. One of the advantages of MAPE in comparison with other accuracy measures is the ease of its interpretation and evaluation. This measure does not depend on the scale of sales that is important in X's case. Moreover, the percentage expressing the deviation of the forecast from the actual data, allows researcher to compare the accuracy of the forecast formed for end products, groups and other levels of the product hierarchy. In addition, measure MAPE allows easy comparison of accuracy achieved by different forecasting methods. However, this measure has two disadvantages: lack of theory on the statistical interpretation of MAPE, and inequality of errors for situations when the forecast Y exceeds and underestimates the value of real demand X by the same percentage value (Makridakis & Hibon 2000). In turn, the coefficient of determination was calculated on a monthly basis for the whole range of forecasted and actual values. The coefficient of determination R-square is an indicator that represents the ratio of the „explained" variation of the variable. R-square value close to "1" indicates that the model explains almost 100% of relevant variable variation (HELPSTAT 2012).

The coefficient of determination that is greater than 0.7 (model explains more than 70% of variation) is usually considered to be satisfactory in practice, MAPE should be less than 10%. Ideally, they should tend in the limit to one and zero, respectively (Ayvazyan 2001). There is a misconception that the accuracy of the forecast depends entirely on the quality of the model, but this is not absolutely correct. The most important factor is the "natural unpredictability" of the forecasted system. Initially, most of the demand forecasting methods described in the literature have been used successfully in practice among consumer markets. Currently there is growing interest in the use of demand forecasting for companies that operate in the industrial sector, despite the fact that the industrial and consumer markets differ significantly from each other. Very often when the company introduces the demand forecasting procedure, there is a tendency when the company begins to imitate the basic concepts, goals, and principles according to the experience of other companies, in order to accelerate the implementation phase at the facility. A similar trend was observed at the stage of forecast accuracy analysis and evaluation. Thus, the analysis of the forecast errors, as well as the definition of the boundaries for satisfactory accuracy rates should be carried out taking into account the specific nature of the company, its customers and products produced and sold. Thereby, next sections provide the description and analysis of the following points

- dynamics of the monthly coefficient of determination;
- dynamics of average monthly mean absolute percentage deviation of the forecast from reality;
- annual average error on determinant hierarchical levels, the forecast figures were based on;
- statistics of false zero forecasts;

4.1.2 Determination coefficient of the forecasts

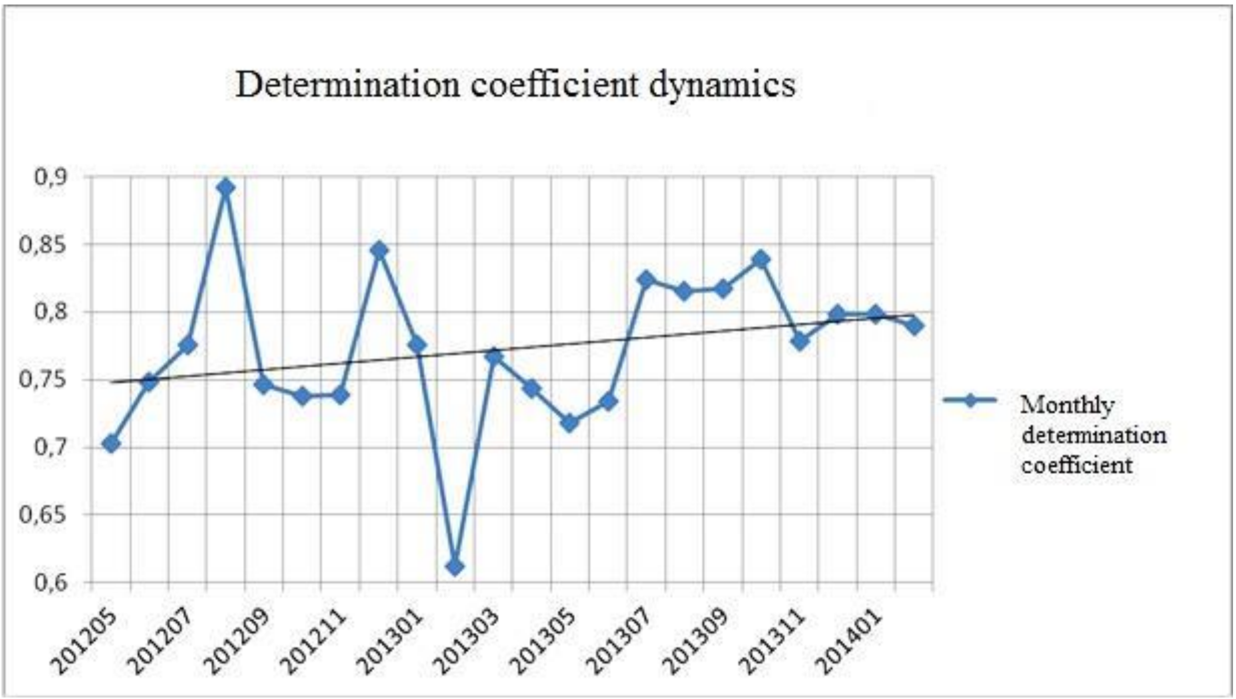


Figure 4.1. Dynamics of the determination coefficient accuracy measure

Figure 4.1 illustrates calculated determination coefficients for the forecasts made by the model at the beginning of each month during the period from May 2012 to February 2014. Current model predicts future sales of all products, which will be sold over the next three months. Looking at the chart, we can say that the coefficient of determination was not constant on the entire time period. Coefficients of determination for different months significantly differ from each other. However, throughout the considered time determination coefficient dropped below a satisfactory level of 0.7 only once in February 2012, it was equal to 0.61. During all other months of percentage variations in sales volumes explained by the model, ranged between 70% and 90% with an average of 78%, which indicates that the current model of company X has a fairly high quality of the forecasts. It should also be noted that the most recent forecasts in average have a higher coefficient of determination, in comparison with later forecasts. Scatter in the values of this accuracy measure forecast significantly reduced by the end of 2013. This could happen due to the fact that in mid-2013 some amendments were made in the forecasting model, namely, it began to take into account the stages of the product life cycle. Since that moment the model started to calculate maximum possible figure of annual sales, which depends on the particular life cycle code assigned to each end product. The value of the maximum possible volume of sales introduces certain corrections in the final calculation of the forecast for the next period.

Looking at the determination coefficient graph, certain months, that are characterized by particularly high and very low values of this accuracy measure, are noticeable. Increased accuracy was reached in August and December 2012, and during the late summer - early autumn of 2013. In turn, the relatively low accuracy rates were achieved in the autumn of 2012, and in late spring 2013 with an abnormally low slump in February 2013. Due to the fact that the accuracy of the forecasting model depends directly on the product hierarchy levels, chosen as determinant levels for forecasts, we can put forward the assumption that in those months when the determination coefficient was low, so large number of forecasts were probably built on the basis of the naive method, which means full equality of the forecast to the product consumption in the previous month. To verify this assumption special diagram was built, it clearly shows the number of naïve forecasts in each month throughout the analyzed time period (Figure 4.2). Looking at the chart, you can see that in February and since April till July number of forecasts, calculated this way is much higher than an average rate. It is also interesting to note that in 2013, the number of products, forecasted by the "Flat" level significantly in average increased in comparison with 2012. The model creates the forecast on the "Flat" level (uses a naive prediction method) when the correlation coefficients and seasonality indices on any of its hierarchy levels do not correspond to the table of boundary limits. As a result, on the basis of the foregoing it can be concluded that the table with boundary limit values directly affects the accuracy of the final forecast.

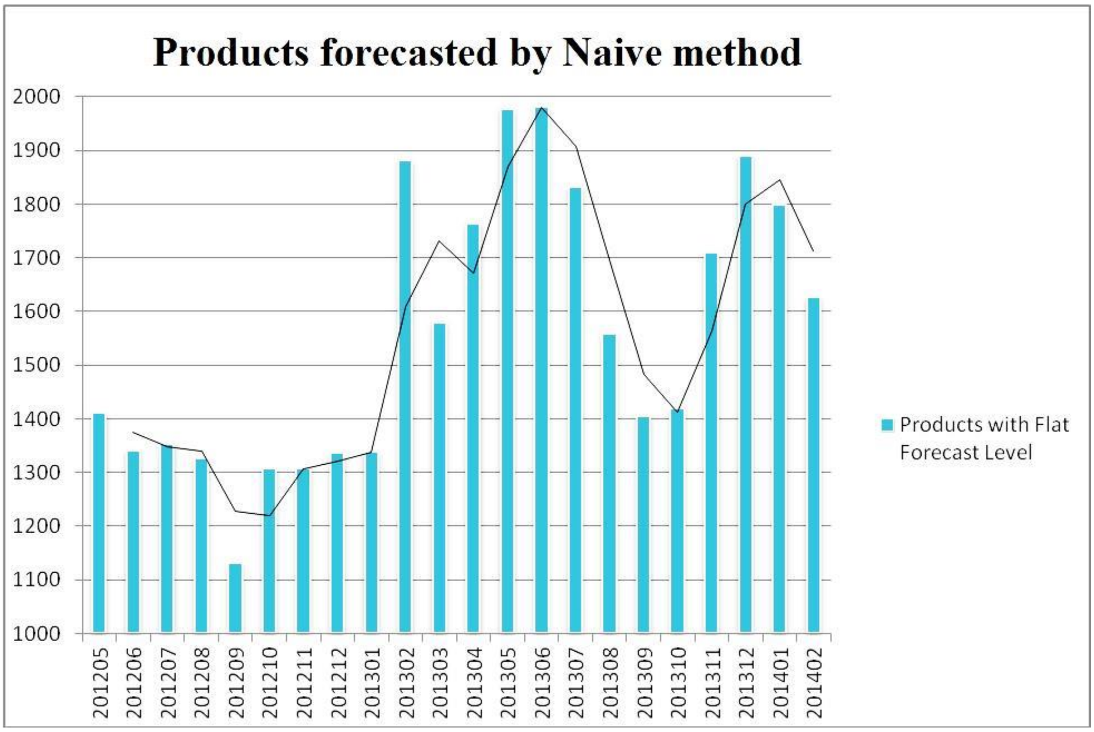


Figure 4.2. Monthly number of products forecasted by the naive method

4.1.3 Analysis of an average monthly MAPE dynamics

The coefficient of determination is quite common measure of a forecast proximity to the reality. In turn, the absolute percentage error can be calculated for each product or group individually throughout the considered time period. Figure 4.3 illustrates the average percentage error of the forecasts, during the time period from May 2012 to February 2014. Recall that MAPE is calculated as the ratio of the deviation of the forecasted sales volumes to the actual values of sales in the next three-month period. Real demand is in the denominator of the ratio. In this case, when the forecast and actual demand simultaneously are equal to zero, MAPE is also equal to zero error. It should be noted that in case the demand is equal to zero and the model figure is greater than zero absolute value of the error rate must be equal to infinity. To simplify the analysis of the error statistics this particular situation has been investigated in isolation from other values of forecast errors.

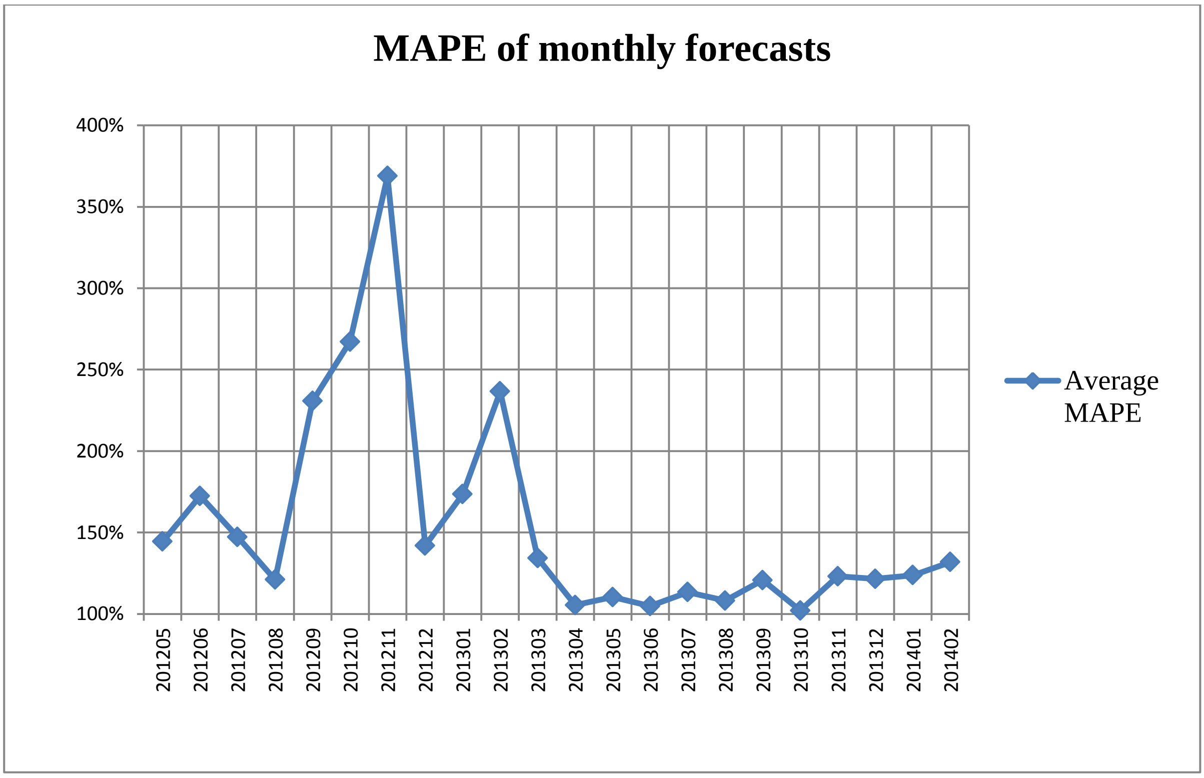


Figure 4.3. Forecast MAPE

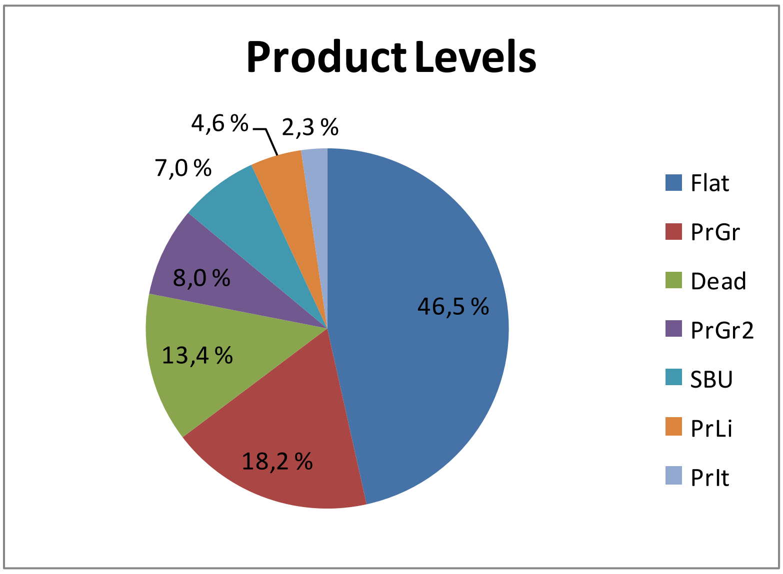


Figure 4.4. Average share of determinant forecast levels

The average percentage error in 2012 from May to December was 200%. In 2013 error significantly decreased by 70% to 130% in average. This observation may become additional substantial evidence that taking into account information about the product life cycle stages had beneficial effect on the accuracy of forecasts. Nevertheless, overestimation of the demand by 130%, excluding the zero false predictions, is a sufficiently large deviation.

In the previous section it was noted that number of products, forecasted by the naive method affects the accuracy of the overall forecast. The pie chart in Figure 4.4 shows the average percentage share of the product hierarchy levels that have been chosen as

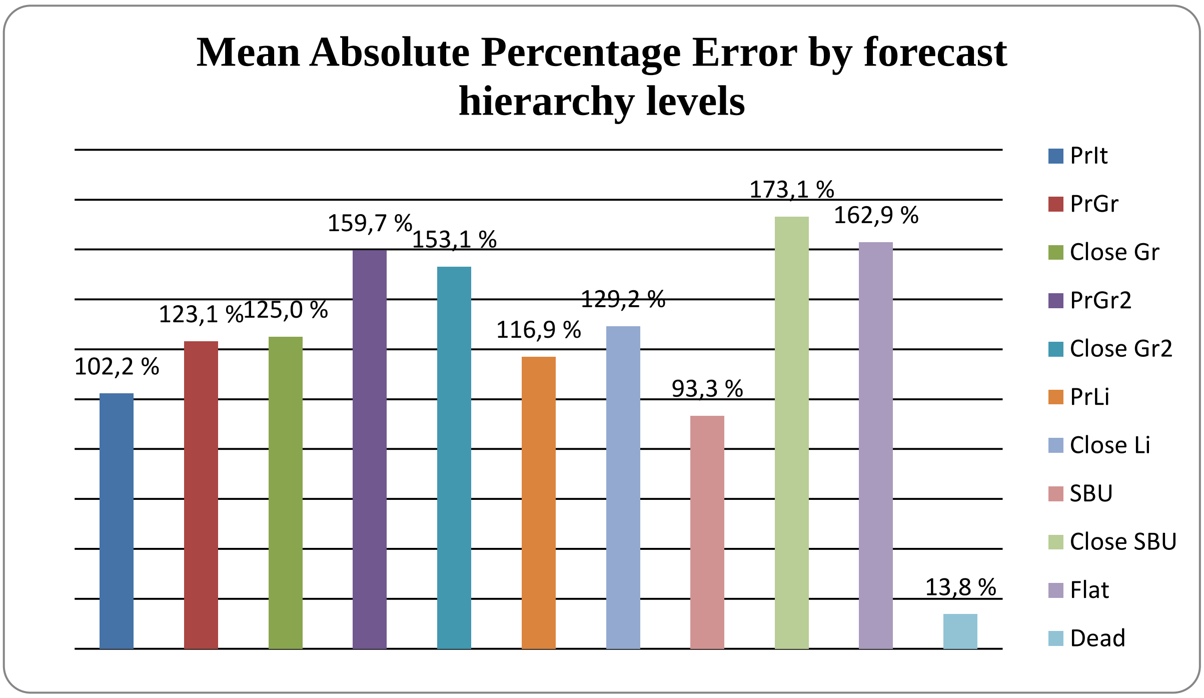


Figure 4.5. Mean absolute percentage error for different determinant hierarchy levels

determined by the model during 2-years' time period. Almost 50% of the company's products were predicted by the naive method. However, the next largest group is the products, whose forecast was built on the basis of seasonal indices belonged to product group level. For such products the correlation coefficients for the time series on the end product level and product group level were considered to be sufficient, and their seasonal indices also satisfied the given constraints presented in the table. The smallest group is a group of products whose forecasts were based on the seasonal indices of end products. These products were only about 2%. However, mean absolute percentage error for this particular group was 102%, with an average error for all levels of 130% (see Figure 4.5). The largest average relative error corresponds to the quality of the forecast based on the level of strategic business units with low values of the correlation coefficient (CloseSBU). Products from this group totaled 7%. The average percentage error of forecast at the SBU hierarchy level exceeded error of the naive method by 10%. Forecasts based on other levels of the product hierarchy have greater accuracy in average compared to naive methods.

Thus, forecasts based on product group's data should be given special attention due to the fact that the correlation coefficients and maximum values of the indices more often than other levels satisfy the table of boundary limits. As well as the average percentage error at the group level is 123%, which is below the average value for all levels in sum.

4.1.4 False zero forecasts

This section focuses on a singular case of the MAPE measure: zero false prediction. Zero false forecast occurs when a model for some reason determines the forecast for future three months sales volumes as zero, when in reality product will be sold during this particular time period. This can happen for example when this particular product did not have any sales during the previous four months at all. Thus, if the model uses the naive method of forecasting, future sales will also be set to zero. If the model identified significant correlation in the sales history on one of the product hierarchical levels, even if the seasonality index will be different from 0.75, each product with zero past sales will get zero as a forecast figure.

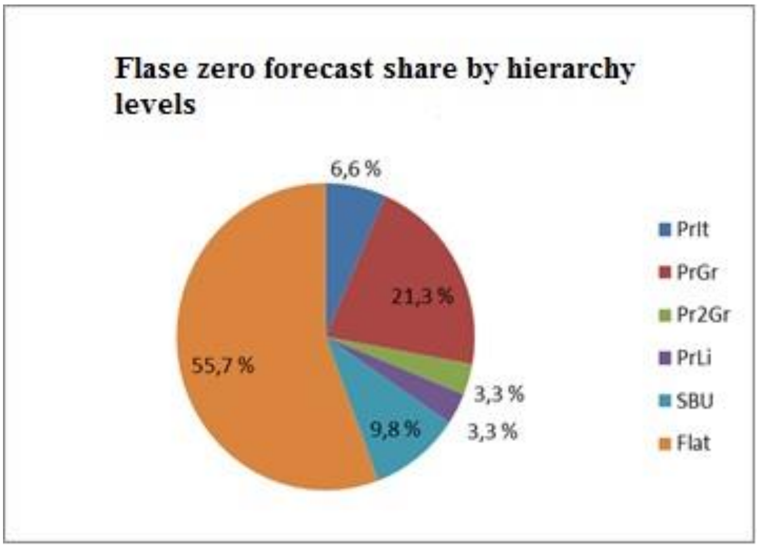


Figure 4.6. Average number of false zero forecasts displayed on the product hierarchy levels

About 10% of all forecasts are false zero on average over the past two years. Due to the fact that about 50% of the forecasts are based on the naive method, the largest absolute value of false zero error is also observed in the hierarchy group "Flat". Figure 4.6 illustrates the average percentage of false zero forecasts for end products sales, committed between May 2012 and February 2014, at all determinant hierarchical levels. Pie chart is divided into sectors in accordance to belonging to one or another determinant hierarchy level. False zero values of the forecast cannot be quantified by average percentage error measure. Accounting such errors is possible only by calculation of the absolute deviation of the forecast. However, the absolute error measures make it impossible to obtain reliable average error values for each month, as well as a comparison of forecast errors among the products and their groups due to the fact that all the products are very different from one another by sale volumes, expressed both in euros and pieces.

4.1.5 Conclusions from the analysis of forecasting model errors

There are no predictions that have absolute 100% accuracy. The error takes place in all kinds of forecasts without any exceptions. However, total forecast error usually consists of systematic and random components. Analysis of erroneous forecasts statistics helps to identify weak points of the model, which may be considered as potential sources of systematic error. Model will have a very high accuracy rate when the system error will be reduced to zero. Demand forecasting for companies from the industrial sector is especially challenging task. Successful forecasting models with high accuracy rates help to reduce average inventory levels of products in stock, as well as significantly improve the level of customer service.

To investigate the accuracy of the current model developed by company X two accuracy measures were selected: the determination coefficient and the mean absolute percentage error. As a result, some graphs were built and analyzed. These diagrams illustrate the dynamics of the determination coefficient on monthly basis, average monthly values of absolute percentage forecast deviation from reality, as well as some other important model data. According to the results of the error analysis the following conclusions were formulated:

1. Forecast has higher accuracy when model captures a significant correlation between sales on one of the hierarchical levels. The maximum index of seasonality in the time series must satisfy the constraints, in order to take into account the seasonality when calculating the overall forecast value. Thus, the figures that were originally placed in the table with the constraints play a key role in determining its accuracy.
2. Only a small share of the products is characterized by the significant seasonal sales behavior at the level of end products, but their forecast has the highest accuracy rate. The most common determinant level is the level of product groups. Therefore, sales data at this level should be given appropriate attention.
3. Adjustment of the forecast, which is based on the product life cycle stage, has a beneficial effect on the accuracy of the final forecast.

4. If the sales history of some end product has a random abrupt in the time series, this occasion directly affects the ability of the model to identify the seasonal profile and capture its seasonal parameters.
5. Mean absolute percentage error of the forecasts, excluding cases of false zero values amounted to 130%.
6. 10% of all the forecast deviations from actuals were caused by the problem of false zero forecasts. The second largest number of false null predictions after the «Flat» level happened on the level of product groups. Current forecasting algorithm is not able to resolve this issue.

These findings formed the basis for further study of possible ways of increasing the accuracy of X's forecasting model. Possible refinements are described in the next section of the chapter.

4.2 Possible directions for model accuracy improvement

This section considers the following ways to improve the accuracy of the current forecasting model:

1. Table of boundary limits;
2. Identification of seasonal behavior;
3. Forecasting the demand for long lead time products;
4. Final stage of forecast calculations;
5. Record of previous forecast errors;
6. The human factor in forecasting;

All forenamed points will be discussed in details from different angles of the research. Finally, practical guidelines for all the improvement directions will be formulated and summarized.

4.2.1 Table of boundary limits

One of the key conditions that effects the accuracy of the forecast is the table of limits (table of the boundary limits is described in the second chapter of this thesis). It contains values of the correlation coefficient for sales and the corresponding maximum individual indices of seasonality that are considered to be significant, calculated for each of the five hierarchical levels of the product. The minimum acceptable level of correlation is 50%. Minimum seasonal index for the level of the end products with a correlation coefficient of 50% -70% originally was calculated using the Solver in Excel. Further, all the other cells in the table of parameters were obtained by multiplying the calculated index for the end products and a special multiplier. Multiplier, in turn, was also calculated using Solver. This means that all the values in the table of boundary limits are linked linearly. Table parameters were not revalued for about two years. Meanwhile, many products are no longer manufactured and were replaced by others. Customer base has been updated during this time as well. The maximum allowable individual indices of seasonality may have changed. Thus, we can assume that one possible way to increase the accuracy of the current forecasting is to update the table of limit values, or even update their calculation algorithm.

4.2.2 Identification of seasonal behavior

Seasonal nature of sales is determined by the correlation coefficient. Coefficient is calculated for two time series of seasonal chain indices of the past years; this coefficient can be considered as key in the whole forecasting model used by the company. “Seasonal fluctuations in sales” is one of the most popular research topics in the demand planning and sales forecasting fields. Indeed, there are seasonal jumps or, on the contrary dramatic slumps of sales in many types of businesses. Today there are two basic ways to capture such seasonal fluctuations - the method of peer review and analysis of the previous sales statistics. The method of peer review is a survey for the employees that have some knowledge in the field of sales of certain product groups. Employees may indicate the upward trend or decline in sales dynamics and approximately numerically evaluate these changes. In X case employees responsible for production planning and procurement on a particular plant do have this knowledge. Series of interviews on three plants of the company proved that personal opinions and intuition of employees still affect the results of the final forecast figures. While carrying out the production planning managers make their personal adjustments in the forecast data manually. However, employee's opinion about the seasonal nature of fluctuations in sales is subjective; this vision may differ among the employees of the same company. Experts are able to indicate trends in different time periods much easier, but much more difficult is to assess them quantitatively. In any case, in order to improve the accuracy and efficiency of current method of identification the relationship between annual sales, it is strongly required to adhere to a certain level of formality in this procedure.

Company X uses statistical methods of analysis to detect the seasonal behavior that provides certainly more objective assessment. Calculated series of indices that show rises or drops, serve as trend in the current model. In turn, the model interprets the multi-directional outliers as random fluctuations. Presumptive period of the seasonal fluctuations for all the products in this model is equal to twelve months. However, some graphs depicting the dynamics of seasonal indices for certain products show that the period of seasonal variations is shorter than one year. This can be explained by the frequency of orders by regular customers. Any periodic dependence between the elements of the time series can be formally determined by the autocorrelation function.

It is assumed that the time series are stationary, which means an independence of autocorrelations from time. Thus, an assumption that possible period of oscillations in sales of all product range is twelve months brings a certain error in the final forecast in advance. It should be noted that if there are several independent areas in business, such as different types of customers: commercial and budgetary, retail and wholesale, - you must develop an appropriate forecasting scheme, which separately considers the specifics of each customers. It is quite natural because different kinds of customers buy the same product in different ways with varying intensity and speed.

It should be also noted that in fact, there are a fairly large number of events and factors that may have an impact on the dynamics of sales in addition to seasonal fluctuations. Irregular outliers affect the possibility to capture interdependence between the time series of sales data. In this case, if the calculated index of seasonality for a particular month exceeds the maximum, which is shown in the table of limits, it will not be accepted by the model and, therefore, increase of sales in this particular month will not be captured, even if the correlation coefficient is significant enough. Due to significant changes in the activity of the enterprise or influence of external disturbances some initial data should be discarded. Use of chain individual indices of seasonality in some cases is not enough for efficient smoothing of irregularities in the time series.

4.2.3 Forecasting the demand for long lead time products

Most of the well-known models described in the literature are able to cope quite successfully with the task of the demand forecasting for various products, which are characterized by a relatively short time of delivery. In this case, end products, which are assembled from materials with the delivery time longer than few weeks, are problematic for preparing the next month forecast. The current model will cope successfully with the forecasting for products that have delivery time a week or ten days. In this case all the materials that are required to produce the desired amount of certain end products with long lead time can be delivered just in a month. Thus, next month forecast for such products is not effective. Seasonal indices for these products should be based on data of lengthened periods, as the ratio of sales over the next six months to sales of the past eight months. However, this solution entails the problem of increasing the inventory levels and as a consequence increases the costs (Effective Inventory Management 2013).

Current model does not take into account lead time parameter. Majority of computer systems store planned lead time as a single parameter. But in fact, this time consists of several components (Shraibpheder 2005):

- time needed for the order preparation and it's submission to the supplier;
- time required for the production, packaging and shipment of goods;
- time for transportation of goods from the supplier to the warehouse;
- time required for receiving the goods, unpacking and preparation for use;

One of the solutions lying on the surface is the registration of lead time parameter in the forecasting model for the whole range of products. This innovation requires significant structural changes in the current model. Therefore, in order not to disrupt the current process radically, demand forecasting for products with long delivery time may be submitted in a separate model.

If the model is strictly based on the precise values of lead times, the company may face a new problem - delivery dates variability, caused by irregular actions of suppliers or other unexpected circumstances. To solve this problem, the company must share information with suppliers and interactively adjust the lead time data (Wallace & Stahl 2008).

4.2.4 Final stage of forecast calculations

Of course, the accuracy of the final forecast directly depends on the hierarchy level where the model has captured a sufficient rate of correlation. The higher determinant level of the product hierarchy, the lower accuracy of the forecast. However, if the correlation cannot be detected at any level, the forecast will be entirely based on the sales data of the past four months. Sales forecast for the next three months is calculated as three times the average product consumption during the previous four months. Minimum significant correlation coefficient between two annual time series of seasonal indexes is 50%. This coefficient is statistically significant with a probability of 90%. On a Cheddock scale closeness of the dependence between parameters that exceed 70% means high dependence between two time series, and values greater than 90% - a very high (HELPSTAT 2012). In turn, a correlation of 30% -50% can be considered moderate. Calculation of the minimum statistically significant correlation coefficient for the annual time series showed that 35% correlation is statistically significant with a probability of 80%. If the correlation rate does not exceed 70% the value of the determination coefficient R-square will be less than 50%. It means that share of one variable variations is smaller compared with other factors that affect the variation of the total dispersion of another variable. Thus, one of the most interesting areas of the model refinement will be the analysis of its accuracy at various acceptable minimum level of correlation between the annual time series of sales, because these values significantly affect the selection of the determinant level of the product that will be used by the model for the calculation of the final index of seasonality for the next period (Orlova 2000).

Looking up the logical algorithm the model based on, we should pay attention to the next phase of the direct calculation of the forecast for the next three months. Model forecast figure is calculated as the product sales of the past four months multiplied by the arithmetic mean of two seasonal indices of the corresponding month in the last previous two years at that level of the hierarchy, where all the coefficients satisfy the table of boundary limits. Many theoretical approaches of forecasting suggest taking into account the fact that the most recent data adequately reflect the current situation. Thus, when calculating the forecast index of seasonality for the future period it is reasonable to give greater weight to the last year's index and lower weight before last year. Direct calculation of the weights for the sales data can be based on the difference in total volume of sales and many other factors. Using the weighted average seasonal index for the past two years may improve the accuracy of forecasting.

While analyzing the model features of X it was noticed that the model gives an erroneous forecast in case of irregular sales. Appendix 4 provides a table with calculations of seasonal indices for false zero forecasts. The first line represents the volume of ST206.3 product sales during 2012, as well as the last four months of 2011. Suppose that forecast is calculated in the beginning of November 2013. The correlation coefficient between sales volumes in 2012 and 2013 meets the restrictions in table of limits; product had sales during four months in both years. Due to the fact that the total sales in the four-month period from 07.2012 to 10.2012 are equal to zero individual seasonal index calculated for November 2012 is 0.75. The model is designed so that in case of any calculation errors Excel takes a value of 0.75, which means that sales in the next three-month period will be equal to three times the average sales in the previous four-month period. Thus, a sudden jump in sales of the product ST206.3 in November will not be captured and simulated by model. This drawback of the model is one of the causes of the forecasting accuracy reduction. In order to avoid erroneous forecasts, it is necessary that the model could recognize single spikes in sales of products, which are characterized by similar irregularities.

4.2.5 Record of previous forecast errors

Any additional information regarding the dynamics of sales may be useful for improvement of the demand forecasting model accuracy. Each monthly forecast contains information about its error. Overall forecasting accuracy rate of the previous forecast is calculated in Excel worksheets of X's forecasting model. The overall accuracy is evaluated by the coefficient of determination for the entire data columns with forecasted and actual values of sales. However, end users of the model cannot apply the knowledge about forecast errors in practice. The process of production planning and procurement, which is based on short-term forecast data, requires much more detailed knowledge of the structure and nature of the forecasting errors hidden in the algorithm. Model structure and its internal architecture determine the accuracy of future predictions. It is important to concentrate not only on forecast calculation, but also on thoughtful analysis of the potential reasons for the deviation of the forecast from the fact. This is important because quickly finding the cause of the deterioration of the forecast quality, the expert will be able to make adjustments to the forecast and improve its accuracy.

Modern forecasts often look very convincing due to the use of different mathematical methods. However, they all are based on the same simple ways of extrapolation of currently available data about the world to the wrong judgments about the future, assuming that the world does not change so much. The reasons for the deviation of the forecast from the actual values can be divided into three groups:

1. System error. There is no single model with 100% accuracy. Systematic error - error caused by the characteristics of the model and the specifics of the analyzed process.
2. The error caused by wrong actual figures that are used by the model as an input. The monthly sales may not match actual demand for the product.
3. Random error. This error is a random component of error measurement that varies without any control from time to time.

Accumulated statistics on model deviations from the actual values may be included in the model structure. Various practical solutions of forecast correction by error value are described in literature. Such kind of probabilistic models use a variety of approaches to take into account all of these types of errors. Current forecasting model does not simulate potential error of future forecast.

4.2.6 The human factor in forecasting

As it was mentioned earlier, human factor plays an important role in the current forecasting procedure in company X. The final stage of the logical forecasting algorithm is manual forecast figures adjustment made by experts when it is needed. Analyzing the results, expert compares the model values with the calculated maximum allowable value of sales in the next period that is based on information about the position of the product on the life cycle curve. Conducted interviews with employees of the company have showed that the database where the life cycle codes information is stored usually is updated less frequently than once a month. In turn, the forecasting model is run monthly. This lag can introduce some additional error in the forecast.

As a result, when the sales forecast for the next period has been formed, it is interesting to track how it is actually used by employees in reality. End users of the forecast, staff responsible for the operational planning of production and procurement make personal adjustments to the forecast before loading the forecast into the system. "Oedipus effect" is typical for any kinds of forecasting activities, it says that the forecast can be "self-fulfilled" or "self-distracted" by taking targeted decisions and actions. This effect is named after the ancient Greek King Oedipus, who learned his future from one oracle: that by killing his father, he marries his mother; and he will have children, cursed by the gods. Oedipus tried to avoid the predicted fate, but fate was stronger: he made a chain of fatal mistakes - and all came to pass. Thus, if a forecasting model makes certain mistakes, Planning Manager no longer trust its values. As a result, making personal adjustments to the forecast, thus they can further increase the deviation of the forecast from the actual values (Mainzer 2004).

4.3 Final list of possible ways to improve the model accuracy

Many Russian and foreign researchers emphasize the importance and necessity of forecasting in companies' performance. For this purpose, for example, US firms spend about 1% of total profit on forecasting research. Profit from the systematic use of forecasts is 40-50 times higher than the cost of the forecasting process implementation itself. Modern development of market interactions and increasing competition between the companies raises the importance of forecasting activity. Without the ability to predict the future it is impossible to perform normally and get the maximum profit. Nowadays price of potential damage from reckless decisions repeatedly increases. Thus, the high accuracy of the forecasting model is a major competitive advantage for the company, which allows management to plan company's expenses efficiently, as well as to provide a high level of customer service (Zharikov & Goryachev 2013).

Analysis of internal structure of the current forecast model of company X and model's error statistics revealed some significant deficiencies that adversely affect forecast accuracy. All the deficiencies of the current model described above have been collected and summarized in a single list of possible ways of increasing the forecast accuracy:

1. Review and update the maximum permissible values of seasonal indices in the table of limits, as well as their calculation algorithm.
2. Improvement of seasonal behavior identification procedures.
   - 2.1. Gathering and analysis of knowledge about the specific features of seasonal behavior of certain products sales among employees. Subsequent formalization of forecast adjustment procedures on the basis of expert knowledge.
   - 2.2. Detailed analysis of seasonal fluctuations, with additional mathematical approaches, for example, such as autocorrelation analysis.

2.3. Consideration of customer's specific characteristics: commercial and budgetary, retail and wholesale.
2.4. To develop the procedures for cleaning the initial data from potential accidental outliers.
3. Development of a mechanism for accounting the lead time of product in the model. Introducing the lead time parameter into the forecast calculation logic.
4. Adjustment of calculation logic in the model.
4.1. Accuracy testing for the forecasts with different rates of the minimum allowable correlation coefficient between time series.
4.2. Using the weighted average index of seasonality for past two years as forecast index.
4.3. Development of a method for detecting single spikes in product sales dynamic, which are characterized by irregular nature of sales.
5. Simulation of the forecast error based on the accumulated statistics on forecast deviations from real sales. To develop the procedures of the forecast correction based on the systematic and random error components.
6. Monitor and control all the adjustments that experts manually introduce into the forecast figures, as well as end-user of the forecast. This kind of actions must be formally documented in order to further improve the forecasting process.

All ways of model accuracy improvement described above, of course, require further research and testing. Within the practical part of this research, several new versions of current forecasting model were developed. One version includes a mechanism designed to identify and eliminate false zero forecasts. The second test model uses weighted average indices of the seasonality of the past two years in the calculation of forecast index. The third version of the current model was supplemented with a mechanism of forecast correction based on the identified systemic errors for specific end products. Each of the test versions has been tested on real data of past sales. Their accuracy then was compared with the results of the current model. All of the above is described in the fifth chapter.

5 DEVELOPMENT OF TEST VERSIONS FOR THE CURRENT FORECASTING MODEL

This chapter describes three test models that are based on some of accuracy improving suggestions listed above. All test models have been tested on sales data of past years; accuracy of their forecasts has been compared with the forecast accuracy of the basic model.

5.1 Error correction add-in

While analyzing the model errors, it was observed that the value of error for certain products has being changed with a seasonal behavior. Classical methods of forecast correction based on the data of the previous errors were described in details by researchers Clements and Hendry (1998), Entov and Nosco (2002). These methods allow adjusting the forecast in a real time. The standard technique includes two adjustment methods based on the information about one-step forecasting error (Turuntseva 2013). Accordingly, the forecast is usually adjusted by the forecasting error on the previous step, or the value of the average error of all the previous one-step forecasts. The first method involves recalculation of the forecast figures depending on the forecast error of the previous step (see Equation 5.1), (Clements & Hendry 1998).

$$\tilde{f}_{T,1} = f_{T,1} + e_{T-1,1}, \qquad (5.1)$$

where $\tilde{f}_{T,1}$ – corrected one-step demand forecast for the next three-month period, at the moment T, $f_{T,1}$ – one-step forecast for the future period $T+1$, $e_{T-1,1}$ – forecast error on the previous step. In turn $e_{T-1,1}=y_T - f_{T-1,1}$ – difference between actual sales volumes during the current period of three months T and it's forecast at the moment T, calculated at the moment $T–1$.

Second classical method of error correction – forecast adjustment by the value of an average error of all known previous one-step forecasts – that means forecast correction by the value equal to (5.2), (Clements & Hendry 1996).

$$\frac{1}{s}\sum_{i=1}^{s} e_{T-i,1}\ , \qquad (5.2)$$

where $e_{T-i,1} = y_{T-i+1} - f_{T-i,1}$, $i = \overline{1, s}$, y_{T-i+1} – actual demand at the moment $T–i+1$, $f_{T-i,1}$ – forecasted demand, calculated at the moment $T–i$ for one step ahead.

Thus, the adjusted forecast is calculated by the flowing formula (5.3).

$$\tilde{f}_{T,1} = f_{T,1} + \frac{1}{s}\sum_{i=1}^{s} e_{T-i,1} \qquad (5.3)$$

According to the results of research conducted by Entov and Nosco (2002) second method can achieve significantly better results compared with the first method. It was proved on the majority of Russian macroeconomic indicators that took part in the research. This paper also noted that both methods lead to the removal of systematic prediction error, if present. But unfortunately the quality of the predictions often worsens after the conducted adjustments (Mazmanova 2000), (Dubova 2004).

The main problem that occurs using these two methods of forecast correction is the practical difficulty of using them in "real time" in case of current forecasting model. Model produces the forecast for the next three-month product consumption every month. Thus, the company utilizes this three-month forecast during one month only. Therefore, to correct the predicted demand by standard methods becomes problematic. Given the fact that the forecasting model used by the company X is based on seasonality identification logic, new approach of taking into account seasonal errors was suggested. This method implements certain adjustments to the forecast figures based on all the available information about the previous forecasts and their errors.

To analyze the seasonality of errors occurred in the model, certain forecast files and all the available data of actual sales were taken in consideration. Due to the limited information available about the forecasts, assessment of the relationship between the errors was conducted with two time series, each with a length of eight months. Only those products that were sold steadily throughout the year participated in the analysis. As a result, the document was created where each product is presented by two sets of data: time series of three-month forecast figures and time series of actual volumes of sales for the corresponding three months. Sales are expressed in euros. Next, time series of coefficients e_t were calculated for each end product. These coefficients represent the ratio of the forecast to reality. With full compliance of the forecast and sales ratio e is equal to 1 (5.4).

$$e_t = \frac{Y_t}{\sum_t^{t+2} X_i}, \quad (5.4)$$

where Y_t − forecast for month t, X_i − actual volumes of monthly sales

If the variation of the coefficients shows seasonal behavior on the whole time period under review (add-in captured close relationship between the time series of coefficients), future forecasts for this particular product is adjusted on the basis of past mistakes. Figure 5.1 shows an example of identified seasonal behavior of coefficients e_t dynamics for the particular product EFPPH4. Correlation coefficient between forecast errors was statistically significant, equal to 0.94 that proves that there is a close seasonal dependence.

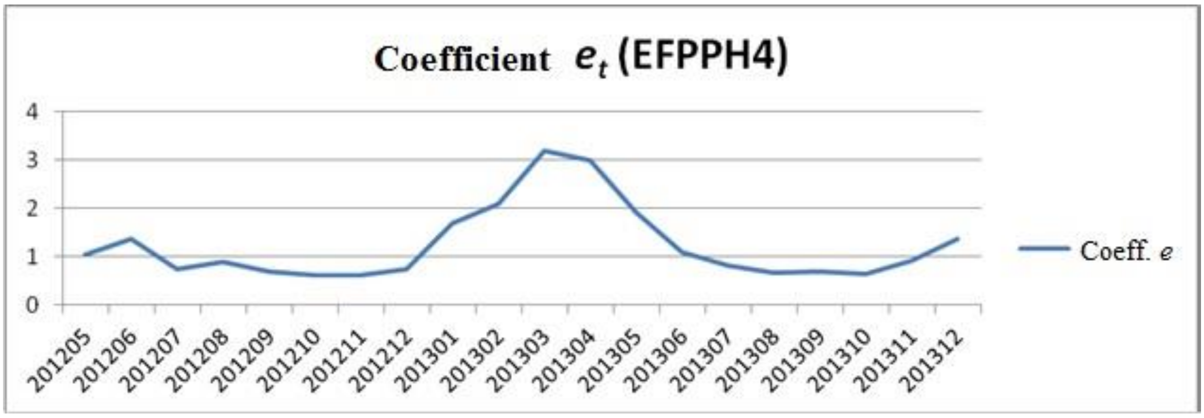


Figure 5.1. Dynamics of the relative error for the product EFPPH

Forecasts made in a particular month T, was adjusted by the average forecast error of five months from the respective T-2 to T + 2 months of the previous year (5.5). For example, the forecast calculated in the beginning of January 2014, will be adjusted by the average forecast error of the period from November 2012 to March 2013.

$$Y'_t = \frac{Y_t * 5}{\sum_{t-2}^{t+2} e_i}, \quad (5.5)$$

where $Y't$ − corrected forecast, Yt – original forecast

Designed for the current model add-in of the forecast correction by the value of a systematic error has been tested on the baseline forecasts of January, February and March 2014. The accuracy of the adjusted forecasts was compared with the original model forecasts using the coefficient of determination measures and measures of absolute percentage error. Usage of this forecast correction add-in helped to increase the accuracy of three basic forecasts markedly, Table 5.1. Determination coefficient became closer to 1 by almost 10% in average, which indicates an increase in the proportion of explained variability of relevant variables. The value of mean absolute percentage error decreased by more than 25%. This model development makes it possible to achieve a significant increase in the forecast accuracy for those products that are characterized by seasonal behavior of their forecast errors.

Table 5.1. Results of error correction model testing

	January 2014	February 2014	March 2014
Original forecast	$\mathbf{R^2}$ = 76% **MAPE** = 55%	$\mathbf{R^2}$ = 68% **MAPE** = 65%	$\mathbf{R^2}$ = 64% **MAPE** = 81%
Corrected forecast	$\mathbf{R^2}$ = 85% **MAPE** = 29%	$\mathbf{R^2}$ = 82% **MAPE** = 28%	$\mathbf{R^2}$ = 71% **MAPE** = 66%

It is important to note that a positive result was observed in the analysis of relatively small amount of available sales data. Increasing the size of data array on actual sales volumes, as well as number of forecasts will help to clarify the results of testing procedures. Since June 2014, it will be possible to correlate two annual series of relative error rates that was considered as is one of the potential directions of further research.

5.2 Forecasting model with weighted indices of seasonality

Accounting seasonal fluctuations causes errors reduction when calculating theoretical values of company performance indicators and their forecasts. Making model more accurate helps to make a difference gap between model and real system smaller. As it was already mentioned before, identification and evaluation of oscillatory processes can have a significant influence on the resulting. Company X uses the average seasonal index of two previous years studying seasonal variations in the sales data. But it is recommended to use method of weighted seasonal indices when modeling a serious economic object (Mazmanova 2000). In this case, the individual indices of seasonality can be averaged by calculating weighted average values. The weights can be set manually by the expert, fit mid-monthly or mid-quarterly levels of each year, or defined in any other way.

As a part of the practical part of this study, it was decided to develop a version of current forecasting model where forecast seasonal coefficients are calculated as a weighted average measure. Weights of significance were set manually on the basis of the assumptions that more recent sales data adequately illustrates current and future situation. Thus, the importance of a seasonal index of last year relates to the significance of index year before as a 3:2. This particular version of the forecasting model has been tested on all available data on actual sales. As a result, twelve new monthly forecasts for 2013^{th} year were composed with weighted indices logic. The accuracy of the new test version of the model was compared with the accuracy of the original model by the coefficient of determination measure and measure of mean absolute percentage error. The results of the comparison are displayed in Table 5.2. Overall accuracy of the model with average weighted index of seasonality significantly deteriorated. Determination coefficient of the forecasts improved only in three months out of 2013 year. And the value of mean absolute percentage error MAPE increased by 30% in average. Thus, we can conclude that the use of mean arithmetic indices of seasonality, that means equal weights of significance for past two years, helped to smooth some accidental outliers in the data array. Thereby, it can be concluded that weights of significance of seasonal indexes last year and year before for each of the end products or groups must be individualized.

Table 5.2. Results of the accuracy tests for the model with weighted average indices of seasonality

Month of 2013th year год	R^2 Ogirinal model	R^2 Weighted average model	*MAPE* Original model	*MAPE* Weighted average model
January	0,78	0,79	1,73	2,14
February	0,61	0,62	2,37	2,91
March	0,77	0,76	1,34	1,70
April	0,74	0,80	1,05	1,31
May	0,72	0,71	1,10	1,37
June	0,73	0,72	1,05	1,31
July	0,82	0,82	1,13	1,35
August	0,82	0,81	1,08	1,33
September	0,82	0,81	1,21	1,51
October	0,84	0,84	1,02	1,30
November	0,78	0,75	1,23	1,60
December	0,80	0,80	1,21	1,58

Currently, one of the most promising areas of time series forecasting research is the use of adaptive methods. When processing the data, usually the last recent time period contains the most valuable information, as it is necessary to know how the current trend will develop in future. Sales forecast for the next three months in company X is based on the most recent data on past sales during the last four-month period. New sales data on a monthly basis substitute the previous calculated forecast. However, as has been described previously, forecasted seasonal indices for all the forecasts are based on old data of the corresponding months of last two years with equal significance rate in the current model. Adaptive methods allow taking into account an informational value of different levels of the time series. Within adaptive methods different values of time series levels depending on their past "fallacy” can be taken into account with the help of special system of weights assigned to these levels.

Evaluation of adaptive model coefficients is usually done on the basis of the recursive method (Dubova 2004). The most important advantage of adaptive methods is an ability to build self-correcting models that take into account the achieved accuracy of forecasts composed on the previous steps. The deviation of the forecast from the actual sales volumes for each of the end products from May 2012 until March 2014 have been already calculated. It is evident that the forecasting error for the same month but different years differs from each other. In the current model, the value of past mistakes is not included in the calculation of the forecast index of seasonality, and the total index is calculated as the average for both years. One possibility to improve the quality of forecasting is to develop an adaptive model, in which the forecasted index of seasonality will be calculated as weighted average index of two years; the weights of significance will be inversely proportional to the values of the corresponding forecast errors. Thus, forecast error data is included into the model structure through feedback link, in accordance with the procedure of transition from one state to another. Feedback link helps to adjust model parameters to ensure greater harmonization with the dynamics of the time series. Thus, adaptation is performed iteratively immediately as there is a new point in the actual time series curve. Figure 5.2 shows the general scheme of forecast figures calculation for each of the end products within the adaptive model.

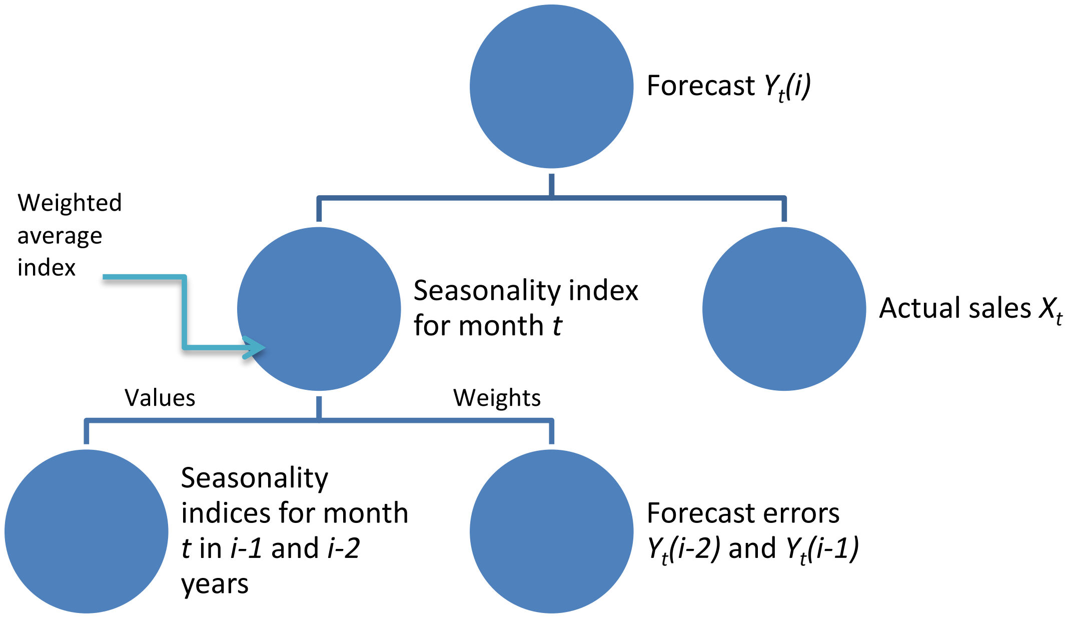


Figure 5.2. General scheme of forecast calculation within the adaptive model

5.3 Model with false zero forecast correction

Seasonality of sales is tracked by the current forecast model using a chain of individual indexes of seasonality. The system of chain indices - series of indices calculated with the variable base index of comparison. Thus, the sequence of chain indices reflects the changes in the levels of sales over time more accurately than the baseline. In the forecasting model of X each month has its own basis, which is equal to the volume of sales over the previous four months. In the section of the study about the potential areas of model accuracy refinement one major drawback of this model was described, which deals with the intermittent sales forecasting. If the demand for a certain product was completely absent for four months, which are the basis for calculating the index of seasonality for the relevant month, it is difficult to assess the jump in sales in the next period if present. Excel treats similar situations as an error of division by zero, and the model assigns the seasonal index value of 0.75, which says that the sale of the next period will remain unchanged compared to the previous. Further, if the nature of the sales of a particular product has been considered as seasonal by model, zero sales multiplied by 0,75 coefficient of seasonality will result in a false zero forecast. The first mistake made by the model - setting the value of chain seasonal index equal to 0.75 for the problematic month. However, even if the index indicates the significant seasonal jump in sales compared to the previous period, the result of multiplying it by zero will always give zero forecasts. Thus, the second mistake in such case is multiplication itself with zero four-month sales in the calculation of the final value of the forecast.

To solve this problem, multiplicative calculation method was replaced by pseudo-additive, which is a combination of certain elements of both additive and multiplicative models (Australian Bureau of Statistics 2005). In this case, if the model decides to create forecast of sales based on the average seasonal index and sales volumes over the past four months equal to zero, forecast is equal to the volume of sales of the corresponding three-month period last year. In fact, sales of the corresponding months in the past two years represent the trend component of the sales forecast for the next three months. Thus, the ratio of current sales to past four-month sales is the coefficient multiplier that reflects the total annual change. New formula (5.6) of the forecast value calculation is listed below. Past sales in this case are average figure of the past two years.

$$A'_3 = B_3 - \left(1 - \frac{A_4}{B_4}\right) * B_3 \,, \qquad (5.6)$$

$B -$ past sales, lower subsripts 3 and 4 mean number of months, $A' -$ forecast

If the A4 and B4 are zero at the same time, their ratio should be taken equal to 1. Test version with a modified algorithm was tested on sales data 2011-2012. Further created forecasts for each month in 2013 have been compared with the results of the original model according to certain criteria of accuracy. Then a new determination coefficient and the number of corrected zero false forecasts were calculated for each month. The results of the comparison are shown in Table 5.3.

Table 5.3. Testing results of model version with false zero forecast correction

Month	R^2 (Original)	Number of corrected zero false forecasts *(Original)*	R^2 (Model version)	Number of corrected zero false forecasts *(Model version)*
January 2013	0,78	103	0,79	77
February 2013	0,61	125	0,61	88
March 2013	0,77	140	0,77	102
April 2013	0,74	156	0,80	121
May 2013	0,72	112	0,73	85
June 2013	0,73	114	0,73	83
July 2013	0,82	136	0,84	98
August 2013	0,82	165	0,84	127
September 2013	0,82	127	0,83	83
October 2013	0,84	129	0,84	108
November 2013	0,78	146	0,80	113
December 2013	0,80	150	0,80	133

Despite the fact that the coefficient of determination hardly changed, the number of false zero predictions decreased by 40% in average. The remaining share of zero figures is the result of random fluctuations in demand. The measure of accuracy R2 remained unchanged compared with the results of the basic model for the reason that tested model overestimated demand quite often. One of the main competitive advantages of the company is to supply high quality products to customers in the most minimal delivery time. Following this goal the continued availability of the relevant products in stock is required. Thus, in this situation, for the company it is better to overestimate the demand, but to predict the presence of sales in the next period, instead of trusting to zero value and input this false information into the system (Lindsey & Pavur 2013). Especially, if the predicted value is significantly greater than the maximum allowable amount of demand for a particular product, the forecast will be adjusted manually by an expert. In case of a false zero values model will not notify an expert about any significant errors.

5.4 Summary

In the practical part of the study three test models have been developed, that are practical implementations of some of the directions of the current model refinement formulated in the fourth chapter. Comparison of the accuracy that test versions showed accuracy of the original model revealed the following:

1. System errors correction model helped to increase the coefficient of determination by an average of 10%, and the value of the mean absolute percentage error by 25% for those products, whose dynamics of relative deviation from the real sales was defined as seasonal.
2. Testing the model with weighted average indices of seasonality did not give any positive results; the accuracy of the forecast figures has been deteriorated. Based on the analysis of possible reasons for the lack of positive results new perspective direction of model development has been found. It is an adaptive model that takes into account data on the mistakes of the past forecasts.
3. The model of false zero forecast correction reduced the number of such errors by 40% in average.

The final chapter presents a synthesis of the work done, and it outlines possible areas of practical application of all the research results.

6 FORECAST CALCULATIONS IN MATRIX FORMAT

In order to improve the accuracy of the model, all previous sections of the research were mostly focused on forecasting methods, their features and forecasting logic. On the other hand, technical aspects of the model-file itself play important role in the whole forecasting process. Model of company X is developed in Excel. Calculation environment (software) imposes some initial restrictions on the model performance. System where the model was created predetermines the speed of the calculations, user interface and variety of other issues. It usually takes one working day for a responsible employee in company X to download an initial sales data, run the model and complete the final forecast for the next three-month period. Model is constantly freezes, making it difficult and time-consuming to progress with the forecasting. The time issue required for preparing the monthly forecast is considered as a disadvantage of the model. Nowadays, matrix calculations became a quite popular among researchers from different scientific fields. Thereby, transferring the current model without any changes in its logic into another calculation environment can be considered as one of the additional improvement directions in context of the proceeding time of forecasting. This chapter provides a description and comparison of computation features in matrix-based software and Excel, presents a practical solution to implement the current forecasting algorithm in matrix form; in the end it gives an assessment to the feasibility of such a transfer.

5.1.Calculation environments

There are different programs that deal with matrix calculations, both open and commercial software. To develop the matrix version of the current forecasting model of company X, Matlab has been chosen as a potential computing environment for this purpose.

MS Excel – is a quite famous and probably the most popular program, a business application developed by Microsoft that deals with spreadsheets using a graphical interface. It is characterized by intuitive interface, outstanding computation capability and excellent chart tools. Almost everyone without any serious knowledge can start to use it; it meets the need of any industries, any enterprise, or any workspace environment. Excel is considered to be user-friendly software. Excel users can name the variable after the cell, just as natural as that, not requiring any effort to define the name like other tools do. Excel cells are aligned by nature, saving the effort of typesetting. It allows for invoking other cells intuitively and calculating automatically, capable to implement the stepwise computation easily. However, on the other hand, the great universality of Excel makes this program suffer from the relatively poor specialty. With a cell as a unit, the support of Excel for structured data is quite poor. It is not flexible, number of functions available is limited and also there are some restrictions on the number of rows in tables. Excel does massively break down at over 10 000 rows. The functional capabilities are rather simple and the representation ability of Excel syntax is not great enough to handle the complex data analysis and specialized scientific computation. Excel is great when it is important to keep a table with the data (input, output, and descriptions) in front of your eyes at all times. Highly formatted output is inherent in this tool. Excel is good for explaining things to business people and in presentations due to its “visual” nature. This is related to the fact that all the numbers are visible on time.

Matlab is a business application developed by MathWorks. It is an interactive computing environment and fourth generation programming language for numerical computation, algorithm development, and data analysis. It also provides some plot graphics and charts. Matlab is widely used in the industrial automation design and analysis, and other fields like the image processing, signal processing, communications, and finance modeling and analysis. MATLAB is best for sophisticated math, especially on large data sets and for things like matrix algebra, differential equation integration, Laplace transforms in process control and others. Getting complex tables of formatted output in Matlab can be a hassle though. The fact that researcher is not able to see all the numbers in RAM always when running to program actually is one of its big advantages. Large groups of numbers are manipulated in Matlab by simple one-line commands, where the variable can be huge matrices with thousands of numbers. Also there is a possibility to develop user-applications based on Matlab calculations through GUI, also known as graphical user interfaces. GUI provides point-and-click control of software applications, eliminating the need to learn a language or type commands in order to run the application (Mathworks 2014). The important question is not which one is the best. Both calculation environments are very good and they both can be used for most of the things, research and business tasks. They are different and each is better for some tasks than others.

5.2.Matrix forecasting model

In general, the entire forecasting process in company X consists of the following stages:

1. Recent sales data loading and cleaning;
2. Forecast calculation in accordance with the model logic;

3. Forecast manual correction;
4. Forecast explosion into the components and materials;

Of course, model transfer from one program to another can possibly confront with certain technical difficulties at each of these stages. At this stage of the study, we are interested mostly in the transfer of all the technical calculations into the matrix format and evaluation of its performance in comparison with original model.

Calculation logic of the current model was described in details in the second chapter of this study. Briefly in general, first of all forecasting model calculates the series of seasonal indices for all products on five hierarchy levels, secondly, it calculates the correlation coefficients, identifies maximum index in the time series, then it compares these pairs of numbers with the limit values and finally assigns forecast index for the next time period. When transferring calculations in the Matlab environment, we should think of matrix format for data storage. Figure 6.1 represents the sequence of calculation steps and flows of data from one stage to another one.

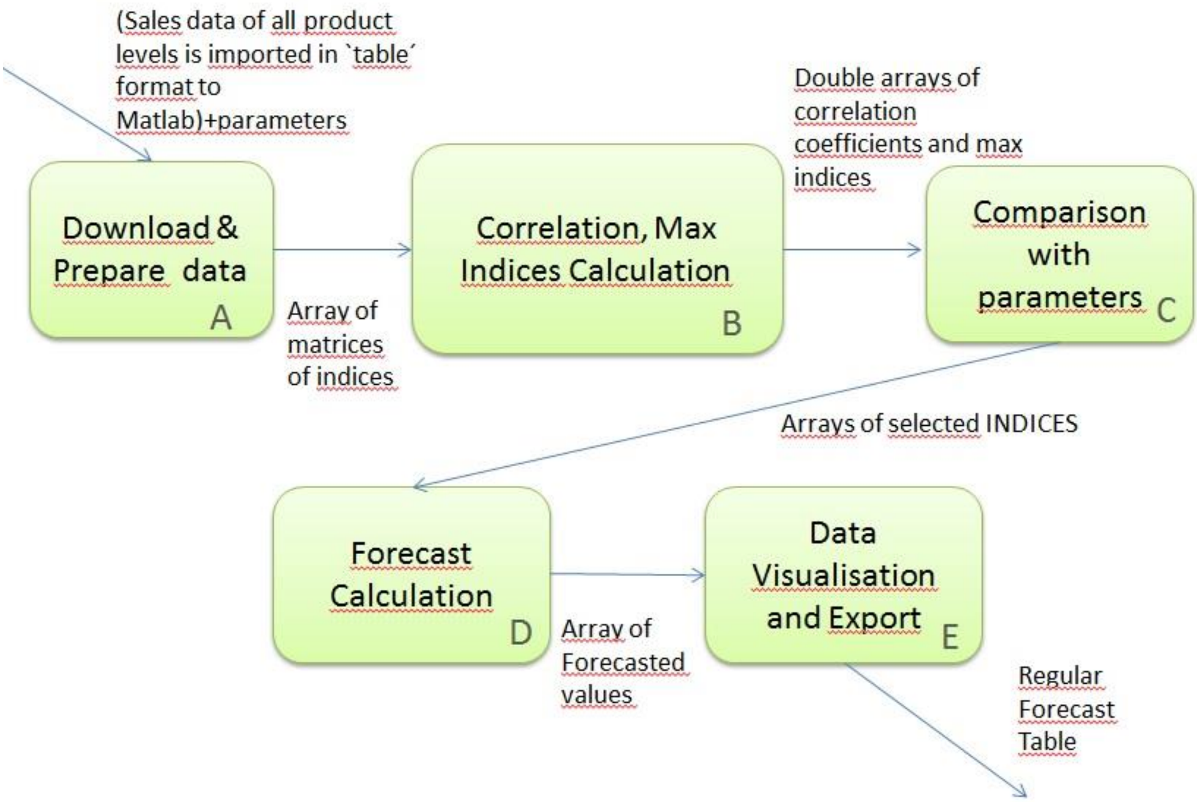


Figure 6.1. Forecast calculation steps in Matlab

Stage of data preparation is a key to the forecasting in Matlab package. The most important task is to assemble correct matrices with the data. All the computations Matlab will perform automatically with the whole matrix. The goal - is to construct a three-dimensional matrix (array of matrices), where each matrix corresponds to a certain product, each column in the matrix is a level of hierarchy, each line represents a certain forecast month. Thereby, each cell in the matrix contains one index of seasonality.

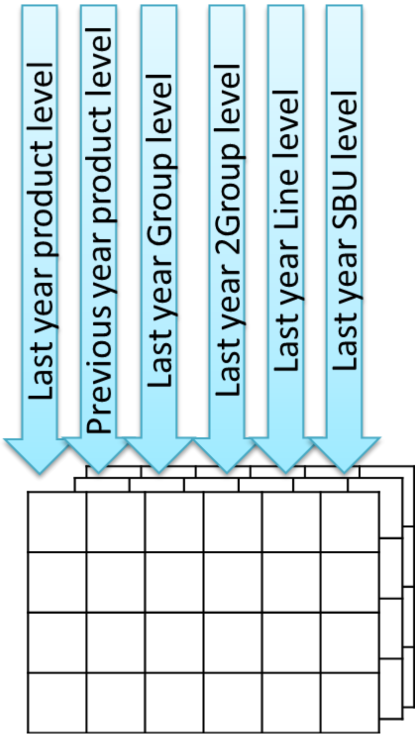


Figure 6.2. Three-dimentional Matrices

Initially, sales data for all produced end products and all higher aggregation levels of the product hierarchy is available to be loaded from database. Detailed information about product trees is necessary in order to assemble the required matrices. Product tree keeps information about a product belonging to the higher levels of the hierarchy. On the basis of product trees indices of seasonality are calculated, which once stored in a matrix format. Final three-dimensional matrix is depicted in the Figure 6.2. Each matrix corresponds to one specific product, and consists of six columns and twelve rows. Each cell stores one index of seasonality for a particular month on the particular hierarchy level. For instance, first column contains an annual time series of seasonal indices on the level of end product in the last year. Now these matrices are ready for the next step of calculations.

As Matlab operates within the Matrix structure, based on the matrices described above correlation coefficients for all the hierarchy levels can be calculated all together at the same time. *R = corr(X)* returns a p-by-p matrix containing the pairwise linear correlation coefficient between each pair of columns in the n-by-p matrix (Mathworks 2014). Basically, we just need the first raw in these matrices. It contains five correlation coefficients corresponding to the hierarchy levels of products. Correlation coefficients (five in a row) of all the products should be collected into the two-dimensional array. Similarly, we get a two-dimensional array with the values of the maximum seasonal indices at each level of the hierarchy. As a result, we continue to work with two two-dimensional arrays 5-by-N size, where N is a number of products in the forecast.

The next step is the comparison with table parameters of boundary limits. Every row of table with limitation parameters was converted into the double array. As a result of comparison series of matrices with correlation coefficients and maximum indices of seasonality with table parameters we can get a double-dimensional matrix full of zeros and ones. Zero means that pair correlation coefficient – maximum index of seasonality did not meet a table of limitation requirements, and the seasonality on this particular hierarchy level is not sufficient. On the other hand, one means that model has captured some signs of seasonality on this particular hierarchy level that can be considered as a significant seasonality. If the product has at least one "1" in the row, that means that it is seasonal and its forecast index will be taken from that level where "1" is placed. The lowest level with "1" is more significant than any lower levels with "1" value in the matrix. If the seasonality was not captured on any of the levels, the forecast index will be equal to 0,75. Final forecast is calculated as a forecast index multiplied by previous four months sales. Finally, the outcome from Matlab forecasting model is one-dimensional array of forecast figures for the next time-period expressed in pieces.

5.3. Evaluation of performance and feasibility of the model transfer into the matrix format

The time issue required for preparing the monthly forecast is considered as a disadvantage of the current forecasting model. Forecast calculation stage itself could take around hour for Excel. In comparison, it took just five minutes for Matlab to get the same forecast figures, using equal forecasting approach. Thereby, from the perspective of time, Matlab can be considered as a potential candidate to transfer the current model. At the same time to download the sales data and prepare it for the forecasting is the most time consuming stage in the whole forecasting process. There is no sense to transfer just the computational stage into the Matlab, leaving the other steps for Excel. Matlab provides an easy access to different databases so the sales data can be downloaded directly from there. And SQL cleaning quires developed in MS Access can be easily adapted for Matlab also through database Explorer. Forecast explosion stage requires further investigation and explosion algorithm development. The main challenge in the development of this algorithm will be correct storage of BOM variable in Matlab workspace.

Improvement in the model performance on computational stage has a price. It entailed some deterioration as well. Namely, user graphical interface completely disappeared in the Matlab model. All the intermediate figures, coefficients and time series are not visible and cannot be accessed just switching the spreadsheets in the model like in Excel. For business purposes, of course, visualization and GUI program plays a very important role. Without knowledge of the programming language and experience in Matlab it is difficult to work with such kind of model. The entire program does not look transparent for company management and forecast end-users; it looks like a black box with the code. It can also entail mistrust on the part of end-users of the forecast.

Fast speed of calculations makes it convenient to solve different optimization tasks. Matlab can optimize (minimize) any function that inputs a vector of parameter variable and returns the value of the minimized criterion. Naturally, in the optimization process algorithm calls an objective function multiple times; thereby its speed is an important parameter. Matlab contains good software tools that can significantly improve an algorithm performance, while supporting the readability and maintainability of the code. Boundary limits for correlation coefficients and maximum values of seasonal indices were not revalued for about one and a half years. Optimization and adjustment of these constant parameters was considered to be one possible way to increase the accuracy of the current forecasting model. Solving this particular problem, Matlab outperforms an Excel program and offers great opportunities for further development this direction. Another very promising accuracy improvement direction is development of outlier detection algorithms for the input sales data. There is a lot of methods on outlier detection including statistical intensive work, which use properties of data distribution, as well as probabilistic and bayesian techniques attempting to find the model of the anomalies, however, these approaches are oriented to univariate data or multivariate samples with only a few dimensions, processing time is also a problem when a probabilistic method is used (Escalante 2005). Matlab is capable to implement the most suitable data filtering method that would help to detect unusual spikes in the product sales dynamic, which are characterized by irregular nature of sales.

Thus, perspectives of transferring the current forecasting model into the Matlab environment is very promising step on the way to improve the model performance. Matrix calculations give a big gain in time. While running different forecast scenarios, it makes it possible to carry out a periodic optimization of constant parameters in the model, detect and remove irregular outliers from the raw sales data in order to study the behavior of the model. On the other hand, in order to start using transferred model, it requires much more effort to adapt the Matlab model for the end user, in the context of the interface, ease of sales data loading, exporting and forecast figures visualization and access.

7 PRACTICAL APPLICATION OF THE RESEARCH RESULTS

Inaccurate sales forecasts are often considered as the source of all the ills of the organization. Information Systems operating in the enterprise carry out daily planning of the entire production to a greater extent on the basis of the forecast. If erroneous values are loaded into the system, there is a threat of inventory levels increase, or their lack, unjustified costs, reduction of customer satisfaction. Price of potential damage from taking the wrong decisions today repeatedly increases, so production management must ensure selection and implementation of optimal solutions only. Creating and development of such a forecasting process that would allow the company to have a high-quality sales forecast - today is no longer a fad, but an urgent necessity that determines the competitive advantages of many enterprises. Studying of the characteristics of the forecasting method, the model is based on and model's adaptation to the specific objectives of the company, is almost always the first step in the challenge of inventory optimization for companies of all industries.

Already in 1979 Markidakis and Hibbon noticed that the forecasting method implemented in the specific model in practice, and the theoretical method described in the literature with the help of mathematics, often significantly differ from each other. The method must be specifically adapted to the needs of the company, as well as the features and requirements of the production and products. Detailed analysis of a particular example of a multiplicative model embodiment was carried out in the third chapter of the research, namely the group of decomposition methods that use linked individual indexes of seasonality. The logic of the model calculations was depicted by a visual schema that is particularly useful in order to track the process of the final forecast figures calculation. Detailed image of a logical algorithm of the model helps to identify possible sources of forecast deviations from the reality.

For effective demand forecasting it is necessary to measure the deviation of the forecasted figures from the actual sales regularly. Prediction accuracy is inversely proportional to a quantitative measure of its falsity. Numerous studies have shown that the universal combination of forecast error measures and a unified approach to interpret them does not exist, the method is chosen individually, taking into account the specific features of an organization dedicated to forecasting activities, (Zotteri 2005). Thus, a systematic analysis of the logical algorithm of forecast calculation together in parallel with a comprehensive analysis of forecast falsity reveal the main causes of errors and identify main directions of model improvement for future work with them.

Forecasting is by definition a look into the future, so it will never be completely accurate. But every organization is committed to ensure a high-quality and reasonable forecast and forecasting model that is flexible enough to respond adequately to certain changes in the demand. Usually logistics department complains about the lack of accurate forecasts. Any home-made forecasting model has its limits of accuracy, it can provide. Therefore, any company that wants to get a competitive advantage in their industry, regardless of the scope of their activities, must deal with the constant development and improvement of their predictive models. Of course, managers can learn some fresh ideas about the future development of a forecasting model from scientific literature on the subject or from published «case-studies» about real companies' experiences. One of the most extensive researches on the field of forecasting was carried out by Armstrong, Mentzer and Moon. One very important thought goes through all their works that structured forecast process is always better then unstructured one. At the same time forecasting procedure should be structured from the moment of the entire data collection till the direct forecast usage by the end users. Mathematical model development is one of the most important stages of the forecast process. Armstrong (2001) described 139 general principles of forecasting. This paper formulates common problems in forecasting, that companies face when they begin to adapt a particular approach to their activities. On the basis of extensive researches of many forecasting practices some general reference principles were formulated in order to help companies to develop a well-structured procedure for forecasting, which will have a significant positive impact on the company performance. However quite common problems and their possible solutions described in the literature are based on analysis of the experience of a large number of companies in various industries. At the same time, many companies involved in the research, probably faced with certain problems in a particular forecasting method implementation. Specific proven solutions to the forecasting problems proposed by real companies are certainly very valuable information to be studied by other companies that perform forecasting activities.

Forecasting algorithm of X's current model has been studied in details and described in this research, as well as an analysis of the fallacy of its forecasts. As a result, some of the weaknesses of the multiplicative model of seasonal products sales forecasting have been identified. Also, specific solutions have been proposed for some of them. The list of possible directions of current model development has touched virtually every step of the logical algorithm. Practical solutions for the following areas of model development were chosen from the entire list for further research:

1. Simulation of the forecast error based on the accumulated statistics on the predicted values deviations from the real sales.
2. Using a weighted average index of seasonality of the past two years in the calculation of the forecast index.
3. Development of a method for single spikes detecting for products, which are characterized by irregular sales.

Two out of three test versions of the forecasting models showed greater forecast accuracy compared to the original model. Systematic error correction and false zero forecasts models have been successful. Developed schemes that allow taking into account these types of errors are fairly simple and transparent in the context of the calculation and interpretation. One of the M3-competition conclusions conducted by Markidakis tells that simple models have an advantage in front of more complex and cumbersome models.

Nowadays, evaluation of the systematic model error is one of the most urgent problems in the constructing an adequate forecasting system. Proposed test model of the error correction is one of the few that allows correcting accurately systematic forecast errors for seasonal products apart from non-seasonal. Different correction approaches described in the literature do not take into account the seasonal factor. Description of specific features of the classical multiplicative decomposition models in future demand predicting can be found more frequently. However, all described examples of the application of these methods relate to forecasting products, which are characterized by a constant (non-zero) demand. Proposed test model of false zero forecast correction allows using multiplicative methods for a wide range of products, including products that are characterized by irregular demand.

Recall that in the framework of the practical part of the study, we have implemented the current forecast algorithm in Matlab calculation space in the format of matrix computations. This provided a substantial gain in the model runtime, made it easy and fast to optimize model parameters in this environment. At the same time model transfer into the Matlab space caused degradation of user interface performance. Excel is widely used in statistical analysis, it is considered to be a very simple and intuitive tool for implementing various forecasting methods. However, it has some limitations on the use. Excel is not the best choice for instance when a company has to predict the sales for thousands of products or the forecasting model has complex structure and logic. Companies of such production scales should be paid to other computing environments to implement their forecasting models. This study offers a promising solution of using matrices in Matlab calculation environment. Further development of custom applications based on Matlab matrix calculations can significantly improve the performance of the model and simplify the task of debugging and optimization of model parameters. Similarly, other prediction methods can be implemented by matrix approach.

However, nowadays there are different professional commercial programs available in the market that focus on the problem of statistical analysis and forecasting issues. The research conducted by Professor of Operations Management Nada R. Sanders showed that 10.8 % of USA companies that responded to the survey use the commercial software for their forecasting purposes. Some of them fully outsource forecasting process to other companies, who are responsible for data gathering and storage, running the model, model parameters optimization and other issues. The highest percentage of responding firms - 48.3 percent – report using spreadsheets, such as Excel, Lotus 1-2-3, or Quattro Pro, for forecasting. Sanders's survey results show that the majority of respondents report being dissatisfied with forecasting software, and identify ease of use and easily understandable results as the features they consider most important. However, users of commercial software packages are found to have the best forecast performance, as measured by mean absolute percentage error (MAPE). In fact, those that use commercial software had the best and most consistent performance in this study. These findings may demonstrate that there are benefits to be gained in accuracy for those that decide to take advantage of the available technology. Another interesting result from this study is that firms that make the financial investment in purchasing software technology feel a greater commitment to use it. Correlation coefficient computed between type of software and degree of reliance on automated forecasts is significantly high (Sanders 2003).

Most forecasting software products have a set of classic built-in methods, mechanism of choosing the best fit method and its optimal internal parameters. For seasonal products, methods of exponential smoothing with seasonal component are generally used. Forecasting model, developed by company X performs quite worthily and focuses especially on product sales with marked seasonal behavior. However, the high use of spreadsheets and the expressed importance of easily understandable results suggest a need for further software simplification and improved results reporting. X's model implementation in the Matlab environment and further development of an application based on it could be a first step towards the production of own software product and forecasting solution.

Since this work has been devoted to improving the accuracy and efficiency of the forecasting model for seasonal products that are characterized by marked repeating patterns in their time series of data, the results may be useful for companies or other systems that deal with seasonal behavior of the forecasting objects. For instance, demand for many products like fashion garments, shoes, sportswear, air-conditioners, heaters, certain types of food like ice cream and cold drinks, and consumption of goods and services related to the tourism industry is highly seasonal, fluctuating, and often hard to predict. Some repeating patterns in sales that could be used in order to improve the accuracy of the forecast are quite difficult to notice without a help of mathematical model. Less obvious example of products with seasonal demand for example is a demand for slippers peaks in the run up to Christmas. Thereby, value of all theoretical conclusions and developments, derived from the study of a particular forecasting model belonged to company X, have fundamental importance and can be used in the forecasting practice of other companies operating in different industrial sectors.

8 CONCLUSIONS

Energy companies are involved into the constant tough competition on the national market and on the international arena if they are exporting their products. In order to survive in the market successfully, companies need to have certain competitive advantages. Some of them rely on the development of innovative technologies and products, while others are forced to look for other ways to strengthen their positions. In particular, some companies try to reduce the production costs due to the transfer of production to regions with lower labor costs. Other companies are forced to find new ways to reduce production costs by reducing all kinds of irrational wastes. In this case, the company management is aware of the need to develop methods of future demand forecasting. However, creation and subsequent operation of a forecasting model does not guarantee to solve this problem. The uncertainty of the market, lack of understanding of the mathematical meaning of various forecasting methods, as well as the need for its continuous development and improvement are the main reasons for the lack of effectiveness of the forecasting methods in practice.

International Finnish Company X specializes in the development, manufacturing and marketing of electrical systems and supplies for the distribution of electrical power as well as electrical applications. X's feature is the use of cleaner technologies to protect the environment. Without the ability to attract new customers through lower prices for the products, company X focuses its efforts on maintaining long-term trustful relationships with their customers through business continuity and minimum delivery times. Such challenge can be achieved only when the relevant product is always in stock. The volume of products must strictly comply with the current consumer demand, eliminating wasteful expenditures. Over the past few years X performs inventory management processes with the help of the demand forecasting model. Comparative analysis of actual and forecast data showed that the model is not perfect. In opinion of company management, the main defect of the current forecasting model is its low accuracy rate. In particular, it refers to a group of products, which are characterized by long production cycle and unstable demand. As a consequence of this defect, there is the need for manual correction of the final forecast results.

Thus, the final objective of the present study was to develop improved versions of the current forecasting model with the higher accuracy rates. The ongoing process of demand forecasting has been analyzed; the degree of similarity of actual data and forecast data for the previous few years has been assessed. In the theoretical part of the study some well-known methods of time series forecasting were considered. Based on the findings of the theoretical analysis, logical forecasting flowchart algorithm was created. As a result of logical algorithm analysis and the comparison of the model figures and actual sales, some significant shortcomings of the current model were identified, that negatively affect the accuracy of the forecasts. Ultimately, a single list of possible ways to improve the forecasting model was created in order to improve its accuracy rate. The list included the following general points:

- Review and update the table with values of the maximum permissible limits of seasonality indexes, as well as their selection algorithm.
- Improve the procedures for identifying the seasonal nature of sales.
- Develop a mechanism taking into account the product lead time of in the model.
- Adjust the calculation logic the model is based on.
- Forecast error simulation based on the accumulated statistics on forecast deviations from the real sales data.
- Gain the control over the manual forecast correction by experts, as well as the actions of the end-users of the forecast.

Fourth chapter provides an exhaustive list of the selected direction for further development and testing. Three test models were created on the basis of selected refinement directions. Developed mechanism of identification and elimination of false zero predictions was included in one of the versions. In the second test model weighted average index of the seasonality of past two years has been used in the calculation stage of the final forecast index instead of simple average. The third version of the current forecasting model was supplemented with a mechanism of systematic error identification and correction for specific end products. As a result of tests conducted on past actual sales data the following findings have been concluded:

1. Model with option of system error correction improved the coefficient of determination by 10% in average, and the value of the mean absolute percentage error by 25% for those products, whose relative deviation from the reality is determined by the model as seasonal.
2. Model with weighted average forecast seasonal indices did not give positive results, the accuracy rates of its forecast figures deteriorated. Based on the analysis of the possible reasons of this fail new perspective direction of model development was formulated: adaptive model that takes into account data on the mistakes of the past forecasts.

3. Model with option of false zero forecast has reduced the number of such errors by 40% in the absolute average.

Required time and ease of use are also considered to be very important requirements that affect the performance success of the forecasting model. In order to improve the time records alternative computing environments were considered for potential transferring of the current model in the sixth chapter of this study. Transfer of all computations to Matlab in matrix format helped to reduce significantly the time required to run forecasting algorithm utilized by the company X. However, this transfer requires further intensive development of the interface and the general procedure of management interaction with the model.

The results of the study can be attributed to a class of improving innovations; they provide a significant increase in the accuracy of the forecasting. Moreover, this class of innovation is not based on the use of absolutely new technologies and approaches (Tukkel et al. 2011). Due to the fact that X company's turnover is 280 million euros, any improvement in the accuracy of demand forecasting and some reduction of unnecessary costs can bring a significant increase in profits, as well as to improve the quality of customer service that the company considered as a priority value.

Many researchers face the challenge of forecast accuracy improvement. The novelty of this work lies in the fact a novel approach of model analysis was first used and described in this study, which reveals its critical poor areas that may affect the accuracy of the final forecast. Visual representation of the calculation logic the model is based on together with an analysis of the accumulated information about past forecast errors can facilitate and accelerate the development of improved test versions.

Nowadays, evaluation of the systematic error is one of the most pressing problems when building an adequate forecasting system. Proposed test model of error correction is one of the few that allow correcting systematic forecast errors sufficiently for seasonal products apart from non-seasonal. False zero forecast model allows using multiplicative forecasting methods for a wide range of products, including products with irregular demand. Possible transfer of the current forecasting model into the Matlab environment is considered to be a quite promising step on the way to improve the model performance in context of time required. Matrix calculations give a big gain in time of the calculations while running the logical algorithm of forecasting.

The future development of this thesis topic will be the practical implementation of the other planned model refinement trends. Development of forecasting seasonal model that can take into account the time of products delivery, as well as an adaptive model with a weighted average indices of seasonality, which are calculated on the basis of past forecast deviations from the reality are particularly interesting for further research. While running different forecast scenarios in matrix format using the Matlab environment, it makes it possible to carry out an optimization of constant parameters in the model, detect and remove irregular outliers from raw sales data in order to study the behavior of the model.

All valuable theoretical findings and developments obtained during the study of a particular forecasting model of company X have fundamental importance and can be used in forecasting activities of other companies operating in plenty of diverse industries.

REFERENCES

Aivazian, S.A. 2001. Прикладная статистика. Основы эконометрики. [Applied Statistics. Foundations of econometrics.] Moscow:Uniti-Dana, p.432. (In Russian)

Archipov, C. 2008. Методы - теория и практика | SCM - Управление цепочками поставок. [Methods - Theory and Practice | SCM - Supply Chain Management]. [online document]. [Accessed 16 April 2014]. Available at http:// scm-book.ru/node/6 (In Russian)

Armstrong, J. 2002. Principles of forecasting. 1st ed. Boston, MA.: Kluwer Academic Publishers.

Arnold Tony J.R., Chapman Stephen N & Clive Lloyd M. 2008. Introduction to materials management, Upper Saddle River, NJ: Pearson/Prentice Hall.

Australian Bureau of Statistics. 2005. Time Series Analysis: The Basics. [Online document] [Accessed 29 May 2014]. Available at http:// www.abs.gov.au/websitedbs (In Russian)

Bails, D., Peppers, L. 1993. Business fluctuations: Forecasting techniques and applications. NY.:Prentice Hall. 126 p.

Box G., Jenkins G. 1976. Time series analysis. 1st ed. San Francisco: Holden-Day, 237 p.

Bushueva, L. I. 2004. Методы прогнозирования объема продаж. Маркетинг в России и за рубежом. [Methods of sales forecasting. Marketing in Russia and abroad]. (In Russian)

Chernysh, E. A., Molchanova, N.P., Novikova, A.A., Saltanova, T.A. 2009. Прогнозирование и планирование в условиях рынка: учебное пособие [Forecasting and planning in market: tutorial]. Moscow: PRIOR. (In Russian)

Chopra, S., Meindl P. 2006. Supply Chain Management: Strategy, Planning and Operation. 3rd edition. Upper Saddle River, NJ: Prentice Hall, 543p.

Chuchueva, I.A. 2012. Модель прогнозирования временных рядов по выборке максимального подобия. [Time series forecasting model based on a sample of maximum similarity]. Doctoral dissertation. Moscow, Moscow State University, 153 p.

Clements M., Hendry D. 1999. Forecasting non-stationary economic time series. 1st ed. Cambridge, MA: MIT Press, 345 p.

Clements, M., Hendry, D. 1996. Intercept corrections and structural change. Journal of Applied Econometrics. vol. 11, № 5, p. 475-494.

Clifton, P., Nguyen, H., Nutt, S. 1998. Market Research: Using Forecasting in Business. International Journal of Forecasting, vol.9, №4, p.579.

Dubova,T.A., Archipova M.Ju. 2004. Статистические методы прогнозирования в экономике: учебное пособие, практикум, тесты, программа курса. [Statistical methods of forecasting in the economy: a tutorial, workshop, tests, course program]. Moscow: Moscow State University of Economics, Statistics and Informatics, 136 p. (In Russian)

Effective Inventory Management. 2013. Forecasting Items with Long Lead Times. [online document] [Accessed 19 May 2014]. Available at http://www.effectiveinventory.com/long-lead-times/ .

Entov, R., Nosko, V., Judin, A., Kadochnikov, P., Ponomarenko, S. 2002. Проблемы прогнозирования некоторых макроэкономических показателей. [Problems of macroeconomic indicators forecasting]. Moscow: IEPP, serial "Scientific work" IEPP, №46. (In Russian)

Escalante, H. J. 2005. A comparison of outlier detection algorithms for machine learning. In Proceedings of the International Conference on Communications in Computing.

Gardner, E. 2006. Exponential smoothing: The state of the art—Part II. International Journal of Forecasting, 22(4), pp.637-666.

Gorban, A.N., Dunin-Barcovsky, V.L., Kirdin, A.N. et.al. 1998. Нейроинформатика [Neuroinformatics.] Novosibirsk: Nauka. Sibirskoye predprijatie RAN, 296 p.

HELPSTAT. 2012. Проверка адекватности регрессионной модели. [Verification of the regression model adequacy] [online document]. [Accessed 10 May 2014]. Available at http://helpstat.ru/2012/01/proverka-adekvsionnoj-modeli/.(In Russian)

Hollins, B., Shinkins, S. 2006. Managing service operations. London: Sage Publications, 312 p.

Keat, P. and Young, P. 1996. Managerial economics. Upper Saddle River, NJ [etc.]: Prentice Hall.

Kerkkänen, A. 2010. Improving demand forecasting practices in the industrial context. Doctoral dissertation. Lappeenranta, Lappeenranta University of Technology, Department of Industrial Management, Acta Universitatis Lappeenrantaensis, 382 p.

Lapygin, U. N., Krylov, V.E., Chernyavskiy, A.P. 2009. Экономическое прогнозирование: учебное пособие [Economic forecasting: tutorial]. Moscow: Eksmo. (In Russian)

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

Lindsey, M., Pavur, R. 2013. Assessing a Modification to Croston's Method to Incorporate a Seasonal Component. Advances in Business and Management Forecasting, № 9, p. 185-195.

Lukashin, U. P. 2003. Адаптивные методы краткосрочного прогнозирования временных рядов. [Adaptive methods of short-term time series forecasting]. Moscow: Finance and Statistics, p. 367. (In Russian)

Mainzer, K. 2004. Thinking in complexity. Berlin: Springer, 482 p.

Makridakis, S., Hibon, M. and Moser, C. 1979. Accuracy of Forecasting: An Empirical Investigation. Journal of the Royal Statistical Society. Series A (General), 142(2), p.97.

Makridakis, S., Hibon, M. 2000. The M3-Competition: results, conclusions and implications. International journal of forecasting, vol. 16, № 4, p.451-476.

Mathworks 2014. MATLAB GUI - MATLAB. [Online] Available at: http://www.mathworks.com/discovery/matlab-gui.html [Accessed: 21 January 2015].

Mazmanova, B.G. 2000. Методические вопросы прогнозирования сбыта. [Methodological issues of sales forecasting]. Marketing in Russia and abroad. № 1, p. 121-130. (In Russian)

Mentzer, J., Moon, M. 2005. Sales forecasting management. 1st ed. Thousand Oaks, Calif.: Sage Publications.

Microsoft 2014. Article. Data analysis in pivot tables using the Excel 2013 [www document]. [Accessed 1 April 2014]. Available URL: http://office.microsoft.com/ru-ru/excel-help

Nikolopoulos K., Assimakopoulos V. 2003. Theta intelligent forecasting information system. Industrial Management & Data Systems, vol. 103, № 9, p.711-726.

Orlova, I.V. 2000. Экономико-математические методы и модели. Выполнение расчетов в среде Exce1. Практикум. [Economic-mathematical methods and models. Performing calculations in Excel. Practical work.] Moscow: ZAO Finstatinform. 126 p. (In Russian)

Raqsoft 2012. Analysis Tools Comparison: R Language, Matlab, esProc, SAS, SPSS, Excel, and SQL. Raqsoft Co. web log [Online] 17 October 2012. Available at: http://datakeyword.blogspot.mx/ [Accessed: 13 January 2015].

Sanders, N., Ritzman, L. 1987. Improving short-term forecasts by combining expert judgment with quantitative techniques. 1st ed. Columbus, Ohio: College of Business, Ohio State University, 320 p.

Sanders, N. and Manrodt, K. 2003. Forecasting Software in Practice: Use, Satisfaction, and Performance. Interfaces, 33(5), pp.90-93.

Saunders, M., Lewis, P. and Thornhill, A. 2009. Research methods for business students. New York: Prentice Hall.

Shraibpheder, G. 2005. Эффективное управление запасами. [Effective Inventory Management] . Moscow: Alpina Business Books. 238 p. (In Russian)

StatSoft. Электронный учебник по статистике. [StatSoft. Electronic textbook on statistics.] 2013. [online document]. [Accessed 28 April 2014]. Available at http:// www.statsoft.ru/home/textbook/default.htm.

Svetunkov, S. G., Svetunkov I. S. 2009. Методы социально-экономического прогнозирования: учебник для вузов. [Methods of social and economic forecasting: a textbook for high schools]. Saint-Petersburg: SPbGUEF, vol. 1, p. 453. (In Russian)

Tatarenko, S. I. 2008. Методы и модели анализа временных рядов : методические указания к лабораторным работам. [Methods and models for time series analysis: guidelines for laboratory work]. Tambov: Izdatelstvo Tambovskogo Tehnicheskogo Universiteta, p. 32. (In Russian)

Tebekin, A. V. 2013. Методы принятия управленческих решений: учебник для бакалавров. [Methods of managerial decision-making: a textbook for undergraduate]. Moscow: Jurait, 572 p. (In Russian)

Tukkel, I.L., Surina A.V., Kultin, N.B. 2011. Управление инновационными проектами: учебник. [Management of innovative projects: a textbook]. Ed.by Tukkel I.L., Saint-Petersburg: BHV-Petersburg, 416 p. (In Russian)

Turuntseva, M.E. 2013. Нужно ли пытаться исправить потенциальные ошибки прогнозирования? [Is it necessary to fix potential predictions errors?]. Nauchni vestnic IAP imeni Gaidara.ru . № 2 p. 38-46 . (In Russian)

Utley, Joanne. 2011. Removing Systematic Bias from Demand Forecasts. Advances in Business and Management Forecasting. UK: Emerald Group Publishing Limited, Ed.Kenneth D. Lawrence, Ronald K. Klimberg. vol. 8, p.3-12.

Waddell, D., Sohal, A. S. 1994. Forecasting: the Key to Managerial Decision Making. Management Decision, vol. 32. No. 1, p. 41–49.

Wallace, T., Stahl, R. Sales & operations planning. 1st ed. Cincinnati, Ohio: T.F. Wallace & Co. 423 p.

Webby, R., O'Connor, M., Lawrence, M. 2001. Judgmental time-series forecasting using domain knowledge. Principles of Forecasting. Norwell, MA: Kluwer Academic Publishers, p.342-348.

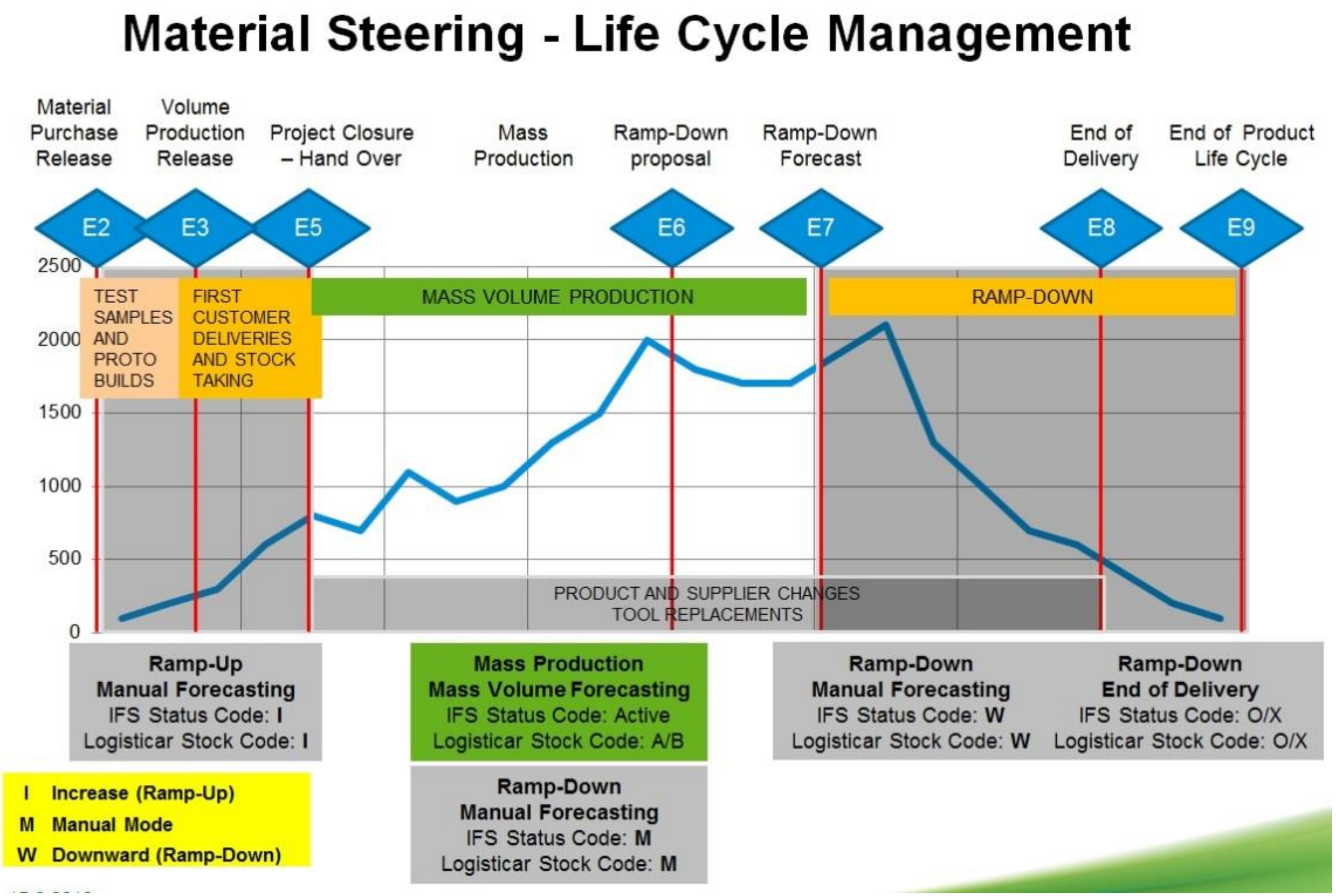
Wright, G., Goodwin, P. 1998. Forecasting With Judgment. 1th ed, John Wiley & Sons Ltd, Baffins Lane, Chichester, West Sussex PO19 1 UD, England.

Yokum, J., Armstrong, J. S. 1995. Beyond Accuracy: Comparison of Criteria Used to Select Forecasting Methods. International Journal of Forecasting, vol.11, p.591-597.

Zharikov, A.V., Goriachev, R.A. 2013. Прогнозирование спроса и объемов продаж: учебно-методическое пособие. [Demand forecasting and sales volumes: a teaching aid]. Nizhni Novgorod: Nizhni Novgorod Stste University. 39 p. (In Russian)

Zotteri, G., Kalchmidt, M., Caniato, F. 2005. The impact of aggregation level on forecasting performance. International Journal of Production Economics, vol. 93-94, № 1, p. 479-491.

APPENDIX 1. MATERIAL STEERING – LIFE CYCLE MANAGEMENT



APPENDIX 2. ADDITIONAL INFPRMATION ABOUT THE TIME SERIES FORECASTING METHODS

(Lapygin et al. 2009), (Chernysh et al. 2009), (Bushueva 2004), (Tatarenko 2008)

1. Naive methods
   a. Mean average

$$Y(t) = C + \varepsilon(t)$$

$$C = \frac{1}{N}\sum_{1}^{N} X,$$

where *Y* – forecast, *X* – element of the time series, *C* – level value, $\mathcal{E}$ – deviation value, *N* – number of elements in the time series

   b. Moving average model

$$\mathrm{C} = \frac{1}{N}\sum\nolimits_{N-k}^{N} X,$$

where *C* – level value, *X* – element of the time series, *N* - number of elements in the time series, *k* - the sequence number of the last month, taken into the forecast calculation horizon

   c. Weighted average model

$$C = \frac{\sum_{N-k}^{N} \alpha_i X_i}{\sum_{N-k}^{N} \alpha_i},$$

where *C* – forecast, *X* – element in the time series, *α* – weight, *N* - number of elements in the time series, *k* - the sequence number of the last month, taken into the forecast calculation horizon

2. Exponential smoothing model
   a. Brown's model

$$Y_{i+1} = \alpha X_i + (1 - \alpha) Y_i\ ,$$

where X_i - element i in the time series, Y_i – forecast i, α – smoothing coefficient, constant for entire time series

b. Holt's model

$$Y_{t+\tau} = a_t + \tau b_t$$
$$a_t = \alpha_1 Y_t + (1-\alpha_1)(a_{t-1} + b_{t-1})$$
$$b_t = \alpha_2(a_t - a_{t-1}) + (1-\alpha_2)b_{t-1},$$

where $Y_{t+\tau}$ - forecast for τ steps ahead, α_t – level coefficient, b_t - proportionality factor, α_1 and α_2 – smoothing parameters

In order to get the forecast, calculated according Holt's method, you must first calculate the values of a_0 and b_0 coefficients for the time series. Thereafter, constant smoothing parameters have to be chosen based on any additional criteria. In the end, the researcher obtains a linear mathematical model which adapts to the current actual data at each step of the forecast performance.

c. Holt-Winters model

$$Y_{t+\tau} = (a_t + \tau b_t)\mathrm{c}_{t-s+\tau}$$
$$a_t = \alpha_1 \frac{Y_t}{c_{t-s}} + (1-\alpha_1)(a_{t-1} + b_{t-1})$$
$$b_t = \alpha_2(a_t - a_{t-1}) + (1-\alpha_2)b_{t-1}$$
$$c_t = \alpha_3 \frac{Y_t}{a_t} + (1-\alpha_3)c_{t-s},$$

where $Y_{t+\tau}$ - forecast for τ steps ahead, α_t – level coefficient, b_t - proportionality factor, $c_{t-s-\tau}$ – seasonal component delayed by s+τ steps, α_1 and α_2 – smoothing parameters

An algorithm for constructing Holt-Winters model tells the following: first you have to calculate the coefficients a_0 and b_0 of the linear trend for the part or the entire time series. Further, based on a certain part of the series containing s observations seasonal components c_t should be calculated by the following formula:

$$c_t = \frac{Y_t}{a_0 + b_0 t}$$

d. Theta-model

Coefficient «Θ» corresponds to the second derivative of the function taken from the curve of the time series.

$X'_{new}(\Theta) = \Theta X''_{data}$, where $X''_{data} = X_t - 2X_{t-1} + X_{t-2}$ at the moment t

Smoothing the curvature of the line allows capturing the trend component in the time series. The smaller the value of the coefficient Θ, the sharper the trend. The greater the value of Θ, the more noticeable are jumps in the time series curve. Thus, the initial time-series is decomposed into two or more Θ-curves, each of them should be extrapolated individually. Ultimately, all the extended lines are summarized with each other. Extrapolation can be performed by any known statistical method.

3. ARIMA models of Box and Jenkins

There are three types of parameters in the model: autoregressive parameters (p), parameter responsible for setting the time delay (d), and moving average parameters (q). In the notation of Box and Jenkins ARIMA model is written as (p, d, q). For example, the model (0, 1, 2) contains 0 autoregressive parameters and 2 moving average parameters, which are calculated for time series after taking the difference in time with a delay of one time period. Forecasting model, where time series levels are defined as a linear function of the previous observations, is called autoregressive (AR). If the current value of the series level X_t depends only on a single previous value X_{t-1}, then this model will be a first-order autoregressive model AR (1). If X_t depends on the previous two levels, the model is a second order auto-regressive AR (2), and so on up to the end of order p.

$$X_t = c + \sum_{i=1}^{p} a_i X_{t-i} + \varepsilon_t ,$$

where $\mathcal{E}_t$ – white noise, $\alpha_1 \ldots \alpha_p$ – model parameters, c – model constant

In contrast to the autoregressive models, each element of the model of moving average (SS) is a subject to the cumulative effects of a number of previous errors. In general terms, it can be written as follows:

$$X_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2},$$

where μ - model constant, $\Theta_1 \ldots \Theta_q$ – model parameters

APPENDIX 3. LIST OF FORECASTING METHODS INCLUDED IN M3-COMPETITION

Method	Competitors	Description
Naïve/simple		
1. Naïve2	M. Hibon	Deseasonalized Naïve (Random Walk)
2. Single	M. Hibon	Single Exponential Smoothing
Explicit trend models		
3. Holt	M. Hibon	Automatic Holt's Linear Exponential Smoothing (two parameter model)
4. Robust-Trend	N. Meade	Non-parametric version of Holt's linear model with median based estimate of trend
5. Winter	M. Hibon	Holt–Winter's linear and seasonal exponential smoothing (two or three parameter model)
6. Dampen	M. Hibon	Dampen Trend Exponential Smoothing
7. PP-autocast[a]	H. Levenbach	Damped Trend Exponential Smoothing
8. Theta-sm	V. Assimakopoulos	Successive smoothing plus a set of rules for dampening the trend
9. Comb S-H-D	M. Hibon	Combining three methods: Single/Holt/Dampen
Decomposition		
10. Theta	V. Assimakopoulos	Specific decomposition technique, projection and combination of the individual components
ARIMA/ARARMA model		
11. B–J automatic	M. Hibon	Box–Jenkins methodology of 'Business Forecast System'
12. Autobox1[a]	D. Reilly	Robust ARIMA univariate Box–Jenkins
13. Autobox2[a]		with/without Intervention Detection
14. Autobox3[a]		
15. AAM1	G. Melard,	Automatic ARIMA modelling with/without
16. AAM2	J.M. Pasteels	intervention analysis
17. ARARMA	N. Meade	Automated Parzen's methodology with Auto regressive filter
Expert system		
18. ForecastPro[a]	R. Goodrich, E. Stellwagen	Selects from among several methods: Exponential Smoothing/Box Jenkins/Poisson and negative binomial models/Croston's Method/Simple Moving Average
19. SmartFcs[a]	C. Smart	Automatic Forecasting Expert System which conducts a forecasting tournament among four exponential smoothing and two moving average methods
20. RBF	M. Adya, S. Armstrong, F. Collopy, M. Kennedy	Rule-based forecasting: using three methods — random walk, linear regression and Holt's, to estimate level and trend, involving corrections, simplification, automatic feature identification and re-calibration
21. Flores/Pearce1	B. Flores,	Expert system that chooses among four methods
22. Flores/Pearce2	S. Pearce	based on the characteristics of the data
23. ForecastX[a]	J. Galt	Runs tests for seasonality and outliers and selects from among several methods: Exponential Smoothing, Box–Jenkins and Croston's method
Neural networks		
24. Automat ANN	K. Ord, S. Balkin	Automated Artificial Neural Networks for forecasting purposes

APPPENDIX 4. EXAMPLE OF SEASONALITY INDICES CALCULATION IN CASE OF FALSE ZERO PREDICTION

Product Id	201109	201110	201111	201112	201201	201202	201203	201204	201205	201206	201207	201208	201209	201210	201211	201212
ST206.3 Sales		50,00			460,00	15,00	7,00								**920,00**	
S.Indexes					9,64	0,04	-	-	-	-	-	0,75	0,75	0,75	0,75	0,25