# MAGISTERARBEIT

Titel der Magisterarbeit

„Causal Demand Forecasting and Safety Stock Planning“

Verfasser
Lalo Valchkov

angestrebter akademischer Grad

Magister der Sozial- und Wirtschaftswissenschaften
(Mag. rer. soc. oec.)

Wien, im Juli 2012

| | |
|---|---|
| Studienkennzahl lt. Studienblatt: | A 066 915 |
| Studienrichtung lt. Studienblatt: | Magisterstudium Betriebswirtschaft |
| Betreuer: | Univ.-Prof. Dr. Stefan Minner |

# Eidestattliche Erklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit selbstständig verfasst, ganz oder in Teilen noch nicht als Prüfungsleistung vorgelegt und keine anderen als die angegebenen Hilfsmittel benutzt habe. Sämtliche Stellen der Arbeit, die benutzten Werken im Wortlaut oder dem Sinn nach entnommen sind, habe ich durch Quellenangaben kenntlich gemacht. Dies gilt auch für Zeichnungen, Skizzen, bildliche Darstellungen und dergleichen sowie für Quellen aus dem Internet.

Wien, im Juli 2012

# Abstract

The inventory management is one of the major tasks a company is faced with. The challenge is that the company orders the products before customers demand them. There are two main reasons why this happens. First of all, there is a lead-time between the ordering time and the delivery time. Second of all, due to certain ordering costs, it is often necessary to order in batches instead of unit for unit. This means that the company needs to forecast the future demand. A demand forecast is an estimated average of the demand size over some future periods. Therefore, the company also needs to determine how uncertain the forecast is. The forecast reflects on the stock level held by the company and its uncertainty influences the company to hold additional stock for the unpredictable demand. An accurate forecast or at least one with the smallest forecast error is crucial for every company, because it ensures two major tasks of the inventory management: to satisfy a given level of demand and to minimize the inventory cost. The accurate demand forecast ensures that there would be no lost sales, which in the worst case leads to lost customers and in the same time, there would be no unnecessary tied capital in inventories.

In order to solve these problems two main forecasting approaches have been developed: the time series forecasting methods and the causal forecasting models.

This work provides a comparison between these two forecasting approaches. In addition two critical assumptions for the causal forecasting methods, namely the homoscedasticity and the nonautocorrelation, will be discussed.

Furthermore, the work investigates the safety stock planning in $(\hat{t},\hat{S})$ inventory policy with zero and positive lead time. The impact which the Order Service Level (OSL), the Unit Service Level (USL) and the forecasting inaccuracy have on safety stock planning will be investigated. For this purpose a sales data for "Schwechater" beer canes sold in Austria for 2005, 2006 and 2007 by ADEG Austria Ltd will be used.

# Exposé

Das Bestandsmanagement ist eine der wichtigsten Aufgaben, mit der ein Unternehmen konfrontiert wird. Die Herausforderung besteht darin, dass das Unternehmen die Produkte beschaffen muss, bevor die Kundennachfrage entsteht. Es gibt zwei Hauptgründe, warum dies geschieht. Zunächst kommt es zu einer Vorlaufzeit zwischen dem Zeitpunkt der Bestellung und der Lieferzeit. Zweitens ist es aufgrund bestimmter Bestellkosten oft notwendig Großmengen statt Kleinmengen zu bestellen. Dies bedeutet, dass in einem Unternehmen der künftige Bedarf prognostiziert werden muss. Eine Bedarfsprognose ist die geschätzte durchschnittliche Nachfragegröße über einige zukünftige Perioden. Deshalb hat das Unternehmen festzustellen, wie unsicher die Prognose ist. Die Prognose spiegelt den Lagerbestand wieder und ihre Ungenauigkeit veranlasst das Unternehmen zusätzliche Bestände für die unvorhersehbare Nachfrage vorrätig zu halten. Eine exakte Prognose oder zumindest eine mit dem kleinstmöglichen Prognosefehler ist von entscheidender Bedeutung für jedes Unternehmen, da sie zwei wichtige Funktionen des Bestandsmanagements erfüllt: ein bestimmtes Nachfrageniveau zu befriedigen und die Lagerkosten zu minimieren. Die genaue Bedarfsprognose sorgt dafür, dass keine entgangenen Umsätze entstehen, die im schlimmsten Fall zu einem Kundenverlust führen können, und zugleich verhindert sie eine unnötige Kapitalbindung im Lagerbestand.

Um diese Probleme zu lösen wurden zwei wesentliche Prognoseverfahren entwickelt: das Zeitreihenprognoseverfahren und die Kausalprognosemodelle.

Die vorliegende Magisterarbeit stellt einen Vergleich zwischen diesen beiden Prognoseansätzen her. Zusätzlich werden zwei bedeutende Annahmen für die Kausalprognosemodelle diskutiert: die Homoskedastizität und die Non-Autokorrelation. Darüber hinaus untersucht die Magisterarbeit die Sicherheitsbestandsplanung der $(\hat{t},\hat{S})$ Lagerhaltungspolitik mit null und positiver Lieferzeit. Die Auswirkungen, die der Order Service Level (OSL), der Unit Service Level (USL) und die Prognoseungenauigkeit auf die Sicherheitsbestandsplanung haben, werden beleuchtet. Zu diesem Zweck werden die Verkaufsdaten von ADEG Österreich AG für „Schwechater Bier" in Dosen am österreichischen Markt aus den Jahren 2005, 2006 und 2007 verwendet.

# Contents

# Tables

# Figures

# 1. Introduction

## 1.1. Problem definition

The inventory management is one of the major tasks a company is faced with. Inventories are the needed quantities that must be kept on stock in order to satisfy a certain level of demand. The problem, however, is that the company orders the products before customers demand them. There are two main reasons why this happens.  First of all, there is a lead-time between the ordering time and the delivery time. Second of all, due to certain ordering costs, it is often necessary to order in batches instead of unit for unit. This means that the company needs to forecast the future demand. A demand forecast is an estimated average of the demand size over some future periods. Therefore, the company also needs to determine how uncertain the forecast is.[1] However, there is only one certain thing about the future and it is that the future is uncertain. Therefore, variations between actual demand and forecast are inevitable.[2] The forecast reflects on the stock level held by the company and its uncertainty influences the company to hold additional stock for the unpredictable demand, which is known as safety stock. An accurate forecast or at least one with the smallest forecast error is crucial for every company, because it ensures two major tasks of the inventory management: to satisfy a given level of demand and to minimize the inventory cost. The accurate demand forecast ensures that there would be no lost sales, which in the worst case leads to lost customers and in the same time, there would be no unnecessary tied capital in inventories. The inventories must be stored for the time before they are taken away from stock by customers. This occurs holding costs that consist of space and equipment costs, insurance costs, personnel wedges, opportunity costs for holding stocks, energy costs and taxes.[3]

Moreover, one of the well-known phenomena in the Supply Chain Management – the Bullwhip Effect, can be caused by the forecast. If the members of the supply chain have access only to a local information system, each member will update the forecast based on the demand of the subsequent member in the supply chain. This could cause that even small order variability at the customer level amplifies the orders for upstream participants, such as wholesalers and

[1] Axsäter Sven, Inventory Control, 2006, p. 7
[2] Krupp J. (A), 1997
[3] Kapkova Albena, Inventory Management for Perishable Goods, Master Thesis, 2006, pp.1

manufacturers, as the order moves up along a supply chain. Even when consumer sales show relatively constant demands, the order placed by a retailer to a wholesaler is likely to fluctuate more than the actual demand perceived by that retailer. The wholesaler's order to the manufacturer and the order of the manufacturer to the supplier fluctuate even more.

The increased variability and uncertainty requires each member of the supply chain to increase the level of stocks in order to maintain established service levels causing increased inventory holding costs due to overstocking throughout the supply chain. This leads to inefficient use of resources and may result in poor customer service and profitability.

Demand forecast updating suggests that demand amplification occurs due to the safety stock and long lead time. As orders are forecasted and transmitted along the supply chain, the safety stocks are built up, and thus the bullwhip effect occurs. Because the bullwhip effect has the detrimental impacts on the performance of the whole supply chain, an accurate forecast is crucial for every member of the supply chain.[4]

On the other hand the variation of the demand, this is the forecasting error, is used to plan the safety stock of a given product. This raises the question what kind of forecasting model should be used in order to fulfill the accuracy requirement.

There two major forecasting categories – quantitative and qualitative. Quantitative methods can be divided into time series and causal methods, and qualitative can be divided into explanatory and normative methods. In this thesis, the quantitative methods will be considered. Although the quantitative methods represent the past and nothing remains the same, the past evidences tend to repeat in the future.[5]

In the "Journal of the Operation Research Society" "time series + safety stock" was given as search criteria and 21 results that match were found. In the same journal for the search criterion "linear regression models + safety stock" there have been found only five results that match. In the "International Journal of Forecasting" the results were 32 articles for the criterion "time series + safety stock" and 16 for "linear regression models + safety stock".

The results from the search show that the forecasting is still considered as a technical process, dominated by statistical methods applied to historical data. Nowadays, however, the demand of a certain product is influenced by psychological, social or political factors. Under psychological

[4] Keller, Die Reduzierung des Bullwhip-Effektes, 2004, pp. 21; Seung-Kuk Paik, Prabir Bagchi, 2007

[5] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, p. 8

factors, for example, can be considered the expected rise or fall of the personal earnings. For example, at the end of 2008 the sales of new automobiles in the USA sunk with 32 % compares with same period one year earlier and with 15% in Europe as a consequence of the financial crises started in 2007, because the people were not sure for their jobs at the time of crises. As a response to the decreased sales the producers sank the prices trying to raise the sales.[6] In such cases, a forecasting based on historical data would be more inaccurate than a forecasting based on customers' earnings expectations and product's changed price.

The both methods require sufficient information about the past, which can be quantified in the form of numerical data. The both methods assumed that some aspects of the past pattern will continue into the future. However, these two methods have their strengths and weaknesses and are used for different purposes. The time-series methods predict the continuation of growth or decrease in sales, for example, based on past values and past errors, while the causal methods try to understand, for example, how prices and advertising affect sales. The causal methods assume that there is cause-effect relationship between the factor and one or more independent variables.[7]

The objective of this thesis is to provide an insight into the time series forecasting methods and linear regression forecasting methods, to compare them and to observe the planning of the safety stock based on the forecasting.

## 1.2. Organization of the Thesis

Apart from this introductory chapter, this thesis contains a further seven chapters.

In chapter two an overview of the basic time-series methods is introduced.

Chapter three reviews the linear regression models.

Chapter four presents classical safety stock planning techniques and the $(\hat{t}, \hat{S})$ inventory policy.

In Chapter five one of the basic assumptions of the linear regression models, namely the homoscedasticity, is discussed.

Chapter six overviews another basic assumption of the linear regression models, namely the autocorrelation.

[6] Spiegel, 25th November, 2008, German Auto Industry Facing the Abyss

[7] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp. 8

After having provided a basic theoretical background of the forecasting methods and the classical safety stock planning techniques, these different forecasting methods are compared in Chapter seven by using a company's product sales record.

Chapter eight delivers a comparison of the forecasting methods with safety stock planning techniques on the $(\hat{t}, \hat{S})$ policy base.

Finally, Chapter nine delivers a concluding remark on the investigations performed.

# 2. Time-Series Forecasting Methods

Time-series methods are the one that can be used more easily to forecast when the necessary data are available. In this case a forecasting relationship can be hypothesized either as a function of time or as a function of independent variables. A very important step in selecting the time-series method is to consider the types of data patterns, so that the most appropriate method can be used. There are four types of data that can be distinguished: horizontal, seasonal, cyclic and trend.

A horizontal pattern exists when data values fluctuate around a constant mean. For example, a product whose sales do not deviate over time would be of this type. A seasonal pattern exists when a series is influenced by seasonal factors such a certain day of the week or a certain quarter of the year. The ice cream, for example, is a product with such a pattern.

A cyclical pattern exists when the data are influenced by long-term economic fluctuations. The main difference between a seasonal and cyclical pattern is that the seasonal has a constant length and regular periodic basis, while the cyclical has varying length and magnitude.

A trend pattern exists when there is a long-term secular increase or decrease in the data. An example for such a pattern is the gross national product.[8]

Appropriate time-series methods exist for every type of data. Because there are plenty of methods, in thesis only the basic average and exponential smoothing methods would be considered.

There are two groups of time series methods - the averaging methods and the exponential smoothing methods. While the averaging methods consider all observations to be equally weighted, the exponential methods apply an unequal set of weights to the past which typically decay in an exponential manner from the most recent to the most distant datapoint.[9]

[8] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp. 9

[9] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, p. 67

## 2.1. Averaging Methods

### 2.1.1. The Mean Method

The Mean Method is the simplest one of all averaging methods. This method takes the average of all data and the forecast for the coming period is the average of the data. If the data $X_1$, $X_2$, $X_3$ …… $X_{N-1}$, $X_N$ is given the forecast for the next period can be compute with the following formula:

$$F_{T+1} = \overline{X} = \sum_{i=1}^{T} X_i / T \qquad (2.1)$$

The forecast for the next periods are calculated in the same way.

Because every forecast has an error, it can be computed as a difference between the forecasted value and the observed value:

$$e_{T+1} = X_{T+1} - F_{T+1}$$

This simple method is appropriate only if the data has no noticeable trend or noticeable seasonality. If the mean is based on a larger and larger past history data set, the forecasting becomes more stable, assuming that the underlying process is stationary. Exactly this is the main disadvantage of this method, because there is very unlikely that the business process would be based on an underlying constant process. When the data series is a step function or exhibits trend and seasonality, then the mean used as a forecast for the next period will be inappropriate, because it will give values away from the observed data. [10]

### 2.1.2. Single Moving Average

This model modifies the influence of the past data by specifying the number of the past data observations that will be included in a mean. It is called moving average because as each new observation becomes available, a new average can be computed by dropping the oldest one and adding the newest one. The new average will then be the forecast for the next period. If the data

[10] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.70

$X_1$, $X_2$, $X_3$ …… $X_{N-1}$, $X_N$ is given the forecast for the next period can be compute with the following formula:

$$F_{T+1} = \overline{X} = \sum_{i=1}^{T} X_i / T \qquad (2.2)$$

In this method, the number of data points remains constant in each average and includes the most recent observations. Compared with the mean method the moving average has the following advantages and disadvantages:

- It deals only with the latest periods of known data, but it requires more storage because all of the latest observations must be stored, not just the average.
- It can forecast better trend and seasonality, because the number of data points does not change over time, but the results are still not good enough.

The number of the periods in the moving average is crucial for the forecasting result. If the method contains only one data, then the last observation will be the forecast for the next period. This is also known as a naive forecast. The use of a small number of data will allow the moving average to follow the pattern, but these forecasts will trail the pattern, lagging behind by one or more periods. With the increasing of the data contained in the average a higher smoothing effect is achieved, but the attention paid to the fluctuations in the data series decreases. However, the method can be helpful in decomposing the series into trend, seasonal and other components, but would not be effective as a forecasting tool for data showing trend or seasonality. Because of the disadvantages mentioned above, this method is not very used in practice. The time-series methods used in practice are the exponential smoothing methods, because they are in generally superior.[11]

## 2.2. Exponential Smoothing Methods

The fundamental principle underlying these methods is the idea that the recent data contain more information than the older data. Thus, the averages are weighted with the greatest weight assigned to recent data. Because the weighting given to historical data decreases geometrically as

[11] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.72

we go back in time and the exponential curve is a continuous approximation to geometrically decaying points, these methods are named exponential smoothing.[12]

### 2.2.1. Single Exponential Smoothing

The exponential smoothing assumes that in order to get a new estimate of the average demand, a fraction of the amount by which the demand of this period exceeds the estimate should be added. This fraction is called a "smoothing constant" and is conventionally noted by the Greek letter alpha – α.

The basic rule of the exponential smoothing can be represented in its general form[15]:

$$F_{t+1} = \alpha X_t + (1- \alpha) F_t \tag{2.3}$$

The single exponential smoothing method requires the specification of an α value which must be between 0 and 1. If a small value is chosen for the smoothing constant, say α = 0.01, the response will slow and gradual, since it is based on the average of many past periods. A high value, say α = 0.5, will cause the estimates to respond quickly, not only to real changes, but also to the random fluctuations.[16]

In his earliest book the founder of the exponential smoothing methods, Robert Brown, states that in his practice he had found that an α value of 0.1 is a satisfactory compromise between a very stable system that fails to track real changes and a system that fluctuates with demand.

Although Brown's suggestion, it is common sense that it is not always suitable to keep α = 0.1. That is why over the time different adaptive methods have been developed. For single exponential smoothing methods with one smoothing parameter, there are two main types of adaptive approaches that have been developed. The first group of approaches changes the smoothing parameter α stepwise rather than continuously.

In his method W. Chow (1965) simply replaces at each period *t* the smoothing parameter α, by $\alpha_{t-1} + \delta$ or $\alpha_{t-1} - \delta$**,** where δ is constant, according to which of these values ($\alpha_{t-1} + \delta$, $\alpha_{t-1} - \delta$) would have given a better forecast in the previous period.

---

[12] Silver, Pyke , Peterson , "Inventory Management and Production Planning and Scheduling", 1998, p.89
[15] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, p.86
[16] Brown R., Statistical Forecasting for Inventory Control, 1959, pp. 53

D.C. Whybark (1978) develops another method. He defines three allowed values for α (0 < $\alpha_B$ < $\alpha_M$ < $\alpha_H$ < 1) and a Boolean variable $\bar{\delta}_t$ defined by:

$$\bar{\delta}_t = \begin{cases} 1 & \text{if } |e_t| > 4\sigma, \\ 1 & \text{if } |e_t| > 1.2\,\sigma \text{ and } |e_{t-1}| > 1.2\,\sigma \text{ and } e_t e_{t-1} > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{2.4}$$

where $e_t$ is the forecast error at period t and σ is the standard deviation of series $\{x_t\}$. The adaptive smoothing coefficient takes the predefined values depending on $\delta_t$, as it is shown in the next equation:

$$\alpha_t = \begin{cases} \alpha_H & \text{if } \bar{\delta}_t, \\ \alpha_M & \text{if } \bar{\delta}_t = 0 \text{ and } \bar{\delta}_{t-1} = 1, \\ \alpha_B & \text{otherwise.} \end{cases} \tag{2.5}$$

The values for { $\alpha_H$ , $\alpha_M$ , $\alpha_B$ } recommended by Chow are {0.2, 0.4, 0.8}.

Another method is developed by J.D. Dennis (1978). According to his method, whenever the number $L_t$ of consecutive errors of the same sign is found to be greater than a prespecified limit $N$, α is increased by $\Delta$. Whenever $L_t$ is found to be less than $N$, α returns to a base value $\alpha_B$. The method is described by the following two equations:

$$L_t = \begin{cases} 1 & \text{if } e_t e_{t-1} <= 0, \\ L_{t-1} + 1 & \text{otherwise,} \end{cases} \tag{2.6}$$

$$\alpha_t = \begin{cases} \alpha_B & \text{if } L_t < N, \\ \min\{\alpha_{t-1} + \Delta, 1\} & \text{otherwise.} \end{cases} \tag{2.7}$$

In the second approach, the smoothing parameter is calculated by backward equations.

The most well known approach is developed by Trigg and Leach (1967), which defined the smoothed absolute error as:

$$\Delta_t = \phi|e_t| + (1-\phi)\Delta_{t-1} \tag{2.8}$$

and the smoothed error as

$$E_t = \phi e_t + (1-\phi)E_{t-1} \ , \tag{2.9}$$

where $\phi$ in [0, 1] is some smoothing constant, and α is updated according to the following equation:

$$\alpha_t = |E_t / \Delta_t|^{17} \tag{2.10}$$

### 2.2.2. Holt's Two-Parameter Method

If the demand follows a trend demand model, two parameters need to be estimated, compared to only one in case of a constant demand. This forecasting method uses two smoothing constants, $\bar{\alpha}$ and γ, with values between 0 and 1 and three equations:

$$S_t = \bar{\alpha} X_t + (1-\bar{\alpha})(S_{t-1} + b_{t-1}) \tag{2.11}$$

$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1} \tag{2.12}$$

$$F_{t+m} = S_t + b_t m \, , \tag{2.13}$$

where *S* is the smoothed value, *b* is the trend and *m* is the number of periods that will be forecasted.

The first equation (2.13) adjusts $S_t$ directly for the trend of the previous period, $b_{t-1}$, by adding it to the last smoothed value $S_{t-1}$. In this way, $S_t$ is brought to the approximate base of the current data value. The next equation (2.14) updates the trend, which is expressed as the difference between the last two smoothed values. This is appropriate because if there is a trend in the data, new values should be higher or lower than the previous ones. Because there may be some randomness, the trend of the last period is smoothed with γ and added to the previous estimate of the trend multiplied by (1-γ). This is similar to the single smoothing but applies to the updating of

[17] Pantazopoulos, S., Pappis, C., 1996

the trend. The last equation (2.15) is used to forecast ahead. The trend $b_t$ is multiplied by the number $m$ of periods needed to be forecasted and added to the base value $S_t$.

The Holt's method requires two estimates – one to get the first smoothed value for $S_1$ and the other to get the trend $b_1$. For the first one $S_1=X_1$ can be chosen. The estimation of the trend can be more problematic, because we need an estimate of trend from one period to another. If one chooses the trend to be equal to the differences between the last and previous period's demand, $b_1=X_2-X_1$, there is obviously nothing disturbing. However, if one decides to compute the trend as an average value of the differences between the demands for several periods, $b_1 = \dfrac{(X_2 - X_1) + (X_3 - X_2) + (X_4 - X_3)}{3}$, and if there is a significant drop or raise in the data, one can have trouble with the forecast. So in this case the estimate of the initial slope will be inaccurate and it could take a long time until the forecast overcomes this influence.[18]

### 2.2.3. Winter's Three-Parameter Trend and Seasonality Method

If the data are stationary, then moving averages or single exponential smoothing methods can be applied. The Holt's linear method can be used, if the data exhibit a linear trend. However, if the data are seasonal, these methods cannot give an accurate forecast. For these cases, one appropriate method is the Winter's method.

This method is based on three smoothing equations – one for stationarity, one for trend and one for seasonality. It is similar to Holt's method, but with one additional equation to deal with seasonality. The basic equations for Winter's method are:

$$S_t = \bar{\alpha} \frac{X_t}{I_{t-L\,t}} + (1-\bar{\alpha})(S_{t-1} + b_{t-1}) \qquad (2.14)$$

$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1} \qquad (2.15)$$

$$I_t = \beta \frac{X_t}{S_t} + (1-\beta)I_{t-L} \qquad (2.16)$$

$$F_{t+p+m} = (S_t + b_t m)I_{t-L+m}, \qquad (2.17)$$

[18] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.97

where $S$ is the smoothed value, $b$ is the trend, $L$ is the length of seasonality, $I$ is the seasonal adjustment factor, $m$ is the number of periods that will be forecasted, $p$ is the periodicity of the demand, $\alpha$ is a smoothing constant for the smoothed value, $0 < \alpha < 1$, $\gamma$ is a smoothing constant for the trend, $0 < \gamma < 1$ and $\beta$ is a smoothing constant for the seasonal factor, $0 < \beta < 1$.

The seasonal smoothing equation (2.18) is comparable to a seasonal index that is found as a ratio of the current value of the series, $X_t$, divided by the current single smoothed value of the series, $S_t$. If $X_t$ is larger than $S_t$, the ratio will be greater than one, while if it is smaller than $S_t$, the ratio will be less than one. This is so, because $S_t$ is an average value of the series that does not include seasonality. On the other hand, the data values $X_t$ contain seasonality and include randomness. In order to smooth this randomness the newly computed seasonal factor is weighted with $\beta$ and the most recent seasonal number corresponding to the same season is weighted with $(1-\beta)$. This prior seasonal factor is computed in period $t-L$, since $L$ is the length of seasonality. The deseasonalizing process can presented also as:

actual data – index = deseasonalized data.[19]

The overall smoothing equation (2.16) differs from Holt's method's equation in that the first term is divided by the seasonal number $I_{t-L}$. This is done in order to eliminate seasonal fluctuations from the data values, $X_t$. If $I_{t-L}$ is greater than 1, which occurs when the value in period $t-L$ than the average in its seasonality, dividing $X_t$ by this number greater than 1 gives a value that is smaller than the original value by percentage just equal to the amount that the seasonality of period $t-L$ was higher than average. The opposite adjustment occurs when the seasonality number is less than 1. The value $I_{t-L}$ is used in these calculations because $I_t$ cannot be calculated until $S_t$ is known from the overall smoothing equation.

One of the major problems in using this method is to determining the values for $\alpha$, $\beta$ and $\gamma$ that will minimize Mean Squares Error (MSE) or Mean Absolute Percentage Error (MAPE). The approach for determining these values is usually trial and error, although it might be possible to use nonlinear optimization algorithms to give optimal parameters values.[20]

Another problem in using this method is that it is quite often difficult to distinguish systematic seasonal variations from independent stochastic deviations. The problem is that there are many

[19] Makridakis, Wheelwright, "The Handbook of Forecasting – A Manager's Guide", 1987, p.188; Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.103

[20] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.105

parameters to estimate and the indices may become very uncertain. Therefore, it can be recommended that only data with obvious seasonal variations should be accepted as seasonal items and in these cases the Winter's method should be used.[21]

## 2.4. Measurements of the Forecasting Accuracy

The crucial part for every forecasting method is the accuracy. A decision whether to use a certain method or not is taken by it. Each data has two main components: the functional relationship governing the system, also known as pattern and randomness, known as error.

The critical task in the forecasting is to separate the pattern from the error component so that the former can be used for forecasting. The procedure for estimating the pattern of a relationship is through fitting some functional form in such a way as to minimize the error component of the data. One form of this estimation is least squares. The name least square is based on the fact that this estimation procedure seeks to minimize the sum of the squares errors in the following equation: error = data – pattern.

The mean squares error (MSE) is a mathematical function whose properties can be established using calculus. The following procedure describes the necessary steps of doing it.

For convenience, the error will be denoted by $e$, the data by $X$ and the pattern by $\overline{X}$, the subscript $i$ ($i$=1, 2, 3, ….., $n$) will be added to denote the $i^{th}$ customer, for example.

The first step is to identify the error by the following equation:

$$e = X_i - \overline{X}\,. \tag{2.18}$$

To examine the squared error both sides must be squared, giving:

$$e^2 = (X_i - \overline{X})^2 \tag{2.19}$$

Summing these squared errors for all $n$ customers yield:

$$\overline{\phi} = \sum_{i=1}^{n} e^2 = \sum_{i=1}^{n} (X_i - \overline{X})^2 \tag{2.20}$$

[21] Axsäter Sven, Inventory Control, 2006, p. 21

The value $\overline{X}$ , which will minimize the sum of the squared errors, is not known, but it can be found applying necessary conditions for minimum, i.e. by taking the derivative of $\bar{\phi}$, setting is equal to zero and solving for $\overline{X}$ , as follows:

$$\frac{d\bar{\phi}}{d\overline{X}} = -2\sum(X_i - \overline{X}) = 0$$

so that

$$\sum_{i=1}^{n} X_i - n\overline{X} = 0$$

which implies

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad . \tag{2.21}$$

This solution (2.17) gives a value that minimizes the sum of the squared errors. As a single point estimate of the pattern of the data, the mean fits the data as closely as possible, given the criterion of minimizing the MSE.

Minimizing the MSE is the most popular method for estimating the accuracy of the forecasting method for several reasons. If one attempts to minimize $\Sigma e_i$ one will involve extra complications because some $e_i$ values will be positive and some will be negative. To avoid having errors canceling each other, one might minimize the absolute errors, $\Sigma|e_i|$. However, this is not as easy computationally as minimizing the MSE. Choosing to minimize $\Sigma e_i^4$ or to the high power has the disadvantage that it has more than one minimum, which meets again the computational and practical problems. On the other hand, increasing the power of the error term gives more weight to extreme values which is attractive, because large errors are less desirable than small errors, but it should be overdone. The use of MSE is a compromise between giving too much weight to extreme errors and giving the same weight to all values using the absolute errors, $\Sigma|e_i|$.[23]

Although the MSE is widely used, it has two main drawbacks. The first refers to fitting a model to historical data. Such fitting does not necessarily imply good forecasting, because a MSE of zero can always be obtained in the fitting phase by using a polynomial of sufficiently high order. This leads to overfitting a model to a data series, which is equivalent to including randomness as

[23] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.19

part of the generating process, and fails to indentify the nonrandom pattern in the data. The second main drawback of the MSE as a measure of accuracy is related to the fact that different methods use different procedures in the fitting phase. The smoothing methods, for example, are highly dependent upon initial forecasting estimates. The regression methods, on the other hand, minimize the MSE by giving equal weight to all observations. Thus, the comparison of the methods only on MSE is of limited value.[24]

Other widely used statistical measures are:

Mean Absolute Deviation

$$MAD = \frac{1}{n}\sum\left|X_i - \overline{X}\right| , \tag{2.22}$$

Standard Deviation of Errors

$$SDE = \sqrt{\sum e_i^2 /(n-1)} , \tag{2.23}$$

Mean Absolute Percentage Error

$$MAPE = \sum_{i=1}^{n}\left|PE_i\right| / n , \tag{2.24}$$

where PE refers to Percentage Error and is equal to:

$$PE_t = \left(\frac{X_t - F_t}{X_t}\right)(100) . \tag{2.25}$$

The most common way to describe variations around the mean is through the standard deviation. However, in the practice is common to estimate the Mean Absolute Deviation (MAD) instead. The original reason that MAD is estimated instead of the standard deviation was that this simplified the computations. Nowadays, it is no problem to estimate the standard deviation directly, but still MAD is more preferable to be estimated. In most cases, the standard deviation and the MAD give a very similar picture of the variations around the mean and it is possible to

[24] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.45

relate them to each other. A common assumption is that the forecast errors are normally distributed. In that case the relationship is:

$$\sigma = \sqrt{\pi / 2} MAD \approx 1.25 MAD . \tag{2.26}$$

The MAD can be updated with exponential smoothing method. In this case the forecast for the absolute error at the end of period *t* is consequently determined as:

$$MAD_t = (1 - \overline{\alpha}) MAD_{t-1} + \overline{\alpha} \left| x_t - \overline{x}_{t-1,t} \right| , \tag{2.27}$$

where $0 < \overline{\alpha} < 1$ is a smoothing constant, but not necessarily the same as in the forecasting smoothing method.[25]

Another ratio that evaluates the forecast is the tracking signal (TS). This ratio has two components – the bias and the MAD, and is given as:

$$TS_t = \frac{bias_t}{MAD_t} . \tag{2.28}$$

The bias is used to determinate whether a forecast method consistently over- or underestimates demand. The sum of forecast errors is used to evaluate the bias, where the following holds:

$$Bias_n = \sum_{t=1}^{n} e_t . \tag{2.29}$$

The bias will fluctuate around 0 if the error is truly random and not biased. If the TS at any period is outside the range $\pm 6$, this is a signal that the forecast is biased and is either underforecasting (TS < -6) or overforecasting (TS > +6). In this case, one can decide to choose a new forecasting method in order to improve the forecast.[26]

[25] Axsäter Sven, Inventory Control, pp.29
[26] S. Chopra, P. Meindl, Supply Chain Management, 2007, p. 204

# 3. Causal Forecasting Models

The causal, also called econometric, forecasting methods are estimating techniques based on the assumption that the variable to be forecast, known as dependent variable, has cause-and-effect relationship with one or more other variables, known as , dependent variables.[27]

Econometric models are just one of a number of different ways of characterizing an economic system. They are typically aggregate linear or almost linear models with a well defined stochastic structure. Model parameters are estimated from the data using well understood and statistically techniques based on these stochastic assumptions. The variables modeled are typically measurable and often based on accounting data. The whole system can be modeled through a simultaneous approach, where the variables being modeled are determined jointly or a recursive approach, where the model is built up sequentially.[28]

A special type of econometric models is the linear regression model, which is under consideration in this thesis. This fails among the basic econometric models, concentrating on just a single equation. The forecast of this model will be expressed as a function of a certain number of factors that determine its outcome.

It is important to bear in mind that although regression analysis deals with relationship between a dependent variable and one or more independent variables, it does not necessarily imply causation. This means that it is not necessarily that the independent variables are the cause and the dependent variable is the effect. If causality between two exists, it must be justified from the theory that underlies the phenomenon that is tested empirically.

The objective of the regression analysis may be to estimate the mean value of the dependent variable, when the values of the independent variables are given, or to test hypotheses about the nature of the dependence, which are suggested by the underlying theory, or to predict the mean value of the dependent variable, when the values of the independent variables are given, or one or more of the preceding objectives combined.[29]

There are two types of linear regression methods – the simple regression and the multiple regression.

[27] www.businessdictionary.com

[28] Fildes Robert, 1985

[29] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, p.118

## 3.1. The Simple Regression Model

The term "simple regression" refers to any regression of a single *Y* measure, which is the dependent variable, on a single $\hat{X}$ measure, which is the independent variable. The mathematical expression of this model is:

$$y_i = \hat{\alpha} + \hat{\beta}x_i + \varepsilon_i \qquad \text{for } i=1,\ldots\ldots, n, \qquad (3.1)$$

where $\hat{\alpha}$ is the intercept, $\hat{\beta}$ is the slope and ε is the error term, also known as disturbance.

This is commonly called the population linear regression equation of *y* on *x*. In this equation, *y* is also known as regressand and *x* as regressor.[30]

The slope coefficient, $\hat{\beta}$, measures the rate of change in the mean value of *Y* per unit change in $\hat{X}$. The intercept coefficient, α, is the mean value of *Y*, if $\hat{X} = 0$.[31]

Like every model the simple linear regression comprise some assumptions. The basic set of assumptions are:

1. Zero mean of the distributance: $E[\varepsilon_i] = 0$

   The first assumption states that the factors, which are not explicit included in the model, and therefore subsumed in $\varepsilon_i$, do not systematically affect the mean value of *Y*. In other words, the positive $\varepsilon_i$ values cancel out the negative $\varepsilon_i$ values so that their average affect on *Y* is zero.

2. Homoscedasiticy: $Var[\varepsilon_i] = \sigma^2$, constant for all *i*

   The second assumption states that the variance of the disturbance for each $\hat{X_\iota}$ is some positive constant number equal to $\sigma^2$. This means that the *Y* populations corresponding to various $\hat{X}$ values have the same variance.

3. Nonautocorrelation: $Cov[\varepsilon_i, \varepsilon_j] = 0$ if $j \neq i$

   This assumption states that the disturbances $\varepsilon_i$ and $\varepsilon_j$ are uncorrelated. This means that the deviations of any two *Y* values from their mean value do not exhibit correlated patterns.

---

[30] Green, William, Econometric Analysis, 1993, p.143

[31] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, p.121

4. Uncorrelatedness of regressor and disturbance: Cov [$x_i$, $\varepsilon_j$] = 0 for all *i* and *j*.

   The fourth assumption states that the disturbance and the explanatory variable are uncorrelated. If the assumption is not true and there is correlation, then the explanatory variable, $\widehat{X_\iota}$, will increase or decrease when the disturbance increases or decreases. In these cases, it will be difficult to isolate their influence on *Y*. This assumption is automatically fulfilled if the explanatory variable is nonstochastic and the first assumption holds.

5. Normality: $\varepsilon_i \sim N[0, \sigma^2]$, i. e. $\varepsilon_i$ are supposed to be normally distributed with mean 0 and variance $\sigma^2$.

   If the disturbances are normally distributed, then the third assumption implies that they are independent as well. This assumption is useful for making exact statements about the behavior of estimators and hypothesis testing procedures.[32]

## 3.1.1. The Significance of the Disturbance Term

The deviation of an individual $Y_i$ around its expected value can be express as:

$$\varepsilon_i = y_i - \left(\hat{\alpha} + \hat{\beta} x_i\right), \tag{3.2}$$

where the deviation $\varepsilon_i$ is an unobservable random variable taking positives or negatives values. This random variable, $\varepsilon_i$, is known as stochastic error term or disturbance term.

The disturbance term is a surrogate for all these variables that are omitted from the model but which collectively affect the dependent variable, *Y*. There are many reasons why these variables are not introduced into the model.

One of the reasons is that the theory determining the behavior of *Y* may be incomplete. It might be known that the $\hat{X}$ influences the dependent variable, but other variables affecting *Y* might be ignorant. Even if it is known what some of the excluded variables are and therefore one considers a multiple regression rather than a simple regression, one may not have quantitative information

[32] Green, William, Econometric Analysis, 1993, pp.143; Gujarati, Damodar, 2009 refer to Gujarati, Damodar, Basic Econometrics, 1988, pp.52

about these variables. It is a common experience in empirical analysis that the data that the forecaster would like to have often are not available.

Another reason is that the influence of all or some of the omitted variables might be so small and at best nonsystematic that as a practical matter and for cost considerations it does not pay to introduce them into the model explicitly. Even if all the relevant variables are introduced into the model, there is bound to be some randomness in individual $Y$, which cannot be explained no matter how one tries.

Although the simple regression model assumes that the independent and the dependent variables are measured accurately, in practice the data may be plagued by errors of measurement. For example, the data on $Y$, which let say is quantity demand, may be rounded to the nearest digit. In this case, the disturbance term represents the errors of measurement.

The last but not the least reason is the principle of Occam's razor. It states that the descriptions should be kept as simple as possible until proved inadequate. This means that the regression model should be kept as simple as possible, but the relevant and important variables should not be excluded just to keep the regression model simple. Therefore, even if one knows what other variables might affect $Y$, their influence may be so small that one can represent them through the disturbance term. [33]

## 3.1.2. The Least Squares Method

If the data of the whole population are given, it will be a straightforward task to estimate values of the parameters $\hat{\alpha}$ and $\hat{\beta}$. All it has to be done is to find the conditional means of $Y$ corresponding to each $\hat{X}$ and then join these means. However, in practice is rarely to have the entire population at one's disposal. Often the forecaster has a sample from this population and his task is to estimate the population regression function on the basis of the sample information.[34]

In this method, the unknown parameters of equation (3.1) are the objects of estimation. It is necessary to distinguish between population quantities, such as $\hat{\alpha}$ and $\hat{\beta}$, and the sample

[33] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, p.123; Gujarati, Damodar, 2009 refer to Gujarati, Damodar, Basic Econometrics, 1988, pp.33

[34] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, pp.125

estimates of them, denoted as $\bar{a}$ and $\bar{b}$. The expected estimate of the independent variable from the sample is represented as:

$$\hat{y}_i = \bar{a} + \bar{b}x_i . \tag{3.3}$$

The disturbance associated with the $i^{th}$ data point is:

$$\varepsilon_i = y_i - \hat{y}_i = y_i - \hat{\alpha} - \hat{\beta}x_i . \tag{3.4}$$

The residual which estimates the disturbance for any values of $\bar{a}$ and $\bar{b}$ is:

$$\bar{e}_i = y_i - \bar{a} - \bar{b}x_i . \tag{3.5}$$

Form here follows that:

$$y_i = \hat{\alpha} + \hat{\beta}x_i + \varepsilon_i = \bar{a} + \bar{b}x_i + \bar{e}_i . \tag{3.6}$$

The population quantities $\hat{\alpha}$ and $\hat{\beta}$ are unknown parameters of the probability distribution of $y_i$ whose values can be estimated with the sample data. The parameters $\bar{a}$ and $\bar{b}$ should be so chosen so that the fitted line, $\bar{a} + \bar{b}x_i$, is close to the actual data points. The measure of closeness constitutes a fitting criterion. Although there are several methods that have been suggested, the most used one is least squares.

The method of least squares states that $\bar{a}$ and $\bar{b}$ should be chosen in such a way that the residual sum of squares, $\sum \bar{e}_i^2$ is as small as possible. The algebraically expression of this statement is:

$$Minimize : \sum \bar{e}_i^2 = \sum \left( y_i - \bar{a} - \bar{b}x_i \right)^2 , \tag{3.7}$$

where $a$ and $b$ are the least squares coefficients that minimize this fitting criterion.

The values of the last squares coefficients can be determined by solving the following two simultaneous equations:

$$\sum_i y_i = \bar{n}\bar{a} + \left( \sum_i x_i \right) \bar{b} , \tag{3.8}$$

$$\sum_i x_i y_i = \left( \sum_i x_i \right) \bar{a} + \left( \sum_i x_i^2 \right) \bar{b} , \tag{3.9}$$

where $\bar{n}$ is the sample size.

These simultaneous equations are known as the least squares normal equation. These two equations are obtained from the partial differential of equation (3.8) with respect to $\bar{a}$ and $\bar{b}$. The first order condition for the minimum are:

$$\frac{\partial\left(\sum_i \bar{e}_i^2\right)}{\partial \bar{a}} = -2\sum_i (y_i - \bar{a} - \bar{b}x_i) = 0,$$

or, equivalently, the mean value of the residual is zero:

$$\sum_i \bar{e}_i = 0, \qquad (3.10)$$

and

$$\frac{\partial\left(\sum_i \bar{e}_i^2\right)}{\partial \bar{b}} = -2\sum_i x_i (y_i - \bar{a} - \bar{b}x_i) = 0,$$

which implies that the residual is uncorrelated with the independent variable:

$$\sum_i x_i \bar{e}_i = 0. \qquad (3.11)$$

Solving the normal equations simultaneously, the result for the least squares coefficients is:

$$\bar{a} = \bar{y} - b\bar{x}, \qquad (3.12)$$

$$\bar{b} = \frac{\sum_i x_i y_i - n\overline{xy}}{\sum_i x_i^2 - n\bar{x}^2} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2}, \qquad (3.13)$$

where $\bar{x}$ and $\bar{y}$ are the sample means of $x$ and $y$.[35]

### 3.1.4. The Coefficient of Determination $r^2$

As already mentioned, there will be some positive $\bar{e}_i$ and some negative $\bar{e}_i$ around the sample regression line, which is presented with the following equation: $y_i = (\bar{a} + \bar{b}x_i) + \bar{e}_i = \hat{y}_i + \bar{e}_i$. There are n pairs of values $(Y_i, \hat{Y}_i)$ and it is of great interest to know how these two values related to each other. The correlation between these $Y_i$ and $\hat{Y}_i$ is usually designed with *R*. It is usually to present this correlation in squared form and this statistic is known as the coefficient of determination.

[35] Green, William, Econometric Analysis, 1993, pp.148; Gujarati, Damodar, 2009 refer to Gujarati, Damodar, Basic Econometrics, 1988, pp.48

The coefficient of determination $r^2$ (for simple regression) or $R^2$ (for multiple regression) is a summary measure which tells how well the sample regression line fits the data or in other words how large are these residuals around the sample regression line. In the language of mathematics, the coefficient of determination tells the proportion of variance in $Y$ that can be explained by $\hat{X}$. The dependent variable $Y$ has a certain amount of variability, which is defined by its variance. The estimated $\hat{Y}$ values also have a certain amount of variance. The ratio of these two variances is r$^2$:

$$r^2 = \frac{Var(\hat{Y})}{Var(Y)}. \tag{3.14}$$

Since the $\hat{Y}$ are defined with reference to the estimated regression equation, this may be expressed as follows:

$$r^2 = \frac{\text{exp}\,lained \quad Var(Y)}{total \quad Var(Y)}. \tag{3.15}$$

Since the variation of the dependent variable is defined in terms of deviations from its mean, the total variation in $Y$ is the sum of squared deviations:

$$TSS = \sum_i \left(y_i - \overline{y}\right)^2. \tag{3.16}$$

In terms of the regression equation, it may be written:

$$y_i = \hat{y}_i + \overline{e}_i = \overline{a} + \overline{b}x_i + \overline{e}_i = \overline{y} - \overline{b}\overline{x} + \overline{b}x_i + \overline{e}_i.$$

Subtracting $\overline{y}$ from both sides gives:

$$\left(y_i - \overline{y}\right) = \hat{y}_i - \overline{y} + \overline{e}_i = \overline{b}(x_i - \overline{x}) + \overline{e}_i. \tag{3.17}$$

Squaring on both sides it yields:

$$\sum_i \left(y_i - \overline{y}\right)^2 = \sum_i \left(\hat{y}_i - \overline{y}\right)^2 + \sum_i \overline{e}_i^{\,2} = \overline{b}^2 \sum_i \left(x_i - \overline{x}\right)^2 + \sum_i \overline{e}_i^{\,2}. \tag{3.18}$$

After these definitions, equation (3.20) can be transformed into following equation:

$$r^2 = \frac{ESS}{TSS} = \frac{\overline{b}^2 \sum_i \left(x_i - \overline{x}\right)^2}{\sum_i \left(y_i - \overline{y}\right)^2}. \tag{3.19}$$

The quantity $r^2$ thus defined is known as the sample coefficient of determination and is the most commonly used measure of the goodness of fit of a regression line. It tells the proportion of variation in the dependent variable explained by the independent variable and therefore provides an overall measure of the extend to which the variation in one variable determine the variation in the other.

Two properties of these coefficients can be noted:

1. It is a nonnegative quantity.
2. Its limits are $0 \le r^2 \le 1$. An $r^2$ of 1 means a perfect fit, whereas an $r^2$ of zero means no relationship between the dependent variable and the explanatory variable.[37]

## 3.1.5. The Significance of the Simple Regression Equation

### 3.1.5.1. The F-Test of Overall Significance

The F-test answers the statistical question if there is a significant relationship between the explanatory variable and the dependent variable. In other words, it tests the significance of the overall regression model. This test helps the forecaster to decide whether the chosen explanatory variable is the proper one.

The F-test can be presented as a ratio between two mean squares as follows:

$$\hat{F} = \frac{MS_{\text{explained}}}{MS_{\text{unexplained}}} = \frac{(SS_{\text{explained}})/(df_{\text{explained}})}{(SS_{\text{unexplained}})/(df_{\text{unexplained}})} = \frac{\sum(\hat{Y} - \overline{Y})^2 / (\overline{k} - 1)}{\sum(Y - \hat{Y})^2 / (\overline{n} - \overline{k})}, \tag{3.20}$$

where $\overline{k}$ is the number of parameters, which for the case of simple regression is $\overline{k} = 2$, MS = mean square, SS = sum of squares, df = degrees of freedom.

The F-test is closely connected to the definition of the coefficient of determination, $r^2$. Thus it can be presented with another formula, as follows:

$$\hat{F} = \frac{r^2 / (\overline{k} - 1)}{(1 - r^2)/(\overline{n} - \overline{k})}. \tag{3.21}$$

[37] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.217; Green, William, Econometric Analysis, 1993, pp.150; Gujarati, Damodar, 2009 refer to Gujarati, Damodar, Basic Econometrics, 1988, pp.64

It should be pointed out that in the case of simple regression, the F-test for overall significance is the same as testing the significance of the slope coefficient.[38]

3.5.1.2. The t-Tests for Individual Coefficients

The values of the coefficients $\bar{a}$ and $\bar{b}$ fluctuate from sample to sample. The pair of values ($\bar{a}$, $\bar{b}$) have a joint sampling distribution and there is a very strong negative correlation between $\bar{a}$ and $\bar{b}$, because if one increases the slope ($\bar{b}$), one automatically decrease the intercept ($\bar{a}$) and the opposite. In order to investigate the stability of $\bar{a}$ and $\bar{b}$, separately, the marginal distribution have to be considered. The sampling distribution of the intercept coefficient has a normal distribution with mean $\hat{\alpha}$ and standard error as follows:

$$se_a = \hat{\sigma}_e \left\{ \frac{\sum X_i^2}{\bar{n} \sum \left(X_i - \overline{X}\right)^2} \right\}^{1/2} . \tag{3.22}$$

The sampling distribution of the slope coefficient has a normal distribution with mean $\hat{\beta}$ and standard error as follows:

$$se_a = \hat{\sigma}_e \sqrt{\frac{1}{\sum (X_i - \overline{X})^2}} . \tag{3.23}$$

The estimate of the standard deviation of the errors in these two equations is given by:

$$\hat{\sigma}_e = \sqrt{\frac{\sum \left(Y_i - \hat{Y}_i\right)^2}{\bar{n} - 2}} . \tag{3.24}$$

This test is a procedure by which sample results are used to verify the truth or falsity of a null hypothesis. The null hypothesis is used to find out whether $Y$ is related to $\hat{X}$at all. If there is no relationship between variables, then there will be no point to include the explanatory variable in the model. Therefore, if $\hat{X}$ belongs in the model, one would fully expect to reject the null hypothesis $H_0$. The key idea of the t-test of significance is that of a test statistic and the sampling distribution of such a statistic under the null hypothesis. The decision to accept or reject $H_0$ is based on the value of the test statistic obtained from the data.

[38] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.223

Using the standard errors formulas in equations (3.28) and (3.29) two t-tests can be set up to test the intercept and slope values as follows:

- t-value for interceptor

$$t_{n-2} = \frac{\bar{a} - \hat{\alpha}}{se_a}, \tag{3.25}$$

where $\bar{a}$ is the estimated value, $\hat{\alpha}$ is the hypothesized value, $se_a$ is the standard error of $\bar{a}$ and the subscript on t indicates the degree of freedom.

- t-value for slope

$$t_{n-2} = \frac{\bar{b} - \hat{\beta}}{se_b}, \tag{3.26}$$

where $\bar{b}$ is the estimated value, $\hat{\beta}$ is the hypothesized value, $se_b$ is the standard error of $\bar{b}$ and the subscript on t indicates the degree of freedom.

Using the null hypothesis, it can be assumed that

$H_0$: $\hat{\beta} = \beta^*$

where $\beta^*$ is a specific numerical value of $\hat{\beta}$ (e.g., $\beta^* = 0$), then

$$t_{n-2} = \frac{\bar{b} - \beta^*}{se_b}, \tag{3.27}$$

can be computed from the sample data. Since all quantities in the latter equation are known, the t-value can be computed and used as test statistic, which follows the t distribution with $n$-2 degrees of freedom.

In order to use the t-test in a concrete application, which is used for the intercept coefficient as well as for the slope coefficient, three things should be known:

1. The degrees of freedom, which are always $n$-2 for the simple regression model.
2. The level of significance, $\hat{\alpha}$, which determines whether the tested parameter is significantly different from zero. The level of significance is a matter of personal choice, although 1%, 5% or 10% levels have been usually used in empirical analysis. If the t-value lies outside the value of 1% critical t-value, then the hull hypothesis can be rejected and the explanatory variable can be considered as highly significant. If the

t-value lies close to the 5% or 10% critical t-value, the forecaster must decide whether the explanatory variable is significant or not.

3. Whether to use a one-tailed or two-tailed test. The t-testing procedure remains the same in both cases, however the problem is whether the probability of an error is equally divided between the two tails of the t-distribution or it is concentrated in only one tail, either left or right.[39]

## 3.1.6. Forecasting Using the Simple Regression Model

The most common use of the regression model is for prediction. Once the demand function is obtained, based on the historical data, the forecaster can predict the demand of a given good. There are two kinds of prediction: (1) prediction of the conditional mean value of $Y$ corresponding to a chosen $\hat{X}$, say $\hat{X}_0$, that is the point on the population regression line itself and (2) prediction of an individual Y value corresponding to $\hat{X}_0$.

Supposing that $\hat{X}_0$ is a known value of the regressor, the $Y_0$ value can be predicted. This is the value $Y$ associated with $\hat{X}_0$. The predicted true value of the independent variable yields:

$$y_0 = \hat{\alpha} + \hat{\beta} x_0 + \varepsilon_0 . \qquad (3.28)$$

In this case, two sources of error should be kept in mind. One source of error will be the sampling error of parameters estimates. But regardless of the precision of the parameter estimates, the forecast error, $\varepsilon_0$, can never be forecasted perfectly.

The forecast will be:

$$\hat{y}_0 = \bar{a} + \bar{b} x_0 . \qquad (3.29)$$

Since $\bar{a}$ and $\bar{b}$are both random variables having a joint probability distribution, the standard error of $\hat{Y}_0$ can be determined as follows:

1. Standard error of $\hat{Y}_0$ as a mean prediction

[39] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.224; Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essential of Econometrics, 1992, pp.153

$$se_{\hat{Y}_0 \ as \ a \ mean} = \hat{\sigma}_e \sqrt{\frac{1}{n} + \frac{\left(x_0 - \bar{x}\right)^2}{\sum x_i^2}} \ . \tag{3.30}$$

Applying the t-test, it yields:

$$t = \frac{\hat{y}_0 - (\hat{\alpha} + \hat{\beta} x_0)}{se_{\hat{Y}_0 \ as \ a \ mean}}, \tag{3.31}$$

where the variable $t$ follows the $t$-distribution with N-2 degree of freedom. Therefore, the t-test can be used to derive confidence intervals for the true $Y_0$ value and test hypothesis about it.

2. Standard error of $\hat{Y}_0$ as an individual prediction

$$se_{individual \ \hat{Y}_0} = \hat{\sigma}_e \sqrt{1 + \frac{1}{n} + \frac{\left(x_0 - \bar{x}\right)^2}{\sum x_i^2}} \ . \tag{3.32}$$

Applying the t-test, it yields:

$$t = \frac{y_0 - \hat{y}_0}{se_{individual \ \hat{Y}}}, \tag{3.33}$$

where the variable $t$ follows the t-distribution with N-2 degree of freedom. Therefore, the t-test can be used to draw inference about true $Y_0$ value.

It should be mentioned that the forecasting ability of the historical sample regression line falls markedly as $\hat{X}_0$ departs progressively from $\overline{X}$. Therefore, a great caution should be paid in extrapolating the historical regression line to predict $Y_0$.[40]

## 3.2. The Multiple Regression Model

The simple regression model, mentioned previously, is a special case of multiple regression model. In multiple regression there is one dependent variable to be predicted, but there are two or

[40] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.226; Green, William, Econometric Analysis, 1993, pp.164; Gujarati, Damodar, 2009 refer to Gujarati, Damodar, Basic Econometrics, 1988, pp.119

more independent variables that has been identified to have influence on the dependent variable. The general form of multiple regression is:

$$y_i = \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + ... + \hat{\beta}_k x_{ik} + \varepsilon_i , \tag{3.34}$$

where *Y* is the dependent variable, $\hat{X}_1 \ldots \hat{X}_k$ are the explanatory variables, $\varepsilon$ is the stochastic disturbance term, *i* corresponds to the $i^{th}$ observation and $\hat{X}_{i1} = 1$ for each observation.

The coefficient $\beta_1$ is the intercept term and as before, it represents the average value of *Y* when $\hat{X}_2 \ldots \hat{X}_k$ are equal to zero. The coefficients $\hat{\beta}_2 \ldots \hat{\beta}_k$ are called partial regression coefficients or partial slope coefficients. The meaning of the partial slope coefficients is that $\hat{\beta}_2$ measures the change in the mean value of *Y* per unit change in $\hat{X}_2$, holding the value of the other independent variables constant. Likewise, every $\hat{\beta}_k$ partial slope coefficients measures the change in the mean value of *Y* per unit change in $\hat{X}_k$. This unique feature of multiple regression enables not only more than one explanatory variables to be included in the model but also the effect of each $\hat{X}$ variable on *Y* to be isolated.

As in the simple regression model, the multiple regression model includes some basic assumptions. The only difference here is the number of the right-hand variables. The assumptions are:

1. The explanatory variables $\hat{X}_1, \hat{X}_2, \ldots \hat{X}_k$ are nonstochastic and their values are fixed in repeated sampling.
2. The error term $\varepsilon$ has a zero mean value: $E[\varepsilon_i] = \begin{bmatrix} E[\varepsilon_1] \\ E[\varepsilon_2] \\ \vdots \\ E[\varepsilon_n] \end{bmatrix} = 0$ .
3. Homoscedasticity, that is, the variance of $\varepsilon$ is constant:

$$\operatorname{var}(\varepsilon_i) = E[\varepsilon\varepsilon'] = \begin{bmatrix} E[\varepsilon_1\varepsilon_1] & E[\varepsilon_1\varepsilon_2] & \cdots & E[\varepsilon_1\varepsilon_n] \\ E[\varepsilon_2\varepsilon_1] & E[\varepsilon_2\varepsilon_2] & \cdots & E[\varepsilon_2\varepsilon_n] \\ \vdots & \vdots & \vdots & \vdots \\ E[\varepsilon_n\varepsilon_1] & E[\varepsilon_n\varepsilon_2] & \cdots & E[\varepsilon_n\varepsilon_n] \end{bmatrix} = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix} = \sigma^2 \hat{I} .$$

4. No autocorrelation exist between the error terms $\varepsilon_i$ and $\varepsilon_j$: $Cov(\varepsilon_i, \varepsilon_j) = 0$ .
5. Uncorrelatedness of regressor and disturbance: Cov [$x_i$, $\varepsilon_j$] = 0 for all *i* and *j*.

6. No exact collinearity exists between the explanatory variables. This means that there is no exact linear relationship between the explanatory variables. This is the only assumption that differs from the simple regression assumptions and it implies that none of the explanatory variables can be expressed as an exact linear function of another variable.
7. The error term follows the normal distribution with mean zero and homoscedastic variance: $\varepsilon_i \sim N[0, \sigma^2 I]$.[41]

### 3.2.1. The Method of Least Squares

As for the simple regression model, the most common method of estimating the parameters of the multiple regression model is the least squares method.

The first step in finding the least squares estimators is to write the sample regression function corresponding to the population regression function, as follows:

$$y_i = \bar{b}_1 x_{i1} + \bar{b}_2 x_{i2} + \ldots + \bar{b}_k x_{ik} + \bar{e}_i = \hat{y}_i + \bar{e}_i, \qquad (3.35)$$

where $\begin{bmatrix} \bar{b}_1 \\ \cdot \\ \cdot \\ \bar{b}_k \end{bmatrix} = the\ \ estimator\ \ of \begin{bmatrix} \hat{\beta}_1 \\ \cdot \\ \cdot \\ \bar{\beta}_k \end{bmatrix}$, $\bar{e}_i$ is the residual term, $\hat{Y}_i$ is the estimate of $Y_i$ and $\hat{X}_{i1} = 1$.

The least squares principle chooses the values of the unknown parameters in such a way that the residual sum of squares (RSS) is as small as possible. Rewriting the latter equation and expressing the error term, it yields:

$$\bar{e}_i = y_i - \hat{y}_i.^{42} \qquad (3.36)$$

Applying the least squares method, it yeields:

$$\text{minimize } \bar{\phi} = \sum \bar{e}_i^{2} = e'\hat{e} = (y_i - \hat{y}_i)'(y_i - \hat{y}_i) = (\breve{y}_i - \tilde{b}\breve{x})(\breve{y}_i - \tilde{b}\breve{x}) \qquad (3.37)$$

[41] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.247; Green, William, Econometric Analysis, 1993, pp.170; Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, pp.182

[42] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.256; Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, p.188

where $e' = (y_i - \hat{y}_i)'$ is the transpose of $\bar{e}$, $\hat{y}_i = \tilde{b}x$, $\breve{y}$ is an $\bar{n}$ *x* k matrix, $\breve{x}$ is an $\bar{n}$ *x* k matrix, $\tilde{b}$ is a *k x* 1 matrix and $\hat{e}$ is a *k x* 1 matrix.

This problem is solved by taking partial derivatives of $\phi$ with respect to $\tilde{b}$, which yields the solution:

$$\bar{b} = (\breve{x}'\breve{x})^{-1}\breve{x}'\breve{y} \tag{3.38}$$

where $(\breve{x}'\breve{x})^{-1}$ is the inverse of $(\breve{x}'\breve{x})$.

In order to verify that this is indeed a minimum, the following equation is required to be a positive definite matrix:

$$\frac{\partial^2 S(\bar{b})}{\partial\bar{b}\partial\bar{b}'} = 2\hat{X}'\hat{X} \tag{3.39}$$

Let $q = c'\hat{X}'\hat{X}c$ for some arbitrary nonzero vector, *c*. Then:

$$q = v'v = \sum_i v_i^2 , \tag{3.40}$$

where $v = \hat{X}c$.

For every element of *v* different from zero, *q* is positive. However, if *v* is zero, *v* would be a liner combination of the columns of $\hat{X}$ that equals zero. This contradicts the assumption that $\hat{X}$ has full rank.[43]

## 3.2.2. The Multiple Coefficient of Determination, $R^2$

In the simple regression case the coefficient of determination $r^2$ measures the goodness of fit of the fitted sample regression line that gives the proportion of the total variation in the independent variable explained by the single explanatory variable. This concept can be extended to multiple regression model containing any number of explanatory variables ($\hat{X}_1 \ldots \hat{X}_k$). The quantity that gives this information is conceptually similar to $r^2$ and is known as the multiple coefficient of determination, and is denoted by the symbol $R^2$.

Also, as in the simple regression model, $R^2$ is defined as:

[43] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.256; Green, William, Econometric Analysis, 1993, p.173

$$R^2 = \frac{ESS}{TSS} = \frac{\sum\left(\hat{Y}_i - \overline{Y}\right)^2}{\sum\left(Y_i - \overline{Y}\right)^2}, \qquad (3.41)$$

where TSS = the total sum of squares of the dependent variable and ESS = the explained sum of squares (explained by all the independent variables).

Like $r^2$, $R^2$ also lies between 0 and 1. The closer it is to 1, the better is the fit of the estimated regression line. The unexplained part refers to the factors that influence the model but are not included in it.[46]

An important property of the coefficient of determination is that the larger the number of explanatory variables is, the higher the value of $R^2$ will be. This is because the definition of $R^2$ = ESS/TSS does not take into account the degrees of freedom. Thus, a measure of goodness of fit that is adjusted for the number of explanatory variables in the model is needed. Such a measure has been created and it is known as the adjusted $R^2$, denoted by the symbol $\overline{R}^2$. It can be derived from the conventional $R^2$, as follows:

$$\overline{R}^2 = 1 - \left(1 - R^2\right)\frac{n-1}{n-k}, \qquad (3.42)$$

where $k$ is the number of the independent variables.

The features of the adjusted $R^2$ are:

1. If k > 1, $\overline{R}^2 \le R^2$. This means that as the number of explanatory variables increases, the adjusted coefficient of determination becomes increasingly less than the unadjusted coefficient of determination.
2. Although the unadjusted $R^2$ is always positive, the adjusted $\overline{R}^2$ can on occasion turn out to be negative.[47]

[46] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.257; Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, pp.194

[47] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, p.205

### 3.2.3. Significance of the Multiple Regression Equation

#### 3.2.3.1. The F-Test for Overall Significance

As in the simple regression case, the F-test answers the statistical question if there is a significant relationship between the explanatory variable and the dependent variable. The F-test can be presented as a ratio between two mean squares as follows:

$$F = \frac{MS_{\text{explained}}}{MS_{unexplained}} = \frac{(SS_{\text{explained}})/(df_{\text{explained}})}{(SS_{unexplained})/(df_{unexplained})} = \frac{\sum(\hat{Y}-\overline{Y})^2/(\overline{k}-1)}{\sum(Y-\hat{Y})^2/(\overline{n}-\overline{k})}, \qquad (3.43)$$

where $\overline{k}$ is the number of parameters including the intercept, MS = mean square, SS = sum of squares, df = degrees of freedom.

It can be seen from the latter equation that the F-test is sensitive to the relative strength of the numerator and denominator. If the unexplained MS, which is the variance of the errors, is large, then $\hat{F}$ becomes smaller. If the explained MS is large relative to the unexplained MS, then $\hat{F}$ becomes larger.

As in simple regression, there is an important relationship between *F*-test and the adjusted coefficient of determination, $\overline{R}^2$. The relationship can be presented, as follows:

$$\hat{F} = \frac{\overline{R}^2/(\overline{k}-1)}{(1-\overline{R}^2)/(\overline{n}-\overline{k})} \qquad (3.44)$$

The direct relationship between the two statistics can be seen from the latter equation. When $\overline{R}^2 = 0$, which means that there is no relationship between the dependent and the independent variables, *F* will be also zero. When $\overline{R}^2$ increases, $\hat{F}$ will increase too and its value will be infinity when the adjusted coefficient of determination equals one.[48]

#### 3.2.3.2. The t-Tests for Individual Partial Regression Coefficients

Although the coefficient of determination gives an overall measure of goodness of fit of the estimated regression line, it does not give information whether the estimated partial regression coefficients are statistically significant, which means that the partial coefficients are tested

[48] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, p.203

against the value zero. A t-test on an individual partial coefficient is a test of its significance in the presence of all other regressors. If the value is zero, it would mean that the independent variable in question is not helping at all in the prediction of the dependent variable in the presence of the other regressors.

As in the simple regression case, for each estimator $\bar{b}_j$ a standard error can be determined and given the normality assumption in the regression model, it follows that $t$ has t-distribution with $(\bar{n}-k)$ degrees of freedom, as follows:

$$t = \frac{\bar{b}_j - \hat{\beta}_j}{se(b_j)} \approx t_{(\bar{n}-k)}, \tag{3.45}$$

where $\bar{b}_j$ = estimated $j^{th}$ coefficient, $\hat{\beta}_j$ = hypothesized $j^{th}$ parameter, $se(b_j)$ = standard error of $\bar{b}_j$, k = number of independent variables.

The use of the t-test is meaningful, because if the F-test indicates a significant regression line, it will be expected that at least one of the t-test would also be significant. However, this is not always true, which raises the question whether the regressors are properly chosen. Additionally, considering the t-test, the stability of the partial coefficient can be tested. Their stability depends upon the intercorrelation among the independent variables. The higher the correlation between them is, the more unstable will be the partial coefficients. Another aspect to be considered is the estimated correlation among the partial regression coefficients themselves. Since the partial regression coefficients are all random variable, which fluctuate form sample to sample and have joint probability distribution, it is possible the correlations among the coefficients to be determined. In the simple regression case it is pointed out that the slope coefficient and the intercept are always going to be negatively correlated because the regression line goes through the mean of $Y$ and the mean of $\hat{X}$ and an increase in the slope automatically means a decrease in the intercept and the opposite. In the multiple regression case, it is more complicated, but, if for instance, two partial coefficients are found to be positively or negatively significant correlated, then the forecasted should be warned that the individual t-test on these coefficients should not be considered in isolation of each other, because these two coefficients are dependent on each other.[49]

### 3.2.3.3. Selecting Independent Variables

The forecaster is often faced with the decision problem which explanatory variable to include into the model and which to omit. Gujarati (1992) introduced one decision method. He suggested that variables should be added as long as the adjusted $R^2$ increases. For this purpose he examines whether the absolute t-value of the coefficient of the added variable is large than one, where the t-value is computed under the hypothesis that the population value of the said coefficient is zero. If the absolute value is larger than one, then the adjusted $R^2$ will increase and the independent variable should be added to the model.[50]

Makridakis, Wheelwright and McGee (1983) introduced alternative approaches. One of these is to consider the intercorrelations among the regressors of all potential candidates and every time when a large correlation is encountered to remove one of the two variables from further consideration. Another approach is to construct a multiple linear regression on all the regressors and to disregard all variables whose t-values are very small. According to them the t-value should be considered as small if the absolute t-value is smaller than 0.5.

However, if the forecaster is not satisfied with the value of the adjusted $R^2$, then the problem may be not in the chosen variables, but in the seasonality. For this case Makridakis, Wheelwright and McGee (1983) also introduced an approach. They suggested that dummy variables, each of which with only two allowed values – zero or one, should be added to model. One dummy variable is added per period, respectively with value one if the variable corresponds to the given period and zero otherwise ($D_1$=1, for the first period and 0 otherwise; $D_2$=1, for the second period and 0 otherwise and so on). In order to avoid multicollinearity ($P$-1) dummy variables should be introduced, where $P$ denotes the number of the periods. If there is seasonality then the value of the adjusted $R^2$ will improve. However, two factors regarding this approach should be considered. The first one is that four seasonal dummy variables for four quarters will result in perfect

---

[49] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.263; Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, pp.196

[50] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, p. 206

nulticollinearity. The second one is that each new dummy variable is a new regressor, which requires the estimation of another regression coefficient and thereby one degree of freedom will be lost. This eventually could result in lowering the adjust coefficient of determination than increasing it.[51]

### 3.2.3.4. Forecasting Using the Multiple Regression Model

In the simple regression case it was already defined the standard error of forecast for $\hat{Y}$ as a mean and as a single point. For the general case of multiple regression model, the standard error are, as follows:

$$se_{(\bar{Y}\ as\ a\ mean)} = \hat{\sigma}_{\varepsilon}\sqrt{c'\left(\breve{X}'\breve{X}\right)^{-1}c} \tag{3.46}$$

$$se_{(\bar{Y}\ as\ a\ point)} = \hat{\sigma}_{\varepsilon}\sqrt{1-c'\left(\breve{X}'\breve{X}\right)^{-1}c} \tag{3.47}$$

where c is the vector [ $X_1^* \ X_2^* \ \ldots . \ X_k^*$ ] of new values for the regressors and $\breve{X}$ is a $\bar{n}$ $x$ $k$ matrix of rank $K$.

For any forecast to be made, a set of values for the regressors has to be provided (the $X_1^* \ X_2^* \ \ldots \ X_k^*$ values). These are then put into the regression equation and a predicted value, $\hat{Y}$, is obtained. The independent variable values often have to be forecasted before dependent variable values can be forecasted. Therefore, it is important to get good forecasts for these independent variables. The equations (3.44) and (3.45) are used to evaluate the accuracy of the prediction. It should be pointed out that these latter equations are based on the assumption that the regressors are measured without an error. When forecasts of $\hat{Y}$ are made, they often depend on forecast of the regressors, so that these regressors values are definitely subject to error, and the standard error formulas underestimate the actual forecast error.[52]

The mathematical expression of the model is as follows:

$$\hat{y}_0 = \bar{a} + \bar{b}_1 x_1 + \bar{b}_2 x_2 ..... + \bar{b}_k x_k = b' x^0 \tag{3.48}$$

[51] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.277

[52] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.296

where $x^0$ is the regressor vector.[53]

Although the regression analysis is a powerful method of estimation and commonly used causal approach to forecasting, it has its disadvantages. One reason, why multiple regression models might be expected lead to less accurate forecasts than alternative methods, is that the models may be too complex in that they include representations of false patterns in the noise associated with past data. For example, Gigerenzer and Todd (2000) have suggested that multiple linear regression (MLR) tends to lead to models, which are overfitted to past observations so that hold-out sample forecasts are relatively inaccurate.

On the other hand, MLR models may be too simple. Because they are, by definition, linear they will be unable to represent non-linear relationships between variables. For example, this may occur in situations where the outcome depends on a non-compensatory combination of cue values (e.g. a low score on one cue might determine the outcome irrespective of whether the other cues have high or low values) or where outcomes depend upon products of some of the cue values rather than their weighted sums.

A third possibility is that changes in the environment mean that the structure of the data used to derive the MLR model differs from that that which applies in the forecasting periods. In these periods, new relationships between the outcomes and cues might apply or special, rare events may occur which mean that the normal relationships are temporarily suspended.[54]

However, as Chiasson, Fildes and Pidd (2006) pointed out, it should be kept in mind that regression models are always simplifications and far from being a limitation of the reality. The essential point is that the regression models are used by people and their value lies in a combination of the models and the people who build and use them. It should be distinguished between the use of a model with the model itself, though the two are clearly linked. Forecasters use regression models in many ways: for example, to explore situations, to support automated decision-making and to enable people to think through options. In all cases, causation is something that the modeler and model user must set their minds to consider, because it does not spring from the model.

Modelers do not and cannot, in practice, pursue inconsequential relationships between butterfly wings and seasonal affective disorder. It is doubtful that practitioners would regard modeling as

---

[53] Green, William, Econometric Analysis, p.195

[54] K. Nikolopoulos, P. Goodwin, A. Patelis, V. Assimakopoulos, 2007

just a means to the discovery of patterns from easily available data without some practical or theoretical goal. Rather, a model may contain causal relationships and some of these may be simple and easily testable (e.g. increasing the service time in a simulation model will lead to longer queues if demand stays the same). Complex forecasting models are used automatically, for example, managing electricity demand. They are also used to gain understanding, as users testify in macroeconomic policy, while simple exponential smoothing models provide the basis for judgmentally incorporating market intelligence in supply chain planning.

It should not also be forgotten that the regression models are no substitute for intelligent thinking with the purpose to find out causal relationships and to forecast.[55]

[55] Chiasson, R. Fildes, M. Pidd, 2006

# 4. Safety Stock Planning

In the previous chapters of this thesis, some basic methods for forecasting of the demand were discussed. However, the accurate demand forecast is only one part of the successful inventory management. The inventory level is a crucial management decision in every company, because the company must have enough items to satisfy the demand. At the same time, the inventory level should not be too high, because the company locks capital in stock and thus reducing its liquidity. The lack of liquidity can be overwhelming for the company particularly in tough times. The reason for this is that the inventory is the balance sheet asset that is most subject to fraudulent overstatement and banks seldom finance more than 50% of inventory value.[56] Other consequences of keeping the inventory level too high are the lost of floor space which is a valuable asset, reducing the cash through increasing the insurance and tax expenses, and the opportunity lost on the funds invested in the excess inventory.[57]

On the other hand, variations between actual demand and forecast are inevitable. That is why, a tool that provides protection against inventory imbalances is need. This tool is called safety stock. However, in order to set the safety stock levels one needs the standard deviation of forecast errors. For this purpose, the relationship between the standard deviation and the true MAD (the average absolute deviation from the mean) can be used. Although this relationship is not simple, it can be shown that for the normal distribution the following equation holds:

$$\sigma = \sqrt{\pi/2}(true \quad MAD) \approx 1.25(true \quad MAD) \,. \tag{4.1}$$

For several other common distributions the theoretical conversion factor form the MAD to $\sigma$ is not very different from 1.25. However, the factor can vary enough so that the safety stock obtained may not provide the required level of service.[58]

In this chapter, the safety stock will be examined and methods to calculate the safety stock using statistical data will be discussed. Additionally, $(\hat{t}, \hat{S})$ inventory policy will be presented and on its base the different types of forecasting methods will be compared in Chapter eight.

[56] W. H. Wiersema, 2008

[57] R. E. Dillon, 1990

[58] Silver E. A., Pyke D. F., Peterson R., Inventory Management and Production Planning and Scheduling, 1998, pp. 111

## 4.1. What is Safety Stock

In order to be given a precise definition of the safety stock, it will be meaningful the different categories of inventory to be presented. Zipkin (2000) categorizes inventories, as follows:

- $\tilde{I}(\tilde{t})$ = inventory at time $\tilde{t}$

  This inventory's category is also known as On-hand stock and represents the stock that is physically on the shelf. This quantity determines whether a particular customer demand is satisfied directly from the shelf and is always positive.[59]

- IO($\tilde{t}$) = inventory on order

  This is the total stock ordered before $\tilde{t}$ but not yet received by $\tilde{t}$.

- IN($\tilde{t}$) = net inventory at time $\tilde{t}$ = $\tilde{I}(\tilde{t})$ – B($\tilde{t}$), where B($\tilde{t}$) = backorders at time $\tilde{t}$

- IP($\tilde{t}$) = inventory position at time $\tilde{t}$ = IN($\tilde{t}$) + IO($\tilde{t}$)

The net inventory captures information in the inventory and in the backorders. Since any available stock is used to fill demand, at any given point, at least the inventory or the backorder function is zero. Therefore, I($\tilde{t}$) = $[\mathrm{IN}(\tilde{t})]^{+}$ and B($\tilde{t}$) = $[\mathrm{IN}(\tilde{t})]^{-}$. [60]

As it can be seen, the net inventory definition treats backorders as negative inventories. Therefore, a buffer stock is needed to satisfy the customer demand. Silver, Pyke and Peterson (1998) define safety stock as "the average level of net inventory just before a replenishment arrives".[61]

Over many reorder cycles, the inventory or the stock on hand will sometimes be positive quantity and sometimes negative quantity. However, this is true only for the case, when the demand that occurs in out of stock situation is backordered and filled as soon as replenishment arrives. This case is suitable for some supply systems, but it is not the general case. For many products, the customer does not want to wait until the replenishment and goes elsewhere to satisfy his need. In

[59] Silver E.dward A., Pyke D. F., Peterson R., Inventory Management and Production Planning and Scheduling, 1998, p. 233

[60] P. Zipkin, Foundtion of Inventory Managenent, 2000, p. 40

[61] Silver E. A., Pyke D. F., Peterson R., Inventory Management and Production Planning and Scheduling, 1998, p. 233

this case, the demand that occurs in out of stock situation is lost, instead of being back ordered. Here the safety stock will be effectively the average of only the positive quantities, since stock outs results in a zero balance, rather than a negative balance.[62]

## 4.2. Safety Stock Planning Techniques

It was already mentioned that safety stock provides a buffer against the uncertainty of demand during lead-time. However, before safety stock planning techniques be presented, it would be meaningful to set some constrains which will affect the techniques. The first one is that the demand is stochastic and it is normally distributed. Although, as Zipkin (2000) states, the basic demand model is the Poisson process, which is the simplest model of random events over time in which demands occur one unit at a time, a Poisson distribution with a large mean can be approximated closely by a normal distribution. This approximation is very useful because it allows the usage of the normal distribution properties. The second one is that the lead time is constant. Although, in the practice this is not always the case, the assumption that the lead time is constant simplifies the techniques.

Another criterion that must be kept in mind when the safety stock techniques are considered is the specification of the service level. Defining service level is most important when an organization does not know its stockout costs or it is difficult to estimate them. It is common for management to set service levels from which recorder points can be ascertained.

There are several ways to measure a service level. Here will be considered two of them. A service level can be measured by either order service level (OSL) or unit service level (USL). USL, which is sometimes known as “fill rate”, counts the average number of units short expressed as the percentage of the order quantity. The OSL measures the percentage of cycles that will be out of stock or the probability of stockouts. It is crucial to define whether the service level is OSL or USL because the safety stock level would be quite different in these two measurements.[63] The techniques that will be considered will use OSL. However, the difference between these two

[62] R. Brown, “Decision Rules for Inventory Management”, 1967, p.83
[63] M. Najdawi, M. Liberatore, 2006

service levels will be discussed and it will also be shown how the considered techniques can be applied in USL case.

### 4.2.1. Constant safety stock (CSS)

This technique maintains a constant safety stock through the planning horizon. The safety stock (SS) is expressed as a product of the safety stock factor $\hat{k}$ and the standard deviation of the demand over the replenishment lead time $\sigma_L$. Therefore, can be presented as follows:

$$SS = \hat{k}\sigma_{LT}. \tag{4.2}$$

The advantage of expressing the safety stock in this way and assuming the normal distribution of the demand is that there are available tables that express the chance of stockout as a function of the safety stock factor, as it is shown on Table 1. Thus one way of setting the safety stock will be first to choose the chance of stock out that best represents company policy and then to use the safety stock factor that corresponds to that chance.

Some scientists use the mean absolute deviation (MAD) instead the standard deviation. However, this changes only the value of the safety stock factor but have no impact on the safety stock value and respectively on the desirable service level. This is so, because, as it was already mentioned in Chapter 2, MAD = 1.25 x σ.[64] This can also be seen on the Table 1.

**TABLE 1: Service Levels and Corresponding *k* Multipliers**

| *Desired Service Level (%)* | *k (MAD)* | *k (sigma)* |
|---|---|---|
| 50.00 | 0.00 | 0.00 |
| 75.00 | 0.84 | 0.67 |
| 80.00 | 1.05 | 0.84 |
| 85.00 | 1.30 | 1.04 |
| 90.00 | 1.60 | 1.28 |
| 95.00 | 2.06 | 1.65 |
| 96.00 | 2.19 | 1.75 |
| 97.00 | 2.35 | 1.88 |
| 98.00 | 2.56 | 2.05 |
| 99.00 | 2.91 | 2.33 |
| 99.50 | 3.20 | 2.57 |
| 99.90 | 3.85 | 3.09 |
| 99.99 | 5.00 | 4.00 |
| 100.00 | Infinite | Infinite |

Table 1: Service Levels and Corresponding $\hat{k}$ Multipliers, (Krupp J. (A), 1997)

[64] Silver E.A., Pyke D. F., Peterson R., Inventory Management and Production Planning and Scheduling, 1998, p. 244; R. Brown, "Decision Rules for Inventory Management", 1967, p.88; Hsu J, El-Najdawi M., 1991

The above presented CSS technique does not include the lead time. Zipkin (2000) and Brown (1967) point out that the standard deviation of the demand increases with the square root of the lead time (LT). This means that a shorter lead time will improve the performance. This leads to the classical CSS technique which considers the replenishment lead time:

$$SS = \hat{k}\sigma_{LT}\sqrt{LT} \,. \tag{4.3}$$

This technique considered the safety stock as a fixed value, which becomes applicable to all future planning periods. This raises the treat that the safety stock does not respond properly to seasonal or trend variations in the future demand and creates exposure to inadequate service. On the other hand, in decreasing conditions, the fixed safety stock can generate excess inventory, which can be dangerous particularly when a product approaches the end of its life cycle.[65]

### 4.2.2. Time Increment Contingency Factor (TICF)

This technique is proposed by Krupp (1982) to calculate safety stock. The basic idea is that safety stock should remain consistent with the time-sensitive fluctuating patterns. He adapts the classical theory by expressing statistical variance of demand in units of time through converting the deviation in a discrete time period to a decimal factor by dividing the deviation in units by the forecast for that time period. The mathematical expression yields the following equations:

$$TICF = \frac{\sum_{i=1}^{\dot{n}} \left| 1 - \frac{\dot{x}_i}{u_i} \right|}{\dot{n}} \,, \tag{4.4}$$

$$SS_{\tilde{t}} = \hat{k}(TICF_{\dot{n}} \times u_{\tilde{t}+1})\sqrt{LT} \,, \tag{4.5}$$

where

$u_i$ = average demand or forecasted for period $i$

$\dot{x}_i$ = actual demand in period $i$

$\dot{n}$ = total number of time periods being considered

[65] Krupp J.(A), 1997; Krupp J. (B), 1997

It should be noted that the expression $u_{\tilde{t}+1}$ represents the forecast period following the period $\tilde{t}$ for which the safety stock is being determined. This ensures that the safety stock that exists at the end of each future planning period $\tilde{t}$ is adequate to cover the next period's forecast. This technique provides the potential of adequate coverage of demand variability which will flex proportionately with forecast variations and will avoid an exposure to inventory excess. However, such flexing may not be desirable in periods of extreme growth or in the cases where the demand is influenced, for example, by promotions, special sales or dating terms. On the other hand, such flexing is crucial where a declining trend may exist. Krupp considers the extreme case in which the product is reaching the end of its life cycle and shows the advantage over the CSS. The fixed safety stock will remain as "dead" inventory at the end of the planning horizon, while the TICF will compensate for this declining trend and ensures that no excess inventory will remain.[66]

## 4.3. Safety Stock Based on Customer Service

As it was already mentioned, the customer service level may measure the order service level (OSL), also known as cycle service level, or unit service level (USL), also known as fill rate. Here will be discussed the criteria which the safety stock factor must satisfy in order to be applied in one of these customer service levels. The criteria and formulas, which will be presented, consider a ($\hat{t}$, $\hat{S}$) system, also known as a ($\hat{R}$, $\hat{S}$) control system. The formulas previously used to determine the safety stock in this chapter considered the lead-time. However, the duration of the lead-time in this system is the sum of the lead-time and the time of the periodic review ($\hat{t}$ or $\hat{R}$). Because of this when the above formulas are used in this system LT must equal $\hat{R}$+LT.[67]

[66] Hsu J, El-Najdawi M., 1991; Krupp J.(A), 1997; Krupp J. (B), 1997

[67] Silver E.A., Pyke D. F., Peterson R., Inventory Management and Production Planning and Scheduling, 1998, pp. 276

## 4.3.1. Safety Stock Factor in Order Service Level (OSL)

At the beginning of this chapter it was mentioned that a Poisson distribution with a large mean can be approximated closely by a normal distribution. Using the normal distribution approximation the demand quantity can be converted to the normalized form, as follows:

$$\hat{k}_i = \frac{(\dot{x}_i - u_i)}{\sigma_i}, \tag{4.6}$$

where

$u_i$ = average demand or forecasted for period $i$

$\dot{x}_i$ = actual demand in period $i$.

Then the safety factor $\hat{k}$ is selected to satisfy the following equation:

$$p_{u\geq}(\dot{k}) = 1 - P_1, \tag{4.7}$$

where

$P_1$ is the desired order service level;

$p_{u\geq}(\dot{k})$ = probability that a unit normal (mean 0, standard deviation 1) variable takes on a value of $\dot{k}$ or larger.

It should be noted that $p_{u\geq}(\dot{k})$ is often expressed as 1-$\Phi(\dot{k})$, where the $\Phi(\dot{k})$ is the cumulative distribution function of the unit normal evaluated at $\dot{k}$ and equals:

$$\dot{F}(\dot{k}) = \frac{1}{\sqrt{2\pi}} \int_k^\infty \exp - \left(\frac{y^2}{2}\right) dy. \tag{4.8}$$

Therefore the equation that has to be satisfied can be presented, as follows:

$$p_{u\geq}[(S - u_{\hat{R}+LT}) / \sigma_{\hat{R}+LT}] = 1 - P_1,$$

where

$\hat{S}$ = the order-up-to level;

$\hat{R}$ = the review interval;

$u_{\hat{R}+LT}$ = the average demand during the lead time;

$\sigma_{\hat{R}+LT}$ = the standard deviation during the lead time and the review interval.

Once the desired cycle service level is defined the techniques described previously can by applied for planning the safety stock during the replenishment time. [68]

### 4.3.2. Safety Stock Factor in Unit Service Level (USL)

Analogically to the OSL the safety stock factor in USL is chosen to satisfy the following equation:

$$G_u(\dot{k}) = \frac{DR}{\sigma_{\hat{R}+LT}}(1-P_2), \tag{4.9}$$

where

DR = the demand during the review interval;

$\sigma_{\hat{R}+LT}$ = the standard deviation during the lead time and the review interval;

$P_2$ = desired unit service level (fill-rate).

The desired unit service level, $P_2$, can be presented as:

$$P_2 = 1-(Fraction \quad backordered) = 1-\frac{ESPRC}{DR} = 1-\frac{\sigma_{\hat{R}+LT}G_u(\dot{k})}{DR} \tag{4.10}$$

where ESPRC = Expected shortage per replenishment cycle.

The $G_u(\dot{k})$ is the standard normal loss function, which equals:

$$G_u(\dot{k}) = -\dot{k}\Phi^0(\dot{k}) + \hat{\phi}(\dot{k}), \tag{4.11}$$

where

$\Phi^0(\dot{k})$ = the standard normal complementary cumulative distribution function = 1-$\Phi(\dot{k})$;

$\hat{\phi}(\dot{k})$ = the standard normal probability density function = $\frac{1}{\sqrt{2\pi}}\exp-\left(\frac{\dot{k}^2}{2}\right)$.

Once the desired fill-rate level is defined the techniques described previously can by applied for planning the safety stock during the replenishment time. However, the given $G_u(\dot{k})$ equation is for the complete backording case. In this case, it is assumed that the demand that occurs during

[68] Silver E.A., Pyke D. F., Peterson R., Inventory Management and Production Planning and Scheduling, 1998, pp. 266; R. Brown, “Decision Rules for Inventory Management”, 1967, pp.89

the stock out is backordered. This means that the customers will wait until the shipment arrives and then the demand will be satisfied. Although this is true for some products, this is not the general case. There are many products for which the customers will not wait and if a given retailer is out of stock the customer will go elsewhere to satisfy his need. This case is known as complete lost sales. In this case, the $G_u(\dot{k})$ equation has to be modified as follows:

$$G_u(\dot{k}) = \frac{DR}{\sigma_{\hat{R}+LT}} \left( \frac{1-P_2}{P_2} \right). \tag{4.12}$$

Considering the equations above, it can be intuitively expected that the required safety stock will increase if:

- the order quantity decreases, which will increase the stockouts
- the uncertainty of the forecast increases (increases of $\sigma_{\hat{R}+LT}$)
- a better fill-rate service level is desired

If any of these changes takes place, $G_u(\dot{k})$ will decrease, which implies an increase of the safety stock factor. This is exactly the desired behavior and this relationship can be seen in Figure 1.

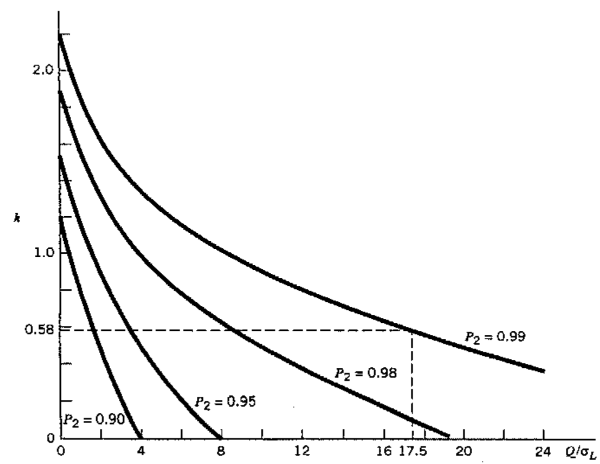


Figure 1: Sensitivity of the $P_2$ Service Measure (Silver, Pyke, Peterson, 1998)

In addition, it should be mentioned that on average $\sigma_{\hat{R}+LT}$ tends to increase with the increase of the demand. Therefore, the increase in $\hat{k}$ with increasing $\sigma_{\hat{R}+LT}$ says that, on the average in USL case, the safety stock factor will be higher for faster-moving items than for slower-moving items.[69]

## 4.4. Safety Stock Suppression

Krupp (1997) presents a method of safety stock suppression in case of overoptimistic forecast. As he mentions, in applying the statistical theory to safety stock calculations, it must be kept in mind that the basic assumption revolves around a unimodal, normal distribution of occurrences around the assumed mean of demand. Therefore, absolute values of each increment of variance are used. This approach is valid only to the extend that that the historical forecast approximates the mean average of demand. However, this is rarely the case. Therefore, a measure is needed to evaluate the degree of variance and the bias between the forecast. He suggests the simple forecast error tracking signals (FETS) technique, which is monitoring the degree and bias of cumulative forecast errors at the stock keeping unit level. The degree and bias of cumulative forecast error can be factored in by using a simple FETS. The value of FETS includes both a magnitude or an order and the bias and ranges from -1.0 to +1.0. Using traditional MAD as a bias, the equation is computed as follows:

$$FETS_n = \frac{\sum_{i=1}^{n}\left(\frac{u_i - \dot{x}_i}{\dot{n}}\right)}{MAD_n}. \tag{4.13}$$

It should be noted that the numerator of the equation does not use absolute values, but considers the netting of positive and negative values. FETS for time-based MAD applications is computed as follows:

$$FETS_n = \frac{\frac{\sum_{i=1}^{n}\left(1 - \frac{\dot{x}_i}{u_i}\right)}{\dot{n}}}{TICF_n}. \tag{4.14}$$

In this application, a FETS of zero signals the optimum circumstance. It defines a condition, where regardless of the magnitude of the individual deviations, plus and minus deviations

[69] Silver E.A., Pyke D. F., Peterson R., Inventory Management and Production Planning and Scheduling, 1998, pp. 268; P. Zipkin, Foundtion of Inventory Managenent, 2000, p. 206

ultimately compensate for each other. In this case, the rolling forecast is considered appropriate. The condition where all actual demands have been greater than forecast is the extreme case of FETS = -1.0. Conversely, a positive FETS = + 1.0 defines a condition where all actual demands have been less than forecast.

In the cases where FETS equals zero or is negative an application of safety stock as calculated through the normal algorithms would be appropriate. However, if forecasts are adjusted to compensate for the bias, safety stock levels would need to be reevaluated and adjusted to a revised forecast based on FETS signals. In the case, where FETS = +1, which represents that actual demands are consistently less than forecast, the statistical variance will be larger than in the case, where FETS = 0. It is so, because the calculation of statistical deviation in units of both quantity and time is based on the absolute values of the variance. It can be intuitively expected that the a lesser safety stock is expected in this case, as actual demands consistently fail to consume even the inventory replenishment planned based on the base forecast. However, because the safety stock is proportional to the statistical variance, the safety stock will be maximized, despite the fact that no actual demand equaled or exceeded forecast. However, in the case where the calculated FETS is less than +1.0 but greater than zero, some degree of safety stock will be still required to meet those customer demands which exceed forecast. Therefore, a totally abandoned safety stock in these cases may not be an option. One alternative is to consider calculating statistical variance only in those cases where demand in fact exceeded forecast. This approach has the disadvantage of reducing the total number of occurrences on which the algorithm is based and potentially assigning disproportionate importance to each of the reduced number of occurrences measured, where exaggerated anomalies would have an increased impact on the calculation. A second alternative is the development of a suppression factor (s) which is applied as a supplemental multiplier to the safety stock calculation, presented earlier in this capter, only in cases where FETS is greater than zero. A straight-line suppression is defined by the following equation:

$$s = 1 - FETS\,. \tag{4.15}$$

If accelerated suppression is desired, this can be derived by considering the square root of the FETS, as follows:[70]

$$s = 1 - \sqrt{FETS}\,. \tag{4.16}$$

[70] Krupp (A), 1997; Krupp (B) 1997

## 4.5. $(\hat{t}, \hat{S})$ Policy

Naddor (1971) presents a periodic review $(\hat{t}, \hat{S})$ policy with lost sales and deterministic demand where the order cycle length is predetermined and constant over time. The order quantity is chosen so that the inventory increases to $\hat{S}$. Because of the deterministic demand and predetermined order cycles the order size will vary in every order. He considers this policy in two cases: with zero lead time and with positive lead time.

### 4.5.1. $(\hat{t}, \hat{S})$ Policy with Zero Lead Time

Figure 2 demonstrates the operating characteristics of the $(\hat{t}, \hat{S})$ policy with zero lead time.

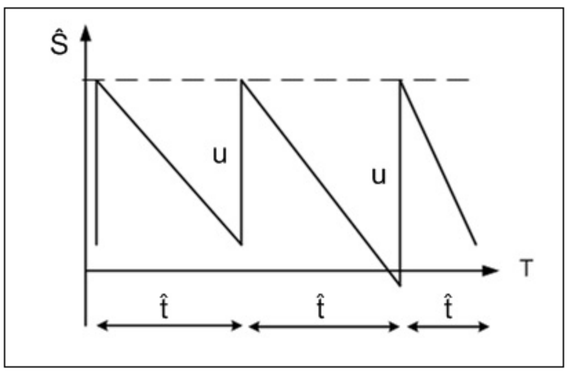


Figure 2: The $(\hat{\boldsymbol{t}}, \hat{\boldsymbol{S}})$ system with zero lead time (Naddor, E., 1971)

The inventory level $\hat{S}$ must be chosen so that no stockout occurs during an order cycle. This means that the inventory level should be equal to the maximal demand of a given order cycle. This can be presented as:

$$\hat{S} = u_{\max}(\hat{t}), \qquad (4.17)$$

where

$\hat{S}$ = order-up-to level

u = demand

$\hat{t}$ = order cycle

However, in order to minimize the costs, the case $\hat{S} > u_{\max}(\hat{t})$ should not be considered. [71]

## 4.5.2. $(\hat{t}, \hat{S})$ Policy with Positive Lead Time

Figure 3 demonstrates the operating characteristics of the $(\hat{t}, \hat{S})$ policy with positive lead time.

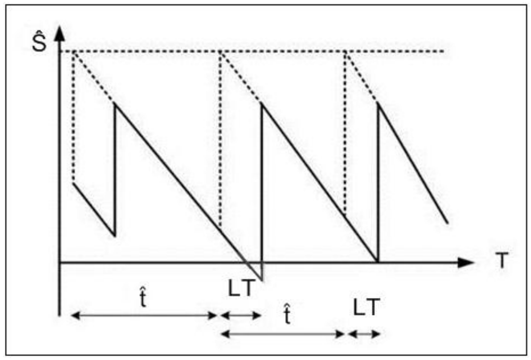


Figure 3: The $(\hat{\boldsymbol{t}}, \hat{\boldsymbol{S}})$ system with positive lead time (Naddor, E., 1971)

The inventory level $\hat{S}$ must be chosen so that no stockout occurs during an order cycle. This means that the inventory level should be equal to the maximal demand of a given order cycle. This can be presented as:

$$\hat{S} = u_{\max}(LT + \hat{t}), \tag{4.18}$$

where

LT = lead time

However, in order to minimize the costs, the case $\hat{S} > u_{\max}(LT + \hat{t})$ should not be considered.[72]

[71] E. Naddor, Inventory Systems, 1971, p. 127
[72] E. Naddor, Inventory Systems, 1971, p. 147

# 5. Homoscedasticity

In Chapter three the assumptions behind the linear regression model were described and explained. In this chapter, the assumption that the disturbances appearing in the population regression function are homoscedastic, which is critical for the linear regression model, will be examined. Here will be considered the consequences when the homoscedasticity assumption is not fulfilled, how it can be detected and the measures that can be taken when the assumption does not hold.

## 5.1. The Consequences of Heteroscedasity

One of the important assumptions of the linear regression models is that the disturbances of the population regression function are homoscedastic. This means that the variance is constant. If the variance of the disturbance indicates variation from sample to sample, the case of heteroscedasticity or nonconstant variance is observed. A better explanation about the difference between homoscedasticity and heteroscedasticity will be provided by example. Let a two-variable linear regression model be considered in which the dependent variable is personal savings and the independent variable is personal disposable income (PDI).

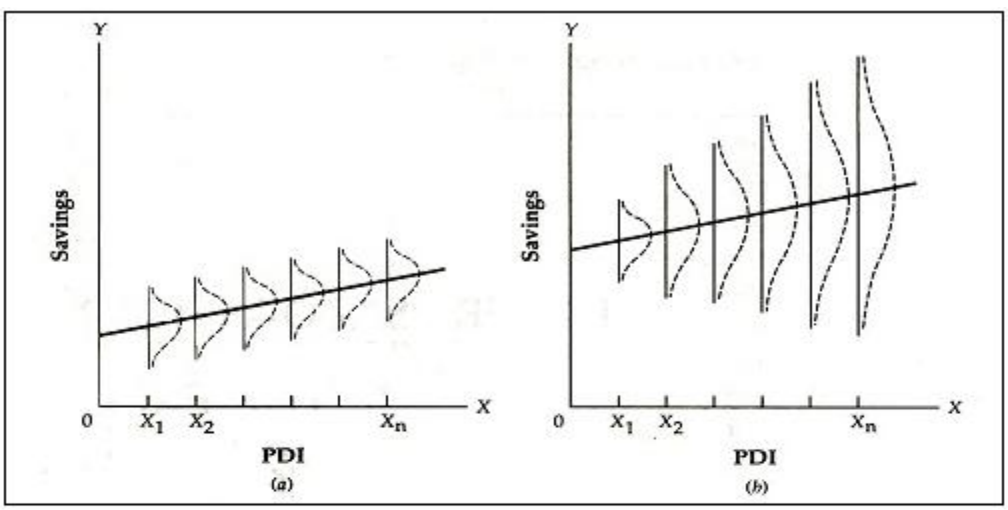


Figure 4: (a) Homoscedasticity; (b) Heteroscedasticity (Gujarati, 1992)

The diagrams of the Figure 4 show the difference between homoscedasticity and heteroscedasticity. Figure 4 (a) shows that as PDI increases, the mean level of savings also increases but the variance of savings around its mean value remains the same at all levels of PDI. On the other hand, Figure 4 (b) shows the heteroscedasticity case. Although the average level increases as the PDI increases, the variance of savings does not remain the same at all levels of PDI.

Researchers have observed that heteroscedasticity is usually found in cross-sectional data and not in time series data. In cross-sectional data one generaly deals with members of a population at a given point in time, such as individual consumers or geographical subdivisions, such as country or city. Moreover, these data may be of different size and there may be some scale effect. On the other hand, the variables in the time series data tend to be of similar order of magnitude because one generally collects data for the same entity over a period of time.[73]

However, it is controversial whether the heteroscedasticity has impact on the forecast or not. Researchers, as Fildes (1985), state that the constant variance is vital for the forecasting performance. Their counter-argument is that if, for some time periods, $e_t$ has larger variance, this has the effect of weighting those observations more heavily when minimizing $\sum \bar{e}_t^2$. This might lead to less weight being given to the most recent observations, which causes the parameter estimates to be inefficient and the standard errors of the parameters to be underestimated. This causes the forecasts to be also mis-estimated.[74]

Gujarati (1992) presents the consequences if all assumptions of the linear regression model hold, except the assumption of the homoscedasticity, as follows:

1. Least Squares (LS) estimators are still linear.
2. They are still unbiased.
3. But they no longer have minimum variance, which means that they are no longer efficient.
4. The formulas to estimate the variance of LS estimators are generally biased. A positive bias occurs if LS overestimate the true variances of estimators and a negative bias occurs if LS underestimate the true variances of estimators. However, if the homoscedasticity does not hold, one cannot tell a priori whether the bias will be positive or negative.

[73] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, pp. 317
[74] Robert Fildes, 1985

5. The bias arises from the fact that $\dot{\sigma}^2$, the conventional estimator of true $\sigma^2$, is no longer an unbiased estimator of $\sigma^2$.
6. As a result, the usual confidence intervals and hypothesis tests based on *t* and *F* distributions are unreliable. Therefore, the conclusion that in the presence of heteroscedasticity, the usual hypothesis-testing routine is not trustworthy which raising the possibility of drawing misleading conclusions, can be made.[75]

## 5.2. Detection of Heteroscedasticity

Although theoretically it is easy to document the consequences of heteroscedasticity, its detection in a concrete situation is not so easy. This is so because the variance can be known only if one has the entire population corresponding to the chosen independent variables. However, one typically has a sample of some members of this population, more specifically a single value of the dependent variable for given values of the independent variable(s), which makes very difficult the determination of the variance. There is no sure method of detecting the heteroscedasticity. However, some tests have been developed which aid the researcher in detecting it. Some of these tests, which according to Gujarati are the basic test for hoteroscedasticity detection, will be presented in the thesis. [76]

### 5.2.1. Park Test

Park (1966) suggests to transform the regression equation into the following form:

$$\ln \sigma_i^2 = \bar{a} + \bar{b} \ln x_i + \nu_i , \qquad (5.1)$$

where $\nu_i$ is a residual term.

However, the latter equation is not operational since the heteroscedastic variance is unknown. For this purpose he suggests to use $\bar{e}_i$ as proxies for $\varepsilon_i$ and running the following equation:

[75] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, pp. 324
[76] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, pp. 326

$$\ln \bar{e}_i^2 = \bar{a} + \bar{b} \ln x_i + \nu_i , \qquad (5.2)$$

where the $\bar{e}_i^2$ is taken from the original linear regression model.

Once the residuals from the original linear regression equation are obtained, squared and their logs are taken against each explanatory variable included in the regression model, the null hypothesis is tested, $\bar{b}$ = 0. If the null hypothesis holds, there is no heteroscedasticity. If a statistically significant relationship exists between $\ln \bar{e}_i^2$ and $\ln x_i$, the null hypothesis of no heteroscedasticity can be rejected. If the null hypothesis is accepted, the $b_1$ in the latter equation can be interpreted as giving the value of the homoscedastic variance.[77]

## 5.2.2. Glejser Test

Glejser (1969) presented another test which is similar to the Park test. After the residuals from the original linear regression model are obtained, he suggests to regress the absolute value of the residuals, $|\bar{e}_i|$, on the independent variable. Some functional forms that he has suggested for this regression are:

$$|\bar{e}_i| = \bar{a} + \bar{b} x_i + \nu_i , \qquad (5.3)$$

$$|\bar{e}_i| = \bar{a} + \bar{b} \sqrt{x_i} + \nu_i , \qquad (5.4)$$

$$|\bar{e}_i| = \bar{a} + \bar{b} \left( \frac{1}{x_i} \right) + \nu_i , \qquad (5.5)$$

$$|\bar{e}_i| = \sqrt{\bar{a} + \bar{b} x_i} + \nu_i , \qquad (5.6)$$

As in the Park test case, if the null hypothesis holds there is no heteresceddasticity. If the null hypothesis is rejected, there is probably evidence of heteroscedasticity.

However, Goldfred and Quandt (1972) point out that the error term $\nu_i$ can itself be heteroscedastic as well as serial correlated. Moreover, some of the equations suggested by Glejser

[77] R. E. Park, 1966; Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, pp. 330

such as equation (5.6) are nonlinear in the parameters and therefore cannot be estimated with the usual LS procedure. [78]

Glejser, however, has maintained that in large samples the preceding models are fairly good in detecting heteroscedasticity. Therefore, this test may be used for large samples as practical matter and may be used in small samples strictly as a qualitative device with the purpose to learn something about the heteroscedasticity.[79]

## 5.2.3. Goldfeld-Quandt test

This method is applicable if it is assumed that the heteroscedastic variance is positively to one of the explanatory variables in the regression model. For simplicity, as in the previous tests, the two-variable model is considered. Following the assumption, the heteroscedastic variance is positively related to the explanatory variable, as follows:

$$\sigma_i^2 = \sigma^2 x_i^2, \qquad (5.7)$$

where $\sigma^2$ is a constant.

This assumption postulates that the heteroscedastic variance, $\sigma_i^2$, is proportional to the square of the independent variable. From this follows that the heteroscedastic variance increases with the increase of the explanatory variable's value. If that is the case, heteroscedasticity is most likely to be present.[80]

## 5.2.4. Remedial Measures

As it was already mentioned, in the presence of heteroscedasticity the LS estimators are no longer efficient. Therefore, it is important redial measures to be taken. Different variance stabilizing

[78] Stephen M. Goldfeld and Richard E. Quandt 1972; Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, pp. 331; Gujarati, Damodar, 2009 refer to Gujarati, Damodar, Basic Econometrics, 1988, pp. 330

[79] H. Glejser, 1969; Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, pp. 331; Gujarati, Damodar, 2009 refer to Gujarati, Damodar, Basic Econometrics, 1988, pp. 330

[80] Stephen M. Goldfeld and Richard E. Quandt 1972; Gujarati, Damodar, 2009 refer to Gujarati, Damodar, Basic Econometrics, 1988, p. 333

models have been developed with the aim to transform the original regression model in such a way so that in the transformed model the heteroscedasticity to be removed.[81]

5.2.4.1. The Method of Weighted Least Squares

This method can be applied if the true error variance is known. Assuming that the error variance for each observation is known the original regression model can be transformed, as follows:

$$\frac{y_i}{\sigma_i} = \bar{a}\left(\frac{1}{\sigma_i}\right) + \bar{b}\left(\frac{x_i}{\sigma_i}\right) + \frac{\varepsilon_i}{\sigma_i}, \tag{5.8}$$

where $\sigma_i$ is the square root of the "known" variance.

The transformed error term can be denoted with $v_i$ and it is equal to:

$$v_i = \frac{\varepsilon_i}{\sigma_i}. \tag{5.9}$$

It can be easily shown that the transformed error term is homoscedastic. In the presence of heteroscedasticity the disturbance is no longer constant it can be presented for each observation as:

$$E\left(\varepsilon_i^2\right) = \sigma_i^2. \tag{5.10}$$

Squaring equation (4.9) and implying equation (4.10), it can be proofed that the error term $v_i$ is homoscedastic, as follows:

$$E\left(v_i^2\right) = E\left(\frac{\varepsilon_i^2}{\sigma_i^2}\right) = \left(\frac{1}{\sigma_i^2}\right)E\left(\varepsilon_i^2\right) = \left(\frac{1}{\sigma_i^2}\right)\sigma_i^2 = 1. \tag{5.11}$$

This proofs that the transformed model (5.8) does not suffer from the heteroscedasticity problem and therefore it can be estimated by the usual LS method. The LS estimators thus obtained are called weighted least squares estimators, because in each observation the dependent and independent variables are weighted by its own heteroscedastic standard deviation. Because of this weighting procedure, the LS method in this context is known as the method of weighted least squares.[82]

---

[81] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, pp. 333

[82] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, p. 333

5.2.4.2. Remedial Measures When True Variance Is Unknown

As noted earlier, the true variance is rarely known, therefore in the absence of knowledge about the true variance additional assumptions about the unknown variance can be made and will be considered by two cases, which for simplicity, will be discussed with the simple regression model.

5.2.4.2.1. Case 1: The error variance is proportional to $\hat{X}_i$. The square root transformation

In this method, after estimating the usual LS regression, the residuals from this regression are plotted against the explanatory variable. If the observe pattern is similar to that shown in Figure 5, the indication is that the error variance is linearly related to the explanatory variable.

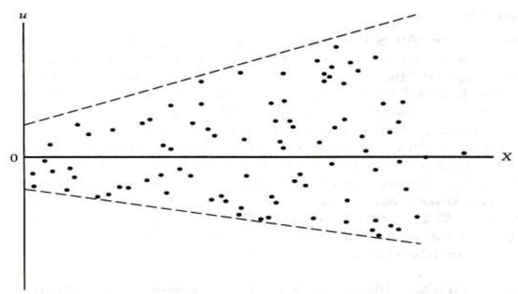


Figure 5: Error variance proportional to $\hat{\boldsymbol{X}}$ (Gujarati, 1992)

The mathematical expression of the statement that the heteroscedastic variance is proportional to the explanatory variable is expressed with the following formula:

$$E\left(\varepsilon_i^2\right)=\tilde{\sigma}^2 x_i \,, \qquad (5.12)$$

where $\tilde{\sigma}^2$ is the constant factor of proportionality.

Using equation (4.12), the simple regression model can be transformed as follows:

$$\frac{y_i}{\sqrt{x_i}}=\bar{a}\frac{1}{\sqrt{x_i}}+\bar{b}\frac{x_i}{\sqrt{x_i}}+\frac{\varepsilon_i}{\sqrt{x_i}}=\bar{a}\frac{1}{\sqrt{x_i}}+\bar{b}\sqrt{x_i}+\nu_i \,, \qquad (5.13)$$

where $\frac{\varepsilon_i}{\sqrt{x_i}}=\nu_i$ .

This method is known as the square root transformation because both sides of the regression model are divided by the square root of $\hat{X}_i$. It is important to be noted that in order to estimate the latter equation, the regression-through-the-origin estimating procedure must be used. In the multiple regression case, the model is divided by the square root of $\hat{X}_i$, which is chosen on the basis of graphical plot. If the appropriate candidates are more than one, the mean value of the dependent variable, $\bar{y}_i$, can be used as the transforming variable. This can be made because the mean value of the dependent variable is a linear combination of the independent variables.

The proof that the transformed model does not suffer from the heteroscedasticity problem and therefore it can be estimated by the usual LS method is:

$$E\left(v_i^2\right) = E\left(\frac{\varepsilon_i^2}{x_i}\right) = \frac{\sigma^2 x_i}{x_i} = \sigma^2 . \quad (5.14)$$

[83]

5.2.4.2.2. Case 2: The error variance is proportional to $\hat{X}_i^2$

After the usual LS regression is estimated, the residuals from this regression are plotted against the explanatory variable. If the observe pattern is similar to that shown in Figure 6, the indication is that the error variance is not linearly related to the explanatory variable, but increases proportional to the square of the explanatory variable.

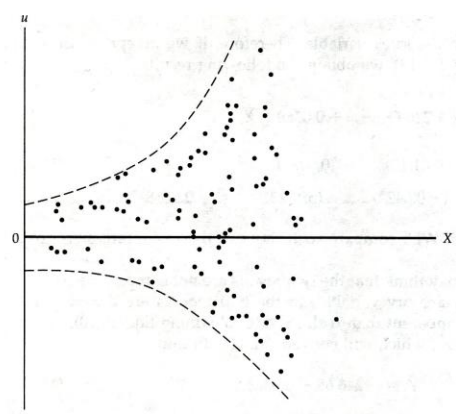


Figure 6: Error variance proportional to $\hat{X}^2$ (Gujarati, 1992)

[83] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, p. 335

The mathematical expression of the statement is expressed with the following formula:

$$E\left(\varepsilon_i^2\right) = \sigma^2 x_i^2 . \tag{5.15}$$

In this case the appropriate transformation of the simple regression model is to divide both sides of the model by explanatory variable, rather than by its square root, as follows:

$$\frac{y_i}{x_i} = \bar{a}\frac{1}{x_i} + \bar{b}\frac{x_i}{x_i} + \frac{\varepsilon_i}{x_i} = \bar{a}\frac{1}{x_i} + \bar{b} + \nu_i , \tag{5.16}$$

where $\frac{\varepsilon_i}{x_i} = \nu_i$.

This equation has an interesting feature, namely, the original slope coefficient becomes the intercept and the original intercept becomes the slope coefficient. However, this change is only for estimation. After the estimation, multiplying by the explanatory variable on both sides, the original model is obtained again.[84]

The proof that the transformed model does not suffer from the heteroscedasticity problem and therefore it can be estimated by the usual LS method is:

$$E\left(\nu_i^2\right) = E\left(\frac{\varepsilon_i^2}{x_i^2}\right) = \frac{\sigma^2 x_i^2}{x_i^2} = \sigma^2 . \tag{5.17}$$

5.2.4.3. Respecification of the Model

If heteroscedasticity is detected, it can be reduced by choosing a different functional form of the population regression function. One possible solution that reduces the heteroscedasticity is to estimate the model in the log form, as follows:

$$\ln y_i = \bar{a} + \bar{b}\ln x_i + \varepsilon_i . \tag{5.18}$$

In this transformation, the heteroscedasticity problem may be less serious because the log transformation compresses the scales in which the variables are measured. In this way, a tenfold difference between two values is reduced to a twofold difference.[85]

[84] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, p. 337

[85] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, p. 339

# 6. Autocorrelation

Another assumption of the linear regression model is that there is no autocorrelation among the disturbances entering the population regression function. This means that the disturbance term relating to any observation is not related to or influenced by the disturbance term relating to any other observation. Said in other words, the assumption states that the expected value of the product of two different error terms is zero.
However, sometimes autocorrelation patterns occur because some important variables that should be included in the model are not included or because used data is smoothed by transforming the raw daily data into weekly, monthly or quarterly data.[103]

## 6.1. The Consequences of Autocorrelation

If the assumption of no autocorrelation is violated the consequences are:

1. The least square estimators are still linear and unbiased but they do not have the minimum variance compared to the producers that take into account autocorrelation.
2. Therefore, the computed variances and standard errors seriously underestimate the true variances and standard errors. This makes the t-test and the F-test not generally reliable.
3. The conventionally computed $R^2$ may be an unreliable measure of the true $R^2$.
4. The conventionally computed variances and standard errors of forecast may also be inefficient.[104]

## 6.2. Detection of Autocorrelation

The detection of the autocorrelation is as difficult as for the heteroscedasticity because the true error variance is unknown. Therefore, we can rely on the proxies obtained from the ordinary least squares to make conclusions about the presence or lack of autocorrelation.[105]

[103] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, pp. 353
[104] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, p. 356
[105] Gujarati, Damodar, 2005 refer to Gujarati, Damodar, Essentials of Econometrics, 1992, p. 356

Although, there is no unique method for detecting the autocorrelation, there are several tests with which help its presence can be proved. However, the deeper analysis of the autocorrelation process which can be applied for the time series and causal forecasting models can be done more appropriate with the time series analysis. This analysis will be presented in the thesis.

### 6.2.1. Time Series Analysis

The key statistics in time-series analysis are the autocorrelation coefficient and the partial autocorrelation coefficient.

The autocorrelation coefficient is the correlation of the time series with itself, lagged by 0, 1, 2, or more periods. The autocorrelation coefficient can be obtained using the following equation:

$$\rho_{Y_t Y_{t-1}} = \frac{\sum_{t=2}^{n}(Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}, \tag{6.1}$$

where $\rho$ is the autocorrelation coefficient.

The autocorrelation of random data have a sampling distribution that can be approximated by normal curve with mean zero and standard error $1 \,/\, \sqrt{\bar{n}}$. This information can be used to develop test of hypothesis similar to F-test and t-tests examined in Chapter 3. These can be used to determine whether some $r_\tau$ comes from a population whose value is zero at $\tau$ time lag. This means that all autocorrelation coefficients must lie within the 95 percent upper and lower confidence limits in order to be confirmed that the data are random.

The partial autocorrelation coefficients are used to measure the degree of association between $X_t$ and $X_{t-n}$ when the effects of other of other time lags – 1, 2, 3, ….up to $\tau$-1 – are partialled out. The partial autocorrelation coefficient of order *m* is defined as the last autoregressive coefficient of an AR ($\bar{m}$) model. The mathematical expression is:

$$X_t = \widetilde{\Phi}_1 X_{t-1} + \widetilde{\Phi}_2 X_{t-2} + \cdots + \widetilde{\Phi}_{\bar{m}-1} X_{t-\bar{m}+1} + \widetilde{\Phi}_{\bar{m}} X_{t-\bar{m}} + e_t, \tag{6.2}$$

where $\widetilde{\Phi}$ is the partial autocorrelation coefficient.

The partial autocorrelation coefficients are used to identify the autocorrelation order. If the process is autoregressive (AR) they will be significantly different from zero only up to the order of the true AR process. If the generating process is moving average (MA) rather than autoregressive, then the partial autocorrelations will decline to zero exponentially.[106]

### 6.2.2. Remedial Measure for Autocorrelated Demand

As the autocorrelation has impact on the variance and therefore the safety stock planning can be under or overestimated, the variance needs to be corrected depending on the grade of the autocorrelation process.

A general optimization approach is proposed by Inderfurth (1995). Assuming a general *m*-order type of autocorrelation the approach has the following mathematic expression:

$$\sigma_n^2(T) = \sigma_n^2 * \left[\dot{\gamma}(T) * T + \dot{\delta}(T)\right], \tag{6.3}$$

where

$\dot{\gamma}(T) = 1 + 2\sum_{\tau=1}^{g(T)} \rho_\tau$ and $\dot{\delta}(T) = -2\sum_{\tau=1}^{g(T)} \tau * \rho_\tau$

with $g(T) = \begin{cases} T - 1 & for \quad T \leq \bar{m} \\ \bar{m} & for \quad T > \bar{m} \end{cases}$ ,

where $\rho_\tau$ is the autocorrelation coefficient at lag $\tau$.[107]

[106] Makridakis, Wheelwright, McGee, Forecasting, Methods and Applications, 1983, pp.364

[107] K .Inderfurth, 1995

# 7. Comparison of the Forecasting Models

In this chapter, the forecast accuracy of the time series forecasting methods will be compared with the causal forecasting methods. Furthermore, the impact of the forecast accuracy on the safety stock inventory planning will be investigated. For these purposes, we will use the sales data for “Schwechater” beer canes sold in Austria for 2005, 2006 and 2007 by ADEG Austria Ltd. (see Table 2)

| Schwechater Beer | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 Unit - 24 Cannes 500 ml | | | | | | | | | |
| | Year 2005 | | | Year 2006 | | | Year 2007 | | |
| Week | Beer | Price in EUR | Temperature in C° | Beer | Price in I | Temperature in C° | Beer | Price in EUR | Temperature in C° |
| 1 | 1314 | 0,59 | 6,7 | 1412 | 0,62 | 1,2 | 844 | 0,65 | 8,2 |
| 2 | 1125 | 0,59 | 7,8 | 1138 | 0,62 | -3,4 | 1051 | 0,65 | 11,6 |
| 3 | 1517 | 0,59 | 4,1 | 1469 | 0,62 | 1,3 | 2256 | 0,65 | 12,9 |
| 4 | 1356 | 0,59 | -2,0 | 6702 | 0,42 | -1,4 | 5176 | 0,45 | 1,8 |
| 5 | 6456 | 0,39 | 0,1 | 4525 | 0,42 | -1,3 | 2907 | 0,45 | 7,0 |
| 6 | 5687 | 0,39 | 2,8 | 1301 | 0,62 | 0,6 | 1099 | 0,65 | 9,9 |
| 7 | 1312 | 0,59 | 1,6 | 1469 | 0,62 | 4,2 | 1243 | 0,65 | 10,4 |
| 8 | 1458 | 0,59 | 0,8 | 1280 | 0,62 | 5,3 | 2658 | 0,65 | 12,8 |
| 9 | 1633 | 0,59 | -1,2 | 5530 | 0,42 | 2,8 | 2608 | 0,65 | 10,3 |
| 10 | 1412 | 0,59 | 1,0 | 4657 | 0,42 | 3,2 | 7319 | 0,45 | 11,9 |
| 11 | 7456 | 0,39 | 12,6 | 1114 | 0,62 | 1,4 | 5701 | 0,45 | 15,6 |
| 12 | 6300 | 0,39 | 13,9 | 1625 | 0,62 | 10,6 | 1457 | 0,65 | 7,2 |
| 13 | 1214 | 0,59 | 13,9 | 1473 | 0,62 | 15,2 | 1158 | 0,65 | 15,1 |
| 14 | 1577 | 0,59 | 18,0 | 1864 | 0,62 | 11,8 | 2767 | 0,65 | 15,0 |
| 15 | 1657 | 0,59 | 12,6 | 2136 | 0,62 | 11,8 | 2959 | 0,65 | 21,9 |
| 16 | 1798 | 0,59 | 11,6 | 1870 | 0,62 | 18,2 | 11504 | 0,45 | 20,2 |
| 17 | 1888 | 0,59 | 15,5 | 1799 | 0,62 | 17,6 | 4291 | 0,45 | 23,9 |
| 18 | 7345 | 0,39 | 20,3 | 1425 | 0,62 | 19,4 | 1102 | 0,65 | 19,0 |
| 19 | 6812 | 0,39 | 13,2 | 7423 | 0,42 | 19,9 | 1048 | 0,65 | 21,9 |
| 20 | 1344 | 0,59 | 18,2 | 6655 | 0,42 | 21,4 | 1259 | 0,65 | 22,0 |
| 21 | 1566 | 0,59 | 24,4 | 1212 | 0,62 | 19,6 | 6361 | 0,45 | 27,7 |
| 22 | 1723 | 0,59 | 25,3 | 1845 | 0,62 | 12,3 | 6193 | 0,45 | 18,6 |
| 23 | 1945 | 0,59 | 16,5 | 1364 | 0,62 | 14,8 | 2032 | 0,65 | 25,4 |
| 24 | 9876 | 0,39 | 22,2 | 1798 | 0,62 | 27,6 | 824 | 0,65 | 25,4 |
| 25 | 7932 | 0,39 | 27,0 | 10037 | 0,42 | 28,5 | 1227 | 0,65 | 25,5 |
| 26 | 1612 | 0,59 | 26,4 | 8563 | 0,42 | 25,7 | 1673 | 0,65 | 23,1 |
| 27 | 1956 | 0,59 | 22,1 | 2036 | 0,62 | 26,2 | 1937 | 0,65 | 21,8 |
| 28 | 2156 | 0,59 | 23,2 | 2455 | 0,62 | 28,1 | 2243 | 0,65 | 23,6 |
| 29 | 2430 | 0,59 | 25,3 | 2712 | 0,62 | 29,6 | 5338 | 0,45 | 30,7 |
| 30 | 2204 | 0,59 | 27,7 | 1847 | 0,62 | 29,9 | 8591 | 0,45 | 26,4 |
| 31 | 1974 | 0,59 | 23,3 | 9625 | 0,42 | 22,7 | 3896 | 0,65 | 22,2 |
| 32 | 9467 | 0,39 | 19,9 | 8533 | 0,42 | 20,4 | 1188 | 0,65 | 26,7 |
| 33 | 8103 | 0,39 | 20,4 | 2685 | 0,62 | 23,9 | 1204 | 0,65 | 22,7 |
| 34 | 1467 | 0,59 | 20,3 | 2422 | 0,62 | 21,8 | 1350 | 0,65 | 23,7 |
| 35 | 1921 | 0,59 | 24,8 | 2512 | 0,62 | 19,2 | 4273 | 0,45 | 20,4 |
| 36 | 2133 | 0,59 | 25,8 | 1864 | 0,62 | 25,4 | 8932 | 0,45 | 14,1 |
| 37 | 1742 | 0,59 | 21,8 | 1956 | 0,62 | 24,7 | 3004 | 0,65 | 19,4 |
| 38 | 1542 | 0,59 | 16,6 | 2135 | 0,62 | 21,1 | 956 | 0,65 | 19,9 |
| 39 | 8463 | 0,39 | 18,9 | 2145 | 0,62 | 21,7 | 13380 | 0,45 | 16,3 |
| 40 | 7453 | 0,39 | 16,6 | 8465 | 0,42 | 18,9 | 13688 | 0,45 | 19,2 |
| 41 | 1455 | 0,59 | 18,4 | 7388 | 0,42 | 18,9 | 1261 | 0,65 | 14,6 |
| 42 | 1688 | 0,59 | 13,9 | 1586 | 0,62 | 16,6 | 568 | 0,65 | 12,2 |
| 43 | 1863 | 0,59 | 18,9 | 6432 | 0,42 | 20,6 | 2050 | 0,65 | 7,1 |
| 44 | 1912 | 0,59 | 13,9 | 5549 | 0,42 | 11,5 | 4380 | 0,65 | 11,0 |
| 45 | 1732 | 0,59 | 11,5 | 1659 | 0,62 | 11,9 | 3803 | 0,65 | 8,5 |
| 46 | 1548 | 0,59 | 7,5 | 1736 | 0,62 | 12,7 | 1831 | 0,65 | 1,6 |
| 47 | 5865 | 0,39 | 0,5 | 1487 | 0,62 | 11,1 | 2282 | 0,65 | 4,5 |
| 48 | 4972 | 0,39 | 1,2 | 5348 | 0,42 | 7,4 | 5956 | 0,45 | 5,3 |
| 49 | 1532 | 0,59 | 3,7 | 4650 | 0,42 | 11,8 | 4234 | 0,45 | 7,7 |
| 50 | 1412 | 0,59 | 2,4 | 1325 | 0,62 | 7,6 | 178 | 0,65 | 1,9 |
| 51 | 2477 | 0,59 | 1,5 | 2827 | 0,62 | 3,8 | 4239 | 0,65 | -1,9 |
| 52 | 1533 | 0,59 | -0,3 | 1833 | 0,62 | 2,6 | 329 | 0,65 | 0,0 |

Table 2: Beer Sales Data[108]

[108] Source: ADEG Österreich Handels AG

For the times series forecast methods we will take as basis the first 126 weeks from the data and then using the Mean method and the Winter's three-parameter trend and seasonality method we will compare the forecast results for the next 30 weeks with a forecast period for 1, 3, 5, 10,15 and 30 weeks.

The causal forecast methods will be compared in an identical way – the first 126 weeks of the data will be taken as basis and then the results from the simple regression model will be compared with the results from the multiple regression model. The variable for the simple regression will be the price of a bier cane and for the multiple regression the price and the temperature will be taken as variables. The temperature data are the maximal daily air temperature converted on a mean weekly basis for the city of Salzburg.[109] We decided to take the temperature of this region, because the company's headquarter is in Bergheim, nearby Salzburg, and most of the company's branches are in the western part of Austria. That is why, it could be assumed that the temperature of Salzburg should be a good variable for the multiple regression method.

The comparison of the forecasting methods is based on the measurements described in Chapter 2.4. Once the forecast methods are compared, the impact of the forecast accuracy on the safety stock planning will be investigated. For this purpose the most accurate time series and causal methods will be compared in $(\hat{t}, \hat{S})$ inventory system with 3 and 5 weeks review periods and with 0, 3 and 5 weeks positive lead time with regards on the order service level and unit service level.

Before starting with the forecast comparison, we need to analyze the data. Figure 7 shows the sales variations on weekly basis.

[109] Source: http://www.zamg.ac.at/klima/jahrbuch/?ts=1271762341

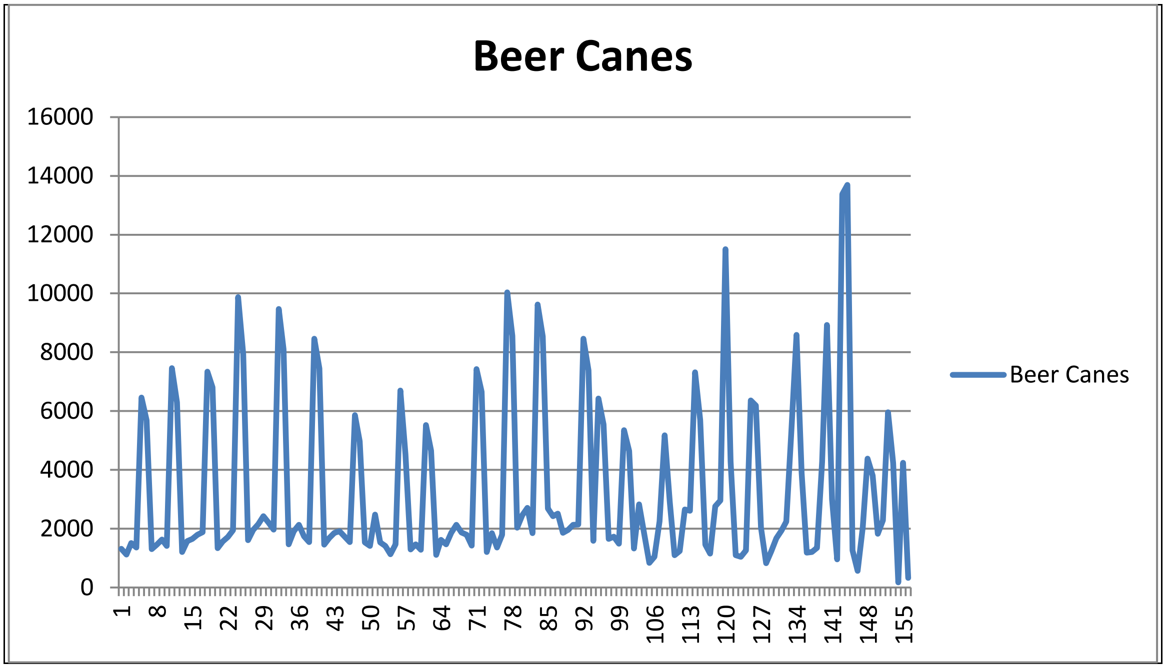


Figure 7: Graphic of the sold beer canes

As it can be seen from the graphic, the sales data has a weekly seasonality which is obviously driven by the discounts that the company has done in these weeks and additionally a yearly based seasonality. This can be clearly seen for the first 2 years (104 weeks). It can be also seen that for the third year, 30 weeks of which will be forecasted, the sales do not have the typical behavior of the previous two years – the yearly seasonality cannot be seen clearly. It can be also seen that the peaks are above the peaks of the previous two years and the lows are below the previous lows. A possible explanation for the change of the seasonality can be the changed customer behavior during the time of crises.

## 7.1. Time Series Forecast Methods

### 7.1.1. The Mean Method

Using Equations 2.1, we will take the average of the first 126 weeks of the sales data and the forecast for the coming week(s) is the average of these 126 weeks. As the forecast goes ahead with T= 1, 3, 5, 10, 15 and 30 weeks, the data of the past weeks will also be included in the past data, which are used for the forecast. Applying these rules, we receive:

For T=1:

| Week | Actual | Forecast | Error | Abs. Error | MSE | MAD | % Error | MAPE | TS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2032 | 3287 | 1255 | 1255 | 1575264 | 1255 | 62 | 62 | 1,00 |
| 2 | 824 | 3277 | 2453 | 2453 | 3796758 | 1854 | 298 | 180 | 2,00 |
| 3 | 1227 | 3258 | 2031 | 2031 | 3906223 | 1913 | 166 | 175 | 3,00 |
| 4 | 1673 | 3242 | 1569 | 1569 | 3545344 | 1827 | 94 | 155 | 4,00 |
| 5 | 1937 | 3230 | 1293 | 1293 | 3170765 | 1720 | 67 | 137 | 5,00 |
| 6 | 2243 | 3220 | 977 | 977 | 2801509 | 1597 | 44 | 122 | 6,00 |
| 7 | 5338 | 3213 | -2125 | 2125 | 3046410 | 1672 | 40 | 110 | 4,46 |
| 8 | 8591 | 3229 | -5362 | 5362 | 6259580 | 2133 | 62 | 104 | 0,98 |
| 9 | 3896 | 3269 | -627 | 627 | 5607760 | 1966 | 16 | 94 | 0,75 |
| 10 | 1188 | 3274 | 2086 | 2086 | 5481953 | 1978 | 176 | 102 | 1,80 |
| 11 | 1204 | 3258 | 2054 | 2054 | 5367228 | 1985 | 171 | 109 | 2,82 |
| 12 | 1350 | 3243 | 1893 | 1893 | 5218662 | 1977 | 140 | 111 | 3,79 |
| 13 | 4273 | 3230 | -1043 | 1043 | 4900981 | 1905 | 24 | 104 | 3,39 |
| 14 | 8932 | 3237 | -5695 | 5695 | 6867514 | 2176 | 64 | 102 | 0,35 |
| 15 | 3004 | 3278 | 274 | 274 | 6414675 | 2049 | 9 | 95 | 0,50 |
| 16 | 956 | 3276 | 2320 | 2320 | 6350096 | 2066 | 243 | 105 | 1,62 |
| 17 | 13380 | 3259 | -10121 | 10121 | 12001592 | 2540 | 76 | 103 | -2,66 |
| 18 | 13688 | 3330 | -10358 | 10358 | 17295033 | 2974 | 76 | 101 | -5,76 |
| 19 | 1261 | 3402 | 2141 | 2141 | 16626059 | 2930 | 170 | 105 | -5,11 |
| 20 | 568 | 3387 | 2819 | 2819 | 16192203 | 2925 | 496 | 125 | -4,16 |
| 21 | 2050 | 3368 | 1318 | 1318 | 15503875 | 2848 | 64 | 122 | -3,81 |
| 22 | 4380 | 3359 | -1021 | 1021 | 14846527 | 2765 | 23 | 117 | -4,29 |
| 23 | 3803 | 3366 | -437 | 437 | 14209329 | 2664 | 11 | 113 | -4,62 |
| 24 | 1831 | 3369 | 1538 | 1538 | 13715826 | 2617 | 84 | 111 | -4,11 |
| 25 | 2282 | 3359 | 1077 | 1077 | 13213563 | 2556 | 47 | 109 | -3,79 |
| 26 | 5956 | 3352 | -2604 | 2604 | 12966239 | 2557 | 44 | 106 | -4,81 |
| 27 | 4234 | 3369 | -865 | 865 | 12513739 | 2495 | 20 | 103 | -5,27 |
| 28 | 178 | 3374 | 3196 | 3196 | 12431700 | 2520 | 1796 | 164 | -3,95 |
| 29 | 4239 | 3354 | -885 | 885 | 12030053 | 2463 | 21 | 159 | -4,40 |
| 30 | 329 | 3359 | 3030 | 3030 | 11935143 | 2482 | 921 | 184 | -3,15 |

Table 3: Mean Method Forecast with T=1

For T=3:

| Week | Actual | Forecast | Error | Abs. Error | MSE | MAD | % Error | MAPE | TS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2032 | 3287 | 1255 | 1255 | 1575264 | 1255 | 62 | 62 | 1,00 |
| 2 | 824 | 3287 | 2463 | 2463 | 3821051 | 1859 | 299 | 180 | 2,00 |
| 3 | 1227 | 3287 | 2060 | 2060 | 3962032 | 1926 | 168 | 176 | 3,00 |
| 4 | 1673 | 3242 | 1569 | 1569 | 3587201 | 1837 | 94 | 156 | 4,00 |
| 5 | 1937 | 3242 | 1305 | 1305 | 3210524 | 1731 | 67 | 138 | 5,00 |
| 6 | 2243 | 3242 | 999 | 999 | 2841871 | 1609 | 45 | 122 | 6,00 |
| 7 | 5338 | 3213 | -2125 | 2125 | 3081006 | 1682 | 40 | 111 | 4,47 |
| 8 | 8591 | 3213 | -5378 | 5378 | 6311302 | 2144 | 63 | 105 | 1,00 |
| 9 | 3896 | 3213 | -683 | 683 | 5661885 | 1982 | 18 | 95 | 0,74 |
| 10 | 1188 | 3274 | 2086 | 2086 | 5530666 | 1992 | 176 | 103 | 1,78 |
| 11 | 1204 | 3274 | 2070 | 2070 | 5417261 | 1999 | 172 | 109 | 2,81 |
| 12 | 1350 | 3274 | 1924 | 1924 | 5274174 | 1993 | 142 | 112 | 3,79 |
| 13 | 4273 | 3230 | -1043 | 1043 | 4952222 | 1920 | 24 | 105 | 3,39 |
| 14 | 8932 | 3230 | -5702 | 5702 | 6921207 | 2190 | 64 | 102 | 0,36 |
| 15 | 3004 | 3230 | 226 | 226 | 6463184 | 2059 | 8 | 96 | 0,50 |
| 16 | 956 | 3276 | 2320 | 2320 | 6395574 | 2076 | 243 | 105 | 1,61 |
| 17 | 13380 | 3276 | -10104 | 10104 | 12024958 | 2548 | 76 | 103 | -2,65 |
| 18 | 13688 | 3276 | -10412 | 10412 | 17379915 | 2985 | 76 | 102 | -5,75 |
| 19 | 1261 | 3402 | 2141 | 2141 | 16706474 | 2940 | 170 | 105 | -5,11 |
| 20 | 568 | 3402 | 2834 | 2834 | 16272771 | 2935 | 499 | 125 | -4,16 |
| 21 | 2050 | 3402 | 1352 | 1352 | 15584940 | 2860 | 66 | 122 | -3,79 |
| 22 | 4380 | 3359 | -1021 | 1021 | 14923907 | 2776 | 23 | 118 | -4,27 |
| 23 | 3803 | 3359 | -444 | 444 | 14283609 | 2675 | 12 | 113 | -4,60 |
| 24 | 1831 | 3359 | 1528 | 1528 | 13785755 | 2627 | 83 | 112 | -4,10 |
| 25 | 2282 | 3359 | 1077 | 1077 | 13280695 | 2565 | 47 | 109 | -3,78 |
| 26 | 5956 | 3359 | -2597 | 2597 | 13029362 | 2566 | 44 | 107 | -4,79 |
| 27 | 4234 | 3359 | -875 | 875 | 12575170 | 2503 | 21 | 104 | -5,26 |
| 28 | 178 | 3374 | 3196 | 3196 | 12490936 | 2528 | 1796 | 164 | -3,95 |
| 29 | 4239 | 3374 | -865 | 865 | 12085994 | 2471 | 20 | 159 | -4,39 |
| 30 | 329 | 3374 | 3045 | 3045 | 11992266 | 2490 | 926 | 185 | -3,13 |

Table 4: Mean Method Forecast with T=3

For T=5:

| Week | Actual | Forecast | Error | Abs. Error | MSE | MAD | % Error | MAPE | TS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2032 | 3287 | 1255 | 1255 | 1575264 | 1255 | 62 | 62 | 1,00 |
| 2 | 824 | 3287 | 2463 | 2463 | 3821051 | 1859 | 299 | 180 | 2,00 |
| 3 | 1227 | 3287 | 2060 | 2060 | 3962032 | 1926 | 168 | 176 | 3,00 |
| 4 | 1673 | 3287 | 1614 | 1614 | 3622850 | 1848 | 96 | 156 | 4,00 |
| 5 | 1937 | 3287 | 1350 | 1350 | 3262831 | 1748 | 70 | 139 | 5,00 |
| 6 | 2243 | 3220 | 977 | 977 | 2878231 | 1620 | 44 | 123 | 6,00 |
| 7 | 5338 | 3220 | -2118 | 2118 | 3107684 | 1691 | 40 | 111 | 4,50 |
| 8 | 8591 | 3220 | -5371 | 5371 | 6324697 | 2151 | 63 | 105 | 1,04 |
| 9 | 3896 | 3220 | -676 | 676 | 5672674 | 1987 | 17 | 95 | 0,78 |
| 10 | 1188 | 3220 | 2032 | 2032 | 5518455 | 1992 | 171 | 103 | 1,80 |
| 11 | 1204 | 3258 | 2054 | 2054 | 5400411 | 1997 | 171 | 109 | 2,83 |
| 12 | 1350 | 3258 | 1908 | 1908 | 5253831 | 1990 | 141 | 112 | 3,79 |
| 13 | 4273 | 3258 | -1015 | 1015 | 4928898 | 1915 | 24 | 105 | 3,41 |
| 14 | 8932 | 3258 | -5674 | 5674 | 6876216 | 2183 | 64 | 102 | 0,39 |
| 15 | 3004 | 3258 | 254 | 254 | 6422112 | 2055 | 8 | 96 | 0,54 |
| 16 | 956 | 3276 | 2320 | 2320 | 6357068 | 2071 | 243 | 105 | 1,66 |
| 17 | 13380 | 3276 | -10104 | 10104 | 11988718 | 2544 | 76 | 103 | -2,62 |
| 18 | 13688 | 3276 | -10412 | 10412 | 17345688 | 2981 | 76 | 102 | -5,73 |
| 19 | 1261 | 3276 | 2015 | 2015 | 16646408 | 2930 | 160 | 105 | -5,14 |
| 20 | 568 | 3276 | 2708 | 2708 | 16180693 | 2919 | 477 | 123 | -4,23 |
| 21 | 2050 | 3368 | 1318 | 1318 | 15492913 | 2843 | 64 | 121 | -3,88 |
| 22 | 4380 | 3368 | -1012 | 1012 | 14835235 | 2760 | 23 | 116 | -4,37 |
| 23 | 3803 | 3368 | -435 | 435 | 14198449 | 2658 | 11 | 112 | -4,70 |
| 24 | 1831 | 3368 | 1537 | 1537 | 13705289 | 2612 | 84 | 110 | -4,19 |
| 25 | 2282 | 3368 | 1086 | 1086 | 13204260 | 2551 | 48 | 108 | -3,87 |
| 26 | 5956 | 3352 | -2604 | 2604 | 12957293 | 2553 | 44 | 105 | -4,88 |
| 27 | 4234 | 3352 | -882 | 882 | 12506234 | 2491 | 21 | 102 | -5,36 |
| 28 | 178 | 3352 | 3174 | 3174 | 12419278 | 2515 | 1783 | 162 | -4,05 |
| 29 | 4239 | 3352 | -887 | 887 | 12018184 | 2459 | 21 | 157 | -4,50 |
| 30 | 329 | 3352 | 3023 | 3023 | 11922106 | 2478 | 919 | 183 | -3,25 |

Table 5: Mean Method Forecast with T=5

For T=10:

| Week | Actual | Forecast | Error | Abs. Error | MSE | MAD | % Error | MAPE | TS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2032 | 3287 | 1255 | 1255 | 1575264 | 1255 | 62 | 62 | 1,00 |
| 2 | 824 | 3287 | 2463 | 2463 | 3821051 | 1859 | 299 | 180 | 2,00 |
| 3 | 1227 | 3287 | 2060 | 2060 | 3962032 | 1926 | 168 | 176 | 3,00 |
| 4 | 1673 | 3287 | 1614 | 1614 | 3622850 | 1848 | 96 | 156 | 4,00 |
| 5 | 1937 | 3287 | 1350 | 1350 | 3262831 | 1748 | 70 | 139 | 5,00 |
| 6 | 2243 | 3287 | 1044 | 1044 | 2900715 | 1631 | 47 | 124 | 6,00 |
| 7 | 5338 | 3287 | -2051 | 2051 | 3087214 | 1691 | 38 | 111 | 4,57 |
| 8 | 8591 | 3287 | -5304 | 5304 | 6217738 | 2143 | 62 | 105 | 1,13 |
| 9 | 3896 | 3287 | -609 | 609 | 5568075 | 1972 | 16 | 95 | 0,92 |
| 10 | 1188 | 3287 | 2099 | 2099 | 5451887 | 1985 | 177 | 103 | 1,98 |
| 11 | 1204 | 3258 | 2054 | 2054 | 5339895 | 1991 | 171 | 109 | 3,00 |
| 12 | 1350 | 3258 | 1908 | 1908 | 5198358 | 1984 | 141 | 112 | 3,97 |
| 13 | 4273 | 3258 | -1015 | 1015 | 4877692 | 1910 | 24 | 105 | 3,60 |
| 14 | 8932 | 3258 | -5674 | 5674 | 6828668 | 2179 | 64 | 102 | 0,55 |
| 15 | 3004 | 3258 | 254 | 254 | 6377733 | 2050 | 8 | 96 | 0,71 |
| 16 | 956 | 3258 | 2302 | 2302 | 6310399 | 2066 | 241 | 105 | 1,82 |
| 17 | 13380 | 3258 | -10122 | 10122 | 11965651 | 2540 | 76 | 103 | -2,51 |
| 18 | 13688 | 3258 | -10430 | 10430 | 17344200 | 2978 | 76 | 102 | -5,64 |
| 19 | 1261 | 3258 | 1997 | 1997 | 16641296 | 2927 | 158 | 105 | -5,06 |
| 20 | 568 | 3258 | 2690 | 2690 | 16171106 | 2915 | 474 | 123 | -4,16 |
| 21 | 2050 | 3368 | 1318 | 1318 | 15483783 | 2839 | 64 | 120 | -3,80 |
| 22 | 4380 | 3368 | -1012 | 1012 | 14826519 | 2756 | 23 | 116 | -4,28 |
| 23 | 3803 | 3368 | -435 | 435 | 14190112 | 2655 | 11 | 112 | -4,61 |
| 24 | 1831 | 3368 | 1537 | 1537 | 13697299 | 2608 | 84 | 110 | -4,10 |
| 25 | 2282 | 3368 | 1086 | 1086 | 13196590 | 2547 | 48 | 108 | -3,78 |
| 26 | 5956 | 3368 | -2588 | 2588 | 12946619 | 2549 | 43 | 105 | -4,79 |
| 27 | 4234 | 3368 | -866 | 866 | 12494886 | 2487 | 20 | 102 | -5,26 |
| 28 | 178 | 3368 | 3190 | 3190 | 12412090 | 2512 | 1792 | 163 | -3,93 |
| 29 | 4239 | 3368 | -871 | 871 | 12010242 | 2455 | 21 | 158 | -4,38 |
| 30 | 329 | 3368 | 3039 | 3039 | 11917767 | 2475 | 924 | 183 | -3,12 |

Table 6: Mean Method Forecast with T=10

For T=15:

| Week | Actual | Forecast | Error | Abs. Error | MSE | MAD | % Error | MAPE | TS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2032 | 3287 | 1255 | 1255 | 1575264 | 1255 | 62 | 62 | 1,00 |
| 2 | 824 | 3287 | 2463 | 2463 | 3821051 | 1859 | 299 | 180 | 2,00 |
| 3 | 1227 | 3287 | 2060 | 2060 | 3962032 | 1926 | 168 | 176 | 3,00 |
| 4 | 1673 | 3287 | 1614 | 1614 | 3622850 | 1848 | 96 | 156 | 4,00 |
| 5 | 1937 | 3287 | 1350 | 1350 | 3262831 | 1748 | 70 | 139 | 5,00 |
| 6 | 2243 | 3287 | 1044 | 1044 | 2900715 | 1631 | 47 | 124 | 6,00 |
| 7 | 5338 | 3287 | -2051 | 2051 | 3087214 | 1691 | 38 | 111 | 4,57 |
| 8 | 8591 | 3287 | -5304 | 5304 | 6217738 | 2143 | 62 | 105 | 1,13 |
| 9 | 3896 | 3287 | -609 | 609 | 5568075 | 1972 | 16 | 95 | 0,92 |
| 10 | 1188 | 3287 | 2099 | 2099 | 5451887 | 1985 | 177 | 103 | 1,98 |
| 11 | 1204 | 3287 | 2083 | 2083 | 5350742 | 1994 | 173 | 110 | 3,01 |
| 12 | 1350 | 3287 | 1937 | 1937 | 5217541 | 1989 | 143 | 113 | 3,99 |
| 13 | 4273 | 3287 | -986 | 986 | 4890962 | 1912 | 23 | 106 | 3,64 |
| 14 | 8932 | 3287 | -5645 | 5645 | 6817675 | 2179 | 63 | 103 | 0,60 |
| 15 | 3004 | 3287 | 283 | 283 | 6368506 | 2052 | 9 | 96 | 0,78 |
| 16 | 956 | 3276 | 2320 | 2320 | 6306813 | 2069 | 243 | 106 | 1,89 |
| 17 | 13380 | 3276 | -10104 | 10104 | 11941419 | 2542 | 76 | 104 | -2,44 |
| 18 | 13688 | 3276 | -10412 | 10412 | 17301017 | 2979 | 76 | 102 | -5,57 |
| 19 | 1261 | 3276 | 2015 | 2015 | 16604088 | 2928 | 160 | 105 | -4,98 |
| 20 | 568 | 3276 | 2708 | 2708 | 16140489 | 2917 | 477 | 124 | -4,07 |
| 21 | 2050 | 3276 | 1226 | 1226 | 15443444 | 2837 | 60 | 121 | -3,76 |
| 22 | 4380 | 3276 | -1104 | 1104 | 14796892 | 2758 | 25 | 116 | -4,26 |
| 23 | 3803 | 3276 | -527 | 527 | 14165634 | 2661 | 14 | 112 | -4,62 |
| 24 | 1831 | 3276 | 1445 | 1445 | 13662374 | 2610 | 79 | 111 | -4,15 |
| 25 | 2282 | 3276 | 994 | 994 | 13155384 | 2546 | 44 | 108 | -3,87 |
| 26 | 5956 | 3276 | -2680 | 2680 | 12925698 | 2551 | 45 | 106 | -4,91 |
| 27 | 4234 | 3276 | -958 | 958 | 12480974 | 2492 | 23 | 102 | -5,41 |
| 28 | 178 | 3276 | 3098 | 3098 | 12377950 | 2513 | 1740 | 161 | -4,13 |
| 29 | 4239 | 3276 | -963 | 963 | 11983116 | 2460 | 23 | 156 | -4,61 |
| 30 | 329 | 3276 | 2947 | 2947 | 11873131 | 2476 | 896 | 181 | -3,39 |

Table 7: Mean Method Forecast with T=15

For T=30:

| Week | Actual | Forecast | Error | Abs. Error | MSE | MAD | % Error | MAPE | TS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2032 | 3287 | 1255 | 1255 | 1575264 | 1255 | 62 | 62 | 1,00 |
| 2 | 824 | 3287 | 2463 | 2463 | 3821051 | 1859 | 299 | 180 | 2,00 |
| 3 | 1227 | 3287 | 2060 | 2060 | 3962032 | 1926 | 168 | 176 | 3,00 |
| 4 | 1673 | 3287 | 1614 | 1614 | 3622850 | 1848 | 96 | 156 | 4,00 |
| 5 | 1937 | 3287 | 1350 | 1350 | 3262831 | 1748 | 70 | 139 | 5,00 |
| 6 | 2243 | 3287 | 1044 | 1044 | 2900715 | 1631 | 47 | 124 | 6,00 |
| 7 | 5338 | 3287 | -2051 | 2051 | 3087214 | 1691 | 38 | 111 | 4,57 |
| 8 | 8591 | 3287 | -5304 | 5304 | 6217738 | 2143 | 62 | 105 | 1,13 |
| 9 | 3896 | 3287 | -609 | 609 | 5568075 | 1972 | 16 | 95 | 0,92 |
| 10 | 1188 | 3287 | 2099 | 2099 | 5451887 | 1985 | 177 | 103 | 1,98 |
| 11 | 1204 | 3287 | 2083 | 2083 | 5350742 | 1994 | 173 | 110 | 3,01 |
| 12 | 1350 | 3287 | 1937 | 1937 | 5217541 | 1989 | 143 | 113 | 3,99 |
| 13 | 4273 | 3287 | -986 | 986 | 4890962 | 1912 | 23 | 106 | 3,64 |
| 14 | 8932 | 3287 | -5645 | 5645 | 6817675 | 2179 | 63 | 103 | 0,60 |
| 15 | 3004 | 3287 | 283 | 283 | 6368506 | 2052 | 9 | 96 | 0,78 |
| 16 | 956 | 3287 | 2331 | 2331 | 6310100 | 2070 | 244 | 106 | 1,90 |
| 17 | 13380 | 3287 | -10093 | 10093 | 11931078 | 2542 | 75 | 104 | -2,43 |
| 18 | 13688 | 3287 | -10401 | 10401 | 17278175 | 2978 | 76 | 102 | -5,56 |
| 19 | 1261 | 3287 | 2026 | 2026 | 16584853 | 2928 | 161 | 105 | -4,97 |
| 20 | 568 | 3287 | 2719 | 2719 | 16125284 | 2918 | 479 | 124 | -4,05 |
| 21 | 2050 | 3287 | 1237 | 1237 | 15430290 | 2838 | 60 | 121 | -3,73 |
| 22 | 4380 | 3287 | -1093 | 1093 | 14783206 | 2758 | 25 | 117 | -4,23 |
| 23 | 3803 | 3287 | -516 | 516 | 14152030 | 2661 | 14 | 112 | -4,58 |
| 24 | 1831 | 3287 | 1456 | 1456 | 13650704 | 2611 | 80 | 111 | -4,11 |
| 25 | 2282 | 3287 | 1005 | 1005 | 13145085 | 2546 | 44 | 108 | -3,82 |
| 26 | 5956 | 3287 | -2669 | 2669 | 12913468 | 2551 | 45 | 106 | -4,86 |
| 27 | 4234 | 3287 | -947 | 947 | 12468400 | 2492 | 22 | 103 | -5,36 |
| 28 | 178 | 3287 | 3109 | 3109 | 12368331 | 2514 | 1747 | 161 | -4,07 |
| 29 | 4239 | 3287 | -952 | 952 | 11973083 | 2460 | 22 | 157 | -4,55 |
| 30 | 329 | 3287 | 2958 | 2958 | 11865658 | 2477 | 899 | 181 | -3,32 |

Table 8: Mean Method Forecast with T=30

Comparing the results (see Figure 8), it can be seen that the increasing of the time horizon does not have a great impact on the forecast behavior. All of the results, except for T= 15 and T=30 which are the mean values of the previous periods, have almost the same values, which widely differ from the actual data.

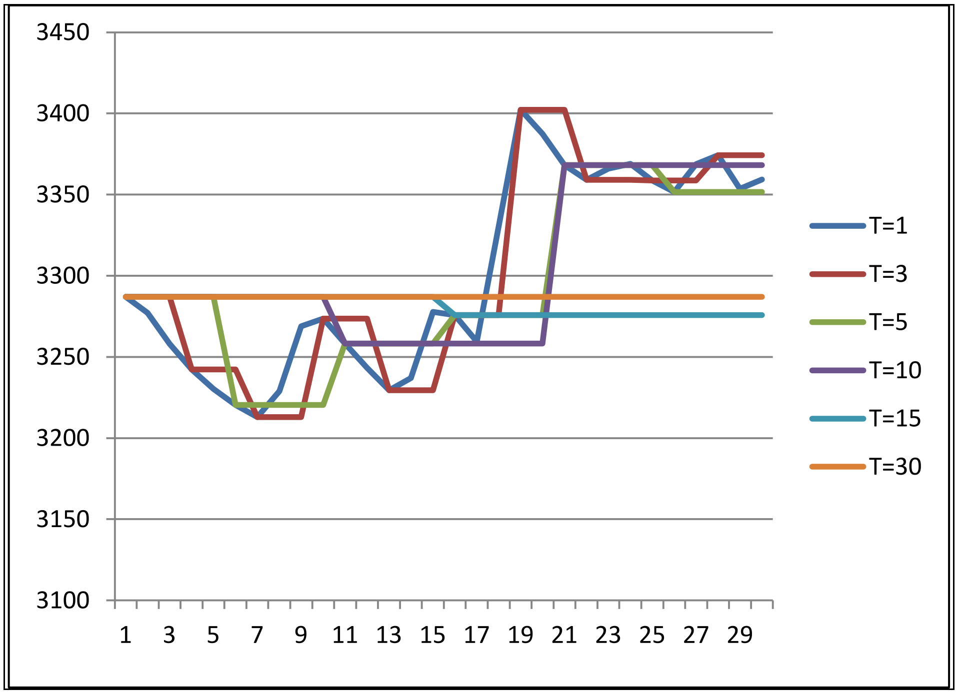


Figure 8: Forecast using the Mean Method

However, even in this case, the results raise the question whether the increasing of the time reduces the accuracy of the forecast. For this purpose we will use the measurements of the forecasting accuracy described in chapter 2.4.

| Average Values | T=1 | T=3 | T=5 | T=10 | T=15 | T=30 |
|---|---|---|---|---|---|---|
| Sum of Abs. Error | 74.470 | 74.701 | 74.339 | 74.238 | 74.284 | 74.296 |
| MAD | 2.482 | 2.490 | 2.478 | 2.475 | 2.476 | 2.477 |
| MAPE | 184 | 185 | 183 | 183 | 181 | 181 |

Table 9: Forecast accuracy of the Mean Method

As Table 9 shows, using different measures for the forecasting accuracy, it is difficult to estimate which time horizon gives the most accurate forecast. If we sum the absolute errors, the most accurate forecast is the one with 10 weeks time horizon. We receive the same result using the

Mean Absolute Deviation. Comparing the Mean Absolute Percentage Error (MAPE) the forecast results with 15 and 30 weeks time horizon give the most accurate results. This can be expected because the forecast of these two is the mean value of the stored data. The better performance of the forecast with 10 weeks time horizon can be explained with the weekly seasonality of the data. In addition, it should be mentioned that none of the results is either underforecasting (TS < -6) or overforecasting (TS > +6). So according to the theory the mean method should give a good forecast. However, as Figure 9 shows the Mean Method does not achieve the task. The statistical results have a good performance only because the forecast fluctuate around the mean.

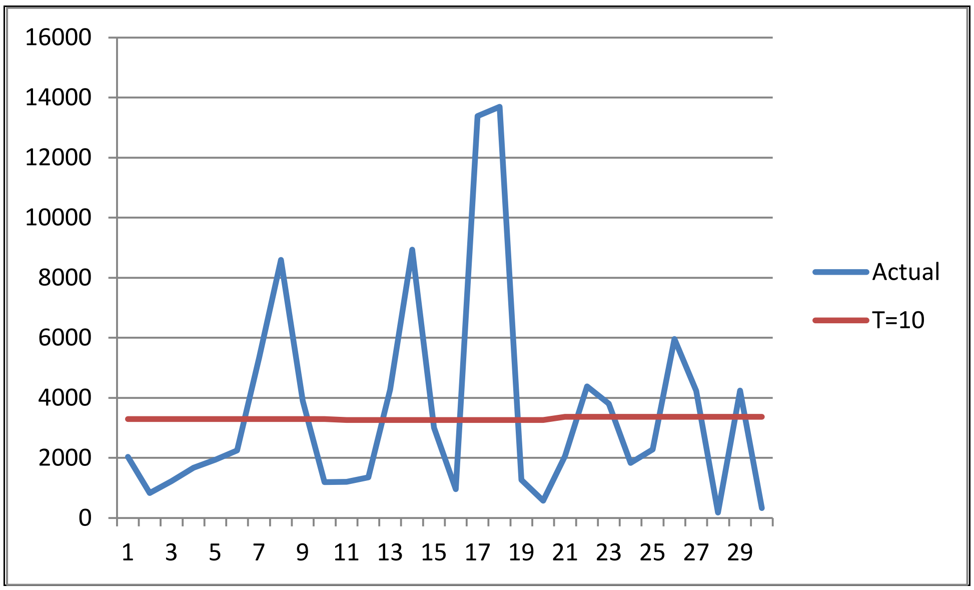


Figure 9: Comparison of the Mean Method forecast with T=10 and the actual data

### 7.1.2. Winter's Three-Parameter Trend and Seasonality Method

As Figure 7 shows, the data has trend and seasonality. In this case, according to the theory the most appropriate time series forecast method should be the Winter's model.

In order to use the model, the following steps must be done: 1) first of all, the demand must be deseasonalized; 2) to ensure that the each season is given equal weight, the length of the cycle must be determined. 3) then the initial level and the trend must be estimated by running a regression between deseasonalized demand and time. The initial level is obtained as the intercept coefficient and the trend is obtained as the variable coefficient from the regression results.

As the length of the data varies from 6 to 8 weeks, a comparison of the model with cycle length of 6, 7 and 8 weeks for the first 126 weeks will be done. The result with the lowest SSE can be considered as the one that corresponds to the data's cycle length.

The results of the comparisons are shown in Table 10 and determine that a cycle with length of 7 weeks will be the appropriate for this forecast model.

| Cycle Length | L=6 | L=7 | L=8 |
|---|---|---|---|
| SSE | 1.120.870.275 | 858.208.145 | 1.198.832.435 |

Table 10: SSE for L= 6, 7and 8 weeks

Following the steps described above, we deseasonalize the demand and after starting a regression between deseasonalized demand and time, the initial level and the trend are obtained (Table 11).

| SUMMARY OUTPUT | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Regression Statistics* | | | | | | | | |
| Multiple R | 0,422567966 | | | | | | | |
| R Square | 0,178563686 | | | | | | | |
| Adjusted R Square | 0,014276423 | | | | | | | |
| Standard Error | 418,5293714 | | | | | | | |
| Observations | 7 | | | | | | | |
| | | | | | | | | |
| ANOVA | | | | | | | | |
| | *df* | *SS* | *MS* | *F* | *Significance F* | | | |
| Regression | 1 | 190388,6835 | 190388,6835 | 1,086899148 | 0,344920755 | | | |
| Residual | 5 | 875834,1737 | 175166,8347 | | | | | |
| Total | 6 | 1066222,857 | | | | | | |
| | | | | | | | | |
| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95,0%* | *Upper 95,0%* |
| Intercept | 3233,671414 | 260,6387947 | 12,40671565 | 6,02914E-05 | 2563,679158 | 3903,663671 | 2563,679158 | 3903,663671 |
| X Variable 1 | -0,078269644 | 0,075075586 | -1,042544555 | 0,344920755 | -0,271257267 | 0,11471798 | -0,271257267 | 0,11471798 |

Table 11: SLR between deseasonalized demand and time

Using equations 2.14 to 2.17 the forecast for the historical data and the level, trend and the seasonal factor for a given period can be found (Table 13). Then minimizing the SSE, the smoothing constants α, β and γ can be estimated (Table 12).

| $\bar{\alpha}$ | β | γ |
|---|---|---|
| 0,00 | 1,00 | 0,22 |

Table 12: Estimated smoothing constants

| T (weeks) | Demand (beer canes) | Deseasonalized Demand | Ď | Š | Level | Trend | Seasonal Factor | Forecast | Error |
|---|---|---|---|---|---|---|---|---|---|
| 0 | | | | | 3.234 | -0,08 | | | |
| 1 | 1.314 | | 3.234 | 0,41 | 3.234 | -0,08 | 0,56 | 1.798 | 484 |
| 2 | 1.125 | | 3.234 | 0,35 | 3.234 | -0,08 | 0,55 | 1.767 | 642 |
| 3 | 1.517 | | 3.233 | 0,47 | 3.233 | -0,08 | 0,88 | 2.850 | 1.333 |
| 4 | 1.356 | 2.681 | 3.233 | 0,42 | 3.233 | -0,08 | 1,93 | 6.227 | 4.871 |
| 5 | 6.456 | 2.702 | 3.233 | 2,00 | 3.233 | -0,08 | 1,88 | 6.089 | -367 |
| 6 | 5.687 | 2.774 | 3.233 | 1,76 | 3.233 | -0,08 | 0,80 | 2.586 | -3.101 |
| 7 | 1.312 | 2.759 | 3.233 | 0,41 | 3.233 | -0,08 | 0,52 | 1.680 | 368 |
| 8 | 1.458 | 3.631 | 3.233 | 0,45 | 3.233 | -0,08 | 0,52 | 1.692 | 234 |
| 9 | 1.633 | 3.608 | 3.233 | 0,51 | 3.233 | -0,08 | 0,50 | 1.627 | -6 |
| 10 | 1.412 | 2.969 | 3.233 | 0,44 | 3.233 | -0,08 | 0,79 | 2.560 | 1.148 |
| 11 | 7.456 | | 3.233 | 2,31 | 3.233 | -0,08 | 1,60 | 5.167 | -2.289 |
| 12 | 6.300 | | 3.233 | 1,95 | 3.233 | -0,08 | 1,91 | 6.168 | -132 |
| 13 | 1.214 | | 3.233 | 0,38 | 3.233 | -0,08 | 1,01 | 3.260 | 2.046 |
| 14 | 1.577 | | 3.233 | 0,49 | 3.233 | -0,08 | 0,49 | 1.600 | 23 |
| 15 | 1.657 | | 3.232 | 0,51 | 3.232 | -0,08 | 0,51 | 1.641 | -16 |
| 16 | 1.798 | | 3.232 | 0,56 | 3.232 | -0,08 | 0,50 | 1.628 | -170 |
| 17 | 1.888 | | 3.232 | 0,58 | 3.232 | -0,08 | 0,71 | 2.310 | 422 |
| 18 | 7.345 | | 3.232 | 2,27 | 3.232 | -0,08 | 1,75 | 5.664 | -1.681 |
| 19 | 6.812 | | 3.232 | 2,11 | 3.232 | -0,08 | 1,92 | 6.195 | -617 |
| 20 | 1.344 | | 3.232 | 0,42 | 3.232 | -0,08 | 0,87 | 2.815 | 1.471 |
| 21 | 1.566 | | 3.232 | 0,48 | 3.232 | -0,08 | 0,49 | 1.594 | 28 |
| 22 | 1.723 | | 3.232 | 0,53 | 3.232 | -0,08 | 0,51 | 1.644 | -79 |
| 23 | 1.945 | | 3.232 | 0,60 | 3.232 | -0,08 | 0,52 | 1.665 | -280 |
| 24 | 9.876 | | 3.232 | 3,06 | 3.232 | -0,08 | 0,69 | 2.218 | -7.658 |
| 25 | 7.932 | | 3.232 | 2,45 | 3.232 | -0,08 | 1,87 | 6.028 | -1.904 |
| 26 | 1.612 | | 3.232 | 0,50 | 3.232 | -0,08 | 1,96 | 6.328 | 4.716 |
| 27 | 1.956 | | 3.232 | 0,61 | 3.232 | -0,08 | 0,77 | 2.494 | 538 |
| 28 | 2.156 | | 3.231 | 0,67 | 3.231 | -0,08 | 0,49 | 1.588 | -568 |
| 29 | 2.430 | | 3.231 | 0,75 | 3.231 | -0,08 | 0,51 | 1.661 | -769 |
| 30 | 2.204 | | 3.231 | 0,68 | 3.231 | -0,08 | 0,53 | 1.725 | -479 |
| 31 | 1.974 | | 3.231 | 0,61 | 3.231 | -0,08 | 1,20 | 3.882 | 1.908 |
| 32 | 9.467 | | 3.231 | 2,93 | 3.231 | -0,08 | 1,99 | 6.441 | -3.026 |
| 33 | 8.103 | | 3.231 | 2,51 | 3.231 | -0,08 | 1,64 | 5.302 | -2.801 |
| 34 | 1.467 | | 3.231 | 0,45 | 3.231 | -0,08 | 0,74 | 2.377 | 910 |
| 35 | 1.921 | | 3.231 | 0,59 | 3.231 | -0,08 | 0,53 | 1.711 | -210 |
| 36 | 2.133 | | 3.231 | 0,66 | 3.231 | -0,08 | 0,57 | 1.828 | -305 |
| 37 | 1.742 | | 3.231 | 0,54 | 3.231 | -0,08 | 0,57 | 1.829 | 87 |
| 38 | 1.542 | | 3.231 | 0,48 | 3.231 | -0,08 | 1,07 | 3.467 | 1.925 |
| 39 | 8.463 | | 3.231 | 2,62 | 3.231 | -0,08 | 2,20 | 7.098 | -1.365 |
| 40 | 7.453 | | 3.231 | 2,31 | 3.231 | -0,08 | 1,83 | 5.910 | -1.543 |
| 41 | 1.455 | | 3.230 | 0,45 | 3.230 | -0,08 | 0,67 | 2.179 | 724 |
| 42 | 1.688 | | 3.230 | 0,52 | 3.230 | -0,08 | 0,54 | 1.756 | 68 |
| 43 | 1.863 | | 3.230 | 0,58 | 3.230 | -0,08 | 0,59 | 1.894 | 31 |
| 44 | 1.912 | | 3.230 | 0,59 | 3.230 | -0,08 | 0,56 | 1.810 | -102 |
| 45 | 1.732 | | 3.230 | 0,54 | 3.230 | -0,08 | 0,94 | 3.048 | 1.316 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 46 | 1.548 | | 3.230 | 0,48 | 3.230 | -0,08 | 2,29 | 7.393 | 5.845 |
| 47 | 5.865 | | 3.230 | 1,82 | 3.230 | -0,08 | 1,93 | 6.244 | 379 |
| 48 | 4.972 | | 3.230 | 1,54 | 3.230 | -0,08 | 0,63 | 2.021 | -2.951 |
| 49 | 1.532 | | 3.230 | 0,47 | 3.230 | -0,08 | 0,54 | 1.741 | 209 |
| 50 | 1.412 | | 3.230 | 0,44 | 3.230 | -0,08 | 0,58 | 1.887 | 475 |
| 51 | 2.477 | | 3.230 | 0,77 | 3.230 | -0,08 | 0,57 | 1.832 | -645 |
| 52 | 1.533 | | 3.230 | 0,47 | 3.230 | -0,08 | 0,85 | 2.761 | 1.228 |
| 53 | 1.412 | | 3.230 | 0,44 | 3.230 | -0,08 | 1,90 | 6.121 | 4.709 |
| 54 | 1.138 | | 3.229 | 0,35 | 3.229 | -0,08 | 1,91 | 6.161 | 5.023 |
| 55 | 1.469 | | 3.229 | 0,45 | 3.229 | -0,08 | 0,82 | 2.662 | 1.193 |
| 56 | 6.702 | | 3.229 | 2,08 | 3.229 | -0,08 | 0,53 | 1.695 | -5.007 |
| 57 | 4.525 | | 3.229 | 1,40 | 3.229 | -0,08 | 0,55 | 1.783 | -2.742 |
| 58 | 1.301 | | 3.229 | 0,40 | 3.229 | -0,08 | 0,61 | 1.972 | 671 |
| 59 | 1.469 | | 3.229 | 0,45 | 3.229 | -0,08 | 0,77 | 2.494 | 1.025 |
| 60 | 1.280 | | 3.229 | 0,40 | 3.229 | -0,08 | 1,58 | 5.097 | 3.817 |
| 61 | 5.530 | | 3.229 | 1,71 | 3.229 | -0,08 | 1,57 | 5.068 | -462 |
| 62 | 4.657 | | 3.229 | 1,44 | 3.229 | -0,08 | 0,74 | 2.402 | -2.255 |
| 63 | 1.114 | | 3.229 | 0,35 | 3.229 | -0,08 | 0,86 | 2.783 | 1.669 |
| 64 | 1.625 | | 3.229 | 0,50 | 3.229 | -0,08 | 0,74 | 2.379 | 754 |
| 65 | 1.473 | | 3.229 | 0,46 | 3.229 | -0,08 | 0,57 | 1.826 | 353 |
| 66 | 1.864 | | 3.229 | 0,58 | 3.229 | -0,08 | 0,70 | 2.271 | 407 |
| 67 | 2.136 | | 3.228 | 0,66 | 3.228 | -0,08 | 1,32 | 4.266 | 2.130 |
| 68 | 1.870 | | 3.228 | 0,58 | 3.228 | -0,08 | 1,60 | 5.168 | 3.298 |
| 69 | 1.799 | | 3.228 | 0,56 | 3.228 | -0,08 | 0,90 | 2.892 | 1.093 |
| 70 | 1.425 | | 3.228 | 0,44 | 3.228 | -0,08 | 0,75 | 2.420 | 995 |
| 71 | 7.423 | | 3.228 | 2,30 | 3.228 | -0,08 | 0,69 | 2.215 | -5.208 |
| 72 | 6.655 | | 3.228 | 2,06 | 3.228 | -0,08 | 0,54 | 1.749 | -4.906 |
| 73 | 1.212 | | 3.228 | 0,38 | 3.228 | -0,08 | 0,68 | 2.182 | 970 |
| 74 | 1.845 | | 3.228 | 0,57 | 3.228 | -0,08 | 1,18 | 3.803 | 1.958 |
| 75 | 1.364 | | 3.228 | 0,42 | 3.228 | -0,08 | 1,38 | 4.450 | 3.086 |
| 76 | 1.798 | | 3.228 | 0,56 | 3.228 | -0,08 | 0,82 | 2.654 | 856 |
| 77 | 10.037 | | 3.228 | 3,11 | 3.228 | -0,08 | 0,68 | 2.203 | -7.834 |
| 78 | 8.563 | | 3.228 | 2,65 | 3.228 | -0,08 | 1,04 | 3.346 | -5.217 |
| 79 | 2.036 | | 3.227 | 0,63 | 3.227 | -0,08 | 0,87 | 2.815 | 779 |
| 80 | 2.455 | | 3.227 | 0,76 | 3.227 | -0,08 | 0,61 | 1.971 | -484 |
| 81 | 2.712 | | 3.227 | 0,84 | 3.227 | -0,08 | 1,05 | 3.376 | 664 |
| 82 | 1.847 | | 3.227 | 0,57 | 3.227 | -0,08 | 1,17 | 3.778 | 1.931 |
| 83 | 9.625 | | 3.227 | 2,98 | 3.227 | -0,08 | 0,76 | 2.467 | -7.158 |
| 84 | 8.533 | | 3.227 | 2,64 | 3.227 | -0,08 | 1,21 | 3.906 | -4.627 |
| 85 | 2.685 | | 3.227 | 0,83 | 3.227 | -0,08 | 1,39 | 4.480 | 1.795 |
| 86 | 2.422 | | 3.227 | 0,75 | 3.227 | -0,08 | 0,82 | 2.645 | 223 |
| 87 | 2.512 | | 3.227 | 0,78 | 3.227 | -0,08 | 0,64 | 2.076 | -436 |
| 88 | 1.864 | | 3.227 | 0,58 | 3.227 | -0,08 | 1,00 | 3.231 | 1.367 |
| 89 | 1.956 | | 3.227 | 0,61 | 3.227 | -0,08 | 1,04 | 3.358 | 1.402 |
| 90 | 2.135 | | 3.227 | 0,66 | 3.227 | -0,08 | 1,25 | 4.023 | 1.888 |
| 91 | 2.145 | | 3.227 | 0,66 | 3.227 | -0,08 | 1,52 | 4.911 | 2.766 |
| 92 | 8.465 | | 3.226 | 2,62 | 3.226 | -0,08 | 1,27 | 4.089 | -4.376 |
| 93 | 7.388 | | 3.226 | 2,29 | 3.226 | -0,08 | 0,80 | 2.596 | -4.792 |
| 94 | 1.586 | | 3.226 | 0,49 | 3.226 | -0,08 | 0,67 | 2.170 | 584 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 95 | 6.432 | | 3.226 | 1,99 | 3.226 | -0,08 | 0,91 | 2.934 | -3.498 |
| 96 | 5.549 | | 3.226 | 1,72 | 3.226 | -0,08 | 0,95 | 3.053 | -2.496 |
| 97 | 1.659 | | 3.226 | 0,51 | 3.226 | -0,08 | 1,12 | 3.612 | 1.953 |
| 98 | 1.736 | | 3.226 | 0,54 | 3.226 | -0,08 | 1,34 | 4.309 | 2.573 |
| 99 | 1.487 | | 3.226 | 0,46 | 3.226 | -0,08 | 1,56 | 5.039 | 3.552 |
| 100 | 5.348 | | 3.226 | 1,66 | 3.226 | -0,08 | 1,13 | 3.637 | -1.711 |
| 101 | 4.650 | | 3.226 | 1,44 | 3.226 | -0,08 | 0,63 | 2.043 | -2.607 |
| 102 | 1.325 | | 3.226 | 0,41 | 3.226 | -0,08 | 1,15 | 3.694 | 2.369 |
| 103 | 2.827 | | 3.226 | 0,88 | 3.226 | -0,08 | 1,11 | 3.595 | 768 |
| 104 | 1.833 | | 3.226 | 0,57 | 3.226 | -0,08 | 0,99 | 3.187 | 1.354 |
| 105 | 844 | | 3.225 | 0,26 | 3.225 | -0,08 | 1,16 | 3.749 | 2.905 |
| 106 | 1.051 | | 3.225 | 0,33 | 3.225 | -0,08 | 1,32 | 4.266 | 3.215 |
| 107 | 2.256 | | 3.225 | 0,70 | 3.225 | -0,08 | 1,24 | 4.008 | 1.752 |
| 108 | 5.176 | | 3.225 | 1,60 | 3.225 | -0,08 | 0,81 | 2.609 | -2.567 |
| 109 | 2.907 | | 3.225 | 0,90 | 3.225 | -0,08 | 0,99 | 3.178 | 271 |
| 110 | 1.099 | | 3.225 | 0,34 | 3.225 | -0,08 | 1,06 | 3.427 | 2.328 |
| 111 | 1.243 | | 3.225 | 0,39 | 3.225 | -0,08 | 0,90 | 2.892 | 1.649 |
| 112 | 2.658 | | 3.225 | 0,82 | 3.225 | -0,08 | 0,97 | 3.117 | 459 |
| 113 | 2.608 | | 3.225 | 0,81 | 3.225 | -0,08 | 1,11 | 3.567 | 959 |
| 114 | 7.319 | | 3.225 | 2,27 | 3.225 | -0,08 | 1,12 | 3.627 | -3.692 |
| 115 | 5.701 | | 3.225 | 1,77 | 3.225 | -0,08 | 0,98 | 3.167 | -2.534 |
| 116 | 1.457 | | 3.225 | 0,45 | 3.225 | -0,08 | 0,97 | 3.119 | 1.662 |
| 117 | 1.158 | | 3.225 | 0,36 | 3.225 | -0,08 | 0,91 | 2.921 | 1.763 |
| 118 | 2.767 | | 3.224 | 0,86 | 3.224 | -0,08 | 0,79 | 2.533 | -234 |
| 119 | 2.959 | | 3.224 | 0,92 | 3.224 | -0,08 | 0,94 | 3.017 | 58 |
| 120 | 11.504 | | 3.224 | 3,57 | 3.224 | -0,08 | 1,04 | 3.358 | -8.146 |
| 121 | 4.291 | | 3.224 | 1,33 | 3.224 | -0,08 | 1,37 | 4.429 | 138 |
| 122 | 1.102 | | 3.224 | 0,34 | 3.224 | -0,08 | 1,15 | 3.717 | 2.615 |
| 123 | 1.048 | | 3.224 | 0,33 | 3.224 | -0,08 | 0,86 | 2.757 | 1.709 |
| 124 | 1.259 | | 3.224 | 0,39 | 3.224 | -0,08 | 0,79 | 2.537 | 1.278 |
| 125 | 6.361 | | 3.224 | 1,97 | 3.224 | -0,08 | 0,80 | 2.583 | -3.778 |
| 126 | 6.193 | | 3.224 | 1,92 | 3.224 | -0,08 | 0,93 | 3.004 | -3.189 |

Table 13: Estimation of the trend, level and the seasonal factor of the Winter's model

Using equation 2.17 the demand for the next 30 weeks can be estimated with forecast period from T= 1, 3, 5, 10, 15 and 30 weeks and accuracy of the forecast can be measured with the statistical measures described in chapter 2.4. Applying these rules, we receive:

For T=1

| Week | Actual | Forecast | Error | Abs. Error | MSE | MAD | % Error | MAPE | TS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2.032 | 5.128 | 3.096 | 3.096 | 9.584.142 | 3.096 | 152 | 152 | 1,00 |
| 2 | 824 | 4.398 | 3.574 | 3.574 | 11.179.015 | 3.335 | 434 | 293 | 2,00 |
| 3 | 1.227 | 3.148 | 1.921 | 1.921 | 8.682.677 | 2.864 | 157 | 248 | 3,00 |
| 4 | 1.673 | 2.385 | 712 | 712 | 6.638.751 | 2.326 | 43 | 196 | 4,00 |
| 5 | 1.937 | 2.259 | 322 | 322 | 5.331.712 | 1.925 | 17 | 160 | 5,00 |
| 6 | 2.243 | 3.404 | 1.161 | 1.161 | 4.667.783 | 1.798 | 52 | 142 | 6,00 |
| 7 | 5.338 | 3.696 | -1.642 | 1.642 | 4.385.989 | 1.775 | 31 | 126 | 5,15 |
| 8 | 8.591 | 4.454 | -4.137 | 4.137 | 5.977.031 | 2.071 | 48 | 117 | 2,42 |
| 9 | 3.896 | 3.620 | -276 | 276 | 5.321.352 | 1.871 | 7 | 104 | 2,53 |
| 10 | 1.188 | 2.730 | 1.542 | 1.542 | 5.026.954 | 1.838 | 130 | 107 | 3,41 |
| 11 | 1.204 | 2.230 | 1.026 | 1.026 | 4.665.628 | 1.764 | 85 | 105 | 4,14 |
| 12 | 1.350 | 2.188 | 838 | 838 | 4.335.412 | 1.687 | 62 | 101 | 4,82 |
| 13 | 4.273 | 3.151 | -1.122 | 1.122 | 4.098.731 | 1.644 | 26 | 96 | 4,27 |
| 14 | 8.932 | 4.052 | -4.880 | 4.880 | 5.506.647 | 1.875 | 55 | 93 | 1,14 |
| 15 | 3.004 | 5.352 | 2.348 | 2.348 | 5.507.232 | 1.906 | 78 | 92 | 2,35 |
| 16 | 956 | 3.680 | 2.724 | 2.724 | 5.626.701 | 1.957 | 285 | 104 | 3,68 |
| 17 | 13.380 | 2.394 | -10.986 | 10.986 | 12.394.909 | 2.489 | 82 | 103 | -1,52 |
| 18 | 13.688 | 2.006 | -11.682 | 11.682 | 19.287.276 | 2.999 | 85 | 102 | -5,15 |
| 19 | 1.261 | 2.006 | 745 | 745 | 18.301.356 | 2.881 | 59 | 99 | -5,11 |
| 20 | 568 | 3.394 | 2.826 | 2.826 | 17.785.730 | 2.878 | 498 | 119 | -4,13 |
| 21 | 2.050 | 5.112 | 3.062 | 3.062 | 17.385.377 | 2.887 | 149 | 121 | -3,06 |
| 22 | 4.380 | 4.841 | 461 | 461 | 16.604.797 | 2.776 | 11 | 116 | -3,01 |
| 23 | 3.803 | 3.087 | -716 | 716 | 15.905.134 | 2.687 | 19 | 111 | -3,38 |
| 24 | 1.831 | 4.782 | 2.951 | 2.951 | 15.605.194 | 2.698 | 161 | 114 | -2,27 |
| 25 | 2.282 | 4.545 | 2.263 | 2.263 | 15.185.872 | 2.680 | 99 | 113 | -1,44 |
| 26 | 5.956 | 1.844 | -4.112 | 4.112 | 15.252.253 | 2.736 | 69 | 111 | -2,92 |
| 27 | 4.234 | 2.780 | -1.454 | 1.454 | 14.765.706 | 2.688 | 34 | 108 | -3,51 |
| 28 | 178 | 4.446 | 4.268 | 4.268 | 14.888.894 | 2.745 | 2.398 | 190 | -1,88 |
| 29 | 4.239 | 4.740 | 501 | 501 | 14.384.141 | 2.667 | 12 | 184 | -1,75 |
| 30 | 329 | 3.242 | 2.913 | 2.913 | 14.187.555 | 2.675 | 885 | 207 | -0,65 |

Table 14: Winter's method forecast with T=1

For T=3

| Week | Actual | Forecast | Error | Abs. Error | MSE | MAD | % Error | MAPE | TS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2.032 | 5.128 | 3.096 | 3.096 | 9.584.142 | 3.096 | 152 | 152 | 1,00 |
| 2 | 824 | 8.796 | 7.972 | 7.972 | 36.567.679 | 5.534 | 967 | 560 | 2,00 |
| 3 | 1.227 | 9.443 | 8.216 | 8.216 | 46.881.285 | 6.428 | 670 | 596 | 3,00 |
| 4 | 1.673 | 2.385 | 712 | 712 | 35.287.707 | 4.999 | 43 | 458 | 4,00 |
| 5 | 1.937 | 4.517 | 2.580 | 2.580 | 29.561.944 | 4.515 | 133 | 393 | 5,00 |
| 6 | 2.243 | 10.212 | 7.969 | 7.969 | 35.218.550 | 5.091 | 355 | 387 | 6,00 |
| 7 | 5.338 | 3.696 | -1.642 | 1.642 | 30.572.361 | 4.598 | 31 | 336 | 6,29 |
| 8 | 8.591 | 8.908 | 317 | 317 | 26.763.369 | 4.063 | 4 | 294 | 7,19 |
| 9 | 3.896 | 10.861 | 6.965 | 6.965 | 29.179.580 | 4.385 | 179 | 282 | 8,25 |
| 10 | 1.188 | 2.730 | 1.542 | 1.542 | 26.499.360 | 4.101 | 130 | 266 | 9,20 |
| 11 | 1.204 | 4.460 | 3.256 | 3.256 | 25.053.864 | 4.024 | 270 | 267 | 10,18 |
| 12 | 1.350 | 6.565 | 5.215 | 5.215 | 25.232.467 | 4.123 | 386 | 277 | 11,20 |
| 13 | 4.273 | 3.151 | -1.122 | 1.122 | 23.388.320 | 3.893 | 26 | 257 | 11,58 |
| 14 | 8.932 | 8.105 | -827 | 827 | 21.766.602 | 3.674 | 9 | 240 | 12,05 |
| 15 | 3.004 | 16.057 | 13.053 | 13.053 | 31.673.691 | 4.299 | 435 | 253 | 13,33 |
| 16 | 956 | 3.680 | 2.724 | 2.724 | 30.157.757 | 4.200 | 285 | 255 | 14,29 |
| 17 | 13.380 | 4.788 | -8.592 | 8.592 | 32.725.838 | 4.459 | 64 | 243 | 11,54 |
| 18 | 13.688 | 6.019 | -7.669 | 7.669 | 34.174.987 | 4.637 | 56 | 233 | 9,44 |
| 19 | 1.261 | 2.006 | 745 | 745 | 32.405.503 | 4.432 | 59 | 224 | 10,04 |
| 20 | 568 | 6.789 | 6.221 | 6.221 | 32.720.107 | 4.522 | 1.095 | 267 | 11,22 |
| 21 | 2.050 | 15.336 | 13.286 | 13.286 | 39.568.204 | 4.939 | 648 | 286 | 12,96 |
| 22 | 4.380 | 4.841 | 461 | 461 | 37.779.314 | 4.735 | 11 | 273 | 13,62 |
| 23 | 3.803 | 6.174 | 2.371 | 2.371 | 36.381.160 | 4.633 | 62 | 264 | 14,43 |
| 24 | 1.831 | 14.344 | 12.513 | 12.513 | 41.389.642 | 4.961 | 683 | 281 | 16,00 |
| 25 | 2.282 | 4.545 | 2.263 | 2.263 | 39.938.942 | 4.853 | 99 | 274 | 16,82 |
| 26 | 5.956 | 3.687 | -2.269 | 2.269 | 38.600.820 | 4.754 | 38 | 265 | 16,69 |
| 27 | 4.234 | 8.338 | 4.104 | 4.104 | 37.795.019 | 4.730 | 97 | 259 | 17,65 |
| 28 | 178 | 4.446 | 4.268 | 4.268 | 37.095.731 | 4.713 | 2.398 | 335 | 18,61 |
| 29 | 4.239 | 9.480 | 5.241 | 5.241 | 36.763.710 | 4.731 | 124 | 328 | 19,65 |
| 30 | 329 | 9.726 | 9.397 | 9.397 | 38.481.733 | 4.887 | 2.856 | 412 | 20,95 |

Table 15: Winter's method forecast with T=3

For T=5

| Week | Actual | Forecast | Error | Abs. Error | MSE | MAD | % Error | MAPE | TS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2.032 | 5.128 | 3.096 | 3.096 | 9.584.142 | 3.096 | 152 | 152 | 1,00 |
| 2 | 824 | 8.796 | 7.972 | 7.972 | 36.567.679 | 5.534 | 967 | 560 | 2,00 |
| 3 | 1.227 | 9.443 | 8.216 | 8.216 | 46.881.285 | 6.428 | 670 | 596 | 3,00 |
| 4 | 1.673 | 9.539 | 7.866 | 7.866 | 50.630.977 | 6.788 | 470 | 565 | 4,00 |
| 5 | 1.937 | 11.293 | 9.356 | 9.356 | 58.011.297 | 7.301 | 483 | 549 | 5,00 |
| 6 | 2.243 | 3.404 | 1.161 | 1.161 | 48.567.437 | 6.278 | 52 | 466 | 6,00 |
| 7 | 5.338 | 7.392 | 2.054 | 2.054 | 42.232.164 | 5.675 | 38 | 405 | 7,00 |
| 8 | 8.591 | 13.362 | 4.771 | 4.771 | 39.797.870 | 5.562 | 56 | 361 | 8,00 |
| 9 | 3.896 | 14.481 | 10.585 | 10.585 | 47.824.538 | 6.120 | 272 | 351 | 9,00 |
| 10 | 1.188 | 13.648 | 12.460 | 12.460 | 58.567.341 | 6.754 | 1.049 | 421 | 10,00 |
| 11 | 1.204 | 2.230 | 1.026 | 1.026 | 53.338.708 | 6.233 | 85 | 390 | 11,00 |
| 12 | 1.350 | 4.377 | 3.027 | 3.027 | 49.657.290 | 5.966 | 224 | 377 | 12,00 |
| 13 | 4.273 | 9.453 | 5.180 | 5.180 | 47.901.507 | 5.905 | 121 | 357 | 13,00 |
| 14 | 8.932 | 16.209 | 7.277 | 7.277 | 48.262.247 | 6.003 | 81 | 337 | 14,00 |
| 15 | 3.004 | 26.760 | 23.756 | 23.756 | 82.667.501 | 7.187 | 791 | 367 | 15,00 |
| 16 | 956 | 3.680 | 2.724 | 2.724 | 77.964.454 | 6.908 | 285 | 362 | 16,00 |
| 17 | 13.380 | 4.788 | -8.592 | 8.592 | 77.720.376 | 7.007 | 64 | 345 | 14,55 |
| 18 | 13.688 | 6.019 | -7.669 | 7.669 | 76.669.829 | 7.044 | 56 | 329 | 13,38 |
| 19 | 1.261 | 8.023 | 6.762 | 6.762 | 75.040.999 | 7.029 | 536 | 340 | 14,37 |
| 20 | 568 | 16.971 | 16.403 | 16.403 | 84.741.228 | 7.498 | 2.888 | 467 | 15,66 |
| 21 | 2.050 | 5.112 | 3.062 | 3.062 | 81.152.518 | 7.286 | 149 | 452 | 16,54 |
| 22 | 4.380 | 9.682 | 5.302 | 5.302 | 78.741.552 | 7.196 | 121 | 437 | 17,48 |
| 23 | 3.803 | 9.261 | 5.458 | 5.458 | 76.613.124 | 7.121 | 144 | 424 | 18,43 |
| 24 | 1.831 | 19.125 | 17.294 | 17.294 | 85.883.224 | 7.544 | 945 | 446 | 19,69 |
| 25 | 2.282 | 22.724 | 20.442 | 20.442 | 99.162.699 | 8.060 | 896 | 464 | 20,97 |
| 26 | 5.956 | 1.844 | -4.112 | 4.112 | 95.999.201 | 7.909 | 69 | 449 | 20,85 |
| 27 | 4.234 | 5.559 | 1.325 | 1.325 | 92.508.690 | 7.665 | 31 | 433 | 21,68 |
| 28 | 178 | 13.337 | 13.159 | 13.159 | 95.389.111 | 7.861 | 7.393 | 682 | 22,82 |
| 29 | 4.239 | 18.959 | 14.720 | 14.720 | 99.571.395 | 8.097 | 347 | 670 | 23,97 |
| 30 | 329 | 16.209 | 15.880 | 15.880 | 104.658.459 | 8.357 | 4.827 | 809 | 25,12 |

Table 16: Winter's method forecast with T=5

For T=10

| Week | Actual | Forecast | Error | Abs. Error | MSE | MAD | % Error | MAPE | TS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2.032 | 5.128 | 3.096 | 3.096 | 9.584.142 | 3.096 | 152 | 152 | 1,00 |
| 2 | 824 | 8.796 | 7.972 | 7.972 | 36.567.679 | 5.534 | 967 | 560 | 2,00 |
| 3 | 1.227 | 9.443 | 8.216 | 8.216 | 46.881.285 | 6.428 | 670 | 596 | 3,00 |
| 4 | 1.673 | 9.539 | 7.866 | 7.866 | 50.630.977 | 6.788 | 470 | 565 | 4,00 |
| 5 | 1.937 | 11.293 | 9.356 | 9.356 | 58.011.297 | 7.301 | 483 | 549 | 5,00 |
| 6 | 2.243 | 20.422 | 18.179 | 18.179 | 103.422.622 | 9.114 | 810 | 592 | 6,00 |
| 7 | 5.338 | 25.870 | 20.532 | 20.532 | 148.872.605 | 10.745 | 385 | 563 | 7,00 |
| 8 | 8.591 | 35.626 | 27.035 | 27.035 | 221.627.488 | 12.782 | 315 | 532 | 8,00 |
| 9 | 3.896 | 32.578 | 28.682 | 28.682 | 288.407.489 | 14.548 | 736 | 554 | 9,00 |
| 10 | 1.188 | 27.293 | 26.105 | 26.105 | 327.712.608 | 15.704 | 2.197 | 719 | 10,00 |
| 11 | 1.204 | 2.230 | 1.026 | 1.026 | 298.016.223 | 14.370 | 85 | 661 | 11,00 |
| 12 | 1.350 | 4.377 | 3.027 | 3.027 | 273.945.012 | 13.424 | 224 | 625 | 12,00 |
| 13 | 4.273 | 9.453 | 5.180 | 5.180 | 254.936.327 | 12.790 | 121 | 586 | 13,00 |
| 14 | 8.932 | 16.209 | 7.277 | 7.277 | 240.508.866 | 12.396 | 81 | 550 | 14,00 |
| 15 | 3.004 | 26.760 | 23.756 | 23.756 | 262.097.679 | 13.154 | 791 | 566 | 15,00 |
| 16 | 956 | 22.076 | 21.120 | 21.120 | 273.594.292 | 13.652 | 2.209 | 669 | 16,00 |
| 17 | 13.380 | 16.757 | 3.377 | 3.377 | 258.171.523 | 13.047 | 25 | 631 | 17,00 |
| 18 | 13.688 | 16.049 | 2.361 | 2.361 | 244.138.410 | 12.454 | 17 | 597 | 18,00 |
| 19 | 1.261 | 18.049 | 16.788 | 16.788 | 246.122.770 | 12.682 | 1.331 | 635 | 19,00 |
| 20 | 568 | 33.937 | 33.369 | 33.369 | 289.491.457 | 13.716 | 5.875 | 897 | 20,00 |
| 21 | 2.050 | 5.112 | 3.062 | 3.062 | 276.152.736 | 13.209 | 149 | 862 | 21,00 |
| 22 | 4.380 | 9.682 | 5.302 | 5.302 | 264.878.124 | 12.849 | 121 | 828 | 22,00 |
| 23 | 3.803 | 9.261 | 5.458 | 5.458 | 254.656.802 | 12.528 | 144 | 798 | 23,00 |
| 24 | 1.831 | 19.125 | 17.294 | 17.294 | 256.508.415 | 12.727 | 945 | 804 | 24,00 |
| 25 | 2.282 | 22.724 | 20.442 | 20.442 | 262.962.882 | 13.035 | 896 | 808 | 25,00 |
| 26 | 5.956 | 11.060 | 5.104 | 5.104 | 253.851.002 | 12.730 | 86 | 780 | 26,00 |
| 27 | 4.234 | 19.454 | 15.220 | 15.220 | 253.028.498 | 12.822 | 359 | 765 | 27,00 |
| 28 | 178 | 35.561 | 35.383 | 35.383 | 288.704.848 | 13.628 | 19.878 | 1.447 | 28,00 |
| 29 | 4.239 | 42.652 | 38.413 | 38.413 | 329.631.722 | 14.483 | 906 | 1.429 | 29,00 |
| 30 | 329 | 32.415 | 32.086 | 32.086 | 352.960.238 | 15.070 | 9.752 | 1.706 | 30,00 |

Table 17: Winter's method forecast with T=10

For T=15

| Week | Actual | Forecast | Error | Abs. Error | MSE | MAD | % Error | MAPE | TS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2.032 | 5.128 | 3.096 | 3.096 | 9.584.142 | 3.096 | 152 | 152 | 1,00 |
| 2 | 824 | 8.796 | 7.972 | 7.972 | 36.567.679 | 5.534 | 967 | 560 | 2,00 |
| 3 | 1.227 | 9.443 | 8.216 | 8.216 | 46.881.285 | 6.428 | 670 | 596 | 3,00 |
| 4 | 1.673 | 9.539 | 7.866 | 7.866 | 50.630.977 | 6.788 | 470 | 565 | 4,00 |
| 5 | 1.937 | 11.293 | 9.356 | 9.356 | 58.011.297 | 7.301 | 483 | 549 | 5,00 |
| 6 | 2.243 | 20.422 | 18.179 | 18.179 | 103.422.622 | 9.114 | 810 | 592 | 6,00 |
| 7 | 5.338 | 25.870 | 20.532 | 20.532 | 148.872.605 | 10.745 | 385 | 563 | 7,00 |
| 8 | 8.591 | 35.626 | 27.035 | 27.035 | 221.627.488 | 12.782 | 315 | 532 | 8,00 |
| 9 | 3.896 | 32.578 | 28.682 | 28.682 | 288.407.489 | 14.548 | 736 | 554 | 9,00 |
| 10 | 1.188 | 27.293 | 26.105 | 26.105 | 327.712.608 | 15.704 | 2.197 | 719 | 10,00 |
| 11 | 1.204 | 24.522 | 23.318 | 23.318 | 347.352.249 | 16.396 | 1.937 | 829 | 11,00 |
| 12 | 1.350 | 26.255 | 24.905 | 24.905 | 370.092.801 | 17.105 | 1.845 | 914 | 12,00 |
| 13 | 4.273 | 40.953 | 36.680 | 36.680 | 445.117.750 | 18.611 | 858 | 910 | 13,00 |
| 14 | 8.932 | 56.717 | 47.785 | 47.785 | 576.424.296 | 20.695 | 535 | 883 | 14,00 |
| 15 | 3.004 | 80.260 | 77.256 | 77.256 | 935.896.253 | 24.466 | 2.572 | 996 | 15,00 |
| 16 | 956 | 3.680 | 2.724 | 2.724 | 877.866.409 | 23.107 | 285 | 951 | 16,00 |
| 17 | 13.380 | 4.788 | -8.592 | 8.592 | 830.569.275 | 22.253 | 64 | 899 | 16,23 |
| 18 | 13.688 | 6.019 | -7.669 | 7.669 | 787.693.789 | 21.443 | 56 | 852 | 16,48 |
| 19 | 1.261 | 8.023 | 6.762 | 6.762 | 748.642.645 | 20.670 | 536 | 835 | 17,43 |
| 20 | 568 | 16.971 | 16.403 | 16.403 | 724.662.792 | 20.457 | 2.888 | 938 | 18,41 |
| 21 | 2.050 | 30.671 | 28.621 | 28.621 | 729.161.901 | 20.845 | 1.396 | 960 | 19,44 |
| 22 | 4.380 | 33.883 | 29.503 | 29.503 | 735.582.788 | 21.239 | 674 | 947 | 20,47 |
| 23 | 3.803 | 24.692 | 20.889 | 20.889 | 722.573.587 | 21.224 | 549 | 930 | 21,47 |
| 24 | 1.831 | 43.027 | 41.196 | 41.196 | 763.178.855 | 22.056 | 2.250 | 985 | 22,53 |
| 25 | 2.282 | 45.442 | 43.160 | 43.160 | 807.163.889 | 22.900 | 1.891 | 1.021 | 23,58 |
| 26 | 5.956 | 20.275 | 14.319 | 14.319 | 784.004.789 | 22.570 | 240 | 991 | 24,56 |
| 27 | 4.234 | 33.345 | 29.111 | 29.111 | 786.355.443 | 22.812 | 688 | 980 | 25,57 |
| 28 | 178 | 57.780 | 57.602 | 57.602 | 876.770.257 | 24.055 | 32.361 | 2.100 | 26,65 |
| 29 | 4.239 | 66.340 | 62.101 | 62.101 | 979.520.795 | 25.367 | 1.465 | 2.078 | 27,72 |
| 30 | 329 | 48.616 | 48.287 | 48.287 | 1.024.591.331 | 26.131 | 14.677 | 2.498 | 28,76 |

Table 18: Winter’s method forecast with T=15

For T=30

| Week | Actual | Forecast | Error | Abs. Error | MSE | MAD | % Error | MAPE | TS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2.032 | 5.128 | 3.096 | 3.096 | 9.584.142 | 3.096 | 152 | 152 | 1,00 |
| 2 | 824 | 8.796 | 7.972 | 7.972 | 36.567.679 | 5.534 | 967 | 560 | 2,00 |
| 3 | 1.227 | 9.443 | 8.216 | 8.216 | 46.881.285 | 6.428 | 670 | 596 | 3,00 |
| 4 | 1.673 | 9.539 | 7.866 | 7.866 | 50.630.977 | 6.788 | 470 | 565 | 4,00 |
| 5 | 1.937 | 11.293 | 9.356 | 9.356 | 58.011.297 | 7.301 | 483 | 549 | 5,00 |
| 6 | 2.243 | 20.422 | 18.179 | 18.179 | 103.422.622 | 9.114 | 810 | 592 | 6,00 |
| 7 | 5.338 | 25.870 | 20.532 | 20.532 | 148.872.605 | 10.745 | 385 | 563 | 7,00 |
| 8 | 8.591 | 35.626 | 27.035 | 27.035 | 221.627.488 | 12.782 | 315 | 532 | 8,00 |
| 9 | 3.896 | 32.578 | 28.682 | 28.682 | 288.407.489 | 14.548 | 736 | 554 | 9,00 |
| 10 | 1.188 | 27.293 | 26.105 | 26.105 | 327.712.608 | 15.704 | 2.197 | 719 | 10,00 |
| 11 | 1.204 | 24.522 | 23.318 | 23.318 | 347.352.249 | 16.396 | 1.937 | 829 | 11,00 |
| 12 | 1.350 | 26.255 | 24.905 | 24.905 | 370.092.801 | 17.105 | 1.845 | 914 | 12,00 |
| 13 | 4.273 | 40.953 | 36.680 | 36.680 | 445.117.750 | 18.611 | 858 | 910 | 13,00 |
| 14 | 8.932 | 56.717 | 47.785 | 47.785 | 576.424.296 | 20.695 | 535 | 883 | 14,00 |
| 15 | 3.004 | 80.260 | 77.256 | 77.256 | 935.896.253 | 24.466 | 2.572 | 996 | 15,00 |
| 16 | 956 | 58.854 | 57.898 | 57.898 | 1.086.916.383 | 26.555 | 6.056 | 1.312 | 16,00 |
| 17 | 13.380 | 40.687 | 27.307 | 27.307 | 1.066.842.550 | 26.599 | 204 | 1.247 | 17,00 |
| 18 | 13.688 | 36.102 | 22.414 | 22.414 | 1.035.484.001 | 26.367 | 164 | 1.186 | 18,00 |
| 19 | 1.261 | 38.094 | 36.833 | 36.833 | 1.052.390.297 | 26.918 | 2.921 | 1.278 | 19,00 |
| 20 | 568 | 67.858 | 67.290 | 67.290 | 1.226.165.970 | 28.936 | 11.847 | 1.806 | 20,00 |
| 21 | 2.050 | 107.308 | 105.258 | 105.258 | 1.695.363.367 | 32.571 | 5.135 | 1.965 | 21,00 |
| 22 | 4.380 | 106.450 | 102.070 | 102.070 | 2.091.862.802 | 35.730 | 2.330 | 1.981 | 22,00 |
| 23 | 3.803 | 70.965 | 67.162 | 67.162 | 2.197.031.488 | 37.096 | 1.766 | 1.972 | 23,00 |
| 24 | 1.831 | 114.697 | 112.866 | 112.866 | 2.636.264.504 | 40.253 | 6.164 | 2.147 | 24,00 |
| 25 | 2.282 | 113.564 | 111.282 | 111.282 | 3.026.162.569 | 43.095 | 4.877 | 2.256 | 25,00 |
| 26 | 5.956 | 47.905 | 41.949 | 41.949 | 2.977.452.299 | 43.051 | 704 | 2.196 | 26,00 |
| 27 | 4.234 | 75.000 | 70.766 | 70.766 | 3.052.650.078 | 44.077 | 1.671 | 2.177 | 27,00 |
| 28 | 178 | 124.403 | 124.225 | 124.225 | 3.494.768.710 | 46.939 | 69.790 | 4.591 | 28,00 |
| 29 | 4.239 | 137.369 | 133.130 | 133.130 | 3.985.413.394 | 49.912 | 3.141 | 4.541 | 29,00 |
| 30 | 329 | 97.197 | 96.868 | 96.868 | 4.165.344.041 | 51.477 | 29.443 | 5.372 | 30,00 |

Table 19: Winter’s method forecast with T=30

Comparing the Winter's forecast results with the actual demand (see Figure 10), it can be seen that the increase of the time horizon has a great impact on the forecast behavior. The results with shorter time horizon forecast better the real demand. Figure 11 shows the comparison of the Winter's model with T= 1, 3 and 5 weeks with the actual demand. As it can be seen from the graphic the forecast with time horizon of one week closely follows the demand, although it fails to predict the peeks in week 8, 14, 18 and 26. The forecast with time horizon of 3 and 5 weeks give results that are above the actual demand.

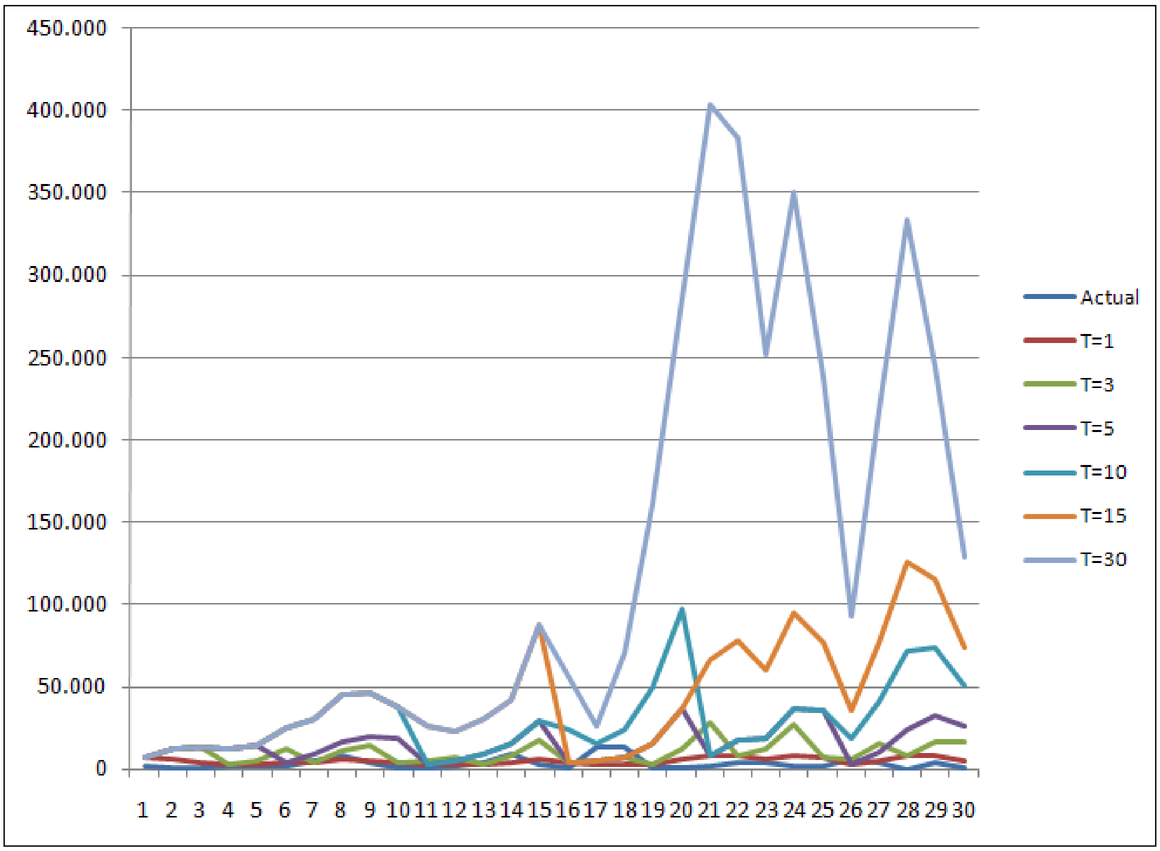


Figure 10: Forecast using the Winter's model

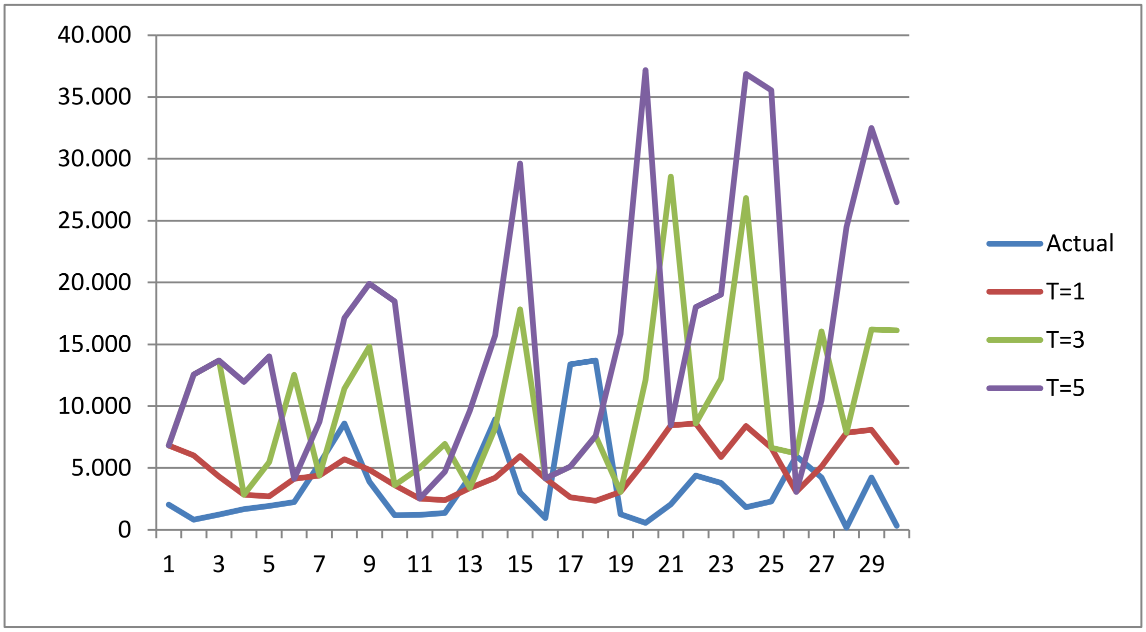


Figure 11: Forecast using the Winter's model with T=1; 3 and 5 weeks

Using the measurements of the forecasting accuracy described in chapter 2.4., shown in Table 20, it can be confirmed again that the increase of the time horizon impacts the forecast accuracy. It should be also noticed that the forecast with time horizon of 1 week has not only the best statistical measurements but it is also the only result which is neither underforecasting (TS < -6), nor overforecasting (TS > +6).

The conclusion that can be made by the graphic comparison and the statistical measurements of the forecasting accuracy verifies the theory behind the Winter's model that the forecast with the shortest time horizon is the most accurate.

| Average Values | T=1 | T=3 | T=5 |
| --- | --- | --- | --- |
| Sum of Abs. Error | 2.675 | 4.887 | 8.357 |
| MAD | 2.407 | 4.567 | 6.780 |
| MAPE | 134 | 307 | 445 |

Table 20: Forecast accuracy of the Winter's model

### 7.1.3. Comparison between the Mean method and the Winter's model

Figure 12 shows a comparison of the best results of the both time series models. As it can be expected for data with seasonal pattern the Winter's model overwhelms the Mean method. This result confirms that the Winter's model at best can forecast the seasonality. However, this is true only for a forecast with a short time horizon. As the results in chapter 4.5.2 show the increase of the time horizon makes the model inaccurate.

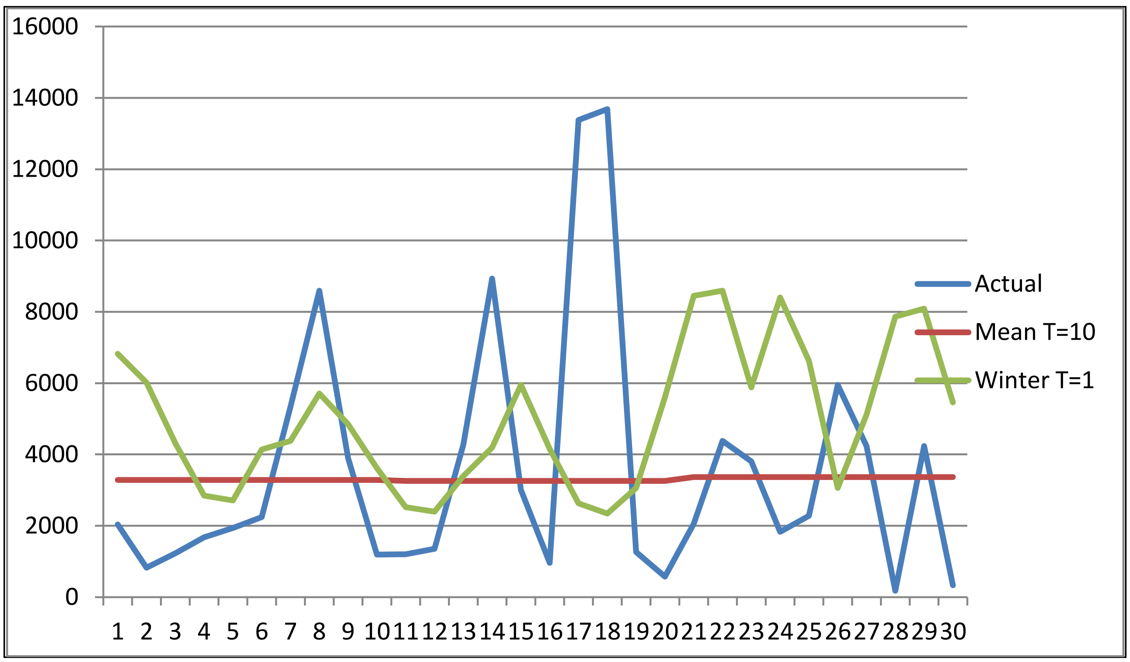


Figure 12: Comparison between the Mean method and the Winter's model

## 7.2. Causal Forecasting Models

### 7.2.1. The Simple Regression Model

Investigating the data shown in Table 2 it can be seen that the higher demand of the beer canes corresponds to the price discounts that has been made in the same periods. Considering this relationship between the demand and the price of the beer cane we will investigate whether the price has a significant relationship to the beer demand that can be used as dependent variable by forecasting the beer demand. For this purpose the simple regression model will be used with the formulas describe in chapter 3.1.

Running the regression model for the first 126 weeks we receive the following result:

| SUMMARY OUTPUT | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | |
| *Regression Statistics* | | | | | | | | |
| Multiple R | 0,896 | | | | | | | |
| R Square | 0,803 | | | | | | | |
| Adjusted R Square | 0,801 | | | | | | | |
| Standard Error | 1163,214 | | | | | | | |
| Observations | 126 | | | | | | | |
| | | | | | | | | |
| ANOVA | | | | | | | | |
| | *df* | *SS* | *MS* | *F* | *Significance F* | | | |
| Regression | 1 | 681988620 | 681988620 | 504,0319981 | 1,65404E-45 | | | |
| Residual | 124 | 167780198,9 | 1353066,12 | | | | | |
| Total | 125 | 849768818,9 | | | | | | |
| | | | | | | | | |
| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95,0%* | *Upper 95,0%* |
| Intercept | 17138,69 | 625,62 | 27,39 | 1,95447E-54 | 15900,41 | 18376,97 | 15900,41 | 18376,97 |
| Price | -25068,95 | 1116,62 | -22,45 | 1,65404E-45 | -27279,07 | -22858,84 | -27279,07 | -22858,84 |

Table 21: Result of the SLR model

At first, the measurement of the goodness of the fit of the regression line, $r^2$, will be checked. The value of 80.3% can be considered as an evidence of significant relationship between the independent variable and the dependent variable. This means that 80.3% of the beer demand is explained by the price of the beer. The correlation coefficient which is measuring the strength of the linear relationship between the demand and the price, multiple $R$, shows also that a strong linear relationship exists. Because the coefficient of the price is negative, it shows also that a decrease of the price causes an increase of the demand. Since the p-value, which for the simple regression model corresponds to the F-test, can be shown to be 1.65*10$^{-45}$ and $|t| > t_{0.001}$ we can reject the null hypothesis proving that there is strong evidence that the regression relationship is significant.

The next step is to prove whether the disturbances of the population regression function are homoscedastic or there exists heteroscedasticity. Analyzing the residual plot of the price, shown in Figure 13, it can be seen that the variance is not constant.

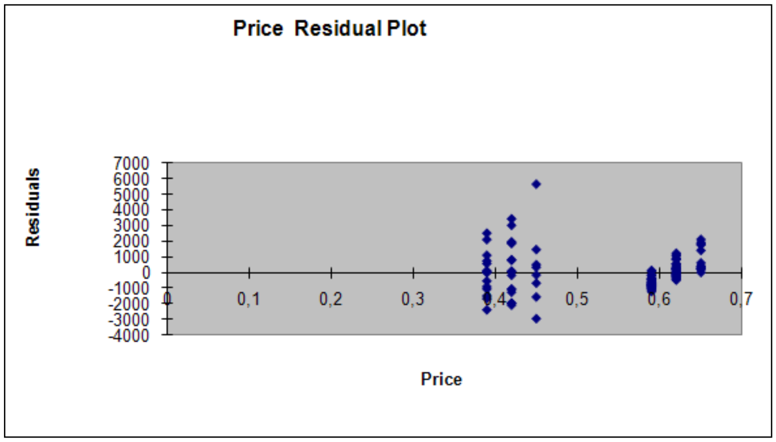


Figure 13: Price Residual Plot (SLR)

Using the Park test described in chapter 5.2.1 we receive a t-value of -3,096 for 124 degrees of freedom. This results that $|t| > t_{0.01}$ which proves the existence of heteroscedasticity.

Because of the existence of heteroscedasticity, we use the logarithmic transformation which takes the natural logarithm of each dependant and independent value.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| SUMMARY OUTPUT | | | | | | | | |
| | | | | | | | | |
| *Regression Statistics* | | | | | | | | |
| Multiple R | 0,909448978 | | | | | | | |
| R Square | 0,827097443 | | | | | | | |
| Adjusted R Square | 0,825703068 | | | | | | | |
| Standard Error | 0,288200754 | | | | | | | |
| Observations | 126 | | | | | | | |
| | | | | | | | | |
| ANOVA | | | | | | | | |
| | *df* | *SS* | *MS* | *F* | *Significance F* | | | |
| Regression | 1 | 49,26825411 | 49,26825411 | 593,1669546 | 4,35056E-49 | | | |
| Residual | 124 | 10,29939963 | 0,083059674 | | | | | |
| Total | 125 | 59,56765374 | | | | | | |
| | | | | | | | | |
| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95,0%* | *Upper 95,0%* |
| Intercept | 5,766505759 | 0,08923076 | 64,51257088 | 2,92645E-97 | 5,579893211 | 5,933118307 | 5,579893211 | 5,933118307 |
| Price (ln) | -3,417280568 | 0,140311143 | -24,35501908 | 4,35056E-49 | -3,694995432 | -3,139565704 | -3,694995432 | -3,139565704 |

Table 22: Result of the SLR model using the logarithmic transformation

As it can be seen from the regression statistics, using the logarithmic transformation and due to this compressing the scales, the independent variable explains the dependent variable better. Considering the F-test we can conclude that there is also a strong evidence that that the regression relationship is more significant.

Using equation 3.29 the demand for the 30 weeks will be forecasted.

| Week | Price | Price (ln) | Actual | Forecast (ln) | Forecast | Error | Abs. Error | MSE | MAD | % Error | MAPE | TS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0,65 | -0,43 | 2.032 | 7,23 | 1.378 | -654 | 654 | 427.313 | 654 | 67,8 | 67,8 | -1,00 |
| 2 | 0,65 | -0,43 | 824 | 7,23 | 1.378 | 554 | 554 | 367.285 | 604 | 167,3 | 117,6 | -0,16 |
| 3 | 0,65 | -0,43 | 1.227 | 7,23 | 1.378 | 151 | 151 | 252.488 | 453 | 112,3 | 115,8 | 0,11 |
| 4 | 0,65 | -0,43 | 1.673 | 7,23 | 1.378 | -295 | 295 | 211.077 | 414 | 82,4 | 107,5 | -0,59 |
| 5 | 0,65 | -0,43 | 1.937 | 7,23 | 1.378 | -559 | 559 | 231.289 | 443 | 71,2 | 100,2 | -1,81 |
| 6 | 0,65 | -0,43 | 2.243 | 7,23 | 1.378 | -865 | 865 | 317.356 | 513 | 61,4 | 93,7 | -3,25 |
| 7 | 0,45 | -0,80 | 5.338 | 8,49 | 4.843 | -495 | 495 | 307.064 | 510 | 90,7 | 93,3 | -4,23 |
| 8 | 0,45 | -0,80 | 8.591 | 8,49 | 4.843 | -3.748 | 3.748 | 2.024.887 | 915 | 56,4 | 88,7 | -6,46 |
| 9 | 0,65 | -0,43 | 3.896 | 7,23 | 1.378 | -2.518 | 2.518 | 2.504.207 | 1.093 | 35,4 | 82,8 | -7,71 |
| 10 | 0,65 | -0,43 | 1.188 | 7,23 | 1.378 | 190 | 190 | 2.257.408 | 1.003 | 116,0 | 86,1 | -8,21 |
| 11 | 0,65 | -0,43 | 1.204 | 7,23 | 1.378 | 174 | 174 | 2.054.952 | 928 | 114,5 | 88,7 | -8,69 |
| 12 | 0,65 | -0,43 | 1.350 | 7,23 | 1.378 | 28 | 28 | 1.883.772 | 853 | 102,1 | 89,8 | -9,42 |
| 13 | 0,45 | -0,80 | 4.273 | 8,49 | 4.843 | 570 | 570 | 1.763.834 | 831 | 113,3 | 91,6 | -8,98 |
| 14 | 0,45 | -0,80 | 8.932 | 8,49 | 4.843 | -4.089 | 4.089 | 2.832.293 | 1.064 | 54,2 | 88,9 | -10,86 |
| 15 | 0,65 | -0,43 | 3.004 | 7,23 | 1.378 | -1.626 | 1.626 | 2.819.665 | 1.101 | 45,9 | 86,1 | -11,97 |
| 16 | 0,65 | -0,43 | 956 | 7,23 | 1.378 | 422 | 422 | 2.654.583 | 1.059 | 144,2 | 89,7 | -12,05 |
| 17 | 0,45 | -0,80 | 13.380 | 8,49 | 4.843 | -8.537 | 8.537 | 6.785.798 | 1.499 | 36,2 | 86,5 | -14,21 |
| 18 | 0,45 | -0,80 | 13.688 | 8,49 | 4.843 | -8.845 | 8.845 | 10.755.425 | 1.907 | 35,4 | 83,7 | -15,81 |
| 19 | 0,65 | -0,43 | 1.261 | 7,23 | 1.378 | 117 | 117 | 10.190.075 | 1.813 | 109,3 | 85,1 | -16,56 |
| 20 | 0,65 | -0,43 | 568 | 7,23 | 1.378 | 810 | 810 | 9.713.401 | 1.762 | 242,7 | 92,9 | -16,57 |
| 21 | 0,65 | -0,43 | 2.050 | 7,23 | 1.378 | -672 | 672 | 9.272.342 | 1.710 | 67,2 | 91,7 | -17,47 |
| 22 | 0,65 | -0,43 | 4.380 | 7,23 | 1.378 | -3.002 | 3.002 | 9.260.425 | 1.769 | 31,5 | 89,0 | -18,59 |
| 23 | 0,65 | -0,43 | 3.803 | 7,23 | 1.378 | -2.425 | 2.425 | 9.113.412 | 1.798 | 36,2 | 86,7 | -19,64 |
| 24 | 0,65 | -0,43 | 1.831 | 7,23 | 1.378 | -453 | 453 | 8.742.225 | 1.742 | 75,3 | 86,2 | -20,53 |
| 25 | 0,65 | -0,43 | 2.282 | 7,23 | 1.378 | -904 | 904 | 8.425.203 | 1.708 | 60,4 | 85,2 | -21,47 |
| 26 | 0,45 | -0,80 | 5.956 | 8,49 | 4.843 | -1.113 | 1.113 | 8.148.826 | 1.685 | 81,3 | 85,0 | -22,42 |
| 27 | 0,45 | -0,80 | 4.234 | 8,49 | 4.843 | 609 | 609 | 7.860.741 | 1.645 | 114,4 | 86,1 | -22,59 |
| 28 | 0,65 | -0,43 | 178 | 7,23 | 1.378 | 1.200 | 1.200 | 7.631.455 | 1.629 | 774,3 | 110,7 | -22,08 |
| 29 | 0,65 | -0,43 | 4.239 | 7,23 | 1.378 | -2.861 | 2.861 | 7.650.493 | 1.672 | 32,5 | 108,0 | -23,23 |
| 30 | 0,65 | -0,43 | 329 | 7,23 | 1.378 | 1.049 | 1.049 | 7.432.178 | 1.651 | 418,9 | 118,4 | -22,88 |

Table 23: Forecast using the simple linear regression model

Although the results in Table 23 show that the simple regression model is underforecasted (TS < -6) from the 6th week on, the comparison of the forecast accuracy between Winter's model and the simple regression model (Table 24) shows that the simple regression model gives better statistical measurements than the best time series model. As Figure 14 shows, the simple regression model can predict better the demand than the Winter's model. However, the simple regression model achieves to predict the peaks in the demand but does not achieve to predict the demand during the peak. This failure explains the underforcasting performance of the model.

| Average Values | Winter's model with T=1 | Simple linear regression |
|---|---|---|
| Sum of Abs. Error | 2.675 | 1.651 |
| MAD | 2.407 | 1.214 |
| MAPE | 134 | 93 |

Table 24: Forecast accuracy Comparison between the Winter's model with T=1 and the simple linear regression model

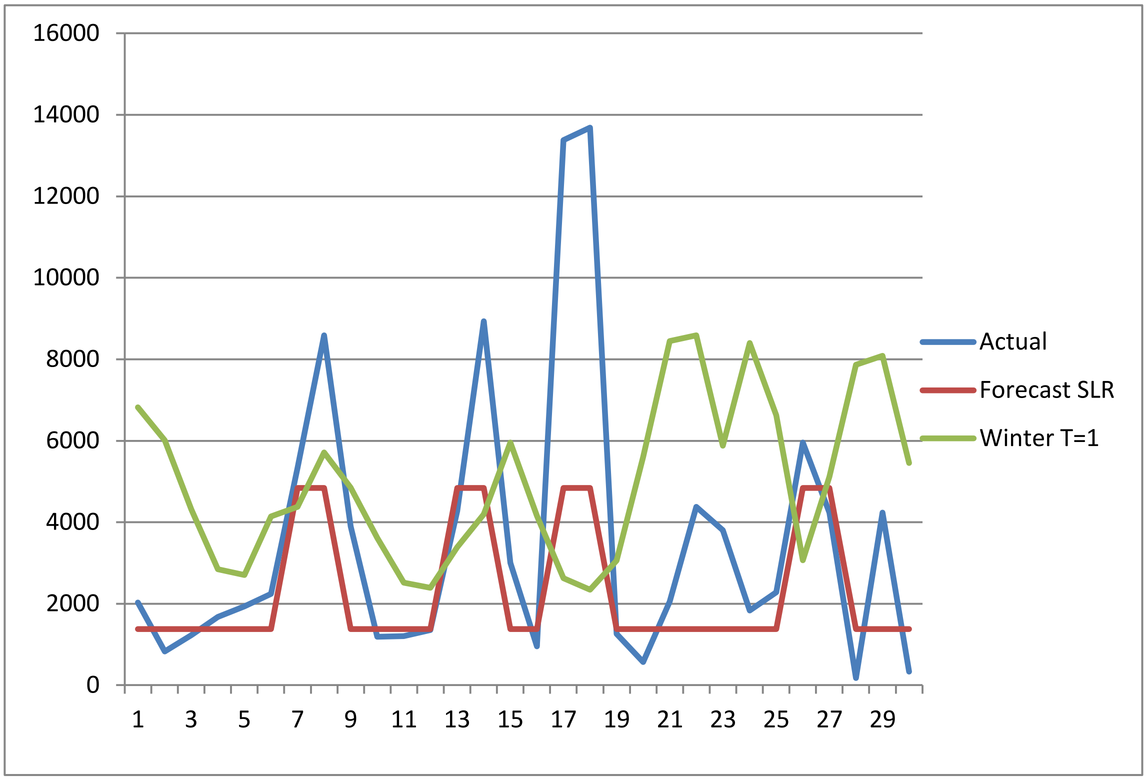

Figure 14: Forecast comparison between the simple regression model and the Winter's model

### 7.2.2. The Multiple Regression Model

In the simple regression model the price of the beer was taken as a variable for the model. Although the results in Chapter 7.2.1 show that there is a strong relationship between the demand and the price, analyzing Figure 7 it can be seen that among the peaks caused by the price there are also three cycles in the data and each of them corresponds to 52 weeks. From this analysis it can be concluded that year seasons also influence the demand of the beer. As the biggest peaks in these three cycles always correspond to the summer months we will include the maximal average weekly temperature as a second variable.

Running the regression model for the first 126 weeks we receive the following result:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| SUMMARY OUTPUT | | | | | | | | | |
| | | | | | | | | | |
| *Regression Statistics* | | | | | | | | | |
| Multiple R | 0,916 | | | | | | | | |
| R Square | 0,839 | | | | | | | | |
| Adjusted R Square | 0,836 | | | | | | | | |
| Standard Error | 1055,670 | | | | | | | | |
| Observations | 126 | | | | | | | | |
| | | | | | | | | | |
| ANOVA | | | | | | | | | |
| | *df* | *SS* | *MS* | *F* | *Significance F* | | | | |
| Regression | 2 | 712692860,4 | 356346430,2 | 319,7541817 | 1,8669E-49 | | | | |
| Residual | 123 | 137075958,4 | 1114438,686 | | | | | | |
| Total | 125 | 849768818,9 | | | | | | | |
| | | | | | | | | | |
| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95,0%* | *Upper 95,0%* | |
| Intercept | 16337,99 | 587,91 | 27,79 | 7,4419E-55 | 15174,25 | 17501,73 | 15174,25 | 17501,73 | |
| Price | -25072,54 | 1013,39 | -24,74 | 1,3895E-49 | -27078,48 | -23066,60 | -27078,48 | -23066,60 | |
| Temperature | 56,30 | 10,73 | 5,25 | 6,4879E-07 | 35,07 | 77,54 | 35,07 | 77,54 | |

Table 25: Result of the MLR model

Because we detect the heteroscedasticity using the price as a variable in the simple linear regression, the first thing that we will exanimate in the multiple linear regression is the existence of heteroscedasticity. Analyzing the residual plot of the both variables, shown in Figure 15 and Figure 16, it can be seen that the variance is again not constant for the price. From the residual plot of the temperature it could not be concluded whether the heteroscedasticity exists or not.

Using the Park test described in chapter 5.2.1 we receive a t-value of -2,427 for 124 degrees of freedom for the price and 1,588 for the temperature. This results that $|t| > t_{0.01}$ for the price which proves the existence of heteroscedasticity and $|t| < t_{0.05}$ for the temperature which reject the existence of the heteroscedasticity.

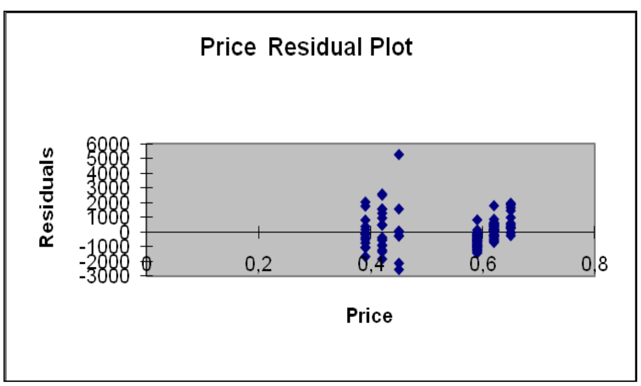


Figure 15: Price Residual Plot (MLR)

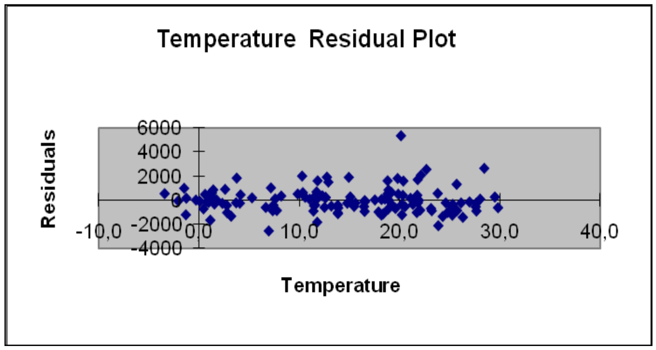


Figure 16: Temperature Residual Plot (MLR)

Because of the existence of heteroscedasticity for one of the both variables, we use the logarithmic transformation which takes the natural logarithm of each dependant and independent value and we receive the following result:

| SUMMARY OUTPUT | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Regression Statistics* | | | | | | | | |
| Multiple R | 0,929090104 | | | | | | | |
| R Square | 0,863208421 | | | | | | | |
| Adjusted R Square | 0,860984167 | | | | | | | |
| Standard Error | 0,25738449 | | | | | | | |
| Observations | 126 | | | | | | | |
| | | | | | | | | |
| ANOVA | | | | | | | | |
| | *df* | *SS* | *MS* | *F* | *Significance F* | | | |
| Regression | 2 | 51,41930031 | 25,70965015 | 388,0890777 | 7,37311E-54 | | | |
| Residual | 123 | 8,148353433 | 0,066246776 | | | | | |
| Total | 125 | 59,56765374 | | | | | | |
| | | | | | | | | |
| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95,0%* | *Upper 95,0%* |
| Intercept | 4,017491099 | 0,315416039 | 12,73711733 | 3,23824E-24 | 3,393144633 | 4,641837565 | 3,393144633 | 4,641837565 |
| Price | -3,425248142 | 0,125315981 | -27,33289166 | 4,31212E-54 | -3,673303347 | -3,177192938 | -3,673303347 | -3,177192938 |
| Temperature F | 0,432241899 | 0,075855027 | 5,698263066 | 8,4479E-08 | 0,282091583 | 0,582392216 | 0,282091583 | 0,582392216 |

Table 26: Result of the MLR model using the logarithmic transformation

Comparing the regression statistics of the MLR with the SLR, it can be seen that the temperature as a second variable improves the measurement of the goodness of the fit of a regression line, which means that the both independent variables together explain better the dependent variable. It can be also seen that there is a negligibility difference between the *R Square* and the *Adjusted R Square*, which shows that there is no overestimation of the model.

Considering the F-test we have even stronger evidence that the model is significant since $\hat{F}(model) = 388.089 > \hat{F}_{0.01} = 4.61$, based on 2 numerator and 123 denominator degrees of freedom. Additionally, as the p-value of the price and the p-value of the temperature are smaller than 0.001 we can reject the null hypothesis proving that there is strong evidence that the regression relationship is significant for both variables.

Using equation 3.48 the demand for the 30 weeks will be forecasted.

| Week | Price(ln) | Temp. F (ln) | Actual | Forecast (ln) | Forecast | Error | Abs. Error | MSE | MAD | % Error | MAPE | TS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | -0,43078 | 4,353 | 2032 | 7,37477014 | 1595 | -437 | 437 | 190772,3 | 436,7749 | 21,49 | 21,49 | -1,0 |
| 2 | -0,43078 | 4,352 | 824 | 7,37434103 | 1595 | 771 | 771 | 392252,7 | 603,6578 | 93,51 | 57,50 | 0,6 |
| 3 | -0,43078 | 4,355 | 1227 | 7,37562707 | 1597 | 370 | 370 | 307034,7 | 525,6361 | 30,12 | 48,38 | 1,3 |
| 4 | -0,43078 | 4,299 | 1673 | 7,35126851 | 1558 | -115 | 115 | 233572,4 | 422,9341 | 6,86 | 38,00 | 1,4 |
| 5 | -0,43078 | 4,266 | 1937 | 7,33684092 | 1536 | -401 | 401 | 219041,8 | 418,5768 | 20,71 | 34,54 | 0,4 |
| 6 | -0,43078 | 4,311 | 2243 | 7,35651979 | 1566 | -677 | 677 | 258838,2 | 461,5847 | 30,17 | 33,81 | -1,1 |
| 7 | -0,79851 | 4,468 | 5338 | 8,68396571 | 5907 | 569 | 569 | 268182,4 | 476,9908 | 10,67 | 30,51 | 0,2 |
| 8 | -0,79851 | 4,376 | 8591 | 8,64421211 | 5677 | -2914 | 2914 | 1295944 | 781,5929 | 33,92 | 30,93 | -3,6 |
| 9 | -0,43078 | 4,275 | 3896 | 7,34087994 | 1542 | -2354 | 2354 | 1767616 | 956,2972 | 60,42 | 34,21 | -5,4 |
| 10 | -0,43078 | 4,383 | 1188 | 7,38758787 | 1616 | 428 | 428 | 1609156 | 903,4479 | 36,01 | 34,39 | -5,3 |
| 11 | -0,43078 | 4,288 | 1204 | 7,34641086 | 1551 | 347 | 347 | 1473792 | 852,8272 | 28,79 | 33,88 | -5,2 |
| 12 | -0,43078 | 4,312 | 1350 | 7,35681795 | 1567 | 217 | 217 | 1354894 | 799,8285 | 16,06 | 32,39 | -5,2 |
| 13 | -0,79851 | 4,230 | 4273 | 8,58097875 | 5329 | 1056 | 1056 | 1336503 | 819,5586 | 24,72 | 31,80 | -3,8 |
| 14 | -0,79851 | 4,051 | 8932 | 8,50360704 | 4933 | -3999 | 3999 | 2383594 | 1046,695 | 44,78 | 32,73 | -6,8 |
| 15 | -0,43078 | 4,202 | 3004 | 7,30945883 | 1494 | -1510 | 1510 | 2376620 | 1077,558 | 50,25 | 33,90 | -8,0 |
| 16 | -0,43078 | 4,216 | 956 | 7,31523988 | 1503 | 547 | 547 | 2246784 | 1044,4 | 57,22 | 35,36 | -7,8 |
| 17 | -0,79851 | 4,115 | 13380 | 8,53132862 | 5071 | -8309 | 8309 | 6175591 | 1471,719 | 62,10 | 36,93 | -11,2 |
| 18 | -0,79851 | 4,197 | 13688 | 8,56684035 | 5255 | -8433 | 8433 | 9783831 | 1858,484 | 61,61 | 38,30 | -13,4 |
| 19 | -0,43078 | 4,064 | 1261 | 7,24982315 | 1408 | 147 | 147 | 9270027 | 1768,399 | 11,65 | 36,90 | -14,0 |
| 20 | -0,43078 | 3,988 | 568 | 7,21670919 | 1362 | 794 | 794 | 8838048 | 1719,679 | 139,79 | 42,04 | -13,9 |
| 21 | -0,43078 | 3,800 | 2050 | 7,13556487 | 1256 | -794 | 794 | 8447221 | 1675,606 | 38,74 | 41,89 | -14,7 |
| 22 | -0,43078 | 3,947 | 4380 | 7,19904227 | 1338 | -3042 | 3042 | 8483841 | 1737,708 | 69,45 | 43,14 | -16,0 |
| 23 | -0,43078 | 3,856 | 3803 | 7,15973977 | 1287 | -2516 | 2516 | 8390299 | 1771,565 | 66,17 | 44,14 | -17,1 |
| 24 | -0,43078 | 3,550 | 1831 | 7,02735833 | 1127 | -704 | 704 | 8061351 | 1727,081 | 38,45 | 43,90 | -17,9 |
| 25 | -0,43078 | 3,689 | 2282 | 7,08776467 | 1197 | -1085 | 1085 | 7785966 | 1701,389 | 47,54 | 44,05 | -18,8 |
| 26 | -0,79851 | 3,727 | 5956 | 8,36366281 | 4288 | -1668 | 1668 | 7593467 | 1700,09 | 28,00 | 43,43 | -19,8 |
| 27 | -0,79851 | 3,825 | 4234 | 8,40591737 | 4473 | 239 | 239 | 7314351 | 1645,993 | 5,66 | 42,03 | -20,3 |
| 28 | -0,43078 | 3,566 | 178 | 7,03432785 | 1135 | 957 | 957 | 7085828 | 1621,384 | 537,60 | 59,73 | -20,1 |
| 29 | -0,43078 | 3,350 | 4239 | 6,94104172 | 1034 | -3205 | 3205 | 7195731 | 1675,996 | 75,61 | 60,28 | -21,3 |
| 30 | -0,43078 | 3,466 | 329 | 6,99106575 | 1087 | 758 | 758 | 6975019 | 1645,392 | 230,36 | 65,95 | -21,2 |

Table 27: Forecast using the multiple linear regression model

The results in Table 27 show that the multiple regression model is underforecasted (TS < -6) from the 13$^{th}$ week on which is improvement compared to the forecasting performance of the simple linear regression. The comparison of the forecast accuracy between multiple regression model and simple regression model (Table 28) shows that including the temperature as a second variable we improve the statistical measurements of the regression model.

| Average Values | Multiple linear regression | Simple linear regression |
|---|---|---|
| Sum of Abs. Error | 1.654 | 1.651 |
| MAD | 1.178 | 1.214 |
| MAPE | 40 | 93 |

Table 28: Forecast accuracy comparison between the multiple regression model and the simple linear regression model

Observing Figure 17 we can see that having better explanation of the variance and better statistical performance the multiple regression model achieves to predict the peaks in the demand better than the simple regression model only till week 21. For the last 9 weeks of the forecasted periods the simple regression model forecasted better the demand of beer. The underperformance of the multiple regression model can be explained with the coefficient of the temperature. The coefficient of the temperature has a positive value which means that an increase in the temperature will increase the demand of the beer or a decrease in the temperature will decrease the demand of the beer. On the other hand the coefficient of the price has a negative value which means that decrease of the price will increase the demand of the beer. Observing the last 9 weeks of the forecasted period which correspond to the last 9 weeks of the year we are facing the problem that the second variable forecasts lower demand as the temperatures are falling down in the late autumn and early winter. At the same time there are 5 holidays in Austria (Christmas and the New Eve are among them) in the last 4 weeks of the forecasted period which may influence the demand of beer. These two factors can explain the underperformance of the multiple regression model and should be investigated in another research.

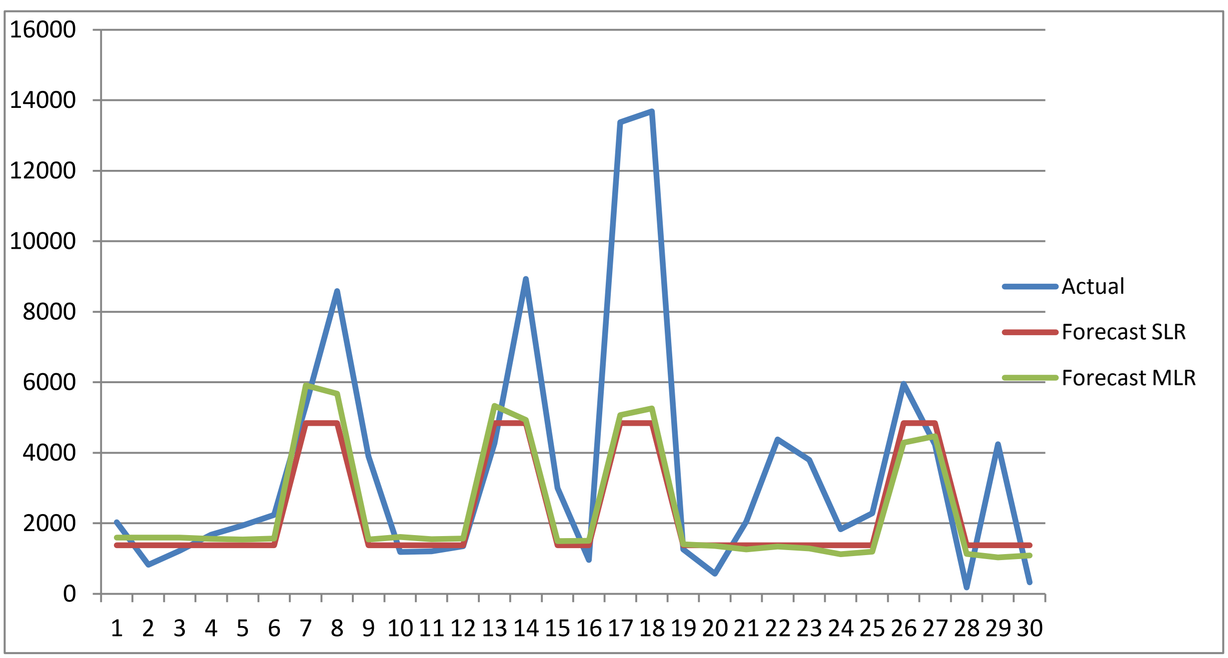


Figure 17: Forecast comparison between the simple and multiple regression models

# 8. Safety Stock Planning under Demand Forecasting and Positive Lead Time

In this chapter the safety stock planning caused by the forecasting results of the Winter's model with T=1 and the multiple regression model will be compared using the ($\hat{t}$, $\hat{S}$) inventory policy with $\hat{t}$= 3, 5 weeks and lead time (LT) = 0, 3, 5 weeks with a given Order Service Level (OSL) and Unit Service Level (USL). For the inventory policy the lost sales case will be considered.

In addition, the presence of autocorrelation in the data will be examined and the remedial measure will be taken if needed.

## 8.1. Examination of Autocorrelation

As the accurate variance is the basis for reliable safety stock planning the data needs to be examined for the presence of autocorrelation. Using equations (6.1) and (6.2) the autocorrelation coefficients and partial autocorrelation coefficients will be calculated and with their help the grade of the autocorrelation process will be determined. Using the Minitab program we compute the autocorrelation coefficients and partial autocorrelation coefficients. The results are plotted on Figure 18 and Figure 19.

The autocorrelation coefficients do not lie within the 95 percent upper and lower confidence limits and therefore it cannot be confirmed that the data are random. It can be seen from the plotted partial autocorrelation coefficients that the third partial autocorrelation coefficient is not significantly different from zero which is evidence of autoregressive process of second grade (AR2).

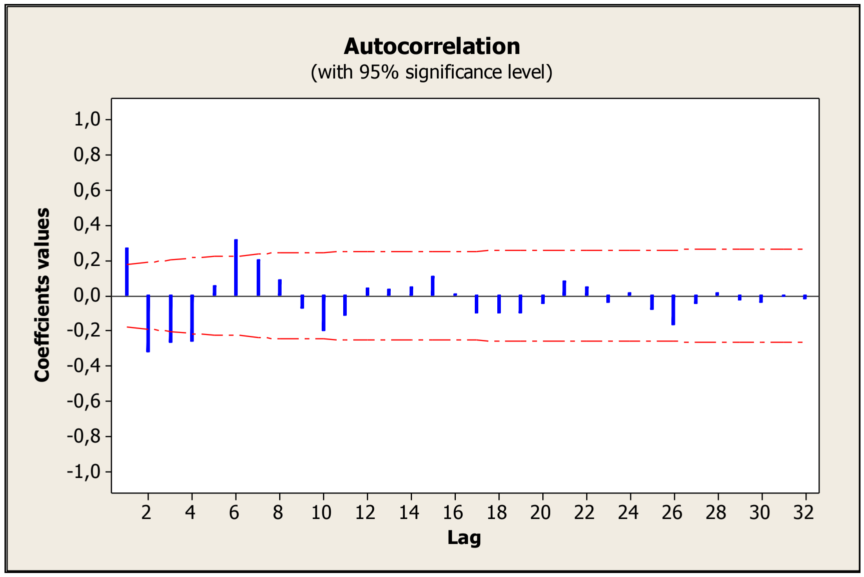


Figure 18: Autocorrelation coefficients

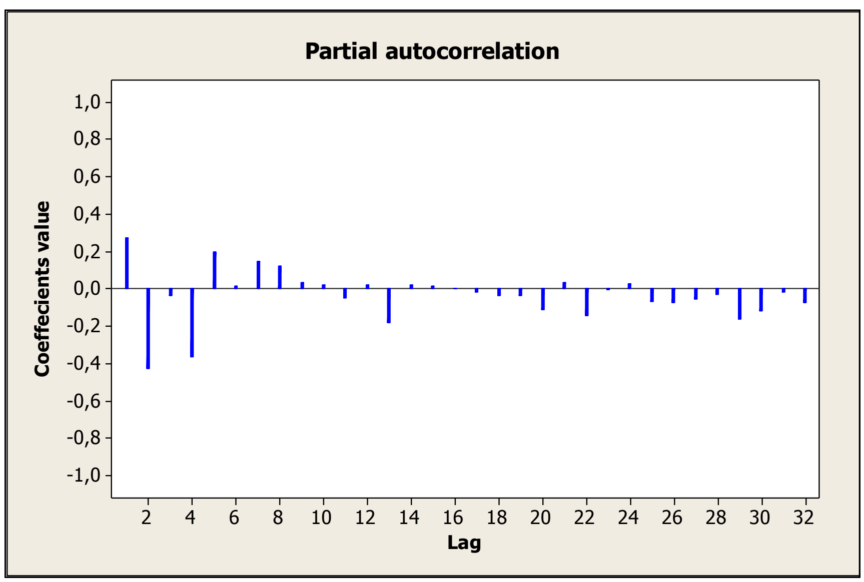


Figure 19: Partial autocorrelation coefficients

## 8.2. Safety Stock Planning with a given OSL

An OSL of 95% and 85% will be used for the safety stock planning for both models. After receiving the forecasted weeks, we can find the standard deviation for each model using equation 2.26. The standard deviation (or sigma) that will be used for the safety stock planning is the one from period 6. We can assume that in this period the deviation should be closed to the true one as it is stabilized and the deviation of the forecast from the actual demand is acceptable, as shown in the previous chapter.

| | | Winter's model | | | MLR | | |
|---|---|---|---|---|---|---|---|
| Week | Actual | Forecast | MAD | Sigma | Forecast | MAD | Sigma |
| 1 | 2.032 | 5.128 | 3.096 | 3.870 | 1.595 | 437 | 546 |
| 2 | 824 | 4.398 | 3.335 | 4.169 | 1.595 | 604 | 755 |
| 3 | 1.227 | 3.148 | 2.864 | 3.580 | 1.597 | 526 | 657 |
| 4 | 1.673 | 2.385 | 2.326 | 2.907 | 1.558 | 423 | 529 |
| 5 | 1.937 | 2.259 | 1.925 | 2.406 | 1.536 | 419 | 523 |
| 6 | 2.243 | 3.404 | 1.798 | 2.247 | 1.566 | 462 | 577 |
| 7 | 5.338 | 3.696 | 1.775 | 2.219 | 5.907 | 477 | 596 |
| 8 | 8.591 | 4.454 | 2.071 | 2.588 | 5.677 | 782 | 977 |
| 9 | 3.896 | 3.620 | 1.871 | 2.339 | 1.542 | 956 | 1.195 |
| 10 | 1.188 | 2.730 | 1.838 | 2.298 | 1.616 | 903 | 1.129 |
| 11 | 1.204 | 2.230 | 1.764 | 2.205 | 1.551 | 853 | 1.066 |
| 12 | 1.350 | 2.188 | 1.687 | 2.109 | 1.567 | 800 | 1.000 |
| 13 | 4.273 | 3.151 | 1.644 | 2.055 | 5.329 | 820 | 1.024 |
| 14 | 8.932 | 4.052 | 1.875 | 2.344 | 4.933 | 1.047 | 1.308 |
| 15 | 3.004 | 5.352 | 1.906 | 2.383 | 1.494 | 1.078 | 1.347 |
| 16 | 956 | 3.680 | 1.957 | 2.447 | 1.503 | 1.044 | 1.305 |
| 17 | 13.380 | 2.394 | 2.489 | 3.111 | 5.071 | 1.472 | 1.840 |
| 18 | 13.688 | 2.006 | 2.999 | 3.749 | 5.255 | 1.858 | 2.323 |
| 19 | 1.261 | 2.006 | 2.881 | 3.601 | 1.408 | 1.768 | 2.210 |
| 20 | 568 | 3.394 | 2.878 | 3.597 | 1.362 | 1.720 | 2.150 |
| 21 | 2.050 | 5.112 | 2.887 | 3.608 | 1.256 | 1.676 | 2.095 |
| 22 | 4.380 | 4.841 | 2.776 | 3.471 | 1.338 | 1.738 | 2.172 |
| 23 | 3.803 | 3.087 | 2.687 | 3.359 | 1.287 | 1.772 | 2.214 |
| 24 | 1.831 | 4.782 | 2.698 | 3.372 | 1.127 | 1.727 | 2.159 |
| 25 | 2.282 | 4.545 | 2.680 | 3.351 | 1.197 | 1.701 | 2.127 |
| 26 | 5.956 | 1.844 | 2.736 | 3.419 | 4.288 | 1.700 | 2.125 |
| 27 | 4.234 | 2.780 | 2.688 | 3.360 | 4.473 | 1.646 | 2.057 |
| 28 | 178 | 4.446 | 2.745 | 3.431 | 1.135 | 1.621 | 2.027 |
| 29 | 4.239 | 4.740 | 2.667 | 3.334 | 1.034 | 1.676 | 2.095 |
| 30 | 329 | 3.242 | 2.675 | 3.344 | 1.087 | 1.645 | 2.057 |

Table 29: Standard deviation for the Winter's model and MLR

Because of the presence of autocorrelation the variance is corrected using equation (6.3). In the equation we will consider the first two autocorrelation coefficients as we have AR(2) process with the T = 30 as we are forecasting 30 periods.

The adjusted standard deviation equals 11.925 for the Winter's model and 2.956 for the MLR model.

After defining the standard deviation we use equations 4.2 and 4.3 with the $\hat{k}$-values from Table 1 for OSL 95% and 85% in order to find the safety stock for the ($\hat{t}$, $\hat{S}$) inventory policy with $\hat{t}$= 3, 5 weeks and lead time (LT) = 0, 3, 5 weeks. Using the ($\hat{t}$, $\hat{S}$) inventory policy as described in Chapter 4.5.1 and 4.5.2, we receive the inventory levels for 95% OSL shown in Table 30 and for 85% OSL shown in Table 31.

| Winter's model | | | | | | MRL model | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| OSL 95% | | | | | | | | | | | |
| f=3 | | | f=5 | | | f=3 | | | f=5 | | |
| LT=0 | LT=3 | LT=5 | LT=0 | LT=3 | LT=5 | LT=0 | LT=3 | LT=5 | LT=0 | LT=3 | LT=5 |
| 44.616 | 66.736 | 82.319 | 59.146 | 82.319 | 95.218 | 11.176 | 19.325 | 32.752 | 16.721 | 32.752 | 37.533 |
| 43.792 | 65.912 | 81.495 | 58.322 | 81.495 | 94.394 | 10.352 | 18.501 | 31.928 | 15.897 | 31.928 | 36.709 |
| 42.565 | 64.685 | 80.268 | 57.095 | 80.268 | 93.167 | 9.125 | 17.274 | 30.701 | 14.670 | 30.701 | 35.482 |
| 40.892 | 63.012 | 78.595 | 55.422 | 78.595 | 91.494 | 11.409 | 15.601 | 29.028 | 12.997 | 29.028 | 33.809 |
| 38.955 | 61.075 | 76.658 | 53.485 | 76.658 | 89.557 | 9.472 | 13.664 | 27.091 | 11.060 | 27.091 | 31.872 |
| 36.712 | 58.832 | 74.415 | 59.522 | 74.415 | 87.314 | 7.229 | 11.421 | 24.848 | 24.938 | 24.848 | 29.629 |
| 40.407 | 56.674 | 69.077 | 54.184 | 69.077 | 81.976 | 16.210 | 18.506 | 19.510 | 19.600 | 19.510 | 24.291 |
| 31.816 | 48.083 | 60.486 | 45.593 | 60.486 | 73.385 | 7.619 | 9.915 | 10.919 | 11.009 | 10.919 | 15.700 |
| 27.920 | 44.187 | 56.590 | 41.697 | 60.885 | 69.489 | 3.723 | 6.019 | 109 | 7.113 | 7.521 | 11.804 |
| 39.934 | 51.132 | 55.402 | 40.509 | 59.697 | 68.301 | 11.967 | 17.161 | -1.079 | 5.925 | 6.333 | 10.616 |
| 38.730 | 49.928 | 54.198 | 59.631 | 58.493 | 74.447 | 10.763 | 15.957 | -2.283 | 24.542 | 5.129 | 13.482 |
| 37.380 | 48.578 | 60.035 | 58.281 | 57.143 | 73.097 | 9.413 | 14.607 | 15.676 | 23.192 | 3.779 | 12.132 |
| 42.257 | 67.869 | 55.762 | 54.008 | 52.870 | 68.824 | 15.905 | 38.734 | 11.403 | 18.919 | -494 | 7.859 |
| 33.325 | 58.937 | 46.830 | 45.076 | 64.775 | 59.892 | 6.973 | 29.802 | 2.471 | 9.987 | 25.189 | -1.073 |
| 30.321 | 55.933 | 68.493 | 42.072 | 61.771 | 56.888 | 3.969 | 26.798 | 36.174 | 6.983 | 22.185 | -4.077 |
| 41.098 | 75.082 | 67.537 | 56.385 | 60.815 | 80.114 | 19.294 | 46.729 | 35.218 | 24.515 | 21.229 | 43.892 |
| 27.718 | 61.702 | 54.157 | 43.005 | 47.435 | 66.734 | 5.914 | 33.349 | 21.838 | 11.135 | 7.849 | 30.512 |
| 14.030 | 48.014 | 61.950 | 29.317 | 33.747 | 53.046 | -7.774 | 19.661 | 32.581 | -2.553 | -5.839 | 16.824 |
| 43.226 | 57.459 | 60.689 | 28.056 | 18.968 | 51.785 | 11.186 | 19.367 | 31.320 | -3.814 | 8.785 | 15.563 |
| 42.658 | 56.891 | 60.121 | 27.488 | 18.400 | 51.217 | 10.618 | 18.799 | 30.752 | -4.382 | 8.217 | 14.995 |
| 40.608 | 54.841 | 71.579 | 64.178 | 16.350 | 90.155 | 8.568 | 16.749 | 28.702 | 15.027 | 6.167 | 49.124 |
| 42.304 | 73.717 | 67.199 | 59.798 | 11.970 | 85.775 | 7.793 | -4.354 | 24.322 | 10.647 | 1.787 | 44.744 |
| 38.501 | 69.914 | 63.396 | 55.995 | 8.167 | 81.972 | 3.990 | -8.157 | 20.519 | 6.844 | -2.016 | 40.941 |
| 36.670 | 68.083 | 84.705 | 54.164 | 74.852 | 80.141 | 2.159 | -9.988 | -1.831 | 5.013 | 19.806 | 39.110 |
| 40.860 | 80.884 | 82.423 | 51.882 | 72.570 | 77.859 | 16.099 | 23.339 | -4.113 | 2.731 | 17.524 | 36.828 |
| 34.904 | 74.928 | 76.467 | 54.955 | 66.614 | 122.133 | 10.143 | 17.383 | -10.069 | 16.934 | 11.568 | 49.475 |
| 30.670 | 70.694 | 87.198 | 50.721 | 62.380 | 117.899 | 5.909 | 13.149 | -3.304 | 12.700 | 7.334 | 45.241 |
| 46.224 | 85.318 | 87.020 | 50.543 | 62.202 | 117.721 | 11.499 | 38.095 | -3.482 | 12.522 | 7.156 | 45.063 |
| 41.985 | 81.079 | 82.781 | 46.304 | 57.963 | 113.482 | 7.260 | 33.856 | -7.721 | 8.283 | 4.007 | 40.824 |
| 41.656 | 80.750 | 82.452 | 45.975 | 57.634 | 113.153 | 6.931 | 33.527 | 20.748 | 7.954 | 3.678 | 40.495 |

Table 30: Inventory levels for 95% OSL using ($\hat{\boldsymbol{t}}$, $\widehat{\boldsymbol{S}}$) inventory policy with $\hat{\boldsymbol{t}}$= 3, 5 and LT = 0, 3, 5 weeks

The inventory levels for the Winter's model are plotted on Figure 20. As it can be seen, the increase of the periodic review and the increase of the lead time lead to increase of the inventory levels. As it can be seen with the ($\hat{t}$, $\hat{S}$) inventory policy and the adjusted variance we can manage the availability by building up high inventory levels. With the increase of the periodic review and the lead time we increase the inventory levels. The question in this case is what would be the better off situation for the company in this case – to keep the high inventory levels in order to meet the demand or to have optimized inventory and to be prepared for lost sales.

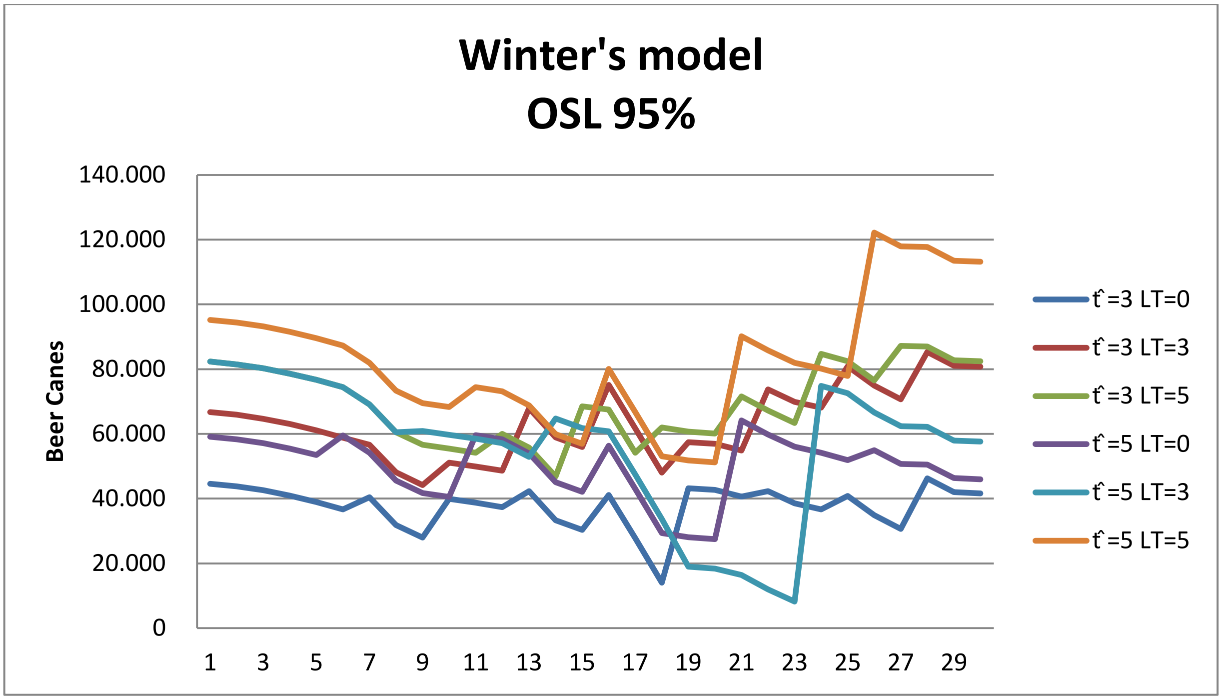


Figure 20: ($\hat{\boldsymbol{t}}$, $\hat{\boldsymbol{S}}$) inventory policy with $\hat{\boldsymbol{t}}$=3, 5 and LT= 0, 3, 5 for the Winter's model with OSL 95%

The inventory levels for the MLR model are plotted on Figure 21. The tendency that the increase of the periodic review and the increase of the lead time lead to increase of the inventory levels can be seen also for the ($\hat{t}$, $\hat{S}$) inventory policy using the forecasted results of the MLR model. It can be also seen that due to the fact that the MLR model has smaller standard deviation the inventory levels are significantly lower in comparison of the inventory levels of the Winter's model. However, as shown on the Figure 17 the MLR model does not achieve to predict the exact demand during the peaks. This fail of the model together with the lower inventory levels lead to more stockouts in comparison with the inventory levels of the Winter's model.

Which of the both models is better for the company's profit must be investigated in another research where the inventory costs and the sales profit must be considered.

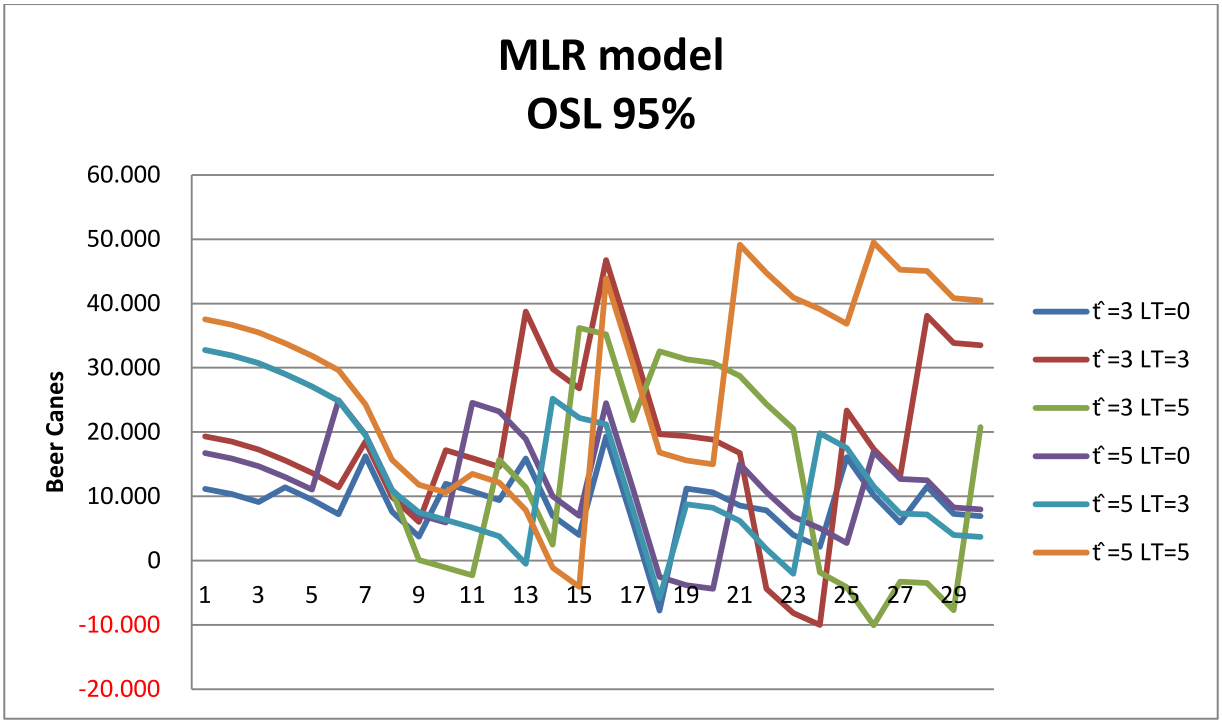


Figure 21: ($\hat{\boldsymbol{t}}$, $\hat{\boldsymbol{S}}$) inventory policy with $\hat{\boldsymbol{t}}$=3, 5 and LT= 0, 3, 5 for the MLR model with OSL 95%

The results of reducing the OSL from 95% to 85% are shown in Table 31. The invenotry levels for Winter's model and MLR using ($\hat{t}$, $\hat{S}$) inventory policy are plotted on Figure 22 and 23. It can be seen that the tendency that was observed for OSL 95% remeins the same and for OSL 85% - the increase of the periodic review and the increase of the lead time lead to increase of the inventory levels. Due to the lower standard deviation of the MLR model the inventory levels are significant lower compared to the inventory levels of the Winter's model with T=1 which again leads to more stockouts in the ($\hat{t}$, $\hat{S}$)inventory policy using the MLR model. Reducing the OSL we reduce also the safety stock factor which downsizes the product availability in the presence of demand variability. As a result more stockouts in comparison with the 95% OSL are observed.

| Winter's model | | | | | | MRL model | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| OSL 85% | | | | | | | | | | | |
| $\hat{t}$=3 | | | $\hat{t}$=5 | | | $\hat{t}$=3 | | | $\hat{t}$=5 | | |
| LT=0 | LT=3 | LT=5 | LT=0 | LT=3 | LT=5 | LT=0 | LT=3 | LT=5 | LT=0 | LT=3 | LT=5 |
| 32.049 | 48.964 | 61.798 | 42.922 | 61.798 | 72.274 | 8.061 | 14.919 | 27.665 | 12.699 | 27.665 | 31.846 |
| 31.225 | 48.140 | 60.974 | 42.098 | 60.974 | 71.450 | 7.237 | 14.095 | 26.841 | 11.875 | 26.841 | 31.022 |
| 29.998 | 46.913 | 59.747 | 40.871 | 59.747 | 70.223 | 6.010 | 12.868 | 25.614 | 10.648 | 25.614 | 29.795 |
| 28.325 | 45.240 | 58.074 | 39.198 | 58.074 | 68.550 | 8.294 | 11.195 | 23.941 | 8.975 | 23.941 | 28.122 |
| 26.388 | 43.303 | 56.137 | 37.261 | 56.137 | 66.613 | 6.357 | 9.258 | 22.004 | 7.038 | 22.004 | 26.185 |
| 24.145 | 41.060 | 53.894 | 43.298 | 53.894 | 64.370 | 4.114 | 7.015 | 19.761 | 20.917 | 19.761 | 23.942 |
| 27.840 | 38.902 | 48.556 | 37.960 | 48.556 | 59.032 | 13.095 | 14.101 | 14.423 | 15.579 | 14.423 | 18.604 |
| 19.249 | 30.311 | 39.965 | 29.369 | 39.965 | 50.441 | 4.504 | 5.510 | 5.832 | 6.988 | 5.832 | 10.013 |
| 15.353 | 26.415 | 36.069 | 25.473 | 40.364 | 46.545 | 608 | 1.614 | 109 | 3.092 | 7.521 | 6.117 |
| 27.367 | 33.360 | 34.881 | 24.285 | 39.176 | 45.357 | 8.852 | 17.161 | -1.079 | 1.904 | 6.333 | 4.929 |
| 26.163 | 32.156 | 33.677 | 43.407 | 37.972 | 51.503 | 7.648 | 15.957 | -2.283 | 20.520 | 5.129 | 13.482 |
| 24.813 | 30.806 | 39.513 | 42.057 | 36.622 | 50.153 | 6.298 | 14.607 | 15.676 | 19.170 | 3.779 | 12.132 |
| 29.690 | 50.097 | 35.240 | 37.784 | 32.349 | 45.880 | 12.790 | 34.328 | 11.403 | 14.897 | -494 | 7.859 |
| 20.758 | 41.165 | 26.308 | 28.852 | 44.254 | 36.948 | 3.858 | 25.396 | 2.471 | 5.965 | 20.103 | -1.073 |
| 17.754 | 38.161 | 21.663 | 25.848 | 41.250 | 33.944 | 854 | 22.392 | 31.087 | 2.961 | 17.099 | -4.077 |
| 28.532 | 57.310 | 20.707 | 40.161 | 40.294 | 57.170 | 16.179 | 37.918 | 30.131 | 20.493 | 16.143 | 38.205 |
| 15.152 | 43.930 | 7.327 | 26.781 | 26.914 | 43.790 | 2.799 | 24.538 | 16.751 | 7.113 | 2.763 | 24.825 |
| 1.464 | 30.242 | 15.121 | 13.093 | 13.226 | 30.102 | -10.889 | 10.850 | 22.407 | -6.575 | -10.925 | 11.137 |
| 30.659 | 39.687 | 13.860 | 11.832 | 18.968 | 28.841 | 8.071 | -294 | 21.146 | -7.836 | 8.785 | 9.876 |
| 30.091 | 39.119 | 13.292 | 11.264 | 18.400 | 28.273 | 7.503 | -862 | 20.578 | -8.404 | 8.217 | 9.308 |
| 28.041 | 37.069 | 51.058 | 47.954 | 16.350 | 38.938 | 5.453 | -2.912 | 18.528 | 11.005 | 6.167 | 37.749 |
| 29.737 | 55.945 | 46.678 | 43.574 | 11.970 | 34.558 | 4.678 | 10.902 | 14.148 | 6.625 | 1.787 | 33.369 |
| 25.934 | 52.142 | 42.875 | 39.771 | 8.167 | 30.755 | 875 | 7.099 | 10.345 | 2.822 | -2.016 | 29.566 |
| 24.103 | 50.311 | 90.492 | 37.940 | 54.330 | 28.924 | -956 | 5.268 | -1.831 | 991 | 14.719 | 27.735 |
| 28.294 | 63.112 | 88.210 | 35.658 | 52.048 | 26.642 | 12.984 | 18.933 | -4.113 | -1.291 | 12.437 | 25.453 |
| 22.338 | 57.156 | 82.254 | 38.732 | 46.092 | 99.189 | 7.028 | 12.977 | -10.069 | 12.912 | 6.481 | 38.100 |
| 18.104 | 52.922 | 92.985 | 34.498 | 41.858 | 94.955 | 2.794 | 8.743 | 1.783 | 8.678 | 2.247 | 33.866 |
| 33.657 | 67.546 | 92.807 | 34.320 | 41.680 | 94.777 | 8.384 | 29.285 | 1.605 | 8.500 | 2.069 | 33.688 |
| 29.418 | 63.307 | 88.568 | 30.081 | 37.441 | 90.538 | 4.145 | 25.046 | -2.634 | 4.261 | 4.007 | 29.449 |
| 29.089 | 62.978 | 88.239 | 29.752 | 37.112 | 90.209 | 3.816 | 24.717 | 20.748 | 3.932 | 3.678 | 29.120 |

Table 31: Inventory levels for 85% OSL using ($\hat{\boldsymbol{t}}$, $\widehat{\boldsymbol{S}}$) inventory policy with $\hat{\boldsymbol{t}}$= 3, 5 and LT = 0, 3, 5 weeks

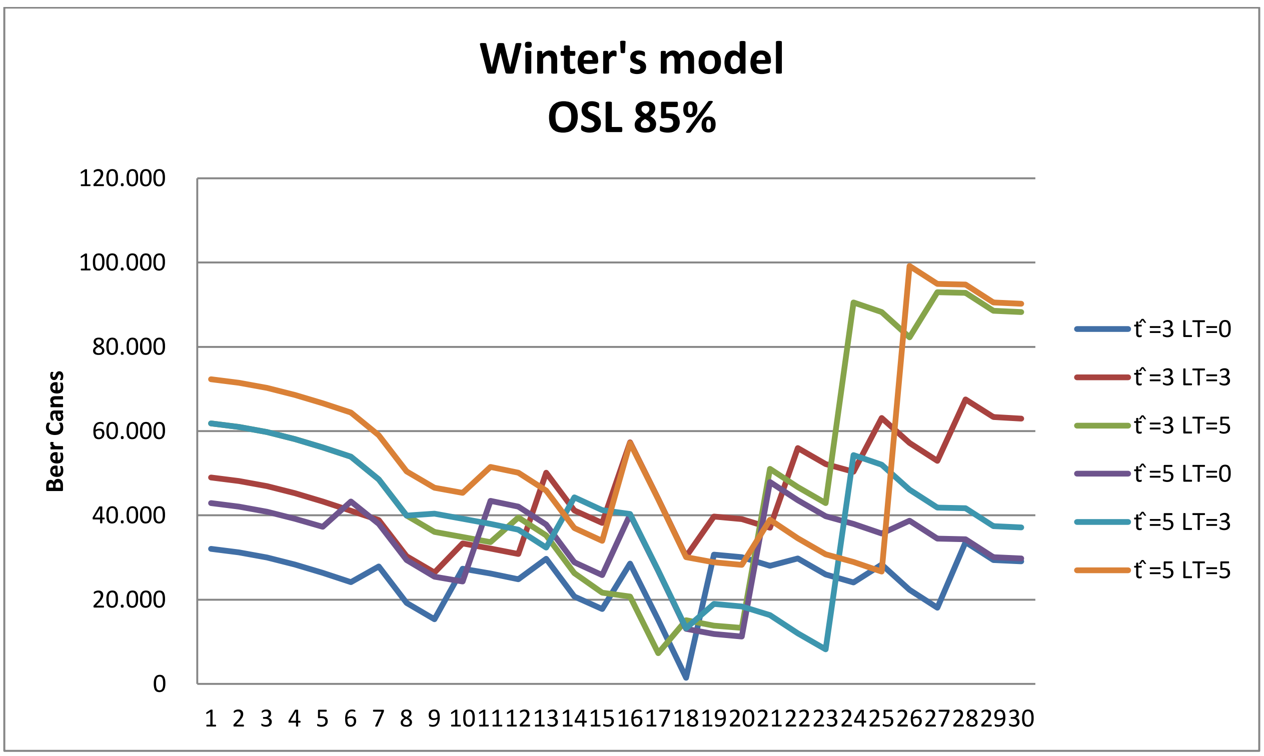


Figure 22: ($\hat{\boldsymbol{t}}$, $\widehat{\boldsymbol{S}}$) inventory policy with $\hat{\boldsymbol{t}}$=3, 5 and LT= 0, 3, 5 for the Winter's model with OSL 85%

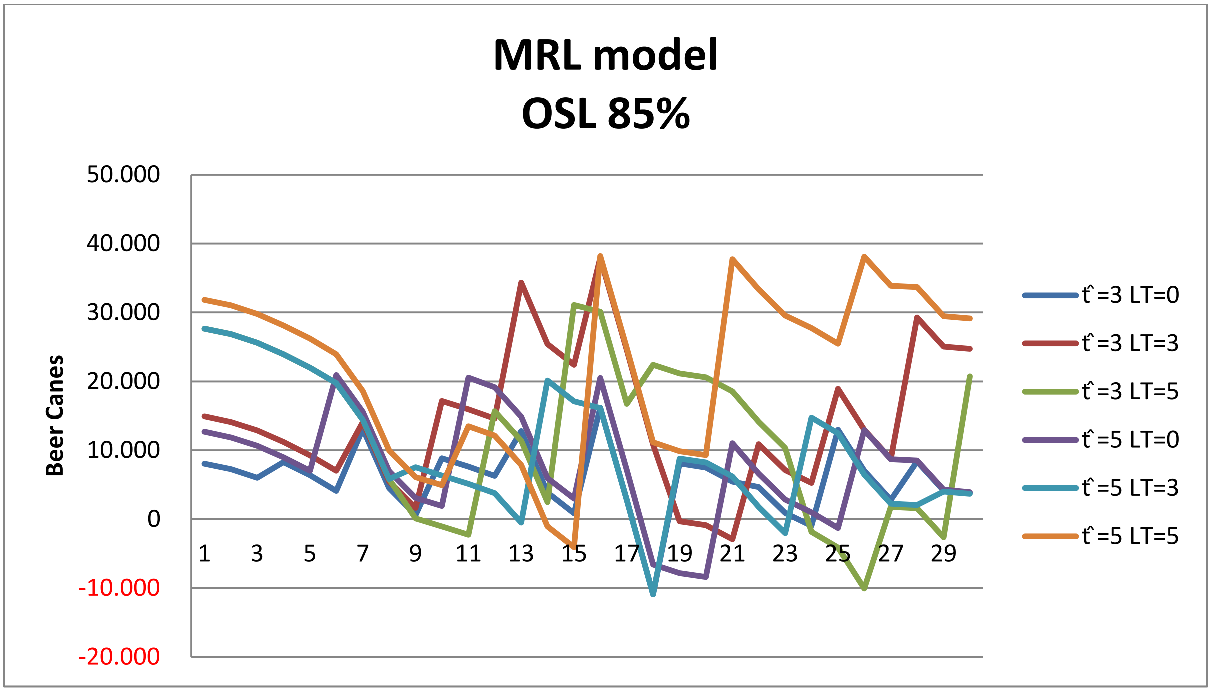


Figure 23: ($\hat{\boldsymbol{t}}$, $\hat{\boldsymbol{S}}$) inventory policy with $\hat{\boldsymbol{t}}$=3, 5 and LT= 0, 3, 5 for the MLR model with OSL 85%

## 8.3. Safety Stock Planning with a given USL

As already described in Chapter 4.2, the USL counts the average number of units short expressed as the percentage of the order quantity while the OSL measures the percentage of cycles that will be out of stock.

However, every given OSL has a corresponded USL and can be transformed with some simple steps as shown in Table 32. The transformation shown in Table 32 uses the data from MLR model for ($\hat{t}$, $\hat{S}$) inventory policy with $\hat{t}$= 3 and LT=0 for OSL= 95% and 85%.

As it can be seen from Table 32 the USL that corresponds to a given OSL has a higher value. It can be also noticed that a 10% reduction of the OSL results only in 0,8% reduction of USL.

| OSL | | 95% | 85% |
|---|---|---|---|
| Safety Stock | $SS = \hat{k}\sigma\sqrt{\hat{t}+LT}$ | 8.499 | 5.325 |
| Sigma | $\sigma_{\hat{t}+LT} = \sigma\sqrt{\hat{t}+LT}$ | 2956 | 2956 |
| Expected shortage per replenishment cycle | $ESC = -ss\left[1-F_s\left(\frac{ss}{\sigma_{\hat{t}+LT}}\right)\right]+\sigma_{\hat{t}+LT}\, f_s\left(\frac{ss}{\sigma_{\hat{t}+LT}}\right)$ | 1,74 | 42,06 |
| USL | $USL = (u-ESC)/u$ | 99,95% | 98,8% |

Table 32: Transforming the OSL into a USL[110]

Using the USL of 99,95% and 98,8% we can compare again the forecasted results of the Winter's model and MLR model in a ($\hat{t}$, $\hat{S}$) inventory policy with $\hat{t}$= 3, 5 weeks and lead time (LT) = 0, 3, 5 weeks. The inventory levels for 99,95% USL are shown in Table 33 and for 98,8% OSL shown in Table 34.

| Winter's model | | | | | | MRL model | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| USL 99.95% | | | | | | | | | | | |
| $\hat{t}$=3 | | | $\hat{t}$=5 | | | $\hat{t}$=3 | | | $\hat{t}$=5 | | |
| LT=0 | LT=3 | LT=5 | LT=0 | LT=3 | LT=5 | LT=0 | LT=3 | LT=5 | LT=0 | LT=3 | LT=5 |
| 47.820 | 69.515 | 82.830 | 62.750 | 60.569 | 94.658 | 11.253 | 18.276 | 29.032 | 16.028 | 27.360 | 33.188 |
| 46.996 | 68.691 | 82.006 | 61.926 | 59.745 | 93.834 | 10.429 | 17.452 | 28.208 | 15.204 | 26.536 | 32.364 |
| 45.769 | 67.464 | 80.779 | 60.699 | 58.518 | 92.607 | 9.202 | 16.225 | 26.981 | 13.977 | 25.309 | 31.137 |
| 47.065 | 65.791 | 79.106 | 59.026 | 56.845 | 90.934 | 11.435 | 14.552 | 25.308 | 12.304 | 23.636 | 29.464 |
| 45.128 | 63.854 | 77.169 | 57.089 | 54.908 | 88.997 | 9.498 | 12.615 | 23.371 | 10.367 | 21.699 | 27.527 |
| 42.885 | 61.611 | 74.926 | 62.592 | 52.665 | 86.754 | 7.255 | 10.372 | 21.128 | 22.064 | 19.456 | 25.284 |
| 44.231 | 60.038 | 69.588 | 57.254 | 47.327 | 81.416 | 13.881 | 15.357 | 15.790 | 16.726 | 14.118 | 19.946 |
| 35.640 | 51.447 | 60.997 | 48.663 | 38.736 | 72.825 | 5.290 | 6.766 | 7.199 | 8.135 | 5.527 | 11.355 |
| 31.744 | 47.551 | 59.452 | 44.767 | 63.420 | 68.929 | 1.394 | 2.870 | 109 | 4.239 | 8.441 | 7.459 |
| 47.683 | 55.664 | 58.264 | 43.579 | 62.232 | 67.741 | 11.993 | 15.061 | -1.079 | 3.051 | 7.253 | 6.271 |
| 46.479 | 54.460 | 57.060 | 63.234 | 61.028 | 73.887 | 10.789 | 13.857 | -2.283 | 21.932 | 6.049 | 12.361 |
| 45.129 | 53.110 | 64.245 | 61.884 | 59.678 | 72.537 | 9.439 | 12.507 | 14.422 | 20.582 | 4.699 | 11.011 |
| 45.462 | 72.401 | 59.972 | 57.611 | 55.405 | 68.264 | 13.832 | 33.702 | 10.149 | 16.309 | 426 | 6.738 |
| 36.530 | 63.469 | 51.040 | 48.679 | 67.647 | 59.332 | 4.900 | 24.770 | 1.217 | 7.377 | 19.547 | -2.194 |
| 33.526 | 60.465 | 72.040 | 45.675 | 64.643 | 56.328 | 1.896 | 21.766 | 30.783 | 4.373 | 16.543 | -5.198 |
| 47.814 | 96.555 | 71.084 | 62.655 | 63.687 | 79.177 | 17.170 | 47.890 | 29.827 | 21.971 | 15.587 | 40.294 |
| 34.434 | 83.175 | 57.704 | 49.275 | 50.307 | 65.797 | 3.790 | 34.510 | 16.447 | 8.591 | 2.207 | 26.914 |
| 20.746 | 69.487 | 63.147 | 35.587 | 36.619 | 52.109 | -9.898 | 20.822 | 23.721 | -5.097 | -11.481 | 13.226 |
| 47.876 | 78.349 | 61.886 | 34.326 | 17.957 | 50.848 | 10.957 | 22.772 | 22.460 | -6.358 | 11.210 | 11.965 |
| 47.308 | 77.781 | 61.318 | 33.758 | 17.389 | 50.280 | 10.389 | 22.204 | 21.892 | -6.926 | 10.642 | 11.397 |
| 45.258 | 75.731 | 70.751 | 64.848 | 15.339 | 87.332 | 8.339 | 20.154 | 19.842 | 14.995 | 8.592 | 42.490 |
| 45.302 | 74.451 | 66.371 | 60.468 | 10.959 | 82.952 | 7.717 | -4.380 | 15.462 | 10.615 | 4.212 | 38.110 |
| 41.499 | 70.648 | 62.568 | 56.665 | 7.156 | 79.149 | 3.914 | -8.183 | 11.659 | 6.812 | 409 | 34.307 |
| 39.668 | 68.817 | 82.855 | 54.834 | 74.351 | 77.318 | 2.083 | -10.014 | -500 | 4.981 | 14.665 | 32.476 |
| 46.543 | 66.535 | 80.573 | 52.552 | 72.069 | 75.036 | 14.435 | 21.059 | -2.782 | 2.699 | 12.383 | 30.194 |
| 40.587 | 60.579 | 74.617 | 58.559 | 66.113 | 130.623 | 8.479 | 15.103 | -8.738 | 14.985 | 6.427 | 45.178 |
| 36.353 | 56.345 | 85.674 | 54.325 | 61.879 | 126.389 | 4.245 | 10.869 | 2.924 | 10.751 | 2.193 | 40.944 |
| 49.429 | 59.188 | 85.496 | 54.147 | 61.701 | 126.211 | 11.679 | 33.681 | 2.746 | 10.573 | 2.015 | 40.766 |
| 45.190 | 54.949 | 81.257 | 49.908 | 60.121 | 121.972 | 7.440 | 29.442 | -1.493 | 6.334 | 7.602 | 36.527 |
| 44.861 | 54.620 | 80.928 | 49.579 | 59.792 | 121.643 | 7.111 | 29.113 | 23.765 | 6.005 | 7.273 | 36.198 |

Table 33: Inventory levels for 99,95% USL using ($\hat{\boldsymbol{t}}$, $\hat{\boldsymbol{S}}$) inventory policy with $\hat{\boldsymbol{t}}$= 3, 5 and LT = 0, 3, 5

[110] S. Chopra, P. Meindl, Supply Chain Management, 2007, pp. 310

| Winter's model | | | | | | MRL model | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| USL 98.8% | | | | | | | | | | | |
| $\hat{t}$=3 | | | $\hat{t}$=5 | | | $\hat{t}$=3 | | | $\hat{t}$=5 | | |
| LT=0 | LT=3 | LT=5 | LT=0 | LT=3 | LT=5 | LT=0 | LT=3 | LT=5 | LT=0 | LT=3 | LT=5 |
| 36.047 | 52.574 | 62.593 | 47.284 | 62.930 | 71.655 | 8.079 | 13.642 | 22.929 | 11.863 | 22.929 | 26.644 |
| 35.223 | 51.750 | 61.769 | 46.460 | 62.106 | 70.831 | 7.255 | 12.818 | 22.105 | 11.039 | 22.105 | 25.820 |
| 33.996 | 50.523 | 60.542 | 45.233 | 60.879 | 69.604 | 6.028 | 11.591 | 20.878 | 9.812 | 20.878 | 24.593 |
| 36.118 | 48.850 | 58.869 | 43.560 | 59.206 | 67.931 | 8.363 | 9.918 | 19.205 | 8.139 | 19.205 | 22.920 |
| 34.181 | 46.913 | 56.932 | 41.623 | 57.269 | 65.994 | 6.426 | 7.981 | 17.268 | 6.202 | 17.268 | 20.983 |
| 31.938 | 44.670 | 54.689 | 47.127 | 55.026 | 63.751 | 4.183 | 5.738 | 15.025 | 17.239 | 15.025 | 18.740 |
| 32.458 | 43.096 | 49.351 | 41.789 | 49.688 | 58.413 | 10.093 | 10.072 | 9.687 | 11.901 | 9.687 | 13.402 |
| 23.867 | 34.505 | 40.760 | 33.198 | 41.097 | 49.822 | 1.502 | 1.481 | 1.096 | 3.310 | 1.096 | 4.811 |
| 19.971 | 30.609 | 39.551 | 29.302 | 43.520 | 45.926 | -2.394 | -2.415 | 109 | -586 | 6.602 | 915 |
| 36.942 | 38.722 | 38.363 | 28.114 | 42.332 | 44.738 | 8.870 | 14.410 | -1.079 | -1.774 | 5.414 | -273 |
| 35.738 | 37.518 | 37.159 | 48.035 | 41.128 | 51.261 | 7.666 | 13.206 | -2.283 | 17.239 | 4.210 | 11.613 |
| 34.388 | 36.168 | 44.683 | 46.685 | 39.778 | 49.911 | 6.316 | 11.856 | 14.004 | 15.889 | 2.860 | 10.263 |
| 33.688 | 55.459 | 40.410 | 42.412 | 35.505 | 45.638 | 10.146 | 27.910 | 9.731 | 11.616 | -1.413 | 5.990 |
| 24.756 | 46.527 | 31.478 | 33.480 | 47.747 | 36.706 | 1.214 | 18.978 | 799 | 2.684 | 14.949 | -2.942 |
| 21.752 | 43.523 | 20.999 | 30.476 | 44.743 | 33.702 | -1.790 | 15.974 | 24.178 | -320 | 11.945 | -5.946 |
| 36.867 | 60.919 | 20.043 | 47.723 | 43.787 | 59.944 | 13.535 | 28.847 | 23.222 | 17.278 | 10.989 | 31.508 |
| 23.487 | 47.539 | 6.663 | 34.343 | 30.407 | 46.564 | 155 | 15.467 | 9.842 | 3.898 | -2.391 | 18.128 |
| 9.799 | 33.851 | 11.769 | 20.655 | 16.719 | 32.876 | -13.533 | 1.779 | 11.180 | -9.790 | -16.079 | 4.440 |
| 36.516 | 43.005 | 10.508 | 19.394 | 17.957 | 31.615 | 8.499 | 2.602 | 9.919 | -11.051 | 9.872 | 3.179 |
| 35.948 | 42.437 | 9.940 | 18.826 | 17.389 | 31.047 | 7.931 | 2.034 | 9.351 | -11.619 | 9.304 | 2.611 |
| 33.898 | 40.387 | 50.851 | 48.849 | 15.339 | 38.561 | 5.881 | -16 | 7.301 | 10.963 | 7.254 | 26.505 |
| 33.735 | 57.509 | 46.471 | 44.469 | 10.959 | 34.181 | 5.311 | 10.333 | 2.921 | 6.583 | 2.874 | 22.125 |
| 29.932 | 53.706 | 42.668 | 40.666 | 7.156 | 30.378 | 1.508 | 6.530 | -882 | 2.780 | -929 | 18.322 |
| 28.101 | 51.875 | 94.095 | 38.835 | 53.776 | 28.547 | -323 | 4.699 | 6.523 | 949 | 10.318 | 16.491 |
| 36.836 | 64.092 | 91.813 | 36.553 | 51.494 | 26.265 | 10.903 | 16.063 | 4.241 | -1.333 | 8.036 | 14.209 |
| 30.880 | 58.136 | 85.857 | 43.360 | 45.538 | 95.930 | 4.947 | 10.107 | -1.715 | 10.490 | 2.080 | 29.754 |
| 26.646 | 53.902 | 96.577 | 39.126 | 41.304 | 91.696 | 713 | 5.873 | 9.696 | 6.256 | -2.154 | 25.520 |
| 37.862 | 68.234 | 96.399 | 38.948 | 41.126 | 91.518 | 9.375 | 23.688 | 9.518 | 6.078 | -2.332 | 25.342 |
| 33.623 | 63.995 | 92.160 | 34.709 | 47.304 | 87.279 | 5.136 | 19.449 | 5.279 | 1.839 | 6.515 | 21.103 |
| 33.294 | 63.666 | 91.831 | 34.380 | 46.975 | 86.950 | 4.807 | 19.120 | 23.430 | 1.510 | 6.186 | 20.774 |

Table 34: Inventory levels for 98,8% USL using ($\hat{\boldsymbol{t}}$, $\widehat{\boldsymbol{S}}$) inventory policy with $\hat{\boldsymbol{t}}$= 3, 5 and LT = 0, 3, 5

Observing the inventory levels we can see that the behavior and the relationships that were analyzed for OSL for the Winter's model and MLR model remain the same. However, analyzing the inventory levels of MLR model more carefully we can see that with the increase of the periodic review or the lead time following the given USL of 99,95% or 98,8% we are more often in the out of stock situation. A comparison between 95% OSL and 99,95 USL with $\hat{t}$=5 and LT=3 is shown in Figure 24.

This is caused by the fact that the USL is defined for $\hat{t}$=3 and LT=0. With the increase of the periodic review time and/or the lead time the demand for a given replenishment cycle also increases which leads to higher values for USL in the transformation process shown in Table 32. This leads to the conclusion that with the increase of the periodic review time and/or the lead time the OSL does not change but the USL increases.[111]

[111] S. Chopra, P. Meindl, Supply Chain Management, 2007, p. 312

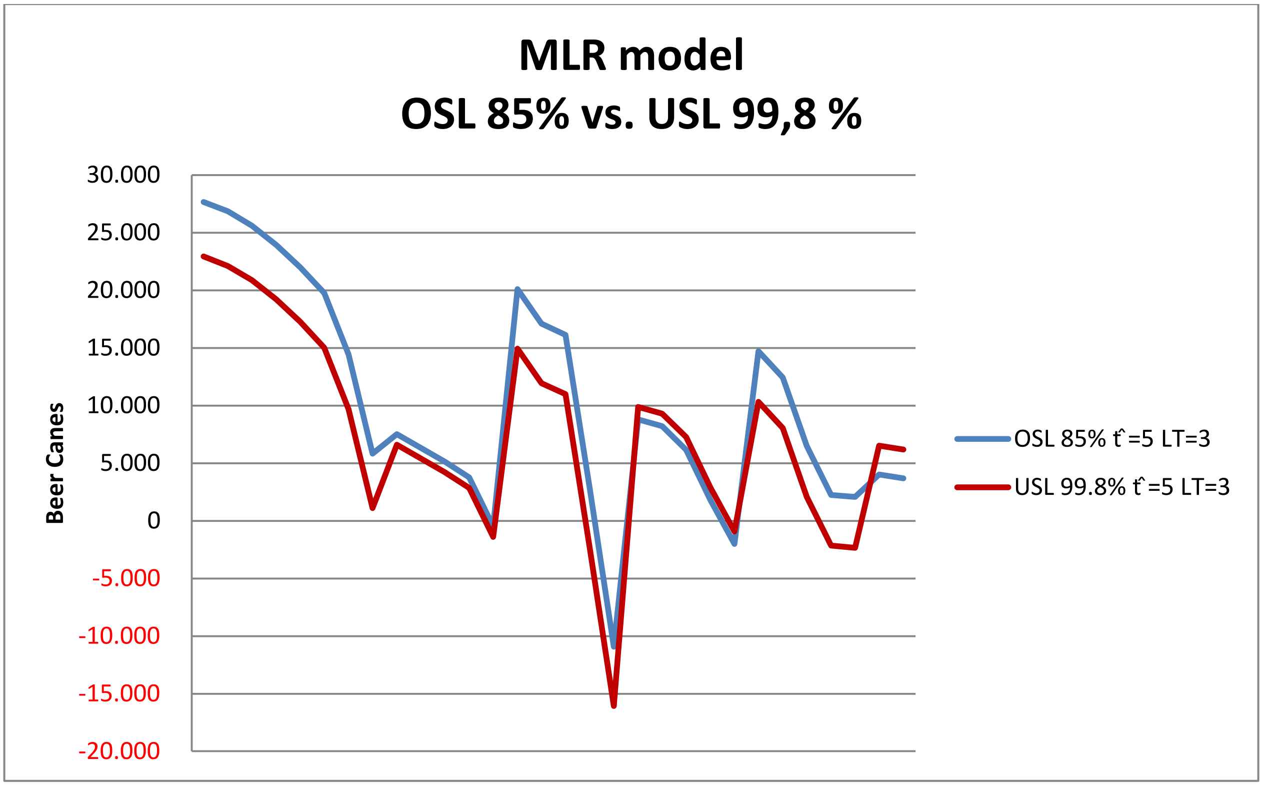


Figure 24: ($\hat{\boldsymbol{t}}$, $\widehat{\boldsymbol{S}}$) inventory policy with $\hat{\boldsymbol{t}}$= 5 and LT= 3 for the MLR model with OSL 85% and USL 99,8%

# 9. Conclusion

This work provided a comparison between the time series forecasting models and the causal forecasting models. Additionally, the work investigates the safety stock planning in ($\hat{t}$, $\hat{S}$) inventory policy and the impact that the forecasting models with their inaccuracy have on it. For the comparison a sales data for “Schwechater” beer canes sold in Austria for 2005, 2006 and 2007 by ADEG Austria Ltd were investigated.

In Chapter 7 the different forecasting models were investigated and compared. It was shown that the causal forecasting models are more accurate than the time series forecasting models, especially for longer time horizon. It was also shown that the increase of the variables included in the model improves the forecast accuracy.

However, it must not be forgotten that the variables included in the regression model should be chosen very carefully and their impact on the model investigated. Nevertheless a method that can be used in case of heteroscedasticity in order to eliminate its effect was also presented.

In Chapter 8 the forecasting results of the Winter’s method and the multiple regression model were compared in the ($\hat{t}$, $\hat{S}$) inventory policy with $\hat{t}$ = 3, 5 weeks and lead time (LT) = 0, 3, 5 weeks with a given OLS and USL. The autocorrelation impact on the safety stock planning and a general optimization approach for the presence of autocorrelation were also presented

The impact of the forecasting inaccuracy on the inventory levels was shown as well as that the increase of the lead time and/or the periodic review time increases the inventory levels. In addition the main relationships and differences between the OSL and USL were also highlighted. It was paid especially attention to the fact that if a given OSL is transformed into a USL, the USL increases as the demand for a given replenishment cycle increase while the OSL remains the same.

As the importance of the inventory management will be more crucial for the company’s success in the future there is no doubt that the safety stock planning will continue to develop. However, the current state already offers some methods that can ensure an accurate planning of the inventory levels.

# List of notations

| | |
|---|---|
| α | A positive constant |
| $\bar{\alpha}$ | A smoothing constant |
| $\hat{\alpha}$ | The intercept for the whole population |
| $\bar{a}$ | The intercept for the sample data |
| β | A smoothing constant |
| β* | A specific numerical value of $\hat{\beta}$ |
| $\hat{\beta}$ | The slope for the whole population |
| b | The trend |
| $\bar{b}$ | The slope for the sample data |
| $\tilde{b}$ | A k *x* 1 matrix of $\bar{b}$ |
| B($\tilde{t}$) | The backorder at time $\tilde{t}$ |
| c | An arbitrary nonzero vector |
| $\bar{\delta}$ | A Boolean variable |
| δ | A positive constant |
| D | A dummy variable |
| DR | The demand during the review interval |
| ε | The error term in the regression model |
| e | Difference between the forecasted value and the observed value |
| $\bar{e}$ | The residual |
| $\hat{e}$ | A k *x* 1 matrix of $\bar{e}$ |
| ESC | Expected shortage per replenishment cycle |
| $\phi$ | A smoothing constant |
| $\bar{\phi}$ | Sum of the squared errors |
| $\tilde{\phi}$ | The partial autocorrelation coefficient |
| $\hat{\phi}(\dot{k})$ | The standard normal probability density function |
| $\Phi(\dot{k})$ | The cumulative distribution function |

| | |
|---|---|
| $\Phi^0(\dot{k})$ | The standard normal complementary cumulative distribution function |
| F | Forecast result |
| $\hat{F}$ | Variable for the F-test for overall significance |
| $G_u(\dot{k})$ | The standard normal loss function |
| $H_0$ | The null hypothesis |
| i | A state counting variable |
| I | The seasonal adjustment factor |
| $\tilde{I}(\tilde{t})$ | The inventory level at time $\tilde{t}$ |
| IN$(\tilde{t})$ | The net inventory at time $\tilde{t}$ |
| IO$(\tilde{t})$ | The inventory on order at time $\tilde{t}$ |
| IP$(\tilde{t})$ | The inventory position at time $\tilde{t}$ |
| j | A state counting variable |
| k | The number of the independent variables |
| $\bar{k}$ | The number of parameters |
| $\hat{k}$ | The safety stock factor |
| L | Number of consecutive errors of the same sign |
| L | The length of seasonality |
| LT | The lead time |
| m | The number of forecasted periods |
| $\bar{m}$ | The order of the autocorrelation process |
| n | The last state of a state counting variable |
| $\bar{n}$ | The sample size |
| $\dot{n}$ | The total number of time periods being considered |
| N | A prespecified limit |
| OSL | Order Service Level |
| P | The number of periods |
| $P$ | The periodicity of the demand |
| $P_1$ | The desired order service level |
| $P_2$ | The desired unit service level |
| $\rho$ | The autocorrelation coefficient |

| | |
|---|---|
| $r^2$ | The coefficient of determination for simple regression |
| $R^2$ | The coefficient of determination for multiple regression |
| $\overline{R}^2$ | The adjusted coefficient of determination for multiple regression |
| $\hat{R}$ | The time of the periodic review |
| $\sigma$ | The standard deviation of a normally distributed random variable |
| $\sigma^2$ | The variance of a normally distributed random variable |
| $\sigma_L$ | The standard deviation of the demand over the replenishment lead time |
| $\tilde{\sigma}^2$ | The constant factor of proportionality |
| $\hat{\sigma}$ | The standard deviation of the errors |
| s | A suppression factor |
| S | The smoothed value |
| $\hat{S}$ | The order-up-to level |
| SS | The safety stock |
| $\tau$ | A state counting variable |
| t | Variable for the t-test |
| $\hat{t}$ | The time of the periodic review |
| T | Time period |
| u | The average demand for a given period |
| USL | Unit Service Level |
| x | The regressor |
| $x^0$ | The regressor vector |
| $\dot{x}$ | The actual demand for a given period |
| $\overline{x}$ | The sample means of x |
| $\overline{X}$ | Mean of the data |
| $\check{X}$ | A $\bar{n}$ $x$ k matrix of rank K. |
| $\hat{X}$ | The independent variable |
| X | Element from the data |
| X* | A value for the regressor |
| y | The regressand |
| $\hat{y}$ | The expected estimate of the independent variable from the sample |

| | |
|---|---|
| $\bar{y}$ | The sample means of y |
| Y | The dependent variable |
| $\hat{Y}$ | The estimated value |
| $\gamma$ | A smoothing constant |

# References

Annals 2006-2008. Central Institute for Meteorology and Geodynamics www.zamg.ac.at/klima/jahrbuch/?ts=1271762341

Axsäter, Sven: "Inventory Control", 2nd Edition, Springer Science + Business Media, LLC, 2006

Brown Robert G.: "Statistical Forecasting for Inventory Control", McGraw-Hill Book Company, 1959

Brown, Robert G.: “Decision rules for inventory management”, Holt, Rinehart & Winston Inc., 1967

Chiasson, M., Fildes, R., Pidd, M.:"Intelligent thinking instead of critical realism?", Journal of the Operational Research Society, 2006, pg. 1373–1375

Chopra, S., Meindl, P., Supply Chain Management, 3rd edition, Prentice Hall, 2007

Chow, W. N., “Adaptive control of the exponential smoothing constant”, Journal of Industrial Engineering 16, 1965, pg. 314-317

Dennis, J. D., “A performance test of a run based adaptive exponential forecasting technique”, Production and Inventory Management 19, 1978, pg. 43-46

Dillon, Robert E.: "Some simple steps to inventory reduction", Production and Inventory Management Journal, First Quarter 1990, pg. 62

Fildes Robert: "Quantitative Forecasting the State of the Art: Econometric Models", The Journal of the Operational Research Society, Vol. 36, No. 7 (Jul., 1985), pg. 549- 580

Gigerenzer, G., Todd, P., ABC Research Group, 2000. Simple Heuristics that Make us Smart. Oxford University Press, Oxford

Goldfeld, Stephen M. and Quandt, Richard E., “Nonlinear Methods in Econometrics”, North-Holland Publishing Company, Amsterdam, 1972, pg. 93-94

Glejser, H., “A New Test for Heteroscedasticity”, Journal of the American Statistical Association, vol.64, 1969, pp. 316-323

Green, William H.: "Econometric Analysis", 2nd Edition, Prentice-Hall Inc, 1993

Gujarati Damodar N: "Basic Econometrics", 2nd Edition, McGraw-Hill Inc, 1988

Gujarati Damodar N: "Essentials of Econometrics", 2 Edition, McGraw-Hill Inc, 1992

Hsu John, El-Najdawi M. K.: "Integrating Safety Stock and Lot-Sizing Policies for Multi-Stage Inventory Systems Under Uncertainty", Journal of Business Logistics; 1991; Vol.12, No 2; pg. 221

Inderfurth Karl: ”Multistage safety stock planning with item demands correlated across products and through time”, Production and operations management; 1995; Vol.4, No 2; pg.127

Kapkova Albena: "Inventory Management for Perishable Goods", Master Thesis, University of Vienna, 2006

Keller Susanne: "Die Reduzierung des Bullwhip-Effektes", 1 Auflage, Deutscher Universitäts-Verlag, Wiesbden 2004

Krupp James (A): "Managing demand variations with safety stock", The Journal of Business Forecasting Methods & Systems; Summer 1997; 16, 2; pg. 8

Krupp James (B): "Safety stock management", Production and Inventory Management Journal; Third Quarter 1997; 38, 3; pg. 11

Makridakis Spyros, Wheelwright C., McGee Viktor E.: "Forcasting, methods and applications", John Wiley & Sons Inc., 1983

Makridakis Spyros, Wheelwright C.: "The Handbook of Forcasting", John Wiley & Sons Inc., 1987

Naddor, E.: "Inventory Systems", BSB B. G. Teubner, 1971

Najdawi, M. K., Liberatore, M. J.: "Determination of Optimal Safety Stock Policies", Studies in Business and Economics, Vol. 12, No. 1, 2006

Nikolopous, K., Goodwin, P., Patelis, A., Assimakopoulos, V.: ”Forecasting with cue information: A comparison of multiple regression with alternative forecasting approaches”, European Journal of Operational Research 180, 2007, pg. 354-368

Pantazopoulus, Sotiris N., Pappis, Costas P.: "A new adaptive method for extrapolative forecasting algotithms", European Journal of Operational Research, 1996, pg. 106-111

Park, R. E.: "Estimation with Heteroscedastic Error Terms", Econometrica, Oct 1966, pg. 888

Silver Edward A., Pyke David F., Peterson Rein: "Inventory Management and Production Planning and Scheduling" 3 Edition, John Wiley & Sons Inc., 1998

Seung-Kuk Paik, Prabir K. Bagchi: "Understanding the causes of the bullwhip effect in a supply chain", International Journal of Retail & Distribution Management, Bradford: 2007, Vol. 35, Iss. 4; pg. 308

Trigg, D.W., Leach A.G.,: “Exponential smoothing with an adaptive response rate”, Operations Research 18, 1967, pg. 13-26

Whybark, D.C.: “A comparison of adaptive forecasting techniques”, Logical Transactions 8, 1978, pg. 13-26

Wiersema, William H.: "Less inventory, more cash", Electrical Apparatus, Nov. 2008, pg. 33

Zipkin, Paul H.: "Foundations of Inventory Management", McGraw-Hill Inc, 2000

www.businessdictionary.com

http://www.spiegel.de/international/business/0,1518,592658,00.html
Spiegel, 25th November, 2008, German Auto Industry Facing the Abyss

Sales data – ADEG Österreich Handels AG; 2005, 2006 and 2007

# CURRICULUM VITAE


Lalo Valchkov
born on the 17th of February 1981
in Pleven, Bulgaria


**Academic Education**

| | |
|---|---|
| since 2011 | Master Degree in Engineering<br>Faculty of Mechanical Engineering<br>Specialization: *Quality Management*<br>Technical University of Sofia – Sofia, Bulgaria |
| 2008 – 2012 | Master Degree in Business Administration<br>1. Specialization: *Investments*; 2 Specialization: *Supply Chain Management*<br>Faculty of Business, Economics and Statistics, University of Vienna – Vienna, Austria |
| 2002 – 2008 | Bachelor Degree in Business Administration<br>Specialization: *Management*<br>Faculty of Business, Economics and Statistics, University of Vienna – Vienna, Austria |
| 2000 – 2008 | Bachelor Degree in Economics<br>Specialization: *Business Administration*<br>University of National and World Economics – Sofia, Bulgaria |
| 1996 – 2000 | First English Language School – Sofia, Bulgaria |

**Career**

| | |
|---|---|
| since 2009 | Employee at Atomic Austria GmbH – Altenmarkt im Pongau, Austria<br>Currant position: Distribution and Transportation Manager |

**Languages**

Bulgarian, English, German, Russian