Smoothing Techniques for Time Series Forecasting

Haifaa Hussein Hameed

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the Degree of

Master of Science
in
Applied Mathematics and Computer Science

Eastern Mediterranean University
July 2015
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Serhan Çiftçioğlu
Acting Director

I certify that this thesis satisfies the requirements as a thesis for the degree of Master of Science in Applied Mathematics and Computer Science.

Prof. Dr. Nazim Mahmudov
Chair, Department of Mathematics

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Applied Mathematics and Computer Science.

Prof. Dr. Rashad Aliyev
Supervisor

Examining Committee

1. Prof. Dr. Rashad Aliyev
2. Assoc. Prof. Dr. Benedek Nagy
3. Asst. Prof. Dr. Ersin Kuset Bodur

ABSTRACT

There are many forecasting techniques available, and selecting the appropriate technique is very important issue to achieve a good forecasting performance.

This thesis intends to present the smoothing techniques for time series forecasting. The forecasting process using simple moving average and weighted moving average methods is investigated. The exponential smoothing forecasting method is analyzed. The simple exponential smoothing method is described.

Some error measures - Mean Absolute Deviation, Mean Absolute Percentage Error, and Mean Square Error are calculated for above forecasting techniques to define the forecast accuracy of these methods.

The double exponential smoothing method is discussed.

Keywords: Forecasting, Time series, Simple moving average, Weighted moving average, Simple exponential smoothing, Double exponential smoothing

ÖZ

Birçok öngörü teknikleri mevcuttur ve tekniğin uygun seçilmesi iyi bir öngörü performansı elde etmek için çok önemli bir konudur.

Bu tez zaman serisi öngörüsü için düzeltme teknikleri sunmayı amaçlıyor. Basit hareketli ortalama ve ağırlıklı hareketli ortalama yöntemleri kullanarak öngörü süreci incelenmiştir. Üstel düzeltme öngörü yöntemi analiz edilir. Basit üstel düzeltme tarif edilir.

Yukarıdaki tekniklerde öngörü doğruluğunu tanımlamak için bazı hata önlemleri - Ortalama Mutlak Sapma, Mutlak Yüzde Hata ortalama, ve Ortalama Hata Kare hesaplanır.

Çift üstel düzeltme yöntemi tartışılır.

Anahtar Kelimeler: Öngörü, Zaman serisi, Basit hareketli ortalama, Ağırlıklı hareketli ortalama, Basit üstel düzeltme, Çift üstel düzeltme

ACKNOWLEDGMENT

I would not have been able to write this thesis in its complete form without the guidance and support of my supervisor Prof. Dr. Rashad Aliyev, and he is the one who is having a lot of experience, understanding and patience. He has spent a lot of his time in advising and supporting me, and hereby, I deliver my gratefulness and appreciation to him.

I would like to thank more than one person, starting with my father and dear mother, Naam Faraj, for the continuous support in my whole life. I would also like to thank my brothers and sisters and the whole family. I thank my beloved husband Kahi Kivee. Without his love, encouragement and help, this thesis would have never been finished. I thank my best friend Dr. Aree Ali, and all other friends for their continuous support.

At the end, I thank everyone who believed in me and who supported me for continuing my studies and finishing my thesis, without you all, this would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

People always attempt to predict the future, and forecasting is a statement used by the people to make a decision about the future.

Forecasting can be characterized as a process of prediction the future event based on past historical data (observations). One of the key ideas of the forecasting is being sure in amount of past data to be gathered for making prediction since the past data should be related with data to be predicted. In other words, the analysis of the identification of the appropriate number of historical data should be realized to effectively predict the future value.

In order to perform a reasonable forecasting process, the data collected for this purpose must be reliable and consistent. At the same time the objective of the forecast should be determined, and the forecasting model fitting the considered problem should be selected. After the forecasting is made, the result should be implemented.

Forecasting can be carried out for different types of problems according to time horizons. The accuracy of forecast decreases if time horizon for forecast increases, and versus visa the accuracy of forecast increases if time horizon for forecast decreases.

The, weather prediction, exchange rates, job assignments and scheduling require a short-range forecasting; the sales and production planning require a medium-range forecasting, and a strategic planning of the company requires a long-range forecasting. Normally, short-range and medium range forecasting provides better forecasting accuracy than long-range forecasting.

The demand for tellers in a supermarket, the demand for foods and soft drinks in groceries, the demand for fruits in markets are the examples of forecasting.

There are two types of forecasting techniques: quantitative and qualitative. The difference between them is that in qualitative type the forecasting process is subjective to be generated by the forecaster, but the quantitative type of forecasting is based on mathematical modeling.

This master thesis considers the time series methods of a quantitative forecasting. These methods are classified into the moving average, weighted moving average, and exponential smoothing methods.

Time series assumes some properties such as the information about the past observations must be available and quantitatively represented in data form, for example, to predict the demand for the product which has been on sales for many years, and based on past observations it is possible to predict the future demand. If the values of past observations are not available, for example, if we need to predict the demand for new products without having old data about quantity of the product sold before, the implementation of a time series method is meaningless.

Time series is applied in a wide range of subjects: economics, finance, meteorology etc. The minimum and maximum temperatures of a weather measured during a day, number of babies born within a month, number of people suffered from different diseases over a year compromise a time series forecasting.

A time series is a collection of observations which are collected sequentially over discrete or continuous time, and the data can be measured in uniformly distributed form. The period between measurements is considered at any regular interval like hours, days, weeks, months etc.

The data of time series can be classified as stationary and non-stationary data according to presence of absence of trend. If there is an upward or a downward trends in data, the time series is stationary. If there is no trend in data, the time series is non-stationary.

There are some components a time series forecasting can be decomposed into. These components are secular trend, seasonal variation, cyclical variation, and irregular variation.

The secular trend shows that data can smoothly increase or decrease for a long-term fluctuation. The population increment in a country, the price inflation, and changes in prices of petrol and gold can be considered as examples of the secular trend. The upward and downward secular trends in a time series are considered.

The seasonal variation shows that the changes in data can be affected by seasonal factors. There is a higher need in water and ice-cream in summer compare to winter, the number of cars sold in winter exceeds the number of cars sold in summer.

The cyclical variation in a time series means that some patterns can repeat themselves throughout a time series.

Time series always involves a random component. The absence of random component in time series is impossible, that is why the forecasting can't be perfect, i.e. reaching the optimal value in forecasting in real-world cases is almost impossible.

Data in irregular variation can be changed, and this can't be predicted in advance, and this property is not associated with the properties of the trend, seasonal, and cyclical components. The accuracy of the prediction in a time series is up to how the irregular variation is reduced, i.e. the less the irregular variation is, the higher the prediction is accurate.

Majority of people do not take major action in response to small fluctuations in the environment. For example, no one cancels a trip after feeling just one drop of rain. The manager of a firm does not employ more staff if his employees have overexertion of the work for one day only.

The model of "smoothing" through the formation of expectations on the basis of weighted averages is used in this thesis. Moving average is a way for smoothing time series by averaging (with or without weights) a fixed number of consecutive terms. Over time, the average account is "moving", leaving each series data points on average in this sequence, and also increases the average to delete old data points.

In a simple moving average all the observations are assigned with the same weight in averaging.

In a weighted moving average all the observations are assigned with different weights in averaging. The most recent observations are given higher weights in comparison with older observations where the weights decline. The difference between the weights of observations can be explained by the fact that recent data are more important for forecasting a new value than old data. The weights assigned to observations are based on intuition of a person.

In both simple and weighted moving average methods the total sum of the weights assigned to observations is equal to 1.

The exponential smoothing forecasting method is an equivalent to the weighted moving average method. The exponential smoothing is originated from the works of Brown (1959, 1962) and Holt (1960), and was intended to create a forecasting tool for stock control systems. The simple exponential smoothing is a technology to ensure a smooth and expected time series model in the border without addressing. It is based on a recursive computing strategy, whereas the forecast is updated with every new incoming notice.

The exponential smoothing is a method intended for short-term forecasts, and this is why the people can implement this forecasting technique in their daily living.

Can we rely on the accuracy of the result of the forecasting after a time series method is implemented? This question is important. Time series method must also provide the probability of accuracy of the forecasting. If this probability is high, the result of the forecasting can be accepted with a high confidence.

Chapter 2

REVIEW OF EXISTING LITERATURE ON SMOOTHING TECHNIQUES FOR TIME SERIES FORECASTING

[1] discusses the single space modeling framework that enables the modeling both linear and non-linear time series to forecast the seasonal time series features. The advantages of the provided procedure compare to traditional alternatives are given. The adaptability of the proposed framework dealing with data with both zero and negative values, and also with Gaussian distribution for the errors for point interval forecasts is a strong feature. Reducing the computational burden in maximum likelihood estimation is another main feature of the framework. The developed framework can be applied in a wide range of problems.

In [2] a new Empirical Information Criterion (EIC) is suggested which is used in a forecasting of a big number of time series whereas its bootstrap version was intended for forecasting a single time series. One of the advantages of EIC is to provide a tool to be used for tuning to specific task of forecasting. The comparison of the proposed criterion with other existing criteria shows that EIC provides better forecasting especially for longer horizons.

In [3] the approach for a time series forecasting with multiple seasonal patterns is introduced. A time series forecasting is carried out for both linear and non-linear seasonal patterns. The multiple seasonal models are applied to the utility demand and vehicle flows in hourly and daily patterns. The proposed approach is also useful to provide a model for some existing seasonal methods.

The paper [4] presents the examples of exponential smoothing techniques. The optimization linear regression model is considered in which the initial parameters and smoothing constants are optimized by minimizing mean square error (MSE). The linear regression method as a case of Holt's exponential smoothing with trend is presented. Another advantage is that it better fits the time series data.

The accuracy of exponential smoothing technique used by many organizations in forecasting is mostly depending on the appropriateness of the constant value of the exponential smoothing. In the paper [5] the optimal value of constant is defined by using the trial and error method. The use of the optimal value of the constant minimizes the mean square error and the mean absolute deviation in order to get an accurate forecasting.

The simple exponential smoothing technique is considered a short-range forecasting method. The evaluation of the accuracy of the forecasting depends on the value of the smoothing constant. The optimal value of the constant is made available using the lowest mean absolute error, the mean absolute percentage error, and the value root mean square error, and the forecasting for one year ahead is performed [6].

Two methods are proposed for Holt's additive exponential smoothing method for the parameter estimation problem [7]. The Bayesian approach is the first method to be considered. The advantage of this method is yielding a good forecast density. The second method allows the evaluation of the maximum likelihood parameter, and this method is based on state-space formulation of the problem.

In [8] the structure of predictive hybrid redundancy is proposed to remove most erroneous values. The double exponential smoothing is used. Five modules of the predictive hybrid redundancy and their roles are given. The computer simulation on MATLAB shows that the performance of the method outperforms the average and median voters of other dynamic and hybrid methods in terms of safety critical systems.

The damped trend exponential smoothing models are considered to perform an excellent forecasting. The forecast error variance based on ARIMA model is calculated in [9]. The relationships between forecast error variances of trend and damped exponential smoothing models and structural parameters of the same forecasting techniques are defined.

In [10] the exponential smoothing prediction model is extended to univariate time series where the observations are irregular. A new alternative prediction method to a Wriht's simple exponential smoothing model is offered. The discounted least squares method is used to estimate the polynomial trend of order m. The maximum likelihood parameters are estimated to make a forecasting in irregular time-series using ARIMA process.

In [11] two algorithms are built in which the first is the adaptive outlier-tolerant algorithm for the selection of the parameter of the exponential prediction smoothing, and the exponential smoothing prediction model for the elimination of a negative influence from outliers for the process of forecasting in case of necessity of sampling data.

In [12] the double exponential smoothing method for the prediction of the number of software failure is discussed. The advantage of the proposed method is its better accuracy compare to classical techniques, and the weight of most resent failure is given higher value that provides a reasonable prediction of the future events. A limited amount of storage and less computational effort are other advantages of the proposed prediction technique which are very important for the application in practise.

In the paper [13] the robust version of the exponential and Holt-Winters smoothing techniques are presented. A simulation to compare the robust and classical forecasts is performed. The good performance of the proposed method for time series is outlined. The real data trend and seasonal effects using the presented method is considered. The importance of the proposed method for different types of data is discussed.

A new adaptive method for modelling a smoothing parameter as a logistic function is presented in [14]. The simulation shows that the new method outperforms the performance of the existing methods which produce unstable forecasts in the presence of empirical situations.

The algorithms which are based on double exponential smoothing to predict the user position, are presented in [15]. The presented algorithms run much faster and show better performance compare to Kalman and extended Kalman prediction models. The easiness of the implementation of the algorithms is described.

The selection of the appropriate forecasting model is realized through Information Criteria (IC). The new exponentially weighted IC, presented in [16], is outperforming the performance of models based in standard IC, in particular, the Akaike's IC and Schwarz's Bayesian methods. As case studies, the sales data in supermarket and the call center arrivals are investigated.

In [17] five different exponentially weighted methods are evaluated to make forecasting for one day ahead. The singular value decomposition (SVD) based exponential smoothing method is developed, and this method shows better performance in load forecasting application.

A multivariate Exponentially Weighted Moving Average (EWMA) control chart is introduced in [18] which is used detect the small changes in process variability. EWMA V-chart is also introduced to monitor process mean. As one of the best alternatives for detecting small changes in the process it is suggested to combine EWMA M-chart and EWMA V-chart.

The approach to design residual-based EWMA chart is proposed in [19]. This approach is suitable for using in autoregressive moving average (ARMA) model where the uncertainty of parameters of this model should be taken into account. The proposed approach is compared with other two methods used for designing residual-based EWMA chart and shows better results in widening of control limits.

Chapter 3

FORECASTING USING SIMPLE MOVING AVERAGE AND WEIGHTED MOVING AVERAGE METHODS

Moving average model is used for a prediction process in which the forecast value is defined as a simple combination of average of recent actual values in time series. Two types of moving average method are discussed in this chapter: the first type is a simple moving average, and the second type is a weighted moving average.

3.1 Forecasting using simple moving average (SMA) method

SMA is a forecasting method where all the weights of recent actual value used for the forecasting are equal. Despite the simplicity of this method, SMA is the most common smooth method, and it is one of the quantitative methods used in determining the trend of the series.

Under this method the demand forecasting for future period equals the total demand for a certain number of past periods divided by the number of periods. The n-period moving average uses the actual value of last n periods used to forecast the next period value. A large number of recent actual values make the forecasting more stable, but a small number of recent actual values make the forecasting more responsive.

This method assumes that demand is stable and involves no seasonal factors. In this method of forecasting for the subsequent period equals the total quantity of production for a certain number of past data divided by the number (length) of the period.

$$\text{Moving Average} = \frac{\text{Total demand for a certain number of recent data}}{\text{Number of data}}$$

For example, for the prediction using four recent actual values, the total sum of values for these periods is found, and then this sum is divided by four, then the oldest value is discarded, and the new data is added to the end of the list.

The simple moving average is represented in the following form:

$$F_{t+1} = \frac{D_t + D_{t-1} + D_{t-2} + \cdots + D_{t-n+1}}{n} \tag{1}$$

where F_{t+1} is a forecast for the next period t+1, n is a number of periods to be averaged, $D_t + D_{t-1} + D_{t-2} + \cdots + D_{t-n+1}$ are the actual values (demand) for the past period, two periods ago, three periods ago and so on, respectively.

The data in Table 1 show the demand on the electric light unit for a particular company for 12 months, and 3-month moving average forecast is calculated.

Table 1: Demand and 3-month moving average forecast

Month	Demand (D_t)	3-month moving average forecast
Jan	35	
Feb	40	
Mar	42	
Apr	50	39
May	58	44
Jun	68	50
Jul	75	58.67
Aug	85	67
Sep	80	76
Oct	65	80
Nov	50	76.67
Dec	45	65

The 3-month simple moving average forecast for the month April was calculated as follows:

$$F_4 = \frac{D_3 + D_2 + D_1}{3} = \frac{42 + 40 + 35}{3} = 39$$

To calculate the 3-month moving average forecast for the month May, the demands for the fourth, third, and second months have been taken into account. So

$$F_5 = \frac{D_4 + D_3 + D_2}{3} = \frac{50 + 42 + 40}{3} = 44$$

This method assumes that the demand is stable and does not imply the seasonal factor. The demand forecasting for the next month (by using 3-month moving average) equals to

$$F\ (\text{next month}) = \frac{D_{12} + D_{11} + D_{10}}{3} = \frac{45 + 50 + 65}{3} = 53.33$$

In Figure 1 the graphical representation of demand and 3-month moving average is represented.

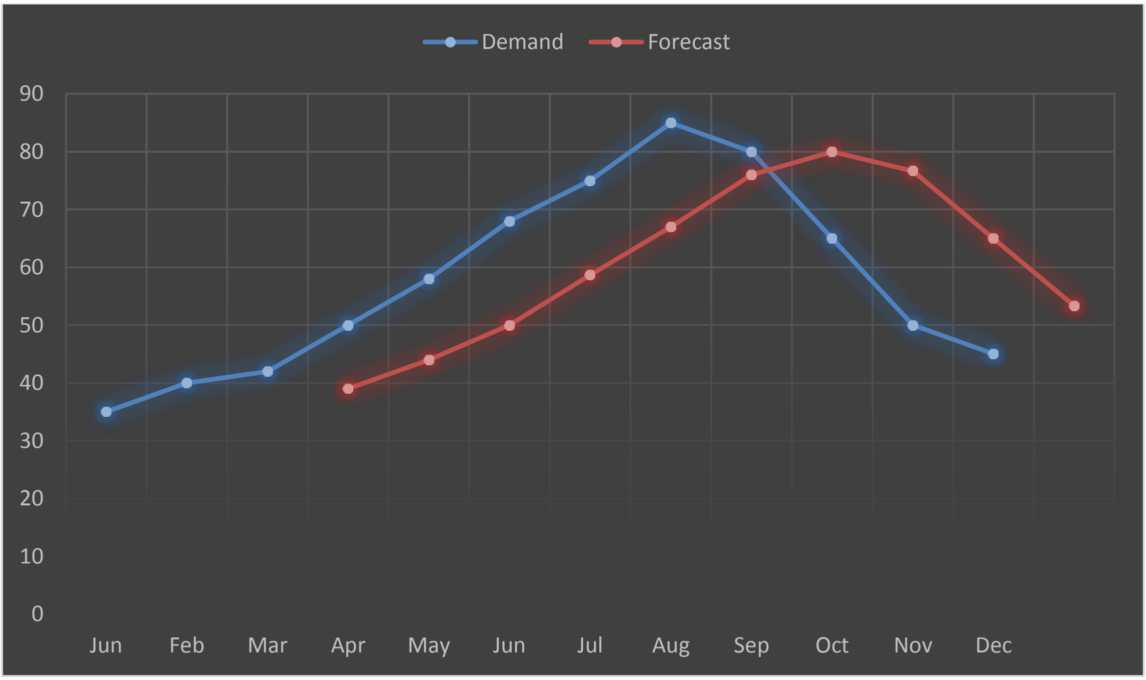

Figure 1: Graphical representation of demand and 3-month moving average forecast

The following is another example. Table 2 represents the demand for a specific product from 1999 to 2014, and 5-year moving average forecast is calculated. .

Table 2: Demand and 5-year moving average forecast

Year	Demand (D_t)	5-year moving average forecast
1999	230	
2000	245	
2001	300	
2002	336	
2003	368	
2004	380	295.8
2005	400	325.8
2006	422	356.8
2007	459	381.2
2008	470	405.8
2009	460	426.2
2010	455	442.2
2011	420	453.2
2012	399	452.8
2013	350	440.8
2014	313	416.8

In order to calculate the demand forecasting for the year 2015 (by using 5-year moving average), the demands for the years 2014, 2013, 2012, 2011, and 2010 are taken into account. So the demand forecasting for the year 2015 is equal to

$$\mathrm{F(for\ the\ year\ 2015)} = \frac{\mathrm{D}_{2014} + \mathrm{D}_{2013} + \mathrm{D}_{2012} + \mathrm{D}_{2011} + \mathrm{D}_{2010}}{5}$$

$$= \frac{313 + 350 + 399 + 420 + 455}{5} = 387.4$$

Figure 2 depicts the graphical representation for demand and 5-year moving average forecast.

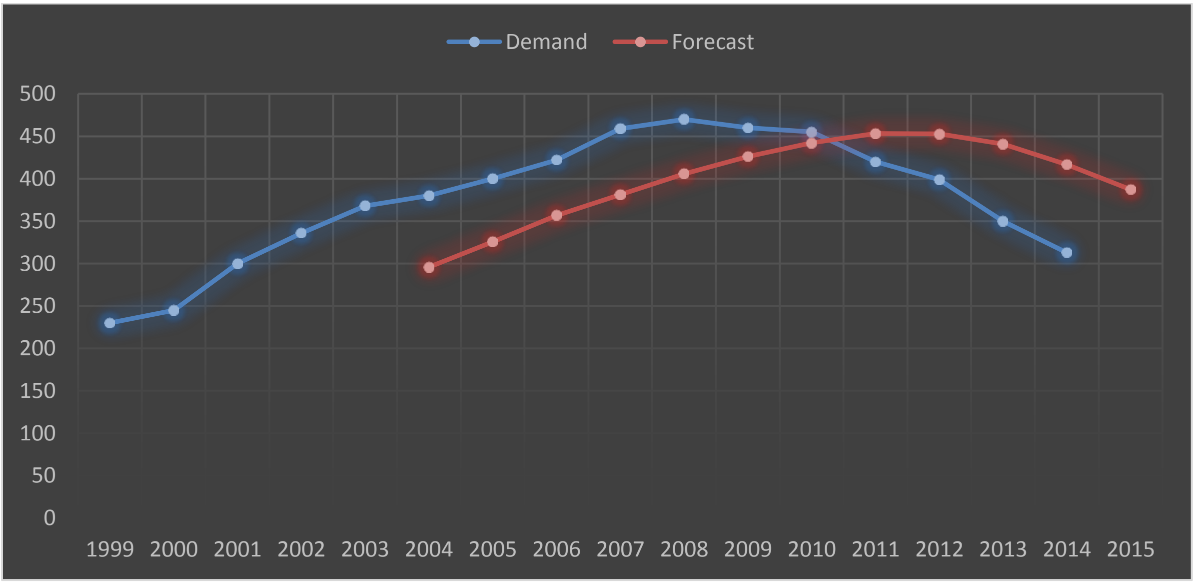

Figure 2: Graphical representation of demand and 5-year moving average forecast

3.2 Advantages and disadvantages of simple moving average method

The simple moving average method has some advantages and disadvantages. The advantages of this method are given below:

- easily computed;

- does not require a lot of data from the past;

- easily understood;

- removes “bad” data after n periods.

The disadvantages of this method are given below:

- outcome prediction depends on the length of the average, so the appropriate period for calculating the forecast should be chosen;

- requires to retain all the data from the past which leads to higher costs to save and retrieve data either manually or by computer;

- this method gives the same weight or significance to all the data used for the calculation of the forecast value;

- trends can’t be forecasted well.

3.3 Weighted Moving Average method

As it was mentioned above, the basic problem of a simple moving average consists in assigning the same weights to all the recent data (demand) to calculate a forecast value, but it can be sometimes required that higher weights should be given on particular recent period's data.

This disadvantage can be overcome by using weighted moving average (WMA), and WMA is also more suitable for the calculation of forecast values if there is a trend, because this method is more responsive to trends. The total weight is equal to 1, using the following equation

$$F_{t+1} = W_1 D_t + W_2 D_{t-1} + W_3 D_{t-2} + \cdots + W_n D_{t-n+1} \tag{2}$$

where

F_{t+1} is a forecast for the next period;

n is the total number of periods in the forecast;

W_i is the weight to be assigned to the demand;

$D_t, D_{t-1}, D_{t-2}, D_{t-n+1}$ are the actual values (demand) in the past period, two periods ago, three periods ago etc., respectively. The total sum of weights must be equal to 1:

$$\sum_{i=1}^{n} W_i = 1 \tag{3}$$

Table 3 shows the demand on the electric light unit for a specific company for 12 months to make a forecast using 3-month weighted moving average, if the weights assigned are $W_1 = 0{,}5$, $W_2 = 0.3$, $W_3 = 0.2$.

Table 3: Demand and 3-month weighted moving average forecast

Month	Demand (D_t)	3-month weighted moving average forecast
Jan	35	
Feb	40	
Mar	42	
Apr	50	40
May	58	45.6
Jun	68	52.4
Jul	75	61.4
Aug	85	69.5
Sep	80	78.6
Oct	65	80.5
Nov	50	73.5
Dec	45	60.5

The forecast for the month April was calculated as:

$$F_4 = 0.5 * 42 + 0.3 * 40 + 0.2 * 35 = 40$$

Let's find the forecast value for the next month which is calculated as follows:

$$\text{Next period} = 0.5 * 45 + 0.3 * 50 + 0.2 * 65 = 50.5$$

Figure 3 depicts the graphical representation for demand and 3-month weighted moving average forecast.

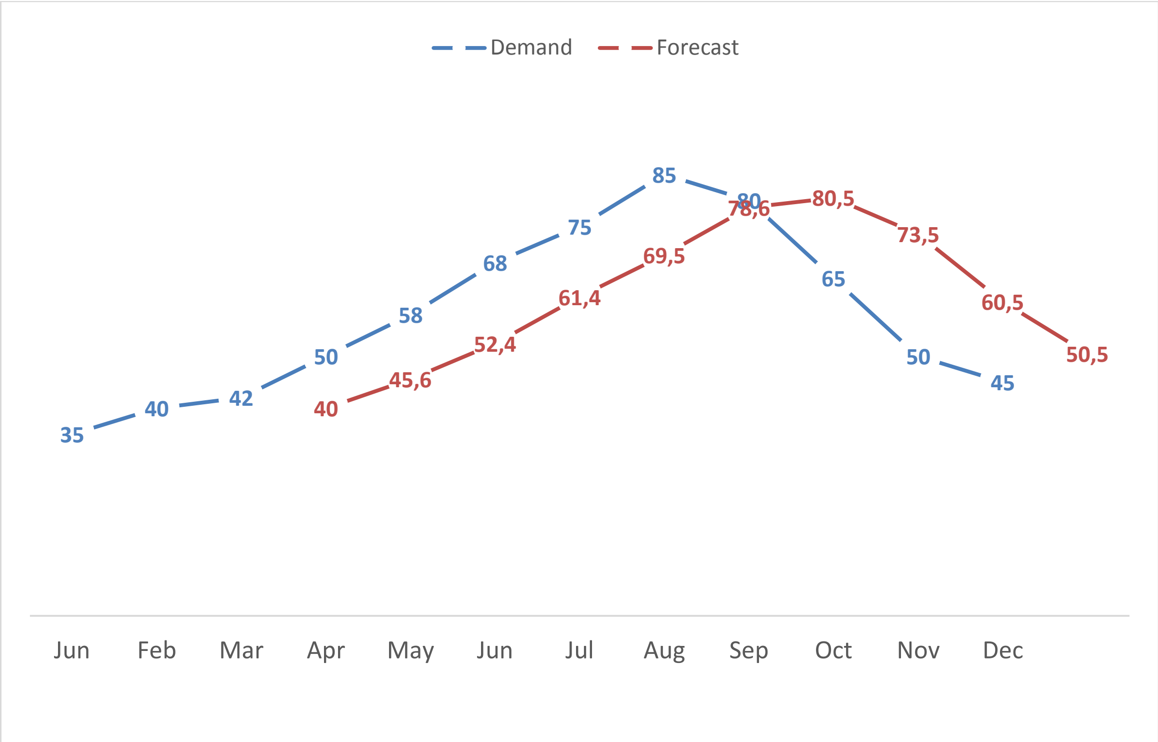

Figure 3: Graphical representation of demand and 3-month weighted moving average forecast

3.4 Determining the forecast accuracy for simple and weighted moving average methods

In most cases the prediction accuracy is essential in choosing an appropriate forecasting method, whatever method of forecasting tends to be fairly inaccurate, In order to realize it, the actual values must be compared with estimated forecast.

The criteria used for evaluating the forecasting accuracy are given below:

- Mean Absolute Deviation (MAD) determines how the forecast accuracy has safer measure. To compute a MAD, the sums of absolute values of the forecast errors are divided by the number of forecasts ($\mathrm{D_t} - \mathrm{F_t}$ is an error between demand value and forecast value):

$$MAD = \frac{\sum|D_t - F_t|}{n} \tag{4}$$

- Mean Absolute Percentage Error (MAPE) is calculated as

$$MAPE = 100 * \frac{\sum\left[\frac{|e_i|}{D_t}\right]}{n} \tag{5}$$

- Mean Square Error (MSE) is calculated as

$$MSE = \frac{\sum(D_t - F_t)^2}{n} \tag{6}$$

Suppose in the study of marketing a product, the data on demand for the 10 weeks have been collected shown in Table 4. To estimate the amount of the forecasted demand for this product, 3-week simple moving average method, and also 3-week weighted moving average method with $W_1 = 0{,}5$, $W_2 = 0.3$, $W_3 = 0.2$ are used.

Table 4: Demand and 3-week simple moving average forecast, and 3-week weighted moving average forecast

Week	Demand	3-week simple moving average forecast	Absolute Error	$(\text{Error})^2$	Absolute Percentage Error	3-week weighted moving average Forecast	Absolute Error	$(\text{Error})^2$	Absolute Percentage Error
1	20								
2	25								
3	33								
4	29	26	3	9	10.345	28	1	1	3.448
5	38	29	9	81	23.684	29.4	8.6	73.96	22.632
6	47	33.33333	13.667	186.78688	29.079	34.3	12.7	161.29	27.021
7	50	38	12	144	24	40.7	9.3	86.49	18.6
8	56	45	11	121	19.643	46.7	9.3	86.49	16.607
9	60	51	9	81	15	52.4	7.6	57.76	12.667
10	59	55.33333	3.667	13.446889	6.215	56.8	2.2	4.84	3.729
	Total		61.334	636.23376	127.9646		50.7	471.83	104.704

The absolute value of percentage error for the weeks 4 and 5 are calculated as:

For the week 4: $100 * \frac{|e_4|}{D_4} = 100 * \frac{3}{29} = 10.345.$

For the week 5: $100 * \frac{|e_5|}{D_5} = 100 * \frac{9}{38} = 23.684.$

In the same manner, the absolute value of percentage error for all other months are calculated.

The values of the error metrics MAD, MAPE, and MSE are calculated below:

MAD (3-week simple moving average) = 61.3333/7=8.762

MAD (3-week weighted moving average) =50.7/7=7.243

MAPE (3-week simple moving average) =127.965/7=18.281

MAPE (3-week weighted moving average) =104.7038/7=14.958

MSE (3-week simple moving average) = 636.23376/7 = 90.891

MSE (3-week weighted moving average) = 471.83/7=67.404

Table 5 shows the values of error metrics for 3-week simple moving average, and 3-week weighted moving average.

Table 5: Values of error metrics for 3-week simple moving average, and 3-week weighted moving average

Metric	3-week simple moving average	3-week weighted moving average
MAD	8.762	7.243
MAPE	18.281	14.958
MSE	90.891	67.404

Forecasting using 3-week weighted moving average method is better than 3-week simple moving average since the first forecasting method provides smaller standard errors (Figure 4).

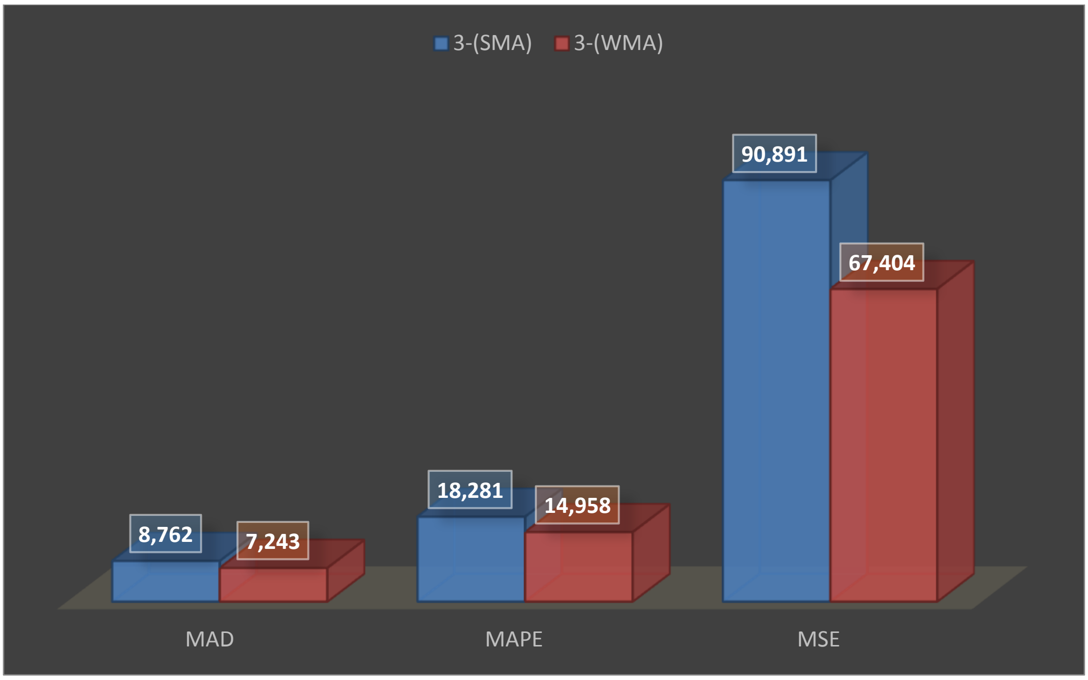

Figure 4: Comparison of metrics between 3-week simple moving average and 3-week weighted moving average

Chapter 4

THEORETICAL ASPECTS AND BASIC CONCEPTS OF EXPONENTIAL SMOOTHING METHOD

4.1 Forecasting using exponential smoothing method

Exponential smoothing is an important quantitative forecasting technique in a time series. Exponential smoothing is differed from other forecasting techniques by attaching maximum and minimum weights to most recent and old observations, respectively. In other words, the weight is declined exponentially when we go back data points in time. The value of the smoothing parameter or smoothing constant α determines the accuracy of forecasting, i.e. the optimal selection of the coefficient α leads to accurate prediction.

Because the forecasting with exponential smoothing is quite reliable and quick, using this model is a big advantage and importance in applications to the wide range of areas.

4.1.1 Simple exponential smoothing (SES)

The simple exponential smoothing (SES) model is best suited for a short-term forecasting, and usually this model is used for the prediction a future value for one month. SES model is a type of weighted moving average, and is generally known as exponentially weighted moving average (EWMA) model.

The simplicity of SES is because that only one parameter is to be estimated, and this parameter is the smoothing constant α.

The SES equation can be represented in the following form:

$$F_t = \alpha y_t + (1 - \alpha)\, F_{t-1} \tag{7}$$

where F_t is a forecast for the time series t, F_{t-1} represents the value of forecasting for the previous period t-1, and α is a smoothing constant. When the compensation value F_{t-1} is put in F_t, we get:

$$F_t = \alpha y_t + (1 - \alpha)[\alpha y_{t-1} + (1 - \alpha) F_{t-2}]$$

$$F_t = \alpha y_t + \alpha(1 - \alpha) y_{t-1} + (1 - \alpha)^2 F_{t-2} \tag{8}$$

$$F_t = \alpha y_t + \alpha(1 - \alpha) y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} + \alpha(1 - \alpha)^3 y_{t-3} + \alpha(1 - \alpha)^4 y_{t-4} + \cdots + (1 - \alpha)^t F_0 \tag{9}$$

The equation (3) can be written as:

$$F_t = \alpha \sum_{r=0}^{n-1} (1 - \alpha)^r y_{t-r} + (1 - \alpha)^t y_0 \tag{10}$$

where F_0 is the initial value of the smooth process, and $(1-\alpha)^r$ is a string of weights of the previous period, and these weights are progressively decreasing over time.

In order to start SES algorithm, we need to know F_1, because the calculation of F_2 needs F_1 to be known: $F_2 = \alpha y_1 + (1-\alpha) F_1$. The smoothed series starts with the smoothed version of the second observation. For any time period t, the expected value is determined as following:

New prediction = α (current observation) + (1-α) (last prediction)

or by the equation:

$$F_{t+1} = \alpha * D_t + (1-\alpha) * F_t \tag{11}$$

where D_t is the actual value (demand) for the period t, $0 \leq \alpha \leq 1$ is called the smoothing constant which determines the relative weight placed on the current observation. The original and smoothed versions of the series are similar when α =1. On the other hand, the series is smoothed flat when α = 0.

While using the algorithm, we need the initial forecast and the actual value of the smoothing constant. Since F_1 is not known, it is necessary to perform the following operations: setting the first estimate equal to the first observation, and then using the average of first few observations for the smoothed value.

The determination of smoothing weights for smoothing models is very important, because it is the spine of the smoothing model and the basis of accuracy in the operation of smoothing.

The researchers have different opinions about the value of the smoothing constant, but the majority offers to set the value of a smoothing constant through experience.

The exponential smoothing method requires the initial smoothed value to be set for the future forecast. The determination of the initial smoothed value is a difficult problem. The determination process becomes more difficult if we are dealing with many observations. Normally the initial smoothed value is taken equal to the first element of the list of observations or is determined as a mean value of the first few elements.

The estimation of the parameter called a fixed smoothing is one of the most important steps in forecasting, and exponential smoothing method depends on the value of the constant smoothing.

Let’s consider the following example. Table 6 represents the actual demand over the years 2010-2014.

Table 6: Actual demand over the years 2010-2014

Year	Demand D_t
2010	1467
2011	1500
2012	1433
2013	1395
2014	1400

If the simple exponential model is used, what is the expected demand for the year 2015? Let's solve the problem with the coefficient α=0.7. The exponential smoothing value for the year 2011 is 1467. So in order to forecast the demand for the year 2015, we do the following:

For 2012 = (0.7) * (1500) + (1-0.7) * (1467) = 1490.1

For 2013 = (0.7) * (1433) + (1-0.7) * (1490.1) =1450. 13

For 2014 = (0.7) * (1395) + (1-0.7) * (1450.13) = 1411.539

For 2015 = (0.7) * (1400) + (1-0.7) * (1411.539) = 1403.462

Figure 5 represents the comparison of actual demand and forecasted demand using SES.

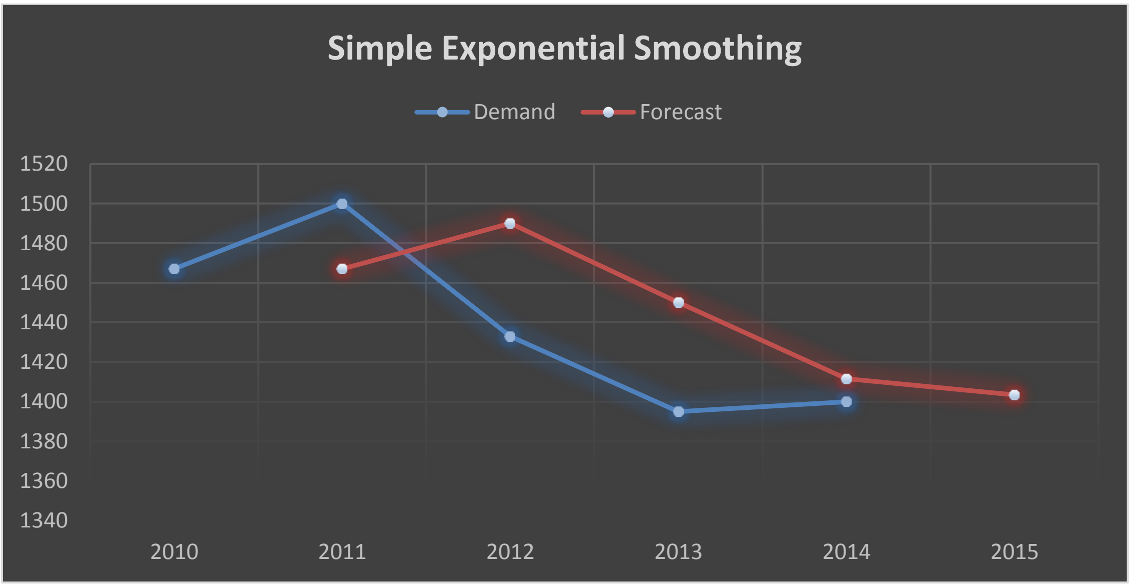

Figure 5: Comparison of actual demand and forecasted demand using SES

4.1.2 Testing the forecast accuracy

It is necessary to determine the error metrics of forecast accuracy affecting the quality of forecasting. The trial-and-error approach should be used to define the optimal value of smoothing constant for reducing the error in forecast accuracy.

One of most important error metrics is cumulative forecasting error (CFE) which is a sum of the forecasting errors, and the error is the difference between the actual time-series and forecast value for the same period. The formula for calculating CFE is

$$CFE = \sum e_t \tag{12}$$

$$\mathrm{e_t} = \mathrm{D_t} - \mathrm{F_t} \tag{13}$$

where D_t is the actual demand of time series for the period t, $\mathrm{F_t}$ is the forecast value of the time series and $\mathrm{e_t}$ represents the error for the period t.

We also need to calculate mean absolute deviation (MAD), and mean absolute percentage error (MAPE), and mean square error (MSE). If MAPE is 10% or below, then this case is considered as a forecasting with a very good accuracy. Mostly, MAPE varies in the range 20-30% which is common case in evaluation of accuracy of forecasting.

Another example is given below. In the study of product market, the following data were collected on the amount of demand for the product during the 12 months. The simple exponential smoothing model is applied. The calculation process of the forecast values for the smoothing constant α=0.5 is shown in Table 7.

Table 7: Calculation of the forecast value for the smoothing constant α=0.5

Month	Demand	Forecast (α=0.5)	Forecast Error $D_t - F_t$	Absolute Value of Forecast Error	$(Error)^2$	Absolute Value of Percentage Error
1	62					
2	64	62	2	2	4	3.125
3	60	63	-3	3	9	5
4	56	61.5	-5.5	5.5	30.25	9.8214
5	50	58.75	-8.75	8.75	76.5625	17.5
6	52	54.375	-2.375	2.375	5.6406	4.5673
7	55	53.1875	1.8125	1.8125	3.2852	3.2955
8	49	54.09375	-5.0938	5.0938	25.9468	10.3955
9	45	51.54688	-6.5469	6.5469	42.8619	14.5487
10	51	48.27344	2.7266	2.7266	7.4343	5.3463
11	57	49.63672	7.3633	7.3633	54.2182	12.9181
12	54	53.31836	0.6816	0.6816	0.4646	1.2622
Totals			-16.6817	45.8497	259.6641	87.78

The values of CFE, MAD, $MAPE$ and MSE are calculated for the smoothing constant α=0.5 as

CFE = $\sum e_t$ = -16.6817

MAD = 45.8497/11= 4.1682

MAPE = 87.78/11= 7.98

MSE = 259.6641/11= 23.6058.

The calculation process of the forecast value for the smoothing constant α=0.7 is shown in Table 8.

Table 8: Calculation of the forecast value for the smoothing constant α=0.7

Months	Demand	Forecast (α=0.7)	Forecast Error $D_t - F_t$	Absolute Value of Forecast Error	$(Error)^2$	Absolute Value of Percentage Error
1	62					
2	64	62	2	2	4	3.125
3	60	63.4	-3.4	3.4	11.56	5.6667
4	56	61.02	-5.02	5.02	25.2004	8.9643
5	50	57.506	-7.506	7.506	56.34004	15.012
6	52	52.2518	-0.2518	0.2518	0.063403	0.4842
7	55	52.07554	2.92446	2.92446	8.552466	5.3172
8	49	54.12266	-5.12266	5.12266	26.24165	10.4544
9	45	50.5368	-5.5368	5.5368	30.65615	12.304
10	51	46.66104	4.33896	4.33896	18.82657	8.5078
11	57	49.69831	7.30169	7.30169	53.31468	12.81
12	54	54.80949	-0.80949	0.80949	0.655274	1.4991
Totals			-11.0816	44.2119	235.4106	84.1447

The values of CFE, MAD, $MAPE$ and MSE are calculated for the smoothing constant α=0.7 as:

CFE = -11.0816;

MAD = 44.2119/11 = 4.0193;

MAPE = 84.1447/11 = 7.6495;

MSE = 235.4106/11= 21.401.

Table 9 depicts the values of error metrics for the smoothing constants α=0.5 and α=0.7.

Table 9: Values of error metrics for the smoothing constants α=0.5 and α=0.7

Standards	α=0.5	α=0.7
CFE	-16.6817	-11.0816
MAD	4.1682	4.0193
MAPE	7.98	7.6495
MSE	23.6058	21.401

Figure 6 represents the graphical representation of comparison of error metrics for α = 0.5 and α = 0.7.

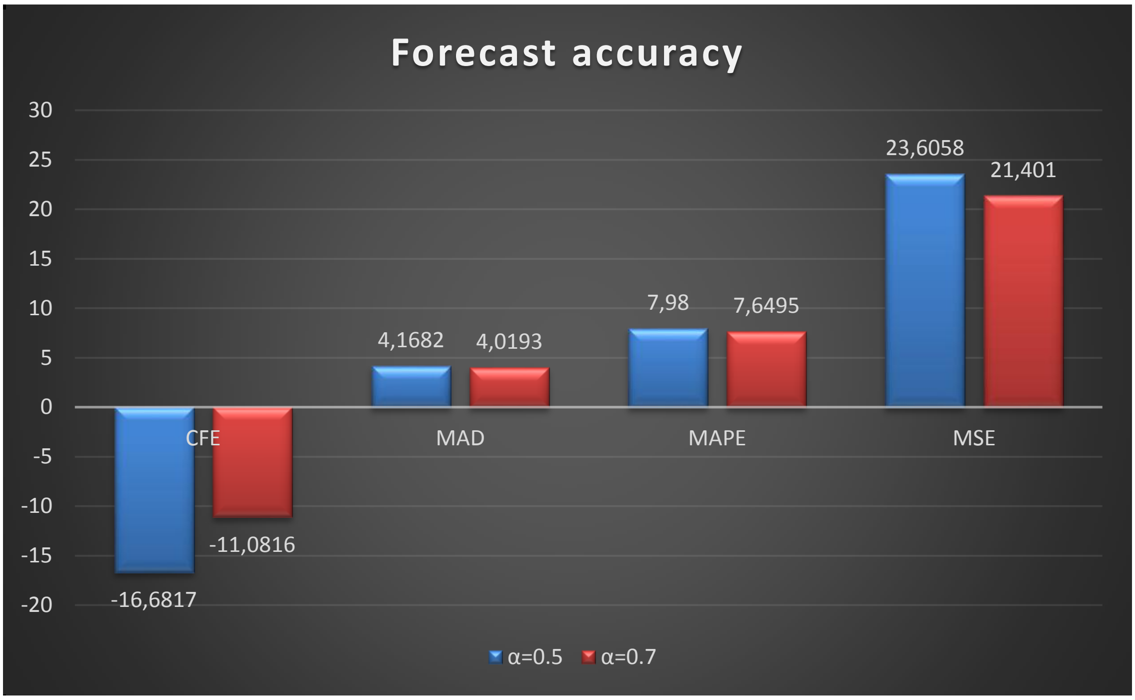

Figure 6: Graphical representation of comparison of error metrics for α = 0.5 and α = 0.7

The exponential smoothing model using α=0.7 provides a better forecast accuracy than the exponential smoothing using α=0.5 since it has a smaller MAD, MAPE, and MSE.

4.2 Double Exponential Smoothing (DES)

Double exponential smoothing (DES) method, also known as Holt-Winters method, is the extension of exponential smoothing for using in trended and seasonal time series, and in particular, in situations when there is a trend in data. DES uses two smoothing parameters to update the level and trend components. In this forecasting process three equations are used: the first equation is for smoothing time series, the second equation is for smoothing trend, and the third equation is for the combination of above two equations. So we have:

$$S_t = \alpha * D_t + (1-\alpha) * (S_{t-1} + b_{t-1}) \qquad 0 < \alpha < 1 \qquad (14)$$

$$b_t = \beta * (S_t - S_{t-1}) + (1-\beta) * b_{t-1} \qquad 0 < \beta < 1 \qquad (15)$$

$$F_{t+1} = S_t + b_t \qquad (16)$$

All the characters used in a single exponential smoothing equation represent the same meaning in a double exponential smoothing equation appearing (while α is smoothing constant for process-stabile-constant), but β is a trend-smoothing constant. S_t is the smoothed constant process value for the period t, and b_t is the smoothed trend value for the period t.

Suppose we have the actual sales data for 12 months represented in Table 10, and the graphical representation of actual sales data for 12 months is represented in Figure 7.

Table 10: Actual sales data for 12 months

Month	Sales
t	D_t
1	150
2	162
3	159
4	178
5	195
6	219
7	200
8	253
9	300
10	286
11	319
12	332

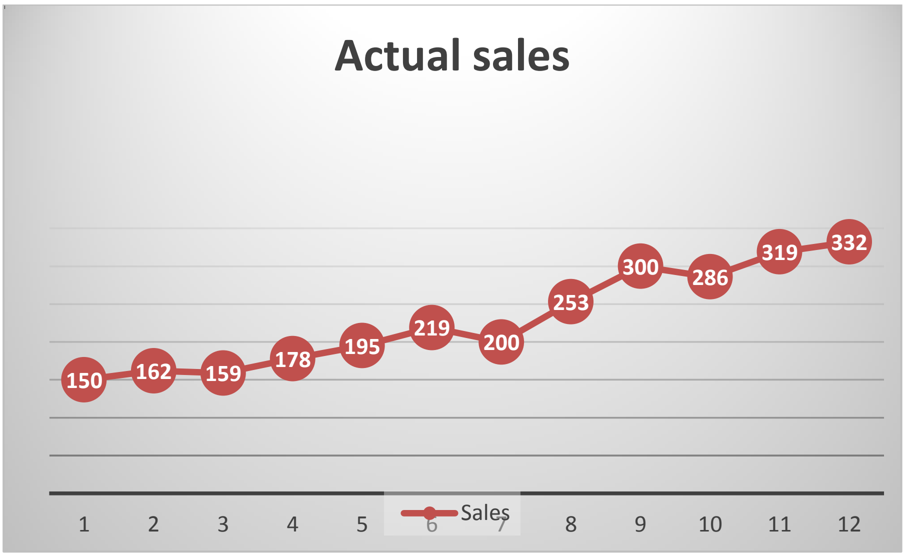

Figure 7: Graphical representation of actual sales data for 12 months

Note that the time series exhibits a growing trend, and then must use the double exponential smoothing. Firstly the initial values for S and b are determined. S_1 and b_1 are not defined, and one way to identify these values is assuming that the initial value is equal to its expectations. Using this as a starting point, set S_2=D_1 or 150. Then subtract D_1 from D_2 to get B_2: B_2= $D_2 - D_1$ = 12. Hence, at the end of period 2, forecast for the period 3 is 162 (F_3 = 150 + 12). We choose the following values for α and β to solve this example: $\alpha = 0.3$, and $\beta = 0.5$. The actual sale for the period 3 was 159.

S_3 =0.3*(159) + (1-0.3)*(150+12) =161.1

b_3= 0.5*(161.1-150) + (1-0.5)*12=11.55

F_4=161.1+11.55=172.65.

Then all the forecasts for 12-month period are completed. The obtained results are depicted in Table 11.

Table 11: Calculation of the forecast values for 12-month period using DES

Month t	Sales D_t	S_t	B_t	Forecast F_t
1	150			
2	162	150	12	150
3	159	161.1	11.55	162
4	178	174.255	12.3525	172.65
5	195	189.1253	13.61138	186.6075
6	219	207.6156	16.05088	202.7367
7	200	216.56653	12.5009	223.6665
8	253	236.2472	16.090785	229.0674
9	300	266.6366	23.240092	252.338
10	286	288.7137	22.658596	289.8767
11	319	313.6606	23.802748	311.3723
12	332	335.8243	22.98324	337.4633

Now, we need to calculate the forecast for the period 13. We find the sum of S_{12} and b_{12}.

F_{13}=335.8243+22.98324=358.8076.

Figure 8 represents the comparison of actual and forecast sales for α=0.3, β=0.5.

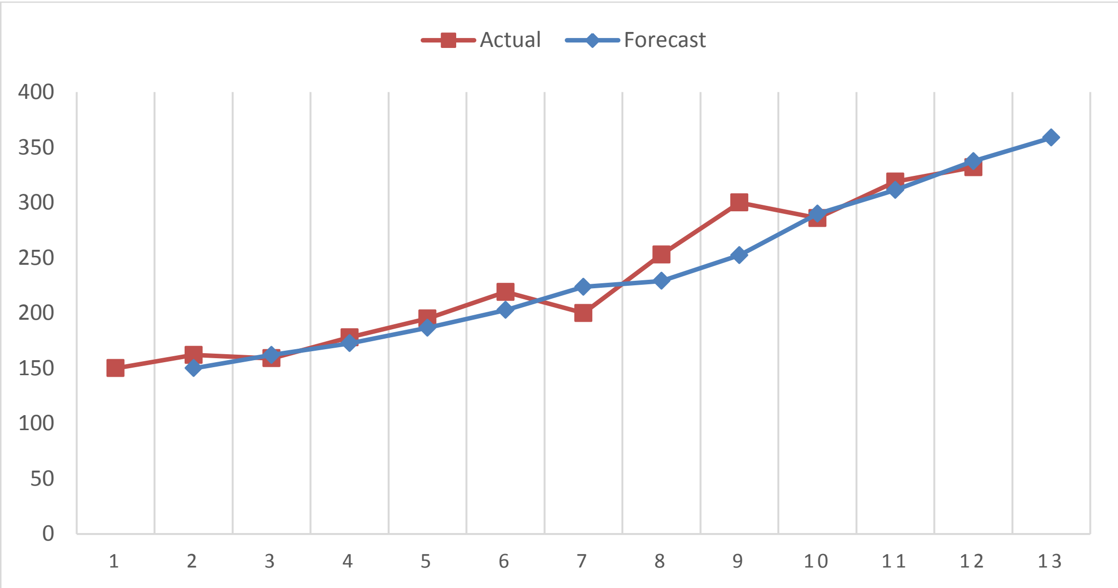

Figure 8: Comparison of actual and forecast sales for α=0.3 and β=0.5

4.3 Advantages and disadvantages of the exponential smoothing method

The exponential smoothing method has some advantages and disadvantages with respect to its capability of forecasting. Firstly the advantages of the exponential smoothing method are described:

- implementation of exponential smoothing is simpler than many other forecasting models for producing good results;

- exponential smoothing is fast computational forecasting method;

- computation using exponential smoothing method is very efficient and easier than noving average;

- exponential smoothing is a perfect model for one-period forecasting;

- limited number of data (observations) is implemented to make a forecasting;

- exponential smoothing method can easily adapt to frequent changes of all types of data in the environment to make the self-correcting;

- exponential smoothing method performs better accuracy in comparioson with moving average method, especially for a short-term time horizon;

- exponential smoothing method can be easily updated for the future realization.

At the same time, exponential smoothing method has some disadvantages which are given below:

- exponential smoothing model is extremely simple and inflexible in terms of using few statistical data for the prediction of the future value;

- exponential smoothing method is not a convenient way for forecasting for a long-term time horizon, i.e. the accuracy of the exponential smoothing is getting worse if the forecast for medium or long term time horizon is required;

- exponential smoothing is not appropriate method for using for all types of data, i.e. this method is not adaptable to particular type of data;

- the smoothing constants are chosen randomly which can’t guarantee the reasonable forecast.

Chapter 5

CONCLUSION

Selecting a good forecasting method is necessary to make the correct decisions.

In this master thesis the smoothing techniques of time series forecasting is analyzed. The short-term, medium-term, and long-term forecasts in terms of time horizons are known. The important feature of time-series forecasting is usage of recorded historical data which are collected sequentially in time to predict the future event.

The simple moving average which is a common average of values in time series, and the weighted moving average which assigns different weights to values in time series, are studied. The forecast accuracy for both techniques is calculated in order to select the appropriate forecasting method.

The basic concepts of simple and double exponential smoothing methods are discussed.

REFERENCES

[1] De Livera, A. Hyndman, R., & Snyder, R. (2010). Forecasting time series with complex seasonal patterns using exponential smoothing. *Monash University, Department of Econometrics and Business Statistics. Working paper 15/09.*

[2] Billah, M., Hyndman, R., & Koehler, A. (2003). Empirical information criteria for time series forecasting model selection. *Monash University, Faculty of Business and Economics, Department of Econometrics and Business Statistics, Working Papers.*

[3] Gould, P., Koehler, A., Ord, J., Snyder R., Hyndman, R., & Vahid-Araghi, F. (2008). Forecasting time series with multiple seasonal patterns. *European Journal of Operational Research, Volume 191, Issue 1,* pp. 207-222.

[4] Rasmussen, R. (2004).On time series data and optimal parameters. *Omega, Volume 32, Issue 2,* pp. 111-120.

[5] Sanjoy, k . (2011). Determination of Exponential Smoothing Constant to Minimize Mean Square Error and Mean Absolute Deviation. *Global Journal of Research in Engineering, Volume 11, Issue 3,* pp. 31-34.

[6] Ostertagova, E., & Ostertag, O. (2012). Forecasting using simple exponential smoothing method. *Acta Electrotechnica et Informatica, Vol. 12, No. 3,* pp. 62-66.

[7] Andrawis, R., & Atiya, A. (2009). A New Bayesian Formulation for Holt's Exponential Smoothing. *Journal of Forecasting*, *28,* pp. 218-234.

[8] Ho Kim, M., Lee, S., & Chang Lee, K. (2008). Predictive Hybrid Redundancy using Exponential Smoothing Method for Safety Critical Systems. *International Journal of Control, Automation, and Systems, vol. 6, no. 1,* pp. 126-134.

[9] Sbrana, G. (2012).Damped trend exponential smoothing: prediction and control. *Journal of Quantitative Economics, Vol. 10, No.2,* pp. 152-159.

[10] Cipra, T., & Hanzak, T. (2008). Exponential smoothing for irregular time series. *Kybernetika, Volume 44, Number 3,* pp. 385-399.

[11] Shaolin, H., Wei, Z., Li Ye & Shunxi, F. (2011). Adaptive Outlier-tolerant Exponential Smoothing Prediction Algorithms with Applications to Predict the Temperature in Spacecraft. *(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 11*, pp. 130-133.

[12] Xie, M., Hong, G.Y., & Wohlin, C. (1997). A Study of the Exponential Smoothing Technique in Software Reliability Growth Prediction. *Quality and Reliability Engineering International, Vol.13,* pp. 347-353.

[13] Gelper, S., Fried, R., & Croux, CH. (2010). Robust Forecasting with Exponential and Holt–Winters Smoothing. *Journal of Forecasting, 29,* pp. 285-300.

[14] Taylor, J.W. (2004). Smooth Transition Exponential Smoothing. *Journal of Forecasting, Vol. 23,* pp. 385-394.

[15**]** LaViola Jr, J. (2003). Double exponential smoothing: an alternative to Kalman filter-based predictive tracking. *Proceedings of the Workshop on Virtual Environments, EGVE 2003,* pp. 199-206.

[16] Taylor, W. (2008). Exponentially Weighted Information Criteria for Selecting Among Forecasting Models. *International Journal of Forecasting, Vol. 24,* pp. 513-524.

[17] Taylor, J.W. (2012). Short-Term Load Forecasting with Exponentially Weighted Methods. *IEEE Transactions on Power Systems, Vol. 27,* pp. 458-464.

[18] Yeh, A.B., Lin, D.K., Zhou, H., & Venkataramani, CH. (2003). A multivariate exponentially weighted moving average control chart for monitoring process

variability. *Journal of Applied Statistics, Vol. 30, No. 5*, pp. 507-536.

[19] Lee, H., & Apley, D.W. (2011). Improved Design of Robust Exponentially Weighted Moving Average Control Charts for Autocorrelated Processes. *Quality and Reliability Engineering International*, *Volume 27, Issue 3*, pp. 337-352.