COMPARISON OF SINGLE AND MODIFIED EXPONENTIAL SMOOTHING METHODS IN THE PRESENCE OF A STRUCTURAL BREAK

İrem EFE\(^1\)  Esin FİRUZAN\(^2\)  Berhan ÇOBAN\(^3\)

**Abstract**

The modeling of the series in the time series analysis, as well as the examination of the relations between them, is the main purpose of the future forecasting. One of the most widely used methods in the literature is exponential smoothing methods. Due to many reasons such as financial crises, natural disasters in the data production processes of the series, permanent structural changes can occur. These changes affect model parameters as well as analysis results. The main purpose of this study is to compare the predictive performances of the newly developed Modified Exponential Smoothing (MSES)(2016) methods with the simple exponential smoothing (SES) when there are structural breaks in the series with different break magnitude and different break location. Mean Absolute Error values of methods are affected by the sample size, break magnitude and location. The breaks in the data set would affect the model estimation negatively. Possible breaks' magnitude and locations should be taken into consideration in the use of the MSES method.

**Keywords:** Simple Exponential Smoothing, Modified Simple Exponential Smoothing, Structural Break, Forecast.

**JEL Classification:** C15, C22, C53

---

BASİT VE MODİFİYE EDİLMİŞ ÜSSEL DÜZELTME YÖNTEMLERİNİN YAPISAL KIRILMA OLMASI DURUMUNDA KARŞILAŞTIRILMASI

**Öz**


**Anahtar Kelime:** Modifiye Edilmiş Basit Üssel Düzeltme, Basit Üssel Düzeltme, Yapısal Kirılma, Öngörümleme

**JEL Sınıflandırması:** C15, C22, C53

---

\(^{1}\) Student of master degree, Department of Statistics, Faculty of Science, Dokuz Eylül University e-mail: ir.tanri@gmail.com

\(^{2}\) Professor Doctor, Department of Statistics, Faculty of Science, Dokuz Eylül University e-mail: esin.firuzan@deu.edu.tr

\(^{3}\) Research Assistant, Department of Statistics, Faculty of Science, Dokuz Eylül University e-mail: berhan.coban@deu.edu.tr

DOI: 10.18092/ulikidince.354325

Makalenin Geliş Tarihi (Review Date): 16-11-2017

Yayına Kabul Tarihi (Acceptance Date): 14-01-2018
1. Introduction

Time series is a collection of data values in which the values of variables from one period to the next are observed sequentially. For example, minute-by-minute foreign exchange, hourly temperatures at a weather station, daily numbers of arrivals at a bank, weekly price of a product, monthly inflation rate and annual turnover of a company. That is, time series arise whenever something is observed over time (Hydman et al. 2008:3). Using these series (\( Y_t \)), researchers can forecast the new series (\( Y_{t+1} \)) that affect and develop industry, science, economy and many other fields such as future exchange rates, air temperature, number of patients in clinics, number of customers coming to the bank etc.

Exponential smoothing (ES) methods are the most widely used techniques in forecasting due to their simplicity, robustness and accuracy as an automatic forecasting procedure. There are many studies in the literature about exponential smoothing and other forecasting methods. Some of those: Gardner (1985), Brown et. al. (1961), Brown et. Al. (1961), Billah et. al. (2006), Gardner (2006).

Although ES is a common method, there are some shortcomings, for example there is no consistent rule in the literature about the choice of initial value and smoothing constant. These are that negatively affect the accuracy of forecast. To remove for these shortcomings, Yapar (2016) develop new method, Modified Simple Exponential Smoothing. Yapar (2016) obtained strong and more accurate forecast in the series without structural breaks.

In this study, we propose to compare forecasting performance of ES that include no trend and seasonality in other words Simple Exponential Smoothing (SES) and newly developed MSES methods when there structural break are in the series.

2. Structural Break

The structural breaks which cause the interruptions of the series long and/or termed changes in their trends are expressed as outlier observations. Structural changes may occur in the data generating process of time series due to policy changes, financial crises and natural disasters. These changes in the series, without any exact definition, are generally called as the change in the model parameter (Çoban and Firuzan 2016:23-35). In such cases, it is difficult to obtain forecasts. Biased and inconsistent estimation results are obtained.

Last decades, researchers developed new methods take in to account of structural breaks in time series analysis. Primarily, Nelson and Plosser (1982:139-162) demonstrated, using statistical techniques developed by Dickey and Fuller (1979, 1981), that current shocks have a permanent effect on the long-run level of most macro-economic and financial series.

Knowledge of the break point enables the inclusion of these shocks into the model as dummy variables. Such inclusion of the break into the model as a dummy variable does not express the models which are built for the variables representing the series, but it is used to remove the effects of the shocks in the series, only.

There are a lot of structural breaks types in time series models. For instance Perron (1989) and Zivot & Andrews (1992) examined for unit root analysis on three different models. Most widely used breaks model can be defined as break in level shift model (1), break in trend (2) and break in regime shift (3). Model A is constructed by taking a structural change in the level (intercept) of the series into consideration; Model B, a structural change in the slope of the series; and Model C, taking into consideration the structural changes both in the level and the slope of the series. Models can be expressed as below:

\[
\text{MODEL A:} \quad Y_t = \mu_t + \mu_0 \phi_t + \alpha Y_{t-1} + e_t, \quad t = 1, 2, ..., n
\]
MODEL B:  
\[ Y_t = \mu_t + \alpha_Y Y_{t-1} + \alpha_x X_{t-1} \phi + e_t, \quad t = 1, 2, \ldots, n \]  

MODEL C:  
\[ Y_t = \mu_t + \mu_x \phi + \alpha_Y Y_{t-1} + \alpha_x X_{t-1} \phi + e_t, \quad t = 1, 2, \ldots, n \]  

In the models above let \( T_B \) be \( 1 < T_B < n \) and indicate the location of break, the dummy variable is defined as below:

\[ \varphi_t = \begin{cases} 
1, & \quad t > [T_Bn] \\
0, & \quad t < [T_Bn] 
\end{cases} \]

Therefore, accuracy parameter estimates and robust forecasting results obtained when model regards to structural break.

Forecasting strategies that are robust to structural breaks have earned renewed attention in the literature. They are built on assigning less weight on past observations and include forecasting with rolling window, exponential smoothing or exponentially weighted moving average and forecast pooling (Giraitis et al. 2015:401).

3. Exponential Smoothing Method (ES)

Exponential smoothing is a forecasting method that weights the observed time series values exponentially. More recent observations are weighted more than more past observations. The exponential weighting is constituted by using one or more smoothing constants, which determine how much weight is given to each observation. Exponential smoothing has been found to be most effective when the parameters describing the time series may be changing slowly over time (Bowerman and O’Connel 1993:379).

ES models assume that the time series have up to three underlying data components: level, trend and seasonality. Estimates, for the final values of these components are used to construct the forecast. An ES model can include one of five types of trend (none, additive, damped additive, multiplicative, or damped multiplicative) and one of three types of seasonality (none additive, or multiplicative). Thus, there are 15 different ES models, the best known of which are simple exponential smoothing (SES) (no trend, no seasonality), Holt’s linear model (additive trend, no seasonality) and Holt-Winters’ additive model (additive trend, additive seasonality) (Yapar 2016:3). We used the SES method in this study.

4. Simple Exponential Smoothing Method (SES)

In the SES method, assume we have observed data up to and including time \( t-I \) and we wish to forecast the next value of our time series, \( x_t \). Our forecast is denoted by \( \hat{S}_t(x) \). It can be seen that the new forecast is simply the old forecast plus an adjustment for the error in the last forecast (Hyndman et.al. 2008:20).

\[ S_t(x) = \alpha x_t + (1-\alpha) S_{t-1}(x) \]  

The forecast \( S_t(x) \) (4) is based on weighting the most recent observation \( x_t \) with a weight value \( \alpha \), and weighting the most recent forecast \( S_{t-1}(x) \) with a weight of \( (1-\alpha) \). Thus, it can be interpreted as a weighted average of the most recent forecast and the most recent observation. (Hyndman, Koehler, Ord, Snyder 2008).
\[ S_t(x) = \alpha \sum_{k=0}^{t-1} (1-\alpha)^k x_{t-k} + (1-\alpha)^t S_0 \quad \alpha \in [0,1] \] (5)

\( S_0 \) (5) represents the initial value. SES always requires a previous value of the smoothing function. When the process is started, there must be some value that can be used as the previous value \( S_{t-1} \).

\( S_t(x) \) (5) represents a weighted moving average of all past observations with the weights decreasing exponentially; hence the name “exponential smoothing” (Brown 1962:101). It can be seen that for large \( \alpha \) recent observations get more weight. The assumption of these weights:

1) \( w_t \in [0,1] \quad t = 1,....,n \)
2) \( \sum_{t=1}^{n} w_t = 1 \)
3) \( w_1 \leq w_2 \leq .... \leq w_n \)

Weights assigned by SES are non-negative and sum to unity. If \( \alpha \) is small, more weight is given to observations from the more distant past. If \( \alpha \) is large, more weight is given to the more recent observations. In the SES process, the weight given data \( k \) periods ago is \( \alpha(1-\alpha)^k \) (Selamlar 2017:18). For instance, if smoothing constant is equal to 0.2 then the weight associated with the last observation is equal to 0.2 and the weights assigned to previous observations are 0.16, 0.128, 0.1024, 0.0819, and so on. These weights appear to decline exponentially when connected by a smooth curve. More weights given to most recent observations and weights decrease geometrically with age (Çapar 2009:27).

The smoothing constant and initial values for any SES method can be estimated by minimizing the sum of the squared errors. For any SES model, they are very important. Although successful research on this subject, forecasters were unable to have a consensus on how to select smoothing constant and initial value. Determining the initial value and making mistakes in the selection of the optimum smoothing parameter adversely affect the estimation results. Different methods are applied to solve these problems. The MSES method has been developed to deal with these problems.

When we applied SES model to data set “YAF2” from the 1001 series of the M-competition data (Makridakis et. al., 1982) can we easily see that smoothing constant choice is very important.

The comparison of different \( \alpha \) levels according to YAF2 series is shown in the figure 1 below.

**Figure 1:** The comparison of different \( \alpha \) levels
When we compare different α levels, it is seen that as the α level increases, the smoothed value approaches the actual value. However, it is not right to generalize this situation with this sample.

5. Modified Simple Exponential Smoothing Method (MSES)

The MSES method, developed in recent years, calculates the most appropriate smoothing constant by giving more weight to the current observations in the series so that the calculated estimates perform better than the classical method.

The MSES method is similar to the SES method, but in this method smoothing parameter is not determined by the user as in the constant SES method. The smoothing parameters are modified so that when obtaining a smoothed value at a specific time point the weights among the observations are distributed taking into account how many observations can contribute to the value being smoothed. Therefore the smoothing parameter for this method is a function of \( t \) unlike exponential smoothing in regardless of the location of smoothed value on the time line, the observation receive weights only depending on their distances from the value being smoothed (Selamlar 2017:48).

When do not have trend and seasonal, model reduces to simple model. For the series \( X_t, t = 1, \ldots, n \), can be written as:

\[
S_t(x) = \left( \frac{m}{t} \right) x_t + \left( \frac{t-m}{t} \right) S_{t-1}(x)
\]

for \( t > m \), \( m = 0, 1, \ldots, n \).

Recognize that MSES has similar from to SES but the smoothing parameters are now dependent on the number of observations.

When the model (6) is applied recursively to all successive observations in the series, the smoothed value at time \( n \) obtained by MSES can be re-written as:

\[
S_t(x) = \sum_{k=0}^{t-(m+1)} \left( \frac{t-k-1}{m} \right) x_{t-k} + \frac{1}{(m)} s_m
\]

where \( s_m \) is the starting or initial value for MSES which can be simply the \( m^{th} \) observation or the average of the oldest \( m \) observations. It can now easily be seen that the smoothed value at time \( n \) is a weighted average of past observations and the initial value \( s_m \).

Figure 2: Comparison of smoothed value in different \( m \) levels

The weights of MSES as given in (7) can be thought of as the probabilities from a Negative Hyper-Geometric distribution with parameters \((t,m,1)\) and a random variable \( X \) (Yapar 2016 :4-5).
According to the YAF2 series, we compare the smoothed values at different m levels to obtain the following graph.

As Figure 2 also shows in Figure 1, the smoothed values approach the actual values as the m level increases.

In a series without structural break, the advantages of MSES over SES

- There is no initial problem since the MSES method selects the average of \( m^{th} \) observation and the first observation.
- Since the errors of the MSES method are smaller than the SES, the predictions are better.
- The MSES also provides more meaningful weighting schemes when the number of iterations is small.
- The smoothing coefficient is calculated according to certain parameters, not according to user preference.
- Since the variance of the new model is smaller, the model is more flexible.
- MSES achieves stronger predictions because it assigns more weight to recent observations than to SES.

The following figures summarize this situation better. In these figures the smoothed values were compared in both methods at the same level of smoothing constant.

Figure 3: The smoothed values comparison of both methods for m=1

Figure 4: The smoothed values comparison of both methods for m=2
As the smoothing constant increases, the smoothed value approaches the actual value. At the same time, the MSES method gives better results than the SES method, regardless of the smoothing value.

6. Simulation

In this study, Monte-Carlo simulations were carried out in R 3.2 program to compare the performances of SES and MSES methods when structural break occurred. Each data series that do not have seasonality and trend component was generated with 20,50,200 sample size and was repeated 1000 times and .

Since it was thought that the magnitude of breaks in the series could affect the power of the model, the performance of the model were investigated with 1,5 and 10 breaks’ magnitudes. Similarly, since it was also thought that the breaks’ occurring in different location of the series affect the power of the model, the breaks were applied in the first quarter (0,25n), the second quarter (0,50n) and in the third quarter (0,75n) ( Firuzan and Çoban 2016).

Since both methods must have equal smoothing coefficients to compare, R 3.2 program, using MSES method, we calculated alpha value by divide the value of m that get the smallest error by sample width for using in the SES method. We obtained the MAE (Mean Absolute Error) values both methods and we computed mean and variance of these values. These values are shown in table 1.

Table 1: Mean and Variance Comparison of MSES and SES Methods for Sample Size 20 under 1000 repetition

<table>
<thead>
<tr>
<th>n,l,m</th>
<th>MSES error mean</th>
<th>SES error mean</th>
<th>MSES error var</th>
<th>SES error var</th>
</tr>
</thead>
<tbody>
<tr>
<td>20,5,1</td>
<td>0,8907</td>
<td>0,9431</td>
<td>0,0228</td>
<td>0,0461</td>
</tr>
<tr>
<td>20,10,1</td>
<td>0,9132</td>
<td>0,9202</td>
<td>0,0231</td>
<td>0,0347</td>
</tr>
<tr>
<td>20,15,1</td>
<td>0,9096</td>
<td>0,9212</td>
<td>0,0238</td>
<td>0,0396</td>
</tr>
<tr>
<td>20,5,5</td>
<td>1,1544</td>
<td>1,4265</td>
<td>0,0333</td>
<td>0,049</td>
</tr>
<tr>
<td>20,10,5</td>
<td>1,2292</td>
<td>1,23</td>
<td>0,0398</td>
<td>0,0328</td>
</tr>
<tr>
<td>20,15,5</td>
<td>1,2727</td>
<td>1,1775</td>
<td>0,0465</td>
<td>0,0337</td>
</tr>
<tr>
<td>20,5,10</td>
<td>1,4311</td>
<td>2,1207</td>
<td>0,0354</td>
<td>0,0761</td>
</tr>
<tr>
<td>20,10,10</td>
<td>1,5168</td>
<td>1,6276</td>
<td>0,0429</td>
<td>0,0339</td>
</tr>
<tr>
<td>20,15,10</td>
<td>1,5587</td>
<td>1,496</td>
<td>0,0488</td>
<td>0,0358</td>
</tr>
</tbody>
</table>
First, when we investigate our data set with a sample size of 20, it is seen that when the breaking magnitude is 1, the error averages are similar in both methods. As the magnitude of the break increases, the average of the errors increases. When the severity of breaking is 5 and 10, while the MSES gives better results in the first quarter of the series, SES is better in the last quarter of the series.

As can be easily seen on the figure 7 and the table 1, the variance of the MSES method increases as the magnitude of the break increases by 20 sample size. In the SES method, the variances of errors in all breaking magnitudes show sudden jumps. When we compare the two methods, it is generally seen that the errors’ mean and variance of the MSES are smaller than SES. If break is at the first quadnart of series, MSES is better; if break is at the last quadnart of series, SES is better.
When we research the data set with a sample size of 50, we can see that there is a difference between the MSES and the average of the errors of the SES methods when the magnitude of the break is 1. As with 20 sample size, as the magnitude of break increases, the average of the errors in both methods increases. It is observed that the error of the MSES method is smaller when the break series is at the beginning, and the error of the SES method is smaller when the break series is at the end.

Figure 8: Comparison of the Mean of the Errors of MSES and SES Methods at 50 Sample Size

It is observed in figure 8 that the error’ mean of the MSES method is smaller when the break series is at the beginning, and the error’ mean of the SES method is smaller when the break series is at the end.

Figure 9: Comparison of the Variance of the Errors of MSES and SES Methods at 50 Sample Size

If we compare the variances of the errors of both methods at 50 sample size, when the magnitude of the breaks are 5 and 10, and at the same time in the middle and last quadrant of the break series, the MSES method can give worse results. When we investigate in general, it is seen that the errors of the MSES method are smaller than the SES method. The following figure 9 and table 2 summarize the situation.

The following table 3 and, summarizes all the analysis in the study.
Table 3: Mean and Variance Comparison of MSES and SES Methods for all of Sample Size under 1000 repetition

<table>
<thead>
<tr>
<th>Mean of the Errors</th>
<th>Variance of the Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>n,l,m</td>
<td>MSES</td>
</tr>
<tr>
<td>20,5,1</td>
<td>0.8907</td>
</tr>
<tr>
<td>20,10,1</td>
<td>0.9132</td>
</tr>
<tr>
<td>20,15,1</td>
<td>0.9096</td>
</tr>
<tr>
<td>20,5,5</td>
<td>1.1544</td>
</tr>
<tr>
<td>20,10,5</td>
<td>1.2292</td>
</tr>
<tr>
<td>20,15,5</td>
<td>1.2727</td>
</tr>
<tr>
<td>50,10,1</td>
<td>0.855</td>
</tr>
<tr>
<td>50,25,1</td>
<td>0.8806</td>
</tr>
<tr>
<td>50,40,1</td>
<td>0.8803</td>
</tr>
<tr>
<td>50,10,5</td>
<td>0.9976</td>
</tr>
<tr>
<td>50,25,5</td>
<td>1.0891</td>
</tr>
<tr>
<td>50,40,5</td>
<td>1.1415</td>
</tr>
<tr>
<td>50,10,10</td>
<td>1.1152</td>
</tr>
<tr>
<td>50,25,10</td>
<td>1.2244</td>
</tr>
<tr>
<td>50,40,10</td>
<td>1.2818</td>
</tr>
<tr>
<td>200,50,1</td>
<td>0.8336</td>
</tr>
<tr>
<td>200,100,1</td>
<td>0.8475</td>
</tr>
<tr>
<td>200,150,1</td>
<td>0.8555</td>
</tr>
<tr>
<td>200,50,5</td>
<td>0.9139</td>
</tr>
<tr>
<td>200,100,5</td>
<td>0.9622</td>
</tr>
<tr>
<td>200,150,5</td>
<td>0.9900</td>
</tr>
<tr>
<td>200,50,10</td>
<td>0.9846</td>
</tr>
<tr>
<td>200,100,10</td>
<td>1.0392</td>
</tr>
<tr>
<td>200,150,10</td>
<td>1.0846</td>
</tr>
</tbody>
</table>

Figure 12: Comparison of the Mean of the Errors of MSES and SES Methods at all of them Sample Size

When the Mean Absolute Error variance is examined, it is observed that it is decreasing meanwhile the sampling size increases parallel to the average results. When examined generally, error variances of the SES method rates in a wider range. When the break is at the end and firm (m=10) the MSES method’s error appears to be higher than the SES method’s.
The purpose of this study compare the method of MSES (Yapar2016), which was develop to increase power and accuracy, help cope with shortcomings in smoothing techniques, with the SES model, which has been widely used for many years.

In order to make this comparison, the series are produced for different scenarios with 1000 replications by Monte-Carlo simulation method in R 3.2. program and obtained the mean and variance of MAE values for each scenario.

The mean absolute error values under the different sample size, different break magnitude and different break locations of the MSES and the SES methods are shown in the tables and graphics. When the mean absolute errors change under the breaks is examined, it is observed that it decreases for both of the tests while the sample size increases and it decreases while the break magnitude increases. Considering both of the methods, while almost the same results are obtained when the breaks is low \((m=1)\), the MSES method’s error appears to be lower than the SES method’s while the magnitude of the break increases. Regardless of this superiority of the MSES method, when the magnitude of the break is medium and high, the location of the break is in the middle and at the end, the error of the MSES method is higher than the SES method’s. It can be said that due to the MSES method’s concentrating on the recent observations, the effect of the firm breaks affects the model estimation negatively.

Mean Absolute Error’s mean and variance are affected by the sample’s size, break’s magnitude and location. The MSES value are lower than the classic SES method in addition to the MSES method usage advantages like initial value and smoothing constant determination. But, the breaks in the data set would affect the model estimation negatively. Especially, the possible breaks being in the end of the series increases the error of estimation of the method. Possible breaks’ magnitude and locations should be taken into consideration in the use of the MSES method.

Reference


