MATHEMATICAL MODELING AND COMPUTING, Vol. 9, No. 1, pp. 50–56 (2022)

Exponential smoothing constant determination to minimize the forecast error

Noor N. A. M., Rahman N. H. A.

Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

(Received 7 July 2021; Accepted 16 November 2021)

One of the fundamental issues in exponential smoothing is to determine the smoothing constants. Researchers usually use the determination available in the statistical software. However, the result may not able to minimize the forecast error. For this study, the optimal values of smoothing constant are based on minimizing the forecast errors, mean absolute percentage error (MAPE) and root mean squared error (RMSE). The double exponential smoothing method or Holt's method is chosen where two constant values must identify specifically the level and trend estimate, respectively. The real data set of tourism emphasize the number of international tourists visit Malacca from year 2003 to 2016 has been studied. The result shows that the values of level and trend obtained from this analysis is small and close to zero. This indicates that the level and trend react slowly towards the data. In addition, simulation also have been computed using the random walk model. The result suggested, by using optimal result available by statistical software is not recommended since the obtained smoothing constants do not minimize the forecast error.

Keywords: *smoothing constant, forecasting, trend, level, error minimizing.*

2010 MSC: 62M10 **DOI:** 10.23939/mmc2022.01.050

1. Introduction

Forecasting using mathematical model is an alternative planning tool to develop assumptions about the future uncertainty [1]. It is important in many types of organizations since the prediction of future events incorporated into decision making process. Time series analysis is one of the forecast approach that uses the information of data being taken sequentially in time. One of the methods is known as exponential smoothing one. This method enables to find the forecast values by averaging past values of a series with in a decreasing series of weight. The most recent observation receives the largest weight while the older observations receive the lesser weight.

The exponential smoothing method arise when Robert G. Brown, an Operations Research (OR) Analyst for the US Navy in 1944 used a tracking model for fire control information on the location of submarines [2]. The simple exponential smoothing was used to estimate the velocity and the lead angle for firing depth chargers from destroyers. After six years, he broaden this method from continuous to discrete time series and developed methods for trend and seasonality. One of his first applications was forecasting spare parts demand in the US Navy inventory system. In 1956, Brown presented his work on exponential smoothing at a conference and then formed his first book on inventory control [3]. He also developed the ideas in his second book emphasized on Smoothing, Forecasting and Prediction of Discrete Time Series [4]. Then, Charles C. Holt with support from Office of Naval Research (ONR) performed similar Brown method in additive trends and different method for seasonal data. Holt's ideas have been widely spread among the public after his original work went published in a known journal, International Journal of Forecasting in 2004 [5].

This work was supported by grant 9587700.

The exponential smoothing method widely used in many fields especially in statistics, management science, marketing and business operation [6]. Three classifications of the exponential smoothing method are: simple exponential smoothing (SES), double exponential smoothing (Holt's method) and triple exponential smoothing (Winter's method). The SES method is used for forecasting a stationary series, Holt's method when the data consist of trend pattern and Winter's method is an exponential smoothing approach for handling seasonal data.

Smoothing constants in the exponential smoothing are the assign weight in each of the data pattern. These constant usually denoted as α, β and γ, represent the level, the trend and the seasonality of the series respectively. These constants are between 0 and 1, where small constant is suitable for stable variations while large smoothing constant value is required when a rapid response to a real change in the pattern of observations exists. Thus, assigning an optimal constant value is crucial step in order to minimize the forecast errors. However, these exponential smoothing constants do not have a general rule and specification on how they should be selected to obtain the best forecast values.

A lot of study have been conducted to find the smoothing constant. Realistically, more constant values for exponential smoothing should be experimented to determine the optimal smoothing constant [7,8]. In addition, [9] mentioned that no research paper has found the optimal value of exponential smoothing constant. [10] agreed that smoothing constants are the key to successful forecasting but there are no consistent guidelines in the forecasting literature on how they should be selected. There are no empirical studies in time series that the smoothing constant should be selected between 0.7 and 0.9 and also no theoretical reasons seem to be available for the discussion [11]. Because of these issues, researcher usually uses the available software to find the optimal smoothing constant in exponential smoothing method.

2. Material and method

2.1. Smoothing method

Smoothing method gives the largest weight to the most recent observation meanwhile the older observations receive the lesser weight. Double exponential smoothing method or Holt's method use smoothing constant in level and trend components to provide pattern information before generate the forecast. The equations can be written as follows:

$$L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + T_{t-1}), \tag{1}$$

$$T_t = \beta(L_t - L_{t-1} + (1-\beta)T_{t-1}), \tag{2}$$

$$\hat{Y}_{t+p} = L_t + pT_t, \tag{3}$$

where

L_t = New smoothes value or estimate of current level;

α = Smoothing constant for the level, [0,1];

Y_t = New observation or actual value of series in period t;

β = Smoothing constant for the trend estimate, [0,1];

T_t = Trend estimate;

p = Periods to be forecast in the future;

$\hat{Y}_{t+p}$ = Forecast p periods into the future.

The smoothing constant is a value between 0 and 1. For this study, the combination will vary from 0.1 to 0.9 and each of α will be pairing with the fixed β and vice versa. All the possible combinations are examined in order to obtain the constant values that minimize MAPE and RMSE. However, there are restrictions for α and β due to state space models where the interval for the α remains the same but for β lies between 0 and the value of α. The interval will be $0 < \alpha < 1$ and $0 < \beta^* < \alpha$, where the restriction for β denoted as β^* instead of β on the condition of $\beta = \alpha\beta^*$ [12]. As in the study, the term use is still parameter β even it is actually β^* in restricted condition. The restrictions are usually stricter than necessary (although in a few cases they are not restrictive enough). As example, the combination of α and restricted β such as $\alpha = 0.2$, $\beta = 0.1$ still can be computed but for the combination $\alpha = 0.1$, $\beta = 0.2$ can not be computed as the parameter of β is out of the range.

2.2. Forecast accuracy

Forecasting accuracy plays an important role when deciding among several alternatives. This study aim is to find the forecast error which is the deviation between the actual value and the forecast value of a given period. Most of the textbooks recommend the use of the mean absolute percentage error (MAPE) and it was the primary measure in M-competition [13]. In addition, MAPE is often used in practice because of its very intuitive interpretation in terms of relative error [14]. In other side, [13] also stated that MSE and RMSE have been popular, largely because of their theoretical relevance in statistical modeling. Thus, this study uses MAPE and RMSE as the measures of forecast accuracy. The formulae for both measurements are as follow:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| 100\%; \quad y_t \neq 0; \tag{4}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}, \tag{5}$$

where y_t is the actual value while $\hat{y}_t$ is the forecast value.

2.3. Simulation

A random walk is a time series model in which the value of an observation in the current time period is equal to the value of the observation in the previous time period plus the value of an error term from a fixed probability distribution. Random walk is a non-stationary series. It is the simplest and yet important model in time series forecasting. This model assumes that in each period, the variable takes a random step away from its previous value, and the steps are independently and identically distributed (i.i.d). Two common types of random walk model are random walk with no drift and random walk with drift. For this study, the model with no drift will be used and the value of generated data is from random normal distribution with mean zero and variance one. The random walk formula is given as:

$$Y_t = Y_{t-1} + \varepsilon_t, \tag{6}$$

where ε_t is identically distributed with mean zero and variance one for $t = 1, 2, 3, \ldots, t$. Note that the parameter 1 on the Y_{t-1} renders the series is non-stationary since the random walk does depend on t as shown below,

$$\begin{aligned}\text{Var}(Y_t) &= \text{Var}(Y_{t-1} + \varepsilon_t)\\ &= \text{Var}(\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_t)\\ &= \sigma^2 + \sigma^2 + \ldots + \sigma^2\\ &= t\sigma^2.\end{aligned} \tag{7}$$

A simulation study is conducted to further identify the optimal value of smoothing constant α and β for the minimum MAPE and RMSE. The steps used to generate data and obtain optimal smoothing constant for the simulation experiment are as follows:

1. The random data is generated and replicate using random walk model where the observation data and the error term is from normal distribution with mean zero and variance one.
2. The series of data is divided into two sets of data which are in-sample and out-sample data.
3. Ets( ) function is used in these loops to give the output of estimated parameter α and β. For forecast values, forecast( ) function is used.
4. Obtain the MAPE and RMSE.
5. Step (4) to (5) are repeated for the different values of α and β with the increament of 0.1. The combination values of α and β also will be examined in addition to get the minimum forecast error.

3. Result and discussion

The tourism data has been used emphasizing the number of international visitors in Malacca from 2003 to 2016. The data is divided into two parts, in-sample data consists of 13 years (2003–2015) for data training and 2016 as out-sample data used to validate and measure the performance of the combination smoothing constant for Holt's method. The analysis was performed by using the R software.

The in-sample data of 13 years or specifically 156 months was implemented in ets( ) function to find the best smoothing constant based on the build function in the R software. The computed model gave the output of α and β, initial values and model criterion. The program used maximizing the likelihood in which this is an alternative use besides minimizing the sum of squared errors [5]. To obtain the point forecast from the ETS model, the forecast( ) function was used to forecast the number of tourists for the next year or 12 months ahead. This study used the same program but with the different combination of α and β, where the main objective was to find the optimal smoothing constants in which can minimize the MAPE and RMSE.

The result of different combination of α and β and the maximum likelihood in R was shown in Table 1. From the result, the value of MAPE and RMSE had been calculated for every estimated α and β and also the combination of α and β. For RMSE, the lowest value obtained among all the combination was 65.8258 while for MAPE was 1.5719. The value of RMSE was quite large because the values were influenced by the larger number of tourists data set which in million. Both RMSE and MAPE gave the same value of optimal smoothing constant which were $\alpha = 0.1$ and $\beta = 0.0156$. The value of β was close to zero showing that the level reacts weakly to each of the new observation. Meanwhile, the value of β was approximately close to zero indicating that the trend was changing steady over time.

For the simulation experiment, the same procedure was conducted but with replication to obtain the significant optimal values of smoothing constants for N times. The simulated data was obtained by using the random walk model with no shift. Each set of the data scattered unevenly in the direction of ups and downs indicating the random walk series can be either in positive and negative ways. The simulation data was proceeded with the parameters estimation and forecasting. The smoothing constants of α and β were estimated by maximizing the likelihood, the default setting in R software. To achieve the aim of obtaining the optimal smoothing constants, the values of α and β combination were conducted until produced the smallest forecast errors, MAPE and RMSE. Table 2 and Table 3 showed the minimum MAPE and RMSE obtained after all combinations of α and β had been tested.

From Table 2 and Table 3, the result showed that $\alpha = 0.9$ and $\beta = 0.1$, were the most suitable smoothing constants for the random walk model that has random data pattern. Moreover, the values of α and β also indicate the significant result with the simulation. The value of 0.9 represents the higher level exists in most of the data while the value 0.1 represents the lower trend exists in the most of the simulation data.

Table 1. The value of RMSE and MAPE for forecasting tourism data in 2016.

Combination of α and β	α	β	RMSE	MAPE
α and β – MLE	0.4291	0.0001	75.5693	17.4624
α — trial and error, β — MLE	0.1	0.0156	65.8258	1.5719
	0.2	0.0001	69.1678	15.7610
	0.3	0.0001	72.0707	16.7427
	0.4	0.0001	74.8997	17.3402
	0.5	0.0001	77.1911	17.7402
	0.6	0.0001	79.4548	18.0934
	0.7	0.0001	81.7895	18.4284
	0.8	0.0001	84.1406	18.7440
	0.9	0.0001	86.4717	19.0674
α — MLE, β — trial and error	0.5447	0.1	79.2974	17.9452
	0.694	0.2	86.9615	19.1683
	0.7726	0.3	97.2834	21.3216
	0.8045	0.4	108.9587	24.7715
	0.8143	0.5	122.2269	28.7330
	0.8116	0.6	136.7576	32.7673
	0.7999	0.7	151.9220	36.7732
	0.8	0.8	163.1919	39.6684
	0.9	0.9	143.7109	34.6329
α — trial and error, β — fixed 0.3	0.3	0.3	73.2408	16.4621
	0.4	0.3	73.9296	1.6874
	0.5	0.3	79.2197	17.8355
	0.6	0.3	86.6489	19.1219
	0.7	0.3	93.5534	20.5563
	0.8	0.3	98.3378	21.5704
	0.9	0.3	100.4639	22.0576
α — trial and error, β — fixed 0.4	0.4	0.4	73.0408	16.8822
	0.5	0.4	84.0014	18.5475
	0.6	0.4	95.8234	21.0511
	0.7	0.4	104.4592	23.4388
	0.8	0.4	108.8574	24.7430
	0.9	0.4	109.3304	24.8302
α — trial and error, β — fixed 0.5	0.5	0.5	99.9317	22.2716
	0.6	0.5	113.7586	26.3788
	0.7	0.5	121.0287	28.4291
	0.8	0.5	122.4290	28.7952
	0.9	0.5	119.3565	27.8763
α — trial and error, β — fixed 0.6	0.6	0.6	139.8725	33.5526
	0.7	0.6	141.7919	34.0973
	0.8	0.6	137.5508	32.9809
	0.9	0.6	129.3045	30.7175
α — trial and error, β — fixed 0.7	0.7	0.7	163.6101	39.7152
	0.8	0.7	151.9071	36.7694
	0.9	0.7	137.5748	32.9873
α — trial and error, β — fixed 0.8	0.8	0.8	163.1921	39.6684
	0.9	0.8	142.7098	34.3656
α — trial and error, β — fixed 0.9	0.9	0.9	143.7109	34.6329

Table 2. The frequency of α and β based on minimizing MAPE.

α	Frequency	β	Frequency
0.1	14	0.1	19
0.2	7	0.2	14
0.3	9	0.3	5
0.4	7	0.4	4
0.5	7	0.5	5
0.6	6	0.6	6
0.7	12	0.7	6
0.8	8	0.8	6
0.9	30	0.9	11
		0.0001	12
		Others (< 0.1)	12

Table 3. The frequency of α and β based on minimizing RMSE.

α	Frequency	β	Frequency
0.1	16	0.1	18
0.2	6	0.2	12
0.3	5	0.3	7
0.4	10	0.4	4
0.5	7	0.5	6
0.6	5	0.6	2
0.7	15	0.7	7
0.8	9	0.8	10
0.9	27	0.9	9
		0.0001	13
		Others (< 0.1)	12

4. Conclusion

Exponential smoothing is one of the useful methods in forecasting and this method is affected by several factors such as smoothing constants in order to produce better forecast. This study identifies the optimal values of smoothing constants by minimizing the forecast errors which are mean absolute percentage error (MAPE) and root mean squared error (RMSE). The research considered the double exponential smoothing (Holt's method) as the aim was to find the optimal value of α and β for trend data.

Based on the result, the default setting in statistical software will not give the best result in minimizing the forecast error. It was shown in both, real and simulation study. For tourism data, the optimal values of smoothing constant, α and β were 0.1 and 0.0156 respectively. Both smoothing constants were close to zero and this indicates the level and trend components were hardly changed over time and respond slowly to a new observation. This shows that the smoothing constants are mostly affected by the recently data pattern.

For simulation data, the random walk model produced the value of $\alpha = 0.9$ while $\beta = 0.1$. The value of α was close to one which indicates the level reacts strongly towards the new observation while the value of β was close to zero which indicates the trend reacts weakly changing over time. Thus, the random walk model produced high value in level and low value in trend as the model takes random step away from the previous data and most of the data scattered constantly.

As the conclusion, the smoothing constants computed directly from the default setting cannot be taken as the optimal values of smoothing constant. This is because, the optimal value may not support the main objective in forecasting which is to minimize the forecast error. For future recommendation, this study will extend to build algorithm and study the effect of the initial values used in the exponential smoothing computation to produce high accuracy in forecasting.

[1] Musii R. S., Nakonechnyy A. Y. Mathematical model for temperature estimation forecasting of electrically conductive plate elements under action of pulsed electromagnetic radiation of radio-frequency range. Mathematical Modeling and Computing. **8** (1), 35–42 (2021).

[2] Gardner Jr. E. S. Exponential smoothing: The state of the art-part II. International Journal of Forecasting. **22** (4), 637–666 (2006).

[3] Brown R. G. Statistical forecasting for inventory control. McGraw/Hill (1959).

[4] Brown R. G. Smoothing, forecasting and prediction of discrete time series. Courier Corporation (2004).

[5] Hyndman R., Koehler A. B., Ord J K., Snyder R. D. Forecasting with exponential smoothing. Springer Science & Business Media (2008).

[6] Meira E., Oliveira F. L. C., Jeon J. Treating and Pruning: New approaches to forecasting model selection and combination using prediction intervals. International Journal of Forecasting. **37** (2), 547–568 (2021).

[7] Krajewski L., Ritzman L., Malhotra M. Operations management: processes and value chains. Pearsons Education (2007).

[8] Karmaker C. Determination of optimum smoothing constant of single exponential smoothing method: A case study. International Journal of Research in Industrial Engineering. **6** (3), 184–192 (2017).

[9] Paul S. K. Determination of exponential smoothing constant to minimize mean square error and mean absolute deviation. Global Journal of Research In Engineering. **11** (3), (2011).

[10] Ravinder H. V. Forecasting with exponential smoothing whats the right smoothing constant? Review of Business Information Systems. **17** (3), 117–126 (2013).

[11] Mu'azu H. New approach for determining the smoothing constant α of a single exponential smoothing method. International Journal of Science and Technology. **3** (11), 717–727 (2014).

[12] Hyndman R. J., Khandakar Y. Automatic Time Series Forecasting: The forecast Package for R. Journal of Statistical Software. **27** (3), 1–22 (2008).

[13] Hyndman R. J., Koehler A. B. Another look at measures of forecast accuracy. International Journal of Forecasting. **22** (4), 679–688 (2006).

[14] De Myttenaere A., Golden B., Le Grand B., Rossi F. Mean absolute percentage error for regression models. Neurocomputing. **192**, 38–48 (2016).

Визначення константи експоненціального згладжування для мінімізації похибки прогнозу

Нур Н. А. М., Рахман Н. Х. А.

Кафедра математики та статистики,
Факультет природничих наук,
Університет Путра Малайзія,
43400 UPM Серданг, Селангор, Малайзія

Однією з фундаментальних проблем у експоненціальному згладжуванні є визначення констант згладжування. Дослідники зазвичай використовують визначення, яке доступне в статистичному програмному забезпеченні. Однак, результат не може мінімізувати похибку прогнозу. Для цього дослідження оптимальні значення константи згладжування базуються на мінімізації прогнозних помилок, середньої абсолютної процентної похибки (САПП) та середньоквадратичної похибки (СКП). Обрано подвійний експоненціальний метод згладжування або метод Хольта, де два постійних значення повинні бути визначені на рівні та оцінці тренда, відповідно. Досліджувався реальний набір туристичних даних, в якому виділено кількість міжнародних туристів, які відвідали Малакку з 2003 року до 2016 року. Результат показує, що значення рівня та тренду, які отримані у результаті цього аналізу, є невеликими та близькими до нуля. Це вказує на те, що рівень і тренд повільно реагують на дані. Крім того, симуляція також була розрахована за допомогою моделі випадкового блукання. Результат показує, що використанням оптимального результату, який доступний статистичним програмним забезпеченням, не рекомендується, оскільки отримані константи згладжування не мінімізують похибку прогнозу.

Ключові слова: *згладжуюча константа, прогнозування, тренд, рівень, мінімізація помилок.*