

Article

Setting the Initial Value for Single Exponential Smoothing and the Value of the Smoothing Constant for Forecasting Using Solver in Microsoft Excel

Wannaporn Junthopas¹ and Chantha Wongoutong^{2,*} ¹ Department of Statistics, Faculty of Science, Khon Kaen University, Khon Kaen 40002, Thailand² Department of Statistics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand

* Correspondence: fscictw@ku.ac.th

Abstract: Although single exponential smoothing is a popular forecasting method for a wide range of applications involving stationary time series data, consistent rules about choosing the initial value and determining the value for the smoothing constant (α) are still required, because they directly impact the forecast accuracy. The purpose of this study is to mitigate these shortcomings. First, a new method for setting the initial value by weighting is derived, and its performance is compared with two other traditional methods. Second, the optimal (α) was automatically solved using Solver in Microsoft Excel, after which α was determined by minimizing the mean squared error (MSE). This was accomplished by comparing the α from Solver with step search by setting the smoothing constant by varying its value from 0.001 to 1 in increments of 0.001 and then choosing the optimal α value from this range that has the lowest MSE. The experimental results show that α from Solver and the optimal α with step search are not different, and the initial value set by the proposed method outperformed the existing ones regarding the MSE.

Keywords: smoothing constant; initial value; single exponential smoothing; Solver in Microsoft Excel



Citation: Junthopas, W.; Wongoutong, C. Setting the Initial Value for Single Exponential Smoothing and the Value of the Smoothing Constant for Forecasting Using Solver in Microsoft Excel. *Appl. Sci.* **2023**, *13*, 4328. <https://doi.org/10.3390/app13074328>

Academic Editor: Jianbo Gao

Received: 16 January 2023

Revised: 17 March 2023

Accepted: 20 March 2023

Published: 29 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Exponential smoothing is one of the most widely used techniques in forecasting due to its simplicity, robustness, and accuracy in an automatic forecasting procedure. Hence, it has been widely used for forecasting future values from time series data [1–5]. One of the basic ideas behind using weighted averages of past observations is that more recent observations carry more weight when determining forecasts than observations in the distant past. Exponential smoothing is divided into three types: single, double, and triple. Single exponential smoothing (SES) is used on data with a stable fluctuating pattern; double exponential smoothing is used on data with a trend pattern; and triple exponential smoothing is used on data with both trend and seasonal patterns [6].

However, there are two problems with exponential smoothing forecasting methods. The first is choosing a suitable value for the smoothing constant (α) and the second is setting a suitable initial value. The forecaster must determine one or more parameters in exponential smoothing for assigning exponentially decreasing weights as the observations become older, because future events usually depend more on the most recent data [7]. Hence, the value of α is important for successful forecasting via exponential smoothing. Nevertheless, there need to be consistent guidelines on how they should be selected. In general, the value of α is selected by applying a suitable calculation process, many of which have been tried out in the past. Many statisticians have recommended that the value of α should be kept small (in the 0.1 to 0.3 range) to minimize the forecasting error function [8,9]. Moreover, Paul [10] recommended selecting the value for α using a nonlinear optimizer [11,12]. However, it is often the case that the value is outside the recommended range. Solver in Microsoft Excel has become increasingly popular as

a nonlinear optimizer [13,14]. This optimization can be performed using the Solver in Microsoft Excel, and many textbooks, such as Chopra and Meindl [15] and Balakrishnan et al. (2013) [16], mentioned and illustrate this approach.

As is well known, SES models are recursive, and thus, an initial value is required to feed the model to obtain the most accurate prediction [6,17]. Therefore, the choice of this value has an important impact on the forecasting performance. Many researchers have proposed methods to determine a suitable initial value for the smoothing parameter in SES models. Brown's [18] original suggestion was simply to use the mean of the data for the initial value while others have suggested using the first observation or the average of the first three observations as the initial value. Ledolter and Abraham [17] recommended backcasting to obtain the initial value. In this method, the smoothing algorithm estimated the initial value by going backward in the series. Another approach with a limited number of data points is to use the Bayesian method to combine the prior estimate of the level with the average of the available data [19–21]. In most cases, the first actual value is considered to be the initial value for the smoothing parameter when using SES [22]. Furthermore, the average of the first five or six observations can be used as the initial value [23], which is used in many statistical packages, such as Minitab, and has been used to set the initial value for SES in many forecasting approaches [24–27].

The accuracy of SES may vary depending on the chosen value of α . Even though much research has been conducted on this subject, forecasters have not been able to reach a consensus on how to select the value for α or the initial value. Therefore, in this study, the effects of α and the initial value are clarified, after which, various methods for setting the initial value are investigated. Their performances and searching for the optimal value for α were then investigated based on the mean squared error (MSE) values, a popular metric that is commonly used for comparing forecasting techniques.

The remaining parts of this paper are as follows: The theoretical framework is covered in Section 2. The proposed methods for setting the initial value are presented in Section 3. Experimentation to show the efficacy of the proposed methods is reported in Section 4. The results and a discussion are provided in Section 5. Finally, the conclusions and remarks are presented in Section 6.

2. Theoretical Framework

2.1. SES

This is a suitable forecasting method based on data without a trend or a seasonal component [18] (Brown, 1959). In SES, α ranging from 0 to 1 is used as a weight to indicate how similar the current and previous observations are: a weight value close to 1 emphasizes the current value whereas one close to 0 emphasizes the previous one [28]. The formula for SES is defined as the following:

$$S_{t+1} = \alpha y_t + (1 - \alpha)S_t, \quad (1)$$

where S_t is the smoothed value of the time series at time t , α is the smoothing constant for $\alpha \in [0, 1]$, and y_t is the actual value of the time series at time t .

In the case of SES, the smoothed statistic is the forecasted value, which is derived as

$$F_{t+1} = \alpha y_t + (1 - \alpha)F_t, \quad (2)$$

where F_{t+1} and F_t are the forecasted values of time series at time $t + 1$ and t , respectively.

Expanding Equation (2) means that

$$F_t = \alpha y_{t-1} + (1 - \alpha)F_{t-1},$$

$$F_{t-1} = \alpha y_{t-2} + (1 - \alpha)F_{t-2},$$

$$F_{t-2} = \alpha y_{t-3} + (1 - \alpha)F_{t-3},$$

$$F_{t-3} = \alpha y_{t-4} + (1 - \alpha)F_{t-4},$$

etc.

By substituting each of these into the following equation, we obtain

$$\begin{aligned} F_{t+1} &= \alpha y_t + (1 - \alpha)F_t = \alpha y_t + (1 - \alpha)[\alpha y_{t-1} + (1 - \alpha)F_{t-1}] \\ &= \alpha y_t + \alpha(1 - \alpha)y_{t-1} + (1 - \alpha)^2 F_{t-1} \\ &= \alpha y_t + \alpha(1 - \alpha)y_{t-1} + (1 - \alpha)^2 [\alpha y_{t-2} + (1 - \alpha)F_{t-2}] \\ &= \alpha y_t + \alpha(1 - \alpha)y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} + (1 - \alpha)^3 F_{t-2} \\ &= \alpha y_t + \alpha(1 - \alpha)y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} + (1 - \alpha)^3 [\alpha y_{t-3} + (1 - \alpha)F_{t-3}] \\ &= \alpha y_t + \alpha(1 - \alpha)y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} + \alpha(1 - \alpha)^3 y_{t-3} + (1 - \alpha)^4 F_{t-3}. \end{aligned}$$

Generalizing the above equation provides

$$F_{t+1} = \sum_{j=0}^{t-1} \alpha(1 - \alpha)^j y_{t-j} + (1 - \alpha)^t F_1. \quad (3)$$

Equation (3) represents the weighted moving average of all past observations with the weights decreasing exponentially; i.e., exponential smoothing [29]. It can be seen that large recent observations are assigned a larger weight, which can be interpreted as the weighted average for the most recent forecast from the most recent observations [1].

The assumption of these weights are as follows:

$$w_t \in [0, 1] \text{ where } t = 1, 2, \dots, n$$

$$\sum_{t=1}^n w_t = 1 \text{ where } w_1 \leq w_2 \leq \dots \leq w_n$$

For forecasting from one step ahead at time $t + 1$, all of the observations in the series y_1, y_2, \dots, y_t are used with an exponential weighting scheme that assigns the maximum weight to the most recent observation. Moreover, the weights decline systematically as the observations that are still included become older. Weighting in an exponential smoothing technique requires a given value of α , and the forecasted values vary depending on the value of this constant, as does the forecasting error. Hence, the critical process is determining the value of α constrained by $\alpha \in [0, 1]$. Since mistakes in selecting the optimal value adversely affect the estimated results, obtaining the value of α is conducted via trial and error while minimizing the sum of the squared error. As the value of α controls the smoothing level adaptation speed, choosing the initial value must be conducted carefully, because a poor choice will require more time for the recursive smoothing formula to adapt and for its effect to dissipate. The impact is very tangible among small time series and those with a small value [30]. Considering Equation (3), the initial value F_1 needs to be found via a prediction algorithm, which plays a critical role in computing all of the subsequent SES. In practice, different methods are applied to solve these problems.

2.2. Setting the Initial Value for SES

Because SES requires the previous forecasted value at each stage, it must be apparent how to start the method. Different initialization and estimation methods for setting the initial value for SES have been discussed in the literature. Still, the state-of-the-art method is to estimate the optimum alpha and the initial value together by minimizing some loss function [31]. Typically, MSE is used, minimizing the squares of one step ahead in a sample forecast error. The first observation as the initial value for SES was suggested by Hyndman and Athanasopoulos [1]. Additionally, many statistical packages, such as Minitab, have been used to set the initial value for SES with the average of the first six observations [23].

The following are two methods that are most commonly used to set the initial value for SES and are used to compare with the proposed method.

1. The first observation is used as the initial value of α when using SES [22].
By simple algebra, of SES, can also be expressed as
 $F_1 = y_1$, $F_{t+1} = \alpha y_t + (1 - \alpha)F_t$ where F_1 and y_1 are the initial value and the actual value, respectively. The choice is preferable when the level changes rapidly (α close to 1, or $1 - \alpha$ close to 0).
2. The average of the first six observations is used for the initial value [23].
By simple algebra of SES, can also be expressed as
 $F_1 = \left(\frac{1}{6}y_1 + \frac{1}{6}y_2 + \frac{1}{6}y_3 + \frac{1}{6}y_4 + \frac{1}{6}y_5 + \frac{1}{6}y_6\right)$, $F_{t+1} = \alpha y_t + (1 - \alpha)F_t$ where F_1 and y_1, y_2, \dots, y_6 are the initial value and the actual values, respectively.

2.3. Performance Metrics

The fundamental challenge for a metric error measure is that it condenses a large amount of data into a single value. Evaluating the performance of a forecasting method is achieved by comparing the actual values with the predicted ones. A typical approach is to use a specific criterion to measure the error of the predicted value. Subsequently, the efficiency of the forecasting method is evaluated based on how close the predicted and actual values are. The most frequently used error index, MSE [6,32], was used in the present study, which is defined as

$$\text{MSE} = \frac{\sum_{t=1}^n (F_t - y_t)^2}{n}$$

where y_t and F_t are the actual and predicted values at time t , respectively, and n is the number of data points.

3. The Proposed Method to Set the Initial Value for SES

In this method, the first six observations are still used for taking the weighted average. Based on the concept of the weighted average of past observations, the closer it gets to the first observation, the greater the weight and the farther (Table 1). Hence, F_1 is assigned more weight (w_1) for y_1 than for y_2, \dots, y_6 .

Table 1. Assigning weights to the first six observations.

| t | y_t | w_t |
|-----|-------|-------|
| 1 | y_1 | 6/21 |
| 2 | y_2 | 5/21 |
| 3 | y_3 | 4/21 |
| 4 | y_4 | 3/21 |
| 5 | y_5 | 2/21 |
| 6 | y_6 | 1/21 |

Performing SES with the proposed method is achieved as follows:

1. Let $F_1 = \left(\frac{6}{21}y_1 + \frac{5}{21}y_2 + \frac{4}{21}y_3 + \frac{3}{21}y_4 + \frac{2}{21}y_5 + \frac{1}{21}y_6\right)$, where $\sum_{t=1}^6 w_t = 1$.
2. Use Solver in Microsoft Excel to find the optimal value for α while minimizing MSE.
3. Provide the initial forecasting value via $F_{t+1} = \alpha y_t + (1 - \alpha)F_t$.
4. Calculate the MSE.

An example of using the Solver in Microsoft Excel is presented in Figure 1. The time series dataset of S1 is set with the initial value of Method 1, and the optimal smoothing value returned is $\alpha = 0.3689$, whereas the MSE value is 896,212.76.

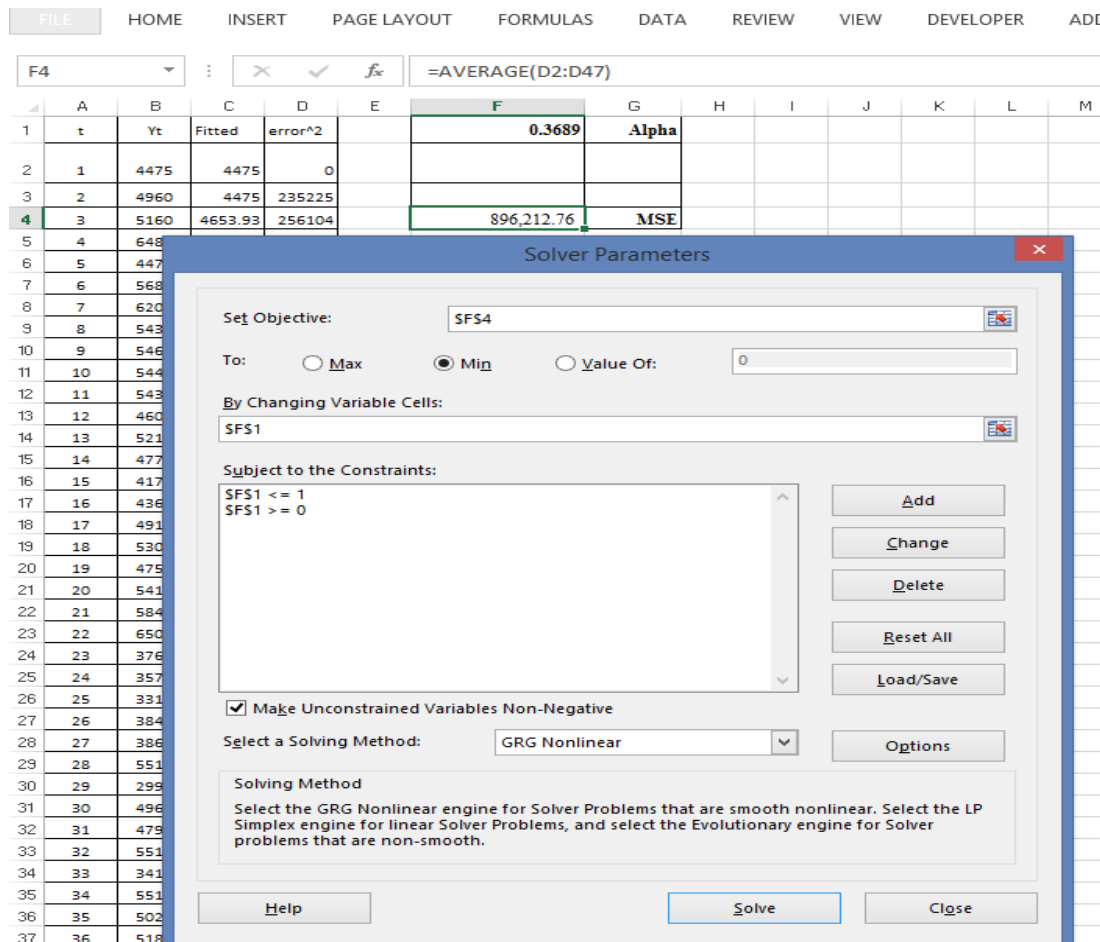


Figure 1. An example of the Solver results pane for the time series of S1 with initial value setting with Method 1.

4. Experimental Study

Fifteen real time series datasets from the M3 competition with a stationary pattern (five small, medium, and large ones) [33] were used in this study to evaluate the performance of the proposed method. For each one, the augmented Dickey–Fuller test was used to check whether the pattern was stationary. Brief details of these datasets are reported in Table 2.

α was set in increments of 0.001 using $\alpha \in [0, 1]$, so there were 1000 sets of conditions for each dataset. The experimental study was conducted via the steps using the R version 3.5.2 [34] to measure the performance of the proposed method compared to the SES method.

The steps of the experimental study are as follows:

1. Each time series dataset with a stationary pattern is performed for three initial values by applying methods 1 and 2, and the proposed method are as follows:

Method 1: $F_1 = y_1$

Method 2: $F_1 = \left(\frac{1}{6}y_1 + \frac{1}{6}y_2 + \frac{1}{6}y_3 + \frac{1}{6}y_4 + \frac{1}{6}y_5 + \frac{1}{6}y_6\right)$

The proposed method: $F_1 = \left(\frac{6}{21}y_1 + \frac{5}{21}y_2 + \frac{4}{21}y_3 + \frac{3}{21}y_4 + \frac{2}{21}y_5 + \frac{1}{21}y_6\right)$.

2. For each initial value setting, use the solver in Excel to find the optimal value for α while minimizing the MSE. Hence, we obtained the optimal value for α and the MSE value of each initial value setting.
3. For each initial value setting, use a grid search with α by varying 0.001, 0.002, ..., 1.00 and compute the MSE in each α . With these 1000 conditions, we searched the optimal value for α and obtained the lowest MSE.

4. The performance of the proposed method compared with the other two initialization methods of the SES method, considering the lowest MSE.

Table 2. The M3 competition time series data used to compare the forecast methods.

| | Dataset | M3 Competition Code | Time Period | Size |
|--------|---------|---------------------|----------------|------|
| Small | S1 | N243 (Y) | 1947–1992 | 46 |
| | S2 | N127 (Y) | 1947–1993 | 47 |
| | S3 | N180 (Y) | 1947–1993 | 47 |
| | S4 | N229 (Y) | 1947–1993 | 47 |
| | S5 | N721 (Q) | 1984Q1–1994Q4 | 44 |
| Medium | M1 | N1368 (Q) | 1974Q2–1991Q3 | 70 |
| | M2 | N1449 (Q) | 1990Q1–2007Q1 | 69 |
| | M3 | N1470 (Q) | 1990Q1–2007Q1 | 69 |
| | M4 | N1472 (Q) | 1990Q1–2007Q1 | 69 |
| | M5 | N903 (Q) | 1976Q1–1993Q4 | 72 |
| Large | L1 | N2125 (M) | 1978M1–1989M12 | 144 |
| | L2 | N1886 (M) | 1979M1–1990M12 | 144 |
| | L3 | N2022 (M) | 1981M1–1992M12 | 144 |
| | L4 | N2025 (M) | 1981M1–1992M12 | 144 |
| | L5 | N2070 (M) | 1982M1–1993M12 | 144 |

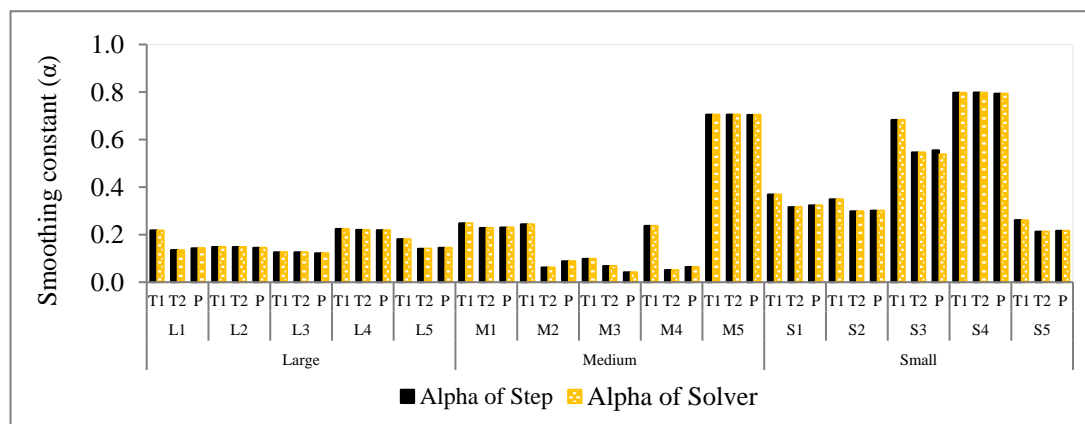
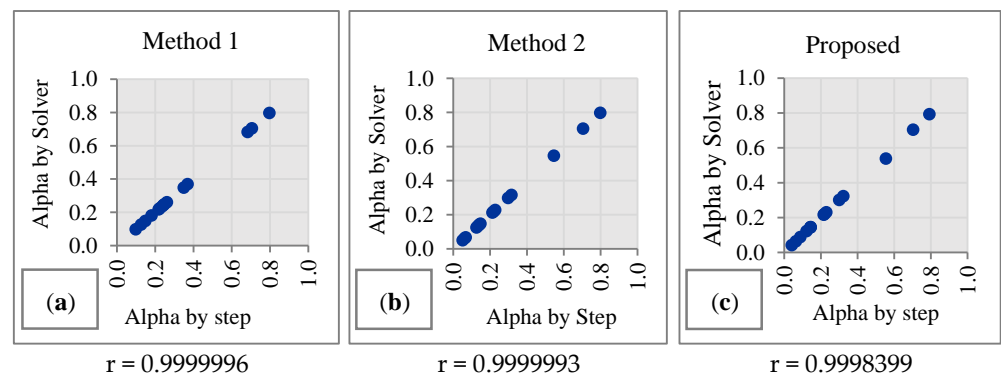
Y, Q, and M are yearly, quarterly, and monthly, respectively.

5. Results and Discussion

Table 3 provides the initial values via the three methods and the optimal value for α using a step search and Solver in Microsoft Excel, which were very similar. As an example for a small-sized dataset (S1), the optimal values for α using a step search and Solver from Microsoft Excel were 0.369 and 0.3689, respectively, when setting the initial value using Method 1, whereas they were 0.316 and 0.3162, respectively, when setting the initial value using Method 2, and 0.323 and 0.3232, respectively, when setting the initial value using the proposed method. This trend was the same for all of the datasets. Figure 2 shows the results for each dataset when using different initialization methods; the optimal value for α differs only slightly. Eventually, Solver from Microsoft Excel and the step search methods obtained very similar optimal values for α for the same initialization method. Overall, the different sizes of small with S1–S5, medium with M1–M5, and large with L1–L5 datasets are represented by the bar chart of Figure 3, and it can be seen that both Solver in Microsoft Excel and step search obtained nearly the same optimal value for α results. Hence, the Solver from Microsoft Excel is an alternative to a powerful method for obtaining the optimal value for α .

Table 3. The optimal values of α using the initial values provided by the three methods.

| Size | Dataset | Method 1 | | Method 2 | | Proposed | |
|--------|---------|----------|--------|----------|--------|----------|--------|
| | | Step | Solver | Step | Solver | Step | Solver |
| Small | S1 | 0.369 | 0.3689 | 0.316 | 0.3162 | 0.323 | 0.3232 |
| | S2 | 0.349 | 0.3486 | 0.298 | 0.2982 | 0.301 | 0.3013 |
| | S3 | 0.683 | 0.6828 | 0.546 | 0.5459 | 0.555 | 0.5383 |
| | S4 | 0.797 | 0.7970 | 0.798 | 0.7975 | 0.793 | 0.7932 |
| | S5 | 0.261 | 0.2608 | 0.213 | 0.2129 | 0.216 | 0.2161 |
| Medium | M1 | 0.248 | 0.2477 | 0.228 | 0.2281 | 0.230 | 0.2301 |
| | M2 | 0.244 | 0.2437 | 0.062 | 0.0621 | 0.088 | 0.0884 |
| | M3 | 0.098 | 0.0982 | 0.068 | 0.0677 | 0.042 | 0.0420 |
| | M4 | 0.237 | 0.2368 | 0.051 | 0.0507 | 0.064 | 0.0642 |
| | M5 | 0.705 | 0.7049 | 0.705 | 0.7048 | 0.704 | 0.7045 |
| Large | L1 | 0.218 | 0.2177 | 0.135 | 0.1348 | 0.143 | 0.1433 |
| | L2 | 0.148 | 0.1481 | 0.148 | 0.1478 | 0.145 | 0.1447 |
| | L3 | 0.126 | 0.1262 | 0.126 | 0.1257 | 0.122 | 0.1221 |
| | L4 | 0.224 | 0.2238 | 0.22 | 0.2195 | 0.219 | 0.2191 |
| | L5 | 0.181 | 0.1809 | 0.141 | 0.1415 | 0.145 | 0.1449 |

**Figure 2.** A bar chart of the optimal values of α (alpha) using a step search or Solver from Microsoft Excel when setting the initial value using T1 (Method 1), T2 (Method 2), or P (the proposed method) for large, medium, and small time series datasets.**Figure 3.** Scatter plots of the optimal value for α obtained using the step search method and Solver from Microsoft Excel when setting the initial value using (a) Method 1, (b) Method 2, or (c) the proposed method for all 15 time series datasets.

The results of a correlation analysis of the performances of the three methods are presented in Figure 3. It can be seen that the correlation coefficient values are close to +1

and lie in a straight line, meaning that the optimal values for α obtained using the step search method or Solver from Microsoft Excel using the initial value from all three methods are in good agreement.

Figure 4 illustrates the actual and predicted values when setting the initial value using Method 1, Method 2, and the proposed method. For example, in S1, when setting the initial value using Method 2 and the proposed method, similar predicted values are achieved, but when setting the initial value using Method 1, the predicted values in the early period are far from those of the other methods. Additionally, they can be interpreted in the same way in the other time series. Moreover, Table 4 reports the lowest MSE values for the three initial value setting methods used to obtain the optimal value for α using the step search method and Solver from Microsoft Excel. For each dataset, the lowest MSE value was obtained from a set of 1000 MSE values. For example, for S1, when setting the initial value using Method 1, Method 2, and the proposed method, the lowest MSE values were 896,212.77, 886,047.38, and 885,070.96, respectively. This trend was the same for the other datasets. It can be seen that although the results are similar, setting the initial value using the proposed method obtained lower MSE values than the other methods. Figure 5 visually supports these findings.

Table 4. The lowest MSE of three initial value setting methods using the optimal value for α from step search and Solver from Microsoft Excel.

| Size | Dataset | Method 1 | | Method 2 | | Proposed | |
|--------|---------|--------------|--------------|--------------|--------------|--------------|--------------|
| | | Step | Solver | Step | Solver | Step | Solver |
| Small | S1 | 896,212.77 | 896,212.76 | 886,047.38 | 886,047.36 | 885,070.96 | 885,070.92 |
| | S2 | 1,450,217.80 | 1,450,217.50 | 1,419,280.76 | 1,419,280.72 | 1,418,605.85 | 1,418,605.71 |
| | S3 | 82,963.99 | 82,963.98 | 82,284.51 | 82,284.51 | 81,506.58 | 81,505.96 |
| | S4 | 393,051.81 | 393,051.81 | 393,111.26 | 393,111.21 | 392,958.71 | 392,958.71 |
| | S5 | 121,187.43 | 121,187.42 | 116,543.72 | 116,543.71 | 116,648.83 | 116,648.82 |
| Medium | M1 | 29,069.60 | 29,069.60 | 28,837.52 | 28,837.52 | 28,735.94 | 28,735.94 |
| | M2 | 934,245.09 | 934,244.89 | 756,669.78 | 756,669.75 | 765,244.84 | 765,244.02 |
| | M3 | 1,649,100.03 | 1,649,099.38 | 1,599,938.85 | 1,599,937.30 | 1,571,685.18 | 1,571,685.15 |
| | M4 | 519,353.65 | 519,353.61 | 424,846.72 | 424,846.34 | 434,269.59 | 434,269.50 |
| | M5 | 165,837.93 | 165,837.93 | 165,830.64 | 165,830.63 | 165,816.74 | 165,816.72 |
| Large | L1 | 1,516,519.54 | 1,516,519.40 | 1,479,599.21 | 1,479,599.13 | 1,470,586.75 | 1,470,586.56 |
| | L2 | 1,880,659.84 | 1,880,659.83 | 1,880,401.03 | 1,880,400.76 | 1,877,245.36 | 1,877,244.59 |
| | L3 | 555,763.08 | 555,762.96 | 555,644.55 | 555,644.38 | 554,925.61 | 554,925.58 |
| | L4 | 28,361.93 | 28,361.93 | 28,278.23 | 28,278.21 | 28,270.66 | 28,270.66 |
| | L5 | 53,326.57 | 53,326.57 | 50,839.83 | 50,839.76 | 50,986.27 | 50,986.27 |

The initial setting methods' lowest MSE frequencies were compared using Chi-squared goodness-of-fit tests (Table 5). The null hypothesis states that the number of times that each method achieves the lowest MSE is the same. For example, for S1, the Chi-squared statistic for the lowest MSE value is 159.65 with a p -value < 0.0001 , which leads to the conclusion that the number of times that each method achieves the lowest MSE is significantly different. The results for the other datasets can be interpreted in the same way. Thus, the proposed method for setting the initial value achieved the lowest MSE values, and it is evident that it quite considerably outperformed the other two methods.

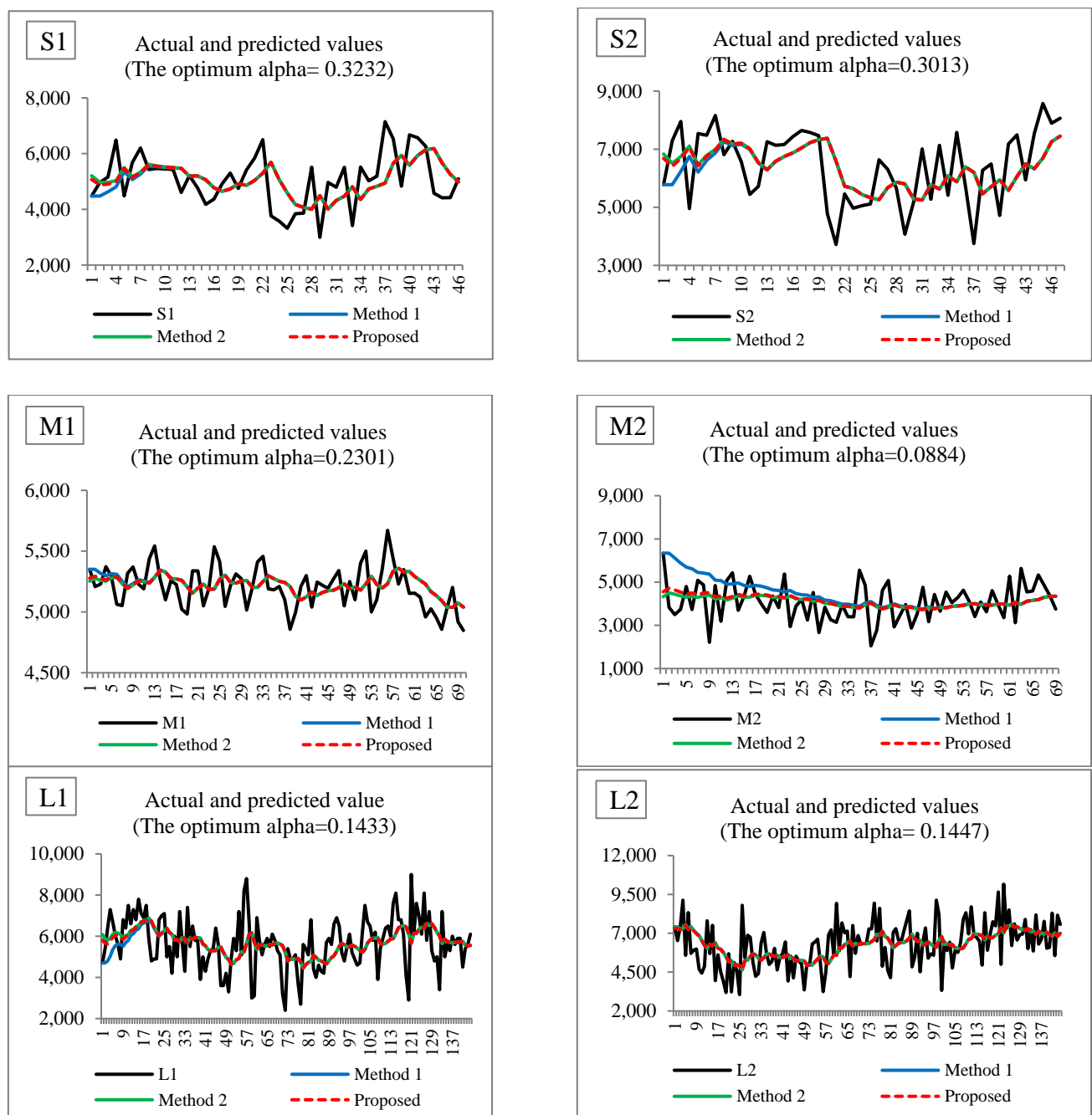


Figure 4. Plots of the actual and predicted values when setting the initial value using Method 1, Method 2, and the proposed method; for example, of small, medium, and large time series data with S1, S2, M1, M2, L1, and L2.

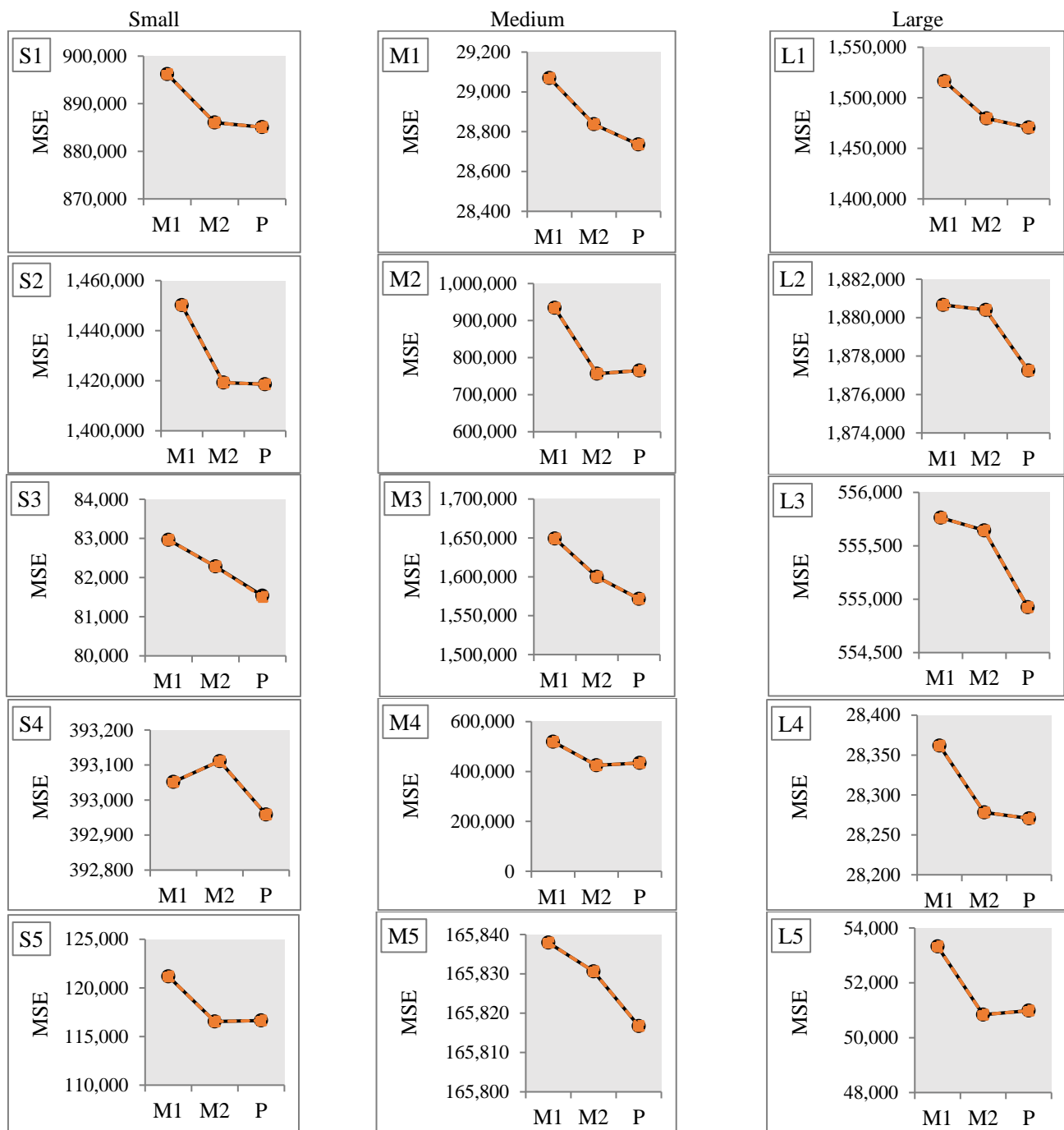


Figure 5. The lowest MSE using the step search method and Solver from Microsoft Excel to find the optimal value for α after setting the initial value using M1 (Method 1), M2 (Method 2), and P (the proposed method) for large (L1–L5), medium (M1–M5), and small (S1–S5) time series datasets.

Figure 6 shows a stacked bar chart of the frequencies of obtaining the lowest MSE value when setting the initial value using the three initial value methods: Method 1, Method 2, and the proposed method for each of the 15 datasets. Notably, Figure 6 shows the MSE values for the proposed method were lower than those for the other two methods, a trend that was the same for all of the datasets. Moreover, in Figure 7, it is worth noting that using an initial value, in Method 1, of α in the range of 0.1–0.4 provided forecasted values far from the actual values, whereas when it exceeded 0.5, all three methods provided similar results. The proposed method and Method 2 usually provided similar MSE values for the

same set of conditions, and their plotted lines tended to overlap each other. Meanwhile, lower MSE values for the three initial value setting methods became more evident as the settings were decreased. These results are consistent with the other findings, and it was concluded that the proposed method performed better than the others.

Table 5. Chi-squared goodness-of-fit test results for comparing the lowest MSE value frequencies of the three initial value setting methods.

| Size | Dataset | Method 1 | Method 2 | Proposed | Chi-Squared Value | <i>p</i> -Value |
|---------|---------|----------|----------|----------|-------------------|-----------------|
| Small | S1 | 425 | 145 | 430 | 159.65 | <0.0001 |
| | S2 | 300 | 97 | 603 | 389.05 | <0.0001 |
| | S3 | 316 | 264 | 420 | 37.86 | <0.0001 |
| | S4 | 155 | 0 | 845 | 1214.15 | <0.0001 |
| | S5 | 248 | 223 | 529 | 173.22 | <0.0001 |
| Medium | M1 | 285 | 218 | 497 | 127.27 | <0.0001 |
| | M2 | 348 | 172 | 480 | 143.26 | <0.0001 |
| | M3 | 146 | 295 | 559 | 262.47 | <0.0001 |
| | M4 | 361 | 234 | 405 | 47.31 | <0.0001 |
| | M5 | 363 | 146 | 491 | 182.50 | <0.0001 |
| Large | L1 | 449 | 165 | 386 | 133.47 | <0.0001 |
| | L2 | 4 | 56 | 940 | 1660.26 | <0.0001 |
| | L3 | 66 | 273 | 661 | 547.42 | <0.0001 |
| | L4 | 78 | 127 | 795 | 962.71 | <0.0001 |
| | L5 | 302 | 221 | 447 | 102.722 | <0.0001 |
| Average | | 256.40 | 175.73 | 565.87 | | |

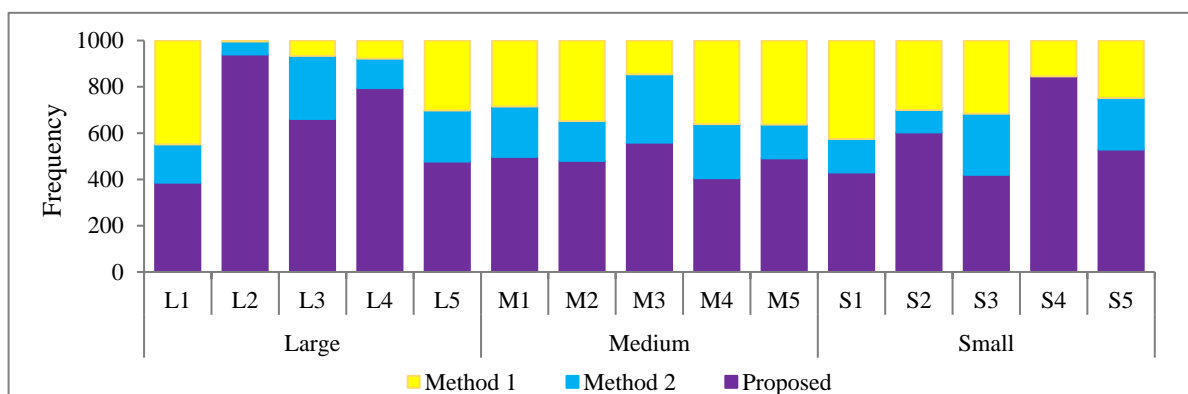


Figure 6. The frequencies of achieving the lowest MSE when setting the initial value using Method 1, Method 2, and the proposed method.

It can be seen that when using values of α ranging from 0 to 1 (0.001, 0.002, ..., and 1) to produce 1000 settings, different average MSE levels were provided (Table 6). For example, for S1, Method 1, Method 2, and the proposed method produced MSE values of 984,214.12, 969,740.55, and 967,322.56, respectively. Similarly, they produced MSE values of 31,757.83, 31,338.43, and 30,365.02, respectively, for M1, and 1,639,370.09, 1,619,541.12, and 1,616,325.69, respectively, for L1. Thus, it can be concluded that, when using the same value of α , the initial value provided by the proposed method provided the lowest MSE for all sizes of datasets and quite considerably outperformed the other two methods in

this endeavor. Bar charts of the average MSE values obtained using the three methods for each dataset are shown in Figure 8. The results clearly show that the proposed method performed much better than Method 1 and Method 2.

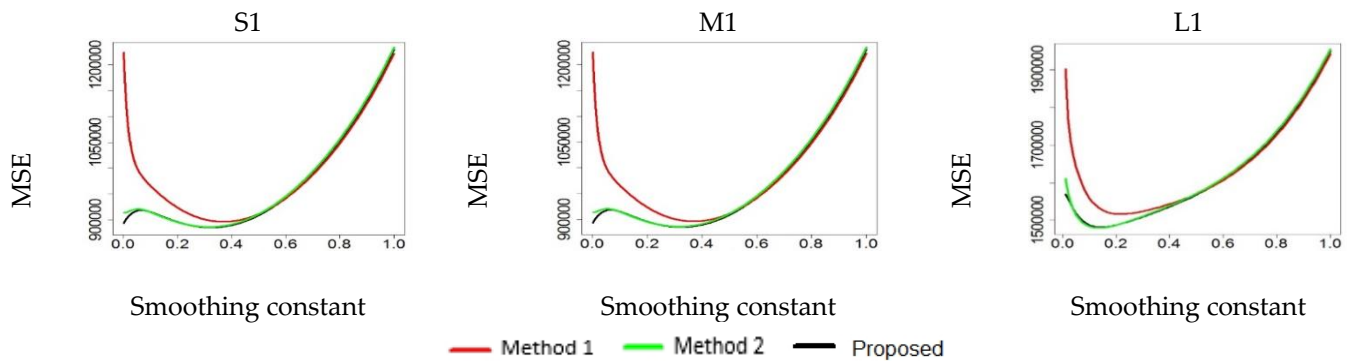


Figure 7. MSE of the three initial value setting methods for 1000 sets of conditions of S1, M1, and L1.

Table 6. The average MSE values for the three initial value setting methods using predetermined α values ranging from 0 to 1 (0.001, 0.002, ..., and 1).

| Size | Dataset | Method 1 | Method 2 | Proposed |
|--------|---------|--------------|--------------|----------------|
| Small | S1 | 984,214.12 | 969,740.55 | 967,322.56 * |
| | S2 | 1,594,382.33 | 1,569,643.61 | 1,564,879.67 * |
| | S3 | 122,698.44 | 105,228.68 | 104,818.88 * |
| | S4 | 460,672.15 | 461,335.23 | 456,778.22 * |
| | S5 | 146,968.36 | 144,174.95 | 143,896.02 * |
| Medium | M1 | 31,757.83 | 31,338.43 | 30,365.02 * |
| | M2 | 1,204,852.71 | 1,064,595.04 | 1,062,339.88 * |
| | M3 | 2,188,507.09 | 2,166,421.78 | 2,160,496.8 * |
| | M4 | 634,169.22 | 584,823.33 | 582,716.40 * |
| | M5 | 181,293.72 | 181,455.69 | 178,257.81 * |
| Large | L1 | 1,639,370.09 | 1,619,541.12 | 1,616,325.69 * |
| | L2 | 2,371,310.64 | 2,371,031.69 | 2,367,682.30 * |
| | L3 | 707,504.85 | 707,387.44 | 706,683.96 * |
| | L4 | 36,068.72 | 35,952.54 | 35,942.25 * |
| | L5 | 69,863.02 | 69,555.56 | 69,277.16 * |

* The best performance in terms of average MSE value.

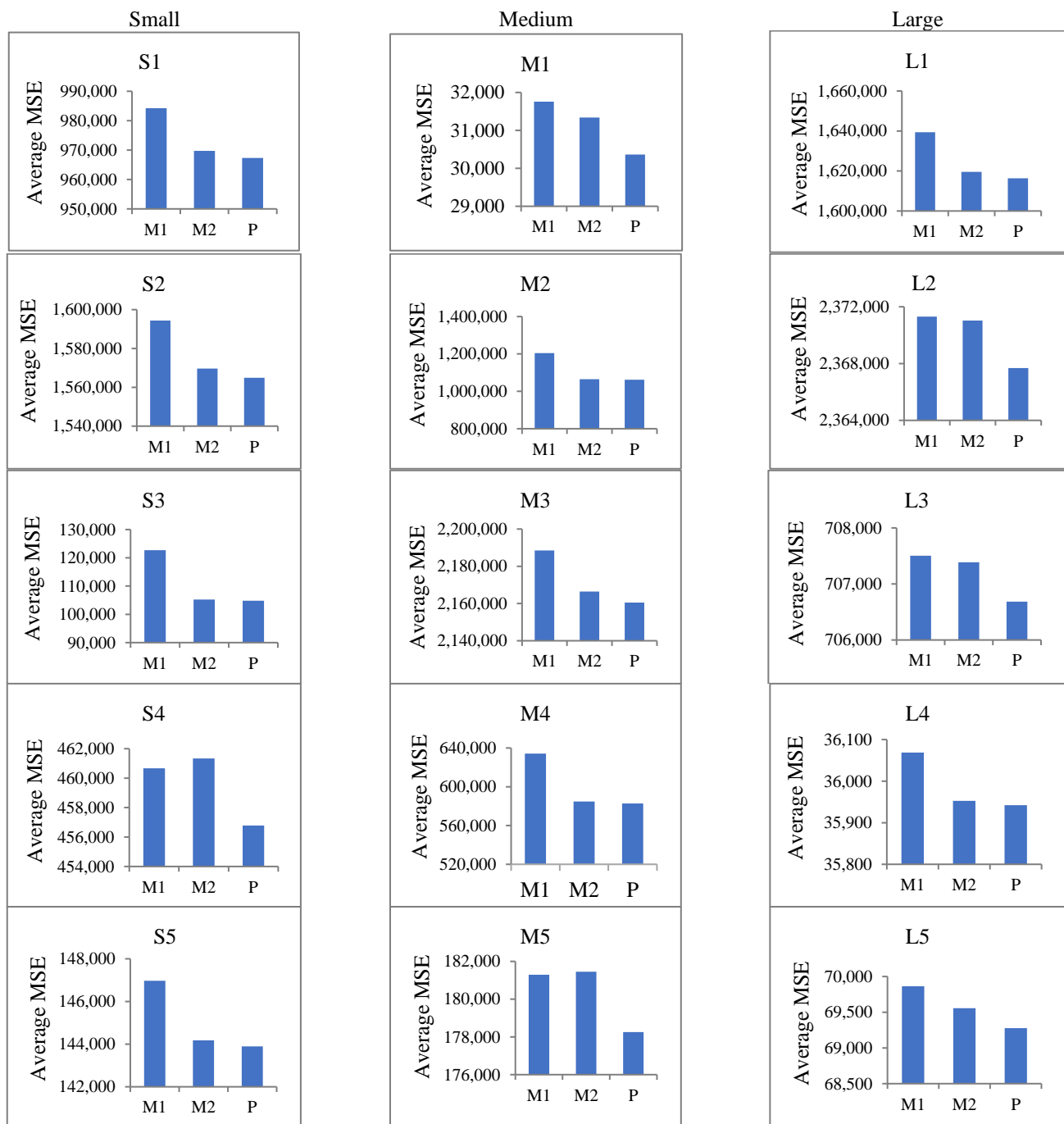


Figure 8. Average accuracy values with MSE for the three initial value setting methods using values of α ranging from 0 to 1 (0.001, 0.002, ..., and 1) for the 15 datasets: small (S1–S5), medium (M1–M5), and large (L1–L5).

6. Conclusions and Remarks

To effectively use the SES method, the forecaster must first choose a proper value for α and then set the initial value to calculate the smoothed values and make the forecast. The MSE is often used as a criterion for selecting an appropriate value for α . For instance, by assigning the values [0,1], one then selects the value that produces the smallest MSE. Importantly, these values considerably affect the accuracy of the forecast.

In this study, the Solver in Microsoft Excel and step search methods were used to determine the value of α that optimally fitted several time series datasets. Since their

performances were not different, Solver from Microsoft Excel is a powerful alternative method for obtaining the optimal value for α . Moreover, the initial value influenced the forecast using the SES method. The simple initial value setting using the first observed value provided a far worse performance than the other two methods based on small, medium, and large time series datasets. It is worth noting that using values for α in the range 0.1–0.4 as the initial value with the first observed provided forecasted values far from the actual value, whereas when it exceeded 0.5, all three methods provided similar results. The limitation of SES work is based on the principle that a prediction is a weighted linear sum of past observations. Moreover, the forecasting accuracy is directly affected by the value of α since it adjusts the weights given to observations and the initial value. Therefore, it is always important to choose a proper value for α .

Author Contributions: Conceptualization, C.W.; Methodology, C.W. and W.J.; Formal analysis, C.W. and W.J.; Writing—original draft, C.W.; Writing—review & editing, W.J.; Visualization, W.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by International SciKU Branding (ISB).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: <https://forecasters.org/resources/time-series-data/>.

Acknowledgments: The authors are grateful for funding from the Thailand Institute of Scientific and Technological Research Institute (TISTR). In addition, the authors also would like to thank the Khon Kaen University, Kasetsart University, and International SciKU Branding (ISB) for providing the facilities to conduct the research.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

- Hyndman, R.J.; Koehler, A.B.; Snyder, R.D.; Grose, S. A state space framework for automatic forecasting using exponential smoothing methods. *Int. J. Forecast.* **2002**, *18*, 439–454. [\[CrossRef\]](#)
- Chen, F.; Ryan, J.K.; Simchi-Levi, D. The impact of exponential smoothing forecasts on the bullwhip effect. *Nav. Res. Logist.* **2000**, *47*, 269–286. [\[CrossRef\]](#)
- Hussain, A.; Rahman, M.; Alam Memon, J. Forecasting electricity consumption in Pakistan: The way forward. *Energy Policy* **2016**, *90*, 73–80. [\[CrossRef\]](#)
- De Baets, S.; Harvey, N. Forecasting from time series subject to sporadic perturbations: Effectiveness of different types of forecasting support. *Int. J. Forecast.* **2018**, *34*, 163–180. [\[CrossRef\]](#)
- Chukwulozie, O.P.; Chinagorom, N.E.; Chukwuneke, C.J.; Nnodim, C.T. The analysis of cigarette production using double exponential smoothing model. *Acad. J. Sci.* **2017**, *7*, 293–308.
- Wongoutong, C. The effect on forecasting accuracy of the Holt-Winters Method when using the incorrect model on a non-stationary time series. *Thail. Stat.* **2021**, *19*, 565–582.
- Xie, M.; Hong, G.Y.; Wohlin, C. A study of the exponential smoothing technique in software reliability growth prediction. *Qual. Reliab. Eng. Int.* **1997**, *13*, 347–353. [\[CrossRef\]](#)
- Schroeder, R.G.; Goldstein, S.M.; Rungtusanatham, M.J. *Operations Management in the Supply Chain: Decisions and Cases*; McGraw-Hill/Irwin: Windsor, NJ, USA, 2013.
- Jacobs, F.R.; Chase, R.B. Operations and supply chain management. In *The Core*, 3rd ed.; Chapter 3; McGraw-Hill Higher Education: Columbus, OH, USA, 2013.
- Paul, S.K. Determination of exponential smoothing constant to minimize mean square error and mean absolute deviation. *Glob. J. Res. Eng.* **2011**, *11*, 31–34.
- Bermúdez, J.D.; Segura, J.V.; Vercher, E. Improving demand forecasting accuracy using nonlinear programming software. *J. Oper. Res. Soc.* **2006**, *57*, 94–100. [\[CrossRef\]](#)
- Chopra, S.; Meindl, P.; Kalra, D.V. *Supply Chain Management: Strategy, Planning, and Operation*; Pearson Education: London, UK, 2013; Volume 232.
- Nariman, N.A.; Hamdia, K.; Ramadan, A.M.; Sadaghian, H. Optimum Design of Flexural Strength and Stiffness for Reinforced Concrete Beams Using Machine Learning. *Appl. Sci.* **2021**, *11*, 8762. [\[CrossRef\]](#)
- Ravinder, H.V. Determining The Optimal Values Of Exponential Smoothing Constants Does Solver Really Work? *Am. J. Bus. Educ.* **2013**, *6*, 347–360. [\[CrossRef\]](#)

15. Chopra, S.; Meindl, P. *Supply Chain Management: Strategy, Planning, and Operation*, 5th ed.; Prentice Hall: Hoboken, NJ, USA, 2013; p. 195, Chapter 7.
16. Balakrishnan, N.; Render, B.; Stair, R. *Managerial Decision Modeling with Spreadsheets*, 3rd ed.; Bermudez, J.D., Segura, J.V., Velcher, E., Eds.; Prentice Hall: Hoboken, NJ, USA, 2013.
17. Ledolter, J.; Abraham, B. Some comments on the initialization of exponential smoothing. *J. Forecast.* **1984**, *3*, 79–84. [[CrossRef](#)]
18. Brown, R.G. *Statistical Forecasting for Inventory Control*; McGraw-Hill: New York, NY, USA, 1959.
19. Cohen, G.D. Bayesian adjustment of sales forecasts in multi-item inventory control systems. *J. Ind. Eng.* **1966**, *17*, 474.
20. Johnson, L.A.; Montgomery, D.C. Operations Research in Production Planning, Scheduling and Inventory Control. *J. Oper. Res. Soc.* **1975**, *26*, 568–569.
21. Taylor, J.W. Volatility forecasting with smooth transition exponential smoothing. *Int. J. Forecast.* **2004**, *20*, 273–286. [[CrossRef](#)]
22. Makridakis, S.; Wheelwright, S.C.; Hyndman, R.J. *Forecasting Methods and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2008.
23. University of Duke. Averaging and Exponential Smoothing Models. 2017. Available online: https://people.duke.edu/~rnau/Decision411_2007/411avg.htm (accessed on 11 January 2023).
24. Cadenas, E.; Jaramillo, O.A.; Rivera, W. Analysis and forecasting of wind velocity in chetumal, quintana roo, using the single exponential smoothing method. *Renew Energy* **2010**, *35*, 925–930. [[CrossRef](#)]
25. Rachmat, R.; Suhartono, S. Comparative Analysis of single exponential smoothing and Holt's Method for quality of hospital services forecasting in general hospital. *Bull. Comput. Sci. Electr. Eng.* **2020**, *1*, 80–86. [[CrossRef](#)]
26. Wongoutong, C. Improvement of the Holt-Winters Multiplicative Method with a New Initial Value Settings Method. *Thail. Stat.* **2021**, *19*, 280–293.
27. Sudibyo, N.A.; Iswardani, A.; Septyanto, A.W.; Wicaksono, T.G. Prediksi inflasi di indonesia menggunakan metode moving average, single exponential smoothing dan double exponential smoothing. *J. Lebesgue J. Ilm. Pendidik. Mat. Mat. Stat.* **2020**, *1*, 123–129.
28. Mas-Machuca, M.; Sainz, M.; Martinez-Costa, C. A review of forecasting models for new products. *Intang. Capital.* **2014**, *10*, 1–25. [[CrossRef](#)]
29. Brown, R.G. *Smoothing, Forecasting and Prediction of Discrete Time Series*; Prentice-Hall: Annapolis, MD, USA, 1962.
30. Gardner, J.R.; Everette, S. Exponential smoothing: The state of the art—Part II. *Int. J. Forecast.* **2006**, *22*, 637–666. [[CrossRef](#)]
31. Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice*; OTexts: Online Textbook, 8 May 2018.
32. Clements, M.; Hendry, D. *Forecasting Economic Time Series*; Cambridge University Press: Cambridge, UK, 1998.
33. Makridakis, S.; Hibon, M. The M3-Competition: Results, conclusions and implications. *Int. J. Forecast.* **2000**, *16*, 451–476. [[CrossRef](#)]
34. The R Project for Statistical Computing. 2019. Available online: <https://www.r-project.org/> (accessed on 11 January 2023).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.