COMPARATIVE ANALYSIS OF EXPONENTIAL SMOOTHING MODELS FOR SALES PREDICTION AND SUPPLY MANAGEMENT IN E-COMMERCE

Aji Saeful1; Fandy Setyo Utomo2*; Yuli Purwati3; Mohd Sanusi Azmi4

Department of Informatics1,2,3, Universitas AMIKOM Purwokerto, Indonesia1,2,3
https://www.amikompurwokerto.ac.id/1,2,3
ajiomic23@gmail.com1, fandy_setyo_utomo@amikompurwokerto.ac.id2*, yulipurwati@amikompurwokerto.ac.id3

Faculty of Information and Communication Technology4, Universiti Teknikal Malaysia Melaka, Malaysia4
https://www.utm.edu.my/4
sanusi@utm.edu.my4

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

The creation is distributed under the Creative Commons Attribution-NonCommercial 4.0 International License.

Abstract—In the growing era of e-commerce, stock management is crucial. Problems arise in forecasting sales in order to achieve effective stock management. This research uses the time series analysis method by focusing on comparing the accuracy of three forecasting methods: Single Exponential Smoothing (SES), Double Exponential Smoothing (DES), and Triple Exponential Smoothing (TES/Holt-Winter). This research provides a solution by comparing the performance of the three methods based on the Mean Absolute Error (MAE) results and prediction graphs. The goal is to determine the most accurate forecasting method using the time series analysis method with several stages, namely data preprocessing, train/test split, modeling, and performance metrics measurement. Based on the test results show MAE SES 1077, DES 96, and TES (Holt-Winter) 101. Although DES has a lower MAE, TES (Holt-Winter) provides better accuracy, especially through prediction graph analysis. Holt-Winter is recognized as the most effective method in forecasting future sales, reliable for proper stock management in the dynamic e-commerce industry. This approach is expected to improve efficiency and accuracy in enterprise stock management, support the growth of online businesses, and contribute to the literature and practice of stock management. The use of time series analysis methods, especially Holt-Winter, is considered an important strategic step to optimize sales prediction, positively impact stock management, and create a competitive advantage in a growing market.

Keywords: DES, mean absolute error, prediction, SES, TES.
INTRODUCTION

In the development of an increasingly modern era, many people do shopping through online shops. Online sales are also called e-commerce, and e-commerce is one of the business sectors that has increased in recent years. This is due to the increasing use of the internet and electronic devices, as well as changes in consumer behaviour that increasingly favour convenience and comfort in shopping [1]. Sometimes, when buyers want to shop online, the stock of goods in the store often runs out. This is due to the need for better management of the stock management system [2]. Stock management can be improved with the help of machine learning, primarily through forecasting techniques for sales transactions [3]. Machine learning empowers systems to acquire knowledge from data, enhance performance without direct programming, and find applications in facial recognition, translation, and predictive analysis [4].

The decision-making process involves three stages, namely collecting information, processing data, and storing the results of data processing. Forecasting is an estimate or Prediction of Projected demand determined by various predictor variables in the future, often using historical data in a time series as its basis [5]. Decision-making on E-commerce is an essential thing that can have an impact on income on E-commerce. E-commerce entities encounter a challenge in effectively managing their inventory of goods. Too little stock can lead to running out of goods and lost sales, while too much stock can lead to a buildup of goods and financial losses [6]. In solving stock management problems, forecasting methods can be used to determine how much stock is needed by the Company based on the number of transactions; forecasting sales transactions is one way to overcome the challenges of managing stock of goods in e-commerce. Forecasting is the process of predicting the value of a variable in the future based on historical data. In the context of e-commerce, forecasting sales transactions can be used to predict the number of goods that will be sold in the future [7][8].

Research conducted by [9] regarding salt sales forecasting by applying the SES and DES methods found that MAPE from SES reached 7.93%. In comparison, MAPE from DES was around 28.14%. A forecasting process that follows the correct steps or an organized procedure will result in efficient forecasting. In a study conducted by [10][11], a comparison between the SES and TES (Holt-Winter) methods shows that both provide an adequate level of accuracy in predicting the number of hospitalized patients. There are three critical aspects in the forecasting process: Analysis of previous data, which is helpful in identifying patterns that occurred in the previous period, and determination of the data used. A method is considered appropriate if the resulting forecast has no significant difference from reality. Projections of current data from the method are used by considering possible shifting factors, including policy changes, societal developments, technological advances, and new inventions [12].

A time series denotes a sequence of observations documented during a particular timeframe, be it weeks, months, or quarters [13]. In forecasting, the SES and DES models can be used, assuming that the data fluctuates around a constant mean value without showing a consistent trend or growth pattern. In the context of forecasting modelling, the use of Mean Absolute Error (MAE) as a model evaluation metric is essential. In general, models with more minor Mean Absolute errors (MAE) are considered to have higher accuracy [14][15]. The Mean Absolute Error (MAE) is utilized as an evaluation metric in forecasting models, measuring the average absolute discrepancy between the actual value and the value predicted by the model [16]. In addition to using MAE to measure model performance, matrix evaluation metrics can be used [17].

Drawing on insights from previous research, this study aims to compare sales forecasting methodologies, specifically SES, DES, and TES (Holt-Winter). The evaluation of model outcomes utilizes the Absolute Error Method (MAE) and prediction charts. The primary goal of this investigation is to determine the optimal accuracy among exponential smoothing techniques for predicting future sales.
volumes, providing a basis for effective stock management. This research seeks to enable companies to make informed adjustments to stock levels, thereby preventing shortages or excessive overstock while contributing to the advancement of forecasting methodologies in the field.

**MATERIALS AND METHODS**

This research using the time series analysis method involves a systematic sequence of steps, from data collection to preprocessing, which includes partitioning the data into two sets: training and testing data. The main focus of this study is on implementing data models that employ SES, DES, and TES (Holt-Winter) techniques. The procedural steps include prediction and performance evaluation phases, further using metrics such as mean absolute error (MAE) and prediction graphs. In processing, researchers use the python language as a tool for data analysis and processing which is run on google colab. All of these steps are detailed in Figure 1, Which offers a graphical depiction of both the procedural steps and outcomes of the research stages conducted.

In Figure 1, the research stages are depicted, starting with the dataset collection. The dataset used comprises sales data from an e-commerce platform spanning from July 4, 2016, to May 30, 2022, obtained from a data science course institution. The use of this dataset enables researchers to analyze long-term sales trends and understand seasonal patterns as well as sales fluctuations.

The next stage is data preprocessing, which includes a stationarity test and seasonal decomposition. The stationarity test is crucial to ensure that the data used has a constant mean and variance over time, which is a primary requirement in many forecasting methods. Meanwhile, seasonal decomposition separates the data into trend, seasonal, and residual components, facilitating further analysis by identifying seasonal patterns that may affect sales.

After preprocessing, the data is split into training and testing datasets. This division is essential for evaluating the model fairly, where the training data is used to build the model and the testing data is used to assess its performance. This helps in avoiding overfitting, where the model fits the training data too closely but fails to predict new data accurately.

In the modeling stage, three forecasting methods are used: single exponential smoothing, double exponential smoothing, and triple exponential smoothing (Holt-Winter). Single exponential smoothing is used for data without a trend or seasonality, with its main benefit being simplicity and ease of implementation. Double exponential smoothing, or Holt's linear trend model, is used for data showing a linear trend, allowing for more accurate forecasting under consistently changing data conditions. Triple exponential smoothing (Holt-Winter) accounts for both trend and seasonality components, making it highly effective for data with strong, recurring seasonal patterns.

Finally, the model’s performance is evaluated using the MAE (Mean Absolute Error) metric and prediction graphs. MAE measures the average absolute error between predicted and actual values, providing an indication of how accurate the model is in making predictions. Prediction graphs, on the other hand, offer a visual comparison of the forecasting results against actual data, making it easier for researchers to intuitively assess model performance and identify potential improvements. Thus, the entire process is designed to ensure that the resulting model is not only accurate but also reliable under various sales conditions.

**RESULTS AND DISCUSSION**

The Results and Discussion section presents the main findings of the conducted research and provides an in-depth analysis of the obtained results. In this section, the outcomes of the modeling process using the methods of single exponential smoothing, double exponential smoothing, and triple exponential smoothing (Holt-Winter) will be explained in detail.
A. Dataset
This research uses an e-commerce platform sales dataset starting from July 4, 2016 to May 30, 2022 obtained from a data science course institute which can be seen at this address https://drive.google.com/drive/folders/1Z6ZnVa5rF6RcvY7ZsRXtI5wf-beW8_x. This dataset has a single numerical attribute that can be adapted and customized in its configuration.

In this context, the dataset consists of a single numerical characteristic that offers adaptability for transformation, consisting of a total of 796,784 transaction data with daily aggregation. After cleaning the data and making it a weekly period, it resulted in 309 clean data. Examples of the dataset are presented in Table 1.

<table>
<thead>
<tr>
<th>No</th>
<th>transaction of date</th>
<th>total transactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2016-07-04</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>2016-07-11</td>
<td>68</td>
</tr>
<tr>
<td>3</td>
<td>2016-07-18</td>
<td>72</td>
</tr>
<tr>
<td>4</td>
<td>2016-07-25</td>
<td>68</td>
</tr>
<tr>
<td>5</td>
<td>2016-08-01</td>
<td>80</td>
</tr>
<tr>
<td>6</td>
<td>2016-08-08</td>
<td>97</td>
</tr>
<tr>
<td>7</td>
<td>2016-08-15</td>
<td>136</td>
</tr>
<tr>
<td>8</td>
<td>2016-08-22</td>
<td>129</td>
</tr>
<tr>
<td>9</td>
<td>2016-08-29</td>
<td>128</td>
</tr>
<tr>
<td>10</td>
<td>2016-09-05</td>
<td>180</td>
</tr>
<tr>
<td>11</td>
<td>2016-09-12</td>
<td>159</td>
</tr>
<tr>
<td>12</td>
<td>2016-09-19</td>
<td>173</td>
</tr>
<tr>
<td>13</td>
<td>2016-09-26</td>
<td>162</td>
</tr>
<tr>
<td>14</td>
<td>2016-10-03</td>
<td>214</td>
</tr>
<tr>
<td>15</td>
<td>2016-10-10</td>
<td>211</td>
</tr>
<tr>
<td>16</td>
<td>2016-10-17</td>
<td>199</td>
</tr>
<tr>
<td>17</td>
<td>2016-10-24</td>
<td>205</td>
</tr>
</tbody>
</table>

Source: (Research Results, 2024)

B. Data Preprocessing
In this research, we emphasize the precision and pertinence of the conducted analysis. This is done to guarantee the reliability of the time series data utilized. We run several tests, including stationary and seasonal decompose. The main objective is to evaluate the data's stability and ensure the data's cleanliness before we put it into the modeling process. By involving these steps, we not only guarantee more valid analysis results but also provide a solid foundation for better understanding and interpretation of the findings of this research [18].

When running the stationary test, the results showed instability in the data. As a reaction step, data transformation is performed to achieve the desired stability. Through the differencing process, followed by the stationary test using the Augmented Dickey-Fuller (ADF) method, the gathered data reveals that both the test statistic value and the p-value strongly indicate the lack of a unit root in the time series, signifying an inclination toward stationarity. The rejection of the null hypothesis is substantiated by a p-value lower than the conventional significance level of 0.05. implying that the data have achieved dormant properties [19]. The successful transformation results can be seen in Figure 2, confirming that the differencing process effectively produces more stationary data.

After the stationary test stage, the focus of the analysis continues with the decomposition test. In this test, the results of the seasonal decomposition test show that there is a seasonality effect in the transaction data. This finding provides important insights into the seasonal patterns that may be influencing the trend of the transaction data, enriching our understanding of the time dynamics in the dataset. By highlighting these seasonal effects, we can more carefully design analysis strategies and make more informed decisions based on the identified seasonal characteristics [20]. The results of the Seasonal Decomposition Check analysis is visible in Figure 3.
C. Train / Test Split

In processing transaction data over a weekly period, a crucial step required is dividing the data into two main subsets: the training data (train) and the testing data (test). This segmentation involves allocating 85% of the data for training purposes and 15% for testing. The training dataset serves as the groundwork for the model to discern patterns and trends within the transaction dataset. By utilizing 85% of this data, the model can learn from weekly variability and characteristics that may be present in transactions. Meanwhile, the testing data covering the remaining 15% provides an opportunity to test the model's performance on periods never seen before. This allows us to evaluate how much the model can generalize and accurately predict transactions in the following weeks.

The train and test split process is important in analyzing weekly transaction data, as it ensures that the developed model cannot only understand trends that have already occurred but can also face challenges from new data in the future. By choosing a data split of 85% train and 15% test, we create a good balance between building a robust model with existing data and testing the model's ability to anticipate changes. The outcome of this procedure not only enhances the model's accuracy but also boosts the reliability of predictions in response to potential variations and fluctuations in weekly transactions. These outcomes are detailed in Table 2.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Train (85%)</th>
<th>Test (15%)</th>
<th>Total (100%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transaction</td>
<td>263</td>
<td>46</td>
<td>309</td>
</tr>
</tbody>
</table>

Source: (Research Results, 2024)

D. Data Modelling

In the modeling stage of this research, we adopt three different methods to improve prediction accuracy on weekly transaction data. The SES approach is used to capture the overall trend in the data by using dynamic weights. Double Exponential Smoothing is applied to broaden the scope of the model by incorporating trend elements. Finally, the Holt-Winters or TES method is used, incorporating trend elements and considering seasonal effects in the prediction. This holistic approach reflects our efforts to produce an accurate and reliable model for predicting weekly transaction movements, with the aim of supporting informed decision-making in this area, without having to detail the underlying basic theory [21].

This research adopts a forecasting approach with the Single Exponential Smoothing (SES) method, which is identified as a simple and effective method for time series data. SES considers exponential weights of past data in the forecasting calculation for the next period. These weights play an important role with decreasing importance over time, enabling the achievement of a high level of accuracy in forecasting trends or patterns in time series data. Despite its simplicity, SES provides satisfactory results, especially for time series that exhibit consistent trends and have no obvious seasonal patterns. The configurations of model parameters are detailed in Table 3.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>alpha</td>
<td>0.91</td>
</tr>
<tr>
<td>step</td>
<td>46</td>
</tr>
</tbody>
</table>

Source: (Research Results, 2024)

The basic formula for forecasting in SES is:

$$F_t + 1 = a * X_t + (1 - a) * F_t$$  \hspace{1cm} (1)

Formula 1 is used to determine the suitable parameters used in the SES model before the model training process is carried out.

This research combines the Single Exponential Smoothing (SES) method with Double Exponential Smoothing (DES) and Triple Exponential Smoothing (TES) or Holt-Winter techniques. This combination is done to expand the forecasting framework, with DES incorporating a trend component to see trend shifts in the time series data, and TES considering a seasonal component in addition to the trend. This approach is designed to improve prediction accuracy on weekly transaction data, allowing for more comprehensive modeling of trends, seasonal shifts, and patterns inherent in the dataset at hand. The specifics of the model's parameter configurations are outlined in Table 4.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>alphas</td>
<td>0.81</td>
</tr>
<tr>
<td>beta</td>
<td>0.11</td>
</tr>
<tr>
<td>step</td>
<td>46</td>
</tr>
</tbody>
</table>

Source: (Research Results, 2024)

Double Exponential Smoothing involves two smoothing parameters, namely alpha (α) for the level component (level smoothing) and beta (β) for the trend component (trend smoothing). The formulas for forecasting and updating levels and trends are as follows:

$$St = a * Yt + (1 - a) * (St - 1 + bt)$$

$$bt = y * (St - St - 1) + (1 - y) * bt$$

$$- 1$$

$$- 1$$
\[ F_t + m = S_t + b_t m \] (2)

Formula 2 is used to determine the suitable parameters used in the DES model before the model training process is carried out.

TES, alternatively referred to as Holt-Winters Exponential Smoothing, represents a more sophisticated forecasting approach compared to DES. It advances upon DES by incorporating a seasonal component in the forecast, rendering it well-suited for capturing recurring patterns within time series data. TES is particularly effective when the data exhibits distinct seasons or cycles that can influence the trajectory of the trend.

In Triple Exponential Smoothing, the forecasting process involves three smoothing parameters: alpha (\(\alpha\)) for the level component (level smoothing), beta (\(\beta\)) for the trend component (trend smoothing), and gamma (\(\gamma\)) for the seasonal part (seasonal smoothing). The forecasting formula and updates of the level, trend, and seasonal components are computed using these three parameters. This model demonstrates optimal performance when a distinct seasonal pattern is present in the time series data, resulting in more accurate forecasting outcomes and adaptability to intricate seasonal variations. Therefore, Triple Exponential Smoothing is well-suited for scenarios where the data displays a seasonal pattern that requires consideration. Details of the model's parameter settings are outlined in Table 5.

### Table 5. Holt-Winter Parameter

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>alphas</td>
<td>0.7</td>
</tr>
<tr>
<td>beta</td>
<td>0.1</td>
</tr>
<tr>
<td>gamma</td>
<td>0.5</td>
</tr>
<tr>
<td>step</td>
<td>46</td>
</tr>
</tbody>
</table>

Source: (Research Results, 2024)

The formula to calculate using the triple Exponential Smoothing method is as follows:

\[
\begin{align*}
    L_t &= a \times (Y_t - S_t - m) + (1 - a) \\
    & \quad \times (L_{t-1} + T_t - 1) \\
    T_t &= \beta \times (L_t - L_{t-1}) + (1 - \beta) \\
    & \quad \times T_{t-1} \\
    S_t &= \gamma \times (Y_t - L_t) + (1 - \gamma) \times S_{t-1} - m \\
    Y_t + h &= (L_t + h \times T_t) + S_t - m + h\%m
\end{align*}
\] (3)

Formula 3 is used to determine the suitable parameters used in the Holt-Winter model before the model training process is carried out.

#### E. Performance Metric

To evaluate and compare the effectiveness of the methods employed, equation (4) is utilized to calculate the MAE for a sample of N data points. The MAE serves as a metric for assessing the average magnitude of prediction errors, regardless of their direction, and is employed as the primary evaluation measure [22]. Furthermore, the evaluation of the model also incorporates prediction graphs, offering a comprehensive visual representation of the model's performance. These graphs provide direct insights into the extent to which the model accurately replicates the trends and patterns inherent in the data [23]. By employing both MAE calculation and Prediction Graph visualization in tandem, the model evaluation becomes more comprehensive, offering a more nuanced and profound perspective on the model's proficiency in predicting observed phenomena.

\[
    MAE = \frac{1}{n} + \sum_{i=1}^{n} |A_i - F_i| \] (4)

Formula 4 is used to obtain the mean absolute error value after training the model. This value is used to test how much error the model makes.

#### F. Results

In this research, we analyze by utilizing forecasting methods such as SES, DES, and TES (Holt-Winter) to forecast future sales in the e-commerce industry based on transaction data. We tune the model parameters according to the previously described configurations. To assess the model's performance, we utilize two evaluation metrics, namely the Mean Absolute Error (MAE) and the visualization of the prediction graph. Each model has been adjusted with the best parameters found through the parameter tuning process. Table 6 provides a thorough overview of the MAE performance scores achieved by each model. The model training was carried out utilizing Google Collaboratory, a cloud-based development platform. By comparing these optimization scores, we can evaluate the effectiveness of each model in predicting the number of sales transactions in e-commerce companies.

### Table 6. Model performance results

<table>
<thead>
<tr>
<th>Model</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES</td>
<td>1077</td>
</tr>
<tr>
<td>DES</td>
<td>96</td>
</tr>
<tr>
<td>TES(Holt-Winter)</td>
<td>101</td>
</tr>
</tbody>
</table>

Source: (Research Results, 2024)

Analysis using forecasting methods using the three methods results in the MAE value of the
results of the training model for more details can be seen in table 6. The table shows that the SES model provides an Average Absolute Error (MAE) value of 1077. On the other hand, the DES model produces an MAE of 96, while the TES or Holt-Winter model achieves an MAE of 101. Table 6 provides an overall picture of the performance of the three models in forecasting the number of sales transactions in e-commerce companies. Based on these findings, it can be concluded that the DES model performed the best, characterized by the lowest MAE, underscoring its accuracy in forecasting future sales trends. Additional model evaluation is done through prediction graphs, which are depicted in Figure 4.

**CONCLUSION**

The conclusion of this study based on the model evaluation results using Mean Absolute Error (MAE) shows that the test results show MAE SES 1077, DES 96, and TES (Holt-Winter) 101. Although the DES model produces a lower MAE when considered with the visual effect of the prediction graph, the Triple Exponential Smoothing (TES) model is superior in predicting sales in the e-commerce industry. Although DES exhibits a high level of numerical accuracy, TES can better capture the dynamics and patterns of actual sales, as reflected by its ability to generate predictions that match the real data. Therefore, the TES model can be effectively used for sales forecasting in e-commerce, offering a balanced rigor between numerical performance and the model's ability to understand and predict patterns of change in complex sales data. In this context, the TES model outperforms the SES and DES models. Suggestions for future research could further develop the TES model by exploring parameter variations, such as different decay rates for trend and seasonal components.

**REFERENCE**


