COMPARISON OF SEASONAL TIME SERIES FORECASTING USING SARIMA AND HOLT WINTER’S EXPONENTIAL SMOOTHING (CASE STUDY: WEST SUMATRA EXPORT DATA)

Lilis Harianti Hasibuan1*, Syarto Musthofa2, Darvi Mailisa Putri3, Miftahul Jannah4

1,2,3,4 Mathematics Study Program, Faculty of Science and Technology, UIN Imam Bonjol Padang, Sungai Bangek Street, Padang, 25171, Indonesia

Corresponding author e-mail: * lilisharianti@uinib.ac.id

ABSTRACT

Export is the activity of selling goods or services from one country to another. This activity usually occurs in a specific region or country. Export data is a type of data that has a seasonal pattern. This study aims to compare SARIMA and Holt Winter’s methods in forecasting export data. In this study, the SARIMA model ((1,1,1) (0,1,1))12 and Holt Winter’s simulation were obtained. The data used is the export data of West Sumatra from 2016 to 2022. The best model is the one with the smallest MAPE or MAD. The SARIMA model yielded a MAPE of 0.437% and MAD of 78.821. Meanwhile, the Holt Winter’s method yielded a MAPE of 0.894% and MAD of 163.320 with α=0.2, β=0.5, γ=0.1. Therefore, the SARIMA outperformed the Holt Winter’s method due to its higher accuracy. It can be concluded that the SARIMA is suitable to use as the forecasting model in this case. In this study, forecast have been made for the next 24 periods, from January 2023 to December 2024.
1. INTRODUCTION

International trade activities stimulate domestic demand, leading to the growth of large-scale industries, along with stable political structures and flexible social institutions [1]. Export refers to the purchase of goods produced by domestic companies in other countries. The most important factor determining exports is the ability of a country to produce competitive goods in the international market [2]. Each province in Indonesia has its own locally produced goods or natural resources. West Sumatran Province is a province that possesses abundant natural resources on land and at sea. This encourages export activities, which contribute to economic growth. However, the export value of West Sumatra exhibits a seasonal pattern. At certain times, the export value experiences increases or decreases. This phenomenon is referred to as seasonal data in statistics.

The export value of West Sumatra increases in August and decreases in February, influenced by several factors. However, these events need to be forecasted to ensure the availability of needs within the Province. The forecasting model used for seasonal data is known as SARIMA [3]. SARIMA is a model constructed from a stochastic process to model data that exhibits seasonality [4]. Another widely used forecasting method for seasonal and trend data is the Holt Winter's method of smoothing [5]. Forecasting is a method to estimate future values using past data. Export data is a time series data, and by modeling past data, it can be used to forecast future data.

The aim of this research is to forecast export data, which is a monthly time series data from 2016 to 2022. The forecast will be conducted for the next two years using SARIMA and Holt Winter's methods. This study specifically compares these two methods to obtain the best results for the seasonal export data. Previous studies on forecasting using these two methods have been conducted by researchers, such as [6], who predicted sales of clothing with seasonal data. In 2021, a comparison between SARIMA and Holt Winter's methods was conducted for the average monthly salary of a company in Cuba [7]. Other studies include forecasting energy consumption for smart grid operations [8], using SARIMA and LSTM on time series data, where the best model among the two methods produced the smallest error. Previous studies used seasonal data with fluctuating patterns at specific times, requiring modeling with seasonal models [7] [9] [10]. It is observed that the export value of West Sumatra increases after festive seasons due to the increased demand for West Sumatra's agricultural products from foreign countries. In the forecasting of export data, which is a time series data, SARIMA (Seasonal Autoregressive Integrated Moving Average) method can accurately predict the actual movement of export data. On the other hand, Holt Winter's method performs well on data with fluctuations around the average. This study will compare SARIMA and Holt Winter's methods, taking into account the seasonal pattern, to generate more accurate and close-to-actual forecasts.

The main objective of this research is to choose the best model through a comparison of the two methods based on the best accuracy values, measured by the smallest MAPE and MAD values for each method used. These results can serve as a reference for decision-making and policy-making by the government in predicting exports in a province. They can also be considered by business entities, particularly those involved in export activities.

2. RESEARCH METHODS

SARIMA is useful in situations when the time series data exhibit seasonality-periodic fluctuations that recur with about the same intensity each year. This characteristic makes the SARIMA model adequate for studies concerning monthly dengue data, given that the number of dengue cases in a population tends to be subject to seasonal variations, with a maximum in the export season and a minimum during the export season [11]. Seasonal Autoregressive Integrated Moving Average or better known as SARIMA method is Time Series forecasting method for stochastic model data with seasonal data pattern [12].

In time series forecasting models, the data tend to follow a pattern which shows a tendency to increase or decrease. Holt-Winter’s Exponential Smoothing method is a forecasting method with an exponential smoothing approach based on forecasting results in the previous period. This method also adds parameters to handle seasonal data patterns. There are two main models in Holt-Winter's Exponential Smoothing method, namely multiplicative model and additive model. The determination of this model was chosen based on the seasonal pattern [13].

In time series forecasting techniques, it is not uncommon for data to show trend data and seasonal data.
Where the data pattern has a tendency to fluctuate over a certain period. In the time series known forecasting method with Holt Winter’s exponential smoothing approach. The parameters of this method are used for seasonal patterns. The Holt Winter method consists of two models, namely the multiplicative model and the additive model. This model is often used for seasonal data.

A comparison between SARIMA and Holt Winter's methods had been studied for the average monthly salary of a company in Cuba. The major findings showed that Holt–Winters method had better performance in reproducing the mean series seasonality when the observations were insufficient, while for longer observations subsets, both were equally competitive in a short term forecasting. On the other side SARIMA model were found to be more reliable for longer lead-time forecasts. Hence, there is no general conclusion that places the models in a hierarchical order regarding their performance, it depends on the data used in the study [7].

2.1 SARIMA Model

Non-stationary series can be modeled using the Seasonal Autoregressive and Integrated Moving Average (SARIMA) model with orders (p, d ,q)(P,D,Q) [6]. SARIMA is a Time Series forecasting method for stochastic model data with seasonal data pattern [14]. This model can (finally) be combined also with seasonal differences of order s. Just as consecutive data points might show AR, MA or mixed ARIMA properties, so might data separated by a whole season show the same properties. The ARIMA notation is extended to such seasonal components thus [14].

\[
\text{ARIMA } (p, d, q) (P, D, Q) \ s \\
\]

with:

- \( p, d, q \) : The non seasonal part of the model
- \((P, D, Q) \ s \) : The seasonal part of the model
- \( s \) : Number of periods per season

\( P, D \) and \( Q \) can be further described in a way similar to the corresponding terms in the nonseasonal framework. The mathematical expression of this model in terms of the backshift notation can be written [12] as:

\[
(1 - \Phi_1 B - \cdots - \Phi_p B^p)(1 - \Phi_1 B^s - \cdots - \Phi_p B^{sp})(1 - B)^d (1 - B^s)^{Q} Y_t \\
= 1 - \theta_1 B - \cdots - \theta_q B^q)(1 - \Theta_1 B^s - \cdots - \Theta_q B^{sq}) \epsilon_t
\]

and is called (Box-Jenkins) ARIMA models. The model provides a powerful approach to time series analysis that pays special attention to correlation between observations, a feature ignored by the more basic forecasting models. However, the simpler techniques should not be ignored.

2.2 Holt-Winter’s

The Holt-Winter’s is a time series forecasting method that uses exponential smoothing to make prediction based on past observations which have seasonality tendency. In exponential smoothing, two commonly used components of a time series are employed, namely level and trend. The method is used when there is a trend in the data. However, to handle data that has a tendency of seasonality, the third parameter should be added. The method which considered three components of a time series yields a set of equations called the “Holt-Winters” (HW) method. Holt-Winter's Method is divided into two types namely Multiplicative Seasonal Method and Additive Method. Multiplicative method should be used when variability is amplified with trend, whereas the additive approach is more suitable in the case of constant variability. Whenever in doubt, both methods can be implemented and the best one chosen based on the error values obtained [15].

2.2.1 Holt Winter’s Method Multiplicative seasonality

The formula Holt Winter’s:

\[
L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + b_{t-1}) \\
b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1} \\
S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-s}
\]
\[ F_{t+m} = (L_t + b_t m)S_{t-s+m} \]

With \( s \) is the number of periods in one cycle of seasons example number of month. To initialize complete cycle of data values. Then set

\[ L_s = \frac{1}{s} (Y_1 + Y_2 + \cdots + Y_s) \quad (3) \]

initialize trend use \( s + k \) time periods

\[ b_s = \frac{1}{k} \left( \frac{Y_{s+1}-Y_1}{s} + \frac{Y_{s+2}-Y_2}{s} + \cdots + \frac{Y_{s+k}-Y_k}{s} \right) \quad (4) \]

If the series is long enough then a good choice is to make \( k = s \) so that two complete cycles are used. However we can, at a pinch, use \( k = 1 \). Initial seasonal indices can be taken as

\[ S_k = \frac{Y_k}{L_s}, k = 1, \ldots, s. \quad (5) \]

The parameters \( \alpha, \beta, \gamma \) should lie in the interval \((0, 1)\), and can be selected by minimizing MAD, MSE or MAPE.

2.2.2. Holt Winter’s Method Additive seasonality

The equations are: [14]

\[
\begin{align*}
L_t &= \alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + b_{t-1}) \\
b_t &= \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1} \\
S_t &= \gamma(Y_t - L_t + (1 - \gamma)S_{t-s}) \\
F_{t+m} &= (L_t + b_t m)S_{t-s+m}
\end{align*}
\]

where \( s \) is the number of periods in one cycle. The initial values of \( L_s \) and \( b_s \) can be as in the multiplicative case. The initial seasonal indices can be taken as

\[ S_k = Y_k - L_s, k = 1, \ldots, s. \quad (7) \]

The parameters \( \alpha, \beta, \gamma \) should lie in the interval \((0, 1)\), and can be selected based on the smallest value of MAD, MSE or MAPE.

3. RESULTS AND DISCUSSION

3.1 SARIMA Model

3.1.1 Model Identification.

The model identification process is conducted by plotting the data. The data used in this analysis is the production data from January 2016 to December 2022. Firstly, we check for the presence of seasonal patterns in the data. Then, we examine whether the data is stationary or not by observing the plots of the Autocorrelation Function (ACF) and Partial Autocorrelation (PACF).

Figure 1. (a) Time Series Plot of West Sumatra's Export Data, (b) Box-Cox Transformation process

Figure 1 (a) It can be observed that West Sumatra’s export data exhibits a seasonal pattern occurring around the month of August each year. In Figure 1 (b), it can be seen that the result of the Box-Cox
Transformation process indicates that the data becomes stationary with respect to variance. Next, it will be examined whether the data stationary with respect to the mean using the Autocorrelation Function (ACF) plot and Partial Correlation Function (PACF).

**Figure 2.** (a) ACF Plot, (b) PACF Plot

Figure 2 shows that the data is already stationary with respect to the mean, as the first three lags are within the bounds of the red lines. Next, we will examine it using the Partial Autocorrelation Function (PACF) plot, as shown in the following figure. It can be observed that the ACF and PACF plots are not stationary with respect to the seasonal component that occurs every year, specifically at lag 12. Therefore, the seasonal data needs to be differenced. After performing differencing on the seasonal data, it can be seen that the seasonal component is now stationary.

**Figure 3.** (a) ACF plot of the seasonal data, (b) PACF plot of the seasonal data

It can be observed that the ACF and PACF plots are stationary with respect to the seasonal data. It is evident that the first three lags fall within the control limits.

### 3.1.2 Estimation Parameter.

The model estimation phase reviews the identified models from the first step and determines the best model. The potential best models are presented in Table 1.

<table>
<thead>
<tr>
<th>Model SARIMA $(p, d, q)(P,D,Q)^{12}$</th>
<th>Parameter Testing</th>
<th>White Noise Test</th>
<th>Residual Normality Test</th>
<th>$R^2_{adj}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(1,1,1)(0,1,1)^{12}$</td>
<td>Significant</td>
<td>White Noise (if ACF and PACF do not exceed the control limits)</td>
<td>Normal (P-Value $0.892 &gt; 0.05$)</td>
<td>0.832</td>
</tr>
<tr>
<td>$(0,1,1)(0,1,1)^{12}$</td>
<td>Significant</td>
<td>White Noise (if ACF and PACF do not cross the control limits)</td>
<td>Normal (P-Value $0.841 &gt; 0.05$)</td>
<td>0.784</td>
</tr>
<tr>
<td>$(1,1,1)(1,1,1)^{12}$</td>
<td>Significant</td>
<td>White Noise (if ACF and PACF do not cross the control limits)</td>
<td>Normal (P-Value $0.826 &gt; 0.05$)</td>
<td>0.657</td>
</tr>
</tbody>
</table>
In the Table 1, all three candidate models meet all the criteria, including being significant, exhibiting white noise, and normality. Since these three models fulfill all these assumptions, the model with the highest $R^2_{adj}$, SARIMA ((1,1,1) (0,1,1)$^{12}$, is chosen [16].

3.1.3 Forecasting with SARIMA.

After determining the best model, we will use that model for forecasting. In SARIMA time series forecasting, there are two methods: the Multiplicative model and the Additive model.

a. Multiplicative Model

![Figure 4](image1.png)

**Figure 4.** Multiplicative Model for Seasonal Export Data, (a) Seasonal Indices, (b) Detrended Data by Season, (c) Percent Variation by Season, (d) Residuals by Season

From Figure 4, it can be observed that the data exhibits seasonality, trend of the data with respect to the seasonality, variability percentage with respect to the seasonality, and residuals with respect to the seasonality. After depicting these aspects, the forecasting will be performed using both models.

![Figure 5](image2.png)

**Figure 5.** Forecasting multiplicative model

In Figure 5, it can be observed that the fitting of the data closely resembles the actual data. Subsequently, forecasting is conducted with MAPE value of 27,240 and MAD value of 41,070. Then, we will compare it with the forecasting using an additive model, as shown in the image below.
It appears that in Figure 5 and Figure 6, the data fitting is very similar to the actual data. After performing the data fitting, the MAPE and MAD will be calculated for both SARIMA models as follows:

Table 2. Calculation of MAPE and MAD for SARIMA model ((1,1,1)(0,1,1)\textsuperscript{12})

<table>
<thead>
<tr>
<th>SARIMA MODEL ((1,1,1) (0,1,1)\textsuperscript{12})</th>
<th>MAPE</th>
<th>MAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiplicative Model</td>
<td>27.240</td>
<td>41.070</td>
</tr>
<tr>
<td>Additive Model</td>
<td>29.530</td>
<td>42.870</td>
</tr>
</tbody>
</table>

From Table 2, it can be seen that the best model is the multiplicative model, which yields the smallest values of MAPE and MAD. Therefore, this model can be used for forecasting seasonal data. Here are the forecasted results generated by the SARIMA model for the export data of West Sumatra.

Table 3. Forecasting Results of the Model (SARIMA)((1,1,1)(0,1,1))\textsuperscript{12} for the Next 3 Months

| Period 2023 | Actual Value \((A_t)\) | Forecast \((F_t)\) | Error Absolute \(|A_t - F_t|\) | Square Of Error \(|(A_t - F_t)|^2\) | Absolute Values Of Errors \(\frac{|A_t - F_t|}{A_t}\) |
|-------------|-----------------|-----------------|------------------|-----------------|------------------|
| January 2023 | 165,470         | 206,727         | 41,257           | 1702,140        | 0.249            |
| February 2023| 222,440         | 209,504         | 12,936           | 167,340         | 0.058            |
| March 2023   | 189,550         | 214,178         | 24,628           | 606,530         | 0.130            |
| Error values |                 |                 |                  |                 |                  |
| MAD          | 78.821          |                 |                  |                 |                  |
| MSE          | 2476,010        |                 |                  |                 |                  |
| MAPE         |                 |                 |                  |                 | 0.437%           |

Table 3 shows the forecast results using the SARIMA ((1,1,1) (0,1,1))\textsuperscript{12} model for the next three months, namely January 2023 to March 2023. The forecasted values have a MAD of 78.821, MSE of 2476,010, and MAPE of 0.437%.

3.1.4 Holt-Winter's Exponential Smoothing method

The steps for forecasting with Holt-Winter's Exponential Smoothing method are as follows: analyzing the data to determine if it contains trend and seasonal components by examining the patterns formed, then using Holt-Winter's Exponential Smoothing with either the seasonal multiplicative or seasonal additive method to predict the data, and finally comparing the error values between the two methods to find the smallest error. Based on the comparison between the multiplicative and additive methods, it is evident that the multiplicative method yields the smallest error. Before conducting the forecasting, the values of alpha, beta, and gamma used in the model should be between 0 and 1. This testing is done through trial and error. The following are the candidate models from Holt Winter.

Table 4. Comparison of MAPE and MAD Values for the Multiplicative Model
From Table 4, it can be seen that the smallest values of MAPE and MAD are obtained when $\alpha=0.2$, $\beta=0.5$, $\gamma=0.1$. Next, with the same parameters, we will examine the additive model, as shown in Table 5 below:

**Table 5. Comparison of MAPE and MAD Values for Additive Model**

<table>
<thead>
<tr>
<th>Data</th>
<th>Additive Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parameter</td>
</tr>
<tr>
<td>Export Data of West</td>
<td>$\alpha = 0.9$, $\beta = 0.1$, $\gamma = 0.1$</td>
</tr>
<tr>
<td>Sumatra</td>
<td>$\alpha = 0.2$, $\beta = 0.5$, $\gamma = 0.1$</td>
</tr>
<tr>
<td></td>
<td>$\alpha = 0.9$, $\beta = 0.1$, $\gamma = 0.9$</td>
</tr>
<tr>
<td></td>
<td>$\alpha = 0.8$, $\beta = 0.3$, $\gamma = 0.4$</td>
</tr>
<tr>
<td></td>
<td>$\alpha = 0.5$, $\beta = 0.6$, $\gamma = 0.8$</td>
</tr>
</tbody>
</table>

From Table 5, it can be seen that the smallest values of MAPE and MAD are obtained with $\alpha=0.2$, $\beta=0.5$, and $\gamma=0.1$. Both the multiplicative and additive Holt-Winter’s models can be compared, and the model with the smallest MAPE and MAD values is the multiplicative Holt-Winter’s model. The plot below shows the Export data for both models with $\alpha=0.2$, $\beta=0.5$, and $\gamma=0.1$, which was selected as the model with the smallest MAPE and MAD.

**Figure 7. Forecasting with Multiplicative Model using $\alpha=0.2$, $\beta=0.5$, and $\gamma=0.1$**
From the two models, the model that will be used is the Multiplicative Holt-Winter model. Here is the forecast table for this model.

Table 6. Forecast Results for the Multiplicative Holt-Winter Model for the Next 3 Months

| Period       | Actual Value ($A_t$) | Forecast ($F_t$) | Error Absolute ($|A_t - F_t|$) | Square Of Error ($|A_t - F_t|^2$) | Absolute Values Of Errors ($|A_t - F_t|$) | Error Values |
|--------------|----------------------|------------------|-----------------------------|-----------------------------------|-------------------------------------------|--------------|
| January 2023 | 165,470              | 238,943          | 73,473                      | 5398,280                          | 0,444                                     | MAD          |
| February 2023| 222,440              | 253,199          | 30,759                      | 946,110                           | 0,138                                     | MSE          |
| March 2023   | 189,550              | 248,641          | 59,091                      | 3491,740                          | 0,312                                     | MAPE         |

From the two methods above, it can be seen that the SARIMA model produces the smallest MAPE and MAD values ((1,1,1) (0,1,1))^12. Therefore, we can perform forecasting for the next 24 periods using the SARIMA model. The forecasting data can be seen in the table below.

Table 7. Forecasting Results of West Sumatra Exports from January 2023 to December 2024

<table>
<thead>
<tr>
<th>Period</th>
<th>Actual Value ($A_t$)</th>
<th>Forecast ($F_t$)</th>
<th>Actual Value ($A_t$)</th>
<th>Forecast ($F_t$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 2024</td>
<td>-</td>
<td>-</td>
<td>January</td>
<td>210,646</td>
</tr>
<tr>
<td>February 2024</td>
<td>-</td>
<td>-</td>
<td>February</td>
<td>218,852</td>
</tr>
<tr>
<td>March 2024</td>
<td>-</td>
<td>-</td>
<td>March</td>
<td>222,469</td>
</tr>
<tr>
<td>April 2024</td>
<td>-</td>
<td>-</td>
<td>April</td>
<td>237,257</td>
</tr>
<tr>
<td>Mei 2024</td>
<td>-</td>
<td>-</td>
<td>Mei</td>
<td>117,674</td>
</tr>
<tr>
<td>Juni 2024</td>
<td>-</td>
<td>-</td>
<td>Juni</td>
<td>210,655</td>
</tr>
<tr>
<td>Juli 2024</td>
<td>-</td>
<td>-</td>
<td>Juli</td>
<td>224,336</td>
</tr>
<tr>
<td>Agustus 2024</td>
<td>-</td>
<td>-</td>
<td>Agustus</td>
<td>312,261</td>
</tr>
<tr>
<td>September 2024</td>
<td>-</td>
<td>-</td>
<td>September</td>
<td>216,921</td>
</tr>
<tr>
<td>Oktober 2024</td>
<td>-</td>
<td>-</td>
<td>Oktober</td>
<td>235,957</td>
</tr>
<tr>
<td>November 2024</td>
<td>-</td>
<td>-</td>
<td>November</td>
<td>210,018</td>
</tr>
<tr>
<td>Desember 2024</td>
<td>-</td>
<td>-</td>
<td>Desember</td>
<td>239,631</td>
</tr>
</tbody>
</table>
To evaluate the model that has been obtained using the entire data without dividing the data into training data and testing data, namely by fulfilling the assumptions of stationary on the mean and stationary on the variance. This assumption can be seen from the ACF plot and PACF plot of the data. After these assumptions are met, it will be estimated with the existing parameters, and the best model is the model that gives the smallest error. This research focuses on real export data which contains seasonal data.

4. CONCLUSION

Based on the discussion, the following conclusions can be drawn:

1. SARIMA and Holt Winter’s models can be used when there is a seasonal pattern in the data.

2. The use of SARIMA \((1,1,1)(0,1,1)\)\(^{12}\) method for forecasting the export data of West Sumatra produces more accurate forecasts compared to Holt Winter’s method, as indicated by the smallest MAPE value of 0.437% obtained by SARIMA.

3. The implementation results of comparing these two methods in forecasting the export data of West Sumatra can be used as a reference for decision-making in determining policies.

ACKNOWLEDGMENT

Authors would like to express gratitude to the Faculty of Science and Technology, UIN Imam Bonjol Padang for the support toward this research project.

REFERENCES


