Comparison of Exponential Smoothing Methods in Forecasting Palm Oil Real Production

B Siregar[1], I A Butar-Butar[1], RF Rahmat[1], U Andayani[1] and F Fahmi[2]

[1]Dept. of Information and Technology, Faculty of Computer Science and Information Technology, University of Sumatera Utara Jl. Dr. Mansur 9 Medan Indonesia
[2]Dept. of Electrical Engg, Faculty of Engineering, University of Sumatera Utara Jl. Dr. Mansur 9 Medan Indonesia

Email: fahmimn@usu.ac.id

Abstract— Palm oil has important role for the plantation subsector. Forecasting of the real palm oil production in certain period is needed by plantation companies to maintain their strategic management. This study compared several methods based on exponential smoothing (ES) technique such as single ES, double exponential smoothing holt, triple exponential smoothing, triple exponential smoothing additive and multiplicative to predict the palm oil production. We examined the accuracy of forecasting models of production data and analyzed the characteristics of the models. Programming language R was used with selected constants for double ES (α and β) and triple ES (α, β, and γ) evaluated by the technique of minimizing the root mean squared prediction error (RMSE). Our result showed that triple ES additives had lowest error rate compared to the other models with RMSE of 0.10 with a combination of parameters $\alpha = 0.6$, $\beta = 0.02$, and $\gamma = 0.02$.

1. Introduction

Palm oil is one of plantation crops that have an important role for the plantation subsector. Industrial development of palm oil benefit in increasing the income of farmers and society, production of the raw material that create added value for the country, the export of CPO which generate foreign exchange, and it provides job opportunities for more than two million workers in the various subsystems.

Palm oil plantations in Indonesia are managed in the form of small plantations and large farms. Large estate consists of a country estate as a Perkebunan Nusantara (PTPN) and private estates [1].

For each plantation, there will be monthly evaluation processing regarding quality objectives. This was done to look at the performance of the production for the month whether it has reached a production target or not. In order to achieve the production target, the availability of information about real production in previous years plays a very important role. Target production in the future can be expected only if the previous year's production data is available and of good quality. In the plantation production targets are calculated by looking at the production of previous years done by manual calculations. Therefore we need a method that can provide information on production forecasts as a reference in determining the company plans more realistic production targets.

Research on Palm oil production forecasts have been done by several methods. Methods existed included the use methods of causal and time-series. Research conducted in 2013 using two methods to compare the production forecasts Palm oil is used in the form of multiple regression methods of causal and time-series methods such as exponential smoothing. In the test, the value of Mean Absolute Percentage Error (MAPE) obtained using regression is 21:08% while the value of MAPE obtained using exponential smoothing method is 12.78% per year [2].

In another study, a research was conducted on the production prediction using neural networks with back propagation algorithm. Usually neural network was utilized in image based analysis such as meter reading [3] or image recognition [4]. The study used seven data as a variable based on the quality of the land as well as rainfall, altitude, and slope, age of the plant, rocks, solium, and the acidity of the soil. Experiment with several layers to get the best parameters: 3 layer, 4 layer and 5 layer. The best results obtained in experiment with 3 layer iteration to 30,000, with the learning rate by 0.9. Training results obtained with R2 = 0.9998 and RMSE = 0.0709 and test results with R2 = 0.8901 and RMSE = 2.2196 [5].

Another research in 2015 predicted Palm oil production using RBF neural network. The best results for Palm oil production forecasts obtained through a combination of parameters: the number of variables = 1, the number of input nodes = 5, the value of learning rate = 0.75 and the maximum epoch value = 3000 with the results of error MAPE = 11.75% [6].

Another research in 2011 utilized exponential smoothing to forecast sea surface temperatures, the researchers predicted the next year based on the results of the model were compared to analyze the characteristics of the model and also comparing the values obtained using three methods: Single Exponential Smoothing, Double exponential smoothing Holt and Holt -winters. Of the three methods produced the best performance on the method and Holt-Winters exponential smoothing Double Holt [7].

In this study we compared several exponential smooothing methods used to predict the actual production of palm oil such as single ES, double exponential smoothing holt, triple exponential smoothing, triple exponential smoothing additive and multiplicative and analyze the performance.

2. Material and Methods

The general architecture of this study can be seen in Figure 1.

2.1 Data Collection

This section shows the retrieval of data, then the data transformation process to the creation of a data plot. The dataset used in this study is a document from Palm oil production. The data used was secondary data, i.e data Palm oil plantation PTPN III Plantations Sei Meranti which have been collected from 2010-2014. The data was stored in the form of csv and tables that no normalization of data until the data can be included in a database or data in comma-separated values extension * .csv.

Collecting data in this study was using survey techniques, performed directly in the office Factory Palm oil (MCC) Sei Plantations Meranti. The data consisted of data from the production of fresh fruit bunches (FFB) Palm oil, palm oil production data and core data obtained from FFB oil that had been processed.

The next stage was the transformation of data, where the data was normalized so that data values can be processed easily. These data were converted into a value in the range of 0.1 up to 0.9 by using equation 1 [7].

$$x' = \frac{0.8\,(x-a)}{b-a} + 0.1 \qquad (1)$$

The next step was to create a data plot. Plot Palm oil production data in 2010-2014 in Plantations Sei Meranti which had been transformed can be seen in Figure 2. Based on Figure 2, Palm oil production starting from the year 2010 to 2014 were likely to increase. Only in early 2014 Palm oil production declined. Based on information from the management of the farm in 2014, there was a drastic climate change, so that not only the plantations Sei Meranti are decreased, but also almost all the plantations that exist in PTPN III experienced the same thing.

Overall it can be concluded that the production of oil in the plantation Palm Sei Meranti has a positive trend (increasing).

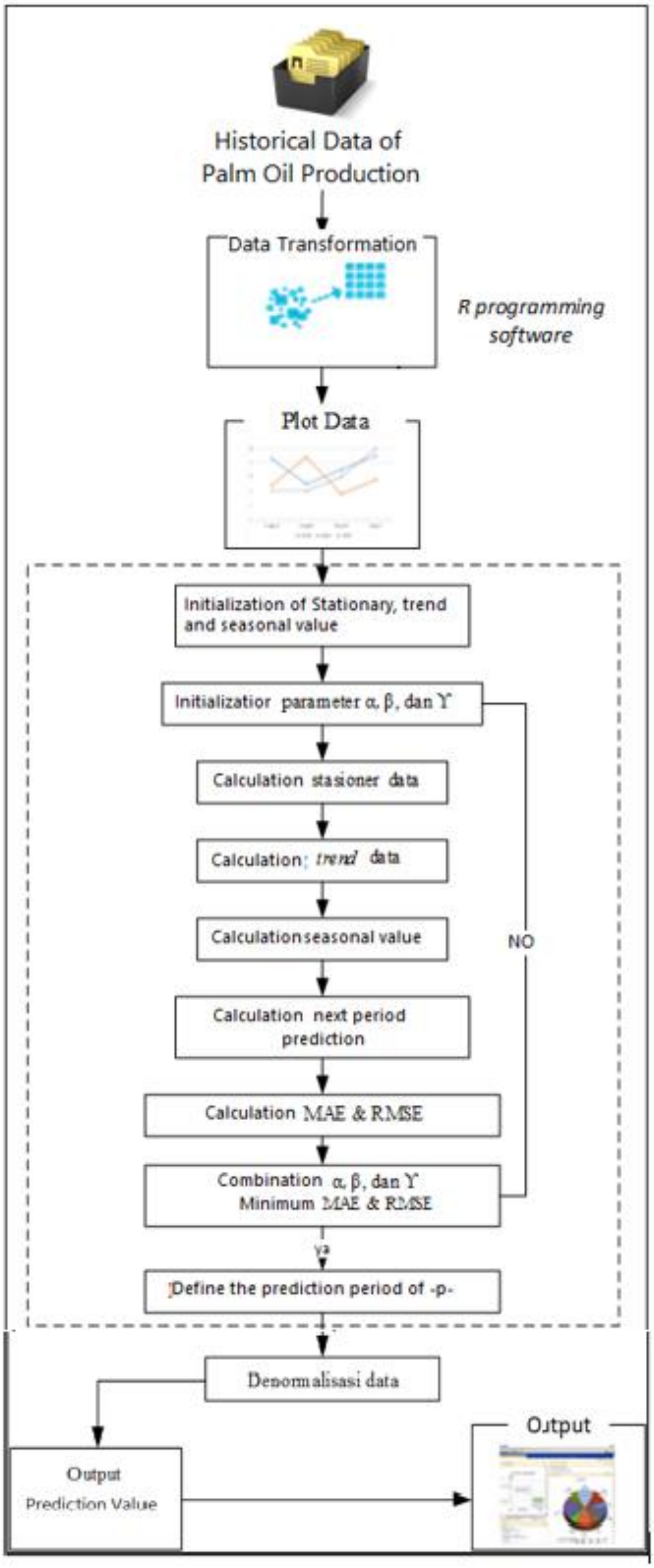

Figure 1. General Architecture

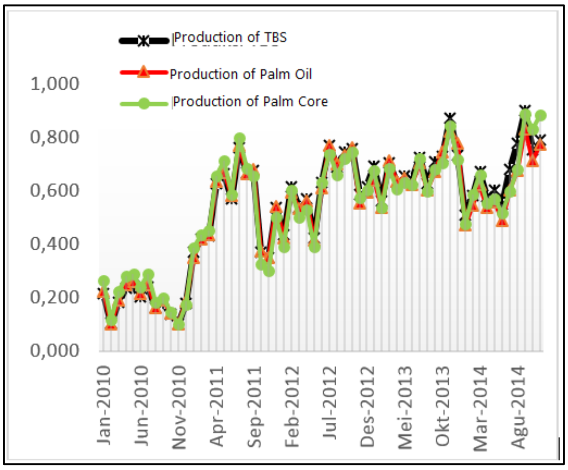

Figure 2. Graph production of Palm oil PTPN III Sei Meranti

2.2 Exponential Smoothing Method

2.2.1. Double exponential smoothing (DES) Holt. In DES Holt method, the smoothed trend component separately using different parameters, namely α and β. In this technique the value of the trend can be smoothed by using different weights. However, these two parameters need to be optimized so the search for the best combination of parameters is more complicated than using only one parameter.
In addition, the components of season in this technique are not taken into account. Here is the process undertaken Double exponential smoothing holt Holt.

Step 1: Determine the initial initialization value. Value is stationary, and the trend is beginning to use equation 2 to equation 3 [8].

$$S_L = X_L \text{ atau } S_L = \frac{1}{L} + (X_1 + X_2 + \ldots + X_L) \tag{2}$$

$$T_L = \frac{1}{L}\left\{\frac{(X_{L+1} - X_1)}{L} + \frac{(X_{L+2} - X_2)}{L} + \ldots + \frac{(X_{L+L} - X_L)}{L}\right\} \tag{3}$$

Step 2: Initialize the value of the parameters a and b where each range is between 0 - 1. In this study, we used the method of trial and error to produce the value of α and β are optimal based on the value Root Mean Square Error (RMSE) of the most minimum. Values α and β are optimal will be determined directly by the application program that has been designed.

Step 3: Calculation of stationary of data can be done using Equation 4 [8].

$$S_t = \alpha Xt + (1 - \alpha)(S_{t-1} + b_{t-1}) \tag{4}$$

\

Step 4: Calculation of trend data can be done using the equation 5 [8].

$$T_t = \beta(S_t + S_{t-1}) + (1 - \beta)b_{t-1} \tag{5}$$

Step 5: Once the data is stationary and trendy value has been established, then the forecasts can be found using equation 6 [8].

$$F_{t+m} = S_t + b_t m \tag{6}$$

Step 6: Then calculate the forecast errors using RMSE and MAPE. If MAPE produced not the smallest then repeat step two and so on until we got the smallest error value.
Step 7: Finding the forecasts for the next few periods, for example, for 12 months. By using the S parameter (level), and T (trend) of the last data specify the length of the forecast period p. By re-using Equation 6, set the value of m = 1, to find the forecasts one period or one month following the next. Suppose for 1 year ahead or equal to 12 periods ahead then set the value of m with m = 1, m = 2 and m = 3 and so on up to m = 12.

2.2.2 *Triple exponential smoothing Additive and Multiplicative.* From the analysis that has been done before, it was found that the amount of oil production Palm Plantations Sei Meranti from the first months of 2010 until the final months of 2014 was influenced by trends and seasonal factors. It is seen from Figure 2, in the seventh month the number of production Palm oil spiked higher then month-a month after the demand will decrease and then going back upward as it approached the seasonal period in the next year. This situation repeated every year, which means influenced by seasonal factors. Because the data is affected by seasonal factors, the other methods that can be implemented are Triple exponential smoothing method.
The following processes are carried out in Triple exponential smoothing (TES) Additive and Multiplicative.

Step 1: Determine the initial initialization value. ie stationary value, trend and seasonally initial value using equations 2, 3 and 7 [8].

$$I_L = \frac{X_k}{S_L}, k = 1,2,..., L \quad (7)$$

Step 2: Initialize the value of the parameters α, β and γ where each range is between 0-1.

Step 3: Calculation of TES data is stationary with additive and multiplicative method can be performed using equations 8 and 9 [8].

Aditive : $S_t = \alpha(X_t - I_{t-L}) + (1-\alpha)(S_{t-1} + T_{t-1})$ (8)

Multiplicative: $S_t = \alpha \frac{X_t}{I_{t-L}} + (1-\alpha)(S_{t-1} + T_{t-1})$ (9)

Step 4: Calculation of trend data can be performed using equations 10 and 11 [8].

Aditive : $T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}$ (10)

Multiplicative: $T_t = \beta(S_t + S_{t-1}) + (1-\beta)T_{t-1}$ (11)

Step 5: For the calculation of the initial value I, in one cycle of the first season (12 first period), can be performed using Equation 12 and 13 [8].

Aditive : $I_t = \gamma(X_t - S_t) + (1-\gamma)I_{t-L}$ (12)

Multiplicative: $I_t = \gamma \frac{X_t}{S_t} + (1-\gamma)I_{t-L}$ (13)

Step 6: With all three parameter values that have been obtained, then the forecasts can be found by Equation 14 [8].

$$F_{t+m} = (S_t + T_t m)I_{t-L+m} \quad (14)$$

Step 7: Then calculate the forecast errors using RMSE and MAPE. If MAPE produced not the smallest then repeat step two and so on until you get the smallest error value.

Step 8: Finding the forecasts for the next few periods, for example, for 12 months. By using the S parameter (level), T (trend), and I (seasonal) of the last data specify the length of the forecast period p by re-using Equation 14.

2.3 Output

After obtaining the approximate value of the desired period, the next step is to normalize the data, which returns the data into the form of the initial unit using Equation 15.

$$x = \frac{(x' - 0{,}1)(d_{maks} - d_{min})}{0{,}8 + d_{min}} \quad (15)$$

Then the results of normalize will be processed into the form of plots and tables to facilitate the user in observing the forecast. Implementation of this method is done using the software R Programming.

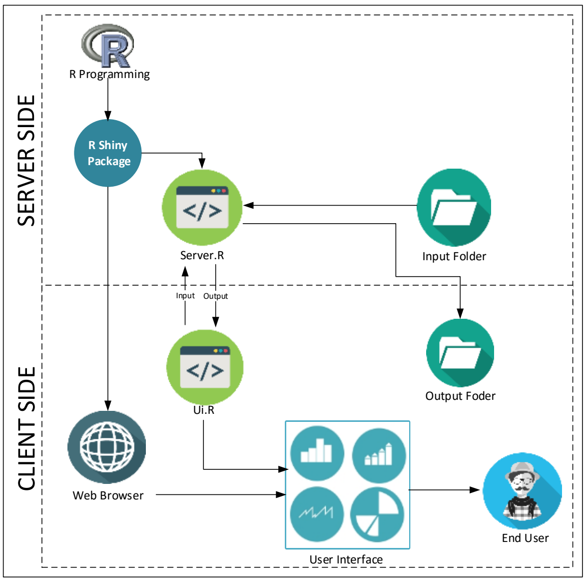

Figure 3. Process in R Programming

Based on Figure 3, there are two parts in the process of forecasting system using the R software, the server side and client side. The entire core code of the implementation process of exponential smoothing is stored in a file named server.R. The process of data input using csv file extension stored in the input folder, which is then called directly by server.R.

While user interface.R handle the process of how the output is displayed to the user. For the output of this system, the R language provides an easy way to integrate server.R and ui.R use shiny library. The results of exponential smoothing configuration can be accessed directly by the web browser by the user.

3. Results

Several tests were conducted to determine the performance of the system in forecasting realization Palm oil production using exponential smoothing method. Based on the draft research steps, then each data will be processed using three methods already mentioned. Testing parameters for all three methods with different combinations of parameters were used. Besides being able to manually select a combination of parameters, the system can also look for optimizing the parameters that produce the smallest error.

By executing the equations exponential smoothing models with the help of R software, a comparison was made between the results obtained form the model with the results of observations for each method. The data on production of fresh fruit bunches (FFB), a constant value of α, β and γ are determined by the system are respectively $\alpha = 0.7$ and $\beta = 0.01$ for Holt DES method, $\alpha = 0.6$, $\beta = 0$, 02 and $\gamma = 0.02$ for TES additive method and $\alpha = 0.5$, $\beta = 0.02$ and $\gamma = 0.02$ for TES multiplicative method. The parameter optimization results using these three methods can be seen in Figure 4.

Figure 4 shows that the model DES holt, TES additives and multipikatif was very precise, especially for some of the initial data. Forecast data generated by additive and multiplicative method of time January 2010 until December 2014 was very precise compared observation data.

Data FFB production is a seasonal cycle of data will be repeated the following year, the multiplicative and additive models try to predict the repetition of the production cycle.

The analysis was conducted based on the accuracy of the model DES, TES TES additive and multiplicative FFB production data using statistical instruments such as the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) [8]. RMSE, MAE and MAPE refers to the magnitude of the error rate (errors) of an estimate, the smaller the value of these instruments, the better the forecasts were made. The results of the evaluation of the forecasting models each dataset with all three methods mentioned can be seen in Table 1.

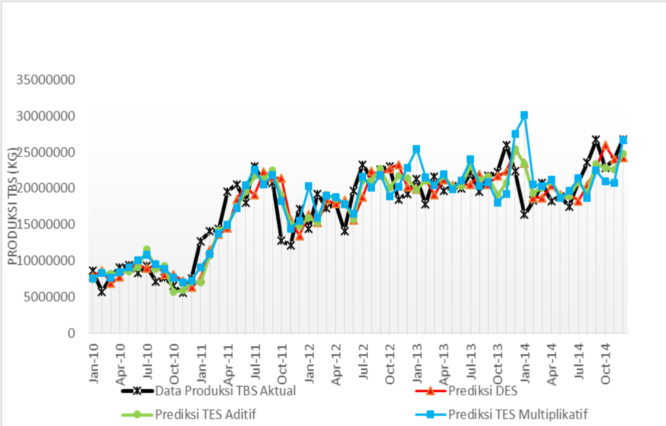

Figure 4. DES Holt, TES Aditive, and TES Multiplicative for production 2010-2014

Table 1 Error estimation for different forecasting models

Dataset	***Exponential smoothing***	**Estimated Error**		
		MAE	**RMSE**	**MAPE**
Production of FFB	Holt (α=0.7;β=0.01)	0.08	0.11	19.37
	Aditive (α=0.6;β=0.02; γ=0.02)	0.07	0.09	17.18
	Multiplicative (α=0.58;β=0.02 ; γ=0.02)	0.09	0.12	20.04
Production of Palm Oil	Holt (α=0.74;β=0.01)	0.09	0.11	20.79
	Aditive (α=0.7;β=0.01 ; γ=0.03)	0.07	0.10	18.06
	Multiplicative (α=0.64;β=0.02 ; γ=0.03)	0.09	0.12	21.31
Production of Palm Kernel	Holt (α=0.75;β=0.01)	0.09	0.11	21.9
	Aditive (α=0.68;β=0.01 ; γ=0.03)	0.08	0.10	18.8
	Multiplicative (α=0.72;β=0.01 ; γ=0.21)	0.10	0.13	22.9

An instrument of error such as MAE, RMSE, and MAPE (%) of the additive model is smaller than Holt and multiplicative models for data in seasonal cycles on the production data Plantations Sei Meranti.

Model triple exponential smoothing (Holt-Winters additive and multiplicative) has good precision, especially the additive method for data that are seasonal cycle such as production data FFB, palm oil, and palm kernel as shown in Table 4.

From Table 4 the exponential smoothing models, which is the method TES additive is the best method which produces the smallest error in forecasting for the data that is periodic with each value RMSE = 0:09 to production data FFB, RMSE = 0.10 for the data production of palm oil, and RMSE = 0.10 for the data production of palm kernel.

After the comparison of the forecasts of the results of field observations from the period January 2010 to December 2014 described Figure 4 and Table 1, then we made further forecasts. In the forecasts for the next two years ie from January 2015 to December 2016 for all three methods of exponential smoothing was provided. Advanced forecasting results can be seen in Figure 5 (a) to (c).

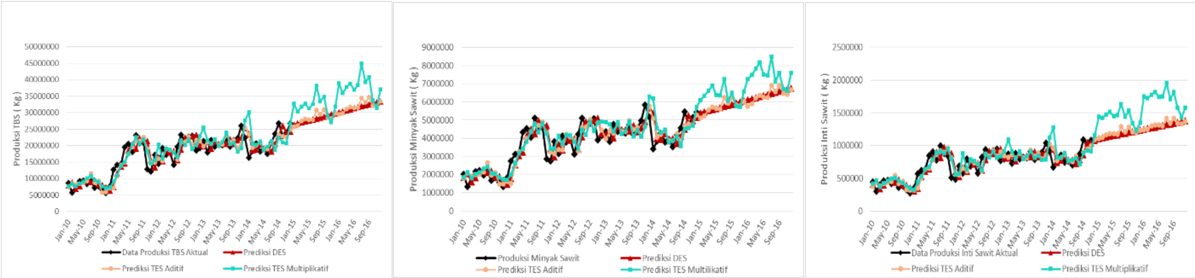

Figure 5. Graph forecasting for 2010-2016 (a) FFB (b) Palm Oil (c) Palm Kernel

4. Discussions

Based on the evaluation of system forecasting result it was obtained that the best results to forecast FFB production, palm oil, and palm kernel produced by the method of triple exponential smoothing additive through a combination of parameters in a row is α= 0.6, β = 0:02, and γ = 0:02 to FFB production with the results of error RMSE = 0:09, α = 0.7, β = 0:01, and γ = 0:03 to palm oil production with the results of error RMSE = 0:10, and α= 0.68, β = 0:01, and γ = 0:03 to palm kernel production with the result of error RMSE = 0.10.

From the test results it is shown that each of the data output has the characteristics of time series data differently that each of the data output forecasts have different parameters.

The use of double exponential smoothing method holt to predict what kind of data Palm oil production is not appropriate because the observation data has a seasonal component. Forecasts for the data containing seasonal components should be solved by using the method of triple exponential smoothing.

For further development, the study suggested the same case can use other methods that produce results forecast by the smaller error value. The forecasting is a reference to be considered in setting production targets that required high accuracy.

5. Conclusion

Compared to other exponential smoothing methods, Triple exponential smoothing additive can provide better forecasting of palm oil real production with suggested parameters as the result from this study.

References

[1] Batubara EA 2015 Implementasi algoritma learning vector quantization pada prediksi produksi Palm oil di PT. Perkebunan Nusantara I Pulau Tiga u*nder publication*

[2] Kacaribu RA 2013 Aplikasi Peramalan Produksi Palm oil dengan Metode Regresi Ganda dan *Exponential Smoothing*. u*nder publication*

[3] Triantoro T, Batubara FR and Fahmi F 2014 Image based water gauge reading developed with ANN Kohonen Proceedings of 2014 *International Conference on Electrical Engineering and Computer Science, ICEECS*

[4] Siswanto A, Tarigan P and Fahmi F 2013 Design of contactless hand biometric system with relative geometric parameters *Proc. of 2013 3rd Int. Conf. on Instrumentation, Communications, Information Technol., and Biomedical Engineering: Science and Technol. for Improvement of Health, Safety, and Environ., ICICI-BME*

[5] Hermantoro and Purnawan RA 2009 Prediksi produksi Palm oil berdasarkan kualitas lahan menggunakan model Artificial Neural Network (ANN) *Agroteknose* **4(2)**: 1-6.

[6] Jannati R 2015 Prediksi produksi panen Palm oil menggunakan jaringan saraf *radial basis function* (RBF) u*nder publication*
[7] Salamena GG 2011 Pengujian model peramalan deret waktu *sea surface temperature* (SST) teluk ambon luar dengan metode exponential smoothing *Oseanologi dan Limnologi di Indonesia* **37(1)**: 43-55.
[8] Maranata L 2013 Peramalan jumlah produksi kakao di sumatera utara dan konsumsi kakao di Indonesia dengan pemulusan eksponensial ganda metode linier satu parameter dari brown. u*nder publication*
[9] Siang JJ 2009 Jaringan saraf tiruan & pemrogramannya menggunakan MATLAB. *Penerbit ANDI*: Yogyakarta.
[10] Makridakis S, Wheelwright SC and McGee VE 1999 Metode dan Aplikasi Peramalan *Erlangga*: Jakarta.