Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2021, 2021:70
https://doi.org/10.28919/cmbn/5973
ISSN: 2052-2541

COMPARISON BETWEEN THE BACKPROPAGATION AND SINGLE EXPONENTIAL SMOOTHING METHOD IN SUGAR PRODUCTION FORECASTING CASE

EKA MALA SARI ROCHMAN*, WAHYUDI AGUSTIONO, NITA SURYANI, AERI RACHMAD

Department of Informatics Engineering, Faculty of Engineering, University of Trunojoyo Madura, Indonesia

Abstract: Sugar is a strategic commodity, because it is one of the basic needs consumed by the community. The decline in sugar production causes turmoil both economically and politically. So companies must be able to balance sugar production in accordance with market demand. Because, if the sugar production exceeds market demand, it will cause accumulation and expenditure costs that exceed the limit. Conversely, if the sugar production is too little, then the company is unable to meet demand. This study was comparing the Single Exponential Smoothing method and the Backpropagation method in predicting the amount of sugar production. The data used was based on production in 2013-2017, with predictions every month starting in the fifth month. The alpha parameter in Single Exponential Smoothing was 0.84, while in Backpropagation used 5 input nodes and 5 hidden layer nodes. From the results of tests that have been carried out, it can be seen that the Backpropagation method produces the smallest MSE and MAPE, amounting to 57187817.49 and MAPE of 1.012% compared to the Single Exponential Smoothing method with MSE of 83602989.43 and MAPE of 1.46%.

Keywords: sugar; predictions; simple exponential smoothing; backpropagation.

2010 AMS Subject Classification: 60G25.

---

*Corresponding author
E-mail address: ekamala.sari@yahoo.com
Received May 3, 2021

1. INTRODUCTION

Sugar is a simple carbohydrate that dissolves in water so that it is easily absorbed by the body and converts it into energy. Sugar cane is the basic ingredient for making sugar because of its low production costs [1][2]. In the Indonesian economy, sugar is a strategic commodity, because it is one of the basic needs consumed by the community. The decline in sugar industry production from year to year results in a gap between national sugar production and consumption [3]. These changes can invite various shocks in society both economically and politically, to reduce this it becomes the responsibility of the government.

The problem faced is the number of goods produced everyday, so companies need to consider financial planning to the fullest. Sugar production must be able to adjust to market demand, because if the amount of sugar production is too much it will cause accumulation; but if the amount of sugar production is too little it can cause the market demand is not fulfilled, causing the company to lose the opportunity for profit. Therefore, companies need to have a system that can predict the amount of goods to be produced in the next period.

Forecasting is an attempt to predict the situation in the future through testing the situation in the past [4][5][6]. Sales forecasting means determining the estimated volume of sales, even determining the potential for sales and the extent of the market being controlled in the future. Time series is a series of data collected in the order of time with the same interval. The time series method that is often used in forecasting is the Autoregressive Integrated Moving Average (ARIMA) method, the Exponential Smoothing and Backpropagation methods. Research on sugarcane production forecasting in Bangladesh has been carried out using the Auto-Regressive Integrated Moving Average (ARIMA) model which can be used to forecast sugarcane production in Bangladesh [2]. This study considers secondary data from the 1971 to 2013 period published from annual sugarcane production in Bangladesh. This study statistically says that this method is quite satisfying in making predictions compared to the AIC and BIC models. But this method in an addition to requiring large amounts of data, and the process also requires a lot of time and resources.

Exponential Smoothing is a technique that can provide accuracy in short-term forecasts and adjustments quickly and at a low cost. This method is widely used because it is simple and easy to be used with results that are not inferior when they are compared to the more complex forecasting models [7]. By using constants to assign weights to the current request and previous estimates arrive at a new estimate, it will affect the forecasting results [8].

The Single Exponential Smoothing method is a procedure that randomly continuously corrects predictions by averaging the past value of a time series data by decreasing (exponential). Research on bidding strategies for wind power producers using the Holt-Winters Exponential smoothing model is optimally designed [7]. This is done to calculate how likely the profit from the market price is based on the seasonal nature of the real time price. The results of this study were to find out the superiority of the proposed approach over a number of general benchmarks which were demonstrated through empirical investigations for Nord Pool over three years during 2008-2011.

In artificial intelligence, artificial neural networks are one method of determining hidden patterns [4] [9]. One method of this artificial neural network is Backpropagation, which is a learning algorithm to reduce the error rate by adjusting its weight based on the desired output and target differences. Stock price predictions use backpropagation with training data and testing data [9] can reduce the level of an error with their weights which have been adjusted based on differences in output and or expected targets.

The purpose of this study is to obtain optimal forecasting results in sugar production using the Backpropagation method and exponential smoothing which looks for the smallest MSE and MAPE values.

2. PRELIMINARIES

This chapter presents data collection, prediction understanding, Single Exponential Smoothing and Backpropagation.

2.1 Data Collection

The data used were data from the results of sugar production at Candi Baru sugar factory in 2013

- 2017. Data production was obtained from May to December each year, due to sugar production at Candi Baru sugar factory only starts in May.

2.2 Forecast

Forecasting is an activity of a business function that estimates the sale and use of the product so that the products can be made in the right quantity [10]. Forecasting is a prediction of an upcoming request based on several forecasters variables, so based on historical time series data. Forecasting uses forecasting techniques that are formal or informal [11].

There are two general approaches used in forecasting, namely qualitative forecasting and quantitative forecasting [12]. Quantitative methods can only be applied if information about past data is available, information can be quantized through numerical forms, and assumptions that apply to certain aspects of past patterns will continue. The types of quantitative forecasting are divided into two, namely time series and causal methods [13].

Time series forecasting is an estimate of the future made based on past values of variables and / or past errors. The causal method assumes that the predicted factor realizes a causal relationship with one or more independent variables. The aim is to find the shape of the relationship and use it to predict future values of the dependent variable [9].

2.3 Single Exponential Smoothing

Exponential Smoothing is a technique that can provide accuracy in short-term forecasts, and adjustments can be made quickly and at a low cost. This method is widely used because it is simple and easy to use with results that are not inferior when it is compared to the more complex forecasting models [14].

$$F_{(t+1)} = \alpha\ D_t + (1-\alpha)\ F_t \quad (1)$$

$F_{(t+1)}$ is the forecast price for the t + 1 period, while α is the smoothing coefficient The actual data for the t period is denoted by Dt and Ft for the forecast results in the t period.

2.4 Backpropagation

Backpropagation is a learning algorithm to reduce the error rate by adjusting the weight based on differences in output and desired targets.[13] Backpropagation is also a systematic method for training multilayer Artificial Neural Networks (ANN) [5][12]. Backpropagation is said to be a multilayer training algorithm because Backpropagation has three layers in its training process, namely the input layer, hidden layer and output layer, where backpropagation is a development of a single layer network that has two layers, namely the input layer and the output layer [16]. The hidden layer on backpropagation can cause the magnitude of the error rate on backpropagation to be smaller than the error rate on a single layer network. That is because the hidden layer on the backpropagation functions as a place to update and adjust the weights, so we get a new weight value that can be directed close to the desired output target.

2.5 Accuracy

2.5.1 Mean Square Error (MSE)

Mean Square Error (MSE) is another method for evaluating forecasting methods. MSE is calculated by summing the square of all forecast errors in each period and dividing it by the forecast period [14].

Consideration of the receipt of a forecasting method is through the following criteria [17]:

$$\text{MSE} = \sum \frac{(D_t - F_t)^2}{n} \tag{2}$$

MSE is the mean of square error, while the actual data is in period t. To show the forecast result for period t and the last one is variable n which shows the number of periods.

2.5.2 Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) is a way to measure an accuracy of a forecasting method using absolute errors in each period and the absolute percentage as for the formula as follows:

$$\text{MAPE} = \frac{1}{n}\sum_{n=1}^{n} \left| \left(\frac{D_{t-F_t}}{D_t}\right) \right| \text{ x } 100\% \tag{3}$$

Similar to calculating MSE on MAPE, there is also a variable that shows the actual data for each forecasting period and which is the forecast result for period t and n is the number of forecasting involved.

3. MAIN RESULTS

This chapter explains several test scenarios in making predictions using single exponential smoothing and backpropagation. The Exponential Smoothing Method produced the best α (alpha) value, the smallest MSE and forecast results. Whereas in the Backpropagation method by predicting the number of monthly production used an architecture where there was an input layer, hiden layer and output layer. Input nodes and hiden nodes of 5 made results in the forecasting and the smallest MSE value.

3.1 System Flow and Single Exponential Smoothing test scenario

This section discusses a description of the flow system for forecasting production using the Single Exponential Smoothing method shown in Figure 1. The figure explains the flow of the system being built. The process of prediction of the amount of sugar production using the single Exponential Smoothing method begins with inputting production data and the number of production sales, then determine the number of future periods to be predicted.

Then the stages of the Single Exponential Smoothing method are calculated. After the calculation process has reached a maximum iteration, the next process is to find the best α value and calculate the accuracy value to see how accurate the prediction results obtained with MSE and MAPE. The Scenario of Single Exponential Smoothing Trial and the result in this study as shown in Table 1 and Table 2.

Table 1. Scenario of Single Exponential Smoothing Trial

No	Change in Iteration
1	10
2	50
3	100
4	500
5	1000
6	10000

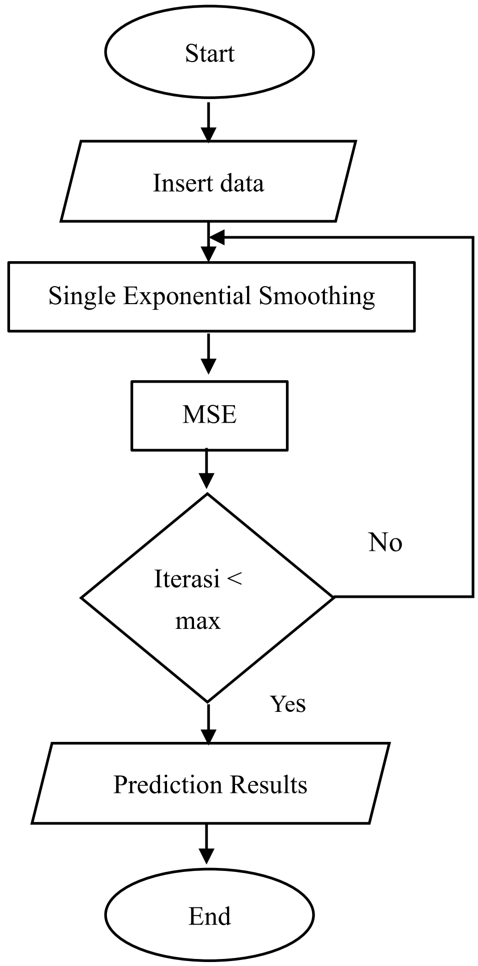

Figure 1. Single Exponential Smoothing Flowchart

Table 2. Single Exponential Smoothing Trial Results

No	Best Alpha	Iteration	MSE	Time (Seconds)
1	0,86	10	3.677.185,65	0,172
2	0,84	50	3.675.816,37	0,119
3	0,84	100	3.675.756,39	0,013
4	0,84	500	3.675.756,38	0,027
5	0,84	1000	3.675.756,38	0,032
6	0,84	10000	3.675.756,37	0,703

From the Single Exponential Smoothing test, the smallest MSE is 3,675,756.37 for 10, 50, 100, 500, 1000 and 10000 iterations. However, because the MSE value has the lowest number of iterations, the Max Iteration which has the biggest iteration is the iteration. 10000 The greater the iteration used, the results will remain.

3.2 System Flow and Backpropagation test scenario

The previous section discussed the flow of Single Exponential Smoothing, this time it discusses the picture of the system flow for forecasting production using the Backpropagation method shown in Figure 2. The prediction of the amount of sugar production using the single Backpropagation method began with inputting production data and the amount of production sales. The data was normalized to be inputted into an input node that can only accept values from 0 – 1. The number of input nodes used was as much as 5 things because sugar production has only begun since the fifth month. Similarly, the number of hidden layer nodes was as much as the number of inputs. After initializing the weights, the feedforward process was carried out, then afterward the backward process was carried out.

The counting process stopped when it reached its max error value or max epoch. If one of the conditions had been met, the system would store the best weight that has been obtained. But if the conditions did not meet then the process returns to the Feedforward stage until the conditions met and stopped. The best weight were processed in the testing process, after that it was denormalized until the end of the process that produced an error value with MSE and MAPE.

Table 3. Backpropagation Trial Scenarios

No	Changes to Parameter	Description
1	*Learning Rate (α)*	Changes in learning rate (alpha) consist of 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8,0,9
2	*Max Epoch*	Max epoch changes consist of 10, 50, 100, 500, 1000 and 10000
3	*Max Error*	Max error changes consist of 0,1, 0,01 and 0,001
4	*K-Fold Cross Validation*	2, 3, 4, and 5 fold

Test scenario used Backpropagation. How it worked by changing parameters and also changing the distribution of training data and testing data. In the implementation of this system trial, it consisted of several parameter changes including learning rate, max epoch, max error, training data sharing and data testing. The trial scenario was carried out by calculating the average MSE of 5 predetermined scenarios. The details of the parameter change scenario can be seen in Table 3.

3.2.1 Learning Rate

The first test is to change the value of the learning rate with a predetermined value of 0.1 - 0.9. This learning rate test aims to get the optimal amount of learning rate towards the MSE value from training and testing. The results of testing the learning rate parameters with max epoch 10, max error 0.1 and Fold 2 are shown in Table 4.

From testing on the value of learning rate that has been done, the optimum value is shown at the learning rate of 0.9 with an MSE value of 1.52 with an estimated time of 0.029. Then the learning rate of 0.9 is the best in this test. The results of the learning rate test in Table 4 obtained a value of 0.9 with an MSE value of 1.52 and the learning rate was made for further testing.

Table 4. Learning rate

No	Learning Rate	MSE	Time (seconds)	Iteration
1	0,1	1,65	0,007000	10
2	0,2	1,64	0,009000	10
3	0,3	1,66	0,008000	10
4	0,4	1,65	0,010001	10
5	0,5	1,66	0,008000	10
6	0,6	1,53	0,009001	10
7	0,7	1,55	0,007001	10
8	0,8	1,58	0,007000	10
9	0,9	1,52	0,006000	10

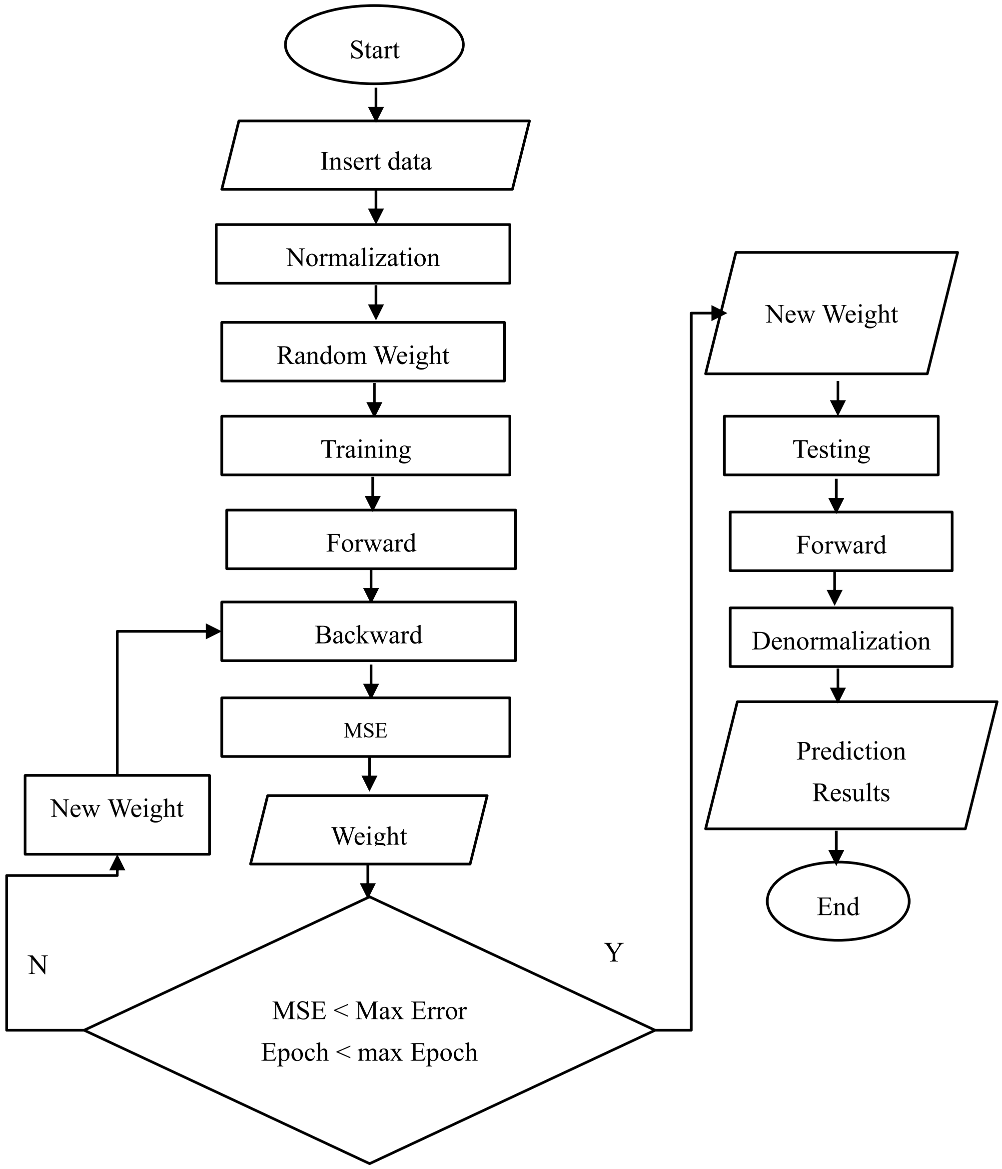

Figure 2. Backpropagation flowchart

3.2.2 Max Epoch

Testing the number of max epoch is aimed to determine whether before max epoch the MSE value has exceeded the limit or stopped at the max epoch that has been determined. The number of max epoch used was 10, 50, 100, 500, 1000 and 10000. The results of the learning rate test in Table 4 obtained a value of 0.9 with an MSE value of 1.5239 and the learning rate was used for further testing. The results of testing the number of mac epoch with a learning rate of 0.9, max error of 0.1 and 2 Fold are shown in Table 5 below.

Table 5. Max Epoch

No	Max Epoch	MSE	Time (seconds)	Iteration
1	10	1,10	0,008001	10
2	50	1,56	0,010000	50
3	100	1,57	0,007001	100
4	500	1,44	0,008001	500
5	1000	1,26	0,008001	1000
6	10000	0,98	0,008000	10000

From the results of trials, it has been carried out, seen, and concluded that the best accuracy was obtained. Testing the number of max epoch was aimed to determine whether before max epoch the MSE value has exceeded the limit or stopped at the specified max epoch. The number of max epochs used is 10, 50, 100, 500, 1000 and 10000. The max epoch test results shown in Table 5 are the optimal number of max epochs with the smallest MSE value at the time of the max epoch 50 number because the MSE value is 0.98 with time an estimate of 0.008. So that subsequent testing uses a learning rate of 0.9 and max epoch 10000 which has been tested previously.

3.2.3 Max Error

Testing the max error was aimed to limit the error value in order to get the optimal error value. Max errors used were 0.1, 0.01 and 0.001. The results of testing the number of max errors with a learning rate of 0.9, epoch 1000 and 2 fold are shown in Table 6 below.

Table 6. Max Error

No	Max Error	MSE	Time (seconds)	Iteration
1	0,1	1,68	0,008001	10000
2	0,01	1,56	0,009000	10000
3	0,001	1,47	0,007001	10000

The max error test results shown in Table 6 are the optimal number of max errors with the smallest MSE value when the max error number is 0.001 because the MSE value is 1.4728 with an estimated time of 0.007001. So that, subsequent testing uses a learning rate of 0.9, max epoch 10000 and max error of 0.001 which has been tested previously.

3.2.4 K-fold

K-fold is one of the popular Cross Validation methods by folding K as much data and repeating (iterating) the experiment as much as K as well where data is separated into two subsets namely learning process data and testing data. In this study, a 5 fold trial scenario is shown in Table 7 – 10.

Table 7. Distribution of Training Data and Data Testing using 2 Fold

No	Train	Test	MSE	Time (seconds)	Iteration
1	K2	K1	1,61	0,007000	10000
2	K1	K2	3,33	0,007001	10000
Average MSE			2,4733		

Table 8. Distribution of Training Data and Data Testing using 3 Fold

No	Train	Test	MSE	Time (seconds)	Iteration
1	K2, K3	K1	1,14	0,008000	10000
2	K1, K3	K2	1,79	0,007001	10000
3	K1, K2	K3	2,26	0,008000	10000
Average MSE			1,7312		

Table 9. Distribution of Training Data and Data Testing using 4 Fold

No	Train	Test	MSE	Time (seconds)	Iteration
1	K2, K3, K4	K1	0,66	0,007000	10000
2	K1, K3, K4	K2	0,76	0,009000	10000
3	K1, K2, K4	K3	1,24	0,008001	10000
4	K1, K2, K3	K4	1,79	0,008000	10000
Average MSE			1,112825		

Table 10. Distribution of Training Data and Data Testing using 5 Fold

No	Train	Test	MSE	Time (seconds)	Iteration
1	K2, K3, K4, K5	K1	0,77	0,009001	10000
2	K1, K3, K4, K5	K2	0,66	0,011000	10000
3	K1, K2, K4, K5	K3	0,78	0,008001	10000
4	K1, K2, K3, K5	K4	0,86	0,007001	10000
5	K1, K2, K3, K4	K5	1,61	0,008000	10000
Average MSE			0,93646		

From all the scenarios that have been tested, the following is the smallest MSE results Table 11 from the parameter arrangement of each scenario:

Table 11. Overall MSE results

No	Number of K-Fold	Average MSE	Stop Iteration
1	2	2,47	10000
2	3	1,73	10000
3	4	1,11	10000
4	5	0.94	10000

From the Table 11 above, it shows the smallest average MSE value of 0.93646. Then, the results of the trial scenario 5 fold is the best division of training data and testing data.

The results of testing the number of max epochs with a learning rate of 0.9, max error of 0.1 and 2 fold are the following from the smallest MSE value. Furthermore, to compare forecasting using the Single Exponential Smoothing and Backpropagation method, the following parameters were used as shown in the Table 12.

Table 12. Parameter of the Bacpropagation Trial results

Parameter	Result
Learning rate	0,9
Max Epoch	10000
Max Error	0,001
Jumlah K-Fold	5

3.3 Analysis

From the results of a trial scenario that had been done, then forecasting the system used parameters such as in Table 13 for the Backpropagation method.

Table 13. Comparison Results of the Single Exponential Smoothing Method, Backpropagation Method and Actual Data

No	Month	SES Forecasting Results (tons)	Backpropagation Forecasting Results (tons)	Actual Data (tons)
1	Okt-13	3.779,37	4.876,42	4.656,9
2	Nov-13	4.153,95	6.173,44	6.526,5
3	Des-13	2.715,61	5.158,77	5.082,7
4	Mei-14	5.735,5	6.175,72	6.148,8
5	Jun-14	4.240,7	5.552,06	3.785,6
6	Jul-14	20.865,28	4.573,85	4.239,7
7	Agu-14	3.805,96	4.264,01	3.834,9
8	Sep-14	4.233,90	4.546,02	3.733,1
9	Okt-14	2.508,21	5.220,97	2.739,8
10	Nov-14	3.689,8	5.778,01	5.735,5
11	Mei-15	880	5.065,49	4.552,9
12	Jun-15	239	5.694,23	5.322
13	Jul-15	4.228,26	5.510,01	2.358
14	Agu-15	7.510,47	5.391,05	3.689,8
15	Sep-15	3.915,39	4.862,88	880
16	Okt-15	4.447,8	4.078,45	239
17	Nov-15	11.020,81	5.011,06	1.218,2
18	Jun-16	4.980,47	5.572,24	5.656
19	Jul-16	5.750,13	4.565,94	4.447,8
20	Agu-16	5.322	4.464,46	2.067,5
21	Sep-16	40.677,34	6.094,18	19.688,6
22	Okt-16	5.071,45	5.397,08	5.436,6
23	Nov-16	6.052,49	5.699,71	5.716,5
24	Des-16	5.783,07	6.136,54	6.651,7
25	Mei-17	4.637,96	6.082,17	44.240
26	Jun-17	5.554,3	4.211,70	4.047,1
27	Jul-17	3.785,6	6.166,84	6.216,3
28	Agu-17	4.447,8	5.946,28	5.753,3
29	Sep-17	5.035,69	4.880,03	4.633,7
30	Okt-17	4.239,7	5.454,75	5.554,3

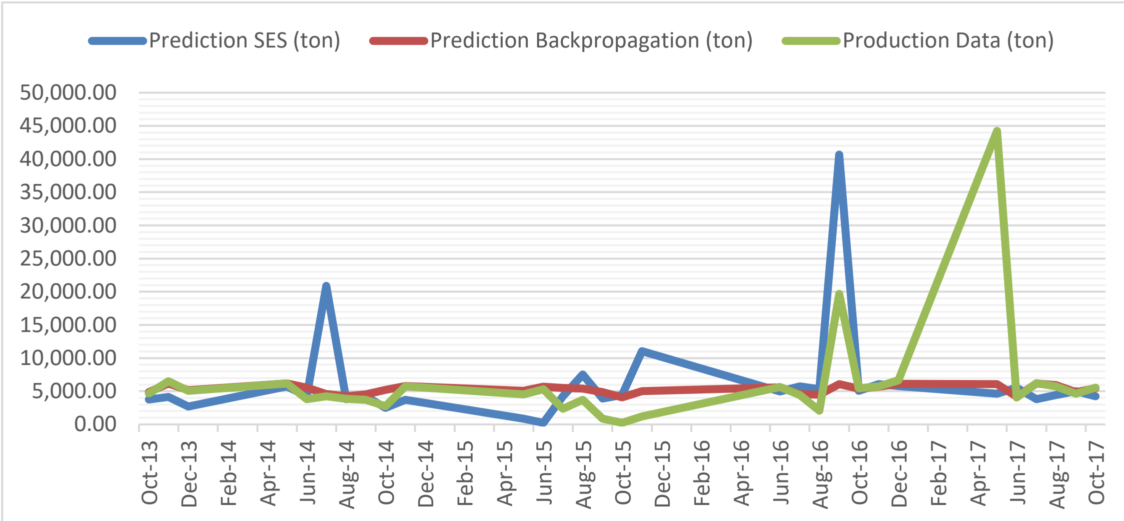

Figure 3. Graphic results of comparison of forecasting results using the Single Exponential Smoothing method and the Backpropagation method with Actual Data

Based on Table 13 regarding the results of forecasting using the Single Exponential Smoothing method can change according to the number of iterations used. When the more iterations are used, then the forecasting results would be the same. In contrast to the results of forecasting using the Backpropagation method, the results are not always the same even with the same amount of training data, the amount of testing data, learning rate, max epoch and max error, because every forecasting process is carried out initial weights randomly with a range of 0 - 1. So that the result in the smallest MSE value can be obtained in different parameters.

4. Conclusion

Based on the results of this study, it can be concluded that the forecasting of sugar production using the Single Exponential Smoothing method has an MSE value of 83602989.43 and a MAPE of 1.46%. While the Backpropagation method has an MSE value of 57187817.49 and MAPE of 1.02%. So, the most appropriate method for forecasting sugar production is the Backpropagation method because it has the smallest MSE and MAPE values.

Acknowledgments

The author would like to thank the faculty of engineering at the University of Trunojoyo Madura for providing opportunities and providing facilities to be able to conduct research.

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

REFERENCES

[1] J. Megha, Y.N. Havaldar, N.L. Pavithra, B.B. Jyoti, V.K. Kumar, Identification of the best model for forecasting of sugar production among linear and non-linear model, Int. J. Curr. Microbiol. Appl. Sci. 8(3) (2019), 2556–2560.

[2] M.M. Hossain, F. Abdulla, Forecasting the sugarcane production in Bangladesh by ARIMA model, J. Stat. Appl. Probab. 4(2) (2015), 297–303.

[3] N. Hernanda, Analisis Peramalan Tingkat Produksi Dan Konsumsi Gula Indonesia Dalam Mencapai Swasembada Gula Nasional, Institut Pertanian Bogor, 2011.

[4] F. Wong, Time series forecasting using backpropagation neural networks, Neurocomputing, 2 (1990), 147–159.

[5] P.T.J. Singh, Time series forecasting using back propagation neural network with ADE algorithm, Int. J. Eng. Techn. Res. 7(5) (2017), 19–23.

[6] C.W. Mohd Noor, R. Mamat, G. Najafi, W.B. Wan Nik, M. Fadhil, Application of artificial neural network for prediction of marine diesel engine performance, IOP Conf. Ser.: Mater. Sci. Eng. 100 (2015), 012023.

[7] T. Jónsson, P. Pinson, H. A. Nielsen, and H. Madsen, Exponential smoothing approaches for prediction in real-time electricity markets, Energies, 7(6) (2014), 3710–3732.

[8] H. Ravinder, Forecasting with exponential smoothing what's the right smoothing constant? Rev. Bus. Inform. Syst. 17 (2013), 117-126.

[9] Z.H. Khan, T.S. Alin, M.A. Hussain, Price prediction of share market using artificial neural network (ANN), Int. J. Computer Appl. 22 (2011), 42-47.

[10] A. Rachmad, D.R. Anamisa, Forecasting the number of patients diseases using backpropagation, MATEC Web Conf. 58 (2016), 03005.

[11] O. Obe, K. Shangodoyin, Artificial neural network based model for forecasting sugar cane production, J. Computer Sci. 6(4) (2010), 439–445.

[12] E.M. Sari Rochman, A. Rachmad, M.A. Syakur, I.O. Suzanti, Method Extreme Learning Machine for

Forecasting Number of Patients' Visits in Dental Poli (A Case Study: Community Health Centers Kamal Madura Indonesia), J. Phys.: Conf. Ser. 953 (2018), 012133.

[13] E.M.S. Rochman, A. Rachmad, Forecasting Application for Simpati Telkomsel Card Using Backpropagation (Case Study in Bangkalan Madura-Indonesia), Adv. Sci. Lett. 23 (2017), 12340-12343.

[14] D.N. Hamdany, Penerapan Metode Single Exponential Smoothing and Double Exponential Smoothing untuk Peramalan Penjualan Export dan Import, Skripsi,Universitas Trunojoyo Madura, 2017.

[15] R. Suryanita, H. Maizir, Ismeddiyanto, V. Trisatria, R. Arditama, Damage level prediction of multi-story steel structure in Sumatra using backpropagation neural network, Int. J. Geomate, 17(60) (2019), 37-42.

[16] E.M. Rohman, Imamah, A. Rachmad, Predicting Medicine-Stocks by Using Multilayer Perceptron Backpropagation, Int. J. Sci. Eng. Appl. 5(3) (2016), 188-191.

[17] M. Arifin, Peramalan Penjualan Produk Berbasis Web Menggunakan Metode Single Exponential Smoothing (SES) (Studi Kasus: Penjualan Sepeda Motor Di Deler Mitra Setia Agung Sumenep), Skripsi, Universitas Trunojoyo Madura, 2012.