

Forecasting sugarcane production in the Asembagus sugar factory

by I Harlianingtyas

Submission date: 25-Nov-2019 10:29AM (UTC+0700)

Submission ID: 1220975662

File name: irma_-_forecasting.docx (77.6K)

Word count: 1736

Character count: 8823

Forecasting sugarcane production in the Asembagus sugar factory

I Harlianingtyas¹, A Salim^{1*}, D Hartatie¹, and Supriyadi¹

¹Department of Agriculture Production, Politeknik Negeri Jember, Jalan Mastrip Pobox 164, Jember 68121, Indonesia

* abdurrahman.salim@polije.ac.id

Abstract. The production of the sugar industry is increasing from year to year because of the national sugar production and consumption. To achieve the level of production that can meet these needs, the government has requested a national sugar self-sufficiency policy. Estimation of sugarcane production has a big influence on the company's policy considerations to determine further production results. This study aims to predict the production of sugar cane that will be obtained by the Asembagus sugar factory for the next five years. The data used to predict are data from 1979 to 2018 obtained from secondary data from Asembagus factory registration results. The method used is the Box-Jenkins ARIMA method by comparing the measurement of forecasting results with the trend and exponential smoothing methods. The appropriate forecasting result is the ARIMA model (1,1,1) because it has a smaller error value than using other methods. Forecasting results for 2019 are 42055.6 tons in 2020 predicted 43222.9 tons in 2021 predicted to produce 44090.7 tons, in 2022 predicted 44788.2 tons and in 2023 predicted sugarcane production of 45389.2 tons.

1. Introduction

Forecasting is an important aid in effective and efficient planning [1]. Forecasting is an activity to predict future events by using and considering data from the past. Many methods in statistics can be constructed to forecast a time series data, such as smoothing, Box-Jenkins, econometrics, regression, transfer functions and so on. These methods are expected to identify the data used to predict conditions in the future so that the errors are as minimal as possible.

Along with the development of increasingly advanced technology that time-series data forecasting methods have been developed much like the ARIMA method. ARIMA is a method commonly used to predict data [2]. The ARIMA method makes full use of past and present data for forecasting [3]. Autoregressive Integrated Moving Average (ARIMA) is a forecasting model that produces predictions based on the synthesis of historical data patterns. The ARIMA method will work perfectly if the data in the time series used are dependent or statistically related to each other [1].

The ARIMA method has been used by comparing forecasting with the Holt-Winters method in predicting pentad OLR anomalies in western Indonesia, with better ARIMA forecasting results [2]. The advantage of the ARIMA method is that it is suitable for simple data forecasting and the application of a relatively easy method of analyzing data containing seasonal or trendy patterns, overcoming the problem of randomness and even the cyclical nature of the time series data being analyzed.

To find out the magnitude of the accuracy of the forecast generated, the authors analyze the comparison of forecasting using the Holt-Winters exponential smoothing method and ARIMA by calculating forecast errors such as Mean Squared Error (MSE), Mean Absolute Deviation (MAD), and Mean Absolute Percentage Error (MAPE) so that the error is as minimal as possible.

2. Method

To predict sugarcane production in the Asembagus sugar factory using time series analysis, because sugarcane production data is time-series data. The data used is sugarcane production data per year from 1979 to 2018. The time series analysis method is used to predict production data for the next 5 years 2019 to 2023. The first step of the analysis is plotting time-series data, and then an upward trend is diagnosed because of the data obtained from year to year experience a gradual increase. Then the trend analysis and exponential smoothing methods were tried out, and then analyzed with ARIMA (p, 1, q) with data so that the equation model for ARIMA was obtained as follows $W_t = X_t - X_{t-1}$ so that the equation model for ARIMA is obtained as follows [4]:

$$W_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (1)$$

In the process of determining the ARIMA model, it is often found that the data is not stationary in the variant or the average, so it needs to be transformed until the rounded value is 1 for the data to be considered stationary in the variant and differencing so that the data is stationary in the average. The next stage is the calculation of Mean Squared Error (MSE), Mean Absolute Deviation (MAD), and Mean Absolute Percentage Error (MAPE) to get the smallest possible error [1].

To calculate MSE using the equation as below :

$$MSE = \frac{1}{n} \sum_{i=1}^n |X_i - F_i|^2 \quad (2)$$

Where for F_i is forecast and X_i is actual data. The equation to calculate MAPE is as follows :

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{X_i - F_i}{X_i} \right| \quad (3)$$

And to calculate MAD using equation follows as below :

$$MAD = \frac{\sum_{i=1}^n |X_i - F_i|}{n} \quad (4)$$

3. Result and discussion

Result of the time-series plot production sugarcane from 1979 to 2018 is illustrated in fig. 1 as follows.

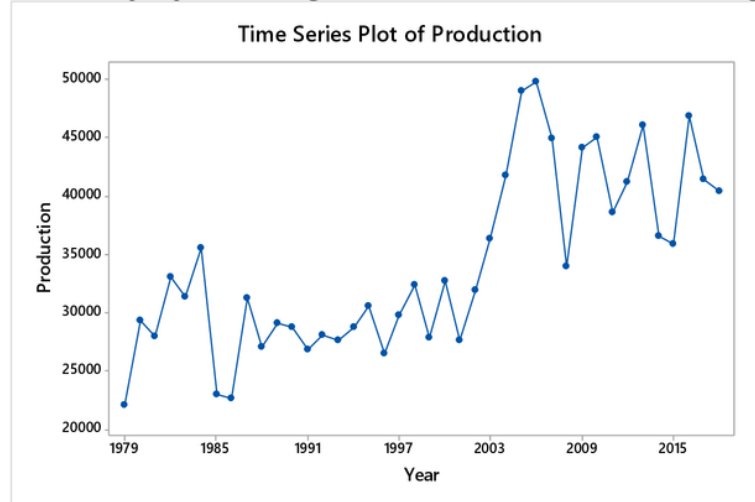


Figure 1. The production time series plot

Based on figure 1, The lines appears that there is an upward trend from sugarcane production data from 1979 to 2018, so forecasting is done using trend analysis, exponential smoothing, and approaching ARIMA Box-Jenkins.

3.1 Fitting model time series

Forecasting that is adjusted to the periodic data plot using the trend analysis forecasting method is based on the measurement of accuracy the forecast method using the value of MAPE, MAD, MSD [5]. Then the measurement of the accuracy of the data is matched with the time series data and shown as a percentage. Each of these forecasting methods has different MAPE, MAD, and MSD values. Forecasting methods that have the smallest MAPE, MAD, and MSD values are effective forecasting methods and have a high degree of accuracy [6].The following are accuracy measurements for trend analysis forecasting and exponential smoothing methods.

Table 1. Measure of Forecasting Result

Measure of Forecasting Results	Forecasting Method			
	Trend Linear	Trend Quadratic	Trend Exponential	Exponential Smoothing
Mean Absolute Percentage Error (MAPE)	13	12	12	12
Mean Absolute Deviation (MAD)	436	402	419	396
Mean Squared Deviation (MSD)	265589	252276	261075	242041

Based on the method of forecasting using trend analysis and Exponential Smoothing, the known values of MAPE, MAD, and MSD of the four forecasting methods that have the lowest values are the exponential smoothing method [5].

3.2 Fitting ARIMA model

The Box-Cox plot used to find out the data is stationary in the variant data, especially for the rounded value. It calls stationary if rounded value approaching 1. The following result of the Box-Cox Plot is a figure in fig.2.

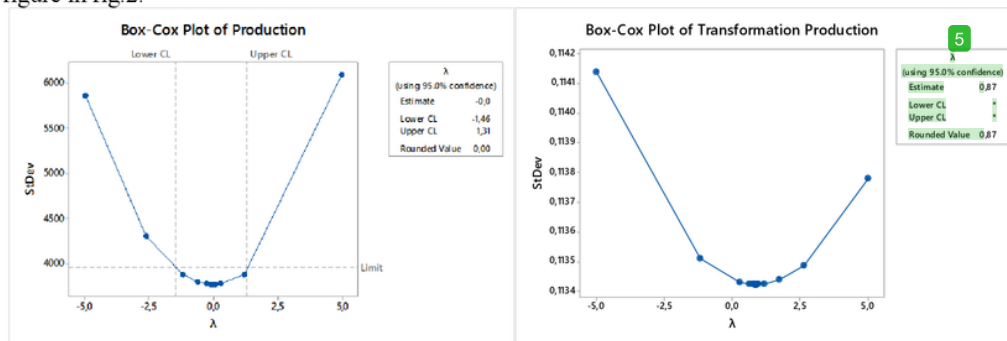


Figure 2. Box-Cox Plot for Production Data (a) and Transformation of Production Data (b)

It can be seen that the rounded value of 0 is said to be stationary in variance if the magnitude of rounded value is 1, so it needs to be transformed until the size of the rounded value becomes 1. Then the data used to estimate the forecasting model is the data after the transformation. The ARIMA model

according to the Box-Jenkins theory, plotting ACF and PACF need to be done first so that it is known whether the data is stationary at variance or average [1].

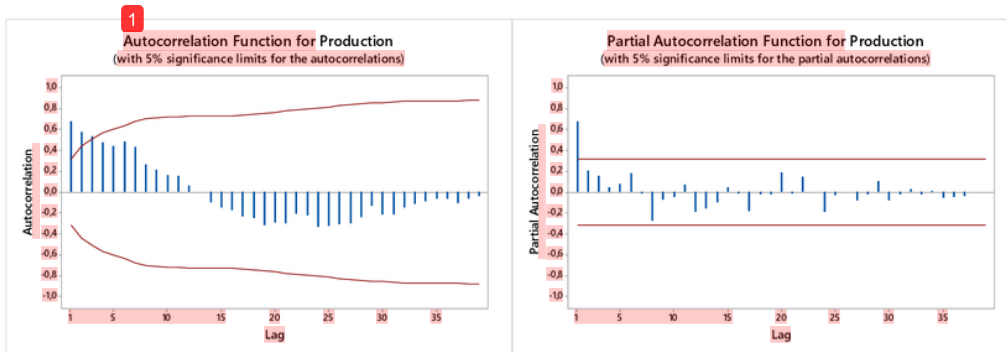


Figure 3. Plot ACF (a) and Plot PACF (b)

Based on the ACF and PACF plots it can be seen that the two graphs are cut off in each lag, so it is assumed that the ARIMA model with the order $p = 1$ and $q = 1$. It is seen that the ACF graph cut off in lag 3 and the PACF graph is cut off in lag 1, so it can be assumed that the data is not stationary in the average. Therefore it is necessary to do differencing on the data that has been transformed earlier. After the data differs, it fulfills stationary assumptions both at variance and average so that it can be estimated by the ARIMA model [1].

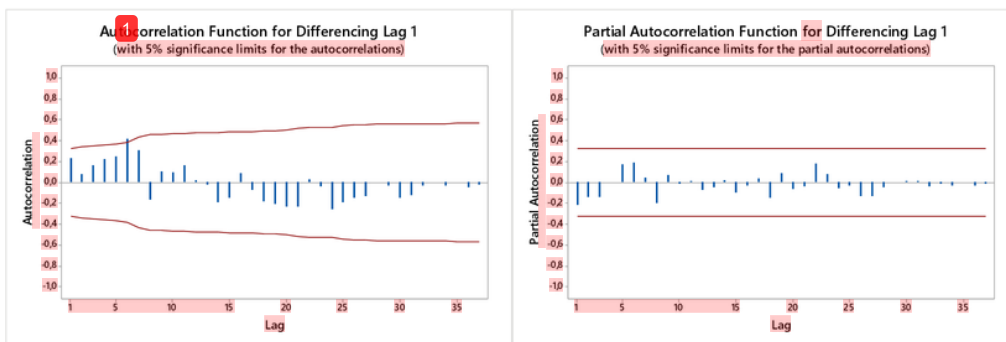


Figure 4. Plot differencing ACF (a) and PACF (b)

After estimating using ARIMA (1,1,0), ARIMA (0,1,1) and ARIMA (1,1,1) we conclude that a significant ARIMA model with a p -value < 0.05 is ARIMA (1,1,1). The selected ARIMA models provide an adequate predictive model for sugar situation. The models have been followed to forecast the production, consumption and gap of sugar from 2019 to 2023. The best fitted ARIMA model is ARIMA (1,1,1) forecast sugarcane production in the next 5 years as shown in table 2. Forecasting result shown that the production of sugarcane increasing every year. ARIMA model projected that production of sugarcane will increase from 4036,09 tons in 2018 to 45389,2 tons in 2023, with an increase ratio of about 12,46% more than its value in 2018 [7].

The best fitted forecasting model can also be seen from the smallest error measurement [6]. Based on the measurement of errors the ARIMA model result that the value of MAPE is 5,386; MSD is 226072,3; and MAD is 255, 6217.

Table 2. Forecasting of sugarcane production

Year	Production (tons)
2019	42055.6
2020	43222.9
2021	44090.7
2022	44788.2
2023	45389.2

The results of forecasting sugarcane production from 2019 to 2023 have increased, this needs to be a concern of the factor management to prepare policies that will be determined in the face of future sugarcane production. These forecast values could be used for formulating food policies especially for sugar [7].

4. Conclusion

The appropriate forecasting result is the ARIMA model (1,1,1) because it has a smaller error value than using other methods with the value of MAPE is <10. The forecasting model is feasible to use. Forecasting results for 2019 are 42055.6 tons in 2020 predicted 43222.9 tons in 2021 predicted to produce 44090.7 tons, in 2022 predicted 44788.2 tons and in 2023 predicted to obtain sugarcane production of 45389.2 tons.

5. Acknowledgments

This research was supported by PNBPN State Polytechnic of Jember 2019.

6. References

- [1] Makridakis S Wheelwright S C McGee V and McGee E 1999 *Forecasting Methods and Applications* (Jakarta: Erlangga)
- [2] Hermawan E 2011 Comparison of the Box-Jenkins and Holt-Winters Method in Predicting the Pentad OLR Water Anomaly in Western Indonesia *Sains Dirgantara Jurnal* 9 1 25-35.
- [3] Anggriningrum D P Hendikawati P and Abidin Z 2013 Comparison of Share Price Predictions by Using Backpropagation and ARIMA Artificial Neural Networks *Unnes Journal of Mathematics* 2 2 105-109
- [4] Box G E P and Jenkins M 1970 *Time Series Analysis* (California: Holden-Day)
- [5] Purfadila S and Andriani D R 2018 *Agricultural Economics and Agribusiness Journal* 2 1 52-61
- [6] Sungkawa I and Megasari R T 2011 *Computer Technology Journal* 2 2 636-645
- [7] Esam A B 2017 *International Journal of Agricultural Economics* 2 4 96-109

Forecasting sugarcane production in the Asembagus sugar factory

ORIGINALITY REPORT

21%

SIMILARITY INDEX

16%

INTERNET SOURCES

10%

PUBLICATIONS

17%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to Nottingham Trent University

Student Paper

10%

2

article.agricecon.org

Internet Source

3%

3

Submitted to Universitas Pendidikan Indonesia

Student Paper

2%

4

etd.lib.nsysu.edu.tw

Internet Source

2%

5

Wiwik Anggraeni, Abdolatul Abdillah, Pujiadi, Lulus Tjondro Trikoratno et al. "Modelling and Forecasting the Dengue Hemorrhagic Fever Cases Number Using Hybrid Fuzzy-ARIMA", 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH), 2019

Publication

1%

6

Submitted to Pascasarjana Universitas Negeri Malang

Student Paper

1%

7

Submitted to University of Warwick

Student Paper

1%

8

Baabak Ashuri, Jian Lu. "Time Series Analysis of ENR Construction Cost Index", Journal of Construction Engineering and Management, 2010

Publication

1%

9

epdf.tips

Internet Source

1%

Exclude quotes On

Exclude matches < 1%

Exclude bibliography On