Pak.j.stat.oper.res. Vol.18 No. 2 2022 pp 411-416 DOI: http://dx.doi.org/10.18187/pjsor.v18i2.3533

Pakistan Journal of Statistics and Operation Research

Decomposition Method with Application of Grey Model GM(1,1) for Forecasting Seasonal Time Series

Nur Hikmah[1], Mujiati Dwi Kartikasari[2*]

* Corresponding Author

1. Department of Statistics, Universitas Islam Indonesia, Indonesia, hikmahnh89@gmail.com
2. Department of Statistics, Universitas Islam Indonesia, Indonesia, mujiatikartikasari@uii.ac.id

Abstract

Forecasting is one of the activities companies need to determine the policies that need to be taken for the continuity of operations. There are many methods for forecasting, one of which is the grey model GM(1,1). The GM(1,1) is one of the successful forecasting methods applied to economics, finance, engineering, and others. However, according to several previous studies, the GM(1,1) is not good enough to forecast data containing seasonal characteristics. Therefore, this study aims to develop a hybrid model so that the GM(1,1) can forecast seasonal time series. The hybrid model combines the decomposition method for seasonality adjustment and the grey model GM(1,1) for forecasting seasonal time series. The results are compared to the seasonal grey model SGM(1,1). Based on the evaluation using error criteria, it is found that the hybrid model is the best.

Key Words: Forecasting, Time Series, Seasonal, Decomposition, Grey Model

Mathematical Subject Classification: 62P01, 62P02

1. Introduction

Forecasting is a process for predicting future values or characteristics (Makridakis, Wheelwright, and Hyndman, 2008). Forecasting is one of the crucial activities for a company because forecasting can help determine the policies to sustain its operational activities. In line with the company's management efforts to reduce dependence on uncertain matters and increase the complexity, competition and level of environmental change, the need for forecasting is also increasing.

Companies sometimes face conditions where the available data is insufficient, so they cannot use existing forecasting methods. Therefore, the grey model GM(1,1) can be a solution for companies to predict future conditions if data availability is incomplete or limited. The GM(1,1) is a model that adopts an essential section of grey system theory, and it has been widely used for forecasting in engineering, economics, finance, and others. Grey system theory was first developed by Julong (1989). This theory is beneficial for analyzing a system with limited data availability and incomplete information, and a short period (Liu and Lin, 2006).

Grey model GM(1,1) is indeed compatible for making predictions if the historical data available is limited. However, GM (1,1) is not good enough to predict seasonal time series data (Tseng, Yu, and Tzeng, 2001; Xia and Wong, 2014). In the real world, many time series data contain seasonality. We construct a hybrid model by combining the decomposition method and grey model GM(1,1) to predict it. The decomposition method decomposes time series data into seasonal components and seasonal adjustment components, consisting of trend-cycle and remainder components. We separately forecast the seasonal component using the average method and the seasonally adjusted component using GM(1,1). The results are combined at the end of the analysis. The hybrid model's analysis results are then compared with the seasonal grey model SGM(1,1). The best model is based on three criteria of error, which are mean absolute percentage error (MAPE), mean absolute error (MAE), and mean squares error (MSE).

2. Decomposition Method

The decomposition method decomposes a time series into trend-cycle, seasonal, and remainder components. There are two kinds of decomposition, which are additive and multiplicative. Additive decomposition is appropriate if the magnitude of seasonal fluctuations does not vary with the level of data. While, if seasonal is proportional to the level of data, then multiplicative decomposition is appropriate. According to Hyndman and Athanasopoulos (2018), the procedures used in additive and multiplicative decomposition are here.

The formulation of additive decomposition is

$$x_t = S_t + T_t + R_t, \tag{1}$$

where x_t is the data, S_t is the seasonal component, T_t is the trend-cycle component, and R_t is the remainder component, all being in the t period. The formulation of multiplicative decomposition is

$$x_t = S_t \times T_t \times R_t. \tag{2}$$

The following are the steps used to perform additive decomposition:

Step 1: Estimate the trend-cycle component $\hat{T}_t$ using the $\mathrm{MA}(2 \times m)$ if m is an even number or using the $\mathrm{MA}(m)$ if m is an odd number. The $\mathrm{MA}(2 \times m)$ is 2nd order moving average to the moving average of the order m and the $\mathrm{MA}(m)$ is moving average of order m.

Step 2: Calculate the detrended of data series using $x_t - \hat{T}_t$.

Step 3: Estimate the seasonal component $\hat{S}_t$ by calculating the average of the detrended value for each season. These seasonal component values are then adjusted to ensure that they add to zero. The seasonal component is obtained by stringing together these monthly values, and then replicating the sequence for each year of data.

Step 4: Calculate the remainder component $\hat{R}_t$ by subtracting the estimated seasonal and trend-cycle components using $\hat{R}_t = x_t - \hat{T}_t - \hat{S}_t$.

Meanwhile, the following are the steps used to perform multiplicative decomposition:

Step 1: Estimate the trend-cycle component $\hat{T}_t$ using the $\mathrm{MA}(2 \times m)$ if m is an even number or using the $\mathrm{MA}(m)$ if m is an odd number.

Step 2: Calculate the detrended of data series using $x_t / \hat{T}_t$.

Step 3: Estimate the seasonal component $\hat{S}_t$ by calculating the average of the detrended value for each season. These seasonal indexes are then adjusted to ensure that they add to m. The seasonal component is obtained by stringing together these monthly indexes, and then replicating the sequence for each year of data.

Step 4: Calculate the remainder component $\hat{R}_t$ by dividing out the estimated seasonal and trend-cycle components using $\hat{R}_t = x_t / (\hat{T}_t \hat{S}_t)$.

3. Grey Model GM(1,1)

Grey model GM(1,1) is a very popular, known as the model with single variable first-order. In accordance with the grey system theory, the procedure for forecasting using GM(1,1) according to Dang et al. (2016) is as follows:

Step 1: Arrange the initial order

$$x^{(0)}(i) = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n-m), \cdots, x^{(0)}(n)\} \tag{3}$$

where $x^{(0)}(i)$ is the time series data at period i.

Step 2: Arrange a new order $x^{(1)}$ by the accumulated generating operation (AGO),

$$x^{(1)}(i) = \{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n-m), \cdots, x^{(1)}(n)\} \tag{4}$$

where

$$x^{(1)}(i) = \sum_{j=1}^{i} x^{(0)}(j). \tag{5}$$

Step 3: The differential equation of single variable first-order is stated as follows:

$$\frac{dx^{(1)}}{dt} + u x^{(1)} = v \tag{6}$$

where u is a development coefficient and v is a grey input, which can be determined by the ordinary least squared method as follows:

$$[u, v]^T = (P^T P)^{-1} P^T Q \tag{7}$$

where

$$P = \begin{bmatrix} -(x^{(1)}(1) + x^{(1)}(2))/2 & 1 \\ -(x^{(1)}(2) + x^{(1)}(3))/2 & 1 \\ \vdots & \vdots \\ -(x^{(1)}(n-1) + x^{(1)}(n))/2 & 1 \end{bmatrix} \tag{8}$$

and

$$Q = \begin{bmatrix} x^{(0)}(1) \\ x^{(0)}(2) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}. \tag{9}$$

By completing (7), calculate the response equation for GM(1,1) by using equation as follows:

$$\hat{x}^{(1)}(i) = \left[x^{(0)}(1) - \frac{v}{u}\right] e^{-u(i-1)} + \frac{v}{u} \quad (i = 2,3,\cdots,n). \tag{10}$$

Step 4: Calculate predicted future values by the first-order inverse accumulating generation operation (1-IAGO),

$$\hat{x}^{(0)}(i) = \hat{x}^{(1)}(i) - \hat{x}^{(1)}(i-1) \quad (i = 2,3,\cdots,n) \tag{11}$$

where $\hat{x}^{(0)}(1) = \hat{x}^{(1)}(1)$.

4. Seasonal Grey Model SGM(1,1)

Many time series data contain seasonal characteristics. Forecasting done by the GM(1,1) with AGO cannot solve seasonal problems and will always produce a series that increases or decreases exponentially. There is a method known as seasonal grey model SGM(1,1) to solve the problem. On SGM(1,1), the AGO is increased to a cycle truncation accumulated generating operation (CTAGO) (Xia and Wong, 2014). According to Ozcan (2017), here is the SGM(1,1) procedure:

Step 1: Arrange the initial order $x^{(0)} = \left\{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(i), \cdots, x^{(0)}(n)\right\}$. Then, arrange a new order $x^{(1)}$ by the cycle truncation accumulated generating operation (CTAGO),

$$x^{(1)}(k) = \mathrm{CTAGO}\left(x^{(0)}(k)\right) = \sum_{j=1}^{q} x^{(0)}(k+j-1) \tag{12}$$

where $k = 1,2,\cdots,n-q+1$ and q is the seasonal period.

Step 2: Specify the seasonal grey forecasting model as follows:

$$\hat{x}^{(1)}(k+1) = d_1\left(x^{(1)}(k) + \lambda\right) + d_2, \quad \forall k = 1,2,\cdots,n-q. \tag{13}$$

The values of d_1, d_2 and λ can be determined using the least square method, which is

$$d = [d_1, d_2]^T = (A^T A)^{-1} A^T Q, \tag{14}$$

where

$$A = \begin{bmatrix} x^{(1)}(1) & 1 \\ x^{(1)}(2) & 1 \\ x^{(1)}(3) & 1 \\ \vdots & \vdots \\ x^{(1)}(n-q) & 1 \end{bmatrix}, \tag{15}$$

$$Q = \begin{bmatrix} x^{(1)}(2) \\ x^{(1)}(3) \\ \vdots \\ x^{(1)}(n-q+1) \end{bmatrix}, \tag{16}$$

$$\lambda = \frac{\sum_{i=1}^{k-1}\left(x^{(1)}(i+1) - d_1 x^{(1)}(i) - d_2\right)}{d_1(k-1)}. \tag{17}$$

Step 3: Calculate the predicted future values by using following equation

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k-q+2) - x^{(1)}(k-q+1) + x^{(0)}(k-q+1) \tag{18}$$

where $k = q, q+1, \cdots, n$.

5. MAPE, MAE, and MSE

MAPE, MAE, and MSE are used to measure the ability of a forecasting method used. The mathematical formulation of MAPE, MAE, and MAE is described as follows:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{x_t - \hat{x}_t}{x_t}\right| \times 100\%, \tag{19}$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|x_t - \hat{x}_t|, \tag{20}$$

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(x_t - \hat{x}_t)^2. \tag{21}$$

where n is the number of data series, x_t is the actual values of the ith data series, and $\hat{x}_t$ is the predicted values of the ith data series.

6. Results

The data used in this study are monthly data of oil palm production in Indonesia from 2014 to 2017. This data is obtained from the publication of the Badan Pusat Statistik entitled "*Kelapa Sawit Indonesia*" (BPS, 2018). Figure 1 describes the data.

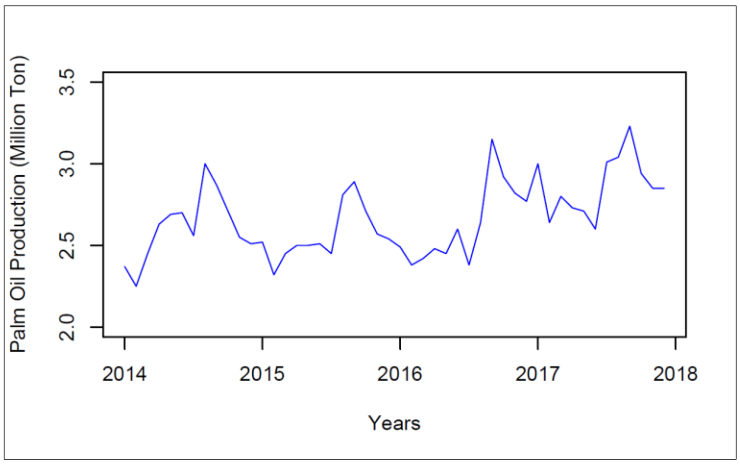

Figure 1: Palm oil production in Indonesia from 2014 to 2017.

Figure 1 explains the time series plot of the amount of oil palm production in Indonesia from January 2014 to December 2017, where the time series pattern repeats in certain months. From the fluctuation pattern in the existing data, it is seen that there are repeated and regular fluctuations, so it can be said that the data contains seasonal characteristics. To ensure the presence of seasonal characteristics in the data, see the autocorrelation function (ACF) and the partial autocorrelation function (PACF) plot in Figure 2.

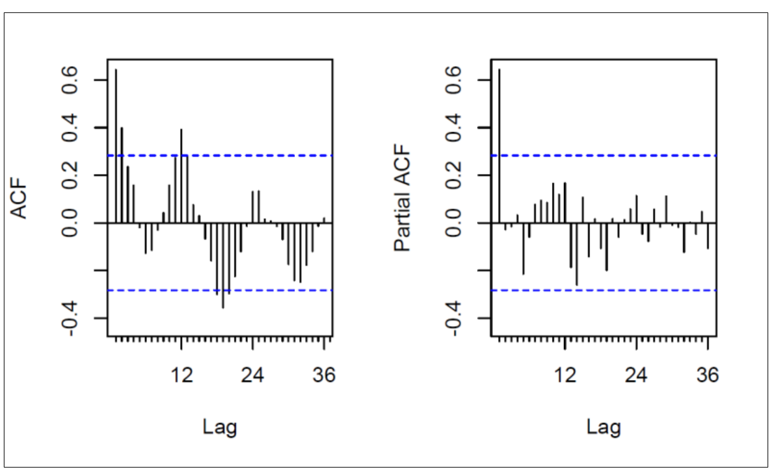

Figure 2: Plot of ACF and PACF from palm oil production.

Figure 2 shows that the data moves up and down or repeatedly forms peaks and valleys over a certain period. Furthermore, Figure 2 also shows that the autocorrelation coefficient peaks are at lag 12, 24, and 36. This form indicates that the data is seasonal. Therefore, we conclude that the data has a seasonal pattern based on these characteristics.

To implement the hybrid model by combining the decomposition method and the grey model GM(1,1), and the seasonal grey model SGM(1,1), the data of palm oil production is divided into two parts as data in-sample and data out-of-sample. The data in-sample starts from January 2014 until November 2017 and the data out-of-sample is December 2017.

In the hybrid model, the decomposition method decomposes the in-sample data into seasonal components and seasonal adjustment components, consisting of trend-cycle and remainder components. Because the palm oil production data have a magnitude of seasonal fluctuations which does not vary with the level of data series, the decomposition method used is additive. The seasonal components in in-sample data are forecasted using the average method, and the seasonal adjustment components in in-sample data are forecasted using GM(1,1). The two results are combined and shown in Figure 3, showing the forecasted palm oil production results in Indonesia using the hybrid model. As a comparison, forecasting palm oil production in Indonesia is also perform using SGM(1,1), and the result is shown in Figure 4.

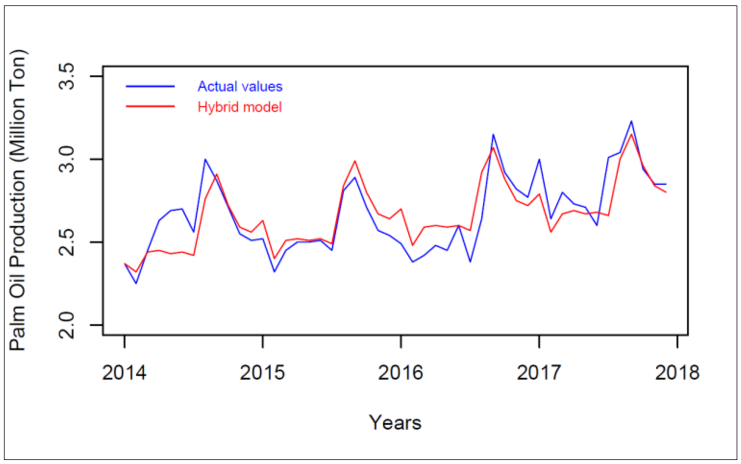

Figure 3: Forecasted results using the hybrid model.

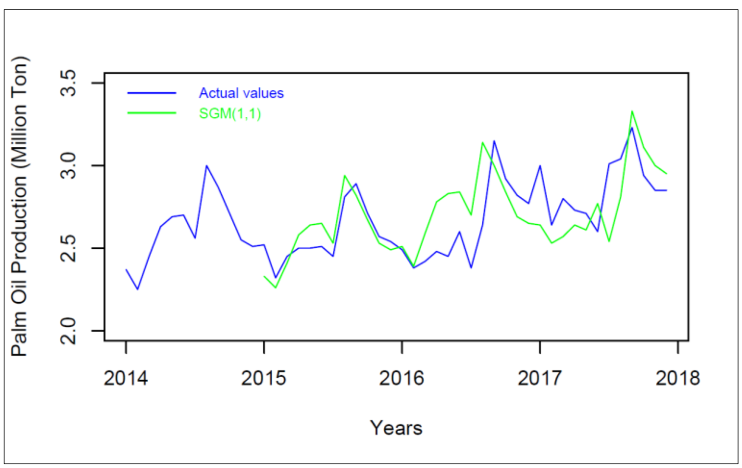

Figure 4: Forecasted results using SGM(1,1).

Figure 3 and Figure 4 represent the forecasted results of palm oil production in Indonesia using the hybrid model and SGM(1,1), respectively. Based on that figure, the forecasted results using the hybrid model have a closer distance to the actual data than the forecasted results using the SGM(1,1). Visually, we conclude that the forecasted results using the hybrid model are better than using the SGM(1,1). To choose the best model accurately, we use MAPE, MAE, and MSE as performance measurement of each model.

Table 1: Performance measurement of the hybrid model and SGM(1,1).

Model	In-sample			Out-of-sample		
	MAPE (%)	MAE	MSE	MAPE (%)	MAE	MSE
Hybrid model	3.858%	0.102	0.015	0.958%	0.027	0.001
SGM(1,1)	5.540%	0.116	0.038	3.692%	0.105	0.011

The results shown in Table 1 indicate that the hybrid model has better forecasting performance than SGM(1,1). The hybrid model has a smaller MAPE, MAE, and MSE than SGM(1,1). In addition, the MAPE of the hybrid model is less than 5% for both data in-sample and data out-of-sample. According to Ma and Zhang (2009), the forecasting accuracy is excellent with that MAPE values.

7. Conclusion

We have implemented the decomposition method with the application grey model GM(1,1) as the hybrid model for forecasting seasonal time series, which is oil palm production in Indonesia. We have also compared the forecasting results using the seasonal grey model SGM(1,1). Based on the comparison of MAPE, MAE, and MSE, we conclude that the decomposition method with application GM(1,1) is better for forecasting seasonal time series than SGM(1,1).

References

1. BPS. (2018). *Kelapa Sawit Indonesia.* Jakarta: Badan Pusat Statistik.
2. Dang, H. S., Huang, Y. F., Wang, C. N., and Nguyen, T., M., T. (2016). An Application of the Short-Term Forecasting with Limited Data in the Healthcare Traveling Industry. *Sustainability*, Volume 8, 1037.
3. Hyndman, R. J. and Athanasopoulos, G. (2018). *Forecasting: Principles and Practice, Second Edition.* Melbourne, Australia: OTexts. Accessed on 13 october 2020.
4. Julong, D. (1989). Introduction to Grey System Theory. *The Journal of Grey System*, Volume 1, Pages 1-24.
5. Liu, S., and Lin, Y. (2006). *Grey Information: Theory and Practical Applications.* London: Springer-Verlag.
6. Ma, H. and Zhang, Z. (2009). Grey Prediction with Marcov-Chain for Crude Oil Production and Consumption in China. In *The Sixth International Symposium on Neural Networks (ISSN 2009)* (pp. 551-561). Berlin, Heidelberg: Springer.
7. Makridakis, S. G., Wheelwright, S. C., and Hyndman, R. J. (2008). *Forecasting: Methods and Applications.* Canda: John Wiley & Sons, Inc.
8. Ozcan, T. (2017). Application of Seasonal and Multivariable Grey Prediction Models for Short-Term Load Forecasting. *Alphanumeric Journal*, Volume 5(2), Pages 329-338.
9. Tseng, F., Yu, H., and Tzeng, G. (2001). Applied Hybrid Grey Model to Forecast Seasonal Time Series. *Technological Forecasting and Social Change*, Volume 67, Issues 2–3, Pages 291-302.
10. Xia, M., and Wong, W. K. (2014). A Seasonal Discrete Grey Forecasting Model for Fashion Retailing. *Knowledge-Based Systems*, Volume 57, Pages 119-126.