

Forecasting of Poverty Data Using Seasonal ARIMA Modeling in West Java Province

Desri Kristina Silalahi

Departement of Electrical Engineering, Telkom University, Indonesia

desrikristina@telkomuniversity.ac.id

ABSTRACT

Article History:

Received : 29-02-2020

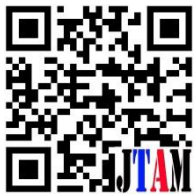
Revised : 30-03-2020

Accepted : 31-03-2020

Online : 02-04-2020

Keyword:

Forecasting;
Poverty Data;
SARIMA Method;
Government Policy.



The government continues to carry out poverty reduction strategies in Indonesia, especially in West Java Province. West Java Province is a province that has the most populous population in Indonesia. This will affect the level of welfare and the amount of poverty. The strategy undertaken is inseparable from accurate poverty data and is available from year to year. Even from the available data, the government can forecast the number of poor people in the coming years. Seasonal Autoregressive Integrated Moving Average (SARIMA) method is one of forecasting methods. SARIMA is the development of the ARIMA model which has a seasonal effect. Based on the results of the study, that poverty data forecasting in the province of West Java using the SARIMA method obtained SARIMA model $(0,1,1) (1,1,1)_4$. This model is the best model for forecasting data with an R-Squared value of 98%, Mean Square Error is 7.705.5800.000 and Mean Absolute Percentage Error is 2,81%. It's means this SARIMA model is very good in predicting poverty data in West Java Province.



<https://doi.org/10.31764/jtam.v4i1.1888>



This is an open access article under the **CC-BY-SA** license

A. INTRODUCTION

The problem of poverty is one of the problems that is still faced by Indonesian state today. Poverty is a complex and multidimensional problem, because it is one indicator to determine the success rate of a country's development (Claro et al., 2016). The development program has been successful if the poverty rate is low so that it can improve the welfare of people's lives (Matondang, 2017). The problem of poverty is a global problem faced by many countries because one of the fundamental problems involves the needs in life. Thus, poverty is a problem that has existed and is still being faced by countries in the world including in Indonesia.

The Indonesian government has made efforts to resolve the problem of poverty, starting from regional, regional, national and international. However, various government policies have not been able to eradicate poverty. Based on the latest data in March 2019 from the Badan Pusat Statistik (BPS- Statistics Indonesia), Indonesia's poor population is 25.14 million.

This number has decreased by 810 thousand population compared to the previous year in the same period. Thus, the percentage of poor people amount as of March 2019 was recorded at 9.41 percent. This percentage has decreased compared to the previous year of 9.82 percent. West Java Province is the province with the most population in Indonesia. This will affect the level of welfare and the amount of poverty is still large in West Java Province. Based on BPS data, that the poor population in West Java has decreased by around 140.2 thousand inhabitants. In September 2018 it was 3,539.40 thousand people to 3,399.16 thousand people in March 2019. BPS categorizes the poor population as residents with per capita expenditure per month below the poverty line.

The government continues to carry out poverty reduction strategies in Indonesia. The strategy undertaken is inseparable from accurate poverty data and is available from year to year. This will help the government coordinate with regional heads to make decisions that must be made in poverty alleviation. Even from the available data, the government can forecast the number of poor people in next years. The government can plan strategies that will be implemented to reduce the population in the future if forecasting result is very large. Forecasting obtained from the analysis of data available from time to time. This analysis is called time series analysis. Time series analysis is used to perform data analysis that considers the effect of time sequence in hours, days, weeks, months, quarters and years.

Time series modelling observes observational behavior taken from time to time sequentially (Hillmer & Wei, 2006). The application of time series analysis for forecasting in various fields as financial, social and economic. Forecasting can resolve problems in their fields to help preparation of plans, strategic and decision making for the future. Forecasting methods that can be used are very diverse. Application of forecasting methods that have been carried out in the economic field of customer transaction forecasting at Store X using the Seasonal Autoregressive Integrated Moving Average (SARIMA) method is more accurate than the Holt-Winters Exponential Smoothing method (Pongdatu & Putra, 2018). In addition, applications in the field of health in predicting the occurrence of Tuberculosis using the SARIMA method produce an accurate model (Patowary & Barman, 2017). The forecasting application using the SARIMA method is very accurate in forecasting electricity consumption for health service buildings (Kaur & Ahuja, 2019) and monthly flow forecasting in the Waterval River in South Africa (Tadesse & Dinka, 2017). The Naïve method is more accurate in forecasting poverty data than the Moving Average method (Kumila et al., 2019). In addition, the Autoregressive Distributed Lag method can be used to Forecasting accurately poverty data in West Nusa Tenggara Province (Rahmasari et al., 2019).

This research predicted poverty data in West Java Province using the Seasonal Autoregressive Integrated Moving Average (SARIMA) method. ARIMA is considered more popular in forecasting because it is more flexible and represents many variations of data at a particular time series (Seymour et al., 2002). While SARIMA is the development of the ARIMA model which has a seasonal effect (W. Gikungu, 2015). The pattern of poverty data each year has a seasonal effect depending on economic conditions and changing positions in the government, so it can be affect government effort in poverty reduction.

B. METHODS

1. Data

The forecasting technique uses the Seasonal Autoregressive Integrated Moving Average (SARIMA) method. If the data has a seasonal pattern, then a more appropriate method is to use SARIMA. Seasonal is the tendency to repeat the pattern of behavior in the season period, usually one year for monthly data (Ukhra, 2014). The ARIMA Seasonal Model is an ARIMA model that is used to complete seasonal time series which consists of two parts, namely the non- section and the seasonal section. The non-seasonal part of this method is the ARIMA model.

The research design uses a comparative ex post facto causal research design, meaning that the data was collected after all the events in question took place. Secondary data used are population poverty data of West Java province from 2002 to 2019. In this study the sampling technique used in sampling is purposive sampling. The analysis used is descriptive statistical analysis and SARIMA method. Data obtained from the official website of Badan Pusat Statistik (BPS- Statistics Indonesia), are presented in the following Table 1 below:

Table 1. Number of Poor Population of West Java Province

Year	Total Population	Year	Total Population
2002	3.399.500	2011	4.716.800
2003	3.615.700	2012	4.852.300
2004	4.168.700	2013	5.249.600
2005	4.224.600	2014	5.457.800
2006	4.435.600	2015	5.712.400
2007	4.238.900	2016	5.137.600
2008	4.375.100	2017	4.654.400
2009	4.430.100	2018	4.898.700
2010	4.650.600	2019	4.938.300

(Source: <https://jabar.bps.go.id>)

2. ARIMA Model Time Series

Time series are events that occur based on certain times in sequence. Forecasting in time series is an action to predict events that will occur in the future based on previous data (Hillmer & Wei, 2006). One of time series models is ARIMA. ARIMA is often also referred to as the Box-Jenkins time series (Ngo, 2013). ARIMA is usually used for short-term forecasting. The ARIMA model in predicting long term is not good and appropriate because it tends to be flat for a long period (Octora & Kuntoro, 2013). ARIMA model is defined as follows (Hillmer & Wei, 2006):

a. Autoregressive (AR) Model

The Autoregressive model is a regression equation model connect to previous values from a dependent variable with its variable (Ekananda, 2014). The general form of the Autoregressive model with the order p written with AR (p) or ARIMA model ($p, 0,0$) is stated as follows:

$$Y_t = \alpha_0 + \theta_1 Y_{t-1} + \dots + \theta_p Y_{t-p} + e_t \quad (1)$$

with Y_t is observation value at the time t ; α_0 is constant; θ_p is autoregressive parameter to p , and e_t is error value t .

b. Moving Average (MA)

Moving average is a relationship current value as the number of small interruptions in previous time or depends on the previous value of the error element. The function of Moving Average will explain an observation t that expressed as a linear combination of a number errors. Process of Moving Average have order q can be stated as ARIMA (0,0, q) or it can be written MA (q). The general form of MA (q) is as follows.

$$X_t = e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (2)$$

with θ_q is moving average parameter p , and e_{i-q} is error value on $i - q$

c. Autoregressive Moving Average (ARMA)

This model is a combination of the Autoregressive (AR) process and the Moving Average (MA) process. So, combination of this models can be well calculated past values of the predicted variable and value of errors in the past. ARMA model can be stated (p,q). The general equation for this model is as follows:

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (3)$$

with Z_t is data on the period t , Z_{t-p} is data on the period $t-p$, e_t is error on the period t , e_{t-q} is error on the period $t-q$, ϕ_p is autoregressive parameter p , θ_q is moving average parameter q .

d. Autoregressive Integrated Moving Average (ARIMA)

The ARIMA model is a development of ARIMA with the addition of one parameter d , so that the ARIMA model has three parameters namely (p, d, q), where p is an autoregressive parameter (AR), q is a moving average (MA) parameter, d is a number of differencing processes. The reason for the development of the ARIMA method is because in some other methods there are assumptions for certain types of data patterns, whereas in ARIMA there are no assumptions regarding a fixed pattern (Hermawan, 2011). The general equation of the ARIMA model (p, d, q) is as follows:

$$Z_t = \phi_0 + (1 - \phi_1)Z_{t-1} + (\phi_2 - \phi_1)Z_{t-2} + \dots + (\phi_p - \phi_{p-1})Z_{t-p} - \phi_p Z_{t-p-1} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (4)$$

with Z_t is data on the period t , Z_{t-p} is data on the period $t-p$, a_t is error on the period t , a_{t-q} is error on the period $t-q$, ϕ_0 is constant, ϕ_p is autoregressive on p , and θ_q is moving average on q .

e. Seasonal ARIMA (SARIMA) Model

This model is a model that contains elements of seasonal patterns. Seasonal is the tendency to repeat the pattern of behavior in the season period, usually one year for monthly data (Wibowo, 2018). This model is a development of the ARIMA model. If the ARIMA is used for non-seasonal data or in other words the ARIMA is not suitable for seasonal data, so to overcome seasonal problems, the development of the previous model, now known as the Seasonal ARIMA model, is developed. The general form of the ARIMA model with seasonal patterns $(p, d, q)(P, D, Q)_s$ can be stated as follows:

$$\Phi_p(B^s)\phi_p(B)(1-B)^d(1-B^s)^D Z_t = \theta_q(B)\Theta_q(B^s)a_t \tag{5}$$

with: p,d,q is AR order, differencing order and MA nonSeasonal order, P,D,Q is AR order, differencing order and MA Seasonal order, $\phi_p(B)$ is AR non-seasonal order, $\Phi_p(B^s)$ is AR seasonal order, $(1-B)^d$ is differencing non-seasonal order, $(1-B^s)^D$ is differencing seasonal order, $\theta_q(B)$ is MA non-seasonal, $\Theta_q(B^s)$ is MA seasonal, a_t is error on t period.

3. Parameters Errors

a. Mean Square Error (MSE)

Mean Square Error (MSE) is to look for errors from the average error in observations (Willmott & Matsuura, 2005). MSE can be used to find out how big is the error in the data from the model used. MSE can be used as an indicator of incompatibility in modeling. RMSE can be calculated using

$$MSE = \frac{\sum_{i=1}^n (x - x_i)^2}{n} \tag{6}$$

with x is forecast value, x_i is observation value to- I , n is the amount of data.

b. Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) is the absolute error in each period divided by the real observed value for that period. Then calculate the average of the absolute percentage error (Kim & Kim, 2016). Just like RMSE, MAPE calculates errors from observational and predictive data that appear in percentage values. MAPE values can be calculated by the following equation:

$$MAPE = \left(\frac{100\%}{n}\right) \frac{\sum_{i=1}^n |x - x_i|}{x_i} \tag{7}$$

The MAPE accuracy criteria can be seen in the following table (Chang et al., 2007):

Forecasting	MAPE Percentage
Excellent forecasting	MAPE < 10%
Good forecasting	MAPE 10% - 20%
Enough forecasting	MAPE 20% - 50%
Inaccurate forecasting	MAPE > 50%

c. R-Squared

R-squared is a method for linear regression models to test the merit of a model (Malkina-Pykh & Pykh, 2019). R-square is 0-1, the closer the value is to 1, the better the model is. The R-square value can be calculated using the following equation:

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - x)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \tag{8}$$

4. Forecasting Stages

The steps in forecasting using the SARIMA method are described in the following flow chart:

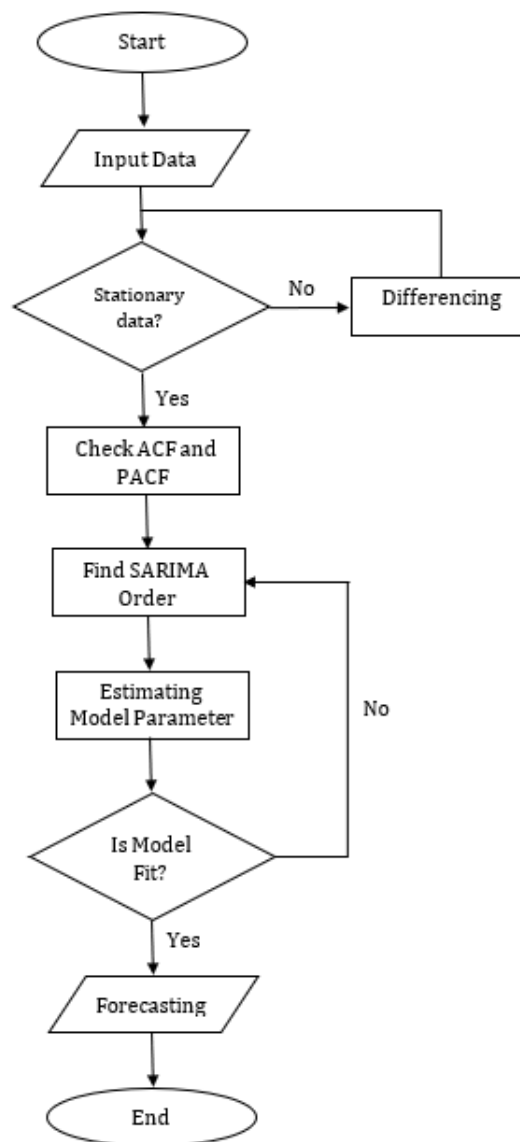


Figure 1. Forecasting Flow Chart

Based on Figure 1, the first step is to identify the model through a time series plot, Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). The data is made into a time series plot, ACF and PACF. Then the data can be determined stationary or not. If the data is not stationary to the mean, then differencing is performed. In determining the order of AR (p), consider the Partial Autocorrelation Function (PACF) graph. If the autocorrelation coefficient is significantly different from zero or there is a lag on the PACF that comes out of the boundary line then that lag will be the AR order. Conversely, to determine the order MA (q) then consider the Autocorrelation Function (ACF) graph if the autocorrelation coefficient is significantly different from zero or there is a lag in the ACF that comes out of the boundary line then that lag will be the MA order.

After checking ACF and PACF, the SARIMA model will be estimated, then the model obtained will be tested. Testing this model to test the feasibility of the model obtained by looking at the value of RMSE, MAPE and R-Squared. Models that are feasible will be used to make predictions for the future

C. RESULT AND DISCUSSION

1. Research data

Figure 2 is a time series plot of poverty data in the province of West Java from 2002 to 2019.

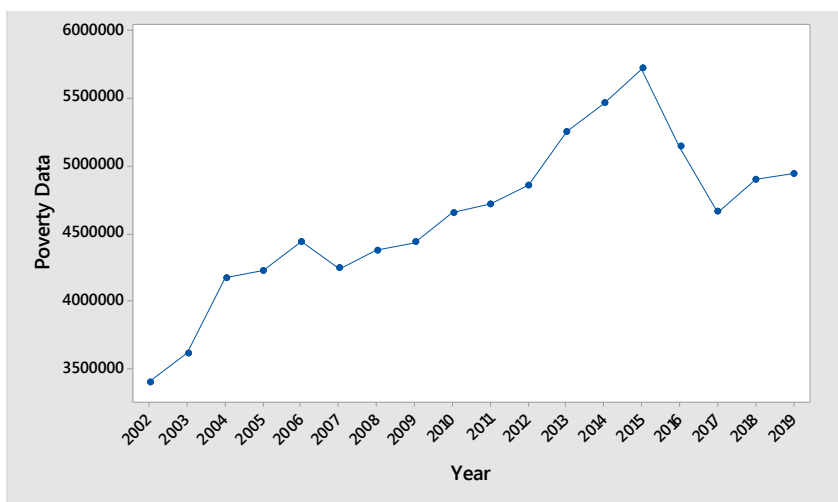


Figure 2. Poverty Time Series Plot Data

Based on the graph, it shows that the data is influenced by both trend and seasonal patterns because the plot shows fluctuations that increase, ie movements from lower left to upper right on the graph and recur in certain years. However, the graph shows that seasonal patterns are stronger than trend patterns so that trend patterns appear less clear. Furthermore, the data will be seen from ACF and PACF. The following chart is ACF and PACF

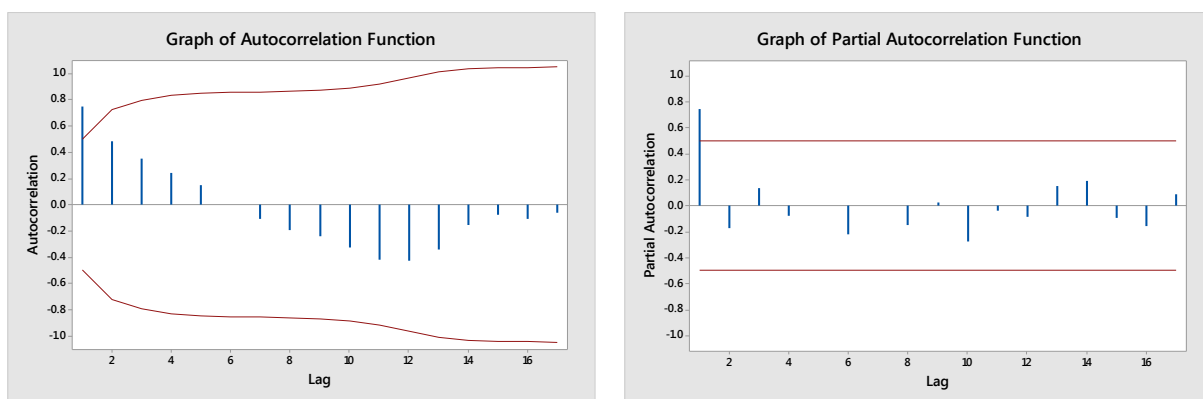


Figure 3. Graph of ACF and PACF

2. Seasonal ARIMA Model

After determining ACF and PACF, Figure 3 shows autocorrelation in the data, which is a bar that crosses the red line so the data is not stationary. Therefore, poverty data needs to be

done differencing to stationing the data. After differencing, ACF and PACF charts will be observed. Differencing is done is the first differencing. The following chart is ACF and PACF after differencing $d = 1$.

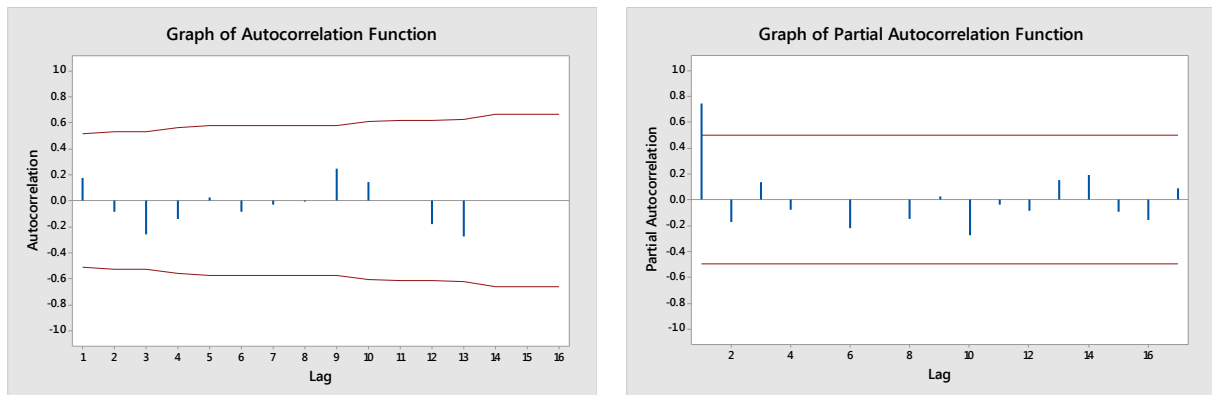


Figure 4. ACF and PACF Graphs after First Differencing

Based on Figure 4, after the first differencing ACF graph is seen cut off, while the PACF graph is seen dying down in lag 1. Thus, the ARIMA model is pure MA with $p = 0$, $d = 1$ and $q = 1$. The ARIMA model can be written ARIMA (0, 1, 1). Furthermore, when viewed from the time series data graph in Figure 2, it appears that the pattern repeated within a period of 4 years. Furthermore, if observed from the ACF results in lag 4 decreased significantly, while PACF in lag 3 increased while decreased in lag 4. Thus, the ARIMA model is a seasonal ARIMA. The next stage is differencing first in lag 4 and then ACF and PACF are determined. The following chart is ACF and PACF after the first differencing with seasonal $S = 4$.

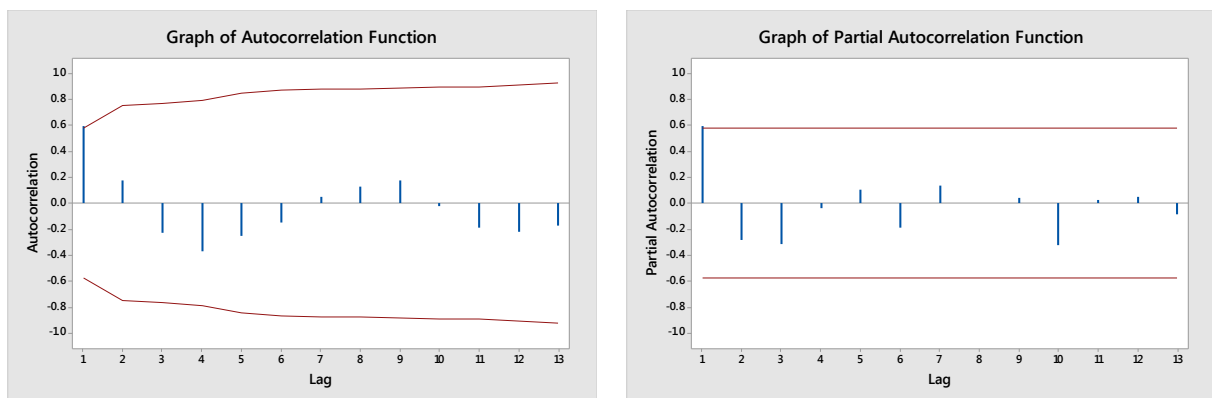


Figure 5. ACF and PACF Graphs after Differencing $d = 1$ with Seasonal $S = 4$

Based on Figure 5, after the first differencing and seasonal lag 4, the data is stationary in the mean and variance. Figure 3 shows the fluctuations in the data along the time axis and means the values $d = 1$ and $S = 4$. The ACF graph looks dying down in lag 1 meaning that the AR value is in seasonal ARIMA. Thus, the ARIMA seasonal model is $p = 1$, $d = 1$ and $q = 1$. Thus, the ARIMA seasonal model is [ARIMA (0,1,1)(1,1,1)₄]. SARIMA modeling results are obtained after conducting trials and errors in model selection. This modeling is considered as the best modeling for poverty data prediction. The following table provides a combination of the SARIMA model.

Table 3. Combination of ARIMA

Combination of Model	MSE	MAPE	R-Squared
ARIMA (0,1,1) (1,1,1) ₄	7.705.5800.000	2,81%	0,98
ARIMA (0,1,2) (1,1,1) ₄	62.002.800.000	2,38%	0,91

Based on Table 3, the MSE value is greater in the ARIMA model (0,1,1) (1,1,1)₄ and the R-Squared value is also greater, then the ARIMA model (0,1,1) (1,1,1)₄ is used as a prediction model because it is able to forecast very accurately. Next, parameter estimation is performed for the ARIMA model (0,1,1) (1,1,1)₄. The following table presents the estimated parameters of the ARIMA model (0,1,1) (1,1,1)₄.

Table 4. ARIMA Model Parameter Value (0,1,1) (1,1,1)₄

Variable	Coefficient	Std. Error	t-value	p-value
SAR (4)	-0,836	0,333	-2,51	0,033
MA (1)	0,991	0,297	3,33	0,009
SMA (4)	0,713	0,512	1,39	0,197
Constant	-119069	1317	-90,40	0,000

Based on Table 4, the parameter value of each variable has a p-value < 0.05 except the Seasonal MA variable (4) has a p-value > 0.05, so it is not used in mathematical modeling. However, other variables can be used in the model. Therefore, mathematically the SARIMA model (0,1,1) (1,1,1)₄ is:

$$Z_t = -11907 + a_t + a_{t-1} + 0,845a_{t-4} - 0,836a_{t-5} - 0,836a_{t-8} + 0,836a_{t-9} \tag{9}$$

ARIMA Seasonal prediction models obtained, then forecasting poverty data from 2002 to 2009 in West Java province. The following graph is the result of poverty data prediction:

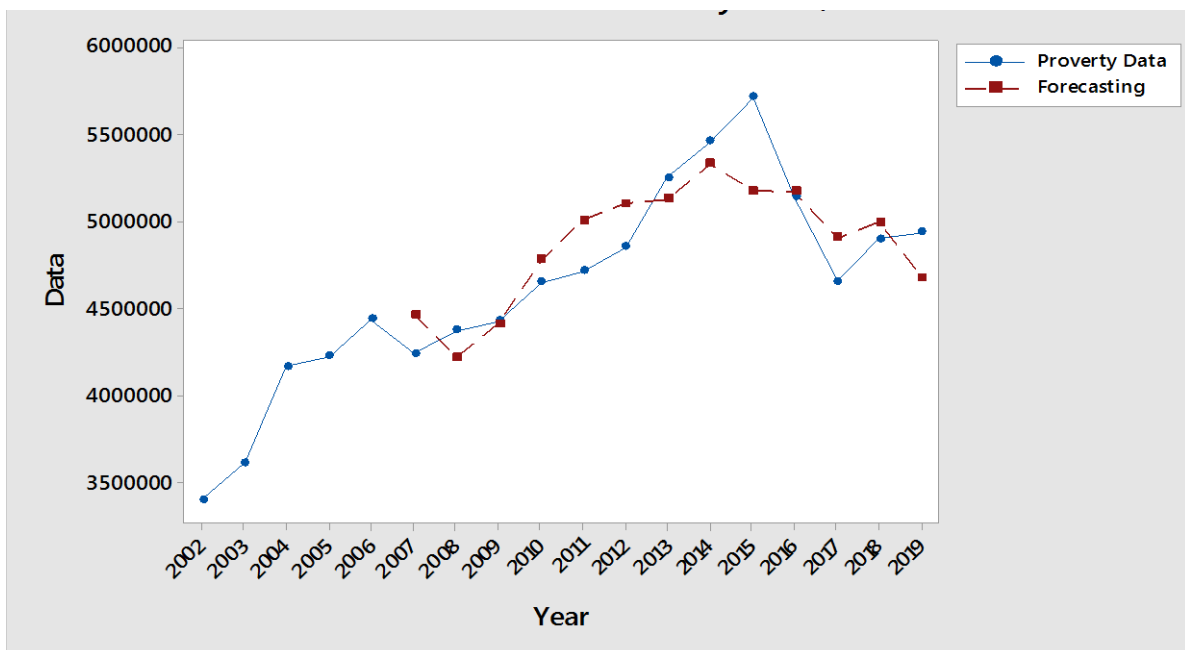


Figure 6. Poverty Data and Forecasting Results in West Java Province

Furthermore, the ARIMA Seasonal prediction model can be done for the next 5 years forecasting, from 2020 to 2024 for poverty data in West Java. Prediction of the amount of poverty data in the future can make a picture of the data used by stakeholders in decision-making in reducing the poverty rate of a region. The policy taken can reduce the poverty rate that occurs in an area. Policies that can be taken by increasing income and better community welfare. Most families are poor has a large number of family members so the condition their economies on the poverty line have deteriorated along with worsening income or welfare inequality (Silastri et al., 2017). Based on SARIMA modeling obtained in equation 9, then the following table forecasting the next 5 years:

Table 5. Poverty Data Forecasting Results in West Java Province

Year	Forecast	Lower	Upper
2020	4.984.285	4.440.100	5.528.469
2021	5.270.816	4.726.607	5.815.025
2022	5.258.144	4.713.910	5.802.377
2023	5.105.535	4.561.278	5.649.793
2024	4.696.876	4.078.319	5.315.434

D. CONCLUSION AND SUGGESTIONS

The conclusion from this research that the SARIMA method can be used for seasonal data that can provide forecasting results that are not much different from the actual data. In forecasting poverty data in West Java Province, the SARIMA model (0,1,1) (1,1,1)₄ is the best model for forecasting the data with an R-Squared value is 98%. Means the SARIMA model (0,1,1) (1,1,1)₄ is very good in predicting poverty data in West Java Province in the next 5 years from 2020 to 2024. Furthermore, to calculate forecasting or predict things in economics, statistics and so on, it would be better to try various methods and use more data. To find the most accurate results from the method performed.

REFERENCES

- Chang, P.-C., Wang, Y.-W., & Liu, C.-H. (2007). The development of a weighted evolving fuzzy neural network for PCB sales forecasting. *Expert Systems with Applications*, 32, 86–96. <https://doi.org/doi:10.1016/j.eswa.2005.11.021>
- Claro, S., Paunesku, D., & Dweck, C. S. (2016). Growth mindset tempers the effects of poverty on academic achievement. *Proceedings of the National Academy of Sciences of the United States of America*. <https://doi.org/10.1073/pnas.1608207113>
- Ekananda, M. (2014). *Analisis Data Time Series Untuk Penelitian Ekonomi, Manajemen, Dan Akutansi*. Mitra Wacana.
- Hermawan. (2011). Perbandingan metode Box-Jenkins dan Holt-Winters dalam memprediksi anomali Olr Pentad di kawasan barat Indonesia. *Jurnal Sains Dirgantara*, 9(1), 27–34.
- Hillmer, S. C., & Wei, W. W. S. (2006). *Time Series Analysis: Univariate and Multivariate Methods*. Addison Wesley Publishing Company, Inc. <https://doi.org/10.2307/2289741>
- Kaur, H., & Ahuja, S. (2019). SARIMA modelling for forecasting the electricity consumption of a health care building. *International Journal of Innovative Technology and Exploring Engineering*, 8(12), 2795–2799. <https://doi.org/10.35940/ijitee.L2575.1081219>
- Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, 32(3), 669–679.

- <https://doi.org/https://doi.org/10.1016/j.ijforecast.2015.12.003>
- Kumila, A., Sholihah, B., Evizia, Safitri, N., & Fitri, S. (2019). Perbandingan metode Moving Average dan metode Naïve dalam peramalan data kemiskinan. *Jurnal Teori Dan Aplikasi Matematika (JTAM)*, 3(1), 65–73. <https://doi.org/https://doi.org/10.31764/jtam.v3i1.764>
- Malkina-Pykh, I. G., & Pykh, Y. A. (2019). Some notes on the application of R-squared for evaluation the goodness-of-fit of nonlinear regression models. *8th International Nonlinear Science Conference*, 1–7. <https://doi.org/10.13140/RG.2.2.35129.21607>
- Matondang, E. (2017). Finding Out the Potency of Nusa Tenggara Timur in Poverty Allevation: The Effect of Local Government's Policy. *Jurnal Bina Praja*, 9(2), 231–242. <https://doi.org/10.21787/jbp.09.2017.231-242>
- Ngo, T. H. D. (2013). The Box-Jenkins Methodology for Time Series Models. *Proceedings of the SAS Global Forum 2013 Conference*.
- Octora, M., & Kuntoro. (2013). Perbandingan metode ARIMA (Box Jenkins) dan metode Winter dalam peramalan jumlah kasus demam berdarah dengue. *Jurnal Biometrika Dan Kependudukan*, 2(1), 88–98.
- Patowary, A. N., & Barman, M. P. (2017). Development of a SARIMA model to forecast Tuberculosis detection rate in the Dibrugarh District of Assam, India. *International Journal of Advanced Research in Computer Science*, 8(7), 984–988. <https://doi.org/http://dx.doi.org/10.26483/ijarcs.v8i7.4424>
- Pongdatu, G. A. N., & Putra, Y. H. (2018). Seasonal Time Series Forecasting using SARIMA and Holt Winter's Exponential Smoothing. *IOP Conference Series: Materials Science and Engineering*. <https://doi.org/10.1088/1757-899X/407/1/012153>
- Rahmasari, U., S, E. H., Jannah, M., Pathullaili, Kurnia, L., & Satria, A. (2019). ARDL method: forecasting data kemiskinan di NTB. *Jurnal Teori Dan Aplikasi Matematika (JTAM)*, 3(1), 52–57. <https://doi.org/https://doi.org/10.31764/jtam.v3i1.767>
- Seymour, L., Brockwell, P. J., & Davis, R. A. (2002). *Introduction to Time Series and Forecasting*. Springer. <https://doi.org/10.2307/2965440>
- Silastri, N., Iyan, R., & Sari, L. (2017). Pengaruh Jumlah Penduduk dan Pendapatan Domestik Regional Bruto (PDRB) terhadap Kemiskinan di Kabupaten Kuantan Singingi. *Jurnal Online Mahasiswa Fakultas Ekonomi Universitas Riau*, 4(1), 105–117.
- Tadesse, K. B., & Dinka, M. O. (2017). Application of SARIMA model to forecasting monthly flows in Waterval River, South Africa. *Journal of Water and Land Development*, 35(X–XII), 229–236. <https://doi.org/10.1515/jwld-2017-0088>
- Ukhra, A. U. (2014). Pemodelan dan peramalan data deret waktu dengan metode SEASONAL ARIMA. *Jurnal Matematika UNAND*, 3(3), 59–67.
- W. Gikungu, S. (2015). Forecasting inflation rate in Kenya using SARIMA model. *American Journal of Theoretical and Applied Statistics*, 4(1), 15–18. <https://doi.org/10.11648/j.ajtas.20150401.13>
- Wibowo, A. (2018). Model peramalan indeks harga konsumen kota Palangka Raya menggunakan Seasonal ARIMA (SARIMA). *Jurnal Matematika*, 17(2), 17–24.
- Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30(1), 79–82. <https://doi.org/10.3354/cr030079>