Application of Clustering and VARIMA for Rainfall Prediction

Asep Rusyana\textsuperscript{1}, Nidya Tatsara\textsuperscript{1*}, Rozatul Balqis\textsuperscript{1}, Syafinatur Rahmi\textsuperscript{1},
Department of Statistics, Faculty of Mathematics and Natural Sciences, Syiah Kuala University, Darussalam, Banda Aceh, Aceh, 23111, Indonesia

E-mail: nidya.t@mhs.unsyiah.ac.id

Abstract. Global warming causes climate change throughout the world which has an impact on changes in erratic rainfall. For fishermen, the estimated rainfall is an important thing to know before sailing. VARIMA is divided into VAR, VMA, VARMA, and VARIMA. This study uses a rainfall dataset consisting of rainfall (mm), temperature (°C), humidity (%) and wind speed (km/hour) with observations per day for one year in 2018. Training data is data that starts from 1 January 2018 until 30 November 2018 and the testing data starts on 1 December 2018 until 31 December 2018. Clustering was carried out in the area around Lampulo Pelabuhan Perikanan Samudera (PPS) which are Banda Raya, Lampineung, Kuta Radja, and Kuta Alam. The best model obtained is VARIMA (1,1,2), which is based on the smallest AIC and BIC values, which are 8.5104 and 9.0581. The VARIMA (1,1,2) model also fulfills the white noise test, so it can be concluded that the data does not have autocorrelation and model can be used for predictions. Estimated values of rainfall, temperature, humidity and wind speed are not much different from actual values, with MAPE of all variables are relatively low.

1. Introduction
Climate change is the impact of global warming. One of the effects of climate change is the irregular pattern of rainfall. Rainfall is the amount of water falling on a flat ground surface during a certain period measured with a height of a millimeter (mm) above a horizontal surface. Each millimeter of rainfall means that in the area of one square meter in the flat is a millimeter high water [1]. The factors that affect high low rainfall are temperature (°C), humidity (%) and wind speed (km/h) [2]. One sector that perceives the effects of irregular rainfall patterns is the maritime sector, especially for fishermen.

Aceh is the westernmost region in Indonesia and surrounded by 295,370 km\textsuperscript{2} of the ocean. Because the position makes Aceh is the most strategic area especially in the maritime sector. Banda Aceh as the capital of Aceh has an international ocean fishing port called Pelabuhan Perikanan Samudera (PPS) Lampulo. Lampulo lies in the Kuta Alam area, see Figure 1. Therefore, some inhabitants in Lampulo in Banda Aceh generally work as fishermen. Information about rainfall prediction is crucial for fishermen before sailing. It is useful to avoid high waves and storms in the ocean so that it can reduce the risk of accidents.

*Corresponding Author: nidya.t@mhs.unsyiah.ac.id
Clustering is a data grouping that has similarities or nearby territories. Information about future conditions can not be determined but can be predicted. One of the models for the multivariate time series prediction is Vector Autoregressive Integrated Moving Average (VARIMA). VARIMA is the generalization of Autoregressive Integrated Moving Average (ARIMA) which is a method of analysis of univariate time series. A univariate time series model sometimes results in less representative predictions. This is due to the influence of other variables that do not count towards the model. Therefore, this issue can be resolved by using a multivariate time series analysis such as VARIMA.

VARIMA model is a vector form of the ARIMA model that works by taking into account data and error data from self-changing and other reviewers simultaneously. So that the assessed changes should have a significant interchanges correlation. VARIMA models consist of 4 parts, namely Vector Autoregressive (VAR), Vector Moving Average (VMA), Vector Autoregressive Moving Average (VARMA) and Vector Autoregressive Integrated Moving Average (VARIMA). ARIMA is not only developed to VARIMA but also other methods such as ARIMA-Autoregressive Conditionals Heteroscedastic (ARCH) and Seasonal ARIMA (SARIMA). Research has been conducted to predict nutmeg prices in South Aceh District by using ARCH[3]. SARIMA has been applied to estimate the number of tourists who land at the Kualanamu International Airport in Medan [4].

Vector Autoregressive (VAR) is the generalization of AR to multiple parallel time series. VAR is a system of dynamic equations, with the estimation of a modified period of time depending on the movement of the variable and other variables involved in the system in the previous period [5]. VAR is used to project time-series variables by analyzing the dynamic impact of the interference factor contained in the variable [6]. The formulation for VAR order p or VAR(p) is as follows [7].

\[ Y_t = \Phi_0 + \sum_{i=1}^{p} \Phi_i Y_{t-i} + a_t \]  \hspace{1cm} (1)

Where,
- \( Y_t \) : \((n \times 1)\), n variables at \( t \) time
- \( \Phi_0 \) : vector averages\((n \times 1)\)
- \( \Phi_p \) : coefficient matrix \((n \times n)\), with \( i = 1, 2, 3... p \)
- \( a_t \) : error vector\((n \times 1)\)

Vector Moving Average (VMA) is the generalization of MA. VMA is a system of dynamic equations, with the estimation of a modified period of time depending on the error of variable and other error of variables involved in the system in the previous period. The formulation for VMA order q or VMA(q) is as follows [7].

![Figure 1. Location map of PPS Lampulo.](image)
\[ Y_t = \Theta_0 + \sum_{i=1}^{q} \Theta_i a_{t-i} + a_t \]  
(2)

Where,
\[ Y_t \] : \((n \times 1)\), \(n\) variables at \(t\) time
\[ \Theta_0 \] : vector averages \((n \times 1)\)
\[ \Theta_q \] : coefficient matrix \((n \times n)\), with \(i = 1, 2, 3... q\)
\[ a_t \] : error vector \((n \times 1)\)

Vector Autoregressive Moving Average (VARMA) is the generalization of ARMA. VMA is a system of dynamic equations, with the estimation of a modified period of time depending on the movement of the variable, error of variable and other variables and error of variables involved in the system in the previous period. The formulation for VARMA\((p,q)\) is as follows[7].
\[ Y_t = \Phi_0 + \Phi_1 Y_{t-1} + ... \Phi_p Y_{t-p} + a_t - \Theta_1 a_{t-1} - ... - \Theta_q a_{t-q} \]  
(3)

Where,
\[ Y_t \] : \((n \times 1)\), \(n\) variables at \(t\) time
\[ \Phi_0 \] : vector averages \((n \times 1)\)
\[ \Phi_p \] : coefficient matrix of VAR \((n \times n)\)
\[ \Theta_q \] : coefficient matrix of VMA \((n \times n)\)
\[ a_t \] : error vector \((n \times 1)\)

VARIMA is a VARMA model that is experiencing the differencing process. The formulation of the VARIMA model is as follows [7].
\[ \Phi_p (B)(I - B)^d Y_t = \Phi_0 + \Theta_q (B)a_t \]  
(4)

Where,
\[ Y_t \] : \((n \times 1)\), \(n\) variables at \(t\) time
\[ \Phi_p \] : coefficient matrix of VAR \((n \times n)\)
\[ \Theta_q \] : coefficient matrix of VMA \((n \times n)\)
\( (I - B)^d \) : differencing components
\[ a_t \] : error vector \((n \times 1)\)
\( B \) : backshift operator

Or it can also be written as follows:
\[ Y_t = (I + \Phi_1) Y_{t-1} + (\Phi_2 - \Phi_1) Y_{t-2} + ... + (\Phi_p - \Phi_{p-1}) Y_{t-p} - \Phi_p Y_{t-p-1} + a_t - \Theta_1 a_{t-1} - ... - \Theta_q a_{t-q} \]  
(5)

The problem of this research is how the rainfall, temperature, humidity, and wind speed overview in the area around PPS Lampulo, Banda Aceh and what are the results of rainfall, temperature, humidity, and wind speed prediction in Banda Aceh for an upcoming period using the Vector Autoregressive Model Integrated Moving Average (VARIMA). The purpose of this research is to know the overview of rainfall, temperature, humidity and wind speed around the PPS Lampulo area in Banda Aceh and find out the prediction of rainfall, temperature, humidity and wind speed in Banda Aceh for an upcoming period.
2. Methodology

2.1. Data and Variable

The data used in this research is secondary data obtained from Meteorological, Climatological, and Geophysical Agency (BMKG) meteorology station class 1 Sultan Iskandar Muda at Blang Bintang, Aceh Besar and Meteorological, Climatological, and Geophysical Agency (BMKG) climatology station at Indrapuri, Aceh Besar.

The data used in this research for descriptive analysis (clustering) is the rainfall data in the area surrounding Lampulo, Banda Aceh which has a rain post that is Kuta Alam, Lampineung, Kuta Radja, and Bandar Raya. While data for the analysis of inference using rainfall data, temperature, humidity and wind speed in Banda Aceh. The period of data on rainfall, temperature, humidity and wind speed used in this research is daily data of 2018. The variables used in this research are presented in Table 1.

| Variable | Description | Unit | Type of Data |
|----------|-------------|------|--------------|
| Y_1      | Rainfall    | mm   | Ratio        |
| Y_2      | Temperature | ºC   | Interval     |
| Y_3      | Humidity    | %    | Ratio        |
| Y_4      | Wind speed  | km/hour | Ratio         |

The data used in this research is divided into training and testing data. Training data is used to create models for predictions i.e. data from 1 January 2018 until 30 November 2018 while testing data is used to measure the accuracy of the model in conducting predictions i.e. data from 1 December 2018 Until 31 December 2018.

2.2. Data Analysis Procedure

The procedure of data analysis in this research has the following stages:

1. A descriptive analysis is the clustering of rainfall, temperature, humidity and wind speed in the area around PPS Lampulo. There are Bandar Raya, Lampineung, Kuta Radja, and Kuta Alam.
2. Time series plotting and analyzing of rainfall, temperature, humidity, and wind speed
3. Multiple correlation testing. The hypothesis for multiple correlation test is:
   H_0: R = I (there is no correlation between rainfall, temperature, humidity and wind speed)
   H_1: R ≠ I (there is a correlation between rainfall, temperature, humidity and wind speed)
4. Stationarity test for rainfall data, temperature, humidity and wind speed. The stationarity test consists of:
   a. Stationarity of variance test using Lambda Box-Cox. If nonstationary then do the transformation
   b. Stationarity of mean test using an Augmented Dickey-Fuller (ADF) test. If nonstationary then differencing. The ADF test hypothesis is: 
      H_0: \rho = 1 (nonstationary data)
      H_1: \rho ≠ 1 (stationary data)
5. Modeling using VARIMA method. This stage consists of:
   a. Determine the model order using the Extended Cross-Correlation Matrix (ECCM). The hypothesis for this test is:
      H_0: \rho_{j+1} (W_{ij}^{(v)}) = ... = \rho (W_{ij}^{(j)}) = 0 (Order model is nonsignificant)
      H_1: \rho (W_{ij}^{(v)}) ≠ 0 for v > j (Order model significant)
   The best models are models with the smallest Akaike's Information Criterion (AIC) and Bayesian Information Criteria (BIC) values.
   b. Make a parameter estimate to know the significance of a parameter in a model
   c. Perform a diagnostic model of white noise test. The following is a hypothesis for the test:
The 9th AIC 2019 on Sciences & Engineering (9thAIC-SE) 
IOP Publishing 
IOP Conf. Series: Materials Science and Engineering 796 (2020) 012063 doi:10.1088/1757-899X/796/1/012063

H₀: Β ≡ 0 (Errors data together have white noise) 
H₁: Β ≠ 0 (Errors data together not white noise) 

d. If the model does not qualify for point c then the process is repeated from point a.

6. Calculated prediction error using Mean Absolute Percentage Error (MAPE). MAPE is used to determine the percentage of deviation of prediction results.

7. Predicted rainfall, temperature, humidity, and wind speed.

8. Concluded the analysis of the results.

3. Result and Discussion
A descriptive or clustering analysis carried out only the rainfall of four regions around Ocean Fishing Port (PPS)Lampulo, Banda Aceh with a rain post. There are Banda Raya, Lampineung, Kuta Radja, and Kuta Alam. Table 2 is the result of a summary of the rainfall in the region.

| Region | Banda Raya | Lampineung |
|--------|------------|------------|
|        | Rainfall (mm) | Rainfall (mm) |
|        | 0-20 | 21-50 | 51-100 | >100 | x | 0-20 | 20-50 | 50-100 | >100 | x |
| Jan    | 4 | 2 | 1 | - | 24 | - | 3 | 2 | - | - | 26 | - |
| Feb    | 3 | 0 | - | - | 25 | 3 | 4 | - | - | - | 24 | 3 |
| Mar    | 4 | 0 | - | - | 27 | - | 2 | - | - | - | 18 | 11 |
| Apr    | 3 | 4 | - | - | 23 | 1 | 6 | 1 | - | - | 17 | 7 |
| May    | 6 | 2 | 1 | - | 22 | - | 14 | - | - | - | 17 | - |
| June   | 2 | 3 | - | - | 25 | 1 | 1 | - | - | - | 17 | 13 |
| July   | 8 | 1 | - | - | 22 | - | 10 | - | - | - | 21 | - |
| Aug    | 7 | 1 | - | - | 23 | - | 6 | - | - | - | 25 | - |
| Sept   | 10 | 3 | - | - | 18 | - | 11 | 2 | - | - | 18 | 1 |
| Oct    | 14 | 2 | - | - | 16 | - | 13 | - | - | - | 19 | - |
| Nov    | 17 | 3 | - | - | 10 | 1 | 14 | - | - | - | 16 | 1 |
| Dec    | 8 | 4 | - | - | 19 | - | 8 | 2 | 1 | - | 20 | - |
| Total  | 86 | 25 | 2 | 0 | 254 | 6 | 92 | 7 | 1 | 0 | 238 | 36 |

Table 2. Rainfall summary in Banda Raya and Lampineung.

| Region | Kuta Radja | Kuta Alam |
|--------|------------|------------|
|        | Rainfall (mm) | Rainfall (mm) |
|        | 0-20 | 20-50 | 50-100 | >100 | x | 0-20 | 20-50 | 50-100 | >100 | x |
| Jan    | 6 | 2 | - | - | 23 | - | 6 | 1 | - | - | 24 | - |
| Feb    | 6 | - | - | - | 22 | 3 | 3 | - | - | - | 15 | 13 |
| Mar    | 4 | 1 | - | - | 26 | - | 6 | 1 | - | - | 24 | - |
| Apr    | 11 | 1 | - | - | 18 | 1 | 9 | - | - | - | 21 | 1 |
| May    | 14 | - | - | 1 | 16 | - | 11 | - | - | 1 | 19 | - |
| June   | 7 | 1 | - | - | 22 | 1 | 10 | 1 | 2 | - | 17 | 1 |
| July   | 12 | - | - | - | 19 | - | 13 | - | - | - | 18 | - |
| Aug    | 10 | 1 | - | - | 19 | 2 | 11 | - | - | - | 20 | - |
| Sept   | 13 | 1 | - | - | 16 | 1 | 7 | 1 | - | - | 22 | 1 |
| Oct    | 12 | 3 | - | - | 16 | - | 8 | 1 | - | - | 1 | 21 |
| Nov    | 13 | 3 | 1 | 1 | 12 | 1 | 13 | 3 | 1 | - | 13 | 1 |
| Dec    | 9 | 2 | 3 | - | 17 | - | 10 | 2 | 1 | - | 18 | - |
| Total  | 117 | 15 | 4 | 2 | 226 | 9 | 107 | 10 | 4 | 1 | 212 | 38 |

Table 3. Rainfall summary in Kuta Radja and Kuta Alam.
According to Table 2 and Table 3 shows that the occurrence of mild rain most occurred in Kuta Radja, the occurrence of moderate rain occurs most in Banda Raya while the rain dense category rarely occurs in all areas. No rain occurs most often in a year that is more than 200 days a year. The absence of most observations occurred in Kuta Alam for 38 days. Based on Figure 2 explaining the average rainfall in the areas of Banda Raya, Lampineung, Kuta Radja, and Kuta Alam are located below the average rainfall of Banda Aceh.

![Figure 2. Rainfall comparison in 2018.](image)

The inference analysis done is the application of VARIMA model. Here are the stages of getting a VARIMA model.

- **Multiple correlation test**
  This correlation test uses a simultaneous correlation test namely Bartlett Sphericity Test. Based on the test was obtained P-value is less than 0.000. The value is smaller than α (0.05) so H₀ is rejected. This indicates that there is an analyzed correlation between the variables so that these variables are appropriate for analysis using the VARIMA model.

- **Stationarity test**
  The stationarity test divided into two tests, variance and mean. The stationarity of variance resulted that the four variables must be transformed because it is nonstationary. The stationarity of the mean test resulted that only Y₁ and Y₄ stationary because P-value is < α (0.05). While Y₂ and Y₃ have P-value > α (0.05). So, it can be concluded that the two variables are nonstationary and should be done a differencing process.

- **VARIMA modeling**
  Before modeling, VARIMA performed vector model identification using ECCM. A significant order is an order with the P-value < α (0.05). Seven models can be tested at the modeling stage. Models with the smallest AIC and BIC values are VARIMA models (1, 1, 2) with AIC value is 8.5104 and the BIC value is 9.0581. Then it can be concluded that the best model is VARIMA (1, 2, 2).

- **Parameter estimation and diagnostic model**
  Parameter estimation is done on the best model of VARIMA (1, 2, 2). The significant parameter is a parameter that has a P-value < α (0.05) so there is a result that two significant parameters are Y₄*4 and Y₄*11. The following are VARIMA (1, 1, 2) formulation that is formed.

### Description:

| Rainfall Category | Description   |
|-------------------|--------------|
| 0 - 20 mm         | Mild         |
| 21 - 50 mm        | Moderate     |
| 51 - 100 mm       | Dense        |
| > 100 mm          | Very dense   |
| x                 | Absence of observation |
| -                 | No rain      |

\[ \hat{Y}_{1,t}^* = 0.5977Y_{1,t-1}^* - 4.497Y_{2,t-1}^* + 0.1310Y_{3,t-1}^* + 0.6328Y_{4,t-1}^* - 0.0054a_{1,t-1} - 1.7746a_{2,t-1} \\
+ 0.2704a_{3,t-1} - 0.1576a_{4,t-1} - 0.0328a_{1,t-2} - 2.265a_{2,t-2} - 1.1540a_{3,t-2} + 0.2503a_{4,t-2} \]

\[ \hat{Y}_{2,t}^* = -0.0054Y_{1,t-1}^* - 0.563Y_{2,t-1}^* - 0.1600Y_{3,t-1}^* + 0.0086Y_{4,t-1}^* - 0.0080a_{1,t-1} + 0.0469a_{2,t-1} \\
- 0.1513a_{3,t-1} - 0.00426a_{4,t-1} - 0.0012a_{1,t-2} + 0.443a_{2,t-2} + 0.1014a_{3,t-2} + 0.0505a_{4,t-2} \]

\[ \hat{Y}_{3,t}^* = -0.0678Y_{1,t-1}^* + 1.750Y_{2,t-1}^* + 0.1000Y_{3,t-1}^* + 0.1091Y_{4,t-1}^* - 0.0053a_{1,t-1} + 1.5861a_{2,t-1} \\
+ 0.6953a_{3,t-1} + 0.1731a_{4,t-1} - 0.0328a_{1,t-2} + 1.442a_{2,t-2} - 0.0348a_{3,t-2} - 0.4284a_{4,t-2} \]

\[ \hat{Y}_{4,t}^* = -0.0046Y_{1,t-1}^* - 0.9510Y_{2,t-1}^* - 0.0672Y_{3,t-1}^* + 1.0066Y_{4,t-1}^* - 0.0047a_{1,t-1} + 0.6970a_{2,t-1} \\
- 0.0087a_{3,t-1} + 0.6175a_{4,t-1} + 0.0019a_{1,t-2} - 0.5520a_{2,t-2} + 0.0004a_{3,t-2} + 0.1241a_{4,t-2} \]

The diagnostic model obtained P-value is 0.999. The decision is $H_0$ cannot rejected, because P-value > $\alpha$ (0.05). It can be concluded that the data has no autocorrelation, and the model has qualified white noise. Thus VARIMA (1, 1, 2) models can be used to predict rainfall, temperature, humidity and wind speed.

- Measuring error prediction
  Error prediction measured using MAPE. The value of MAPE obtained are $Y_1 = 38\%$, $Y_2 = 10\%$, $Y_3 = 6.2\%$ and $Y_4 = 19\%$. The value indicates that the predicted error is relatively low.

- Predictions
  Results comparison of the predictions data and actual data (testing data) of rainfall, temperature, humidity and wind speed are shown in Figure 3.

Figure 3. Comparison of predicted data and actual data (a) rainfall, (b) temperature, (c) humidity, (d) wind speed.
Based on Figure 3 shows that the predicted data of rainfall, temperature, humidity, and wind speed are close to the actual data so that the approach with the VARIMA model is appropriate. The prediction results of rainfall, temperature, humidity and wind speed in January 2019 using the VARIMA model (1,1,2) are presented in Table 4.

Table 4. Predicted results of rainfall, temperature, humidity and wind speed in January 2019.

| Date   | Rainfall | Temperature | Humidity | Wind Speed | Date   | Rainfall | Temperature | Humidity | Wind Speed |
|--------|----------|-------------|----------|------------|--------|----------|-------------|----------|------------|
|        | mm       | ºC          | %        | km/h       |        | mm       | ºC          | %        | km/h       |
| 1/1/2019 | 5.9     | 21          | 108      | 6          | 17/1/2019 | 1.8       | 25          | 82       | 6          |
| 2/1/2019 | 65      | 28          | 83       | 6          | 18/1/2019 | 2.9       | 25          | 82       | 4          |
| 3/1/2019 | 5.9     | 35          | 56       | 4          | 19/1/2019 | 70.3      | 25          | 82       | 4          |
| 4/1/2019 | 5.9     | 20          | 79       | 6          | 20/1/2019 | 46.8      | 25          | 82       | 4          |
| 5/1/2019 | 2.9     | 20          | 86       | 3          | 21/1/2019 | 11.7      | 25          | 82       | 4          |
| 6/1/2019 | 5.9     | 23          | 106      | 4          | 22/1/2019 | 5.8       | 25          | 82       | 4          |
| 7/1/2019 | 100.2   | 31          | 95       | 5          | 23/1/2019 | 5.8       | 25          | 82       | 4          |
| 8/1/2019 | 2.9     | 25          | 67       | 7          | 24/1/2019 | 5.8       | 25          | 82       | 4          |
| 9/1/2019 | 5.9     | 33          | 90       | 5          | 25/1/2019 | 5.8       | 25          | 81       | 4          |
| 10/1/2019 | 4.9    | 25          | 80       | 7          | 26/1/2019 | 5.8       | 24          | 81       | 4          |
| 11/1/2019 | 29.4   | 24          | 80       | 2          | 27/1/2019 | 5.8       | 24          | 81       | 4          |
| 12/1/2019 | 64.7   | 25          | 82       | 5          | 28/1/2019 | 5.8       | 27          | 81       | 4          |
| 13/1/2019 | 82.3   | 25          | 82       | 4          | 29/1/2019 | 5.8       | 27          | 81       | 4          |
| 14/1/2019 | 35.2   | 25          | 82       | 4          | 30/1/2019 | 5.8       | 27          | 81       | 4          |
| 15/1/2019 | 5.9    | 25          | 82       | 6          | 31/1/2019 | 5.8       | 24          | 66       | 4          |
| 16/1/2019 | 5.9    | 25          | 82       | 5          |        |          |            |          |            |

4. Conclusion

Based on the research that has been done can be concluded the occurrence of mild rain category occurred in the Kuta Radja, the most moderate category occurrence of rain occurred in Banda Raya while the rainy category frequently occurs in all areas. No rain occurs most often in a year that is more than 200 days a year. The absence of most observations occurred in Kuta Alam for 38 days. Besides, the average rainfall in the area around PPS Lampulo is below the average rainfall of Banda Aceh. The best model for rainfall predictions, temperatures, humidity and wind speeds is VARIMA (1, 1, 2) with the AIC value equal to 8.5104 and BIC value equal to 9.058. The results of rainfall prediction, temperature, humidity, and wind speed are close to the actual data, so it can be concluded that the VARIMA model is appropriate and the predicted error is relatively low and.

Information on the results of rainfall prediction rainfall that will be received by fishermen will be more accurate so that the number of marine accident events caused by extreme weather be decreased.

Acknowledgments

The author's acknowledgements to Kemristekdikti for providing funds for the implementation of this research, BMKG meteorology station class 1 Sultan Iskandar Muda at Blang Bintang, Aceh Besar and BMKG climatology station at Indrapuri, Aceh Besar for permission to use this research data.

References

[1] Swarinoto YS and Sugiyono 2011 *Utilization of Air Temperature and Air Humidity in Regression Equations for Prediction Simulation of Total Rainy Monthly in Bandar Lampung. BMKG Bandar Lampung*
[2] Tjasyono B 2004 *Climatology* (Bandung: ITB)

[3] Rusyana A, Ferdhiana R and Putri ME 2019 *Application of ARCH model on nutmeg price forecasting in South Aceh District* IOP Conf. Ser. Mater. Sci. Eng. **523** 012003 https://iopscience.iop.org/article/10.1088/1757-899X/523/1/012003/pdf

[4] Rusyana A, Nurhasanah, Marzuki and Flancia M 2016 *SARIMA model for forecasting foreign tourists at the Kualanamu International Airport* In: The 12th International Conference Mathematics, Statistics, and their Applications (ICMSA). (Banda Aceh, Indonesia: IEEE) p. 153–158 https://ieeexplore.ieee.org/document/7954329/

[5] Enders W 1995 *Applied Econometric Time* (New York: Series Willey and Sons, Inc.)

[6] Nahrowi ND 2006 *A Popular and Practical Approach to Econometrics for Economic and Financial Analysis* (Jakarta: Lembaga)

[7] Wei WWS 2006 *Time Series Analysis Univariate and Multivariate Methods* 2nd ed. (Boston: Pearson Education, Inc.)