Development of ARIMA technique in determining the ocean climate prediction skills for pre-service teacher

R Rosmiati\(^{1,2,*}\), S Liliasari\(^3\), B Tjasyono\(^4\) and T R Ramalis\(^5\)

\(^1\)Program Studi Pendidikan IPA, Sekolah Pascasarjana, Universitas Pendidikan Indonesia, Jl. Dr. Setia Budhi No. 229, Bandung 40154, Indonesia
\(^2\)Departemen Pendidikan Fisika, STKIP BIMA, Jl. Piere Tendean Kelurahan Mande, Kota Bima 84111, Indonesia
\(^3\)Departemen Pendidikan Kimia, Universitas Pendidikan Indonesia, Jl. Dr. Setia Budhi No. 229, Bandung 40154, Indonesia
\(^4\)Departemen Meteorologi, Institut Teknologi Bandung, Jl.Ganesha No. 10, Bandung 40132, Indonesia
\(^5\)Departemen Pendidikan Fisika, Universitas Pendidikan Indonesia, Jl. Dr. Setia Budhi No. 229, Bandung 40154, Indonesia

*rosmiati.stkipbima@gmail.com

Abstract. The lack of community understanding of the predictions of the ocean climate is key to addressing the impacts and crisis of the ocean climate. The aims of this study is to explain the stages of ocean climate prediction skills using the ARIMA technique for pre service teacher students as beginner learners and provide solutions to build a conscious and responsive attitude to the ocean climate. The data used is Sea Surface Temperature (SST) Niǹo 3.4, analysis the stages of the Autoregressive Integrated Moving Average (ARIMA) model used MINITAB 16.0 and Microsoft Excel that has been developed then applied by 3 respondents. Based on the results of the study, more detailed stages of the arima model and n 3 respondents both got a correlation value of 0.96 for Nino SST 3.4 original value and model value, correlation 0.92 for 1 year validation and correlation 0.99 ENSO prediction for 1 year (12 months) ahead. Thus the development of the ARIMA model stage is very effective to be used to predict climate such as ENSO events and other ocean climate phenomena.

1. Introduction

The Indonesian maritime continent consists of five major islands (Papua, Kalimantan, Sumatra, Sulawesi and Java) which are flanked by the continents of Asia and the continent of Australia and the Pacific Ocean and Indian Ocean [1]. Indonesia as a maritime continent that has a sea area of around 70% so that it is rich in sea salt aerosols which play a role in the formation of cloud drops. Indonesia is very instrumental in the formation of clouds and convective rain because Indonesia is a tropical equatorial monsoon region with the most active convection compared to the equatorial regions of America and Africa [2-4]. The Indonesian maritime continent is the main heat source of Earth's atmosphere which is characterized by a high amount of rainfall, this is an important lesson to understand the science of the climate system.
The livelihoods of the Indonesian people are very dependent on climate conditions that occur because the majority of the Indonesian population work as farmers and fishermen. In addition, most residential settlements are located in coastal areas so that they are affected by the phenomenon of ocean climate such as ENSO, IOD and Monsoon. Climate phenomena that occur every year really need an understanding of the long-term response for the community to disaster risk due to the ocean climate [5,6]. Climate prediction is a key factor in anticipating the impact of disasters due to the ocean climate [7,8]. Ocean climate prediction is needed to provide early information in planning the lives of Indonesian people in the future due to uncertainty in future climate conditions, especially in areas that are often affected by climate anomalies.

So far, climate prediction skills have only been studied by certain groups, such as ocean engineering, oceanography, statistics and meteorology, have not been taught and provided among educators such as prospective teachers. The lack of community understanding of the predictions of the ocean climate is key to addressing the impacts and crisis of the ocean climate [9-12]. This condition is strongly influenced by the role of prospective teachers as agents of change optimally in building community knowledge. Some developed countries ocean climate prediction skills explicitly become learning objectives and contained in the curriculum and are familiar knowledge [13,14] and have even become a subject of interest by many people with different educational background. Ocean climate predictions that have been developed in various countries have been proven to be successful in improving the competency of prospective teacher or community students in accordance with the objectives of the ocean climate prediction program.

One new approach that is widely used for prediction is the Autoregressive Integrated Moving Average (ARIMA) is a model that completely ignores independent variables in making prediction. ARIMA is very well used for short-term prediction. ARIMA is said to be a complex model, because this model is a combination of AR and MA, this model can also be used for seasonal time series patterns (not seasonal) simultaneously [15,16]. The ARIMA method has advantages over other methods, namely the Box Jenkins method which is arranged logically and is statistically accurate. This method incorporates a lot of information from historical data, and results in an increase in accuracy of prediction and at the same time maintaining a minimum number of parameters and using an iterative approach that indicates the possibility of a useful model [17-19]. Pre-service teacher as beginner learners about predictive skills using the ARIMA model require authentic explanations in understanding the stages of prediction skills. Prediction methods such as the ARIMA model must meet the requirements so that they can be relied on and used in general. It is hoped that Pre-service teacher in Indonesia who will become agents of change can play a role in building community knowledge through predictive skills in the ocean climate.

This study explains the application of the ARIMA method using MINITAB software in stages with Nino 3.4 data. The important question in this study is how to prepare prospective teacher students on the predictive skills of the ocean climate. The purpose of this study is to explain the stages of ocean climate prediction for Pre-service teacher as beginner learners and provide solutions to build a conscious and responsive attitude to the ocean climate. This analysis includes: (i) stages of the Box-Jenkins method which has been developed. (ii) the ability of students to implement the stages of the Arima model that has been developed. To collect data is predict ENSO regional SST Nino 3.4 for the next 12 months using the ARIMA Method tools software MINITAB and Microsoft Excel.

2. Methods

2.1. Types and data sources
The data index from January 1982 to September 2018 (monthly data) in the form of Nino 3.4 SST (Sea Surface Temperature) is obtained from http://www.cpc.ncep.noaa.gov/data/indices/sstoi.indices.
2.2. Data analysis technique
ARIMA methods with the help of statistical software namely MINITAB 16.0 and Microsoft Excel. The steps for applying the ARIMA method in a row are:

- Data Stationary Check
- Model identification Through ACF and PACF plots
- Determination of parameters p, d and q
- Determining the equation of the ARIMA model
- Validation Prediction
- Prediction by using Microsoft Excel

3. Results and discussion

3.1. The ARIMA method stage uses MINITAB

3.1.1. Data stationary examination
- Time series plot. Click Start - time series - time series plot, Simple – OK.

![Time series plot of nino 3.4 anomaly.](image1)

**Figure 1.** Time series plot of nino 3.4 anomaly.

- Determining ACF and PACF. Click Start - time series - Autocorrelation & Partial Autocorrelations.

![ACF and PACF Nino 3.4 Anomaly before differences](image2)

**Figure 2.** ACF and PACF Nino 3.4 Anomaly before differences
3.1.2. Model identification through ACF and PACF plots. What is noticed is ACF and PACF after Differences.

* The blue line down touches the red line on the fifth lag both on ACF and PACF.

**Table 1.** Number of differences and position of the blue line.

| Number of Differences | Position of the Blue Line |
|-----------------------|---------------------------|
|                       | ACF | PACF |
| 1                     | 5   | 5    |

Because the position of the blue line in the 5th lag on ACF and PACF, the ARIMA value entered starts from number 0 for the parameter \( p \) or AR, the value 1 constant in the middle is the number of differences equal to 1 and numbers 1 - 5 are \( q \) or MA. Can be seen consecutively in table 2.

3.1.3. Determination of Parameters \( p, d \) and \( q \). Perform ARIMA Modelling. Click the Start - Time series – ARIMA. Enter the Value in the non-seasonal column. \( AR = p, d = 1 \) (number difference), MA = \( q \).

**Table 2.** Sum of Squares to Error (SSE) and Mean Squared Error (MSE) ARIMA model.

| Arima Model | SSE     | MSE     |
|-------------|---------|---------|
| 0, 1, 1     | 30,7141 | 0.0721  |
| 1, 1, 1     | 28,7598 | 0.0677  |
| 2, 1, 1     | 28,7278 | 0.0678  |
| 3, 1, 1     | -       | -       |
| 4, 1, 1     | -       | -       |
| 5, 1, 1     | -       | -       |
| 0, 1, 2     | 29,2405 | 0.0688  |
| 1, 1, 2     | 28,7169 | 0.0677  |
| 2, 1, 2     | -       | -       |
| 3, 1, 2     | 27,4616 | 0.0651  |
| 4, 1, 2     | 27,4633 | 0.0652  |
| 5, 1, 2     | 28,2009 | 0.0671  |
| 0, 1, 3     | 28,8525 | 0.0680  |
| 1, 1, 3     | -       | -       |
| 2, 1, 3     | -       | -       |
| 3, 1, 3     | 28,2491 | 0.0671  |
| 4, 1, 3     | 28,3234 | 0.0674  |
Table 2. Cont

| No | AR Value | MA Value | Constant  |
|----|----------|----------|-----------|
| 5  | 1, 3     | 27, 8627 | 0, 0665   |
| 0  | 1, 4     | 28, 7423 | 0, 0679   |
| 1  | 1, 4     | -        | -         |
| 2  | 1, 4     | -        | -         |
| 3  | 1, 4     | -        | -         |
| 4  | 1, 4     | 26, 7709 | 0, 0639   |
| 5  | 1, 4     | 27, 2024 | 0, 0651   |
| 0  | 1, 5     | 28, 6346 | 0, 0679   |
| 1  | 1, 5     | 27, 2267 | 0, 0647   |
| 2  | 1, 5     | -        | -         |
| 3  | 1, 5     | -        | -         |
| 4  | 1, 5     | 25, 9174 | 0, 0644   |
| 5  | 1, 5     | -        | -         |

* ARIMA model (4.1.4) because the SSE and MSE values are the smallest

3.1.4. Determination of the equation of the ARIMA model. The general ARIMA formula is:

$$Z_t = (1 + \theta_1)Z_{t-1} + (\theta_2 - \theta_1)Z_{t-2} + \ldots + (\theta_{p-1} - \theta_{p-2})Z_{t-p} + \alpha_t + \phi_1\alpha_{t-1} + \ldots + \phi_q\alpha_{t-q}$$

The ARIMA parameters (4, 1, 4) in table 3.

Table 3. AR and MA values of ARIMA models (4, 1, 4).

| No | AR Value | MA Value | Constant  |
|----|----------|----------|-----------|
| 1  | 0, 5136  | 0, 1738  | -0, 0000724 |
| 2  | 0, 5205  | 0, 5255  |           |
| 3  | 0, 0104  | 0, 2650  |           |
| 4  | -0, 1985 | 0, 0346  |           |

$$Z_t = ((1.5136 Zt-1) + ((0.5205-0.5136) Zt-2) + ((0.0104-0.5205) Zt-3) + (- 0.2098 Zt-4) + (- 0.1985 Zt-5) + (- 0.0000724) - (0.1738 at-1) - (0.5255 at-2) - (0.265 at-3) - (0.0346 at-4)) in the MS EXCEL$$

Enter the ARIMA formula on Microsoft EXCEL can be seen in Figure 4.

Figure 4. The method of entering the ARIMA formula in Ms Excel.
3.1.5. Validation prediction.

- White noise test. Click Graph - Probability Plot - single – OK.

- Data validation and the last 1 year model. In this validation test validation tests are also carried out for the next few months with the aim of being able to find out the accuracy of the models for various prediction ranges.
3.1.6. Prediction. Using Microsoft Excel - Drag down the Nino 3.4 anomaly data and the model until the value is error. The results data are predictions for the next 1 year, this study uses the next 9 months.

Table 4. Predictions using ARIMA 9 Months ahead with a correlation of 0.98.

| Forecast period | Correlation (R) |
|-----------------|-----------------|
| 9 Month         | 0.98            |
It is seen that the results of the original Nino 3.4 SST data correlation and the ARIMA model data get a value of 0.96 for the same three respondents, meaning that the resulting value is very good because it is close to 1. As revealed by Dunning D and Helzer [20]; Mussavi [21] good correlation is a correlation that approaches 1. After correlation the original data and the next important stage model are validation, the three respondents get the same correlation results 3, 6, 9 and 12, respectively get a value of 0.68; 0.48; 0.81; 0, 92. Because the initial value of the correlation is the same magnitude. In some studies, the less time, the greater the correlation value, but in some other studies, on the contrary, namely the amount of time, the greater the correlation value, because the more time used, the more numbers that enter and affect the correlation value. After the original value correlation and validation, the final stage is a prediction, the prediction results from 3 respondents can be seen in Figure 10.

4. Conclusion

Based on the discussion on the development of arima techniques for determining the ocean climate prediction skills for pre-service teachers, it can be summarized as follows: at the stages of the ARIMA model using MINITAB and Microsoft EXCEL software in great detail, the purpose is to facilitate pre-service teacher as a beginner learner in predicting climate, especially climate. From the 3 respondents who implemented the development of the ARIMA method it was obtained 1. Correlation of original Nino SST 3.4 data with the ARIMA model data averaging 0.96 correlation with the validation correlation value approaching 1 is at 12 months, predictions for the next 1 year or 12 months obtained a correlation value of 0.99. So the development of the ARIMA model stage is very effective to be used to predict climate such as ENSO events and other ocean climate phenomena.
Acknowledgments
Thus the author can convey about the results of research contained in this article. On this occasion the
authors would like to express the appreciation and gratitude to BUDI-DN and LPDP as sponsors and
all the parties that has big impacts to the success and completion of this research. The author realizes
that this article is far from perfect, therefore constructive criticism and suggestions are desirable for
the perfection of this article. Hopefully this research will be useful for the readers.

References
[1] Yamanaka M D 1998 Climatology of Indonesian Maritime Continent Kyoto Univ.
[2] Tjasyono B 2017 Sains Kebumian & Antariksa (Surabaya: UNESA Press)
[3] Yamanaka M D, Hashiguchi H, Mori S, Wu P-M, Syamsudin F, Manik T, Hamada J I,
Yamamoto M K, Kawashima M and Fujiyoshi Y 2008 HARIMAU radar-profiler network
over the Indonesian maritime continent: A GEOSS early achievement for hydrological cycle
and disaster prevention J. Disaster Res 3 78–88
[4] Worku L Y, Mekonnen A and Schreek III C J 2019 Diurnal cycle of rainfall and convection
over the Maritime Continent using TRMM and ISCCP Int. J. Climatol. 39 5191–200
[5] Rosmiati R and Satriawan M 2019 The ocean climate phenomenon: the challenges of earth
physics lectures in Indonesia Journal of Physics: Conference Series IOP Publishing vol
1157 p 32038
[6] Rojas-Downing M M, Nejadhashemi A P, Harrigan T and Woznicki S A 2017 Climate change
and livestock: Impacts, adaptation, and mitigation Clim. Risk Manag. 16 145–63
[7] Kim H-M, Webster P J and Curry J A 2012 Seasonal prediction skill of ECMWF System 4 and
NCEP CFSv2 retrospective forecast for the Northern Hemisphere Winter Clim. Dyn. 39
2957–73
[8] Weisheimer A and Palmer T N 2014 On the reliability of seasonal climate forecasts J. R. Soc.
Interface 11 20131162
[9] Berry B J L, Bihari J and Elliott E 2016 The limits of knowledge and the climate change debate
Cato J. 36 589
[10] Reyes-García V, Fernández-Llamazares Á, Guèze M, García A, Mallo M, Vila-Gómez M and
Vilaseca M 2016 Local indicators of climate change: The potential contribution of local
knowledge to climate research Wiley Interdiscip. Rev. Clim. Chang. 7 109–24
[11] Lineman M, Do Y, Kim J Y and Joo G-J 2015 Talking about climate change and global
warming PLoS One 10 e0138996
[12] Wei M, Qiao F and Deng J 2015 A quantitative definition of global warming hiatus and 50-year
prediction of global-mean surface temperature J. Atmos. Sci. 72 3281–9
[13] Xu Z, Tang Y, Connor T, Li D, Li Y and Liu J 2017 Climate variability and trends at a national
scale Sci. Rep. 7 1–10
[14] McCaffrey M S and Buhr S M 2008 Clarifying climate confusion: Addressing systemic holes,
cognitive gaps, and misconceptions through climate literacy Phys. Geogr. 29 512–28
[15] Alsharif M H, Younes M K and Kim J 2019 Time series ARIMA model for prediction of daily
and monthly average global solar radiation: The case study of Seoul, South Korea Symmetry
(Basel). 11 240
[16] Shen S and Shen Y 2016 ARIMA model in the application of Shanghai and Shenzhen stock
index Appl. Math. 7 171–6
[17] Murat M, Malinowska I, Gos M and Krzyszczak J 2018 Forecasting daily meteorological time
series using ARIMA and regression models Int. agrophysics 32
[18] El-Mallah E S and Elsharkawy S G 2016 Time-series modeling and short term prediction of
annual temperature trend on Coast Libya using the box-Jenkins ARIMA Model Adv. Res. 1–11
[19] Svetunkov I and Boylan J E 2020 State-space ARIMA for supply-chain forecasting Int. J. Prod.
Res. 58 818–27
[20] Dunning D and Helzer E G 2014 Beyond the correlation coefficient in studies of self-assessment accuracy: Commentary on Zell & Krizan (2014) Perspect. Psychol. Sci. 9 126–30

[21] Mussavi M, Niknafs P and Bijari B 2013 Determining the correlation and accuracy of three methods of measuring neonatal bilirubin concentration Iran. J. Pediatr. 23 333