Journal of Engineering Science and Technology
Special Issue on INCITEST2019, February (2020) 77 - 87

STUDY OF MULTIPLE SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE SUBSEQUENCES AGGREGATE LONG-TERM TIME SERIES MODEL FOR FLOOD PREDICTION BASED ON THE SEASONAL RAINFALL DATA IN INDONESIA

SRI SUPATMI[1], IRFAN DWIGUNA SUMITRA[2]

[1]Department of Computer Engineering, Universitas Komputer Indonesia, Bandung, Indonesia
[2]Department of Information System, Universitas Komputer Indonesia, Bandung, Indonesia
*Corresponding Author: srisupatmi@email.unikom.ac.id

Abstract

This paper proposed a new prediction method for flood vulnerability during the wet season in Indonesia. This method called Multiple Autoregressive Integrated Moving Average Subsequences Aggregate Long-Term Time Series model or MSARIMASA model. The long-term time series data and divided into data actual-sampling for experiment and data fits-sampling for evaluating. MSARIMASA model built from the aggregate data actual-sampling and subdivided into training long-term time-series data as well as authorizing long-term time-series data to reduce the effect in the rainfall prediction based on time series employing the aggregate method by its period. A fixed number of subsequence patterns generated and fitting well using SARIMA models. It is also verified by the mean absolute percentage error (MAPE) and mean forecast error (MFE) to identify the best-fitted model of MSARIMASA. The step of predicting future rainfall in the aggregate grouped employing the MSARIMASA models needs to determine the subsequence with the best fitted SARIMA model. The disaggregates process to spread this group value to the future rainfall using a ratio from the previous rainfall data. The MSARIMASA model was compared using MAPE and MFE with the SARIMA model and the ARIMA model. The results for the proposed model indicated better MAPE and MFE than the SARIMA and ARIMA models.

Keywords: Aggregation step, ARIMA, A subsequence of long-term time series, Forecast, Multiple SARIMA.

1. Introduction

Indonesia is a country located on the equator. Therefore, Indonesia has a tropical climate and two seasons, dry and rainy season. Some areas in Indonesia experienced a flood during the rainy season. Rainfall is an important component in the climate system because it plays a critical role in agricultural and flood in Indonesia. Besides, rainfall is a major area of concern within the field of forecasting data. The data forecasting of rainfall may become a crucial factor that affects the process of decision making in many fields [1]. The sectors include transportation such as air transportation, agricultural sectors, as well as fishery sectors [1, 2].

The accuracy of the rainfall forecasting in early gives the stakeholders such as farmers, governments more time to determine the decision regarding the water supply for their field for the farmer and prevention warning of flood for the government, in this case, is Bureau of Meteorology, Climatology, and Geophysics in Indonesia. During the wet season, in Bandung Regency.

During the wet season, in Bandung Regency often occurred by a flood. Bandung Regency is a part of West Java province in Indonesia. The flood [3], which occurred in Bandung Regency, was related to the existence of several changes in climatological elements. These elements include the influence of La Nina in the Ocean Pacific, an increase in the surface temperature in the sea level in the Java Sea, wind encounters, and bending of the wind direction over the Bandung area, also the tropical cyclones in the Indian Ocean [4].

Many researchers studied the rainfall forecast employing different models and applied them in a certain area. Nita et al studied rainfall forecasting using the Moving Average algorithm and obtain the accuracy MAPE 15.66% [5]. Wang et al studied the ARIMA model to predict the precipitation in Lanzhou, China, and accounting data, both the internal and inter-monthly variations. The result obtained the forecast accuracy of 21% of error [6]. The ARIMA model [7-10] also used by many researchers to analyse the hydrological process unconsidering the influence of seasonal factor [11-13]. The other study of the ARIMA model did the stationary test for monthly rainfall forecasting and the effects from the inter-monthly variant within a year [14-17]. In addition, some researchers studied seasonal ARIMA (SARIMA) model to forecast the monthly rainfall [18-23].

This study aims to forecast the future seasonal rainfall employing the proposed model of time series, namely multiple seasonal auto repression moving average subsequence aggregate time series or MSARIMASA model. The proposed model applies in Bandung Regency, West Java Province in Indonesia, and the rainfall time series in January 2009 to December 2013 dataset comes from [24]. The forecasted rainfall can be useful to predict the next flood event in Bandung Regency. The effective flood forecasting models could be helpful for early warning and disaster prevention. The part of this research described Non-seasonal ARIMA models that mainly written as ARIMA (p, d, q), representing the time series z_t denoted as:

$$\varphi_p(B)\nabla^d z_t = \theta_q(B)\alpha_t \tag{1}$$

where p is the autoregressive non-seasonal order, d is the differencing non-seasonal degree, q is the moving-average non-seasonal order, and α_t denotes as the random error terms at time t.

The lag operator B defined by

$Bz_t=z_{t-1}$ (2)

The differencing ∇ denoted as

$\nabla=(1-B)$ (3)

Equation (3) computed as

$\nabla z_t = z_t - z_{t-1}$ (1)

The differencing order of ARIMA model denoted as

$\nabla^d=(1-B)^d$ (5)

$\varphi_p(B)$ and $\theta_q(B)$ are AR and MA operators and respectively. $\varphi_p(B)$ *and* $\theta_q(B)$ defined in Eqs. (6) and (7), respectively.

$$\varphi_p(B) = 1 - \sum_{i=1}^{p} \varphi_i B^i \quad (6)$$

$$\theta_q(B) = 1 - \sum_{i=1}^{q} \theta_i B_i \quad (7)$$

The autoregressive coefficients denote as $\varphi_1, \varphi_2, \dots, \varphi_p$, and the moving-average coefficient denote as $\theta_1, \theta_2, \dots, \theta_q$.

Seasonal ARIMA [1-3] models denoted as SARIMA$(p,d,q)x(P,D,Q)_s$ representing the time series Y_t denoted as

$$\Psi_P(B^s)\varphi_p(B)\nabla_s^D \nabla^d Y_t = \Theta_Q(B^s)\theta_q(B)\alpha_t \quad (8)$$

$\Psi_P(B^s)=1-\Psi_1(B^1)-\dots-\Psi_P(B^{sP})$ (9)

$\Theta_Q(B^s)=1-\Theta_1(B^1)-\dots-\Theta_Q(B^s)$ (10)

$\nabla_s^D=(1-B^s)^D$ (11)

where P is the autoregressive seasonal order, D is the differencing seasonal degree, Q is the moving-average seasonal order, and s is for seasonal time series (i.e., monthly time series $S = 12$).

The various parameters of ARIMA and SARIMA models can be identified employing a methodology of the Box-Jenkins model [2, 4, 5].

Steps of the Box-Jenkins model:

i. Verification of the stationery of the time series data by the Box-Jenkins model to define the proper differencing
ii. Auto-correlation function (ACF) and partial autocorrelation function (PACF) used for the parameters appraisal of AR and MA.
iii. The maximum likelihood procedure forecast all of the ARIMA and SARIMA model coefficients [6] .

In this study, the SARIMA model is simulated by Minitab programming, and it is used to determine the basic method for the proposed model (MSARIMASA model).

2. Method

This section describes the algorithm for generating the proposed model and predicting future values. The time-series data consisted of actual-sampling and fits-sampling. The actual sampling data was employed to create the proposed model and predict the aggregate future values. The fits-sample data used to evaluate the accuracy using MFE and MAPE. In this study, 95% of the time series data were employed to actual-sample data, and the remainder was applied to fits-sample data. The structure of the algorithm consisted of two parts. The first part was to create the proposed model and the second one was to predict future values.

The actual-sample denotes as

$$Z_i = Z_1, Z_2, Z_3, \ldots, Z_n \quad (12)$$

The fits-sample was written as

$$Z_{(n+i)} = Z_{(n+1)}, Z_{(n+2)}, Z_{(n+3)}, \ldots, Z_m \quad (13)$$

The forecast for the future values presents as

$$\check{Z}_{(n+i)} = \check{Z}_{(n+1)}, \check{Z}_{(n+2)}, \check{Z}_{(n+3)}, \ldots, \check{Z}_m \quad (14)$$

2.1. Proposed model

The proposed model consisted of four steps aggregate the actual-sample, aggregate actual-sample data were subdivided into training and validating time series data, generating the arithmetic subsequences, and the last was fit the SARIMA model based on Box-Jenkins method as well as obtaining the MFE and MAPE for authorizing portion. The detail explanation is shown below.

The first step of the proposed model was to aggregate the actual-sample, and it denotes as:

$$X_i = X_1, X_2, X_3, \ldots, X_R \quad (15)$$

The aggregation time series data denotes in Eq. (16)

$$X_a = \sum_{i=1}^{b} (a+1)b + i \quad (16)$$

The period of the time series data is b and the number of aggregate groups on the actual-sample data is R. For $a = 1, 2, 3, \ldots, R$ and $n = Rb$

The second step is to subdivided the aggregate actual-sample data into training and validating time series data. Training time series data comes from 95% of the actual-sample data and the remainder employed to the validating time series.

The training time series data determined as:

$$X_1, X_2, X_3, \ldots, X_{\bar{R}} \quad (17)$$

where $\bar{R}$ is the number of training time series data and must be less than R. The validating time series data can be written as:

$$X_{(\bar{R}+1)}, X_{(\bar{R}+2)}, X_{(\bar{R}+3)}, \ldots, X_R \quad (18)$$

The third steps described as follow: The starting of various initial values are generating with the arithmetic subsequence. The beginning is determining the maximum common difference as c and we obtain the number of the arithmetic subsequence patterns as

$$A_S = dx \sum_{f=1}^{c} f = \frac{c(c+1)}{2} \quad (19)$$

where A_S is the arithmetic subsequence. The aggregate actual-sample data for each subsequence is generated in Eqs. (20-22).

For the common difference $f = 1$, we can write

〖X_(*i*)〗 _(*i*=1,2,3,…,*r*)^(*r*≤*R*) (20)

For the common difference, $f = 2$, it generate the arithmetic subsequence in two times and denotes as:

*X*_(2*i*-1)〗 _(*i*=1,2,3,…,*r*)^(2*r*-1≤*R*)= 〖*X*_2*i*〗 _(*i*=1,2,3,…,*r*)^(2*r*≤*R*) (21)

For the maximum common difference $f = c$, it generates the arithmetic subsequence in c time and can be written in the equation as below:

*X*_(*ci*-*j*) 〗 _(*i*=1,2,3,…,*r*)^(*cr*-*j*≤*R*)= 〖*X*_*ci* 〗 _(*i*=1,2,3,…,*r*)^(*cr*≤*R*) (22)

For $j = 0, 1, 2,\ldots, (c - 1)$, respectively

The subsequence of the training time series data generating by determining the start of the index at $i = 1$ to $R^{-}$and the index for the validating time series data start at $i = R^{-}+ 1$ to R.

The last step is to fit the SARIMA models based on the Box-Jenkins method which comes from the training time series data. The validating portion calculates using MFE and MAPE. The subsequence of the aggregate actual-sample denotes as *O*_((*h*,*f*)), where h is the starting index and f is a common difference. Let $SARIMA_{(h,f)}$, 〖*MFE*〗_((*h*,*f*)), and $MAPE_{hf}$ be the best fitted SARIMA model of $O_{(h,f)}$ and MFE as well as MAPE of *O*_((*h*,*f*)) respectively. The future values forecasting by the 〖*SARIMA*〗_((*h*,*f*)) and explained in the next section.

2.2. Forecasting proposed model

This section describes how to forecast future value employing the proposed model. First, we have to determine the aggregate for the forecasted values. The formula of the aggregate forecasted values shown in Eq. (23) below.

$$\bar{X}_{R+1}, \bar{X}_{R+2},, \ldots, \bar{X}_{\bar{\bar{R}}} \quad (2)$$

$\bar{\bar{R}}$ is the number of the aggregate group on time series data.

From Eq. (16) where $n = R_b$, we can determine the forecasting for the future values by using the equation:

$$\check{Z}_{Rb+1}, \check{Z}_{Rb+2}, \check{Z}_{Rb+3}, \ldots, \check{Z}_{(R+1)b}, \check{Z}_{(R+1)b+1}, \check{Z}_{(R+1)b+2} \ \ldots, \check{Z}_{m} \quad (3)$$

where m is a number of forecasts for the future values and denotes as $m{=}R\bar{\bar{b}}$

To obtain the forecast for future values requires three major steps: identify the aggregate forecasted, select the subsequence, and disaggregate method to obtain the future values of rainfall.

By determining in each order from the Eq. (23) to identify the aggregate forecast values of each future value, we can get

$$\bar{X}_{R+1} = \check{Z}_{Rb+1}, \check{Z}_{Rb+2}, \check{Z}_{Rb+3}, \ldots, \check{Z}_{(R+1)b} \quad (4)$$

$$\check{Z}_{m} = \check{Z}_{(\bar{\bar{R}}+1)b+1}, \check{Z}_{(\bar{\bar{R}}+1)b+2}, \ldots, \check{Z}_{\bar{\bar{R}}b} \quad (5)$$

By substituting Eqs. (25) and (26) into Eq. (24), we obtain a new formula for aggregate forecasted values in Eq. (27)

$$\left(\check{Z}_{Rb+1}, \check{Z}_{Rb+2}, \check{Z}_{Rb+3}, \dots, \check{Z}_{(R+1)b}\right), \dots, (\check{Z}_{(\bar{\bar{R}}+1)b+1}, \check{Z}_{(\bar{\bar{R}}+1)b+2}, \dots, \check{Z}_{\bar{\bar{R}}b} \quad (6)$$

The subsequence of the aggregate actual-sample data $O_{(h,f)}$ are identified by $\bar{X}_{R+1}, \bar{X}_{R+2}, \bar{X}_{R+3}, \dots, \bar{X}_{\bar{\bar{R}}}$. Considering to Eqs. (20-22), we can determine the subsequence of the aggregate actual sample data for each value of starting index h and common difference *f*.

For h = 1 and f = 1

$$O_{(1,1)} = \bar{X}_{i+R} \; {}^{r+R\leq\bar{\bar{R}}}_{i=1,2,3,\dots,r} \quad (7)$$

For $h = m$ and $f = m$

$$O_{(m,m)} = \bar{X}_{mi+R} \; {}^{mr+R\leq\bar{\bar{R}}}_{i=1,2,3,\dots,r} \quad (8)$$

The next steps for forecasting the proposed method determine the subsequence which includes the aggregate forecasted value index. This step also extracts the corresponding SARIMA models, MFEs, and MAPEs. The minimum MFEs and MAPEs among all SARIMA models of this index are to be selected to predict $\bar{X}_{R+1}, \bar{X}_{R+2}, , \dots, \bar{X}_{\bar{\bar{R}}}$.

The last steps were the disaggregating procedure to obtain the predicted of future values. To obtain disaggregate predicted values$\check{Z}_{n+1}, \check{Z}_{n+2}, \check{Z}_{n+3}, \dots, \check{Z}_{m}$ become the aggregate forecasted values $\bar{X}_{R+1}, \bar{X}_{R+2}, , \dots, \bar{X}_{\bar{\bar{R}}}$ doing by the following procedure.

The disaggregate start from $\bar{X}_e$ where $e \in \left\{(R+1), (R+2), (R+3), \dots, \bar{\bar{R}}\right\}$ and for the forecasted by some of the subsequence $O_{(h,f)}$ can be expressed by

$$O_{(h,f)} = Z_{(\alpha_i - \beta)} \; {}^{\alpha r-\beta\leq R}_{i=1,2,3,\dots,r} \quad (9)$$

The value of the number aggregate α has membership from $0, 1, 2, \dots, (r-1)$ and β have membership from $1, 2, 3, \dots, r$.

The disaggregate of $O_{(h,f)}$ for $h = 1$, $h = r$ and $f = 1, 2, 3, \dots, r$ are denoted as the below formulas:

$$O_{(1,1)} = X_{(\alpha-\beta)} = Z_{(\alpha-\beta-1)b+1}, Z_{(\alpha-\beta-1)b+2}, Z_{(\alpha-\beta-1)b+3}, \dots, Z_{(\alpha-\beta)b} \quad (10)$$

$$O_{(r,r)} = X_{(i\alpha-\beta)} = Z_{(\alpha i-\beta-1)b+1}, Z_{(\alpha i-\beta-1)b+2}, Z_{(\alpha i-\beta-1)b+3}, \dots, Z_{(i\alpha-\beta)b} \quad (11)$$

The disaggregate of $O_{(h,f)}$ also can be written in Eq. (33)

$$O_{(h,f)} = O_{(1,1)}, O_{(1,2)}, \dots, O_{(r,r)} \quad (12)$$

From the last step, we can obtain the average value of order u and denotes as ∇_u

$$\nabla_u = \frac{\sum_{i=1}^{\left\lfloor\frac{R+\beta}{\alpha}\right\rfloor} Z_{(i\alpha-\beta-1)b+u}}{\left\lfloor\frac{R+\beta}{\alpha}\right\rfloor} \quad (13)$$

where $u = 1, 2, 3, \dots, b$; $\alpha \in \{0, 1, 2, 3, \dots, (r-1)\}$; $\beta \in \{1, 2, 3, \dots, r\}$.

The disaggregate procedure can be activated by employing the weighted average ($\bar{\nabla}_u$) of the value in u order.

$$\bar{\nabla}_u = \frac{\nabla_u}{\sum_{i=1}^{b} \nabla_i}, \quad u = 1, 2, 3, \dots, b \quad (14)$$

Accordingly, $\bar{X}_e$ is able to be disaggregated

$$\bar{X}_e = \check{Z}_{(e-1)b+1}, \check{Z}_{(e-1)b+2}, \check{Z}_{(e-1)b+3}, \dots, \check{Z}_{(e-1)b+b} \tag{15}$$

$$\check{Z}_{(e-1)b+u} = \bar{\nabla}_u \bar{X}_e \tag{16}$$

where $u = 1, 2, 3, \dots, b;\ e \in \{(R+1), (R+2), (R+3), \dots, \bar{\bar{R}}\}$.

3. Performance Model Measurement

In each of the forthcoming definitions, Z_t is the fits-sample values and $\check{\boldsymbol{Z}}_{\boldsymbol{t}}$is the predicted values. The differencing of the real value and the predicted value denoted as d_t and the formula of d_t written as bellow:

$$\boldsymbol{d_t = Z_t - \check{Z}_t} \tag{17}$$

Performance model measurement for this study consists of two models: Mean Forecast Error (MFE) [5] and Mean Absolute Percentage Error (MAPE) adapted from references [3] and [5]. For the formulas described below:

$$MFE = \frac{1}{n}\sum_{t=1}^{n} d_t \tag{18}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{d_t}{z_t}\right| x\ 100\% \tag{19}$$

where n is the number of both real value and predicted value.

4. Results and Discussion

The proposed model was tested employing data from the Statistics of Jawa Barat [24]. The data consisted of six-time series data with the monthly and yearly data period. The implementation of the proposed model employing minitab2018 programming. The time-series data started in January 2009 and ends on December 2013. The number of experiments was 72. The time series plot of rainfall in the Bureau of Meteorological, Climatology, and Geophysics for Bandung Regency shown in Fig. 1.

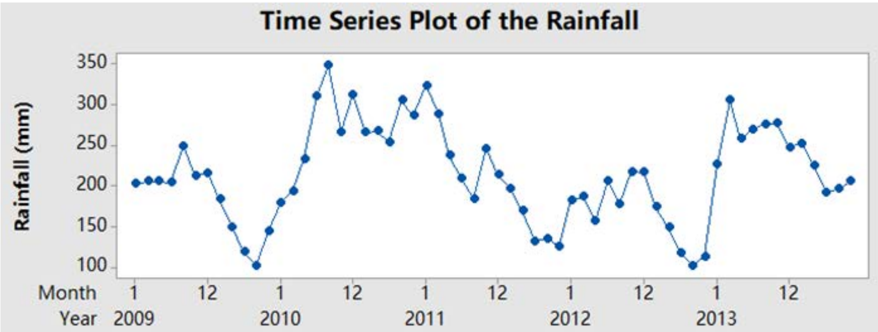

Fig. 1. Rainfall time series from January 2009 to December 2013.

The period of the aggregate SARIMA arranged to 12 in monthly of a year, the highest of common difference c arranged in 1 to 5. The performance of the proposed algorithm in MFEs and MAPE employing the parameters of SARIMA shown in Table 1. The subsequence column in Table 1, according to the original subsequence in X_i in Eq. (20) referred to the SARIMA parameter of the best SARIMA model to adjust the correlate with the subsequence and third, fourth columns showed the MFE and MAPE of the best SARIMA model on the data validation.

The best-fitted model for the proposed model, which obtained two lowest MFE and MAPE came from the subsequence fourth and eighth respectively (see Fig. 2.).

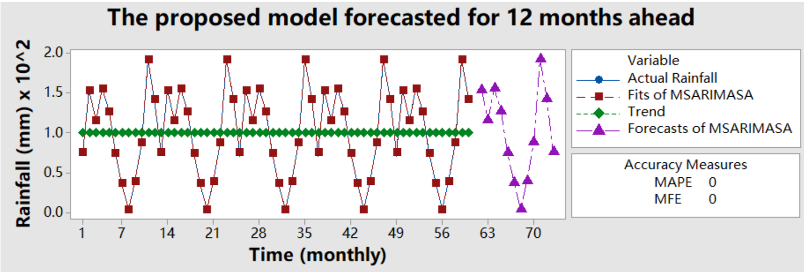

Fig. 2. The proposed model (MSARIMASA model) forecasted 12 months ahead.

Figures 2 and 3 show the proposed model with the fits-sample data. It showed the proposed model could forecast the next rainfall in 12 months ahead and 60 months ahead with the best subsequence aggregate of the SARIMA model in subsequence 8th. The forecasted data from the proposed model similar to the fits-sample data and the MFE, MAPE of the proposed model are 1.17 and 0.67 %, respectively (shown in Table 2). Figure 4 shows the model to compare rainfall time series data in Bandung Regency. The MFE and MAPE for the proposed model, ARIMA and SARIMA shown in Table 2.

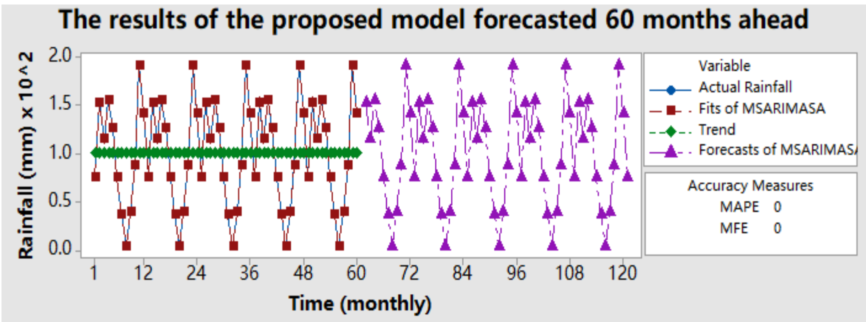

Fig. 3. The proposed model forecasted for 60 months ahead.

The proposed model (MSARIMASA model) provided reliable rainfall predicting comparing with the other algorithms presented in previous work employing another dataset. In literature [11], 1-month rainfall forecasting obtained the accuracy of MAPE reached 15.66%, and in reference [12] the MAPE reached 24.803% for univariate time series model of precipitation forecast. The MAPE obtained with the proposed model (MSARIMASA model) varies from 1.67 to 0.00% for forecasting of 60 months ahead. This is in a good agreement with literature [25].

Table 1. Performance of the proposed model (MSARIMASA model).

No.	**Subsequence (c *to* 5)**	**SARIMA orders $(p,d,q)x(P,D,Q)_s$**	**MFE**	**MAPE (%)**
1	$X_1, X_2, X_3, X_4, X_5, \dots, for\ c = 1$	$(0,0,0)x(1,0,0)_{12}$	1.677	0.83
2	$X_1, X_3, X_5, X_7, X_9, \dots, for\ c = 2$	$(0,0,0)x(1,0,1)_{12}$	1.425	0.55
3	$X_2, X_4, X_6, X_8, X_{10}, \dots, for\ c = 2$	$(1,0,0)x(1,0,0)_{12}$	1.897	0.92
4	$\boldsymbol{X_1, X_4, X_7, X_{10}, X_{13}, \dots, for\ c = 3}$	$\boldsymbol{(1,0,0)x(1,0,1)_{12}}$	**0.23**	**0.16**
5	$X_2, X_5, X_8, X_{11}, X_{14}, \dots, for\ c = 3$	$(0,0,0)x(1,1,0)_{12}$	1.777	0.87
6	$X_3, X_6, X_9, X_{12}, X_{15}, \dots, for\ c = 3$	$(0,0,0)x(1,1,1)_{12}$	1.967	1.25
7	$X_1, X_5, X_9, X_{13}, X_{17}, \dots, for\ c = 4$	$(0,0,0)x(1,0,0)_{12}$	1.997	1.43
8	$\boldsymbol{X_2, X_6, X_{10}, X_{14}, X_{18}\ \dots, for\ c = 4}$	$\boldsymbol{(0,0,0)x(1,0,1)_{12}}$	**0.00**	**0.00**
9	$X_3, X_7, X_{11}, X_{15}, X_{19}, \dots, for\ c = 4$	$(1,0,0)x(1,0,0)_{12}$	2.432	1.67
10	$X_4, X_8, X_{12}, X_{16}, X_{20}, \dots, for\ c = 4$	$(1,0,0)x(1,0,1)_{12}$	0.237	0.18
11	$X_1, X_6, X_{11}, X_{16}, X_{21}, \dots, for\ c = 5$	$(0,0,0)x(1,1,0)_{12}$	0.536	0.20
12	$X_2, X_7, X_{12}, X_{17}, X_{22}, \dots, for\ c = 5$	$(0,0,0)x(1,1,1)_{12}$	0. 679	0.17
13	$X_3, X_8, X_{13}, X_{18}, X_{23}, \dots, for\ c = 5$	$(0,0,0)x(1,0,0)_{12}$	0.723	1.02
14	$\boldsymbol{X_4, X_9, X_{14}, X_{19}, X_{24}, \dots, for\ c = 5}$	$\boldsymbol{(0,0,0)x(1,0,1)_{12}}$	**0.198**	**0.13**
15	$X_5, X_{10}, X_{15}, X_{20}, X_{25}, \dots, for\ c = 5$	$(1,0,0)x(1,0,0)_{12}$	1.786	0.79

Table 2. Comparative results of rainfall prediction.

	Proposed Model	**SARIMA model**	**ARIMA model**
Parameters	Table 1	$p = 1; d = 1; q = 0$ $P = 1; D = 1; Q = 1$ $seasonal = 12$	$p = 1; d = 1; q = 0$ $seasona = 12$
Average MFE	1.17	3.23	5.89
Average MAPE (%)	0.67	2.15	5.60

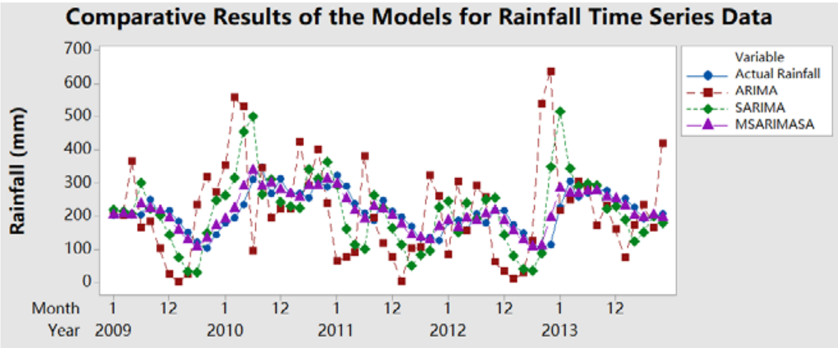

Fig. 4. Comparative performance results with 60 time series monthly data (parameters proposed a model in subsequence 4th in Table 1, ARIMA and SARIMA in Table 2 respectively).

5. Conclusion

The finding of the research indicated that the proposed model (MSARIMASA) was the best model to predict the rainfall with average MFE and MAPE respectively 1.17 and 0.67%. However, these results might not apply to all types of data; different parameters of each model also affected the value of the forecasted data. Besides, the results of the rainfall forecast for the next flood forecasted can be useful for early warning and disaster prevention.

Acknowledgments

We thank Prof. Hou Rongtao for his support in this research. In additionally, thank you to the Rector of Universitas Komputer Indonesia Assoc. Prof. Dr. Ir. H. Eddy Soeryanto Soegoto, M.T. for the support and facilities for the accomplishment of this research.

References

1. Wahyuni, I.; and Mahmudy, W.F. (2017). Rainfall prediction in Tengger, Indonesia using hybrid Tsukamoto FIS and genetic algorithm method. *Journal of ICT Research and Applications*, 11(1), 38-54.
2. Kashid, S.S.; and Maity, R. (2012). Prediction of monthly rainfall on homogeneous monsoon regions of India based on large scale circulation patterns using Genetic Programming. *Journal of Hydrology*, 454, 26-41.
3. Liu, Y.; Ngan, H.; and Ni, L. M. (2007). Power-aware node deployment in wireless sensor networks. *International Journal of Distributed Sensor Networks*, 3(2), 225-241.
4. "Ilmu Geografi.com," (2015). Retrieved on November 1, 2018, from https://ilmugeografi.com/bencana-alam/penyebab-banjir-di-bandung.
5. Fowdur, T.P.; Beeharry, Y.; Hurbungs, V.; Bassoo, V.; Ramnarain-Seetohul, V.; and Lun, E.C.M. (2018). Performance analysis and implementation of an adaptive real-time weather forecasting system. *Internet of Things*, 3, 12-33.
6. Wang, H.R.; Wang, C.; Lin, X.; and Kang, J. (2014). An improved ARIMA model for precipitation simulations. *Nonlinear Processes in Geophysics*, 21(6), 1159-1168.
7. Horne, G.P.; Yang, M.C.; and Ware, W.B. (1982). Time series analysis for single-subject designs. *Psychological Bulletin*, 91(1), 178
8. Zhang, G.P.; (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing, 50*, 159-175.
9. Cochrane, J.H.; (2005). Time series for macroeconomics and finance. Manuscript, University of Chicago, Retrieved on January 13, 2019, from http://www.smartquant.com/references/TimeSeries/ts13.pdf.
10. Zhang, G.P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159-175.
11. Jin, J.L.; Ding, J.; and Wei, Y.M. (1999). Threshold autoregressive model based on genetic algorithm and its application to forecasting the shallow groundwater level, *Journal of Hydraulic Engineering*, 27, 51-55.
12. Toth, E.; Montanari, A.; and Brath, A. (1999). Real-time flood forecasting via combined use of conceptual and stochastic models. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 24(7), 793-798.
13. Niu, X.F.; Edmiston, H.L.; and Bailey, G.O. (1998). Time series models for salinity and other environmental factors in the Apalachicola estuarine system. *Estuarine, Coastal and Shelf Science*, 46(4), 549-563.
14. Wen, Q.I.; and Mianping, Z.H.E.N.G. (2006). Winters and ARIMA model analyses of the lake level of salt Lake Zabuye Tibetan Plateau. *Journal of Lake Sciences*, (1), 3.

15. Ahmad, S.; Khan, I.H.; and Parida, B.P. (2001). Performance of stochastic approaches for forecasting river water quality. *Water Research*, 35(18), 4261-4266.
16. Lehmann, A.; and Rode, M. (2001). Long-term behaviour and cross-correlation water quality analysis of the river Elbe, Germany. *Water Research*, 35(9), 2153-2160.
17. Mohamed, T.M.; and Ibrahim, A. (2016). Time series analysis of nyala rainfall using ARIMA method, *Journal of Engineering and Computer Science (JECS)*, 17(1), 5-11.
18. Kaushik, I.; and Singh, S.M. (2008). Seasonal ARIMA model for forecasting of monthly rainfall and temperature. *Journal of Environmental Research and Development*, 3(2), 506-514.
19. Nirmala, M.; and Sundaram, S.M. (2010). A seasonal ARIMA model for forecasting monthly rainfall in Tamilnadu. *National Journal on Advances in Building Sciences and Mechanics*, 1(2), 43-47.
20. Etuk, E.H.; and Mohamed, T.M. (2014). Time series analysis of monthly rainfall data for the Gadaref rainfall station, Sudan, by SARIMA methods. *International Journal of Scientific Research in Knowledge*, *2*(7), 320.
21. Graham, A.; and Mishra, E.P. (2017). Time series analysis model to forecast rainfall for Allahabad region. *Journal of Pharmacognosy and Phytochemistry*, 6(5), 1418-1421.
22. Osarumwense, O.I. (2013). Applicability of box Jenkins SARIMA model in rainfall forecasting: A case study of Port-Harcourt South South Nigeria. *Canadian J. Computing in Mathematics, Natural Sciences, Engineering Medicine*, 4(1), 1-4.
23. Kibunja, H.W.; Kihoro, J.M.; Orwa, G.O.; and Yodan, W.O. (2014). Forecasting precipitation using SARIMA model: A case study of Mt. Kenya region. *Mathematical Theory and Modelling*, 4, 50-59.
24. Statistics of Jawa Barat. (2018). Temperature, rainfall, and rainy day in Bandung, 2009-2014, Retrieved on May 8, 2018, from https://jabar.bps.go.id/statictable/2015/04/02/40/temperatur--curah-hujan--dan-hari-hujan-di-bandung--2009-2014.html.
25. Abdullah, A.G.; Sopian, W.W.; Arasid, W.; Nandiyanto, A.B.D.; Danuwijaya, A.A.; and Abdullah, C.U. (2018). Short-term peak load forecasting using PSO-ANN methods: The case of Indonesia. *Journal of Engineering, Science, and Technology*, 13(8), 2395-2404.