A study of periodic and stochastic modeling of monthly rainfall from Purajaya station

A. Zakaria

***Abstract*— A study was done using monthly rainfall data with long data 25 years (1977-2001) of Purajaya station. The goal of this research is to study periodic and stochastic models of data series of the monthly cumulative rainfall. Based on daily rainfall data, monthly cumulative rainfall data series was calculated. The series of rainfall is assumed to be free of trend. Periodicities of rainfall data were presented using 125 periodic components. Stochastic series of rainfall data are assumed as residues between rainfall data with periodic rainfall model. Stochastic components were calculated using the approach of autoregressive model. Stochastic Model presented in this research is using the fourth orders autoregressive model. Validation between data and the model is done by calculating the correlation coefficient. For this study, the correlation coefficient between the data and the model of the cumulative monthly rainfall is 0.9992. From the results of this study can be inferred that the model of the monthly rainfall from Purajaya station gives highly accurate approach.**

***Index Terms*—monthly cumulative rainfall, periodic and stochastic, FFT, autoregressive model.**

I. Introduction

Irrigation water requirements for designed, detailed information regarding rainfall in conjunction with time is required. To prove a series of record-keeping of rain is very difficult, so sometimes to predicted, or add the data recording of rainfall, a synthetic rainfall data is required. Various methods have been developed by researchers in the field of engineering and science to prove this information. Most widely used methods that now there is a method of deterministic and stochastic methods [2][4][6][7][8][9]. When the preceding method does not prove the influence of random input data parameter, the method was last applied the concept of probability, where the characteristics of rainfall based on the time neglected, and this calculation is only benefit when data are processed fairly long. But the method is not much used anymore because this method is not sufficient to answer the problems that exist.

In nature, the rainfall is having periodic and stochastic parts, because rainfall is influenced by climate parameters such as temperature, wind direction, humidity and so on, which also is periodic and stochastic. These parameters are transferred into the periodic and stochastic rainfall components. The rainfall can be counted to determine both, periodic and stochastic components. Determines all factors are known and it is assumed that the rainfall is as a function of the periodic and stochastic variation of climate. Next, analysis of periodic and stochastic rainfall time series will produce a model that can be used to calculate the periodic and stochastic and can also be used to predict the monthly rainfall variation on that to come.

For a few years ago, some researchers have been conducting rainfall research in order to study periodic and stochastic modeling of time series data, such as in [3][4][6] [7][8] and [9].

The goal of this research is to produce a model of periodic and stochastic monthly rainfall from the station Purajaya. With periodic and stochastic models, synthetic monthly rainfall can be simulated more accurate than the simulation only applying periodic model. This Model can be used to produce synthetic rainfall data accurately and realistic for planning of irrigation or water resources project for the Purajaya area.

II. Research Methodology

A. Area Study

The study of the research is an area of Purajaya. This area is one of the towns in the Western Province of Lampung, Lampung, Indonesia.

B. Rainfall data collection

Daily rainfall Data of Purajaya is taken from the Indonesian agency for meteorology and Geophysics of Lampung Province. The Data used for the study of rainfall with a period of 25 years (1977-2001). Mathematical procedures taken to formulate a model that predicted will be discussed in the next. The purpose of the principle of this analysis is to determine a realistic model for calculating and outlines the rainfall time series into various components of frequency, amplitude, and phase of rainfall.

In General, the time series data can be parsed into a deterministic component, which can be formulated into a value in the form of a component which is the exact solution and components that are stochastic, which this value is always presented as a function that consists of several functions of the data series of the time. Time series data, presented as an $X(t)$ model that consists of several functions [4][6][7][8][9] as follows ,

$$X(t) = T(t) + P(t) + S(t) \tag{1}$$

A. Zakaria, he is now with the Department of Civil Engineering, Lampung University, Bandar Lampung, Lampung 35145 Indonesia (e-mail: ahmadzakaria@unila.ac.id).

Where, *T* (*t*) is the component of trend, t = 1, 2, 3, ... , *N*. *P*(*t*) is the periodic components, and *S*(*t*) is the stochastic components.

In this research, the rainfall data is assumed to have no trend. So this equation can be presented as follows,

$$X(t) \approx P(t) + S(t) \tag{2}$$

(2) is an equation to obtain the representative periodic and stochastic models of monthly cumulative rainfall series.

C. Spectral Method

Spectral method is one of several transformation methods is generally used in many applications. This method can be presented as Fourier transformation method [5][7][8][9] as follows,

$$P(f_m) = \frac{\Delta t}{2\sqrt{\pi}} \sum_{n=-N/2}^{n=N/2} P(t_n).e^{\frac{-2.\pi.i}{M}.m.n} \tag{3}$$

Where $P(t_n)$ is a series of rainfall in the time domain and $P(f_m)$ is a series of rainfall in the frequency domain. Based on the frequency of rainfall resulting from (3), amplitudes as a function of the frequency of rainfall can be generated. Maximum amplitude can be determined from the amplitudes of the result as significant amplitude. The frequencies of rainfall from significant amplitudes used to simulate synthetic rainfall are assumed as dominant rainfall frequencies. A series of dominant frequencies resulting in this study, in the form of angular frequencies were used to determine periodic rainfall components.

D. Periodic Component

Periodic component of *P* (*t*) with regard to a displacement that oscillate for a specific interval [2]. The existence of *P* (*t*) is identified by using the Fourier transformation method. The part that shows the presence of an oscillating *P* (*t*), using a period *P*, some periods with peak amplitudes can be estimated by using Fourier transformation. Frequency obtained from the spectral method is clearly shows the existence of variations that are periodic. The periodic component of $P(f_m)$ can also be written in the form of angular frequency. Then an equation can be expressed in the form of the Fourier series [3][4][6][7][8][9] as follows,

$$\hat{P}(t) = S_o + \sum_{r=1}^{r=k} A_r \sin(\omega_r .t) + \sum_{r=1}^{r=k} B_r \cos(\omega_r .t) \tag{4}$$

The (4) can be organized into the following equation,

$$\hat{P}(t) = \sum_{r=1}^{r=k+1} A_r \sin(\omega_r .t) + \sum_{r=1}^{r=k} B_r \cos(\omega_r .t) \tag{5}$$

Where $P(t)$ is the periodic component, $\hat{P}(t)$ is the periodic component of the model, P_o = Ak + 1 is the mean of the rainfall series ω_r is the angular frequency, A_r and B_r are the coefficients of Fourier components.

E. Stochastic Components

Stochastic components formed by a random value that can not be calculated precisely. Stochastic models, in the form of autoregressive model or Marcov scheme can be written as the following mathematical functions [2][4][6][9],

$$S(t) = \varepsilon + \sum_{r=1}^{p} b_r .S(t-r) \tag{6}$$

The (6) can be decomposed into,

$$S(t) = \varepsilon + b_1.S(t-1) + b_2.S(t-2) + ... + b_p.S(t-p) \tag{7}$$

Where, b_r is the parameter of the autoregressive model. The ε is the constant of random numbers. r = 1, 2, 3, 4…, p is the order of stochastic components.

To get the parameter of the autoregressive model and the constant of random number, least squares method can be applied.

F. Least Squares Methods

Analysis of periodic components

In the approximation method, as a solution approach from periodic components of *P* (*t*), and to determine the function of (5), a procedure that used to get periodic component model is least squares method. From (5) can be calculated the sum squares of error between the periodic data and model [3] [7] [8] [9] as follows,

$$\text{Sum Squares of Error} = J = \sum_{t=1}^{t=m} \left\{P(t) - \hat{P}(t)\right\}^2 \tag{8}$$

Here *J* is the sum squares of error. It depends on the value of A_r and B_r. J coefficient can only be minimum if it satisfies an equation as follows,

$$\frac{\partial J}{\partial A_r} = \frac{\partial J}{\partial B_r} = 0 \text{ with } r = 1, 2, 3, 4, 5, ..., k \quad (9)$$

By using the least squares method, the A_r and B_r coefficients of Fourier components can be obtained. Based on the Fourier coefficients can be generated the following equations,

a. mean of rainfall,

$$P_o = A_{k+1} \quad (10)$$

b. amplitude of the harmonic components,

$$C_r = \sqrt{A_r^2 + B_r^2} \quad (11)$$

c. phase of harmonic components,

$$\varphi_r = \arctan\left(\frac{B_r}{A_r}\right) \quad (12)$$

The mean of rainfall, the amplitude and phase of harmonic components can be inserted into an equation as follows,

$$\hat{P}(t) = S_o + \sum_{r=1}^{r=k} C_r.Cos(\omega_t.t - \varphi_r) \quad (13)$$

(13) is a periodic model of monthly cumulative rainfall, which fetched based on the rainfall data from station Purajaya.

Analysis of stochastic components

Based on the results of the simulations obtained from periodic rainfall models, stochastic components S (t) can be generated. The stochastic components are the difference between rainfall data series with calculated rainfall series obtained from periodic model. Stochastic series as a residual rainfall series, which can be presented as follows,

$$S(t) \approx X(t) - P(t) \quad (14)$$

(14) can be solved by using the same way with the way that used to get periodic rainfall series components. Following (8), stochastic models (7) can be arranged to be as follows,

$$\text{Sum squares of Error} = J = \sum_{t=1}^{t=m} \left\{S(t) - \hat{S}(t)\right\}^2 \quad (15)$$

Where J is the sum squares of error. It depends on the ε and b_r values, where the coefficients can only be minimum if it satisfies the equation as follows,

$$\frac{\partial J}{\partial \varepsilon} = \frac{\partial J}{\partial b_r} = 0 \text{ with } r = 1, 2, 3, 4, 5, ..., p \quad (16)$$

In the next, by using (16) stochastic parameters ε and b_r of the residual rainfall data can be calculated.

III. Results and Discussion

To test the characteristics of rainfall from rainfall data, time series with long daily rainfall data of 25 years (1977-2001) of Purajaya station. Characteristics of the annual average and maximum daily rainfall series can be calculated. Fig. 1 shows the daily rainfall time series from Purajaya station.

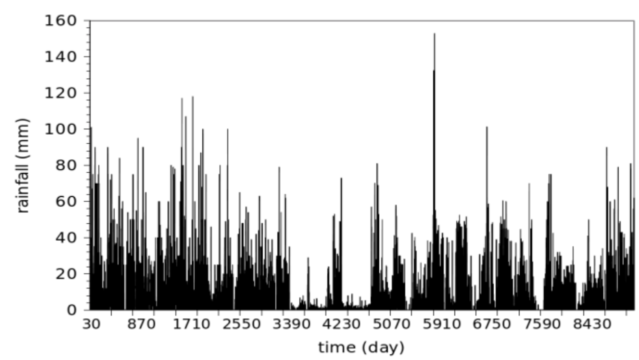

Fig. 1. Daily rainfall time series for 25 years from Purajaya station.

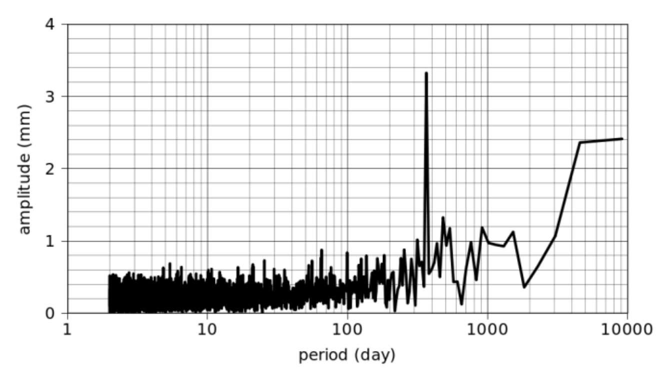

Fig. 2. Variation of daily rainfall periods for 25 years data from Purajaya station.

Daily rainfall data based on the average daily rainfall had annual varies from 2.00 mm in 1986 be 12.5 mm in 1977. The annual maximum daily rainfall varies from 35 mm in 1986 became 152.9 mm in 1992. This variation is likely caused by changes in nature due to climate change. The cumulative annual rainfall of Purajaya shows the value of a minimum of 552.5 mm in 1989 and a maximum of 4308.9 mm in 1996, with a cumulative average annual rainfall of 2553.5 mm.

Based on the daily rainfall data and by using the method of FFT (Fast Fourier Transform), a periodogram of daily rainfall time series can be generated. The transformation result of the daily rainfall data for over 25 years can be calculated by using the method of FFT such as presented in Fig. 2.

From Fig. 2 can be seen that the maximum amplitude of the periodogram is 3.3255 mm for a period of 365.2 days or one year. This shows that the annual components of daily rainfall data are very dominant in comparison with other periods. The spectrum above was presented in the rainfall amplitudes versus rainfall periods. The spectrum presented in Fig. 2 is produced using the FFT method of Matlab.

To estimate the presence of the periodic component of the monthly cumulative rainfall time series and to get dominant frequencies of the rainfall, Fourier transformation method can be applied. To produce the dominant Frequencies, the procedure applied is use an algorithm proposed [1] where the amount of the data N is analyzed as the square of 2, for example $N = 2^k$.

Based on the daily rainfall data of all 25 years from Purajaya station, monthly cumulative rainfall series is calculated. The monthly cumulative rainfall series can be seen as in Fig. 3 below,

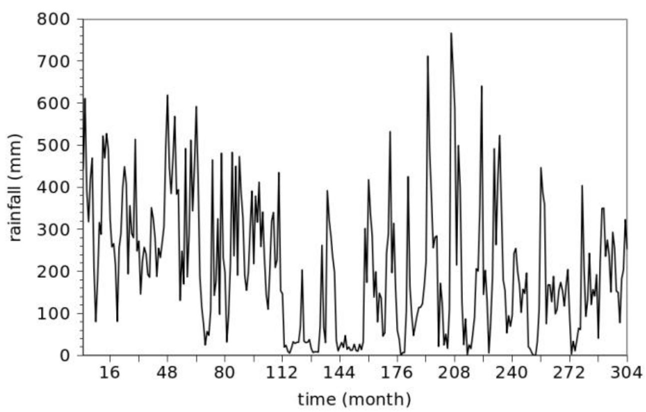

Fig. 3. Variation of monthly cumulative rainfall from Purajaya station.

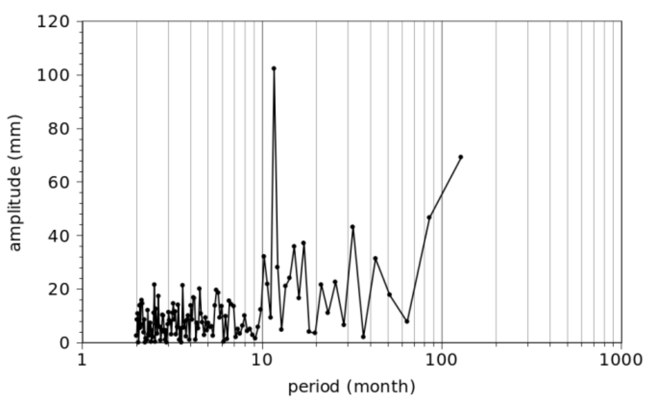

Fig.4. Variation of monthly cumulative rainfall periods.

Based on the monthly cumulative rainfall data can be calculated monthly cumulative rainfall spectrum such as presented in Fig. 4. Using the monthly cumulative rainfall data from Purajaya station, 125 dominant frequencies were estimated. By applying the 125 frequencies as in Fig. 4, then Fourier series (5) can be fitted to the monthly cumulative rainfall data in order to simulate a synthetic rainfall time series as presented in Fig. 5.

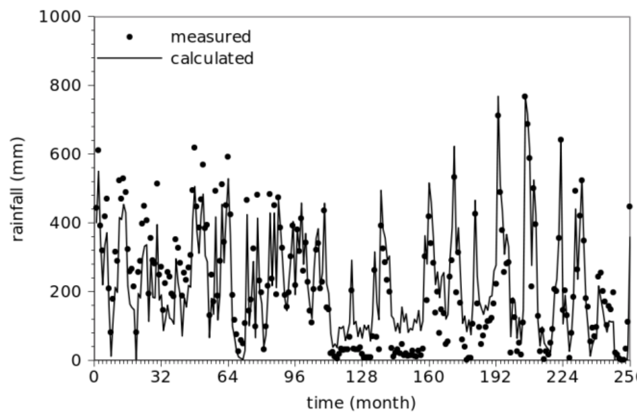

Fig. 5. Variation of monthly cumulative rainfall from station Purajaya between measured and calculated rainfalls using periodic model

The 10 periodic components ordered from the maximum amplitudes of 125 periodic components used in this research can be seen in Table 2 below,

TABLE I
10 PERIODIC COMPONENTS OF THE CUMULATIVE MONTHLY RAINFALL MODEL

r	Frequencies (°)	amplitudes (mm)	phases (°)
1	29.53	103.81	38.81
2	2.81	47.03	176.76
3	9.84	45.09	173.63
4	22.50	44.22	347.01
5	8.44	40.9	95.35
6	19.69	38.95	182.68
7	11.25	37.40	90.9
8	28.13	32.67	58.82
9	21.09	32.53	80.26
10	33.75	32.28	38.12

Based on the results obtained such as shown in Fig. 5, residual monthly rainfall between the data series and the synthetic rainfall series can be calculated by using (2). The residual series of rainfall are errors of the periodic rainfall model. The periodic errors are assumed as stochastic components of the monthly cumulative rainfall. The stochastic components of the monthly cumulative rainfall can be seen as in Fig. 6 below,

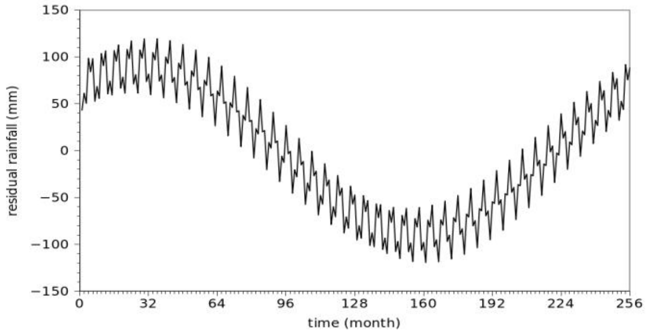

Fig. 6. Variation of stochastic component of the monthly cumulative rainfall from Purajaya station

By using the least squares method such as that presented in (15) and (16), the autoregressive model can be fitted to the stochastic components of the monthly cumulative rainfall as shown in Fig. 6. Parameters of the autoregressive model and the independent random number can be produced. For the stochastic components of the monthly cumulative rainfall from Purajaya station, the 4th order autoregressive model parameters are used to approximate the stochastic model. The autoregressive model parameters and the independent random number can be seen in Table 1 as follows,

TABLE 2
PARAMETERS OF 4TH ORDER AUTOREGRESSIVE MODEL

parameters	values
ε	0,4791
b_1	1,0252
b_2	0,0375
b_3	−1,0176
b_4	0,9555

Based on the parameters of autoregressive model and the independent random number in the Table 1, can be simulate stochastic monthly rainfall as presented in the Fig. 7 and Fig. 8 as follow,

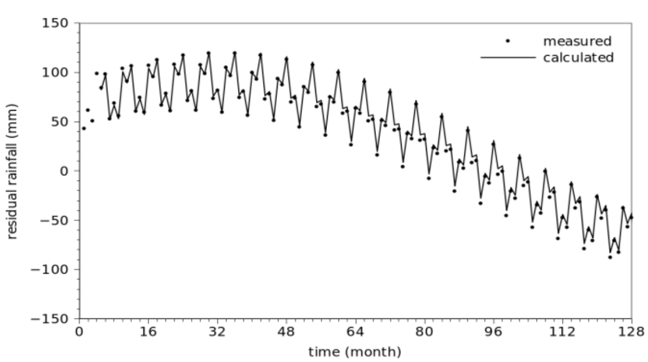

Fig. 7. Variation of measured and calculated stochastic monthly rainfalls from station Purajaya (1-128).

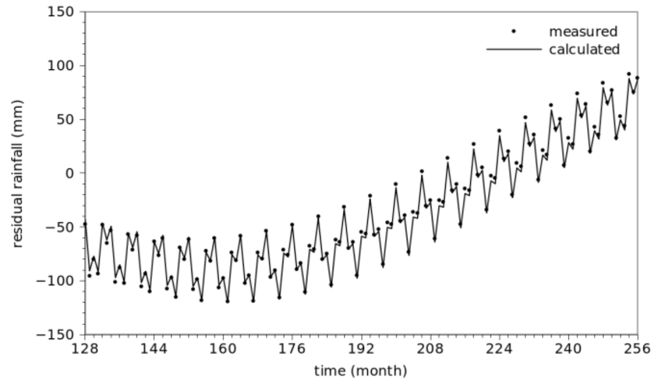

Fig. 8. Variation of measured and calculated stochastic monthly rainfalls from station Purajaya (128 – 256).

By using the periodic and stochastic model, measured and calculated cumulative monthly rainfall series from station Purajaya can be presented in the following Figures,

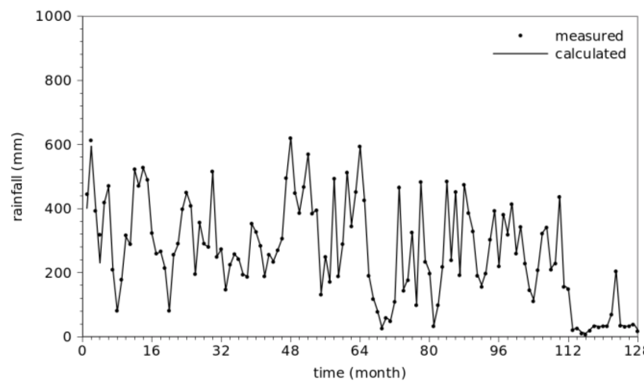

Fig. 9. Variation of measured and calculated periodic + stochastic monthly rainfalls from station Purajaya (1-128).

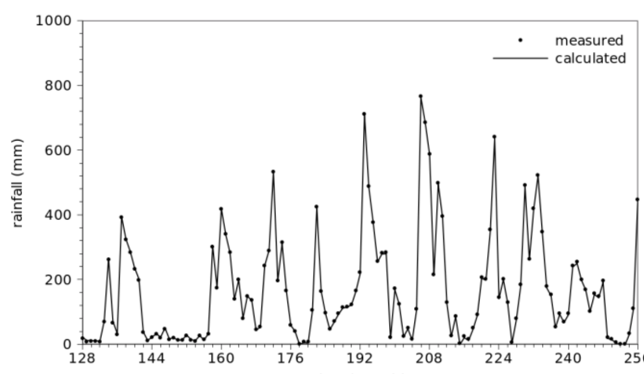

Fig. 10. Variation of measured and calculated periodic + stochastic monthly rainfall from station Purajaya (128-256).

A comparison between the measured monthly cumulative rainfall and the calculated monthly cumulative rainfall of the periodic and stochastic modeling as shown in Fig. 9 and Fig. 10 indicates that, the calculated monthly cumulative rainfall of the periodic and stochastic model gives highly accurate result.

For modeling of the periodic rainfall provides the value of correlation coefficient R is 0.9200. For modeling of the stochastic rainfall is using 4th orders autoregressive model gives the value of correlation coefficient R is 0.9945. For modeling of stochastic and periodic monthly cumulative rainfall giving the value of correlation coefficient, between the data and the model is 0.9992. Value of the coefficient correlation is almost close to 1. This shows that the model of periodic and stochastic cumulative monthly rainfall is almost close to the pattern of rainfall monthly cumulative rainfall data that are measurable. The variation of the orders for the correlation coefficient of stochastic model R(S), the correlation coefficient of periodic and stochastic model R (P+S) and the error of monthly cumulative rainfall can be seen in the Fig. 11.

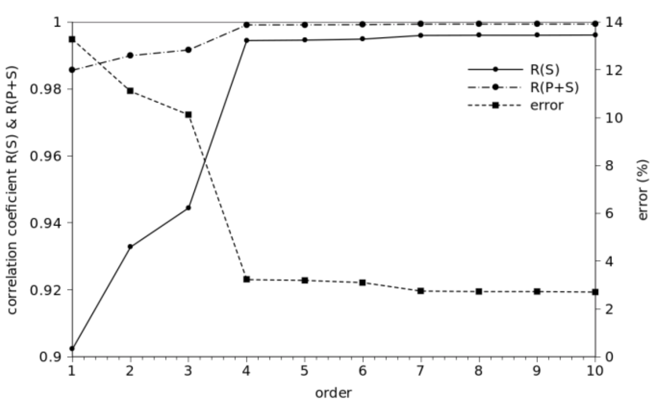

Fig. 11. Variation of the orders for correlation coefficients R(S), R(P + S), and error (%) of the stochastic and periodic models.

Based on the results presented in Fig. 11 shown that, using the 4th order autoregressive model can give better accuracy results than the 3rd order accuracy result. For the accuracy of the 5th order up to the accuracy of the 10th order, did not provide more significant results, if it is compared with the accuracy of 4th order. So in this research, stochastic model is using the 4th accuracy. The correlation coefficient (R) and error (%) for periodic model delivers each of 0.92 and 28%. For the periodic and stochastic models can provide the value of the coefficient of correlation and error each of 0.9992 and 3.23%.

Monthly cumulative rainfall modeling in this research can provide more complete result when it is compared to monthly rainfall modeling such as that done by [4] and [6], where in modeling of monthly rainfalls [4] only use a few periodic and stochastic parameters. In his work [4], he is using 6 harmonic components with the stochastic components for accuracy of 3rd order. Where for [6], in their research, they use only 3 harmonic components with the stochastic component for accuracy of 1st order. In my research, more complex solution is conducted than the previous research. In my research 125 periodic components are used but the harmonic modeling of monthly rainfall in this research still can be done quickly. It is because of by applying the method of Fast Fourier Transform (FFT), prediction of the frequency components of harmonic monthly cumulative rainfall can be generated quickly.

The characteristic of stochastic monthly cumulative rainfalls in my research can be seen as presented in Fig. 7 and Fig. 8. The stochastic component series are the difference between the monthly cumulative rainfall data with the periodic model series. From the figures are presented that the stochastic component fluctuates from - 119.5 mm up to 119.3 mm. In this research, the correlation coefficient of stochastic models with the accuracy of the 4th order is equal to 0.9945, while for the accuracy of the 1st order is equal to 0.9023. The result is better when compared with the other results presented by [6] which uses stochastic model for the accuracy of the 1st order and give the coefficient correlation for stochastic model of 0.9001.

In my results, by using the 125 periodic components and 4th order accuracy stochastic components yield the simulation model of monthly cumulative rainfall accurately, with a correlation coefficient is equal to 0.9992. The correlation coefficient presented in Fig. 11 is evidence that periodic and stochastic models (P + S) of monthly cumulative rainfall has a very good correlation and accurate results when compared with only using periodic model (P) that generates correlation coefficient of 0.9200. This result also presents much better when it is compared to the research is done by [6], where the model only using the 3 periodic components with the 1st order accuracy of stochastic component with the correlation coefficient is 0.9961.

IV. CONCLUSION

The spectrum of the monthly cumulative rainfall time series generated by using the FFT method is used to simulate the synthetic monthly cumulative rainfall. By using the least squares method, the monthly cumulative rainfall time series can be produced synthetic rainfall quickly. By using 125 periodic components and 4th order stochastic components, the monthly cumulative rainfall model or Purajaya station can be produced accurately with the correlation coefficient of 0.9992. In the future work, using the same method, other cumulative rainfall series can be used to study periodic and stochastic rainfall components, and other spectrum methods can be used to generate more accurate rainfall spectrums.

REFERENCES

[1] J. Cooley, W. James, Tukey, W. John, "An Algorithm for the machine calculation of Complex Fourier Series," Mathematics of Computation. pp. 199-215, 1965.

[2] N. T. Kottegoda, *Stochastic Water Resources Technology*, The Macmillan Press Ltd., London, 1980.

[3] A. Zakaria, "Preliminary study of tidal prediction using Least Squares Method," M.S. Thesis, Bandung Institute of Technology, Bandung, Indonesia, 1998.

[4] M. Rizalihadi, "The generation of synthetic sequences of monthly rainfall using autoregressive model," Jurnal Teknik Sipil Universitas Syah Kuala, vol. 1, no. 2, pp. 64-68, 2002.

[5] A. Zakaria, "Numerical Modelling of Wave Propagation Using Higher Order Finite Difference Formulas," Ph.D. dissertation, Curtin University of Technology, Perth, WA, 2003.

[6] S.R. Bhakar, R.V. Singh, N. Chhajed, and A.K. Bansal, "Stochastic modeling of monthly rainfall at Kota region," ARPN Journal of Engineering and Applied Sciences, vol. 1, no. 3, pp. 36-44, 2006.

[7] A. Zakaria, "The generation of synthetic sequences of monthly cumulative rainfall using FFT and least squares method," *in Proc. Seminar Hasil Penelitian & Pengabdian kepada masyarakat Universitas Lampung*, Lampung University, 2008, vol. 1, pp. 1-15.

[8] A. Zakaria, " A study periodic modeling of daily rainfall at Purajaya region," *in Proc. Seminar Nasional Sain & Teknologi III*, 18-19 October 2010, Lampung University, vol. 3, pp. 1-15.

[9] A. Zakaria, " Studi pemodelan stokastik curah hujan harian dari data curah hujan stasiun Purajaya," *in Proc. Seminar Nasional Sain Mipa dan Aplikasinya*, 8-9 December 2010, Lampung University, vol. 2, pp. 145-155.