ABUAD Journal of Engineering Research and Development (AJERD)
ISSN: 2645-2685

Volume 3, Issue 1, 41-47

Synthetic Generation of Daily Rainfall Amounts of Gombe Town

Andy Obinna IBEJE

Department of Civil Engineering, Imo State University, Owerri, Nigeria
engineeribeje@gmail.com

Corresponding Author: engineeribeje@gmail.com

Date Submitted: 08/01/2020
Date Accepted: 23/03/2020
Date Published: 30/06/2020

Abstract: *Usually the length of a rainfall record is insufficient to contain extreme values and critical sequences required for the design of components of water resources system. Stochastic rainfall model which has to do with the occurrence of wet day and depth of rainfall is developed for Gombe town located in semi-arid region of Nigeria, which depends on rainfall as the main source of irrigated agriculture but has limited rainfall record for the design of irrigation system. The first-order Markov-chain model was used to generate the sequence of rainfall occurrence using the method of transitional probability matrices, while daily rainfall amount was generated using a gamma distribution. The model parameters were estimated from 1991-2001 historical rainfall records. When the observed rainfall and the synthetically generated ones were compared, the statistical characteristics were satisfactorily preserved with a correlation coefficient of 0.99. Evidently results obtained have shown that the model could be used to generate rainfall data satisfactorily for the design of irrigation facilities in Gombe, Nigeria.*

Keywords: *Model, rainfall occurrence, distribution, generation.*

1. INTRODUCTION

The pattern of rainfall and the pattern of extreme high or low precipitation are very important for agriculture as well as economy of a country. It is well established that rainfall pattern is changing on both global and regional scales [1] due to climate change. The implications of these changes are particularly significant for Gombe where hydrological disasters like flooding have occurred severally in the past. A rainfall occurrence model is an important tool used in hydrologic planning especially in irrigated agriculture and flood control. As such, farmers are more informed as to when to expect rainfall in order to commence cultivation. In addition, most hydrologic variables such as runoff, evaporation are derivable from rainfall events, thus identifying appropriate model of daily rainfall is very crucial [2].

Estimating future occurrence and value of rainfall amount has become a huge problem in hydrologic design. The problem is that the period of available rainfall record is usually short in comparison with the period required for a reliable direct estimate. However, many hydrologic events including rainfall seem to follow to a continuous statistical distribution. It is based on this premise that such a distribution exists, that prediction of events as rainfall with limited data is possible.

The method of fragments has been used for successful generation of monthly [3, 4]. This approach has been criticized for the repetitions of the same yearly pattern when the available historical data is very small. [5] reported that this deficiency has been overcome by using synthetic fragments. However, this resulted in the problem of generating the accurate number of months with zero rainfall. [5] augured that disaggregation schemes were effective in producing monthly data although the number of parameters to be estimated with many sites is a huge problem. Simplifications to overcome this problem were proposed by [6]. According to [7], models for generating daily rainfall are well developed using the transition probability matrices which retains most of the characteristics and therefore is reported to be the best performing model.

The two-part model has been severally reported [8-10] to perform well across a range of climates at the daily level but has not been tested adequately at monthly or annual levels. The two-part model involved the use of the two-state first-order Markov chain to predict the daily rainfall occurrences (part one) and to statistically fit distribution to the rainfall amounts for wet days (part two). For instance, in the study of sequence of daily rainfall occurrence, [8] used the first-order Markov chain model in rainfall data of northern Nigeria. Also [9] successfully fitted the first-order of the Markov chain model to rainfall occurrences in Italy. Furthermore, [11] used two states first-order Markov chain to simulate daily occurrences of rainfall and the gamma distribution was used to model the distribution of rainfall amount in Bangladesh. In another study, [10] used the first-order Markov chain model to predict annual rainfall intervals using transitional probability matrices and Monte Carlo simulation. The class intervals were treated as probability states in the transition probability matrices while using the frequency distribution of the class intervals. The results revealed that the model provided forecasts of high accuracy for Owerri city in Nigeria. Based on these reviews, Markov chain can be adjudged as a good method for effective modelling rainfall occurrence. In actual sense, a Markov chain represents a system of elements making transition from one state to another over time. In the case of fitting distribution to the rainfall amounts, Gamma distribution is reported to be most appropriate to model for the synthetic generation of daily rainfall amounts [11, 10]. Hence, the objective of the study was focused on the use of the two-state first-order Markov chain to model daily rainfall occurrences and to fit the Gamma distribution to the rainfall amounts for the predicted wet days in Gombe, Nigeria.

2. MATERIAL AND METHOD

2.1 Study Area

Figure 1 shows the study area which is Gombe in Gombe state, Nigeria. The area has variable rainfall pattern with extreme cases of drought and sporadic flood. It records an average annual rainfall of 1072.6mm [12]. On daily basis, the temperature ranges between 25^{0}C and 34^{0}C [13]. Daily rainfall amount of Gombe from 1991 to 2001 were obtained from the Nigerian Meteorological Agency (NIMET) office at Gombe Airport, Gombe, to serve as primary data used in this study.

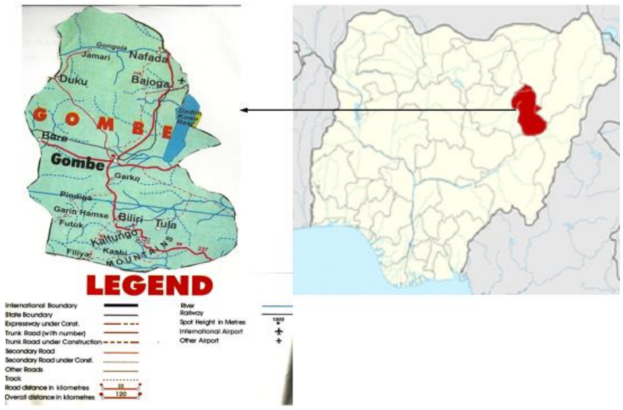

Figure 1: Location map of Gombe

2.2 Model Identification

Akaike's Information Criterion (AIC) was used to investigate if first-order Markov chains adequately fit the synthesis of daily rainfall occurrence in the present study [7]. AIC Test statistic is given by

$$AIC(m) = -2L_m + 2S^m(S-1) \qquad (1)$$

where m represents the order of the Markov chain to be tested, s the number of states, and n the sample size. Chi-square test was also used to test the goodness of fit of the rainfall amounts to gamma probability distribution [14]. This was performed by plotting cumulative probability curves and applying classical statistical inferences to the daily rainfall data. According to [14], for the boundaries (I_{i-1}, I_i), the theoretical frequency, n_i, corresponding to the observed frequency, O_i, of bin i of the histogram, the Chi-square statistic is obtained as

$$\chi^2 = \sum_{i=1}^{N} \frac{(O_i - n_i)^2}{n_i} \qquad (2)$$

The null hypothesis stating that the observed sample is drawn from the theoretical distribution (in this case gamma distribution) is accepted if $\chi^2 < \chi^2_{N-k-1,1-\alpha}$

where, , $\chi^2_{N-k-1,1-\alpha}$is the chi-square value for N-K-1 degree of freedom and α significant level.

2.3 Formulation of Transition Matrices

If $P(X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = P(X_{n+1} = x_{n+1} | X_n = x_n)$ (3)

where $x_0, x_1, \dots, x_{n+1} \in \{0, \quad 1\}$ and let

$$X_n = \begin{cases} 0 \; if \; the \; n^{th} day \; is \; dry \\ 1 \; if \; the \; n^{th} \; day \; is \; wet \end{cases}$$

Then the probability of wetness of any day was assumed to depend on only on whether the previous day was wet or dry, i.e. the rainfall occurrence process was assumed to be a Markov process. The Markov chain of the first-order is one for which each next state depends only on immediately preceding one. Markov chains of second or higher order are the processes in which the next state depends on two or more preceding ones. The first-order Markov chain was used to describe the occurrence of wet and dry days and its transition probability matrix P is defined as

$$P = \begin{bmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{bmatrix} \qquad (4)$$

where, P_{01} = the conditional probability of a wet day following a dry day, P_{11} = the conditional probability of a wet day following a wet day.

The complementary probabilities for dry day occurrences we obtained as

$$P_{00} = 1 - P_{01} \tag{5}$$

$$P_{10} = 1 - P_{11} \tag{6}$$

The transition probabilities were considered on a monthly base thus the model required 24 parameters for the rain event generation for one year (12 for P_{01} and 12 for P_{11}). These probabilities were calculated on all the available recordings in the data set as:

$$P_{01} = \frac{N_{01}}{N_0} \tag{7}$$

$$P_{11} = \frac{N_{11}}{N_1} \tag{8}$$

where, N_{01} is the number of wet days after a dry day in the month; N_0 is the total number of dry days in the data set, for the month; N_{11} is the number of wet days after a wet day in the month; N_1 is the total number of wet days in the data set, for the month.

2.4 Synthetic Generation of Daily Rainfall

By comparing a random number generated from a uniform distribution between 0 and 1 to the value of the transition probabilities P_{01} or P_{11}, the occurrence of a wet day was determined. The preceding day was considered dry and the current day as a wet day if the random number is smaller than P_{01}. Alternatively, the current day was predicted dry if the random number was greater than P_{01}. The decision process was similar if the preceding day was wet. The two parameters of gamma probability distribution function (pdf) were also estimated using the maximum likelihood method [11].When a wet day was generated, a rainfall amount on that day was determined by generating a new random number from a uniform distribution and solving the inverse gamma cumulative distribution function for daily rainfall amount [10]. The pdf of gamma distribution was given as:

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^{\alpha}\Gamma(\alpha)} \tag{9}$$

where $\Gamma(\alpha)$ is the gamma function. The shape and scale parameters are represented by α and β respectively. α and β are specific parameters for each month. The total number of parameters needed to describe the rainfall amount is 24 (12 for α and 12 for β for each month). The parameters α and β are estimated, on a monthly basis as follows

$$\alpha = \frac{M^2}{V} \tag{10}$$

$$\beta = \frac{V}{M} \tag{11}$$

where M is mean and V is the variance of the daily rainfall amounts for wet day.

2.5 Model Validation

The means and standard deviations of daily rainfall amounts of both the generated and observed rainfall amounts were compared. The correlation coefficient between the observed and the generated daily rainfall series was estimated to know if the two series were indistinguishable from each other.

3. RESULTS AND DISCUSSION

3.1 Descriptive Statistics of Flood Data

Figure 2 shows the average and standard deviation of daily rainfall for each month for the data which is taken from Gombe Airport. The variability of mean monthly rainfall is significantly high; between a standard deviation of 27 in July to 4 in December. This is very evident for the wettest month of July (rainy season) except for the months of November, December, January and February which are very dry (dry season). The monthly variation of rainfall during the drier months is significantly lower than the wet months.

 ISSN: 2645-2685

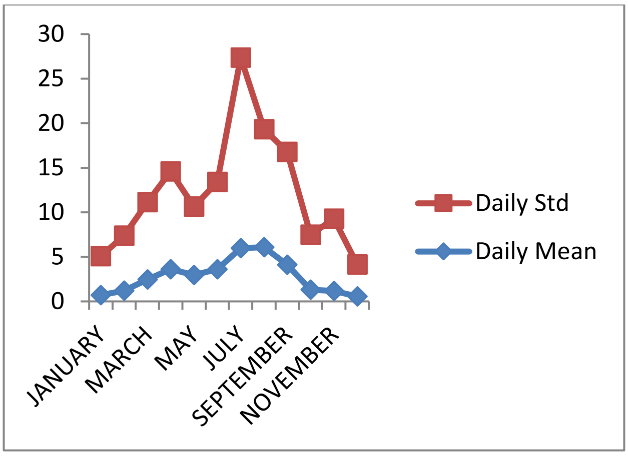

Figure 2: Daily mean and standard deviation of Gombe rainfall data for each month

3.2 Transition Probabilities of Observed Daily Rainfall for Each Month

Figure 3 illustrates the estimates of transitional probabilities. The conditional probabilities of wet day rainfall demonstrate the persistence of daily rainfall events. It is observed that the periodical changes of P_{01} and P_{11} are higher in March, August and November. However, this is in variance with the results obtained by [11] in his research. He observed that P_{01} and P_{11} were higher in July, August and September. This can be ascribed to the difference in the climate between the two study areas. Both in the wet and dry seasons, a wet day is more likely to be followed by a wet day, while except in month of May, the probability of a wet day following dry day is much smaller than a dry day following a dry ($P_{11}> P_{01}$). The conditional probability of wet day following a wet day is greater than the conditional probability of wet day following a dry day for every month ($P_{11}> P_{01}$).

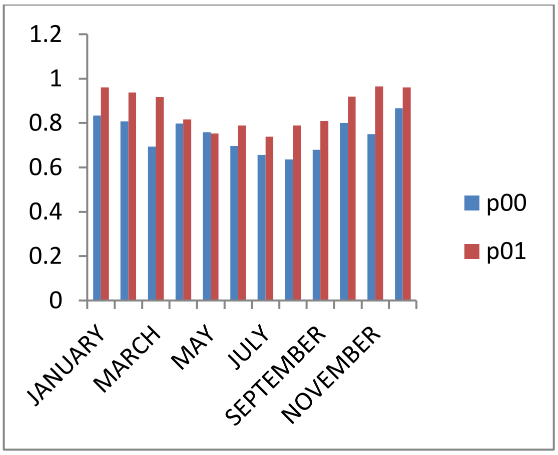

Figure 3: Conditional probability of daily rainfall for each month

3.3 Estimated Parameters of the fitted Gamma Probability Distribution

The results of Gamma distribution are shown in Figure 4. It can be seen that the value of shape parameter of Gamma distribution is varied in the range from 0.01 (December) to 2.99 (August), while scale parameter varies greatly in a range from 3.99 (December) to 22.5 (July), which indicates the significant variation of daily rainfall.

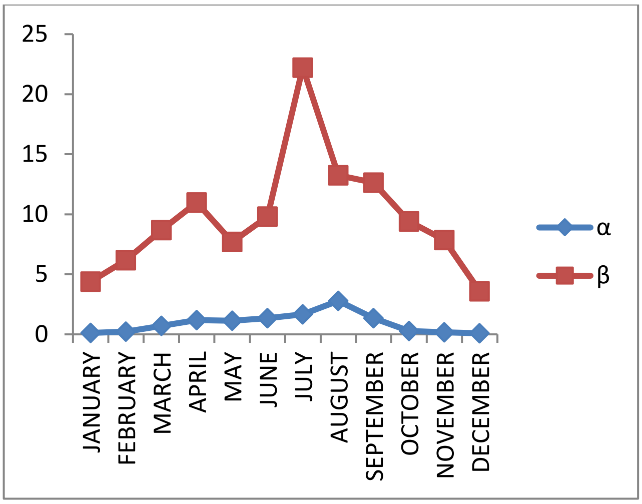

Figure 4: Monthly shape and scale parameters of the fitted gamma distribution

3.4 Statistical Tests on the fitted Gamma Distribution and First-order Markov Chains

According to the result of the two tests, the lowest statistic value is the best fitting model. The result of the two statistical tests is shown in Table 1. From Table 1, it was observed that first-order Markov chains adequately fit the synthesis of daily rainfall occurrence with the Akaike's Information Criterion (AIC) value of 0.11163 and the observed daily rainfall amounts satisfactorily were drawn from Gamma distribution with calculated Chi-Square (C-S) value of 2.263 less than the tabulated C-S value of 11.07. This is consistent with the findings of [11] who performed similar work in Bangladesh.

Table 1: Goodness of Fit Test

Goodness of fit	Akaike's Information Criterion (AIC)	Chi-Square (C-S).
Statistic value	0.11163	2.263

3.5 Results of Model Validation

The comparisons between simulated and recorded daily mean rainfall conditions showed a very good match with a correlation coefficient of 0.99 (Figure 5). The diagonal line in Figure 5 represents a one-to-one correspondence of the observed and simulated daily mean rainfall values. Again, this falls in line with the results obtained by [15] who simulated groundwater heads using MODFLOW in Ohaji Local Government Area of Imo State, Nigeria.

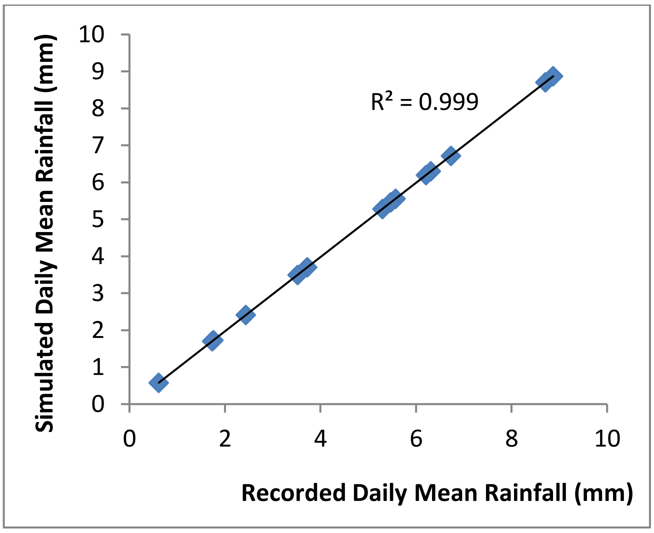

Figure 5: Simulated and recorded daily mean rainfall (mm)

The average amounts of rainfall for wet days from June to September were observed to be larger than that in the months from November to May (Figure 6). This result is slightly different from the results obtained by [11] in Bangladesh. In his study, he observed large average amounts of rainfall for wet days from June to October in the simulated data. This may

 ISSN: 2645-2685

have been as a result of the difference in the climate of Bangladesh, an Asian country and that of Gombe, a semi-arid area in Nigeria where this study was conducted.

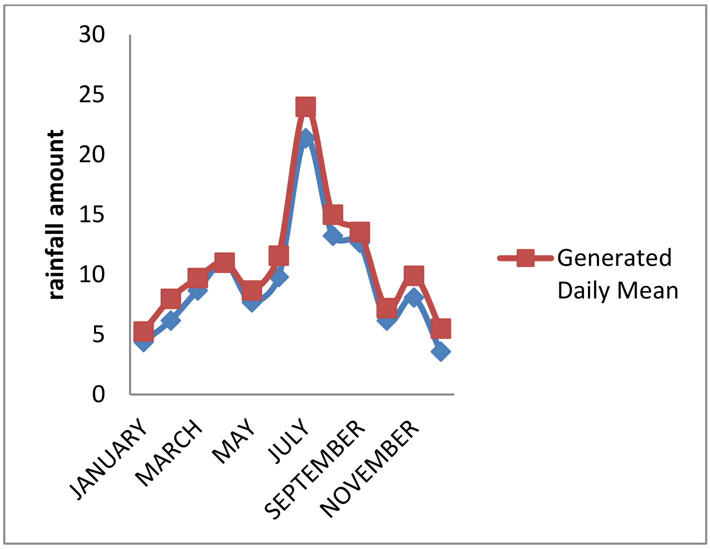

Figure 6: Comparison of recorded and generated mean rainfall (mm) per wet day

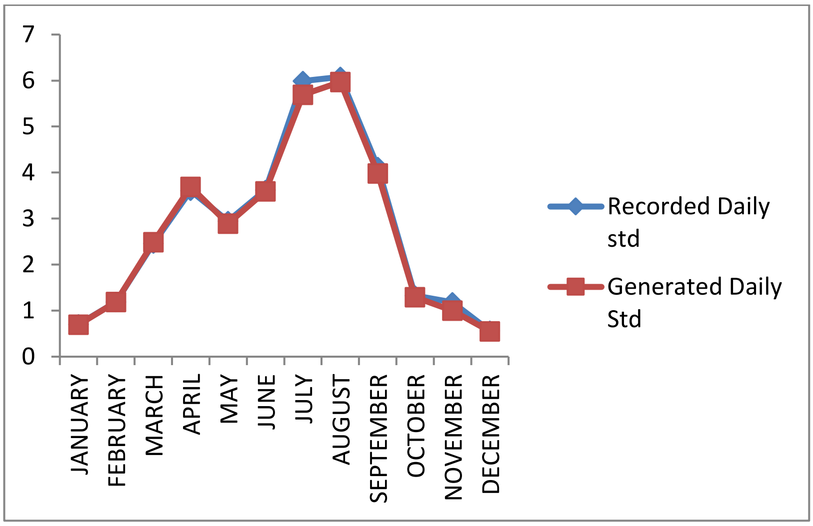

Figure 7: Comparison of recorded and generated standard deviation of rainfall (mm)

The variation of rainfall for wet days of historic rainfall data was similar to the variations in the synthetic rainfall data for every month (Figure 7). The standard deviation of daily rainfall for wet days was higher in June through September than every other month. The highest variability was observed in August both for recorded and synthetic rainfall data.

4. CONCLUSION

The developed Markov chain model was used to generate daily rainfall. The rainfall occurrence was determined by the first-order two state Markov chain and the rainfall amount were described by the gamma distribution. The estimated parameters of the Markov chain model given by the maximum likelihood estimation for each period showed significant variation of daily precipitation. The variability of mean monthly rainfall was observed to be significantly high. The highest variability showed in August both for recorded and generated data. Based on these results, the developed model can be concluded to be capable of representing many of the characteristics that existed in the observed data. In conclusion, on the basis of the comparison made in this paper, the Markov chain model is recommended as an appropriate stochastic simulation model for describing the Gombe rainfall data. There is need for further research on the rainfall parameters and the scenario of the rainfall probability that show the net irrigation requirement of different crops in different seasons

REFERENCES

[1] Danuso, F. (2002). Climak: a stochastic model for weather data generation. Italy. *J. Agron.* 6(1),57-71.

[2] Deni, S. M. & Jemain, A. A. (2009). Fitting the distribution of dry and wet spells with alternative probability models. *Met. and Atmosph. Physics.* 104:13-27.2.
[3] Kannan, S. & Farouk, K. (2015). Stochastic Simulation of Precipitation Using Markov Chain - Mixed Exponential Model. *Applied Mathematical Sciences*. (9) 65: 3205 – 3212.
[4] Piantadosi J., Boland J. W. & Howlett P. G. (2008). Generating synthetic rainfall on various timescales daily, monthly and yearly. *J. Environ. Model Assess*. 14, 431-438.
[5] Srikanthan, R. and McMahon, T.A. (2001). Stochastic generation of annual, monthly and daily climate data: A review. *Hydrology and Earth System Sciences*. 5(4): 653–670.
[6] Ahamed, S., Piantadosi, J, Agrawal, M. & Boland, J. (2013). *Generating Synthetic Rainfall Using a Disaggregation Model*. In Proceedings, 20th International Congress on Modelling and Simulation, Adelaide, Australia, 1–6 December. www.mssanz.org.au/modsim2013.
[7] Sørup, H.J.D., Madsen, H. and Arnbjerg-Nielsen K. (2011). *Markov chain modeling of precipitation time series: Modeling waiting times between tipping bucket rain gauge tips.* In Proccedings, 12th International Conference on Urban Drainage, Porto Alegre/Brazil, 11-16 September.
[8] Nnaji, A. O. (2001). Forecasting seasonal rainfalI for agricultural decision making in northern Nigeria. *J. of Agril. and Forest M*et. 107, 193-205.
[9] Kottegoda, N. T., Natale, L. & Raiteri, E. (2004). Some considerations of periodicity and persistence in daily rainfalls. *J. Hydrol*. 296,23-37.
[10] Ibeje, A.O, Osuagwu, J. C & Onosakponome, O. R. (2018). A Markov Model for Prediction of Annual Rainfall. *International Journal of Scientific Engineering and Applied Science*. 4(11), 34-47.
[11] Barkotulla, M. A. B. (2010) Stochastic Generation of the Occurrence and Amount of Daily Rainfall Pak.J.Stat.Oper.Res. 4(1):61-73.
[12] Ibeje, A. O. (2020). *Web-Based Forecasting of Short-Term Reservoir Inflow*. Published Ph.D. Thesis, Federal University of Technology Owerri. Mauritius: Lambert Academic Publishing.
[13] Ibeje, A. O. & Okoro, B.C. (2013). Short-Term Forecasting of Dadin-Kowa Reservoir Inflow Using Artificial Neural Network. *The International Journal of Engineering and Science*. 2(6), 63-73.
[14] Taha, H. A. (2008). Operations Research: An Introduction 8th Ed Prentice Hall of India, India.
[15] Ibeje, A.O. (2019). Coupled Groundwater Simulation-Optimization Modelling of Agrarian Sub-catchment. *International Journal of Scientific & Engineering Research*. 10(12), 777-781.