Analysis of the Stochastic Characteristics and Modelling of Monthly Rainfall Time Series of Abeokuta, Nigeria

John Jiya Musa[1]*, Ibrahim Jibril[2], Martins Yusuf Otache[1], Abayomi Ibrahim Kuti[1] and Pius Olufemi Dada[3]

[1]Department of Agricultural and Bioresources Engineering, Federal University of Technology, P M B 65, Minna, Nigeria. E-mail: johnmusa@futminna.edu.ng

[2](Research Assistant) Niger State College of Agriculture, P M B 109, Mokwa, Nigeria.

[3]Department of Agricultural and Bioresources Engineering, Federal University of Agriculture, Abeokuta, Ogun State, Nigeria.

Abstract

This study attempts to analyse the stochastic characteristics of rainfall for Abeokuta, Nigeria and its probable simulation. To this end, rainfall time series data of Abeokuta was obtained from the Nigerian Meteorological Agency (NIMET) for a period of 28 years. The analysis of the stochastic characteristics entails the assessment of temporal trend and periodicity while modeling of the time series was done by employing the seasonal multiplicative Autoregressive Moving Average (ARIMA) modeling technique. Results obtained indicate that there was no discernible trend in the annual series though spectral analyses show that there is high periodicity of fluctuating frequency in the monthly time series. Based on the extent of periodicity and degree of randomness, a multiplicative seasonally differenced ARIMA model was found appropriate for the simulation of the monthly rainfall regime; model choice was on the basis of Akaike Information Criterion (AIC) and Autocorrelation functions. On the basis of these two criteria, ARIMA(0,0.1) x (0,1,1) x $(0,1,1)_{12}$ and ARIMA (1,0,1) x $(1,1,1)_2$ respectively were adjudged probable candidate models for simulation studies. However considering the fact that rainfall phenomenon exhibits high spatio-temporal variability, the seasonal persistence can only be explained relatively by the autoregressive component rather than solely the moving average component; thus, the second model is preferable. Despite this though, for effective generalization of simulation results, Artificial Neural Network and Wavelet models are recommended and in this regard too, conditional probability of rainfall occurrence should be considered.

Keywords: *Stochastic characteristic, seasonality, trend, ARIMA, Simulation*

*** Corresponding Author**

Introduction

Weather has generally been considered as the state of the atmosphere at a given time at any given location. It may also be referred to as the aspects of the atmospheric state which is visible, experienced and affects human activities (Ogolo and Adeyemi, 2009). Weather and climate science is founded on observing and understanding our complex and evolving environment. It is therefore an inherent requirement that climate scientists make available the best possible information on the current state of the climate, and on its historical context. This draws on globally distributed observations and monitoring systems and networks. This is also dependent on robust data processing and analysis to synthesis vast amounts of data, properly taking into account observational uncertainty resulting from both measurement limitations and sampling (Morice *et al.*, 2012).The weather conditions of any given location is often described in terms of the meteorological elements which include the state of the sky, temperature, winds, pressure, precipitation, and humidity. These factors initiate and influence the atmospheric processes (Ayoade, 1993).

According to Williams (2008), among all the climatic elements, rainfall is the most variable, both temporally and spatially and such variations can have significant impacts on economic activity. Rainfall is one of the most important components of hydrologic cycle which begins with change in temperature and relative humidity (Tsoho, 2008).Theoretically, an increase in heavy rainfall events can be expected in response to global warming. Satellite observation and related history indicates that the amount of perceptible water in the vertical column over the ocean increases non-linearly with increasing sea surface temperature. Thus, the amount of water vapor available for rainfall is greater for higher base temperature typically of the tropics or warmer temperature due to enhanced greenhouse condition (Whetton *et al.*, 2001). There is general agreement that many areas of currently high precipitation is expected to experience precipitation increases, whereas many of the areas at present with low precipitation and high evaporation, now suffering water scarcity, are expected to have rain decreases in the future (IPCC, 2007).

The hydrologic effects of climate change will have an important influence on all types of basins and many areas will likely follow predicted changes in precipitation (Melkamu, 2013). In the Mediterranean region, continental precipitation is increased by 5–10% over the 20th century in the northern hemisphere and decreased in other regions (for example, North and west Africa and parts of the Mediterranean), increases in the east African basins (IPCC, 2007). Hayileyesus (2011) presented a thesis on evaluation of climate change impacts on hydrology on selected catchments of upper Blue Nile basin based on the precipitation scenarios generated. He indicated that, change of precipitation in percent for the first future time series: 2031-2040 with respect to the base period.

To enhance an understanding of the dynamics of hydroclimatic processes, hydrological modeling becomes an undeniable fact, as models are now major tools in the study of hydrological processes, mainly used for different purposes such as water management or flood forecasting. The estimation of hydrological model parameters is a difficult task. Reasons for these are the highly non-linear nature of hydrological processes. This means that changes of some parameters might be compensated by others. Unfortunately traditional manual calibration of models with reasonable parameter values often leads to weak results. Hence, nowadays automatic procedures based on numerical methods are used (Bardossy and Singh, 2008). Against this backdrop, the central theme of this study is the assessment of the stochastic characteristics of rainfall time series of Abeokuta.

Materials and Method

Study Area

The entire study area is bounded by Oyo state to the north and Lagos State to the South. It is located in southern Nigeria, bordered geographically by latitudes 6.26^0 N and 9.10^{0}N and longitudes 2.28^{0}E and 4.8^0 E. The land area is about 23,000km^2 with a generally low relief and gradient in the North-South direction. The two major vegetation zones that can be identified the area are the high forest vegetation in the north and central parts, and the swamp/mangrove forests that cover the southern coastal and floodplains, next to the lagoon. It has two distinct seasons throughout the year. The monthly rainfall distribution in the study area shows a distinct dry season extending from November through March and a rainy season spanning April to October.

Data Collection

The rainfall data of Abeokuta, Ogun State was obtained from the Nigerian Meteorological Agency (NIMET). The data obtained covered a period of twenty eight years (1982-2009).

(iii) Analysis of Stochastic Characteristics

For this study, stochastic characteristics examined were limited to temporal trend, serial dependence and periodicity.

Trend Analysis

The Mann-Kendall non-parametric test was considered for trend detection in the annual time series data because of its robustness and unique advantages over other methods. Trend examination according to the Mann-Kendall approach was done by employing equations (1) – (4) in terms of the overall Z test statistic; in this regard, at 5 % significance level.

$$S=\sum_{k=1}^{n-1}\sum_{j=k+1}^{n}\operatorname{sgn}(X_j - X_k) \qquad (1)$$

where, X_j and X_k are the annual values in years j and k, j > k, respectively, and

$$\operatorname{sgn}(X_j - X_k) = \begin{cases} 1 & \textit{if } X_j - X_k > 0 \\ 0 & \textit{if } X_j - X_k = 0 \\ -1 & \textit{if } X_j - X_k < 0 \end{cases} \qquad (2)$$

$$VAR(S) = \frac{1}{18}[n(n-1)(2n+5) - \sum_{p=1}^{q} t_p(t_p - 1)(2t_p + 5)] \qquad (3)$$

Source: Musa et al (2016). Besaw et al (2010); and Ghanabarpour et al. (2007).

$$Z = \begin{cases} (S-1)/Var(S)^{1/2} & S > 0 \\ 0 & S = 0 \\ (S+1)/Var(S)^{1/2} & S < 0 \end{cases} \qquad (4)$$

Source: Longobardi and Villan (2009) and Tesfaye et al (2006).

The Mann-Kendall test was carried out in accordance with the works of Otache *et al.* (2011); Ahaneku and Otache (2014) and Chatfield (2004), with the aid of the excel template of 'MAKESEN's version 1.

Serial dependence and Periodicity

The Durbin-Watson test was considered in order to check for serial correlation; i.e., over time rather than over realizations. The analysis of periodicity was done by evaluating the spectral density using the Fast Fourier transforms with Tukey lag window.

Model development

The development of stochastic models for simulation studies requires that the data series should be near second order stationarity. To this end, both seasonal and non-seasonal differencing was considered and corresponding models were developed accordingly. The choice of a particular candidate model was based on Akaike Information Criterion and Autocorrelation functions.

Results and Discussion

Model development and Analysis of Stochastic Characteristics

The time series is treated as trend free since the Mann-Kendall test result showed no significance in the trend with a Z-value of 0.45. However, a clear seasonal pattern is observed in the ACF plot of the raw rainfall data with discernible periodicity of order six as shown in Fig. 1. A single seasonal differencing was found out to be suitable in achieving near stationarity in accordance with the work of Kumar and Vanajakshi (2015). Using Box, *et al.* (2008) methodology and considering the ACF and PACF of the differenced data, seasonal autoregressive integrated moving average models of orders $(0,0,1) \times (0,1,1)_{12}$ and $(0,0,1) \times (0, 1, 2)_{12}$ and ARIMA$(1,0,1) \times (1, 1, 1)_{12}$ were identified. Model diagnostic check was in accordance with the works of Williams and Hoel (2003) and Burnham and Anderson (2004), and the Corrected Akaike Information Criterion (AICc) computed for the two models were 1296.36 and 1297.11 , respectively (see Table 1). Model $(0,0,1) \times (0,1,1)_{12}$ has the least AICc value as shown in high-lighted row of Table 1 and thus, taken as a better model in this regard. Table 2 shows another model characteristics on the basis of the behaviour of the ACF and PACF of the seasonally differenced monthly series (see Figs 3 and 4).

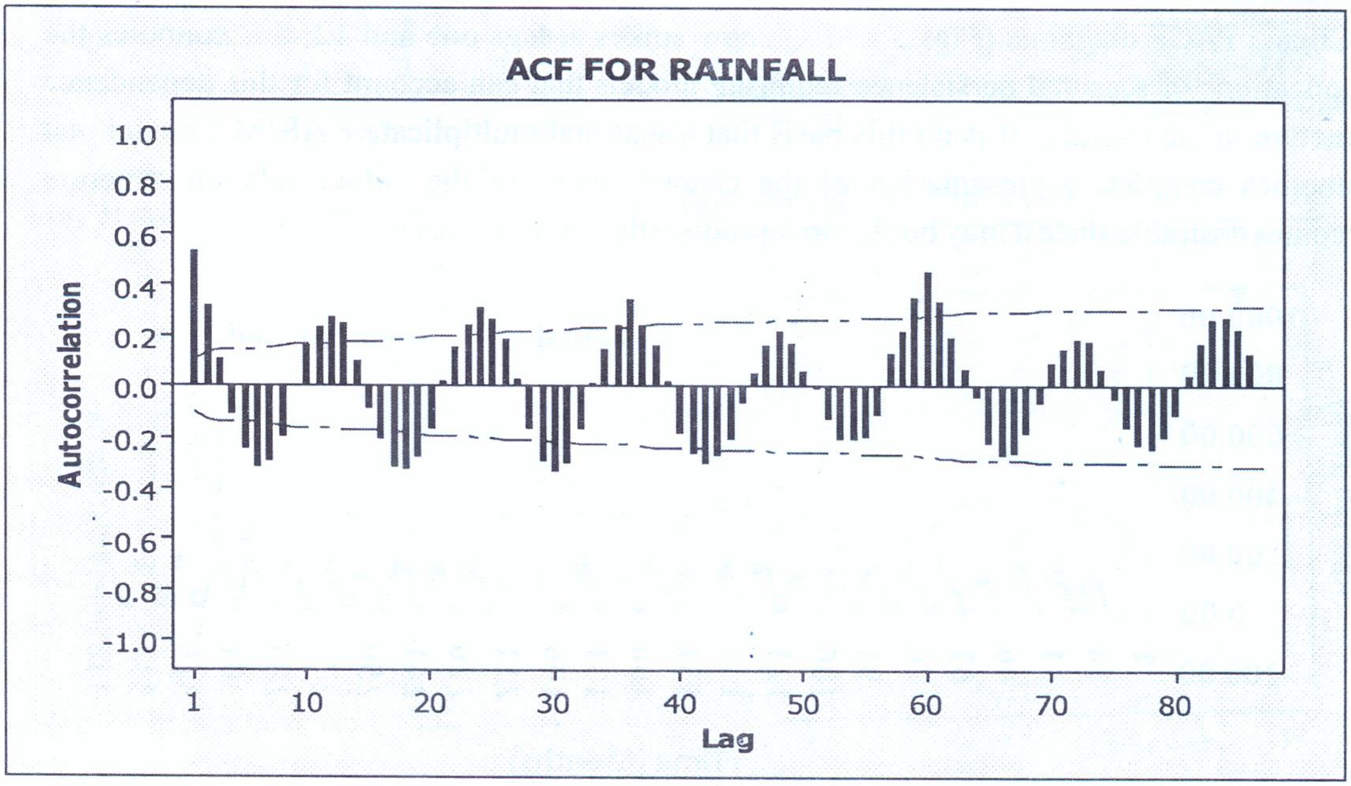

Fig. 1: Correlogram showing seasonal pattern in monthly rainfall of Abeokuta, Nigeria

Table 1: SARIMA - Model order selection for rainfall

Non-Seasonal model order (p, d, q)	Seasonal model order (P, D, Q)s	Sum of squares	AICc - value
(0, 0, 1)	$(0, 1, 1)_{12}$	1799603	1296.36
(0, 0, 1)	$(0, 1, 2)_{12}$	1784306	1297.11

Table 2: SARIMA $(1, 0, 1) \times (1, 1, 1)_{12}$ and estimated values

Model Type	Model Order	Parameter Estimated	P- Value	Constant
NSMA (θ)	1	-0.3047	0.000	0.7801
SMA (Θ)	12	0.9661	0.000	,, ,,

Figure 2 shows the ARIMA $(0,0,1) \times (0,1,1)_{12}$ model prediction capability. It is glaringly evident that there is a good correlation between the observed and simulated values. This implies that the model has the capability of reproducing the rainfall regime over time.Table 3 shows the analysis of the simulations. Here, the Mean Absolute Percentage Error (MAPE) was considered in accordance with Lewi's error scaling system (i.e. MAPE value < 10 %).Both the ACF and PACF diagrams (Figs 3 and 4) show spikes at lags one and 12; this connotes the implications of seasonal persistence requiring models that can account for this dependence structure in the overall. It is on this basis that a seasonal multiplicative ARIMA model that embodies complete representation of the characteristics of the autocorrelation structure becomes desirable since it may be able to reproduce the random error variations.

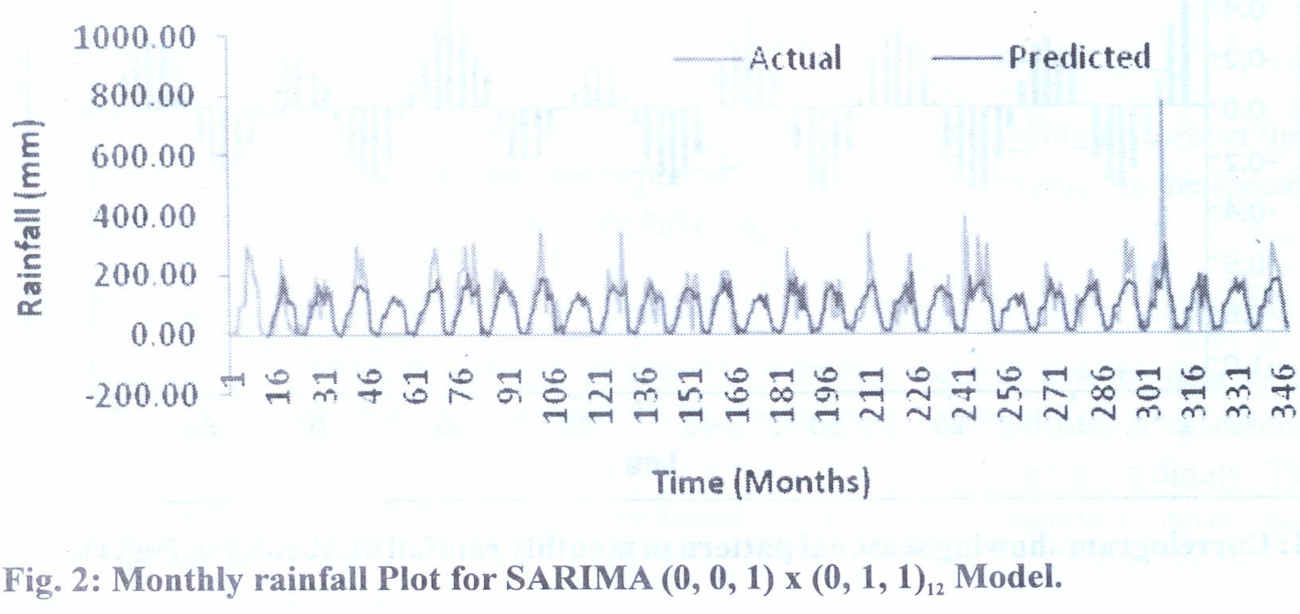

Fig. 2: Monthly rainfall Plot for SARIMA $(0, 0, 1) \times (0, 1, 1)_{12}$ Model.

Table 3: Summery of Error Values for Observed and Predicted Rainfall

Model Type	Model Order	Forecast Err.	MSE	RMSE	MAPE (%)
SARIMA	$(0,0,2)x(0,1,1)_{12}$	3.09	58.01	26.68	7.62
SARIMA	$(0,0,1)x(0,1,1)_{12}$	0.21	53.61	21.61	2.29

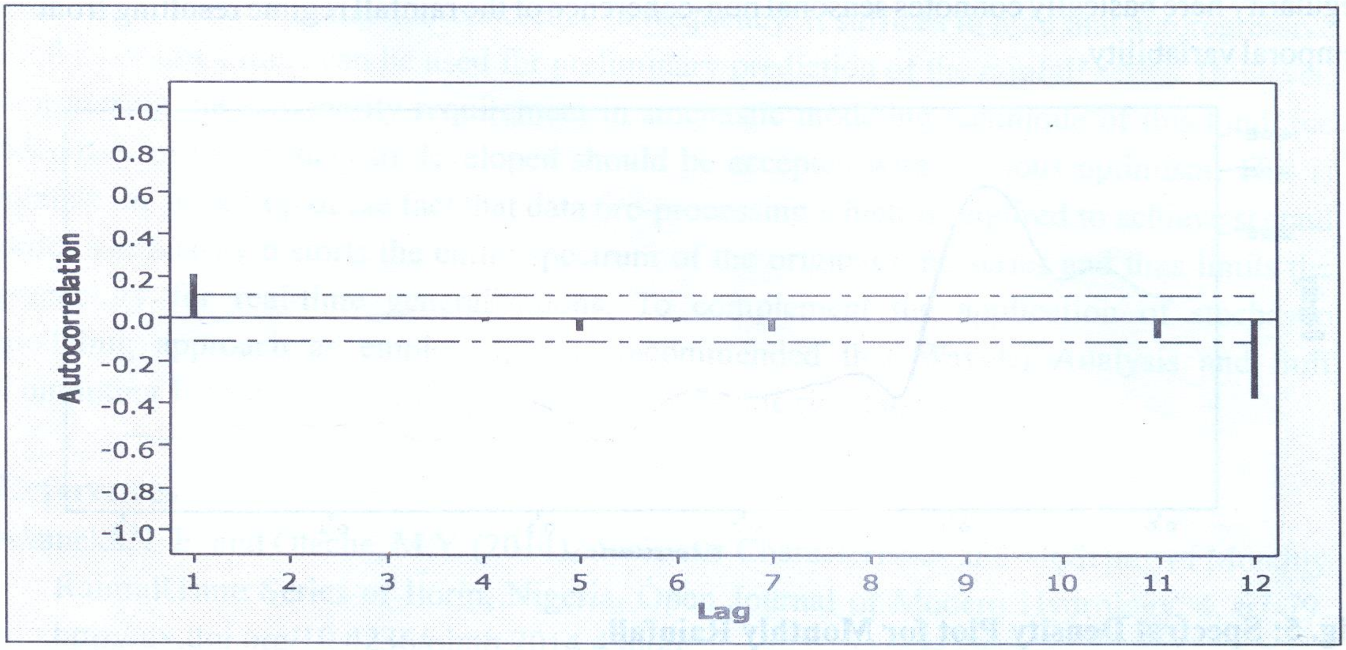

Fig. 3: Autocorrelation plot for seasonally differenced rainfall series

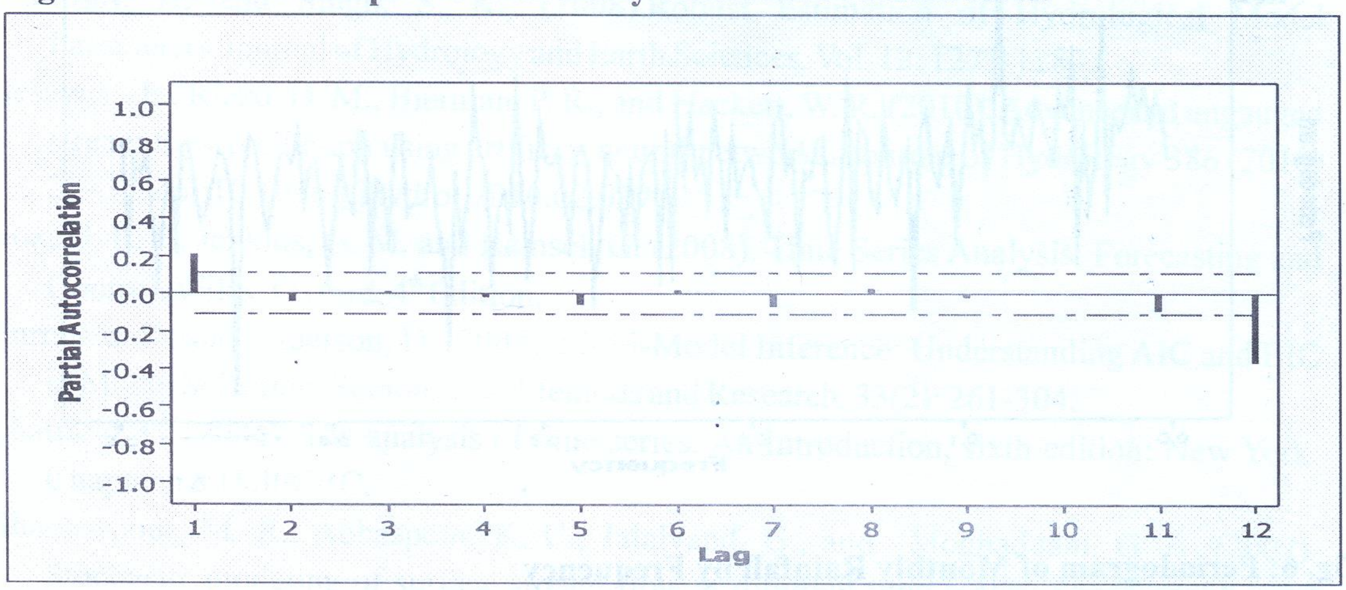

Fig. 4: Partial Autocorrelation plot for the seasonally differenced rainfall series

From the spectral density diagram, it can be seen that the higher the frequency, the lower the density. This indicates a positive autocorrelation in the time series which agrees with the Durbin-Watson test carried out for serial dependence. This is presented in Figs (5) and (6), respectively. From Fig. (5), the periodic nature of the time series is evident; here, low frequencies of high density dominate indicating a probable rainfall regime of short wavelength. In addition, Fig. (6) shows the fluctuations of the frequency pattern; the non-regularity here basically connotes seasonal non-coherence of the rainfall regime resulting from temporal variability.

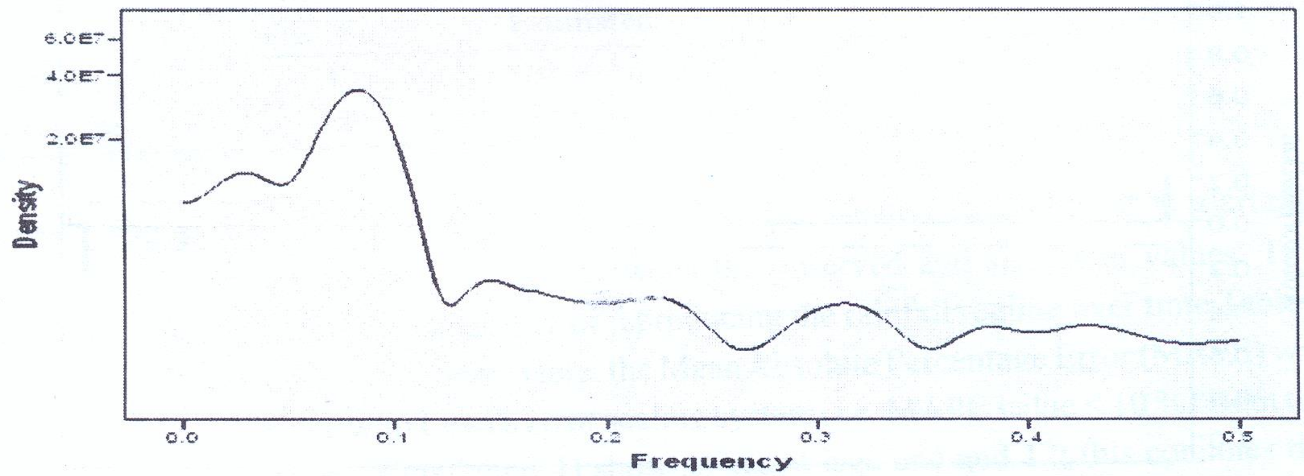

Fig. 5: Spectral Density Plot for Monthly Rainfall

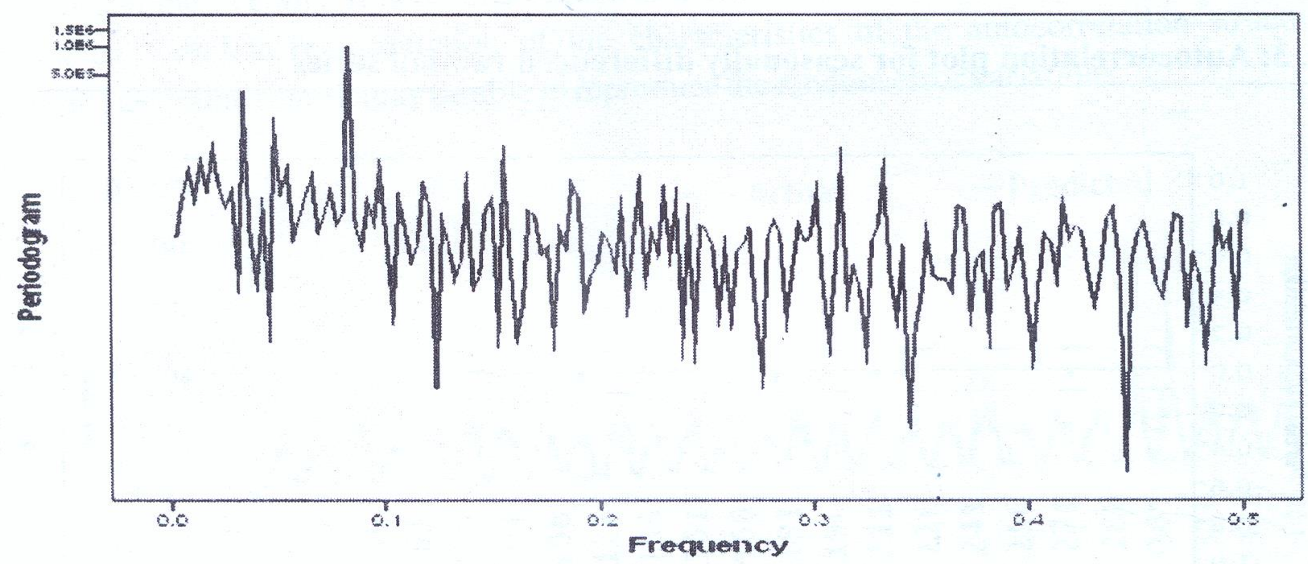

Fig. 6: Periodogram of Monthly Rainfall by Frequency

Conclusions

The accessibility to records of hydrological processes in which rainfall is a major determinant is imperative for proper guide and timely preparation against extreme events. Several methods have been used to predict hydrological behaviors, but have shown some weaknesses due to their stochastic nature. The results obtained in this study demonstrate that there was no trend in the annual series of the rainfall regime though with seeming discernible seasonal periodicity of order six (6) and seasonal non-coherence in frequency. It suffices to note that autoregressive models of low orders can be used for preliminary prediction of the rainfall series. However, considering the stationarity requirement in stochastic modeling technique of this kind, the adequacy of the models so developed should be accepted with cautious optimism. This is against the backdrop of the fact that data pre-processing which is required to achieve second order stationarity distorts the entire spectrum of the original time series and thus limits the possibility for real-time generalizations. To complement the application of stochastic modeling approach as employed, it is recommended that Wavelet Analysis and Soft Computing Techniques should be explored for real-time simulations.

References

Ahaneku, I. E. and Otache, M.Y. (2014).Stochastic Characteristics and Modeling of Monthly RainfallTime Series of Ilorin, Nigeria. Open Journal of Modern Hydrology, 4, 67-79. http://dx.doi.org/10.4236/ojmh.2014.43006.

Ayoade, J. O. (1993).Introduction to the Tropics. Spectrum Books Ltd. Lagos, Nigeria.

Bardossy, A. and Singh, S. K., (2008).Robust Estimation of Hydrological Model Parameters.Journal of Hydrology and Earth Sciences. Vol. 12: 1273-1283.

Besaw, L. E., Rizzo, D. M., Bierman, P. R., and Hackett, W. R. (2010). Advances in ungauged stream flow prediction using artificial neural networks. Journal of Hydrology 386 (2010) 27–37. doi:10.1016/j.jhydrol.2010.02.037

Box, G. E. P., Jenkins, G. M. and Reinsel, G. (2008). Time Series Analysis: Forecasting and Control. Wiley, Oxford, 4th Edition.

Burnham, K. and Anderson, D. (2004). Multi-Model Inference: Understanding AIC and BIC in Model Selection. Sociological Methods and Research, 33(2): 261-304.

Chatfield, C. (2004). The analysis of time series: An Introduction, sixth edition: New York, Chapman &Hall/CRC.

Ghanbarpour, M. R., Abbaspour, K. C., Jalalvand, G., and Moghaddam, G. A. (2007). Stochastic modeling of surface stream flow at different time scales: Sangsoorakh karst basin, Iran. Journal of Cave and Karst Studies, v. 72, no. 1, p. 1–10. DOI:

10.4311/jcks2007ES0017

Haileyesus, B. (2011).Evaluation of Climate Change Impacts on Hydrology of Selected Catchments of Abbay Basin. Unpublished M.Sc. ThesisAddis Ababa University.School of Post Graduate Studies. Pp. 60-67.

IPCC, (2007). Climate Change: The Physical Science Basis. Summary for Policy Makers.Contribution of WG1 to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.Available at:http://www.ipcc.ch/ipccreports/ar4- wg1.htm.

Kumar, V. S. and Vanajakshi, L. (2015).Short-term Traffic Flow Prediction Using Seasonal ARIMA Model with Limited Input Data. European Transport Resources Review 7:21. DOI: 10.1007/s12544-015-0170-8.

Longobardi, A., and Villan, P. (2009). Trend analysis of annual and seasonal rainfall time series in the Mediterranean area. International Journal of Climatology, Pages 1-9. DOI: 10.1002/joc.2001

Morice, C. P., Kennedy, J. J., Rayner, N. A. and Jones, P. D. (2012). Quantifying Uncertainties in Global and Regional Temperature Change using an Ensemble of Observational Estimates: The HadCRUT4 dataset, J. Geophys. Res., 117, D08101, doi:10.1029/2011JD017187.

Otache, M.Y., Ahaneku, I.E. and Mohammed, S.A. (2011). ARMA Modeling of Benue River Flow Dynamics: Comparative Study of PAR Model. Open Journal of Modern Hydrology, Scientific Research, 1, 1-9.

Ogolo, E. O. and Adeyemi, B. (2009).Variations and Trends of Some Meteorological Parametersat Ibadan, Nigeria.Pacific Journal of Science and Technology. 10(2):980-987.

Tesfaye, Y. G., Meerschaert, M. M., and Anderson, P. L. (2006). Identification of periodic autoregressive moving average models and their application to the modeling of river flows. Water Resources Research, VOL. 42, W01419, doi:10.1029/2004WR003772

Tsoho, U. H. (2008). Growth and History of the Establishment of MakeraAssada in SokotoMetropolis to the year 2007.B. A.Project, History Department, UsmanuDan-fodiyo University Sokoto.

Whetton, P., Fowler, A. and Haylock, M. (2001). Implication of the Climate Change due to Enhanced Greenhouse effect on Floods and Droughts in Australia.Climate Change Group CSIRO Division of Atmosphere Research. Journal of Climate Change 25:289-317.

Williams, L. (2008). Nigeria: The Bradt Travel Guide. Pp. 110.ISBN 1-84162-239-7.

Williams, B. M. and Hoel, L. A. (2003). Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results. Journal of Transport Engineering 129(6): 664-672.