Physical Science International Journal

22(4): 1-9, 2019; Article no.PSIJ.50701
ISSN: 2348-0130

Seasonal Autoregressive Integrated Moving Average (SARIMA) Model for the Analysis of Frequency of Monthly Rainfall in Osun State, Nigeria

Samuel Olorunfemi Adams[1*], Bello Mustapha[1] and Auta Irinew Alumbugu[2]

[1]Department of Statistics, University of Abuja, Nigeria.
[2]Department of Mathematics and Statistics, Federal Polytechnic Nasarawa, Nigeria.

Authors' contributions

This work was carried out in collaboration among all authors. Author SOA designed the study, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Author BM managed the analyses of the study. Author AIA managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/PSIJ/2019/v22i430139
Editor(s):
(1) Dr. Manuel Malaver de la Fuente, Universidad Marítima del Caribe, Av. El Ejército, Catia La Mar, Estado Vargas, Venezuela.
(2) Dr. Christian Brosseau Distinguished, Professor, Department of Physics, Université de Bretagne Occidentale, France.
Reviewers:
(1) Olumide Adesina, Olabisi Onabanjo University, Nigeria.
(2) Janilson P. Assisi, Federal Rural University of the Semi-Arid (UFERSA), Brazil.
(3) Ramesh M. Mirajkar, Dr. Babasaheb Ambedkar College, India.
Complete Peer review History: http://www.sdiarticle3.com/review-history/50701

Original Research Article

Received 26 May 2019
Accepted 06 August 2019
Published 10 August 2019

ABSTRACT

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model is proposed for Osun State monthly rainfall data and the analysis was based on probability time series modeling approach. The Plot of the original data shows that the time series is stationary and the Augmented Dickey-Fuller test did not suggest otherwise. The graph further displays evidence of seasonality and it was removed by seasonal differencing. The plots of the ACF and PACF show spikes at seasonal lags respectively, suggesting SARIMA (1, 0, 1) (2, 1, 1). Though the diagnostic check on the model favoured the fitted model, the Auto Regressive parameter was found to be statistically insignificant and this led to a reduced SARIMA (1, 0, 1) (1, 1, 1) model that best fit the data and was used to make forecast.

**Corresponding author: E-mail: samuel.adams@uniabuja.edu.ng, adamsamuel22003@yahoo.com;*

Keywords: SARIMA model; time series; Dickey-Fuller test; autoregressive parameter; rainfall data.

1. INTRODUCTION

The highly variable nature of rainfall as compared with the relatively stable nature of the temperature appears to have imbued more relevance to the former as the major component in the study of climate in a particular region. There is need to understand the dynamical processes that determine changes that occur in climate system, though this has been very difficult and challenging to climate scientists till today [1]. The change has significantly contributed to the increase of global disasters caused by weather, climate and water related hazards as both developed and developing countries of the world are bearing the burden of repeated floods, temperature extremes and storms in which Nigeria is not left out [2]. Water resources are essential renewable resources that are the basis for existence and development of a society. Proper utilization of these resources requires assessment and management of the quantity and quality of the water resources both partially and temporally. Water crises caused by shortages, floods and diminishing water quality, among others are increasing in all parts of the world. The growth in population demands for increased domestic water supplies and at the same time results with a higher consumption of water due to expansion in agriculture and industry [2]. Mismanagement and lack of knowledge about existing water resources and the changing climatic conditions have consequences of an imbalance of supply and demand of water. A few literature exist in the time series analysis of monthly rainfall in some states in Nigeria, they include; Akwa-Ibom State, Uyo [3], Ile-Ife, Osun State [4], Umuahia, Abia State [5], Ilorin, Kwara State [6], Portharcourt, Rivers State [7], Akure, Ondo State [8–10], Ikeja Lagos [11], Ota, Ogun State [12] and Ogbomosho, Oyo State [13]. This study therefore attempts to identify and construct the best SARIMA model that best fits, explains the underlying generating process and satisfactorily forecast into the future of the monthly frequency of rainfall in Osun State.

2. SARIMA MODELING

Rainfall data are time structured and time series analyses are often employed in the analysis of the data. The data were subjected to seasonal autoregressive integrated moving average (SARIMA). Modeling. An ARIMA model is an algebraic statement that describes how a time series is statistically related to its own past. The seasonal ARIMA model incorporates both non-seasonal and seasonal factors in a multiplicative model given as;

$$ARIMA\ (p,d,q)\ \times\ (P,D,Q)_s \tag{2.1}$$

Where; p = non-seasonal AR order, d = non-seasonal differencing, q = non-seasonal MA order, P = seasonal AR order, D = seasonal differencing, Q = seasonal MA order, and S = time span of repeating seasonal pattern.

Without differencing operations, the model could be written more formally as;

$$\emptyset(B^s)\varphi(B)(X_t - \mu) = \theta(B^s)\theta(B)W_t \tag{2.2}$$

The non-seasonal components are:

$$AR{:}\ \varphi(B) = 1 - \varphi_i B -, \ldots, -\varphi_p B^p \tag{2.3}$$

$$MA{:}\ \theta(B) = 1 + \theta_i B +, \ldots, +\theta_q B^q \tag{2.4}$$

The seasonal components are:

$$Seasonal\ AR{:}\ \emptyset(B^s) = 1 - \emptyset_i B^s -, \ldots, -\emptyset_p B^{ps} \tag{2.5}$$

$$Seasonal\ MA{:}\ \theta(B^s) = 1 + \emptyset_i B^s +, \ldots, +\emptyset_q B^{qs} \tag{2.6}$$

Note that on the left side of equation (2.2) the seasonal and non-seasonal AR components multiply each other, and on the right side of equation (2.2) the seasonal and non-seasonal MA components multiply each other (14).

The SARIMA modeling approach is concerned with finding a parsimonious seasonal ARIMA model that describes the underlying generating processed of the observed time series. Box and Jenkins (15) established a three step modeling procedure: Identification, estimation and diagnostic checking steps. The identification step is to tentatively choose one or more ARIMA/SARIMA model(s) using the estimated ACF and PACF plots. The ACF plot of the AR (Auto Regressive)/ SAR (Seasonal Auto Regressive) process shows an exponential decay while its PACF plot truncates at lag p/seasonal lag p and diminishes to zero afterwards. The ACF plot of the MA process truncates to zero after lag q while its PACF decays exponentially to zero. The two processes:

AR (p)/SAR(P) and MA (q)/SMA(Q), could be combined to form the ARMA (p, q)/SARMA (P, Q) process which has ACF and PACF that decays exponentially to zero. The maximum likelihood estimation method could be used to estimate the parameters of the identified model(s) in the identification stage. The last diagnostic checking stage involves assessing the adequacy of the identified and fitted models through possible statistically significant test on the residuals to verify its consistency with the white noise process e.g. the Ljung-Box test [16]. Finally, the best fitting model would be selected among other satisfactory, competing models e.g. the information criteria statistics on the basis of the AIC or BIC rule of thumb, the Models with the smallest information criterion is the best and forecast is made with the model of best fit.

3. APPLICATION OF SARIMA MODEL

The data used in this study which is the frequency of the monthly rainfall in Osun State Nigeria, was collected from the National Bureau of Statistics, Nigeria from the year 1981-2015. The behavior of the data was observed and the estimation of the expected models was carried out using the method of likelihood with the plots of the Autocorrelation function (ACF) and Partial Autocorrelation Function (PACF) of the difference and non-difference series. After several iterations, some models were suggested; SARIMA$(1,0,1)(1,1,1)_{12}$, $(1,0,2)(1,1,1)_{12}$, $(1,0,1)(2,1,1)_{12}$, $(102)(1,1,2)_{12}$, $(2,0,1)(2,1,1)$ $(1,2,1)_{12}$, $(1,0,1)(1,1,2)_{12}$ as presented in the Table 1; Comparing the SARIMA$(1,0,1)(1,1,1)_{12}$, $(1,0,2)(1,1,1)_{12}$, $(1,0,1)(2,1,1)_{12}$, $(102)(1,1,2)_{12}$, $(2,0,1)(2,1,1)(1,2,1)_{12}$, $(1,0,1)(1,1,2)_{12}$ the suggested models were compared based on the criteria; AIC, Standard error, log likelihood, square sigma estimated and coefficient respectively, clearly, SARIMA $(1,0,1)\times(1,1,1)_{12}$ proved to be the appropriate model with minimum Akaike information criterion (AIC) of 4721.14. The selected model was then used to describe and forecast the time series observations.

From the plots in Fig.1 it could be seen that the time series observation displays a wave like pattern an evidence of seasonality and no trend is observed which implies that the time series might be stationary. The sinusoidal or periodic pattern in the ACF plot is again suggesting that the series has a strong seasonal effect also, the PACF plot is neither suggesting otherwise. In order to verify the stationarity claim of the visual displays, the Augmented Dickey-Fuller [17] test was performed.

Table 2 depicts the Augmented Dickey-Fuller Test, the hypothesis.

H_0: The series is unit root non stationary vs H_1: the series is unit root stationary P-value < 0.05, indicating a strong evidence against the null hypothesis at 5% level of significance. In order to eliminate the seasonal effect from the time series observation, it was subjected to a seasonal differencing and the data is re-examined visually as seen in Fig. 4.

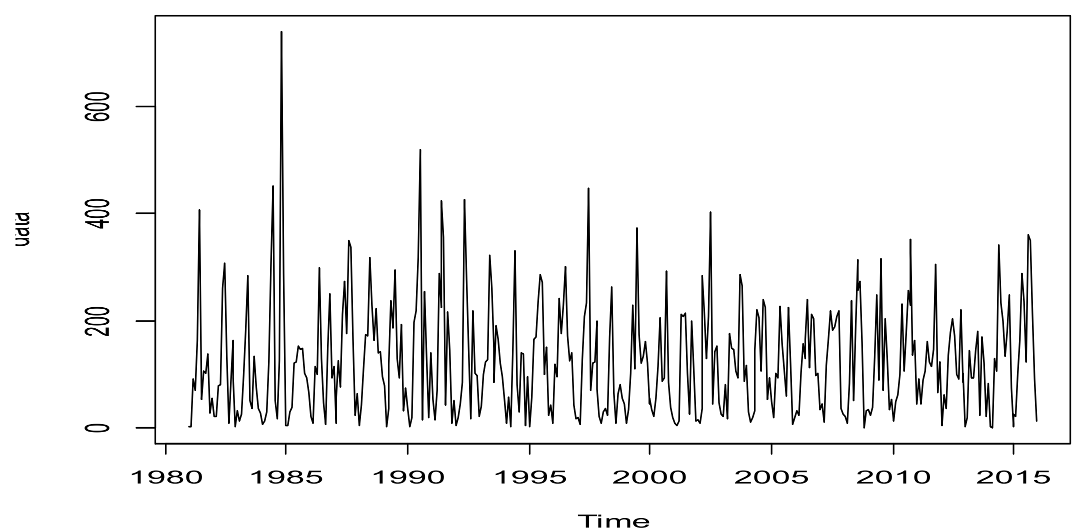

Fig. 1. Frequency of monthly rainfall in Osun State

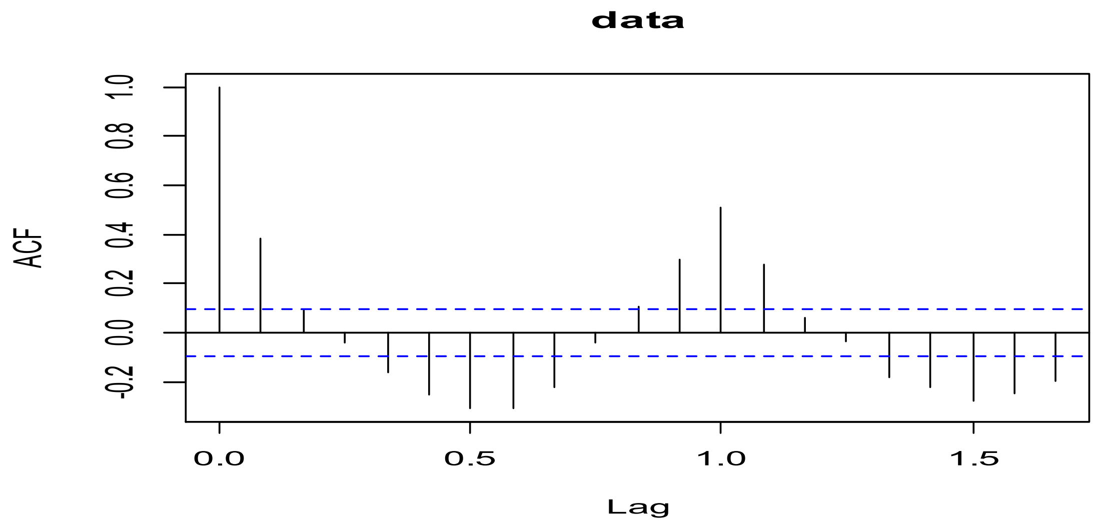

Fig. 2. ACF plot of the frequency of monthly rainfall in Osun State

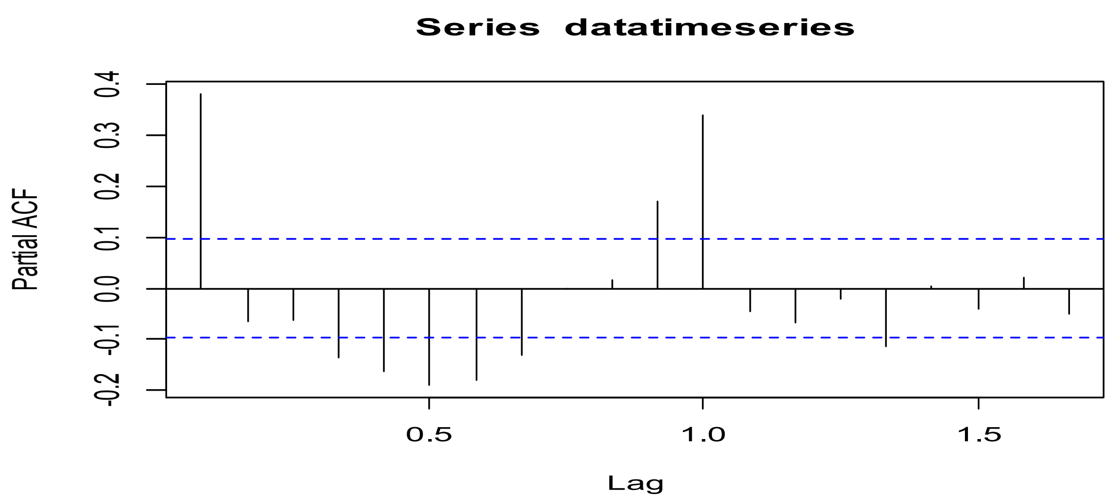

Fig. 3. PACF plot of the frequency of monthly rainfall in Osun State

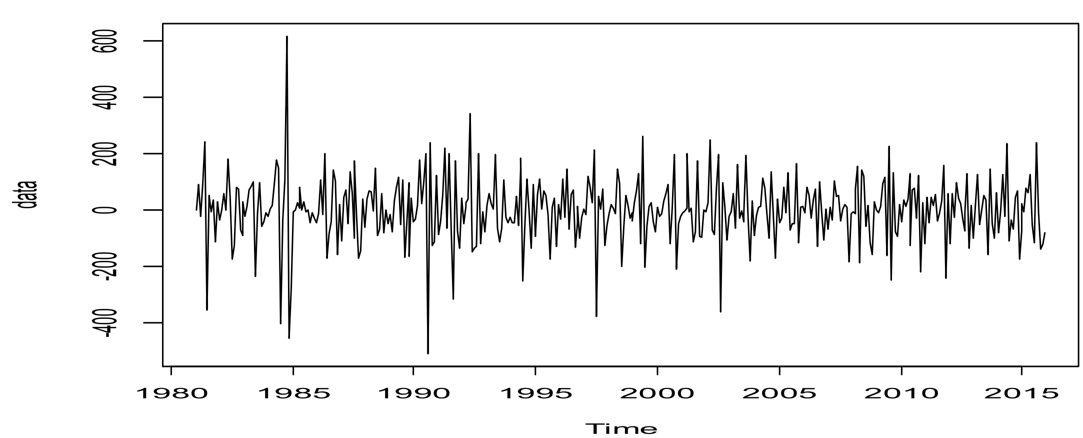

Fig. 4. Plot of diff (1) of monthly rainfall in Osun State

Table 1. Summary of some estimated candidate SARIMA models

Candidate model	Coefficient	S.E	Sigma^2	Log likelihood	AIC
AR1	1.1418	0.0200			
AR2	-0.1553	0.0213			
MA1	-1.0000	0.0131	5697	-2354.59	4723.17
SAR1	0.1062	0.0233			
SAR2	0.0640	0.0231			
SMA1	-0.9292	0.0322			
AR1	0.1897	0.3071			
MA1	-0.0374	0.3127			
SAR1	0.0893	0.0593	5775	-2355.57	4721.14
SMA1	-0.9121	0.0388			
AR1	-0.3062	0.8815			
MA1	0.4603	0.8803			
MA2	0.0794	0.1361	5771	-2355.5	4723.01
SAR1	0.0910	0.0594			
SMA1	-0.9135	0.0387			
AR1	0.0645	0.3733			
AR2	0.0176	0.3882			
MA1	0.0868	0.3992	5738	-2355.1	4724.20
SAR1	0.5672	0.3728			
SMA1	-1.3955	0.3971			
SMA2	0.4279	0.3571			
AR1	0.2082	0.3220			
MA1	-0.0586	0.3293			
SAR1	0.1020	0.0591	5742	-2354.99	4721.97
SAR2	0.0618	0.0571			
SMA1	-0.9317	0.0420			
AR1	0.1698	0.3141			
MA1	-0.0181	0.3184			
SAR1	0.5540	0.3822	5739	-2355.1	4722.2
SMA1	-1.3815	0.4055			
SMA2	0.4152	0.3643			

Table 2. Unit root and stationarity tests of Osun state monthly rainfall

Test	Test statistics	Lag order	p-value
Dickey-Fuller	-13.626	0	0.01

Table 3 depicts the Augmented Dickey-Fuller Test, the hypothesis;

H_0: The series is unit root non stationary vs H_1: the series is unit root stationary P-value < 0.05, indicating a strong evidence against the null hypothesis at 5% level of significance.

Table 3. Unit root and stationarity tests of Osun state monthly rainfall

Test	Test Statistics	Lag Order	p-value
Dickey-Fuller	-12.085	1	0.032

4. FORECASTING WITH THE FITTED MODEL

One of the objectives of fitting SARIMA model to data is to be able to forecast its future values. The model that best fits the data is SARIMA $(1,0,1)\times(1,1,1)_{12}$. The fitted model is therefore used to forecast for four [4] years i.e. January 2016 – December 2020.

The model validation is concerned with checking the residual of the model to determine if the model contains any systematic pattern which can be removed to improve on the selected model may appear to be the best among a number of models considered it become necessary to do diagnostic checking to verify that the model is adequate.

The plots Fig. 7 comprise of the time plot of the residuals, ACF and the PACF plot of the residuals respectively. The plot clearly shows

Table 4. Forecast of monthly rainfall in Osun state from 2016 to 2020

Month/years	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2016	15.36211	-82.25182	112.9768	-133.925716	164.6507
Feb 2016	30.07109	-68.75338	128.8956	-121.067900	181.2101
Mar 2016	87.03243	-11.83149	185.8963	-64.166884	238.2317
Apr 2016	135.47231	36.60574	234.3389	-15.731056	286.6757
May 2016	217.30619	118.43930	316.1731	66.102330	368.5101
Jun 2016	221.34064	122.47370	320.2076	70.136708	372.5446
Jul 2016	180.92120	82.05425	279.7881	29.717249	332.1251
Aug 2016	156.42678	57.55984	255.2937	5.222836	307.6307
Sep 2016	215.29227	116.42532	314.1592	64.088317	366.4962
Oct 2016	202.80709	103.94012	301.6741	51.603108	354.0111
Nov 2016	68.80167	-30.06542	167.6688	-82.402493	220.0058
Dec 2016	53.45305	-45.41474	152.3208	-97.752177	204.6583
Jan 2017	22.87171	-77.59039	123.3338	-130.771816	176.5152
Feb 2017	33.48983	-67.02443	134.0041	-120.233470	187.2131
Mar 2017	84.42676	-16.09105	184.9446	-69.301961	238.1555
Apr 2017	136.37262	35.85437	236.8909	-17.356773	290.1020
May 2017	207.18579	106.66748	307.7041	53.456301	360.9153
Jun 2017	220.21331	119.69499	320.7316	66.483810	373.9428
Jul 2017	182.39962	81.88129	282.9179	28.670109	336.1291
Aug 2017	149.68948	49.17116	250.2078	-4.040030	303.4190
Sep 2017	212.07930	111.56098	312.5976	58.349791	365.8088
Oct 2017	199.83608	99.31773	300.3544	46.106535	353.5656
Nov 2017	67.77333	-32.74517	168.2918	-85.956453	221.5031
Dec 2017	58.44459	-42.07480	158.9640	-95.286561	212.1758
Jan 2018	23.24208	-78.49374	124.9779	-132.349433	178.8336
Feb 2018	34.57036	-67.20853	136.3493	-121.087027	190.2277
Mar 2018	83.61879	-18.16338	185.4010	-72.043613	239.2812
Apr 2018	135.06850	33.28591	236.8511	-20.594546	290.7316
May 2018	202.12505	100.34239	303.9077	46.461902	357.7882
Jun 2018	219.42295	117.64028	321.2056	63.759786	375.0861
Jul 2018	186.30029	84.51762	288.0830	30.637129	341.9635
Aug 2018	136.99492	35.21224	238.7776	-18.668248	292.6581
Sep 2018	203.84346	102.06079	305.6261	48.180291	359.5066
Oct 2018	199.05576	97.27305	300.8385	43.392541	354.7190
Nov 2018	66.53447	-35.24841	168.3174	-89.129019	222.1980
Dec 2018	61.65131	-40.13266	163.4353	-94.013840	217.3165
Jan 2019	23.94533	-78.40262	126.2933	-132.582346	180.4730
Feb 2019	35.09976	-67.27187	137.4714	-121.464135	191.6637
Mar 2019	83.59608	-18.77772	185.9699	-72.971136	240.1633
Apr 2019	135.20616	32.83206	237.5803	-21.361512	291.7738
May 2019	201.22227	98.84812	303.5964	44.654520	357.7900
Jun 2019	219.49075	117.11659	321.8649	62.922989	376.0585
Jul 2019	186.99776	84.62360	289.3719	30.430001	343.5655
Aug 2019	135.52501	33.15085	237.8992	-21.042751	292.0928
Sep 2019	203.03491	100.66076	305.4091	46.467149	359.6027
Oct 2019	199.01391	96.63972	301.3881	42.446099	355.5817
Nov 2019	66.56296	-35.81142	168.9374	-90.005152	223.1311
Dec 2019	62.48961	-39.88596	164.8652	-94.080322	219.0595
Jan 2020	24.25371	-78.62394	127.1314	-133.084081	181.5915
Feb 2020	35.43322	-67.46663	138.3331	-121.938523	192.8050
Mar 2020	83.76050	-19.14149	186.6625	-73.614511	241.1355
Apr 2020	135.35698	32.45470	238.2593	-22.018483	292.7324
May 2020	201.04265	98.14032	303.9450	43.667115	358.4182
Jun 2020	219.66535	116.76301	322.5677	62.289804	377.0409

Month/years	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jul 2020	187.51743	84.61510	290.4198	30.141889	344.8930
Aug 2020	134.83001	31.92767	237.7323	-22.545539	292.2056
Sep 2020	202.67424	99.77190	305.5766	45.298687	360.0498
Oct 2020	199.17805	96.27568	302.0804	41.802451	356.5537
Nov 2020	66.70668	-36.19591	169.6093	-90.669246	224.0826
Dec 2020	62.98184	-39.92203	165.8857	-94.396044	220.3597

that the residuals appear to be randomly scattered, no evidence exists that the error terms are correlated with one another as well as no evidence of existence of an outlier. The residuals or errors are therefore conceived as an independently and identically distributed (i.i.d) sequence with a constant variance and a zero mean. The ACF and the PACF plot of the residuals show no evidence of a significant spike indicating that the residuals seem to be uncorrelated. Therefore, the SARIMA (1,0,1) $(1,1,1)_{12}$ model appears to fit well and can be used to predict the frequency of rainfall in Osun State Nigeria.

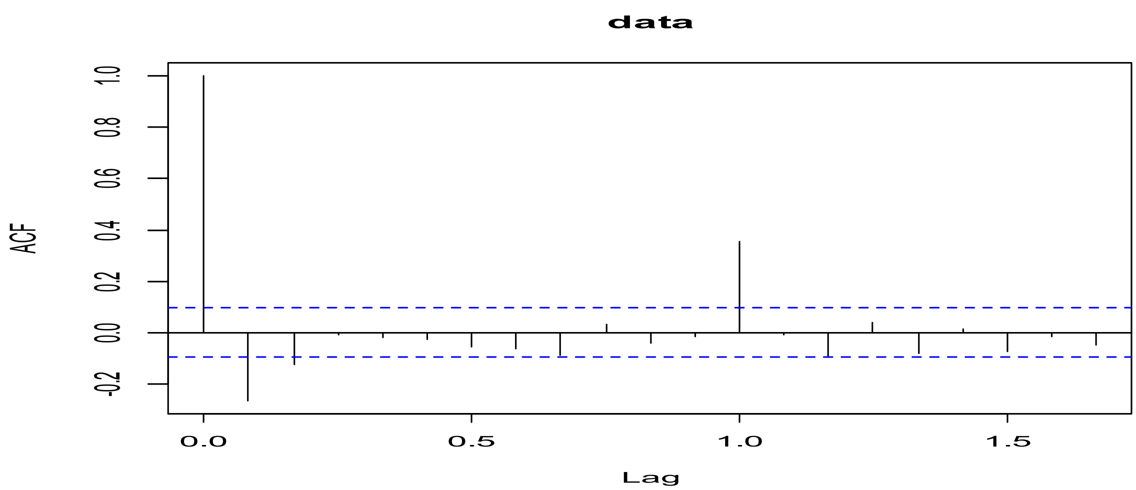

Fig. 5. ACF plot of diff (1) of the frequency of monthly rainfall in Osun State

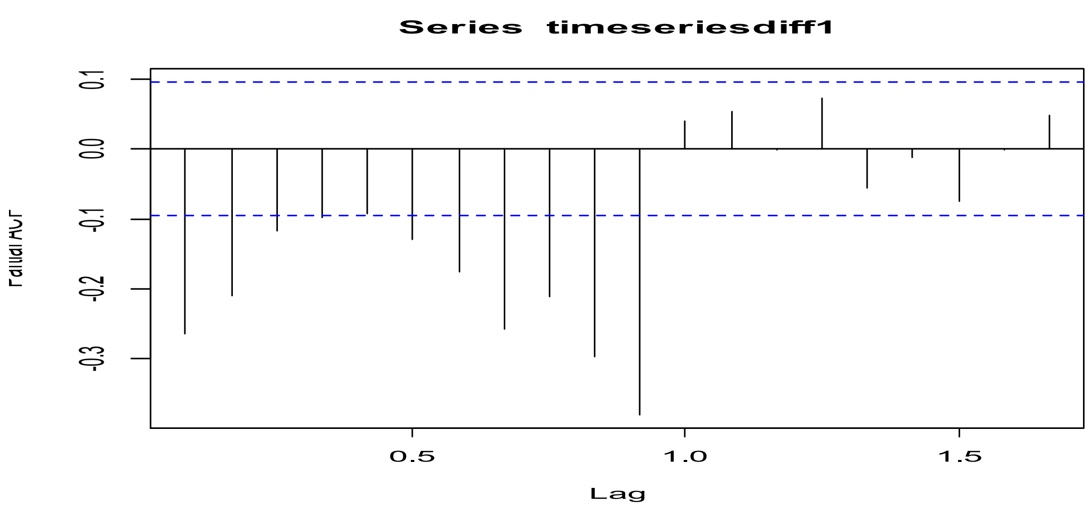

Fig. 6. PACF plot of diff (1) of the frequency of monthly rainfall in Osun State

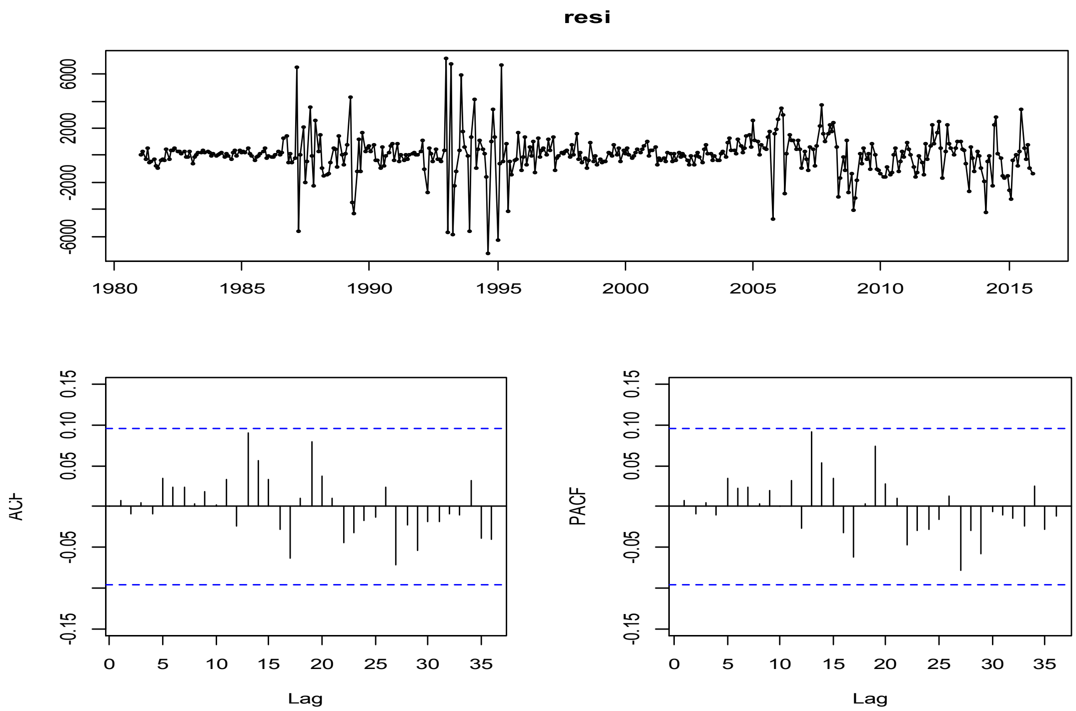

Fig. 7. Model verification plot

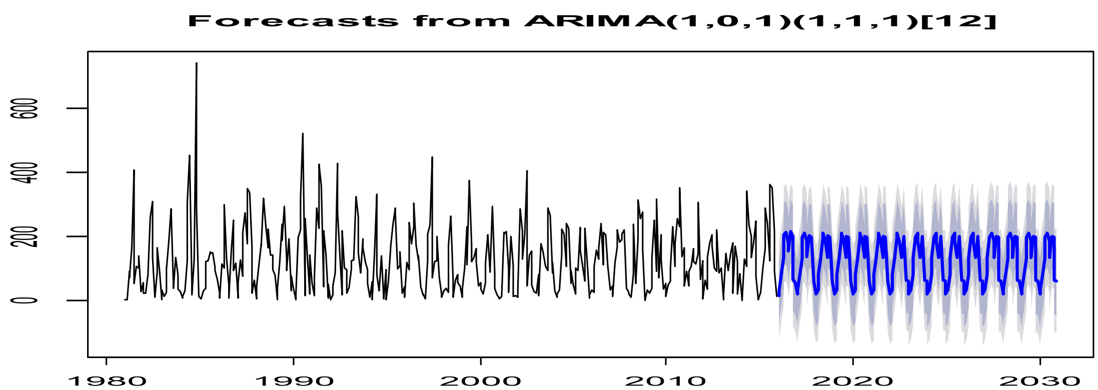

Fig. 8. Forecast plot of monthly rainfall in Osun state from 2016 to 2020

Table 5. Fitted SARIMA (1,0,1)(1,1,1)$_{12}$ model

Sigma^2	Log likelihood	AIC
5771	-2355.5	4723.01

5. CONCLUSION

In this study, the frequency of monthly rainfall from 1981 to 2015 was analysed using seasonal time series modeling approach. The plot of the original data shows that the time series is stationary and has evidence of seasonality, Augmented Dickey Fuller test confirmed the stationarity claim. Seasonal differencing was done to remove the seasonal effect and SARIMA (1,0,1)(2,1,1)$_{12}$ model of was obtained, this resulted was found to be statistically insignificant and this consequently led to a new SARIMA (1,0,1)(1,1,1)$_{12}$ that best fit the data and was used to make forecast. The study has revealed periods of high and low rainfall in Osun State, the rainfall forecast is important for future plans regarding agriculture, and to commodity traders within the stock market, and hence the impact on the economy of Osun state.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Schwarz G. Estimating the dimension of a model. Annals of Statistics. 1978;6:461–464.
2. Adejuwon JO, Balogun EE, Adejuwon SA. On the annual and seasonal pattern of rainfall fluctuations in sub-Saharan West Africa. International Journal of Climatology. 1990;10(8):39–848.
3. Afahakan IE, Udofia KM, Umoren MA. Analysis of rain rate and rain attenuation for earth-space communication links over Uyo, Akwa-Ibom State. Nigerian Journal of Technology (NIJOTECH). 2016;35(1):137–143.
4. Ajayi GO, Ofoche EBC. Some tropical rainfall rate characteristics at Ile-Ife for microwave and millimeter wave applications. Journal of Climate and Applied Meteorology. 1984;23(4):562–567.
5. Akpanta AC, Okorie IE, Okoye NN. SARIMA modelling of the frequency of monthly rainfall in Umuahia, Abia State of Nigeria. American Journal of Mathematics and Statistics. 20155(2):82–87.
6. Edwin A, Martins O. Stochastic characteristics and modelling of monthly rainfall time series of Ilorin, Nigeria, *Open* Journal of Modern Hydrology. 2014;4:67–69.
7. Etuk HE, Moffat UI, Chims EB. Modelling monthly rainfall data of Port Harcourt, Nigeria by seasonal box-Jenkins method. International Journal of Sciences. 2013; 7(2):12–19.
8. Ibiyemi TS, et al. Rain rate and rain attenuation prediction with experimental rain attenuation efforts in south-western Nigeria. 20th In Telecommunications Forum (TELFOR); 2012.
9. Obiyemi OO, Ibiyemi TS. Experimental investigation of rainfall effect on digital satellite television reception in Nigeria: Initial results. In Adaptive Science & Technology (ICAST), IEEE 6th International Conference; 2014.
10. Obiyemi OO, Ojo JS, Ibiyemi TS. Performance analysis of rain rate models for microwave propagation designs over tropical climate. Progress In Electromagnetics Research M. 2014;39:115–122.
11. Okorie IE, Akpanta AC. Threshold excess analysis of Ikeja monthly rainfall in Nigeria. International Journal of Statistics and Applications. 2015;1(5):15–20.
12. Omotosho T, et al. One year results of one minute rainfall rate measurement at Covenant University, Southwest Nigeria. In Space Science and Communication (Icon Space), IEEE International Conference; 2013.
13. Semire FA, et al. Analysis of cumulative distribution function of 2-year rainfall measurements in Ogbomoso, Nigeria. International Journal of Applied Science and Engineering. 2012;10(3):171–179.
14. Akaike HA. New look at the statistical model identification. Annals of statistics: IEEE Trans. Automat. Contr. 1974;19:716–723.
15. Box GEP, Jenkins GM. Time series analysis, forecasting and control. San Francisco: Holden Day; 1976.
16. Ljung GM, Box GEP. On a measure of lack of fit in time series models. Biometrika. 1978;65:297.
17. Dickey DA, Fuller WA. Autoregressive time series with a unit root. Journal of American Statistical Association. 1979;74:427–431.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle3.com/review-history/50701