Journal of the Nigerian Association of Mathematical Physics
Volume34,(March, 2016), pp167 – 176

Model for Forecasting Rainfall in Nigeria

Obisesan K.O

Department of Statistics, University of Ibadan, Nigeria.

Abstract

Forecasting rainfall has only been done with mere comparison with cases from different locations. Much has not been done in obtaining appropriate statistical model in evaluating rainfall events in Nigeria.

In this work, Seasonal Autoregressive Integrated Moving Average (SARIMA)is used by considering monthly data for six stations: Oyo, Enugu, Kwara, Kaduna, Rivers and Tarabafor the period 1994-2014. Six SARIMA models were developed for the abovenamed stations as follows:(1,0,3)x(1,0,1)12, (3,0,3)x(1,0,1)12, (1,0,1)x(1,1,1)12,(1,0,1)x(2,0,1)12, (1,0,2)x(1,0,1)12, (3,0,3)x(1,0,1)12 respectively.

Basic exploratory data analysis was carried out, the normality of the series was also tested, the time plot of the monthly rainfall in each station was checked for trend and seasonal pattern in the series, the timeplot shows no obvious trend or seasonality, the Augmented Dickey-Fullertest, Kwaiatkowski-Phillips-Schmidt-Shin test, Phillips-Perron Unit RootTest were used to establish the stationarity in the series and Hylleberg, Engle, Granger and Yoo(HEGY's Test) was used to test the seasonal unit rootin the series, the test shows that all but Kwara series were seasonally stationary. The Kwara series was seasonally differenced. The ACF and PACFplot were used to identify the possible order of the models, while the modelswith the lowest Akaike Information Criterion (AIC) were selected. Severaldiagnostic statistics and plots of the residuals were used to examine thegoodness of fit of the selected model to the data.

These models were used to forecast the Monthly rainfall series for the fouryears (2015 to 2018)and the forecast shows a decline in pattern of rainfall series in the six states in the short run. Delayed rainfall in some partand early cessation of rainfall in some other parts expected. This corroborates the seasonal rainfall prediction (SRF) made by Nigerian MeteorologicalAgency(NIMET) for 2015.

Key words: Time Series, Forecasting, SARIMA,Rainfall, Nigeria

1.0 Introduction

Climate change is a change in the statistical distribution of weather patterns when that change lasts for an extended period of time (i.e., decadesto millions of years). Climate change may refer to a change in averageweather conditions, or in the time variation of weather around longer-termaverage conditions (i.e., more or fewer extreme weather events). Climatechange is caused by factors such as biotic processes, variations in solar radiation received by Earth, plate tectonics, and volcanic eruptions. Certainhuman activities have also been identified as significant causes of recent climate change, often referred to as "global warming". Humans are exposedto climate change through changing weather patterns (temperature, precipitation, sea-level rise and more frequent extreme events) and indirectlythrough changes in water, air and food quality and changes in ecosystems,agriculture, industry and settlements and the economy.

According to a literature assessment [1], the effects of climate change to date have been small, but are projected to progressively increase in all countries and regions. Extreme rainfall and extremeclimate events and other climatic fluctuations have been shown to have ahigh influence on the social and economic activities of the country and theperformance of the country's economy. The economic estimates of the impact of climate change are categorized into the market and non-market impacts, the market impacts includes effect of climate on sensitive sectors like agriculture, forestry, fisheries and tourism, the non-market impacts covers the effects of climate change on health, leisure activities & human settlements.

Corresponding author: Obisesan K.O, E-mail:Obidairo@gmail.com, Tel.: +2348052416373

A study of the rainfall pattern in Nigeria is necessary, there is need formore accurate forecasting techniques to be applied in predicting climaticpatterns. Due to the fact that rainfall estimates are important componentof water resources applications, an accurate estimate of rainfall is needed.

This also concerns producing valid estimates using appropriate methods andin order to develop a comprehensive solution to the forecasting problem,including addressing the issue of uncertainty in predictions, more so manymethods of statistical analysis make use independent observations. Hence, to tackle these problems appropriate statistical models must be developed.

2.0 Review of Literature

Prediction of rainfall series is a very challenging task especially in this modern world where we are faced with major environmental problem of globalwarming which has rendered the previously employed methods to redundant.

Harvey et al.[2] studied how patterns of rainfall correlate with generalweather conditions and frequency of the cycles of rainfall. They used rainfalldata from Brazil for a particular region which often suffers from drought toaccess the cyclical behavior of rainfall, they observed that cyclical components are stochastic. Chiew et al.[3] also conducted a comparison of six rainfall-runoff modeling approaches to simulate daily, monthly and annualflows in eight unregulated catchments in Australia. They concluded that atime-series approach can provide adequate estimates of monthly and annualyields in the water resources of the catchments.

Al-Ansari et.al [4,5], examined therainfall record of all the Jordanian Badia stations, for the period 1967 -1998to determine periodicity and interrelations between stations using powerspectral, harmonic analysis, and correlation coefficient techniques. Theyused an ARIMA model to forecast rainfall trends in individual stations upto 2020, their results showed that the rainfall intensity has been decreasingwith time for most of the stations. Amha[6] studied the monthly rainfall in Tigray region based on Mkellestation,he employed a univariate Box-Jenkins method to analyze rainfall inthe region and found that SARIMA model is suitable for forecasting futurevalues of monthly rainfall data and used this model to forecast 12-monthrainfall pattern in the study area.Further he concluded that there is no tendency of decreasing or increasing pattern of monthly rainfall over the forecastperiod from January 2010 to September 2011. Saleh Zakaria et al.[7] also used the Box-Jenkins methodology to build Autoregressive IntegratedMoving Average (ARIMA) models for weekly rainfall data from four rainfall stations in the North West of Iraq: Sinjar, Mosul, Rabeaa and Talafar for the period 1990-2011, four ARIMA models were developed for theabove stations as follow: $(3; 0; 2)x(2; 1; 1)_{30}$; $(1; 0; 1)x(1; 1; 3)_{30}$; $(1; 1; 2)x(3; 0; 1)_{30}$ and $(1; 1; 1) x(0; 0; 1)_{30}$ respectively, The performance of the resulting successful ARIMA models were evaluated using the data year (2011),these models were used to forecast the weekly rainfall data for the up-coming4years (2012 to 2016). Abdul-Aziz et al.[8] also modelled andforecasted rainfall patterns in Ghana using seasonal ARIMA process, Theresults showed that rainfall pattern in Ashanti region significantly changesover time. Therefore, the forecast figures for the months show an increase in the rainfall figures for the subsequent year(s).

Etuk et al. [9] also modelled the monthly rainfall pattern in Port-Harcourt, Nigeria, using seasonal ARIMA model. $(5, 1, 0) x (0; 1; 1)_{12}$ theadequacy of the modelled was established and found the seasonality measurein SARIMA to be highly useful in measuring rainfall.Obisesan and Afolabi[10] used rainfall measurements from 1971-2012 by engaging generalized linear model for evaluating rainfall series however they did not consider the structure of the series hence the need for SARIMA models.

2.1 Objectives

The aim of this work is to analyze rainfall pattern in 6 selected states from1994 to 2014 in each of the six geo-political zone in Nigeria, using the Seasonal Autoregressive Moving Average (SARIMA) methodology and to forecast rainfall series for the respective states from 2015-2018.

3.0 Methodology

3.1 Autoregressive Integrated Moving Average [P,D,Q]

Many time series encountered in practice exhibit nonstationarybehavior.Usually, the nonstationarity is due to a trend, a change in the local mean,or seasonal variation. Since the Box-Jenkins methodology is for stationarymodels only, we have to make some adjustments before we can model thesenonstationaryseries.We use one of two methods for reducing a nonstationary series with trendto a stationary series (without trend):

1. Use the first difference of the series,$W_t = X_t – X_{t-1}$: Note that this canbe rewritten as
   $W_t = (1-B)X_t$ ………………………………………….. (1)

A more general form of this equation is:

$\emptyset\_p(B)(1- B)_d X_t = \theta\_q(B)a_t$ ………………………………………….. (2)

Where d is the order of differencing.This is known as the ARIMA(p,d,q)model.

2. Fit a least squares trend and Wt the Box-Jenkins model to the residuals.

If the model exhibits an occasional change of mean, first differences willresult in a stationary model.

3.2 Seasonal Time Series

3.2.1 Seasonality

Seasonality in time series is a regular pattern of changes that repeats over s time period where s denotes the number of time period until the patternrepeat again.

In seasonal arima model, seasonal AR and MA term predict Xtusingdata values and errors at times with lags that are multiples of s (the spanof seasonality).

- With monthly data (s=12), a seasonal first order autoregressive modelwould use x_{t-12} to predict x_t.
- A seasonal second order autoregressive model would use x_{t-12} and x_{t-24} to predict xt.
- A seasonal first order MA(1) model with s=12 would use e_{t-12} as a predictor. A seasonal second order MA(2) model use e_{t-12} and e_{t-24}.

3.2.2 Differencing

Almost by definition, it may be necessary to examine differenced data whenwe have seasonality. Seasonality usually causes the series to be non-stationarybecause the average value at some particular times within the seasonal span (month for example) may be different than the average values at other times.

- Seasonal differencing is denoted as a difference between a value and a valuewith lag that is multiple of s.
- With s=12, which may occur with monthly data, seasonal difference is$(1- B_{12})x_t = x_t x_{t-12}$:

The difference from the previous year may be about the same for each monthof the year giving us a stationary series. Seasonal differencing removes seasonal trend and also can also get rid of a seasonal random walk type ofnon-seasonality.

- Non-seasonal differencing: If trend is present on the data, we may alsoneed non-seasonal differencing, often (not always) a first difference (non-seasonal) with detrend" the data. That is we use (1-B)xt = x_t- x_{t-1} inthe presence of trend.

3.3 Seasonal Autoregressive Integrated Moving Average of Order (P,D,Q)X(P,D,Q)S

The Seasonal Autoregressive integrated moving average models incorporatesboth the seasonal and non-seasonal factor in a multiplicative model. Oneshorthand notation for the model is:

ARIMA (p; d; q) x (P;D;Q)s, with 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 ofrepeating seasonal pattern.

Note that on the left side of equation (1) the seasonal and non-seasonal ARcomponents multiply each other, and on the right side of equation (1) theseasonal and non-seasonal MA components multiply each other.

$$(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_2 B^p)(1-\beta_1 B^s-\beta_2 B^{2s}-\cdots-\beta_p B^{ps})(1-B)^d(1-B^S)^D y_t$$

$$= c + (1 - \Phi_1 B - \Phi_2 B^2 - \cdots - \Phi_2 B^q)(1 - \theta_1 B^s - \theta_2 B^{2s} - \cdots - \theta_q B^{qs})\epsilon_t$$

where $(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_2 B^p)$= autoregressive part of order p AR(p),

$(1 - \beta_1 B^s - \beta_2 B^{2s} - \cdots - \beta_p B^{ps})$ =seasonal autoregressive part of order P AR(P),

$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_2 B^p)$= moving average part of order q MA(q),

$(1 - \theta_1 B^s - \theta_2 B^{2s} - \cdots - \theta_q B^{qs})$= seasonal moving average part of order Q MA(Q),

$(1 - B)^d$= differences of order d,

$(1 - B^S)^D$=seasonal differences of order D

3.4 Test for Stationarity

3.4.1 Augmented Dickey-Fuller Test

The Augmented Dickey-Fuller Test (ADF) is a test for unit root in a timeseriessample.It is an augmented version of the Dickey-Fuller test for largerand more complicated set of time series models.The Augmented Dickey-Fuller (ADF) statistics, used in the test, is a negative number. The morenegative it is, the stronger the rejection of the hypothesis that there is a unit roots at some level of confidence.

Testing Procedure:
The testing procedure for the ADF test is the same as for the [Dickey-Fullertest] but it is applied to the model.

$$\Delta y_t = \alpha + \beta_t + \gamma y_{t-1} + \delta_t \Delta_{y_t - 1} + \cdots + \delta_{p-1} \Delta_{y_{t-p+1}} + \epsilon_t \quad \ldots (3)$$

Where α is a constant, $\beta$the coefficient on a time trend and p the lag orderof the autoregressive process.Imposing the constraintsα= 0 and β = 0 test equation making y_{t-1} endogenous and thus invalidating the Dickey-Fuller t-test.While the augmented Dickey-Fuller test addresses this issueby introducing lag of k, yt as regressor in the test equation, the Phillips-Perron test makes a non-parametric correction of the t-test statistics.

3.4.2 Kwaiatkowski-Phillips-Schmidt-Shin

The Kwaiatkowski-Phillips-Schmidt-Shin (KPSS) test are used for testing anull hypothesis that an observable time series is stationary around a deterministic trend. Some of these models employ several Durbin-Watson type finite sample tests for unitrootsand others employ a test of the null hypothesisthat an observable series is trend stationary(stationary around a deterministic trend). The series is expressed as the sum of deterministic trend, randomwalk, and stationary error, and the test is the Lagrange multiplier test of theintended to complement unit root tests,such as the Dickey-Fuller tests. Bytesting both the unit root hypothesis and the stationary hypothesis,onecandistinguish series that appear to be stationary, series that appear to havea unit root, and series for which the data (or the tests) are not sufficientlyinformative to be sure whether they are stationary or integrated.
The KPSS test is given as:

$$KPSS = T^{-2} \sum \hat{S}_t^2 / \hat{\sigma}_T^2(q)$$

Where $\hat{S}_t = \sum_{t-1}^{T} \hat{e}_t$ and the long run variance is computed as $\hat{\sigma}_T^2(q) = C_o + \sum_{j=1}^{q} W_j(q) C_j$, with the condition that C_j is the j^{th} order sample autocovariance of y_t and $W_j(q)$ are the Bartlett window weights given by $W_j(q) = 1 - j/(q+1)$ for $q < T$. ………….. (4)

3.4.3 Model Identification

The identification stage is the most important and also the most difficult:it consists to determine the adequate model from ARIMA family models.The most general Box-Jenkins model includes difference operators, autoregressive terms, moving average terms, seasonal difference operators, seasonal autoregressive terms, and seasonal moving average terms.Thisphaseis founded on the study of autocorrelation and partial autocorrelation. The first step in developing a Box-Jenkins model is to determine if theseries is stationary and if there is any significant seasonality that needs to bemodeled. At the model identification stage, our goal is to detect seasonality,if it exists, and to identify the order for the seasonal autoregressive andseasonal moving average terms.
Once stationarity and seasonality have been addressed, the next step isto identify the order (the p and q) of the autoregressive and moving averageterms. The primary tools for doing this are the autocorrelation plot and thepartial autocorrelation plot. The sample autocorrelation plot and the sample partial autocorrelation plot are compared to the theoretical behavior of these plots when the order is known.

3.4.4 Model Estimation

There are different ways of estimating the parameter of any model identifiedfrom the time plot of the partial autocorrelation.
However, the main approaches to fitting Box-Jenkins models are non-linear least squares and maximum likelihood estimation. Maximum likelihood estimation is generally the preferred technique.
In maximum likelihood methods, the likelihood function is maximized inorder to obtain the parameter estimates.The likelihood of a set of data is theprobability of obtaining that particular set of data given its distribution.

3.4.4 Diagnostics

Model diagnostics for ARIMA models is similar to model validation for non-linear least squares fitting. After fitting the time series model it is importantto check the adequacy of the model this can be achieved in various ways.
This is an iterative process used to assess model adequacy by checkingwhetherthe model assumption is that$\in_t$'s as uncorrelated random shock(errors) with zero mean and constant variance.Model adequacy isconcerned with the analysis of the residual $\in_t$. Since this residual series is theproduct of parameter estimation, the model diagnostics checking(evaluation)is usually contained in the estimation phase of any time series package.

Table 1: **Summary of rainfall series by states from(1994-2004)**

Rainfall(mm)	Enugu	Kaduna	Kwara	Oyo	Rivers	Taraba
Mean	153.6492	102.8925	106.0175	113.8798	192.6937	88.64643
Median	156.7000	56.45000	86.00000	101.8500	169.8000	62.05000
Maximum	443.2000	546.8000	510.8000	385.8000	795.0000	404.2000
Minimum	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
Std.Dev	131.7984	121.9125	104.6373	98.25640	145.5597	95.25801
Skewness	0.258242	1.059752	0.841634	0.569657	0.641969	0.811596
Kurtosis	1.739434	3.241149	3.079849	2.398848	3.093680	2.707568
Jarque-Bera	19.48571	47.77970	29.81754	17.42392	17.40139	28.56280
Probability	0.000059	0.000000	0.000000	0.000165	0.000166	0.000001
Sum	38719.60	25928.90	26716.40	28697.70	48558.80	22338.90
Sum Sq.Dev.	4360075.	3730530.	2748190.	2423235.	5318097.	2277596.
Observation	252	252	252	252	252	252

4.0 Analysis

The Table-4.1 above shows the summary of monthly rainfall.Fromthetable the mean rainfall of Enugu, Kaduna, Kwara, Oyo, Rivers and Tarabaare 153.6492, 102.8925, 106.0175, 113.8798, 192.6937 & 88.64643 millimeters(mm) respectively, the number of observations for each station is (n=252).The table also shows the median, maximum and minimum rainfall values inthe respective locations.

The Jarque-Bera statistic in the table is used to test the null hypothesisthat the variables have a normal distribution and comparing the probabilityvalues of each of the variables respectively with 0.05 we can conclude thatthe variables are not normally distributed, attesting to the fact that environmental data are always positively skewed.

4.1 Time Plot

The Visual inspection of the time plot is the first step in the analysis of timeseries data, this helps to detect trend and seasonality in the data.

The inspection of the figure below shows no obvious trend or seasonality,therefore we need to test for stationarity.

4.2 Unit Root Test

4.2.1 Augmented Dickey-Fuller test

The Augmented Dickey-Fuller test is used to test for stationarity of theseries, the test was carried on series to establish the stationary.

Hypothesis:

Ho:There is unit root

H_1:The series is stationary

Table 2: **ADF test with test-statistic and p-value**

State	Intercept	p-value
Oyo	-6.8268	0.01
Enugu	-6.299	0.01
Kwara	-7.7547	0.01
Kaduna	-9.4545	0.01
Rivers	-7.3079	0.01
Taraba	-8.5875	0.01

Decision Rule:Reject Ho if p-value<0.05, otherwise we fail to reject.

Conclusion: Since 0.01<0.05 we reject Ho and conclude that the series arestationary.

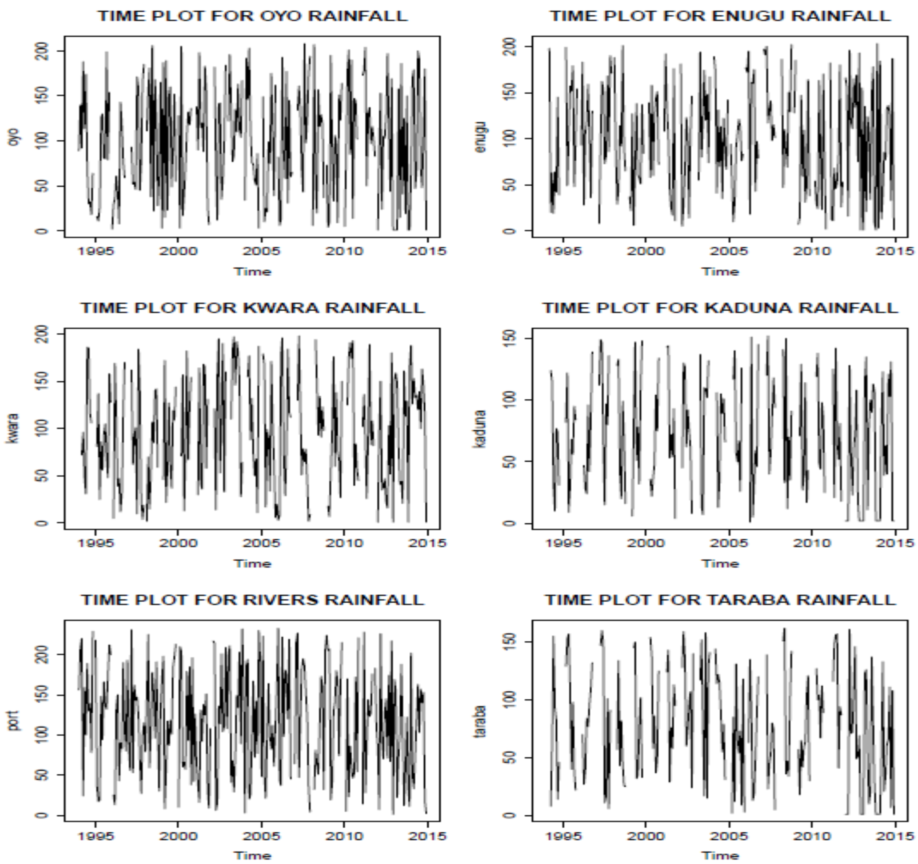

Figure 1: Time Plot of Rainfall series(1994-2014)

4.2.2 Kwaiatkowski-Phillips-Schmidt-Shin Test

The Kwaiatkowski-Phillips-Schmidt-Shin test was also carried out to testthe trend stationarity of the series.
Hypothesis:
Ho:There is trend stationary
H1:The series is not trend stationary

Table 3: KPSS test with test-statistic and p-value

State	Intercept	p-value
Oyo	0.0336	0.1
Enugu	0.0299	0.1
Kwara	0.0503	0.1
Kaduna	0.0195	0.1
Rivers	0.0302	0.1
Taraba	0.0257	0.1

Decision Rule:Reject Ho if p-value<0.05, otherwise we fail to reject.
Conclusion: Since 0.1>0.05 we accept Ho and conclude that the series aretrend stationary

4.2.3 Phillips-Perron Unit Root Test

The Phillips-Perron Unit Root Test is also used to test unit root in a series.
Hypothesis:
Ho:There is unit root
H1:The series is stationary

Table 4: **Phillips-Perron Unit Root Test**

State	Dickey-Fuller Z(alpha)	p-value
Oyo	-242.495	0.01
Enugu	-256.8357	0.01
Kwara	-178.5706	0.01
Kaduna	-150.1758	0.01
Rivers	-227.5671	0.01
Taraba	-149.7511	0.01

Decision Rule:Reject Ho if p-value<0.05, otherwise we fail to reject.
Conclusion: Since 0.01<0.05 we reject Ho and conclude that the series isstationary

4.2.4 Seasonal Unit Root Test

Since we have been able to establish the non-seasonal stationary of eachseries, hence we need to test for seasonal unit root of the series.

4.2.5 Hylleberg,Engle,Granger and Yoo(HEGY's Test)

The Hylleberg, Engle, Granger and Yoo(HEGY's Test) is used to test seasonal unit root of the series.
Hypothesis:
Ho:There has seasonal unit root
H1:The series is seasonal stationary
The F-test is used to test the seasonal unit root test.

Series	F-cal($F_{\pi_3\pi_4}$)	Critical values(@ 5% α level)
Oyo	5.1077	3.07
Enugu	3.7321	3.07
Kwara	2.4500	3.07
Kaduna	4.0223	3.07
River	9.1948	3.07
Taraba	10.3500	3.07

Decision Rule: Reject Ho if Fcal>Ftab, otherwise we fail to reject.
Conclusion: Since the respective Fcal>Ftab, we can conclude that theseries(Oyo, Enugu, Kaduna, River and Taraba) are seasonal stationary butforKwaraFcal<Ftab therefore accept Ho and conclude that series Kwarais seasonally non-stationary.

4.3 Differencing

The various test carried out shows that the series are stationary and theHEGY's test shows that Oyo, Enugu, Kaduna, River and Taraba are seasonally stationary but the series (Kwara) is seasonally non- stationary therefore there is need for seasonal difference.

4.4 Parameter Estimation

As identifying a tentative model is completed. The next step is to estimatethe parameters in the model using maximum likelihood estimation. Findingthe parameters that maximize the probability of observations is main goalof maximum likelihood. This was carried out using the R software, the AICvalues were generated and the order with minimum value was chosen.

Oyo(South West) Rainfall model SARIMA(1,0,3)x(1,0,1)$_{12}$

coeff.	ϕ_1	θ_1	θ_2	θ_3	Φ_1	Θ_1	Intercept
Estimate	0.9522	-0.9986	0.0397	-0.0411	0.9993	-0.9835	88.2874
Standard error	0.0246	0.0675	0.0876	0.0601	0.0025	0.0293	0.9240
σ^2=3693	$loglik = -1400.65$						

Enugu(South East) Rainfall model SARIMA(3,0,3)x(1,0,1)$_{12}$

coeff.	ϕ_1	ϕ_2	ϕ_3	θ_1	θ_2	θ_3	Φ_1	Θ_1	Int.
Estimate	-0.6056	-0.1359	0.6592	0.6323	0.2390	-0.6117	0.9963	-0.9675	82.969
Standard error	0.3013	0.3825	0.2964	0.3184	0.3964	0.3182	0.0196	0.0890	8.765
σ^2=3693	$loglik = -1399.23$								

Kwara(North Central) Rainfall model SARIMA(1,0,1)x(1,1,1)$_{12}$

coeff.	ϕ_1	θ_1	Φ_1	Θ_1	Intercept
Estimate	-0.5755	0.6941	0.0820	-1.0000	
Standard error	0.2458	0.2150	0.0696	0.0832	
σ^2=3222	$loglik = -1327.24$				

Kaduna(North West) Rainfall model SARIMA(1,0,1)x(2,0,1)$_{12}$

coeff.	ϕ_1	θ_1	Φ_1	Φ_2	Θ_1	Intercept
Estimate	0.9332	-1.0000	1.0937	-0.0941	-0.9675	46.9940
Standard error	0.0251	0.0028	0.0680	0.0679	0.0365	0.4053
σ^2=1098	$loglik = -1254.84$					

Rivers(South South) Rainfall model SARIMA(1,0,2)x(1,0,1)$_{12}$

coeff.	ϕ_1	θ_1	θ_2	Φ_1	Θ_1	Intercept
Estimate	0.9332	-1.0000	1.0937	-0.0941	-0.9675	46.9940
Standard error	0.0251	0.0028	0.0680	0.0679	0.0365	0.4053
σ^2=1098	$loglik = -1254.84$					

Taraba(North East) Rainfall model SARIMA(3,0,3)x(1,0,1)$_{12}$

coeff.	ϕ_1	ϕ_2	ϕ_3	θ_1	θ_2	θ_3	Φ_1	Θ_1	Int.
Estimate	0.8100	0.5934	-0.9058	-0.7538	-0.7508	0.9989	0.9990	-0.9761	54.6532
Standard error	0.0316	0.0514	0.0311	0.0312	0.0303	0.0329	0.0046	0.0559	6.6168
σ^2=1616	$loglik = -1300.38$								

4.5 Model Diagnostics

4.5.1 Box- Jenkin test

Hypothesis:

H_o:The residuals are not independently distributed.

H_1:The residuals are independently distributed.

Table 5: **Box-Ljung test and p-value**

State	X-squared	p-value
Oyo	24.793	0.2095
Enugu	15.767	0.731
Kwara	16.5761	0.6803
Kaduna	18.1107	0.5801
Rivers	13.191	0.869
Taraba	17.711	0.6064

Decision rule: Reject Ho if p-value<0.05, otherwise we fail to reject

Conclusion: Since p-value <0.05 we reject Ho and conclude that the residualsfor each series are independently distributed.

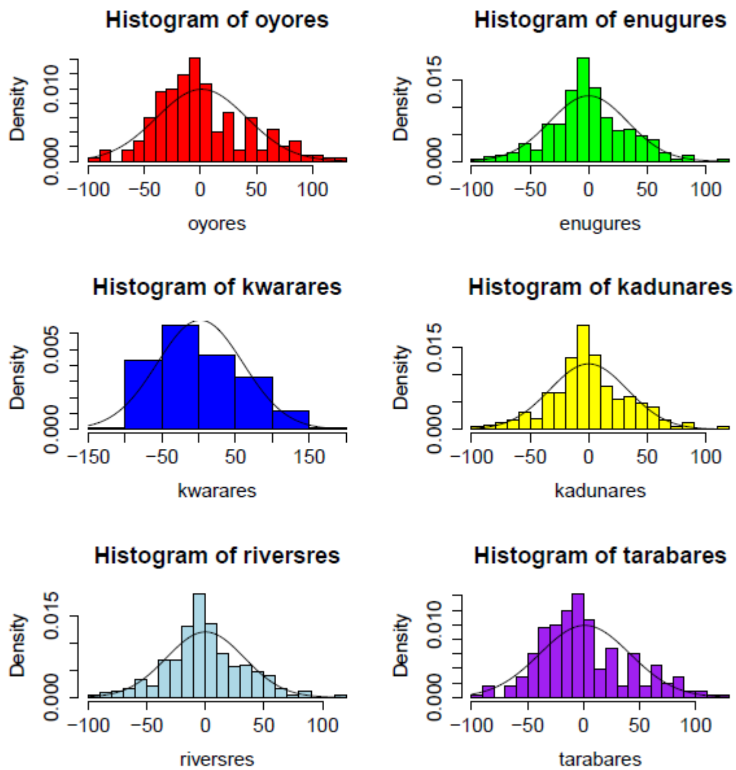

Figure 2: Histogram for the Rainfall model Residuals

4.6 Forecasting

One of the most important objectives of time series analysis is to predictits future values. It is all about making prediction into the future from itspast values on the basis of the model that effectively describe the evolutionof a series. Four years prediction for each stations in the geo-political zone was carried out and the values in millimeters are presented in the time plotabove.

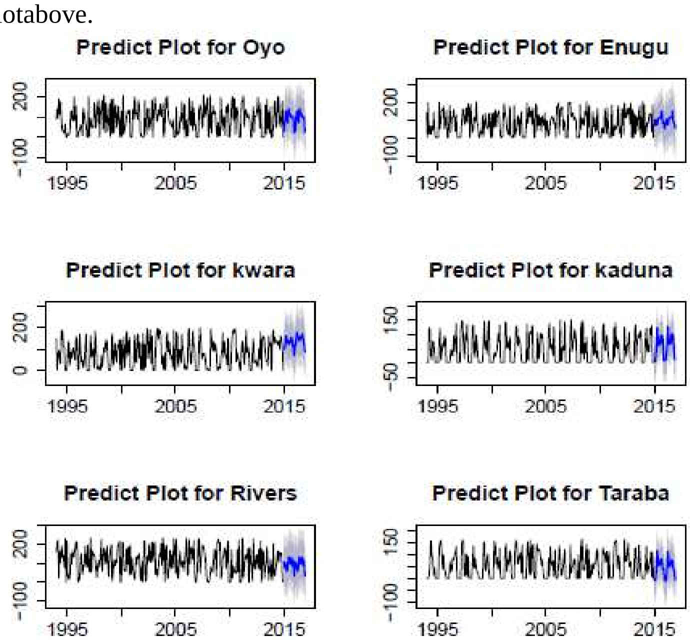

Figure 3: Predict Plots for each States

5.0 Summary and Conclusion

5.1 Summary

This research work aims to model the rainfall series in Nigeria (a case studyof six stations in each geo-political zones). Basic exploratory data analysiswere carried out and the normality of the series was also tested, the timeplot of the monthly rainfall in each stations was checked for trend and seasonal pattern in the series.

After the visual inspection of the time plot, it is important to check bytesting the stationarity and seasonal unit in the series. Therefore, theAugmented Dickey-Fuller test, Kwaiatkowski Phillips-Schmidt-Shin test,Phillips-Perron Unit Root Test were tested to establish the stationarityinthe series and Hylleberg, Engle, Granger and Yoo (HEGY's Test) was usedto test the seasonal unit root in the series.

Furthermore, the ACF and PACF plot were used to identify the possibleorder of the models, while the models with the lowest AkaikeInformationCriterion (AIC) were selected. Also several diagnostic statistics and plots ofthe residuals was used to examine the goodness of fit of the selected modelto the data.

5.2 Conclusion

The monthly rainfall series of the six states in each geo-political zones hasbeen studied using the SARIMA model methodology. A monthly rainfallrecord spanning the period of 1994-2014 for six states (Oyo, Enugu, Kwara,Kaduna, Rivers, Taraba) in Nigeria has been used to develop and test themodels. Basic descriptive was carried out and the time plots were examinedfor seasonal and non-seasonal trend, the adf-test, kpss-test and pp-test wereused to test for non-seasonal stationarity of the series, the test shows thatthe series are non-seasonal stationary. Also the HEGY-test was used to testthe seasonal stationarity of the series and it was established that all but series (Kwara) was seasonal stationary, therefore kwara-series was seasonally differenced.

The main objective of this study is to forecast rainfall series using SARIMA model, to avoid fitting over parametized model, AIC was employed in selecting the best model. The model with a minimum value of the informationcriterions is considered as the best [11, 12]. Theresidual analysis was also carried out to test the accuracy of the models selected, the models were found to adequate. Furthermore the forecast shows adecline in pattern of rainfall series in the various stations, delayed rainfall insome part and early cessation of rainfall in some other parts which corroborate the seasonal rainfall prediction (SRF) made by Nigerian MeteorologicalAgency (NIMET) for 20

6.0 References

[1] Confalonieri,U.[M.L.Parry et al.] (2007)."Human health in: Climate Change2007:Impacts,Adaptation and Vulnerability.Contribution of Working GroupII to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*".Cambridge University Press,Cambridge,U.K., and New York,N.Y.,U.S.A.*pp.391-431

[2] Harvey,R.,Andrew C. and Souza R.C,(1987)."Assessing and Modelling theCyclical Behavior of Rainfall in North-East Brazil",*Journal of Climate andApplied Meteorology,*Vol.26,1339-1344pp

[3] Chiew, F.H.S., Stewardson, M.J., and McMahon, T.A., "Comparison ofsix rainfall-runo_ modeling approaches". *J. Hydrol.*, 147, (1993), 1-36.

[4] Al-Ansari, N.A., Ezz-Aldeen, M., Knutsson, S. and Zakaria, S., 'Water harvesting and reservoir optimization in selected areas of south SinjarMountain,Iraq', *Accepted for publication in ASCE J. Hydrologic Engineering, (2012).*

[5] Al-Ansari, N.A. Shamali, B. and Shatnawi, A., 'Statistical analysis at threemajor meteorological stations in Jordan, Al Manara' *Journal for scientificstudies and research 12, (2006),* 93-120.

[6] Amha, G.,(2010). 'Modelling and Forcasting monthly rainfall in Tigrayregion.Case of Mekelle station. (Unpublished M.Sc. Thesis),*Department of Statistics,Addis Abba University,Ethiopia.*

[7] Saleh Zakaria, Nadhir Al-Ansari, Sven Knutsson, Thafer Al-Badrany,'ARIMAModels for weekly rainfall in the semi-arid Sinjar District at Iraq,'*Journalof Earth Sciences and Geotechnical Engineering*, vol.2, no. 3, 2012, 25-55.

[8] A.R. Abdul-Aziz, M. Anokye , A. Kwame, L. Munyakazi and N.N.N. Nsowah-Nuamah (2013) Modeling and Forecasting Rainfall Pattern in Ghana as aSeasonalArima Process: The Case of Ashanti Region. *International Journal of Humanities and Social Science* Vol. 3 No. 3

[9] Ette Harrison Etuk, ImohUdoMo_at, Benjamin EleChims (2013). Modelling Monthly Rainfall Data of Port Harcourt, Nigeria by Seasonal Box-Jenkins Methods" *International Journal of Sciences Volume 2, Issue July,* 2013.

[10] Obisesan, K.O**.** and Yussuf O.A. (2015): On the use of probability with Generalized Linear Models for Evaluating Extreme Rainfall in Nigeria. International Journal of Applied Science and Mathematics, Vol2(.4). pp117-122

[11] Akaike, H. (1973). 'Maximum likelihood identification of Gaussian autoregressive moving-average models."*Biometrika*, 60, 255-266.

[12] Akaike, H. (1974). "A new look at the statistical model identification."*IEEE Transactions on Automatic Control*, 19, 716-723.