International Journal of Development Mathematics Vol 3 Issue 2 Page No. 409 - 427

INTERNATIONAL JOURNAL OF DEVELOPMENT MATHEMATICS

ISSN: 3026-8656 (Print) | 3026-8699 (Online)

journal homepage: https://ijdm.org.ng/index.php/Journals

Modeling Rainfall in Three Northern States with SARIMA (1981 – 2020)

Dahida Terkuma[a,*], Haruna U. Yahaya[b], Thaddeus Pever[b], Obele T. Elorhor[c] and Sameerah Haruna[c]

[a]Department of Mathematics, Nigerian Army University Biu, Born State, Nigeria
[b]Department of Statistics, Faculty of Science, University of Abuja, Abuja, Nigeria
[c]Department of Economics, Nigerian Army University Biu, Borno State, Nigeria

ARTICLE INFO

Article history:
Received 10 April 2026
Received in revised form 27 May 2026
Accepted 05 June 2026

Keywords:
SARIMA; Rainfall Modeling; Time Series; Drought; Aridity

MSC 2020 Subject classification:
62M10, 62P12

ABSTRACT

This study compares the forecasting models of monthly rainfall data of three northern Nigerian states which are Katsina, Borno and Sokoto. The analysis is based on using monthly amount of rainfall data from January 1981 to December 2020.The data were stationary at first difference with p-values less than 0.05, Seasonal autoregressive Integrated moving Average (SARIMA) models were computed and diagnosed. Diagnostic test was run on the models where the results showed that the models were adequate and serially uncorrelated using Box-Ljung test plot respectively. Furthermore, SARIMA(10, 1, 10) x(1, 2, 1)$_{12}$, SARIMA(10, 1, 10) x(1, 2, 1)$_{12}$ and SARIMA(1, 1, 1) x(1, 2, 1)$_{12}$ were found to be adequate from the result of the Box-Ljung test of residuals respectively. The results showed that, the three selected models were best for each variable on monthly rainfall in the three northern Nigeria states.

1. Introduction

Rainfall is widely regarded as the most consequential element of the climate system in tropical regions, where prevailing economic activities are predominantly agrarian (Oladipo, 1987; Hyuwa, 2005). Its total amount, intensity, duration, variability, and spatial distribution collectively determine agricultural productivity, water availability, and ecosystem function. In the semi-arid northern states of Nigeria, which lie within the Sudano-Sahelian belt, the dependence of both smallholder and large-scale agriculture on rainfall is near absolute, making accurate characterisation and forecasting of rainfall dynamics a scientific and policy priority.

Nigeria's rainfall regime is strongly seasonal, governed by the oscillation of the Inter-Tropical Discontinuity (ITD) between the moist southwesterly monsoon and the dry northeasterly harmattan (Ayoade, 1973; Ati, 1996). Year-to-year variability in this oscillation translates directly into interannual fluctuations in crop yields, groundwater recharge, and the frequency of drought events. It is well established that global warming is amplifying both the variability and regional asymmetry of rainfall patterns (Hulme *et al.*, 1998; Kayano & Sansígolo, 2008), with sub-Saharan Africa and the Sahel among the most vulnerable zones. The knowledge of climate variability across different temporal and spatial scales is therefore indispensable for understanding the nature of climate systems and their societal impacts (Oguntunde *et al.*, 2012).

Droughts are classified as meteorological, agricultural, hydrological, or social (Rashid *et al.*, 2015). Meteorological drought, a prolonged precipitation deficit relative to the long-term mean, is the primary driver of all other drought

* Corresponding author. Tel.: +2348066541055, +2347085779910
E-mail address: dahidaterkuma1@gmail.com (Terkuma, D.)
https://doi.org/10.62054/ijdm/0302.26

types. The African continent has experienced numerous severe drought events leading to widespread famine (Masih & Maskey, 2014). Studies in Nigeria have demonstrated an increasing trend in climate variability and drought frequency (Abiodun *et al.,* 2013; Shiru *et al.,* 2019). Studies show the standardised precipitation evapotranspiration index (SPEI), proposed by Vicente-Serrano *et al.* (2010), is particularly effective for semi-arid and arid environments.

General Circulation Models (GCMs) from CMIP5 (Taylor *et al.,* 2012) are widely used for drought projections, though they require statistical downscaling before application at local scales (Fowler *et al.,* 2007). The selection of suitable GCMs and multi-model ensemble generation (Salman *et al.,* 2018) is recommended to minimise projection uncertainties. Statistical time series models such as SARIMA remain a practical, validated, and computationally accessible alternative for local-scale rainfall forecasting, as demonstrated by Martinez *et al.* (2011), Mceihattan (1975), and Inderjeet & Sabita (2008), among others.

Earlier investigations documented declining annual rainfall totals in northern Nigeria from records terminating in the late 1980s and early 1990s (Oladipo, 1993; Ati *et al.,* 2002). Adefolalu (1986) examined rainfall trends over 28 meteorological stations for 1911–1980 and found evidence of declining rainfall using 40-year moving averages. Eludoyin *et al.* (2009) observed notable fluctuations in monthly rainfall distribution between the decades 1985–1994 and 1995–2004. However, a systematic statistical modelling exercise incorporating the most recent two decades of data, and using SARIMA models, for the key northern states of Katsina, Borno, and Sokoto has remained absent from the literature.

Several researchers have successfully applied the SARIMA framework to rainfall forecasting in sub-Saharan Africa and beyond. Abdul-Aziz *et al.* (2013) proposed SARIMA$(0,0,0)\times(2,1,0)_{12}$ for Ghana's Ashanti Region (BIC = 8.592). Ampaw *et al.* (2013) identified SARIMA$(0,0,1)\times(1,1,1)_{12}$ for New Juaben Municipality, Ghana (BIC = 8.077). Wiredu et al. (2013) reported SARIMA$(0,0,1)\times(0,1,1)_{12}$ for the Navrongo Municipality (AIC = 591.03). Chonge *et al.* (2015) fitted SARIMA$(0,0,0)\times(0,1,2)_{12}$ to Kenyan rainfall data. Within Nigeria, Jabrin *et al.* (2014) fitted SARIMA$(0,0,0)\times(1,1,1)_{12}$ for Kano State, while Eni and Adeyeye (2015) identified SARIMA$(1,1,1)\times(0,1,1)_{12}$ as adequate for Warri. Ogunrinde (2012) applied Box–Jenkins methodology to Lagos and Maiduguri. Bari *et al.* (2015) demonstrated the effectiveness of SARIMA$(0,0,1)\times(1,1,1)_{12}$ for Bangladesh with 95% confidence forecasting intervals.

This paper addresses the identified gap with the following specific objectives: (i) to model monthly rainfall for Katsina, Borno, and Sokoto using the SARIMA framework fitted to 1981–2020 data; (ii) to identify statistically optimal model specifications for each state using information criteria and diagnostic tests; and (iii) to generate short-term rainfall forecasts that can inform drought mitigation and agricultural planning.

2. Literature Review and Theoretical Framework

2.1 Introduction

This section is divided into three sections where section 2.1 introduce the chapter, 2.2 explains the literature review of the research work, section 2.3 shows some theoretical frameworks and section 2.4 identifies review of methodological and empirical issue.

2.2 Literature review

The drought prone areas of northern Nigeria lie within the Sudano-Sahelian region roughly north of Latitude 100N, delimited by latitude 140N; and between longitudes 20 44' E and 140 42'E. The area stretches from Kebbi through Niger, to Bauchi, Kaduna, Gombe and Borno up north, covering states like Yobe, Jigawa, Kano, Katsina, Zamfara

and Sokoto states. Only northern part of Adamawa state experiences drought so the study area covers a total area of about 240,000k2.
The climate of northern Nigeria north of latitude 100N is of tropical continental (AW) type characterized by distinct wet (April to October) and dry from November to March seasons respectively as dictated by the oscillation of the Inter-tropical Discontinuity (ITD).

This study examined trend in aridity with a view of identifying adaptive strategies for sustainable agriculture. Existing literature indicates increasing trend toward aridity in recent times (Dregne & Chou, 1992; Nicholson, 2003; Hanafi & Jauffret, 2008; López *et al.*, 2008; Gaughan & Waylen, 2012). Lázaro *et al.* (2001) mentioned that in order to understand the behaviour of ecosystems in semi-arid areas, rainfall must be analyzed over time. Intra- and inter-seasonal rainfall variability are imperative in studying moisture efficiency or moisture quality in the semi-arid areas of the Sudano-Sahelian belt of Nigeria (Usman, 2000).

2.3 Theoretical Framework

Droughts are classified as meteorological, agricultural, hydrological and social. But the meteorological droughts aggravate all other kinds, and therefore, understanding it is crucial to develop mitigation measures.it occurs due to a reduction of precipitation or atmospheric water balance from long term mean. As a result, it can occur in many climate regime due to natural variability of climate (Rashid *et al.* 2015).the increased variability in climate due to global warming has increased the frequency and severity of drought in many parts of the globe in recent decades (Sung & Chung, 2014), which are expected to swell with climate warming.
The African continent is considered as very prone to droughts due to high variability of rainfall. A large number of extreme droughts have been observed in the recent past which caused famines and loss of millions of lives (Masih & Maskey, 2014) in Africa. Studies reported more erratic behavior of climate due to climate change, and thus, potentially increase its severity indifferent countries including Nigeria (Boudoin & Naik, 2017). Numerous studies on the variability of climate and other changing nature of the droughts using various indices have been conducted in Nigeria (Shiru *et al.* 2019). Though major drought events have not occurred in Nigeria for the past decades, studies showed an increasing tendency in climate variability and thus, potentially increase in drought frequency and affected areas (Abiodum & Lawal, 2013).

2.4 Review of Methodology and Empirical Issues

Monthly rainfall records for four decades (1981 – 2020) for three arid states in northern Nigeria were sourced from the Nigerian Meteorological Agency Abuja. Various researchers used SARIMA model in different research work. Such include; An Empirical Study of the Usefulness of SARIMA models in Energy Science by Leila & Masoud, (2012), A SARIMA forecasting model to predict the number of cases of dengue in Campinas, State of Sao Paulo, Brazil, Martinez *et.al.* (2011). Forecasting Precipitation Using SARIMA model, Mceihattan, (1975).
This research work focuses on statistical analysis on rainfall aridity in three selected northern Nigerian states, using monthly rainfall data to forecast subsequent aridity in the future with the use of SARIMA model.

3. Data and Methodology

3.1 Data

Monthly rainfall records spanning four decades (January 1981 to December 2020) for Katsina, Borno, and Sokoto were obtained from the Nigerian Meteorological Agency (NiMet), Abuja. Each series comprised 480 monthly observations. Prior to analysis, logarithmic transformation (natural log) was applied to all three series to stabilise variance and reduce the positive skewness inherent in raw rainfall distributions. The log-transformed series are denoted LogKatsina, LogBorno, and LogSokoto throughout this paper.

All analyses were conducted in R (version 4.1) using the forecast package.

3.2 Study Area

The three states examined Katsina, Borno, and Sokoto, lie within the Sudano-Sahelian belt of northern Nigeria, broadly north of latitude 10°N, bounded by latitude 14°N and longitudes 2°44'E to 14°42'E (SAWA *et al.*, 2020). The combined study area spans approximately 240,000 km². The climate is of the tropical continental (AW) type, characterized by a distinct wet season (April to October) and dry season (November to March). The terrain consists of high plains at 450–750 m above sea level, occasionally interrupted by granitic inselbergs and volcanic plateaux such as the Kufena Hills of Zaria and the Jos Plateau.

Nicholson *et al*. (2000) documented that long-term rainfall in the semiarid and sub-humid zones of West Africa during 1968–1997 was on average 15–40% lower than during 1931–1960. This secular decline, attributed partly to natural variability and partly to land use change and global warming, underscores the urgency of reliable rainfall forecasting in the region. Gaughan & Waylen (2012) further noted that understanding intra- and inter-seasonal rainfall variability is imperative for studying moisture efficiency and agricultural productivity in semi-arid belts.

3.3 SARIMA Modelling Framework

The Box–Jenkins methodology, Ljung, G. M., & Box, G. E. P. (1978). Biometrika, 65(2), 297–303. was adopted, proceeding through three stages: model identification, parameter estimation, and diagnostic checking. For a seasonal time series of period s = 12, the multiplicative SARIMA (p, d, q)×(P, D, Q)s model is expressed as:

$$\Phi\left(L^s\right)\varphi\left(L\right)\nabla^{D_S}\nabla^d X_t = \Theta\left(L^s\right)\theta\left(L\right)\varepsilon_t$$

where $\varphi\left(L\right)$ and $\Phi\left(L^s\right)$ are the non-seasonal and seasonal autoregressive polynomials of orders p and P respectively; $\theta\left(L\right)$ and $\Theta\left(L^s\right)$ are the non-seasonal and seasonal moving average polynomials of orders q and Q; $\nabla^d = \left(1-L\right)^d$ is the non-seasonal difference operator of order d; $\nabla^{D_S} = \left(1-L^s\right)^D$ is the seasonal difference operator of order D; and ε_t is a white noise process. The orders $\left(p,d,q\right)$ govern the non-seasonal component while $\left(P,D,Q\right)$ govern the seasonal component.

3.4 Model Identification and Selection Criteria

Stationarity was assessed visually via time plots and formally confirmed using the Augmented Dickey–Fuller (ADF) test. Under the ADF framework, rejection of the unit root null hypothesis at the 5% significance level confirms stationarity. Seasonality was identified through ACF/PACF plots and seasonal decomposition. A spike with sharp

cutoff in the PACF at lag k indicates an AR(k) process; gradual decay in the ACF indicates an MA process. Similarly, spikes at seasonal lags in the PACF and ACF indicate seasonal AR and MA orders respectively.

Model selection among candidate specifications employed four criteria applied simultaneously: (i) the Akaike Information Criterion (AIC; lower is preferred); (ii) residual variance σ^2 (lower is preferred); (iii) log-likelihood (higher is preferred); and (iv) number of statistically significant coefficients (higher is preferred). Diagnostic adequacy of the selected model was assessed using the Ljung–Box portmanteau test (Schwarz, 1978; Suhartono & Lee, 2011), where a p-value > 0.05 indicates that residuals are consistent with a white noise process.

4. Results

4.1 Descriptive Statistics

Table 1 presents the descriptive statistics for the raw monthly rainfall observations across the three states.

Table 1: Descriptive Statistics of Monthly Rainfall Data, Katsina, Borno, and Sokoto (1981–2020)

Statistic	Katsina	Borno	Sokoto
Mean (mm)	45.39	48.39	55.54
Median (mm)	1.50	2.35	3.45
Maximum (mm)	413.40	342.70	374.90
Minimum (mm)	0.00	0.00	0.00
Std. Deviation	71.39	76.07	82.58
Skewness	1.903	1.743	1.698
Kurtosis	6.590	5.294	5.501
Jarque-Bera	485.95	309.16	315.78
JB Prob.	0.0000	0.0000	0.0000
Sum	19,336.50	20,613.20	23,659.60
Observations	425	425	425

The descriptive results reveal several important features of the rainfall series. All three states share a pronounced right-skewness values of 1.903, 1.743, and 1.698 for Katsina, Borno, and Sokoto respectively, reflecting the fact that the vast majority of monthly observations record little or no rainfall (dry season months), while a small number of wet-season months record extremely large amounts. This characteristic is archetypal of monsoonal rainfall regimes in the Sahel. Kurtosis values of 6.590, 5.294, and 5.501, all substantially exceeding three, indicate leptokurtic distributions with heavier tails than the Gaussian, driven by the extreme wet-season peaks. The Jarque–Bera test formally rejects normality for all three series at the 1% level ($p = 0.000$ in all cases, which is less than 0.01, i.e. $p < 0.01$). These properties collectively justify the logarithmic transformation applied prior to SARIMA modelling, as they violate the normality assumption and introduce heteroscedasticity that would otherwise destabilize model

estimates. Sokoto records the highest mean monthly rainfall (55.54 mm) and greatest dispersion (SD = 82.58 mm), while Katsina records the lowest mean (45.39 mm).

4.2 Time Series Analysis - Katsina

4.2.1 Time Plot and Logarithmic Transformation

Figure 1 displays the time plot of raw monthly rainfall for Katsina over the 1981–2020 period. The plot reveals the characteristic seasonal pattern of the Sahel: near-zero rainfall from November through March, a rapid onset of precipitation in April to May, peak rainfall in July –August, and a sharp recession through September to October. No systematic upward or downward long-run trend is discernible, suggesting that the series fluctuates around a stable long-run mean. However, the strong within-year amplitude variation and occasional extreme values confirm the presence of heteroscedasticity that necessitates transformation.

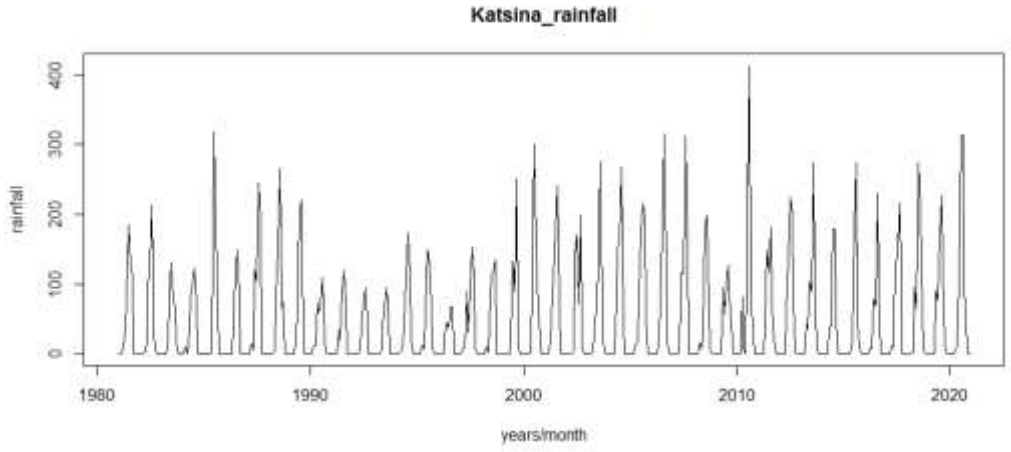

Figure 1: Time Plot of Raw Monthly Rainfall, Katsina (January 1981 – December 2020)

Figure 2 presents the time plot following natural logarithmic transformation (LogKatsina). The transformation markedly compresses the high-amplitude wet-season spikes and expands the near-zero dry-season values, resulting in a more homoscedastic series. The seasonal periodicity remains clearly visible, but the variance is substantially stabilized, which is a prerequisite for reliable ARIMA-class modelling.

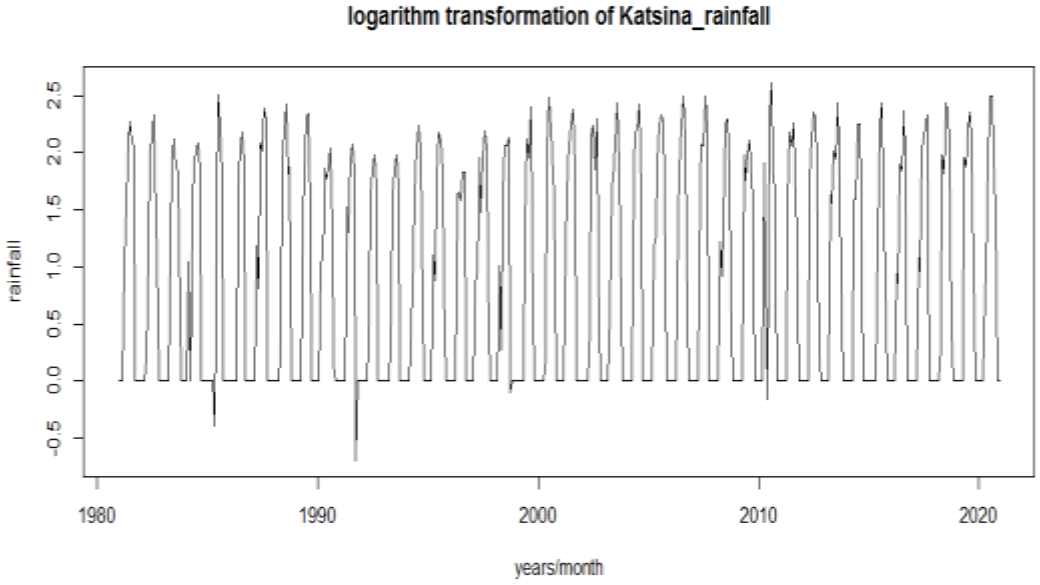

Figure 2: Time Plot of Log-Transformed Monthly Rainfall, Katsina (LogKatsina)

4.2.2 Stationarity Test — Katsina

Table 2 presents the ADF unit root test results for LogKatsina after first differencing.

Table 2: Augmented Dickey–Fuller Unit Root Test - LogKatsina (First Difference)

Test Component	t-Statistic	p-value
ADF Test Statistic	−4.168	0.0008
Critical Value: 1% level	−3.444	—
Critical Value: 5% level	−2.867	—
Critical Value: 10% level	−2.570	—

Note: Lag length = 11, selected automatically based on AIC (maxlag = 17). Null hypothesis: series has a unit root.

The ADF test statistic of −4.168 significantly exceeds the 5% critical value of −2.867 in absolute terms (p = 0.0008), leading to rejection of the null hypothesis of a unit root. LogKatsina is therefore confirmed stationary at first difference, i.e., integrated of order one, I(1). Figure 3 displays the differenced series, which oscillates stochastically around zero without any discernible trend or explosive behavior, providing visual corroboration of the ADF result.

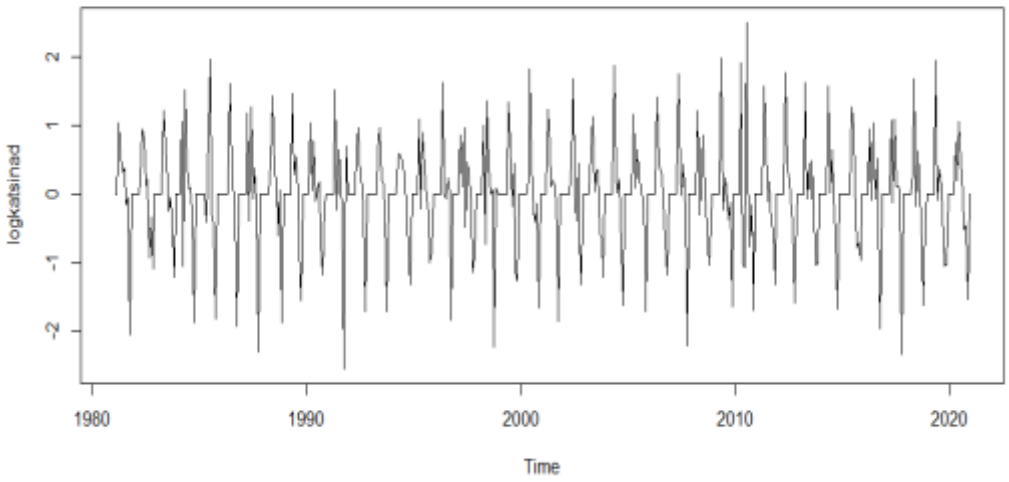

Figure 3: First-Differenced LogKatsina Series (∇LogKatsina)

4.2.3 Correlogram and Seasonal Decomposition — Katsina

Figure 4 presents the seasonal decomposition of LogKatsina, separating the series into trend, seasonal, and irregular components.

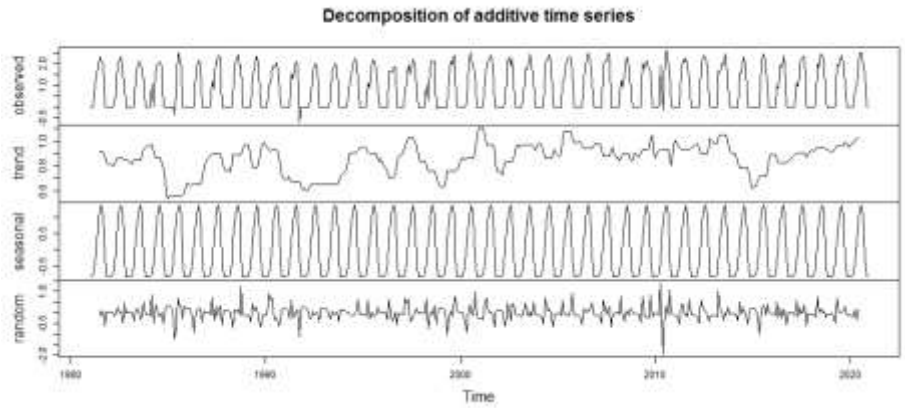

Figure 4: Seasonal Decomposition of LogKatsina

The decomposition clearly isolates the dominant seasonal cycle, which repeats consistently across all 40 years of the sample. The trend component shows mild fluctuation without a systematic directional movement, consistent with the horizontal movement observed in the raw time plot. The irregular component captures residual variation not explained by trend or seasonality and appears approximately stationary in variance, confirming that the log transformation has achieved adequate variance stabilization.

4.2.4 Model Selection and Parameter Estimation - Katsina

Table 3 presents the selection criteria for the four candidate SARIMA models estimated for Katsina.

Table 3: SARIMA Model Selection Criteria - Katsina

Model	Sig. Coef.	σ^2	AIC	Log-Likelihood
SARIMA(1,1,1)×(1,2,1)$_{12}$	4	0.2123	670.42	−330.21
SARIMA(1,1,10)×(1,2,1)$_{12}$	3	0.2109	685.53	−328.76
SARIMA(10,1,1)×(1,2,1)$_{12}$	3	0.2087	680.49	−326.25
SARIMA(10,1,10)×(1,2,1)$_{12}$ ✓	6	0.1610	632.61	−293.31

✓ *Selected model. Selection criteria: highest significant coefficients, lowest AIC and σ², highest log-likelihood.*

Inspection of Table 3 reveals that SARIMA(10,1,10)×(1,2,1)$_{12}$ dominates all competing specifications across every selection criterion simultaneously. It records the highest number of statistically significant coefficients (6 of the estimated parameters significant at 5%), the lowest AIC (632.61 vs. 670.42–685.53 for alternatives), the lowest σ^2 (0.161), and the highest log-likelihood (−293.31). The improvement in AIC over the next-best model is approximately 38 units, which is a substantial gain. This model was therefore selected for Katsina.

Figure 5 presents the diagnostic plots of the fitted SARIMA(10,1,10)×(1,2,1)$_{12}$ model for Katsina.

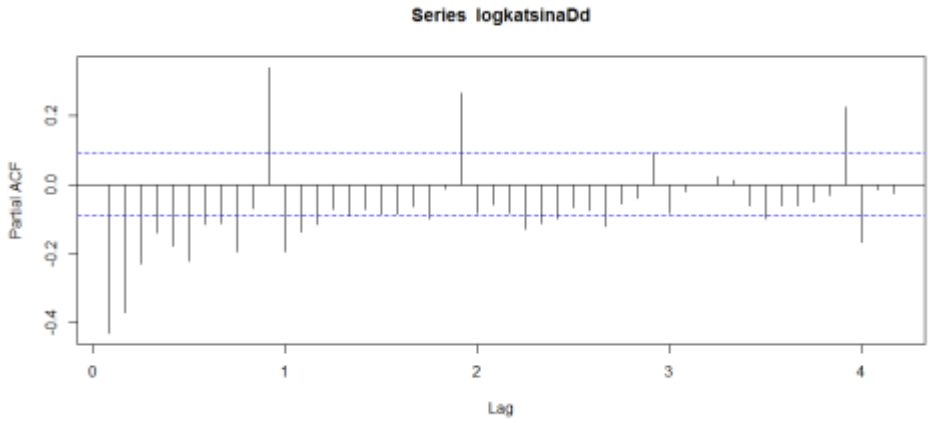

Figure 5: Residual Diagnostics for SARIMA(10,1,10)×(1,2,1)$_{12}$ - Katsina

The Ljung–Box portmanteau test yields Q* = 200.92 (df = 178, p = 0.1148). Since the p-value (0.1148) exceeds the 5% significance threshold, the null hypothesis that residuals constitute a white noise process cannot be rejected. The residual ACF/PACF in the diagnostic plot shows no significant autocorrelation at any lag, and the histogram of residuals approximates the normal distribution. These results collectively confirm that SARIMA(10,1,10)×(1,2,1)$_{12}$ adequately fits the Katsina rainfall data and that no systematic structure remains unexplained.

Figure 6 displays the out-of-sample rainfall forecasts generated by the selected model for Katsina.

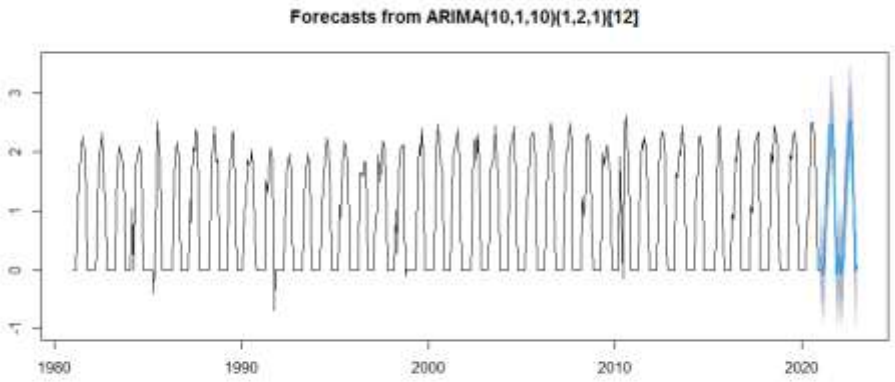

Figure 6: Out-of-Sample Rainfall Forecast with 95% Confidence Bands - SARIMA$(10,1,10)\times(1,2,1)_{12}$, Katsina

The forecast in Figure 6 preserves the characteristic seasonal cycle of the Katsina rainfall regime, with projected peak values in July to August and near-zero projections during the dry months. The confidence bands widen as the forecast horizon extends, reflecting the natural accumulation of forecast uncertainty. Importantly, the central forecast values suggest a modest upward shift in peak rainfall relative to the most recent years of the sample, consistent with the slightly increasing trend in the irregular component identified during seasonal decomposition. This finding implies a mild amelioration of aridity conditions in Katsina over the short-term forecast horizon.

4.3 Time Series Analysis — Borno

Figure 7 presents the raw time plot the log-transformed series for Borno (LogBorno). As with Katsina, the raw series exhibits strong annual seasonality, no discernible long-run trend, and marked heteroscedasticity, motivating logarithmic transformation. After transformation, variance is substantially stabilized, making the series amenable to SARIMA modelling.

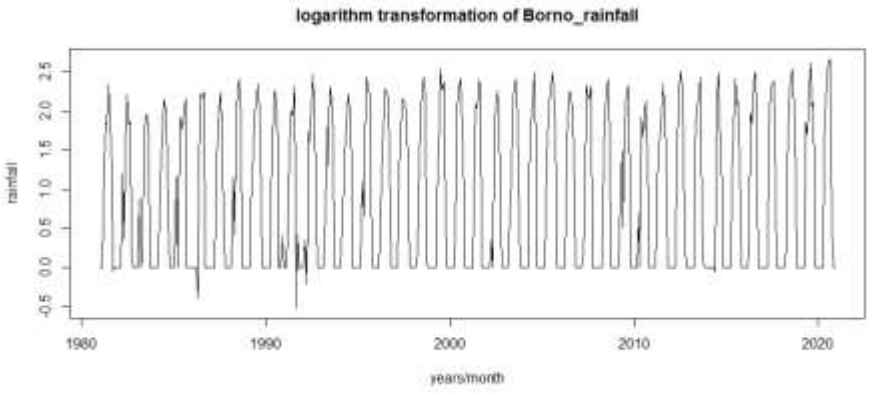

Figure 7: Time Plot of Raw Monthly Rainfall, Borno (January 1981 - December 2020)

4.3.1 Stationarity Test - Borno

Table 4 presents the ADF stationarity test for LogBorno.

Table 4: Augmented Dickey–Fuller Unit Root Test - LogBorno (First Difference)

Test Component	t-Statistic	p-value
ADF Test Statistic	−18.917	0.0000
Critical Value: 1% level	−3.444	—
Critical Value: 5% level	−2.868	—
Critical Value: 10% level	−2.570	—

Note: Lag length = 11, selected automatically based on AIC (maxlag = 17). Null hypothesis: series has a unit root.

The ADF statistic of −18.917 is overwhelmingly significant (p = 0.000), far exceeding the 1% critical value of −3.444 in absolute terms. The null hypothesis is decisively rejected, confirming that LogBorno is stationary at first difference, I (1). Figure 8 shows the first-differenced series oscillating stably around zero, with no remaining trend. The correlogram after first differencing displays the same diagnostic signatures observed for Katsina: a spike at lag 1 of the PACF (AR(1)) and gradual decay in the ACF (MA(1)), along with evidence of annual seasonality at multiples of lag 12.

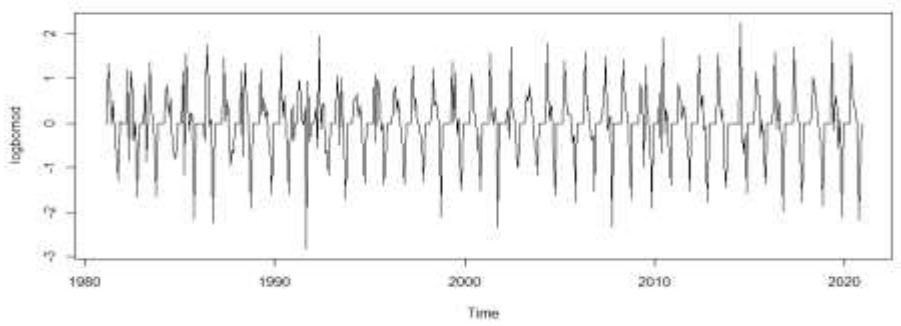

Figure 8: First-Differenced LogBorno Series (∇LogBorno)

4.3.2 Model Selection, Diagnostics, and Forecast - Borno

Table 5 presents the model selection results for Borno.

Table 5: SARIMA Model Selection Criteria - Borno

Model	Sig. Coef.	σ^2	AIC	Log-Likelihood
SARIMA(1,1,1)×(1,2,1)$_{12}$	4	0.2247	695.74	−342.87
SARIMA(1,1,10)×(1,2,1)$_{12}$	2	0.2111	692.19	−332.09
SARIMA(10,1,1)×(1,2,1)$_{12}$	3	0.2194	703.54	−337.77
SARIMA(10,1,10)×(1,2,1)$_{12}$ ✓	4	0.1610	632.61	−293.31

✓ *Selected model.*

SARIMA(10,1,10)×(1,2,1)$_{12}$ is again the dominant specification for Borno, recording the lowest AIC (632.61), the lowest σ^2 (0.161), and the highest log-likelihood (−293.31). The margin of improvement over competing models is consistent with the Katsina result, reinforcing the appropriateness of this higher-order specification for capturing the complex autocorrelation structure in the Borno rainfall series. The Ljung–Box diagnostic test yielded Q* = 200.92 (df = 178, p = 0.1148), confirming white noise residuals. Figure 9 shows the residual diagnostic plots, and Figure 9 presents the rainfall forecast.

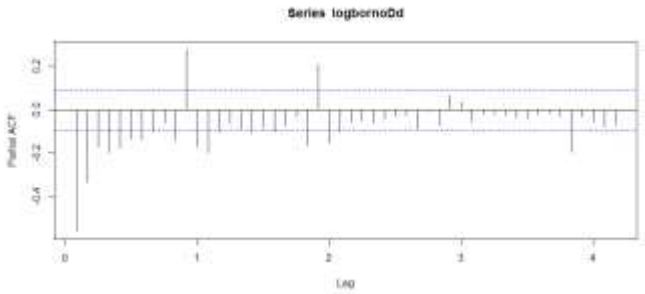

Figure 9: Residual Diagnostics for SARIMA(10,1,10)×(1,2,1)$_{12}$ - Borno

4.4 Time Series Analysis - Sokoto

Figures 10 and 11 present the raw and log-transformed time plots for Sokoto. Among the three states, Sokoto records the highest mean monthly rainfall (55.54 mm) and the greatest standard deviation (82.58 mm), indicating both higher average rainfall and greater volatility. As with the other two states, no systematic long-run trend is discernible, and the seasonal cycle is the dominant structural feature.

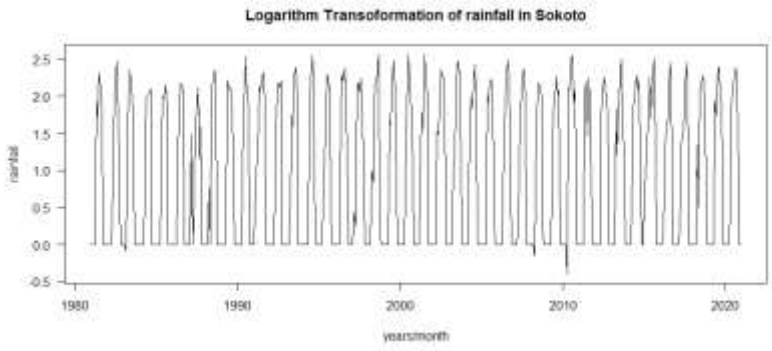

Figure 10: Time Plot of Raw Monthly Rainfall, Sokoto (January 1981 - December 2020)

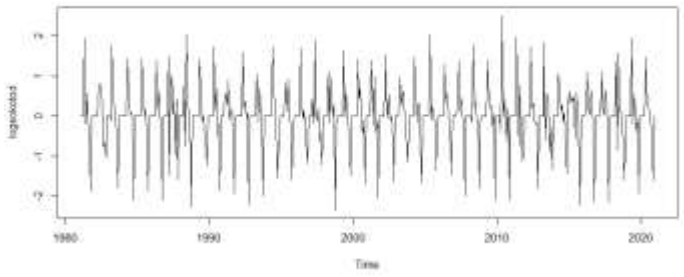

Figure 11: Time Plot of Log-Transformed Monthly Rainfall, Sokoto (LogSokoto)

4.4.1 Stationarity Test — Sokoto

Table 6 presents the ADF stationarity test results for LogSokoto.

Table 6: Augmented Dickey–Fuller Unit Root Test - LogSokoto (First Difference)

Test Component	t-Statistic	p-value
ADF Test Statistic	−13.001	0.0000
Critical Value: 1% level	−3.444	—
Critical Value: 5% level	−2.868	—
Critical Value: 10% level	−2.570	—

Note: Lag length = 12, selected automatically based on AIC (maxlag = 17). Null hypothesis: series has a unit root.

The ADF statistic of −13.001 (p = 0.000) again strongly rejects the unit root null, confirming that LogSokoto is I(1). The ADF lag length of 12 - one more than for Katsina and Borno reflects the slightly higher autocorrelation structure in Sokoto's rainfall series, consistent with its greater variance. Figure 12 shows the first-differenced series. The correlogram for Sokoto reveals a spike at lag 1 of the PACF suggesting AR(1), an additional significant spike at lag 10 suggesting AR(10), and gradual decay in the ACF consistent with MA(1). This richer autocorrelation structure drives the consideration of higher-order non-seasonal components in the candidate model set.

4.4.2 Model Selection, Diagnostics, and Forecast - Sokoto

Table 7 presents model selection results for Sokoto.

Table 7: SARIMA Model Selection Criteria - Sokoto

Model	Sig. Coef.	σ^2	AIC	Log-Likelihood
SARIMA$(1,1,1)\times(1,2,1)_{12}$ ✓	4	0.1954	644.90	−310.28
SARIMA$(1,1,10)\times(1,2,1)_{12}$	2	0.1963	652.82	−312.41
SARIMA$(10,1,1)\times(1,2,1)_{12}$	3	0.1965	653.18	−312.59

✓ *Selected model.*

For Sokoto, the parsimonious SARIMA$(1,1,1)\times(1,2,1)_{12}$ outperforms the more complex alternatives, recording the highest number of significant coefficients (4), the lowest AIC (644.90), the lowest σ^2 (0.1954), and the highest log-likelihood (−310.28). This result is notable: unlike Katsina and Borno, where the high-order SARIMA(10,1,10) specification was necessary. Sokoto's rainfall dynamics are adequately captured by a first-order model. This may reflect a slightly smoother autocorrelation structure in Sokoto's series despite its higher overall variance, perhaps due to its geographic position further west in the Sahel where large-scale atmospheric forcing is more directly influential.

The Ljung–Box test for SARIMA$(1,1,1)\times(1,2,1)_{12}$ on Sokoto data yields $Q^* = 9.749$ (df = 6, p = 0.1356), confirming white noise residuals. Figure 12 presents the diagnostic plots.

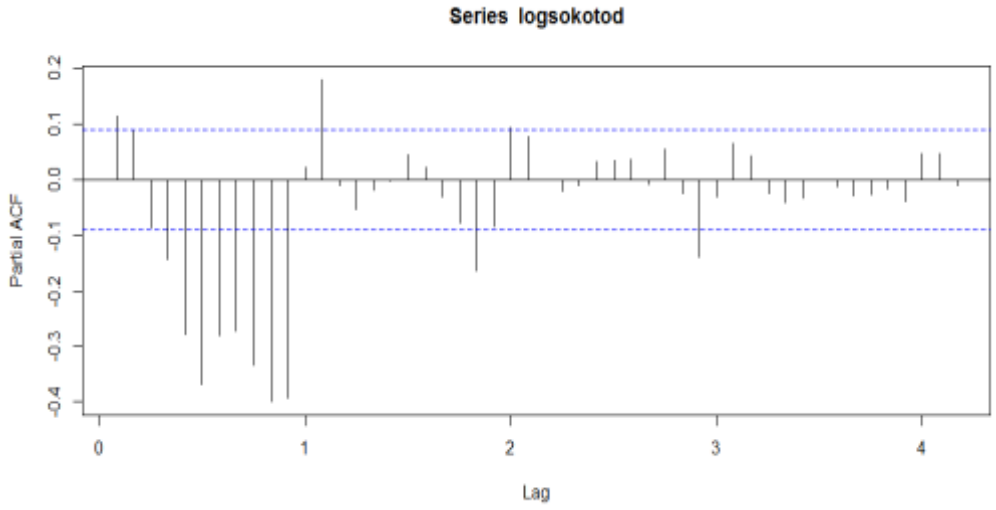

Figure 12: Residual Diagnostics for SARIMA$(1,1,1)\times(1,2,1)_{12}$ - Sokoto

The diagnostic plots confirm that residuals are homoscedastic, uncorrelated, and approximately normally distributed. Figure 13 presents the out-of-sample rainfall forecast for Sokoto.

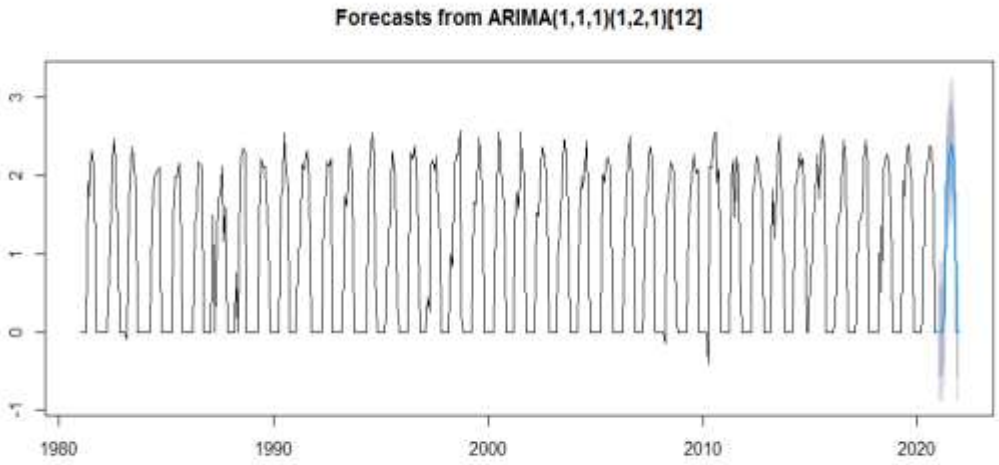

Figure 13: Out-of-Sample Rainfall Forecast with 95% Confidence Bands - SARIMA$(1,1,1)\times(1,2,1)_{12}$, Sokoto

The Sokoto forecast in Figure 13 displays the characteristic seasonal envelope of the region, with the highest projected values in July to August. The central forecast implies a gradual upward drift in annual rainfall for Sokoto, consistent with the broader pattern observed across all three states. The confidence intervals are noticeably narrower than those of Katsina and Borno, reflecting the lower model complexity and correspondingly tighter estimation uncertainty of the parsimonious SARIMA$(1,1,1)\times(1,2,1)_{12}$ specification.

4.5 Discussion

Table 8 consolidates the final model specifications, key diagnostic statistics, and forecast directions for all three states.

Table 8: Summary of Selected SARIMA Models and Diagnostic Results - All Three States

State	Best Model	AIC	σ^2	LB p-val	Forecast
Katsina	SARIMA$(10,1,10)\times(1,2,1)_{12}$	632.61	0.161	0.1148	↑ Rising
Borno	SARIMA$(10,1,10)\times(1,2,1)_{12}$	632.61	0.161	0.1148	↑ Rising
Sokoto	SARIMA$(1,1,1)\times(1,2,1)_{12}$	644.90	0.195	0.1356	↑ Rising

Several patterns emerge from the cross-state comparison. First, all three states share identical seasonal model structures, SAR(1), D = 2, SMA(1) reflecting the common Sahelian climate forcing that governs seasonal rainfall onset, peak, and recession across the region. Second, the non-seasonal orders differ: the high-order ARMA(10,10) structure required for Katsina and Borno reflects more complex short-memory autocorrelation dynamics in those series, while Sokoto's simpler ARMA(1,1) structure suggests smoother month-to-month transitions. Third, the σ^2 values are broadly comparable across states (0.161 – 0.195), indicating similar levels of unexplained residual variation after model fitting. Fourth, and most importantly from a policy standpoint, all three states' best models generate forecasts indicating a gradual upward trend in monthly rainfall, implying a modest reduction in aridity across the Sudano-Sahelian belt over the near-term horizon. A single seasonal difference (D=1) did not fully remove seasonal nonstationarity. The rainfall series in these semi-arid zones exhibits evolving seasonal patterns, due to changing cessation of rains, requiring a second seasonal difference.

This finding is consistent with recent studies documenting a partial recovery of Sahelian rainfall since the drought nadir of the 1970s – 1980s (Nicholson *et al.,* 2000), though the improvement remains modest relative to the long-run mean.

5. Conclusion

This study applied the Box–Jenkins SARIMA methodology to 40 years (1981–2020) of monthly rainfall data for Katsina, Borno, and Sokoto, three drought-prone states in northern Nigeria's Sudano-Sahelian belt. All three log-transformed series were confirmed I(1) by the Augmented Dickey–Fuller test, and strong annual seasonality (s = 12) was identified through ACF/PACF and seasonal decomposition analysis. Candidate SARIMA models were rigorously compared using AIC, σ^2, log-likelihood, and significant-coefficient count. The selected models SARIMA$(10,1,10)\times(1,2,1)_{12}$ for Katsina and Borno, and SARIMA$(1,1,1)\times(1,2,1)_{12}$ for Sokoto passed all diagnostic checks, with Ljung–Box p-values of 0.1148 and 0.1356 respectively, confirming white noise residuals.

Rainfall forecasts from all three models indicate a gradual upward trend in monthly precipitation, implying a mild reduction in aridity over the projection horizon. Peak rainfall is consistently projected in July to August, with onset variability between April and May, consistent with observed patterns over the study period. Katsina recorded the lowest mean monthly rainfall across the sample (45.39 mm) and Sokoto the highest (55.54 mm), and this ordering is reflected in the forecast trajectories. Based on the study findings, the following recommendations are advanced for policy and practice:

(i) Sustained monitoring and regular updating of SARIMA forecasts using NiMet data should be institutionalized to support anticipatory agricultural and water resource planning in the three states.

(ii) The seasonal rainfall calendar peak months of July–August, onset between April and May, and dry Months November – March should be formally integrated into crop planting schedules, market calendars, and irrigation planning frameworks.

(iii) National and state governments should intensify climate change mitigation efforts, as increasing temperatures and shifting precipitation patterns are primary drivers of aridity intensification (Shiru *et al*., 2019; Rashid *et al*., 2015).

(iv) Drip and micro-irrigation systems should be deployed in the most water-stressed areas to maximize agricultural output during low-rainfall periods, reducing dependence on highly variable seasonal rainfall.

(v) Soil moisture conservation practices, including mulching, contour bunding, and conservation tillage should be promoted to reduce evapotranspiration losses during the dry season.

(vi) Future research should extend the SARIMA framework to incorporate exogenous predictors (SARIMAX), explore ensemble approaches using CMIP6 GCM outputs, and apply the methodology to other vulnerable states in the region.

References

Abdul-Aziz, A. R., Anokye, M., Kwame, A., Munyakazi, L., & Nsowah-Nuamah, N. N. N. (2013). Modeling and forecasting rainfall pattern in Ghana as a seasonal ARIMA process: The case of Ashanti Region. *International Journal of Humanities and Social Sciences 3(12).*

Abiodun, B. J., & Lawal, K. A. (2013). Characterisation and dynamics of the West African monsoon. *Theoretical and Applied Climatology, 115(3–4), 709–722. https://doi.org/10.1007/s00704-013-0910-1*

Adefolalu, D. O. (1986). Rainfall trends for periods 1911–1980 over Nigeria. Theoretical and Applied Climatology, 39, 81–89.

Ampaw, E. M., Akuffo, B., Opoku, L. S., & Lartey, S. (2013). Time series modeling of rainfall in New Juaben Municipality of the Eastern Region of Ghana. Contemporary Research in Business and Social Sciences.

Ati, O. F. (1996). Inter-Tropical Discontinuity and rainfall variability in northern Nigeria. Journal of Arid Environments. 33(4), 519–524. https://doi.org/10.1006/jare.1996.0087[citation:2][citation:10]

Ati, O. F., Stigter, C. J., & Oladipo, E. O. (2002). A comparison of methods to determine the onset of the growing season in northern Nigeria. *International Journal of Climatology*, 22(6), 731–742.

Ayoade, J. O. (1973). Seasonal rainfall variability in Nigeria. *The Geographical Journal,* 139(3), 430–437.

Bari, S. H., Rahman, M. T. U., Hoque, M. A., & Hussain, M. M. (2015). Analysis of seasonal and annual rainfall trends in the Sylhet region of Bangladesh. Atmospheric Research.

Boudoin, M.-A., Vogel, C., Nortje, K., & Naik, M. (2017). In search of resilience through climate change risk assessments: Insights from the South African food–energy–water security nexus. *Jàmbá: Journal of Disaster Risk Studies, 9*(1), a360. https://doi.org/10.4102/jamba.v9i1.360

Box, G. E. P., & Jenkins, G. M. (1976). Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco.

Chonge, G., Simwa, R., & Nderitu, J. (2015). Time series modelling of rainfall in Uasin Gishu County, Kenya. Mathematical Theory and Modeling, 5(9), 37–51.

Dregne, H. E., & Chou, N. T. (1992). Global desertification dimensions and costs. In H. E. Dregne (Ed.), *Degradation and restoration of arid lands* (pp. 249–282). Texas Tech University.

Eni, D. D., & Adeyeye, F. J. (2015). Seasonal ARIMA modeling and forecasting of rainfall in Warri town, Delta State, Nigeria. *Journal of Geoscience and Environment Protection*, 3, 91–98.

Eludoyin, O. S., Adelekan, I. O., Webster, R., & Eludoyin, A. O. (2009). Monthly rainfall distribution in Nigeria. *Journal of Geography and Regional Planning*. 2(5), 121–130. https://doi.org/10.5897/JGRP.9000101[citation:4][citation:8]

Fowler, H. J., Blenkinsop, S., & Tebaldi, C. (2007). Linking climate change modelling to impacts studies. International Journal of Climatology, 27(12), 1547–1578.

Gaughan, A. E., & Waylen, P. R. (2012). Spatial and temporal precipitation variability in the Okavango–Kwando–Zambezi catchment, southern Africa. *Journal of Arid Environments,* 82, 19–30.

Hanafi, A. and Jauffret, S.A. (2008), "Long-term vegetation dynamics useful in monitoring and assessing desertification processes in the arid steppe, southern Tunisia", *Journal of Arid Environments*, *vol. 72 No. 4, pp. 557–572*

Hulme, M., Doherty, R., Ngara, T., New, M., & Lister, D. (1998). African climate change: 1900–2100. Climate Research, 17, 145–168.

Hyuwa, G. N. (2005). Climatic elements and agriculture in northern Nigeria. *Journal of African Studies.*

Inderjeet, S., & Sabita, J. (2008). SARIMA modelling of rainfall and temperature for Uttar Pradesh, India. Journal of Agrometeorology, 10(2), 115–122.

Jabrin, A., Aminu, U., & Abdullahi, M. K. (2014). SARIMA method for modeling and forecasting rainfall in Kano, Nigeria. *International Journal of Science and Nature,* 5(3), 413–419.

Kayano, M. T., & Sansígolo, C. (2008). *Interannual to decadal variations of precipitation and daily maximum and*

minimum temperatures in southern Brazil. Theoretical and Applied Climatology, 94, 357–372.

Lázaro, R., Rodrigo, F. S., Gutiérrez, L., Domingo, F., & Puigdefábregas, J. (2001). Analysis of a 30-year rainfall record (1967–1997) in semi-arid SE Spain for implications on vegetation. *Journal of Arid Environments, 48*(3), 373–395. https://doi.org/10.1006/jare.2000.0755

Leila S. and Masoud Y. (2012). An Empirical Study of the Usefulness of SARFIMA models in Energy Science. *International Journal of Energy Science.IJES vol2(2), page 12*.

Ljung, G. M., & Box, G. E. P. (1978). Estimating the dimension of a model. Annals of Statistics, Biometrika, 65(2), 297–303

LutLópez-Moreno, J. I., Vicente-Serrano, S. M., Gimeno, L., & Nieto, R. (2008). Stability of the seasonal distribution of precipitation in the Mediterranean region: Observations since 1950 and projections for the 21st century. *Geophysical Research Letters, 35*(10), L10703. https://doi.org/10.1029/2008GL034774z,

Martinez, E. Z., da Silva, E. A. S., & Fabbro, A. L. D. (2011). A SARIMA forecasting model to predict dengue in Campinas, Brazil. Revista da Sociedade Brasileira de Medicina Tropical, 44(4), 436–440.

Masih, I., & Maskey, S. (2014). A review of droughts on the African continent: A geospatial and long-term perspective. Hydrology and Earth System Sciences, 18, 3635–3649.

Mceihattan. (1975). Forecasting Precipitation Using SARIMA Model. Mathematical Theory and Modeling, 4(11), 51–56.

Nicholson, S. E., Some, B., & Kone, B. (2000). An analysis of recent rainfall conditions in West Africa. *Journal of Climate,* 13(14), 2628–2640.

Oguntunde, P. G., Abiodun, B. J., & Lischeid, G. (2012). Rainfall trends in Nigeria, 1901–2000. *Journal of Hydrology*, 411(3–4), 207–218.

Ogunrinde, A. T. (2012). Application of Box-Jenkins methodology to rainfall in Lagos and Maiduguri. Unpublished MSc Dissertation.

Oladipo, E. O. (1987). A comprehensive approach to drought and desertification in Northern Nigeria. Natural Hazards, 8, 235–261.

Oladipo, E. O. (1993). Some aspects of the spatial characteristics of drought in Northern Nigeria. Natural Hazards, 8, 171–188.

Rashid, M., Beecham, S., & Chowdhury, R. K. (2015). Assessment of trends in point rainfall using continuous wavelet transforms. Advances in Water Resources, 82, 1–15.

Salman, S. A., Shahid, S., Ismail, T., Rahman, N. B. A., Wang, X., & Chung, E. S. (2018). Unidirectional trends in daily rainfall extremes of Iraq. Theoretical and Applied Climatology, 134, 1165–1177.

SAWA, B. A., Adebayo, A. A., Musa, A. A., & Abubakar, I. U. (2020). Climate and vegetation of the Sudano-Sahelian belt of Nigeria. *African Journal of Environmental Science and Technology.*

Schwarz, G. (1978) Estimating the dimension of a model. *Annals of Statistics,vol* 6, pages 461–464.

Shiru, M. S., Shahid, S., Alias, N., & Chung, E. S. (2019). Changing characteristics of drought in Nigeria during 1901–2010. Atmospheric Research, 223, 60–70.

Suhartono, & Lee, M. H. (2011). Forecasting of tourist arrivals using subset, multiplicative or additive seasonal ARIMA model. MATEMATIKA, 27(2), 169–182.

Sung, J. H., & Chung, E.-S. (2014). Development of a drought index considering seasonality and persistence. *Journal of Hydrology, 519*(Part A), 3051–3065. https://doi.org/10.1016/j.jhydrol.2014.09.072

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. Bulletin of the American Meteorological Society, 93, 485–498.

Usman, M. T. (2000). *Intra-seasonal rainfall variability and moisture efficiency in the Sudano-Sahelian belt of Nigeria* [Unpublished doctoral dissertation or report].

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A multiscalar drought index: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate,* 23(7), 1696–1718.

Wiredu, S., Nasiru, S., & Asamoah, Y. G. (2013). Proposed SARIMA model for forecasting rainfall in the Navrongo *Journal of Environment and Earth Science.* Municipality of Ghana.