



SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL OF NIGERIA RAINFALL PATTERN

Emmanuel Kole^{1*} and Adetunji K. Ilori²

¹Department of Mathematics and Statistics Kaduna Polytechnic, Kaduna.

²Statistics Programme, National Mathematical Centre, Abuja.

*Corresponding Author email: koleemmanuel@kadunapolytechnic.edu.ng

ABSTRACT

This study analyzes and forecasts rainfall patterns in Kano State, Nigeria, using a Seasonal Autoregressive Integrated Moving Average (SARIMA) model. A monthly rainfall dataset spanning January 2016 to December 2022 was obtained from the relevant government agency in Nigeria. Descriptive statistics from January 2016 to December 2022 reveal significant monthly rainfall variations, with a mean of 70.23 mm and a maximum of 323.90 mm, highlighting pronounced seasonality. Time series decomposition confirms a strong annual cycle, with distinct wet and dry periods. The Augmented Dickey-Fuller (ADF) test indicates that the original rainfall series is non-stationary, necessitating seasonal differencing to achieve stationarity. Different ARIMA models were compared using the Akaike Information Criterion (AIC), with the ARIMA(0,0,1)(0,2,4)[12] model emerging as the best fit (AIC = 682.82). Parameter estimation and diagnostic checks confirm the model's suitability, and residual analysis supports its predictive accuracy. Forecasting for the next year suggests continued seasonal rainfall patterns, aiding climate monitoring, agricultural planning, and water resource management in the region. The study underscores the effectiveness of SARIMA models in capturing seasonal trends and informing decision-making processes related to climate variability.

Keywords: Rainfall forecasting, SARIMA model, Time series analysis, Kano State Rainfall, Stationarity test

1. INTRODUCTION

Nigeria's tropical climate is characterized by distinct rainy and dry seasons, which vary significantly across different regions. The southeast region is marked by high temperatures and humidity, while the southwest and inland areas are generally dry (Onafeso, O.D., 2023). In contrast, the northern and western parts of the country have a savanna climate with pronounced wet and dry seasons. The far north has a steppe climate with Forecasting limited rainfall.

In Kano State, the tropical savanna climate is influenced by the West African monsoon, resulting in most rainfall occurring between May and October. Annual rainfall totals range from 800 mm in the north to 1100 mm in the south, leading to variations in climate and vegetation. Notably, research has shown that Kano State is experiencing increased rainfall due to climate change, with annual totals rising at a rate of 21.26 mm per annum. (Okemini Nwagbara, M., 2015)

Kano State has experienced extreme droughts, particularly during the droughts of the early 1970s and the 1980s that devastated northern Nigeria. Conversely, flood disasters as a result of heavy rainfall have been recorded. In August, 1988 for example, severe flooding in Kano State resulted in the loss of 146 lives, destruction of 180,000 houses, washing away of 14,000 farms, 194 displacement of 200,000 people, collapsed of the Bagauda dam and damage to residences and infrastructure worth 560 million naira (NEST, 1991)

This increase in rainfall has both beneficial and detrimental effects on the climate. On the one hand, it provides more water for agriculture and replenishes water sources. On the other hand, it poses a risk of flooding and other disasters. To mitigate these risks, it is essential to develop sustainable agricultural practices and invest in infrastructure such as dams (Mehta, P., 2024).

Due to the popularity of time series, numerous intelligent time series models have been created to increase the precision and effectiveness of time series forecasting. Time series analysis includes techniques for examining time series data in order to extract significant statistics and other features inherent in the data, as well as to create suitable models that describe future values based on previously observed values (Adhikari and Agrawal, 2014).

He, Zhang, and Chew (2022) combine machine learning methods with seasonal-trend decomposition to introduce a new method for modeling and forecasting rainfall time series. To isolate and analyze the seasonal and trend elements found in rainfall data, the researchers implement seasonal-trend decomposition methods. To improve the precision of rainfall forecasts, they then apply machine learning algorithms to the dissected data. By addressing the intricacies and non-linear patterns frequently found in meteorological time series data, this methodology seeks to provide a more reliable forecasting framework. The results of the study indicate that integrating machine learning models with decomposition approaches can greatly enhance predictive performance, providing insightful information for agricultural planning and water resource management.

The use of time series models for rainfall forecasting is examined by Sidiq (2018), who highlights how well these models capture temporal patterns and trends. The study models and forecasts future rainfall patterns using a variety of statistical approaches. The Autoregressive Integrated Moving Average (ARIMA) model, which is popular for time series forecasting because of its ability to manage stationarity, trends, and seasonality adjustments, is a major area of study. The study emphasizes how crucial data pre-processing is for increasing model accuracy, including parameter selection and stationarity checks. Sidiq (2018) uses the Box-Jenkins methodology to systematically find the best ARIMA parameters using partial autocorrelation (PACF) and autocorrelation (ACF) functions. The findings show that ARIMA models are useful instruments for climate researchers since they can accurately forecast short-term rainfall changes.

This study aimed to develop and evaluate a Seasonal Autoregressive Integrated Moving Average (SARIMA) model for forecasting rainfall patterns in Kano State, Nigeria. By fitting and comparing different SARIMA models based on the Akaike Information Criterion (AIC), this study seeks to identify the most suitable model for capturing seasonal variations and long-term trends in rainfall data. The findings will provide valuable insights for climate monitoring, agricultural planning, and water resource management in the region.

2. METHODOLOGY

Data Collection

The monthly rainfall dataset was extracted from relevant government agencies in Nigeria from January 2016 to December 2022 for Kano State.

Seasonal ARIMA Models

Seasonality in a time series refers to a consistent pattern of variations that recurs over S time intervals, where S denotes the number of time intervals until the pattern reoccurs (Dufour, J. -M., Hecq, A., and Wan, A. 2022).

For instance, there exists seasonality in monthly data where elevated values consistently appear in certain months, while lower values generally emerge in other specific months. In this instance, $S = 12$ (months per year) indicates the duration of the cyclical seasonal behavior. For quarterly data, $S = 4$ time intervals each year.

In a seasonal ARIMA model, seasonal AR and MA terms predict x_t using data values and errors at times with lags that are multiples of S (the span of the seasonality).

With monthly data (and $S = 12$), a seasonal first order autoregressive model would use x_{t-12} to predict x_t . For instance, if we were selling cooling fans we might predict this August’s sales using last August’s sales. (This relationship of predicting using last year’s data would hold for any month of the year.)

A seasonal second order autoregressive model would use x_{t-12} and x_{t-24} to predict x_t . Here we would predict this August’s values from the past two Augusts.

A seasonal first order MA(1) model (with $S = 12$) would use w_{t-12} as a predictor. A seasonal second order MA(2) model would use w_{t-12} and w_{t-24}

The seasonal ARIMA model incorporates both non-seasonal and seasonal factors in a multiplicative model. One shorthand notation for the model is

$$ARIMA(p, d, q) \times (P, D, Q)S \tag{1}$$

Where

- p is the non-seasonal AR order,
- d is the non-seasonal differencing
- q is the non-seasonal MA order
- $P =$ seasonal AR order
- $D =$ seasonal differencing
- $Q =$ seasonal MA order
- and S is the seasonal period

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

$$\phi(B^S)\phi(B)(x_t - \mu) = \theta(B^S)\theta(B)w_t \tag{2}$$

Where

$\phi(B)$ and $\phi(B^S)$ are the non-seasonal and seasonal AR components, respectively

$\theta(B)$ and $\theta(B^s)$ are the non-seasonal and seasonal MA components, respectively and w_t is the white noise error term.

Akaike's Information Criterion (AIC)

The Akaike Information Criterion (AIC) is a measure used to assess the goodness of fit of a statistical model while penalizing for the number of estimated parameters. It helps in model selection by balancing the model's accuracy and complexity. The model for AIC is:

$$AIC = 2k - 2\ln(L) \tag{3}$$

Where:

k is the number of parameters in the model.

L is the likelihood of the model (or the maximized value of the likelihood function).

In the case of a regression model,

L represents the likelihood of observing the data given the model's parameters. A lower AIC value indicates a better model, as it suggests a model that provides a good fit with fewer parameters.

3. DATA ANALYSIS AND DISCUSSION OF RESULTS

TABLE 1: Descriptive Statistics of Kano State Rainfall (mm) (January, 2016 – December, 2022)

Statistic	Value
Minimum	0.00
First Quartile	0.00
Median	5.30
Mean	70.23
Third Quartile	135.15
Maximum	323.90
Variance	9977.55
Standard Deviation	99.89

Table 1 presents the descriptive statistics for Kano State rainfall (January 2016 – December 2022) demonstrating significant fluctuation in monthly rainfall. The minimum rainfall recorded is 0.00 mm, indicating dry months, while the maximum reaches 323.90 mm, showing periods of heavy rainfall. The median rainfall is relatively low at 5.30 mm, implying that more than half of the months get little rainfall. Due to significant rainfall occurrences, the distribution is biased to the right, with the mean rainfall of 70.23 mm being higher than the median. It is also confirmed by the first quartile (0.00 mm) and third quartile (135.15 mm) that rainfall is generally modest with sporadic peaks. Significant variations in monthly rainfall throughout the observed period are shown in the large variance (9977.55) and standard deviation (99.89 mm).

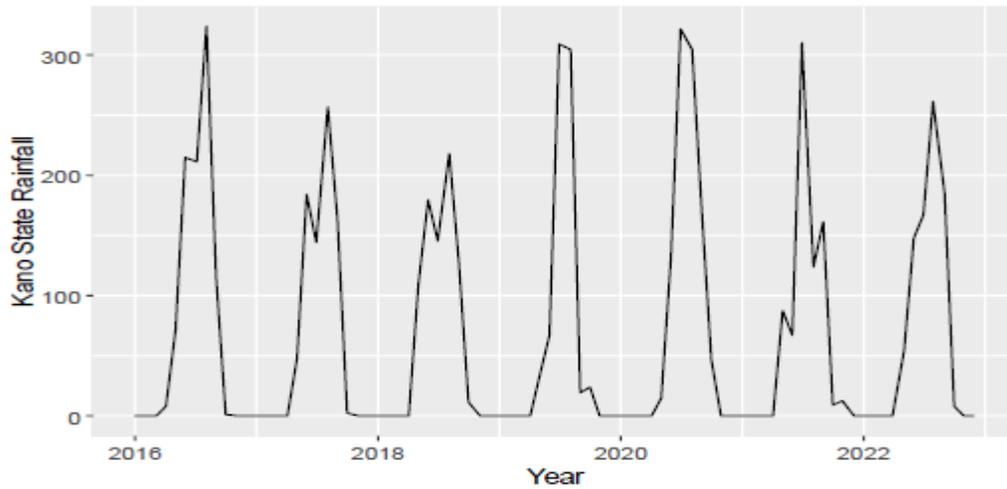


Figure 1: Time series plot of Kano State rainfall (mm)

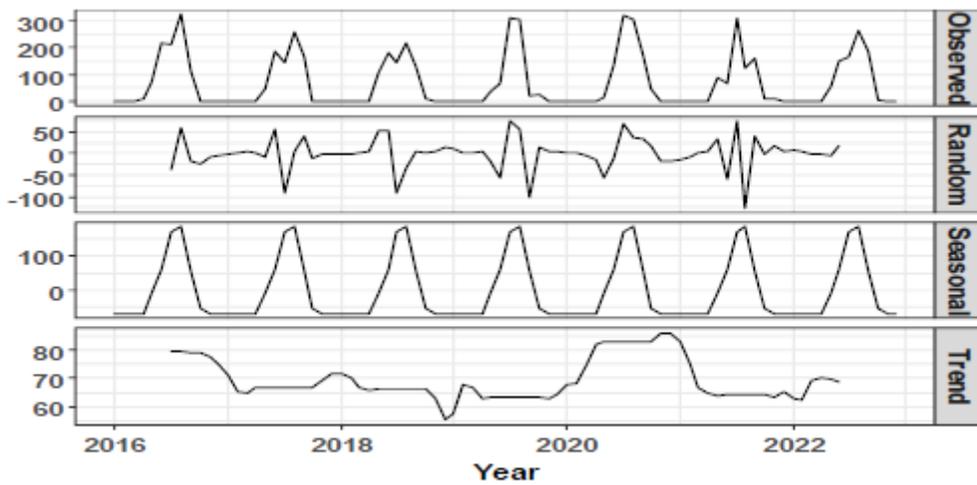


Figure 2: Decomposition Time Plot of Kano State Rainfall

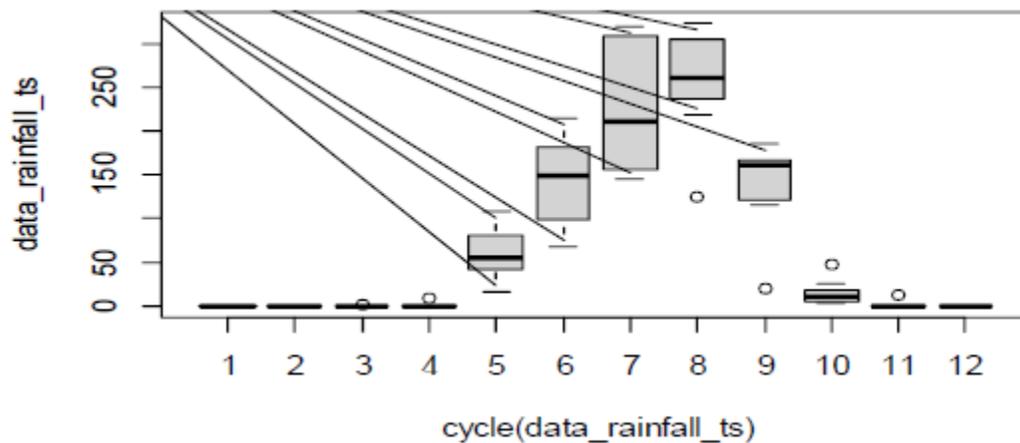


Figure 3: Box plot of Kano State rainfall

Figure 1 presents a time series plot of monthly rainfall in Kano State from Jan, 2016 to Dec, 2022. The pattern exhibits clear seasonality, with distinct peaks and troughs recurring annually. The peaks, representing periods of high rainfall, appear consistently each year, indicating a strong seasonal cycle. The lowest values, close to zero, suggest dry months with little to no rainfall. The amplitude of the peaks varies slightly across years, reflecting interannual variability in rainfall intensity. Overall, the plot highlights the cyclical nature of rainfall in Kano State, likely influenced by climatic and meteorological factors. The decomposition time plot of Kano State rainfall from 2016 to 2022 as shown in figure 2 provides a detailed breakdown of the rainfall patterns into four key components: observed data, random variations, seasonal fluctuations, and long-term trends. The first panel displays the observed rainfall data, highlighting periodic fluctuations that suggest a strong seasonal pattern. The peaks and troughs indicate recurring high and low rainfall periods within each year. The second panel represents the random component, which captures irregular variations that are not explained by the trend or seasonal effects. These fluctuations appear inconsistent over time, reflecting the unpredictable nature of short-term rainfall changes. The third panel illustrates the seasonal component, revealing a clear annual cycle in rainfall patterns. This suggests that rainfall follows a consistent seasonal trend, with specific months experiencing higher precipitation levels, likely corresponding to the region's wet and dry seasons. The final panel shows the trend component, which represents long-term changes in rainfall. The trend remains relatively stable over the years, with minor fluctuations, indicating that while seasonal patterns dominate the data, there are also underlying shifts in rainfall levels over time. This decomposition provides valuable insights into the behavior of rainfall in Kano State, highlighting the significance of seasonality, the presence of a stable trend, and the influence of irregular variations on precipitation levels.

Stationarity Test

The Augmented Dickey-Fuller (ADF) test was used to investigate the stationarity of the Kano State rainfall series. Unit root tests are frequently used to evaluate the stationarity of the series. The ADF test aids in identifying non-stationarity in the data by detecting the presence of a unit root (Moosa, I.A., Bhatti, R.H. 1997). According to the test results, the original rainfall series needed to be transformed because it was non-stationary. Initially, seasonal differencing was used to get rid of seasonality and trends. Following this adjustment, the series became stationary at a 5% significance level, as indicated by the Dickey-Fuller test statistic of -3.7178 and p-value of 0.0295. Furthermore, Figures 7 and 8 Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots provided additional evidence that the Kano State rainfall series had stabilized.

Series data_rainfall_ts

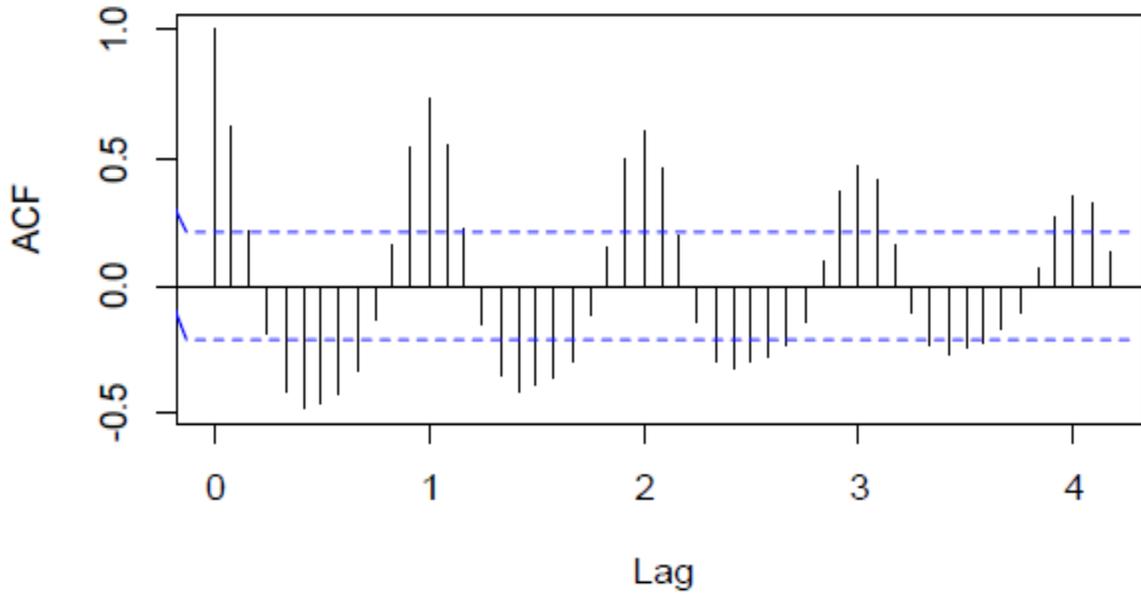


Figure 4: ACF Plot of the raw data

Series data_rainfall_ts

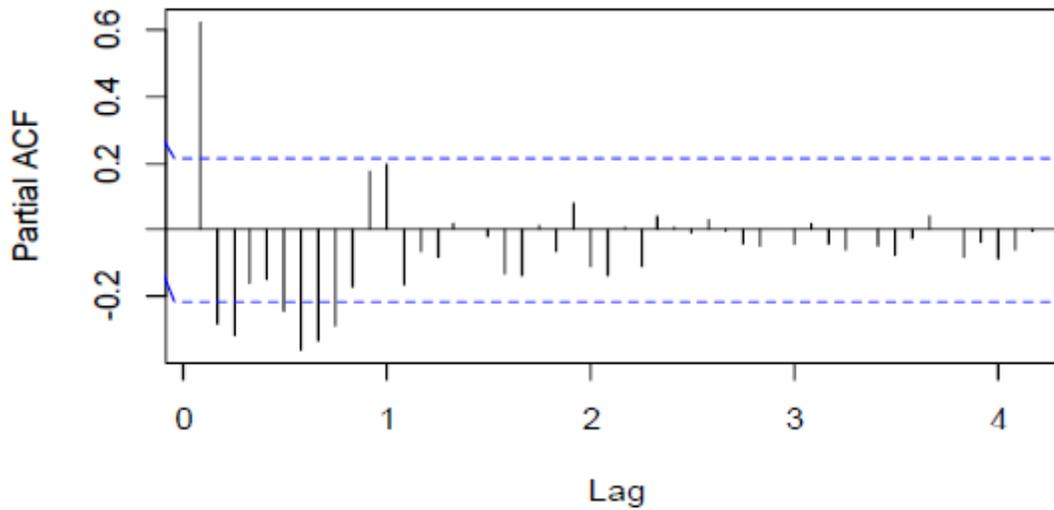


Figure 5: PACF Plot of the raw data

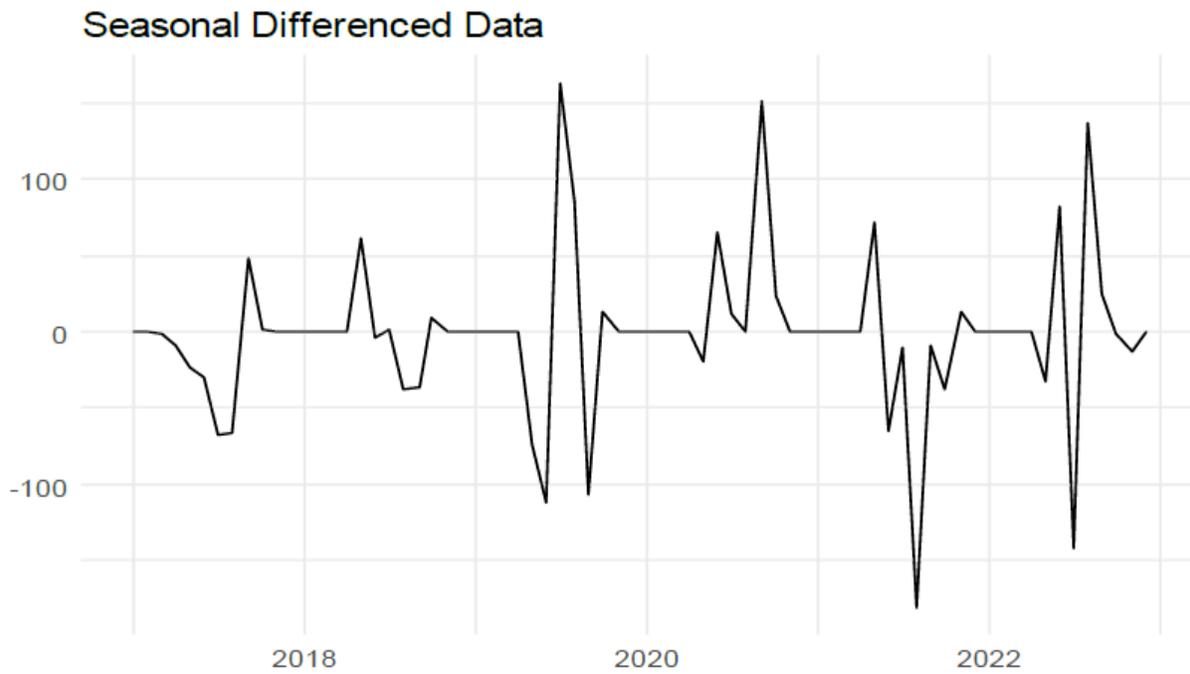


Figure 6: Seasonal differenced data

Series seasonal_diff

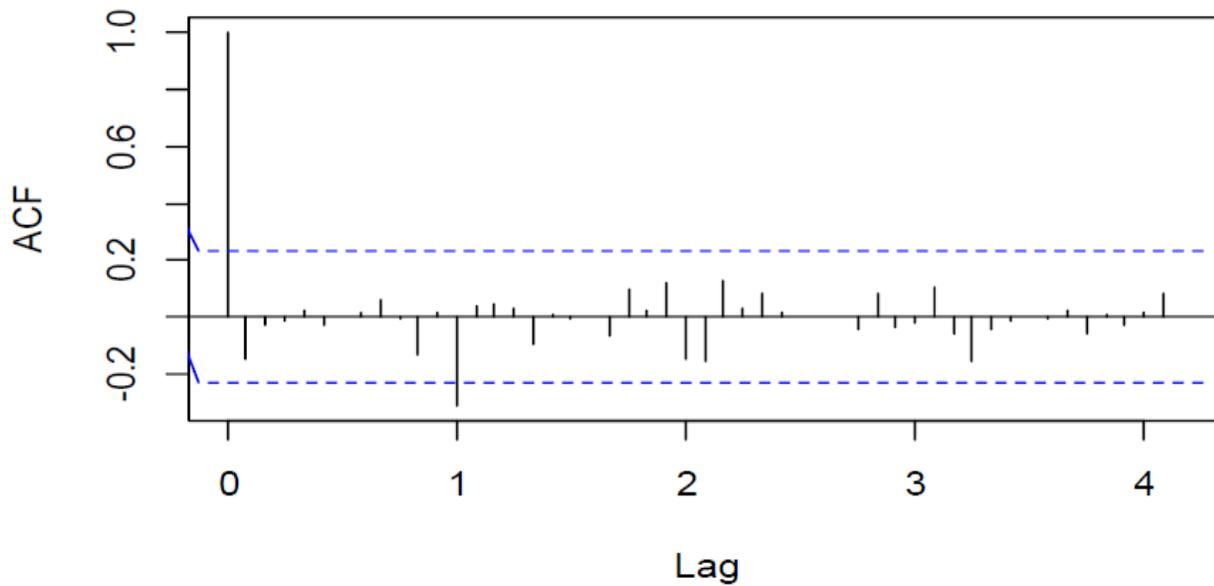


Figure 7: ACF Plot of the seasonal differenced data

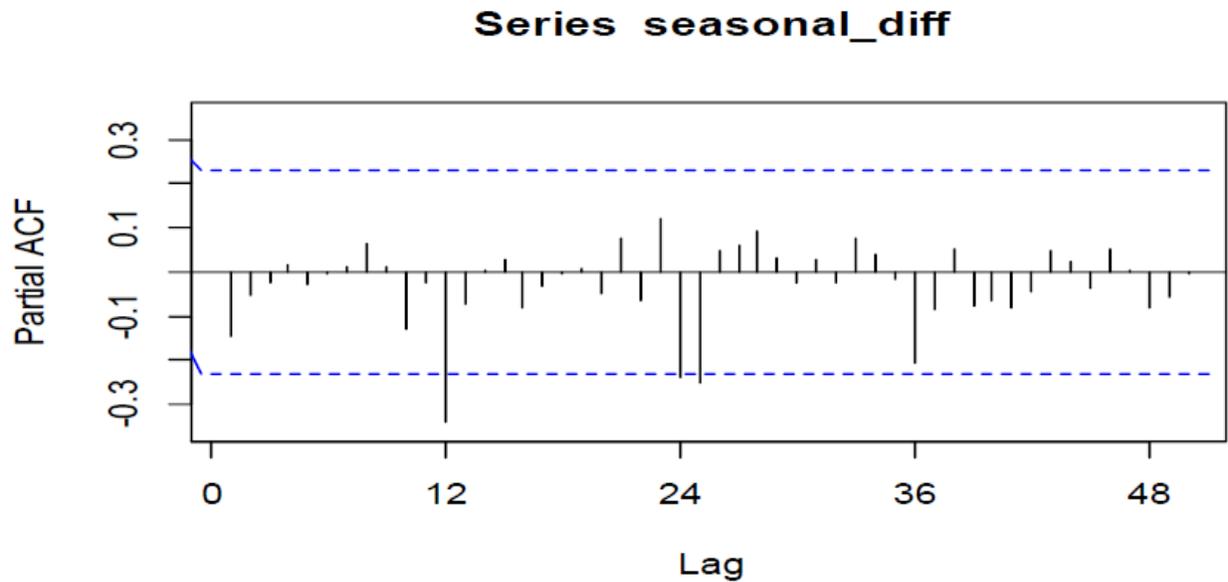


Figure 8: PACF Plot of the seasonal differenced data

Model Selection

Table 2 compares different ARIMA (Auto Regressive Integrated Moving Average) models used to forecast rainfall patterns in Kano State. The models are evaluated based on their Akaike Information Criterion (AIC) values, a metric used to assess model fit, where lower AIC values indicate a better balance between goodness of fit and model complexity. Among the models presented, the ARIMA (0,0,1) (0,2,4) (12) has the lowest AIC value of 682.82, making it the best combination for modeling Kano State rainfall. This suggests that it provides the best predictive performance while avoiding overfitting. The seasonal order (12) indicates that the model accounts for annual seasonality, which aligns with the expected seasonal variations in rainfall. The ARIMA (0,0,2) (1,2,2) (12) model follows closely with an AIC of 682.89, suggesting that it also performs well but slightly less efficiently than the first model. The ARIMA (0,0,1) (0,2,1) (12) model has a marginally higher AIC (683.04), meaning it is a slightly less optimal choice. The ARIMA (1,0,1) (1,2,2) (12) model has the highest AIC value (684.58), indicating it is the least favorable model among the four considered. The presence of additional autoregressive (AR) and moving average (MA) terms may have increased complexity without significantly improving model performance.

TABLE 2: ARIMA Model for Kano State Rainfall

Model	AIC
ARIMA (0, 0,1) (0, 2, 4) (12)	682.82
ARIMA (0, 0,1) (0, 2, 1) (12)	683.04
ARIMA (0, 0,2) (1, 2, 2) (12)	682.89
ARIMA (1, 0,1) (1, 2, 2) (12)	684.58

Kano State Rainfall

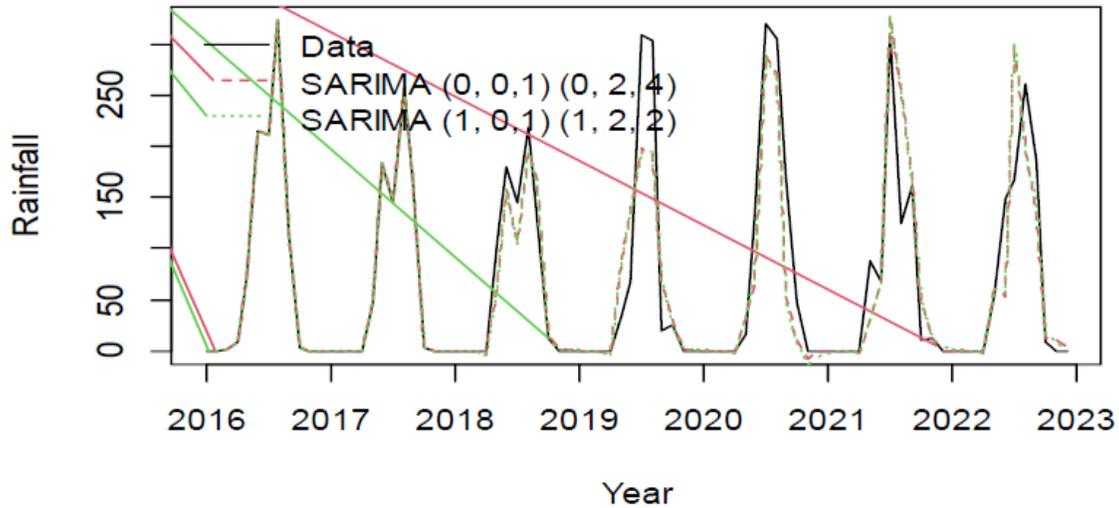


Figure 9: Performance of ARIMA (0, 0,1) (0, 2, 4) (12) and ARIMA (1, 0,1) (1, 2, 2) (12) with raw data

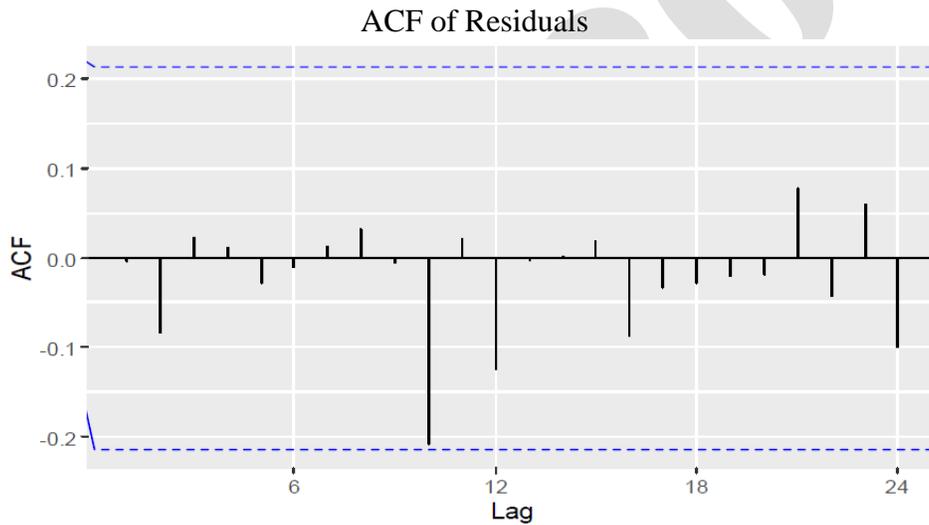


Figure 10: ACF plot of SARIMA (0,0,1) (0,2,4)[12] Residuals

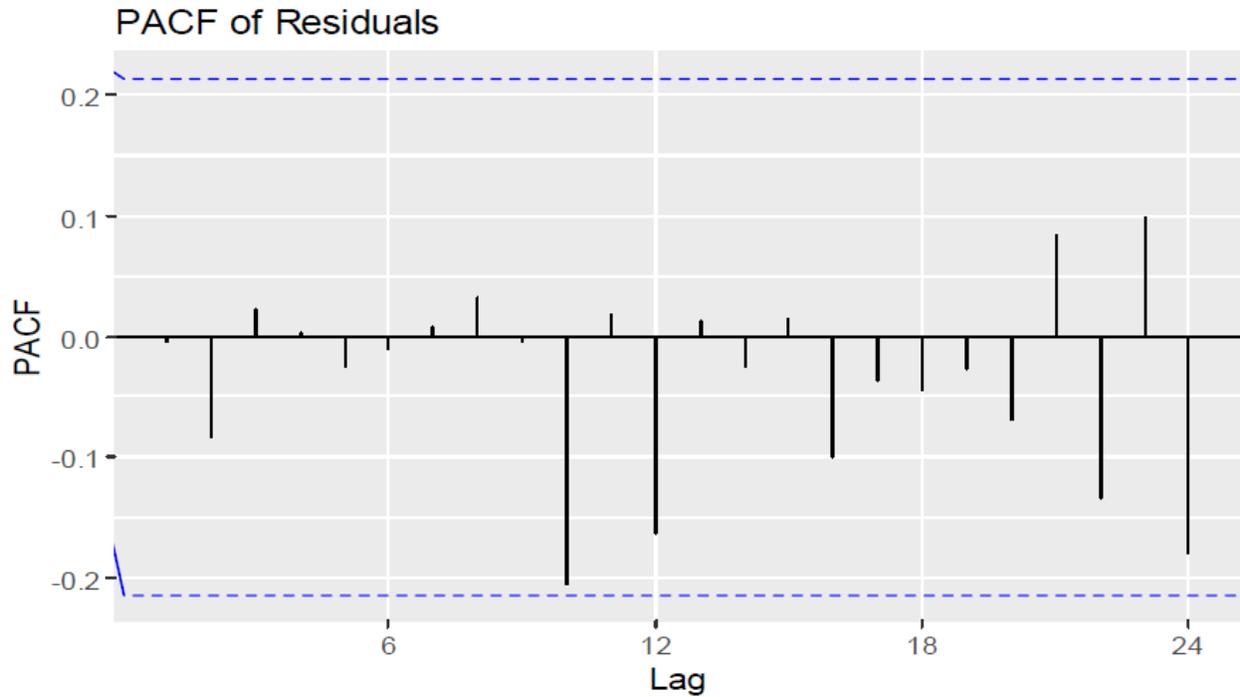


Figure 11: PACF plot of SARIMA (0,0,1) (0,2,4)[12] Residuals

Model Parameter

Table 3 presents the estimated parameters of the ARIMA (0,0,1)(0,2,4)[12] model, which was used to analyze and forecast rainfall patterns in Kano State. The table includes the estimated coefficients for the moving average (MA) and seasonal moving average (SMA) terms, along with their respective standard errors (S.E.).

TABLE 3: Parameter Estimation

Parameter	Coefficient	Standard Error (S.E)
MA(1)	-0.2119	0.1348
SMA(1)	-1.6071	0.656
SMA(2)	0.5081	0.5411
SMA(3)	0.2802	0.2947
SMA(4)	-0.0136	0.2377

An ARIMA(0,0,1)(0,2,4)[12] model follows the form:

$$Y_t = -0.2119e_{t-1} - 1.6071e_{t-12} + 0.5081e_{t-24} + 0.2802e_{t-36} - 0.0136e_{t-48} + e_t$$

Diagnostic Check

The estimated model was subjected to a diagnostic check to check for the suitability of the forecasted model and residual testing was also carried out on the models.

The performance metrics of different types of SARIMA models are compared by calculating forecasting accuracy statistics using Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). As shown

in Table 4, SARIMA (0, 0,1) (0, 2, 4) [12] recorded lowest value for performance metrics considered. Figure 6 shows ACF plot of the residuals of the models.

TABLE 4: Performance Matrix

Model	RMSE	MAE
SARIMA (0, 0,1) (0, 2, 4) [12]	37.2697	18.07851
SARIMA (0, 0,1) (0, 2, 1) [12]	48.2051	23.28614
SARIMA (0, 0,2) (1, 2, 2) [12]	37.5719	18.24945
SARIMA (1, 0,1) (1, 2, 2) [12]	37.6950	18.28221

Forecasting

Using the identified model, the future rainfall in Kano State, Nigeria for the next one year was forecasted. Figure 7 shows how the nature of the rainfall for the next one year will look like.

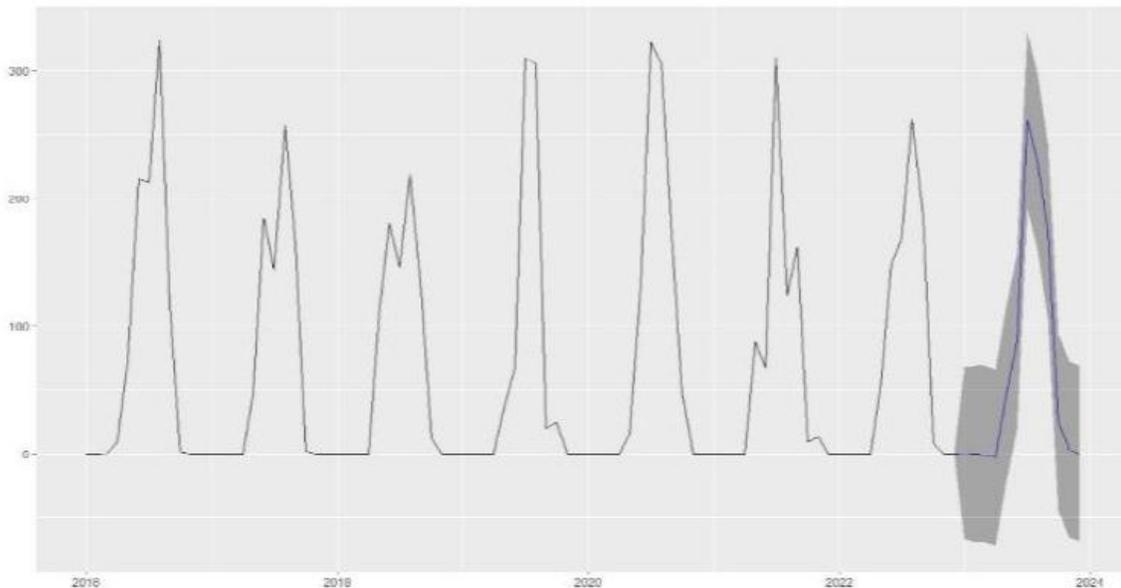


Figure 12: Forecasted Kano State rainfall with SARIMA (0, 0,1) (0, 2, 4) [12]

4. CONCLUSION AND RECOMMENDATION

The analysis of Kano State rainfall from January 2016 to December 2022 reveals significant variations in monthly precipitation, characterized by pronounced seasonal fluctuations. Descriptive statistics highlight a highly skewed distribution, with the median rainfall (5.30 mm) being significantly lower than the mean (70.23 mm), indicating that dry months are predominant while occasional heavy rainfall drives the overall average upward. The large variance (9977.55) and standard deviation (99.89 mm) further confirm the unpredictability of rainfall patterns in the region.

The time series plot of monthly rainfall confirms a well-defined seasonal pattern, with recurring peaks and troughs each year. This seasonality is further supported by the decomposition analysis, which separates the data into observed, random, seasonal, and trend components. The seasonal component

clearly illustrates an annual cycle, while the trend component remains relatively stable, suggesting that while interannual variability exists, no major long-term shifts in rainfall patterns are observed.

A stationarity test using the Augmented Dickey-Fuller (ADF) test initially indicated that the rainfall series was non-stationary. However, after applying seasonal differencing, the series became stationary, making it suitable for time series modeling. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots further confirmed that the transformed data was appropriate for modeling.

To forecast rainfall, different ARIMA models were evaluated based on their Akaike Information Criterion (AIC) values. The ARIMA (0,0,1)(0,2,4)[12] model was identified as the best-performing model, having the lowest AIC value (682.82). This model effectively captures the seasonal rainfall patterns while maintaining a balance between complexity and predictive accuracy. Parameter estimation further confirmed the dominance of seasonal moving average terms, reinforcing the importance of seasonal components in explaining rainfall variations.

Diagnostic checks and residual tests confirmed the suitability of the chosen model, as it demonstrated the lowest forecasting error based on Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) values. Using this model, rainfall for the next year was forecasted, providing insights into expected precipitation trends.

Overall, the study highlights the strong seasonal nature of rainfall in Kano State, driven by climatic and meteorological factors. The identified ARIMA model serves as a reliable tool for short-term forecasting, which can aid in water resource management, agricultural planning, and disaster preparedness. Future studies could explore additional external variables, such as temperature and atmospheric pressure, to enhance the model's predictive capabilities.

Conflict of interest: The authors declare that there is no conflict of interest

REFERENCES

- Adhikari, R., and Agrawal, R. K. (2014). A combination of artificial neural network and random walk models for financial time series forecasting. *Neural Computing and Applications* 24 (6), 1441-1449.
- Dufour, J.-M., Hecq, A., & Wan, A. (2022). EcoSta special issue on theoretical econometrics. *Econometrics and Statistics*, 25, 1–2. <https://doi.org/10.1016/j.ecosta.2022.02.001>
- He, R., Zhang, L., & Chew, A. W. Z. (2022). Modeling and predicting rainfall time series using seasonal-trend decomposition and machine learning. *Knowledge-Based Systems*, 251, 109125. <https://doi.org/10.1016/j.knosys.2022.109125>
- Mehta, P. (2024). The Impact of Climate Change on the Environment, Water Resources, and Agriculture: A Comprehensive Review. In: Rai, P.K., Rai, S. (eds) *Climate, Environment and Agricultural Development. Advances in Geographical and Environmental Sciences*. Springer, Singapore. https://doi.org/10.1007/978-981-97-8363-2_1

- Moosa, I.A., Bhatti, R.H. (1997). Testing for Stationarity and Unit Root. In: International Parity Conditions. Palgrave Macmillan, London. https://doi.org/10.1007/978-1-349-25523-8_7
- Nigerian Environmental Study/Action Team, NEST, 1991. Nigeria's Threatened Environment: A National Profile. Nigerian Environmental Study/Action Team (NEST). A NEST Publication. Ibadan, Nigeria.
- Okemini Nwagbara, M. (2015) Case Study: Emerging Advantages of Climate Change for Agriculture in Kano State, North-Western Nigeria. *American Journal of Climate Change*, **4**, 263-268. doi: 10.4236/ajcc.2015.43021.
- Onafeso, O.D. (2023). The Climate of Nigeria and Its Role in Landscape Modification. In: Faniran, A., Jeje, L.k., Fashae, O.A., Olusola, A.O. (eds) Landscapes and Landforms of Nigeria. World Geomorphological Landscapes. Springer, Cham. https://doi.org/10.1007/978-3-031-17972-3_2
- Sidiq, M. (2018). Forecasting rainfall with time series model. *IOP Conference Series: Materials Science and Engineering*, **407**(1), 012154. <https://doi.org/10.1088/1757-899X/407/1/012154>