Journal of Applied Econometrics and Statistics
Vol. 2, No. 1, 2023, pp. 41-62
https://DOI:10.47509/JAES.2023.v02i01.03

ARF INDIA
Academic Open Access Publishing
www.arfjournals.com

Time Series Modelling and Forecasting of Seasonal Rainfall Patterns in Bida Basin, Nigeria

Mayaki, J.[1], Sule, I. M.[2,4] and Bako, D.[3]

[1]*Department of Statistics, Federal University Minna, Niger State, Nigeria. E-mail: j.mayaki@futminna.edu.ng*
[2]*Department of Geography, Federal University Minna, Niger State, Nigeria. E-mail: isaiah.sule@futminna.edu.ng*
[3]*Department of Mathematics, Federal University Minna, Niger State, Nigeria. E-mail: deborah.bako@futminna.edu.ng*
[4]*WASCAL CC& HH, Federal University Minna, Niger State, Nigeria.*

Received: 14 February 2023; Revised: 24 March 2023;
Accepted 29 March 2023; Publication: 26 June 2023

Abstract: The study titled Time Series Modelling and Forecasting of Seasonal Rainfall Patterns in Bida Basin, Nigeria, modelled and forecasted time series data from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 reanalysis precipitation data from 1981-2020. The mean method, naïve method and seasonal naïve method were used as benched mark forecasting method to linear regression methods, exponential methods and ARIMA methods. The R-statistical package was used for analysing the data. In choosing among exponential methods and ARIMA methods, the est() and auto.arima() functions were employed respectively to select model that best capture the features of the time series data while the R-squared was employed to select from linear regression models. The scaled methods for prediction error evaluations and residual analysis were performed on the various models considered. The linear regression, seasonal ARIMA and exponential smoothing methods were better than the three benched mark forecasting methods. However, root mean squared error (RMSE) evaluation showed that the linear regression model with trend and seasonality had the closest predicted value to the actual rainfall values. With the mean absolute error (MAE) and the mean absolute scaled error (MASE), ARIMA $(0,0,0)(0,1,1)_{[12]}$ has the least value. Based on the results from the residual analysis and error dependent evaluation methods, linear regression model was selected for modelling and forecasting Bida basin rainfall since it had the closest prediction value to the actual value.

Keywords: Rainfall, Modelling, Forecasting, Bida Basin, Pattern

1. INTRODUCTION

Rainfall has been identified as a key climatic variable in climate sciences and hydrology; as changes in its pattern may have diverse effects on human

for citation :
Mayaki, J., Sule, I. M. and Bako, D. (2023). Time Series Modelling and Forecasting of Seasonal Rainfall Patterns in Bida Basin, Nigeria. *Journal of Applied Econometrics and Statistics*, Vol. 2, No. 1, pp. 41-62. https://DOI: 10.47509/JAES.2023.v02i01.03

wellbeing, ecosystems, flora, and fauna (Yila *et al*., 2023). It plays key roles in the hydrologic cycle; with changes in its patterns directly influencing the water resources of a given area by altering the spatiotemporal distribution of runoff, soil moistness, and groundwater reserve (Yila *et al*., 2023). Water resources are so important and the need for them are increasing essentially for many purposes such as transportation, power generation, domestic consumption, agricultural activities and industry (Elouissi *et al*., 2016). However, studies have revealed global impacts of climate change and variability on water availability due to alterations in rainfall patterns, with Africa predicted to experience the worst effects (Oti *et al*., 2020). The uncertainty of the rainfall events will have several effects on the water resources and water demands especially household water use, agriculture use and hydropower generation among several others (Oti *et al*., 2020). consequently, it becomes imperative to examine the trends of rainfall in order to understand climate variability and change because it is very variable spatio-temporarily at all scales (Yila *et al*., 2023).

The work of Abe *et al*. (2022) which analysed and forecasted rainfall patterns in Gombe Nigeria revealed an alternation of wet and dry period. Also, time series analysis of rainfall data of Kano, Katsina and zaria meteorological stations by Ekpoh (2007) revealed a decrease in mean annual rainfall for the three station. Changes in the pattern of rainfall is largely due to threat posed by climate change witnessed globally and therefore, understanding the pattern of rainfall and predicting it future occurrence is necessary for the planning of various human activities. Rainfall being a random phenomenon, the paper consider some forecasting procedure to choice the one that provide better estimates for the future.

The Bida basin, which is the focal area of this study is quite important in the sustenance of livelihoods of Nigerians and plays a vital role in the nation's food security. This is because its nutrient-rich soil provides livelihoods to majority of its rural and urban dwellers. The nutrient-rich flood plains (fadamas) of the basin allow for large scale cultivation of rice and sugar production while the upland areas support the cultivation of a numerous varieties of other staple foods and cash crops such as maize, sorghum, millet, yam, cassava, sweet potatoes, hot pepper, sesamum, cowpea, groundnut, among others. Its rivers and plains also provide opportunities for fishing and irrigation agriculture. However, agricultural production system in the basin is mainly rainfed; which is highly vulnerable to extreme rainfall patterns and/or seasonal variability and exposes the farmers who are largely dependent on the system to production risks and uncertainties (Olayide *et al.,* 2016). This is because optimum yield of crops depends on required rainfall thresholds and any condition. At the same time, other livelihoods, particularly, fishing are affected by the by the rainfall characteristics. Hence, the role of rainfall on agricultural and other livelihood activities in the basin cannot be overemphasized. Understanding the pattern of rainfall and predicting its future occurrence is expedient for planning of various human activities. Unfortunately, there are no evidence of rainfall forecasts for the basin. It therefore becomes very imperative to study and predict future rainfall patterns in the basin. Thus, this study considered some forecasting procedures to in a bid to choose the one that offers better estimates for the future.

2. MATERIALS AND METHODS

2.1. Description of Study Area

The Bida Basin, otherwise referred to known as the Mid-Niger or Nupe Basin, is located in north-central Nigeria between latitude 8° 30Œ§N to 9° 30Œ§N and longitudes 5° 00Œ§E to 7° 00Œ§E. It covers an area of approximately 24,200 km^2 (Obaje *et al*., 2011; Okonkwo et al., 2018). It has a NW–SE trending intracratonic sedimentary basin extending from Kontagora in Niger State of Nigeria to areas slightly beyond Lokoja in Kogi state (Obaje *et al*., (2011). Bida Basin has a tropical climate with two distinct seasons, the wet and dry seasons. The wet season last for about 6-7 months. The rainy season commences in the month of April and ceases in the month of October while the dry season lasts between October and April. The area also experiences a short Harmattan season between the months of December and February. The annual rainfall of the area ranges between 1100 – 1400mm with about 60 percent of it falling between July and September. The daily maximum temperature varies between 27°C and 36°C. Maximum daily temperature is recorded between March and May while the minimum temperature of 23! to 25! is recorded between December and March and from July – September. The vegetation of the Basin is largely Southern Guinea Savanna type which is characterized by abundant trees and shrubs interspersed by grasses especially where there are minimal human interferences with the natural vegetation, particularly on hills which are not often cultivated, while forest-like riparian vegetation is common around river channels. The major rivers draining the basin are Rivers Niger and Kaduna.

2.2. Data

The data used for this study is the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 reanalysis precipitation data spanning 30 years from 1981 to 2020. Reanalyses utilise a wide variety of observation

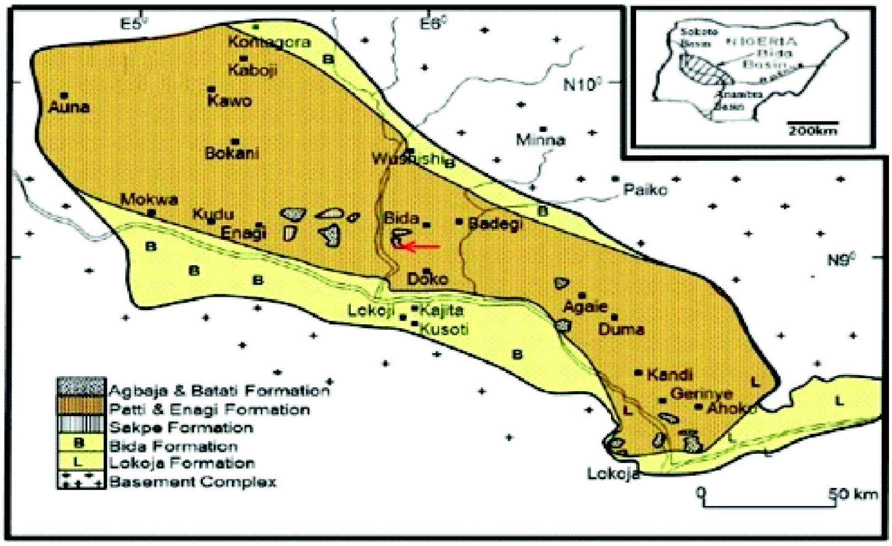

Figure 1: Nigeria showing Bida Basin

databases assimilated in a complex fashion into a numerical weather prediction model to produce a spatially and temporally coherent synthesis of various meteorological variables over the recent historical period (Essou *et al.*, 2016). It has a grid resolution of 31 km (0.218125°) and data are available on hourly basis and comprises of analysis and short forecasts, which run twice daily from 06 and 18 UTC.

2.3. Analytical Methods

2.3.1. Decomposition Model

In time series, we have four components namely, Trend, Seasonal, Cyclic and Random. These components are usually decomposed to see their effect on the time series data. Two decomposition methods for checking how these components affect the time series observations were used.

(a) **Additive Decomposition**: As the name implies, the component are taken as the sum of the decomposed patterns

$$X_t = T_t + S_t + \varepsilon_t$$

Where, X_t is the time series observation, T_t is the trend effect at time t, S_t is the seasonal effect at time t and ε_t is the random effect at time t.

(b) **Multiplicative Decomposition:** Here, the time series data were taken as a product of the decomposed pattern

$$X_t = T_t \times S_t \times \varepsilon_t$$

2.3.2. Simple Forecasting Methods

The paper used three simple forecasting method as benchmarks for other forecasting methods used in the paper. The methods are, average method, Naïve Method and seasonal naïve method.

(a) Average Method: With average method, the mean of the historical data is used as the forecast of all future values. Given that the historical data are $y_1, \ldots, y_T$, then the forecast can be written as

$$\hat{y}_{T+h|T} = \bar{y} = (y_1, \ldots, y_T)/\mathrm{T}.$$

Where $\hat{y}_{T+h|T}$ denote the estimate of $y_{T+h|T}$ based on the observed data $y_1, \ldots, y_T$.

(b) Naïve Method: here all forecast are set to be the value of last observed data. That is

$$\hat{y}_{T+h|T} = y_T$$

(c) Seasonal Naïve Method: This method is useful for highly seasonal series. With seasonal naïve method, the forecast value is set to be equal to the last observed value from the corresponding season of the year. It is usually written as

$$\hat{y}_{T+h|T} = y_{T+h-km},$$

Where m=the seasonal period, k=[(h-1)/m]+1.

2.3.3. Linear Regression Models

(a) Linear regression with trend component: To fit a linear trend model to capture the relationship between rainfall and time, we set the output variable y_t as rainfall observation at time t and the predictor as the time index t in the regression model:

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t$$

where y_t is the Rainfall at time, β_0 is the intercept, β_1 is the slope (trend), and ε_t is the noise term at time t.

(b) Linear regression with seasonal component: A seasonal pattern in a time series means that observations at certain seasons have consistently higher or lower observation than the other seasons. To ascertained the month with higher rainfall or lower rainfall, linear regression with season is employed. The most common way to capture seasonality in a regression model is by turning each of the month to a categorical variable that denotes the season for each observation. This categorical variable is then turned into absence presence (dummy) variables, which are included as predictors in the regression model. The model is given as:

$$y_t = \beta_0 + \sum_{i=2}^{12} \beta_i Season_{i,t} + \varepsilon_t \quad \boldsymbol{i} = 2,3,4,5,6,7,8,9,10,11,12$$

Usually, the first season is omitted for the singularity problem.

(c) Linear Regression with trend and seasonal component: The linear regression with trend and seasonal component is a model with both trend and seasonal dummies as below

$$y_t = \beta_0 + \beta_1 t + \sum_{i=2}^{12} \beta_i Season_{i,t} + \varepsilon_t$$

2.3.4. Exponential Smoothing Methods

(a) Simple Exponential Smoothing: This method uses the weighted average of all the past time series observation to perform a short-term forecast of the future value. Using the weighted average implies that as the observation get older the weight decreases exponentially. The method is implied that, weight is given to the current information and yet not ignoring the older information. We use simple exponential smoothing method when the series trend and seasonality have been removed. It is defined mathematically as:

$$\text{Level } L_t = \alpha y_t + (1-\alpha) L_{t-1} + \alpha(y_t - L_{t-1})$$

where

L_t = smoothed statistics, that is, the simple weighted average of current value y_t

L_{t-1} = previous smoothed statistic

α = smoothing constant of the data, note that $0 < \alpha < 1$

t = time period

and the forecast is

$$\hat{Y}_{t+1} = \hat{Y}_t + \alpha \varepsilon_t$$

where ε_t is the forecast error at time t

(b) Double Exponential Smoothing: sometime referred to as Holt's trend or second-order exponential smoothing. This method is used for a data with linear tend but no seasonal pattern. The double exponential smoothing is given by

$$\text{Level } L_t = \alpha y_t + (1-\alpha)(L_{t-1} - T_{t-1})$$

$$\text{Trend } T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$$

where

L_t = smoothed statistics, that is, the simple weighted average of current value y_t

L_{t-1} = previous smoothed statistic

α = smoothing constant of the data, note that $0 < \alpha < 1$

t = time period

T_{t-1} = best estimate of trend at time t

β = trend smoothing factor, note that $0 < \beta < 1$

The k-step ahead forecast is given by

$$\widehat{Y}_{t+k} = L_t + kT_t$$

(c) The Seasonal Holt-Winter's Model: Named after the inventors, the Holt-Winter's seasonal forecasting model for a given time series data, say $\widehat{Y}_t$ and with component in additive method is:

$$\text{Level}: L_t = \alpha(Y_t - S_{t-s}) + (1-\alpha)(L_{t-1} + T_{t-1})$$

$$\text{Trend}: \mathrm{T}_t = \beta'(\mathrm{L}_t - L_{t-s}) + (1-\beta')(\mathrm{T}_{t-1})$$

$$\text{Seasonal}: \mathrm{S}_t = \gamma*(\mathrm{Y}_t - L_t) + (1-\gamma*)\mathrm{S}_{t-s}$$

$$\text{Forecast}: \widehat{\mathrm{Y}}_{t+k} = L_t + kT_t + S_{t+k-s}$$

α is the level smoothing constant $0 \le \alpha \le 1$

β' is the trend smoothing constant $0 \le \beta' \le 1$

$\gamma*$is the seasonal smoothing constant $0 \le \gamma* \le 1$

L_t is the estimate of the level of the series at time t.

Y_t is the actual value of the series at time t

T_t is the estimate of the slope of the series at time t.

S_t is the seasonal component.

s is the length of seasonality and

k is the number of periods ahead to be forecast

2.3.5. Seasonal Autoregressive Integrated Moving Average (SARIMA) Model: This is the extension of the ARIMA(p,d,q) model by addition of four new parameters P,D,Q and s known as seasonal parameters. The model is denoted by:

$$SARIMA(p,d,q)(P,D,Q)_s,$$

where
P = the order of AR(p) process, d = the order of integration, q = the order of the MA(q) process and the seasonal parameters, P = the order of seasonal AR(P) process, D = the seasonal order of integration, Q = the order of seasonal MA(Q) process and s = the number of observations per cycle.

2.4. Model Selection

In choosing the exponential model or autoregressive integrated moving average (ARIMA) model that capture the features of the time series data, the following model selection criteria's: (a) the Akaike information criterion (AIC), (b) the Akaike information criterion corrected (AICc) and (c) The Bayesian information criterion (BIC) where used by est() and auto.arima() functions in r-package to select the models while (d) R-square was used to select from three linear regression models, model with trend only, model with seasonality only and model with both trend and seasonality.

(a) The Akaike information criterion (AIC) is defined as

$$AIC = 2L + 2k$$

Where L is the log likelihood function of the model and k is the total number of parameters, initial states that have been estimated and the residual variance.

(b) The Akaike information criterion corrected (AICc) is defined as

$$AICc = AIC + \frac{k(k+1)}{T-k-1}$$

(c) The Bayesian information criterion (BIC) is defined as

$$AICc = AIC + k[\log(T) - 2]$$

(d) **R-Squared** is defined as proportion of variation in dependent variables that is explained by changes in the independent variable. It is expressed mathematically as:

$$R - squared = \frac{Regression\ Sum\ of\ Squares}{Total\ Sum\ of\ Squares} = \frac{SSR}{TSS}$$

2.5. Forecast Accuracy Evaluation of the Models

The forecast error from each model was check in this article using the following forecast error measures, (a) Root Mean Squared Error (RMSE) (b) Mean Absolute Error (MAE) and (c) Mean Absolute Scaled Error (MASE).

(a) **The Root Mean Squared Error (RMSE)** is the root of the average squared difference between prediction and actual data. Ti measures the magnitude of the error. It is defined mathematically as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} \varepsilon_t^2}$$

(b) **Mean Absolute Error (MAE)** is a measure of the average magnitude of errors in a set of predictions, without considering their direction.

$$MAE = \frac{1}{T}\sum_{t=1}^{T} |\varepsilon_t|$$

(c) **Mean Absolute Scaled Error (MASE)** for a seasonal time series id defined as

$$MASE = \frac{1}{T}\sum_{t=1}^{T} \frac{|y_t - \hat{y}_{t|t-1}|}{\frac{1}{T-m}\sum_{t=m+1}^{t} |y_t - y_{t-m}|}$$

It is the ratio of MAE of the model to that of seasonal naïve method. Here, value less than 1 indicate that the model has lower average error than naïve forecasts for the training period and poor forecasting otherwise.

2.6. Data Analysis

R-software (version 4.2.2) was used for the analysis of the data with the help of "zoo", "fpp2", "forecast", "TTR", "seasonal", "xts", "TSstudio", "dplyr", "lubridate", "trend", "plotly", "stats", and "tseries".

3. RESULTS AND DISCUSSIONS

The monthly rainfall forecasting of Bida basin was achieved using various time series forecasting techniques. The Bida monthly rainfall data for the period of 1981 to 2020 was used. The data was split into the train series and the test series. The train series was the Bida monthly rainfall data from January1981 to December 2017 and was used to develop the models while the test series represent Bida monthly rainfall data from January, 2018 to December 2020 was used for forecasting. The prediction was made using various time series forecasting methods.

Table 1: Descriptive Statistics

Month	*Mean*	*Standard Deviation*
January	2.32	3.33
February	6.62	8.75
March	24.8	20.4
April	56.7	33.2
May	139.	41.3
June	205.	52.6
July	262.	47.1
August	318.	69.9
September	288.	63.9
October	131.	48.3
November	10.3	10.5
December	2.01	4.12

The descriptive statistics table 1 above and the figure 1, the barplot represents Bida average monthly rainfall for the months of the years from 1981 to 2020. As was observed, on the average, each month is different from the other by its standard deviation however, the pairs months of July/ October, August/September are relatively close to each other.

Figure 3, the polar plot representation of the series also confirmed a repeated seasonal patterns along with a year to year trend. The polar plot revealed a distinct pattern where the month of August has the highest rainfall, followed by September, June, October, July and May in that order. In other words, yearly rainfall in Bida is usually between May to October. Figure 4 revealed normal plot, cycle plot and the box plot respectively. Figure 4 further confirmed that in the month of August rainfall is at the peak,

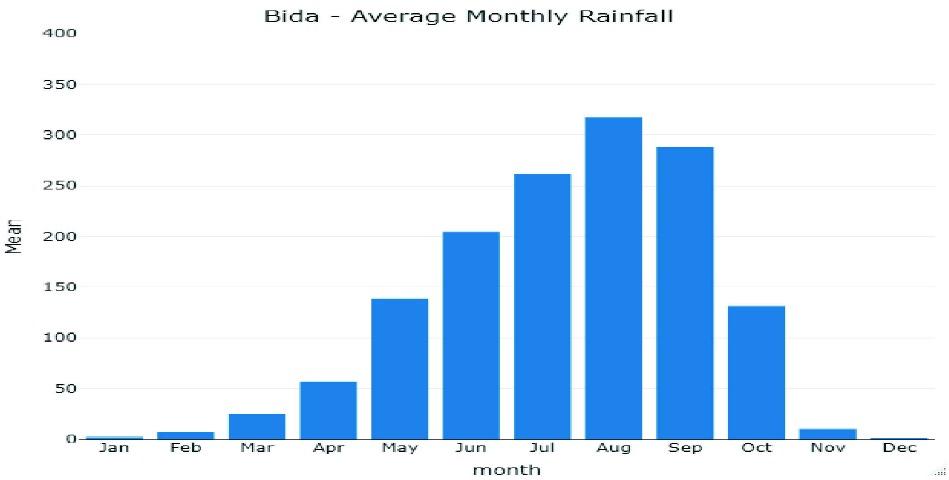

Figure 2: Bida Average Monthly Rainfall

followed by September, June, October and July in that order. Although the onset of rain in Bida basin is usually April through to October based on the statistics.

Time-plot and Model Identification

The series was decomposed using the multiplicative model as suggested by both Ljung-Box test and Box-Pierce test. One of the assumptions that generally concerned all the forecasting approaches is that the trend, cyclic and seasonal components are stable, and that the past patterns will continue. From the decomposition of multiplicative time series in figure 2 below, it revealed the presence of a trend component in the series and exhibits a slightly downward trend from about 1993, although not obvious. The seasonal component was very strong and persistent with ups and downs cycles at certain month(s) each year.

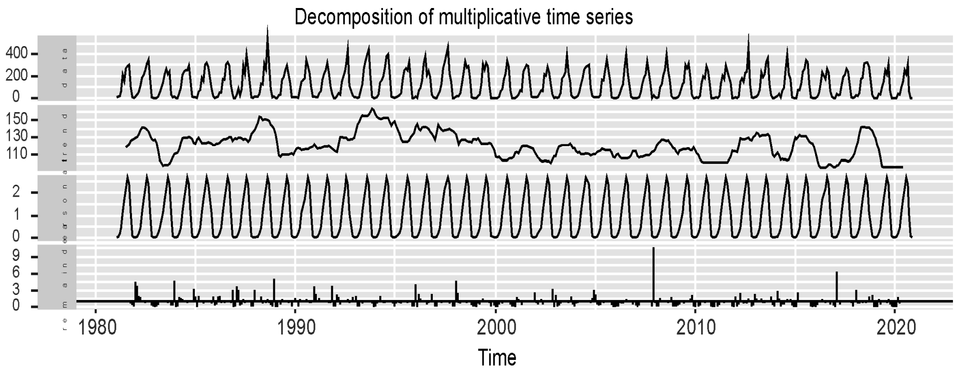

Figure 3: Decomposition of Multiplicative Time Series of Monthly Rainfall in Bida from 1981 to 2020

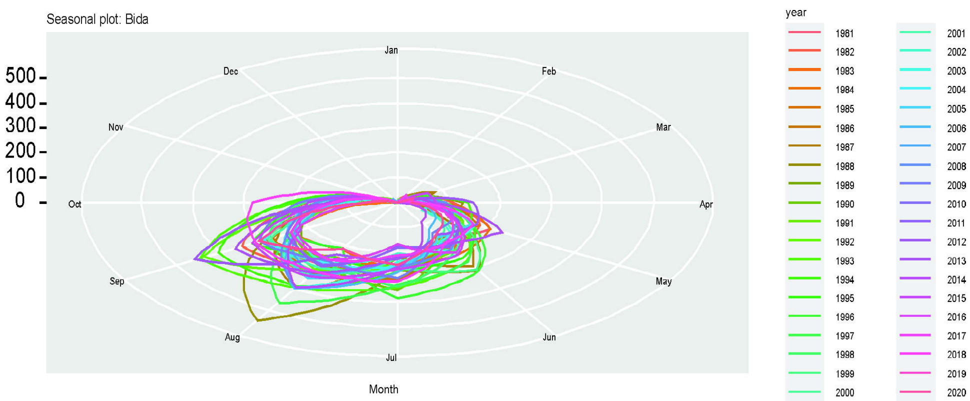

Figure 4: Polar Seasonal Plot for Bida Monthly Rainfall Data from 1981 to 2020

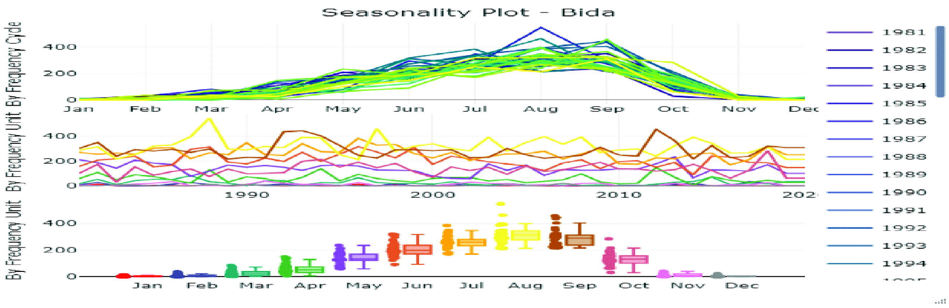

Figure 5: Seasonality Plot for Bida Monthly Rainfall Data from 1981 to 2020

Table 2: Models Selected and Portmanteau Test

Model Category	*Selected Model*	*Ljung-Box Test*	*Degree of freedom*	*P-Value*
Linear	$y_t = \beta_0 + \beta_1 t + \sum_{i=2}^{12} \beta_i Season_{i,t} + \varepsilon_t$	1.1461	1	0.2844
Exponential	EST(A, N, A)	0.6251	1	0.4292
ARIMA	SIAR(0,0,0)(0,1,1)$_{12}$	1.7350	1	0.1878

The figures above, that is, figure 5-10 are the residuals plots from the models, the 5-7 there are significant seasonal sparks from lag 12 to 36 and from the Ljung test was statistical significant which implied that mean, naïve and seasonal naïve method do not capture the pattern in the data. The figures 8, 9 and 10 also revealed a spark at lag 12 with was statistical significant was the Ljung-Box test and where the models further examining.

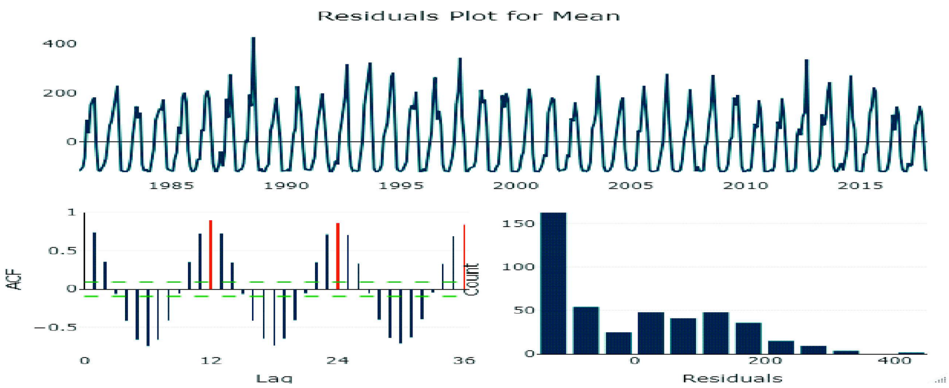

Figure 6: Residual Plot for Mean or Average Forecasting Method

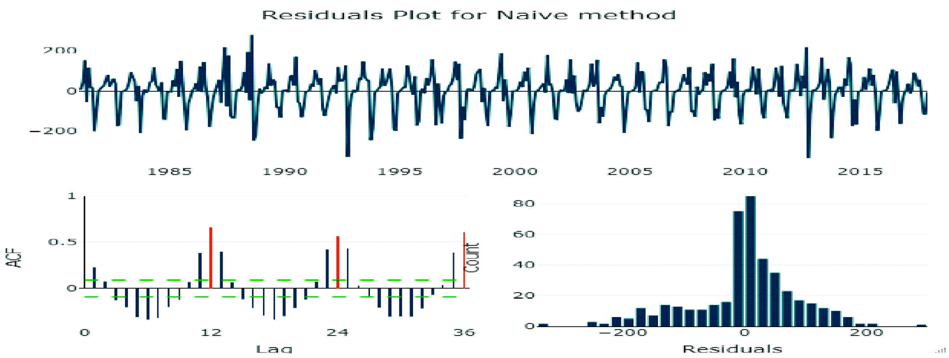

Figure 7: Residual Plot for Naive Forecasting Method

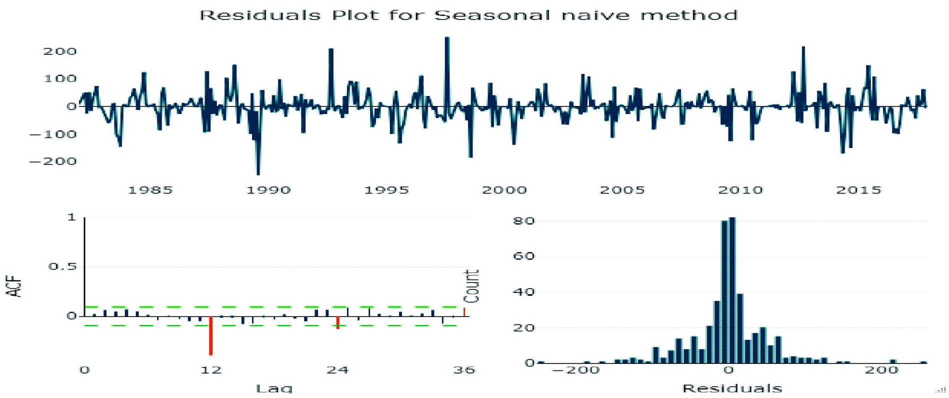

Figure 8: Residual Plot for Seasonal Naive Forecasting Method

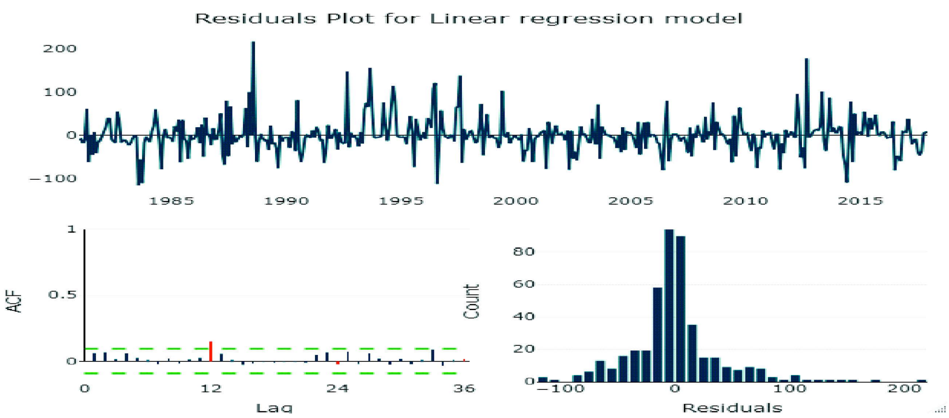

Figure 9: Residual Plot for Linear Regression Forecasting Method

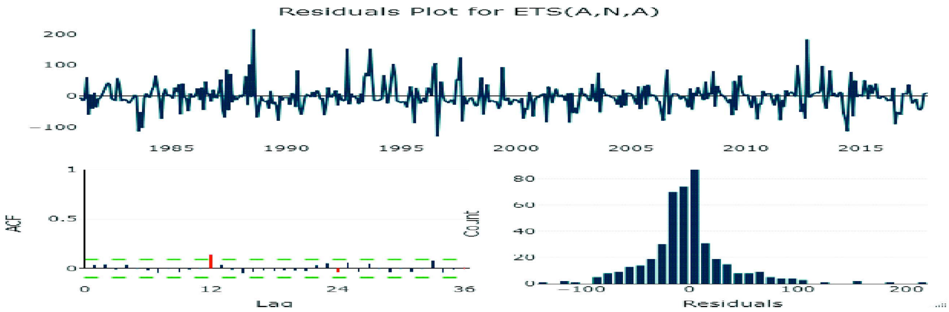

Figure 10: Residual Plot for Seasonal Holt-Winter Forecasting Method

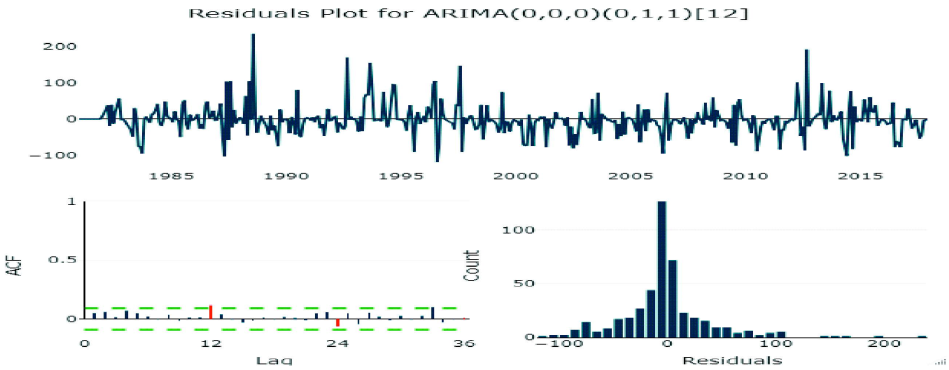

Figure 11: Residual Plot for Seasonal ARIMA Forecasting Method

Table 2 above reported the models selected from linear regression models, Exponential smoothing methods and seasonal ARIMA models. The selection of the linear regression model where based on their adjusted R-squares. Three models where considered, model with trend only, model with seasonal dummies only and model with both trend and seasonal dummies. The model with both trend and seasonal dummies had the highest adjusted R-squared and was selected. The exponential smoothing and seasonal arima models where automatically using est() and auto.arima functions in r statistical package respectively. The three categories of the models considered passed the portmanteau test and are all good for estimation of rainfall parameters.

Table 3 is the predictive accuracy evaluation of the forecasting methods used in this paper. The mean, naïve and seasonal naïve methods are used as benchmarks are the selected three models. Three scaled evaluation techniques where employed, the root mean square error (RMSE), the mean absolute error (MAE) and the mean absolute scaled error (MASE). The linear regression model with trend and seasonality show least root mean squared error, this implies that, prediction from the linear regression model is the closest to the actual values than all other models considered. With the MAE and MASE SARIMA(0, 0, 0)(0,1,1)$_{12}$ is the best, since it has the least value compare to other models considered.

Furthermore, comparing models from linear regression model, exponential smoothing and SARIMA model with the benchmark forecasting methods, it was revealed by MASE that the values from linear regression model (0.9167), exponential model (0.9262) and SARIMA model (0.9062) are less than 1, which implies that all the three models compared with the seasonal naïve performed better than the seasonal naïve method. However, we adopted the model with least root mean squared error for prediction since it is the model with closest prediction values to the actual value among the models considered.

The model results presented in Table 4 are based on the European Centre for Medium-Range Weather Forecasts (ECMWF) ERAS reanalysis precipitation data. It was showed that the monthly average precipitation is at least 12.5234mm. However, the result revealed a negative trend in rainfall, this implies a decrease in rainfall with about 0.0464mm yearly and was statistically significant at less than 1% significance level.

Table 3: Predictive Accuracy Evaluation

Forecast Method	*RMSE*	*MAE*	*MASE*
Mean	117.4384	107.0264	3.2795
Naïve	161.0406	110.8286	3.3961
Seasonal Naïve	48.1396	33.2272	1.0182
Linear Regression	45.1776	29.9149	0.9167
Exponential	45.3875	30.2258	0.9262
SARIMA	45.3145	29.5724	0.9062

Table 4: Result of Linear Regression Models

coefficient	*Estimate*	*Std. Error*	*t value*	*Pr(>|t|)*
(Intercept)	12.5234	7.2597	1.725	0.0852 .
trend	-0.0464	0.0147	-3.158	0.0017 **
season2	4.2115	9.2234	0.457	0.6482
season3	22.2271	9.2235	2.410	0.0164 *
season4	56.2452	9.2235	6.098	2.39e-09 ***
season5	137.6794	9.2235	14.927	< 2e-16 ***
season6	205.9020	9.2237	22.323	< 2e-16 ***
season7	258.6203	9.2238	28.038	< 2e-16 ***
season8	321.0989	9.2240	34.811	< 2e-16 ***
season9	284.1761	9.2242	30.808	< 2e-16 ***
season10	129.0725	9.2244	13.993	< 2e-16 ***
season11	8.3838	9.2246	0.909	0.3639
season12	0.2291	9.2248	0.025	0.9802

Furthermore, the result revealed a no significant rainfall in the month of February, November and December as rainfall in these months are not statistically significant. In contrast, the significant onset of rainfall in Bida basin is March and retreat by November, with the month of August with the highest rainfall followed by September, July, June, May, October, April and March respectively.

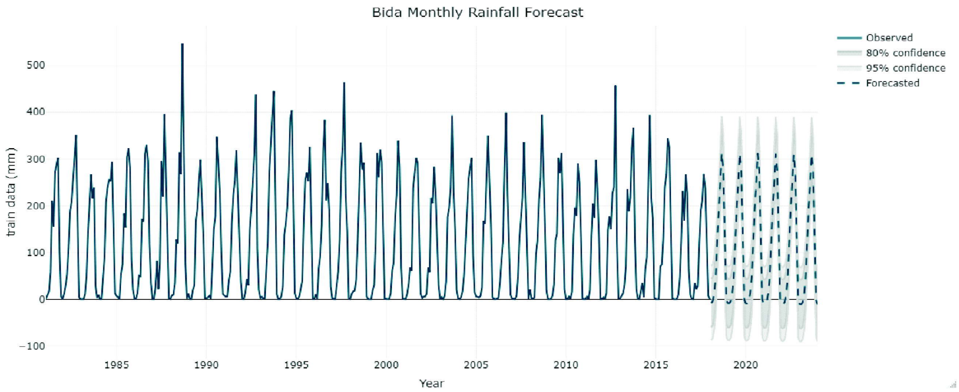

Figure 12: Actual and Forecasted Monthly Rainfall from Linear Regression Model

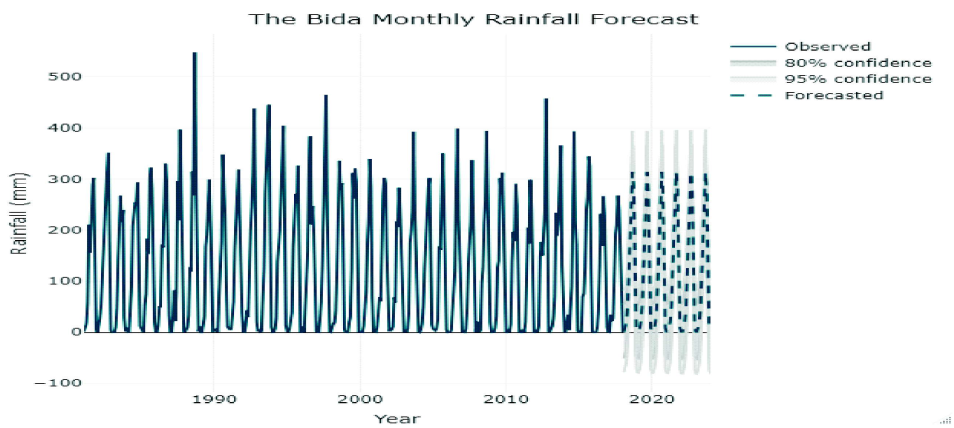

Figure 13: Actual and Forecasted Monthly Rainfall from SARIMA (0,0,0) (0,1,1)$_{12}$ Model

Figure 12 and 13 are the observed and 80% and 95% confidence interval for the forecasted values from 2018 to 2023 and revealed that the forecasted values are within 80% and 95% confidence interval. The forecasted values followed pattern revealed by the model. The Forecasted values are showed in Appendix I.

4. CONCLUSION

This study explored some time series forecast methods, the mean method, naïve, seasonal naïve, simple, double and triple exponential smoothing methods, the linear regression methods and seasonal autoregressive integrated moving average (ARIMA) method. The performance of the models were appraised by using monthly rainfall data for Bida from the year 1981 to 2020. The forecast accuracy measures were based on the forecast errors and residual plots.

The empirical findings revealed that linear regression method with both trend and seasonality performed better in terms of closest of predicted values to the actual values and was used for forecast. The negative and statistically significant trend coefficient corroborate the finding of Oti et al. (2020) that the global impact of climate change is bringing about alterations in rainfall patterns with Africa experiencing the worst effect. Kkpoh (2007) also revealed a decreasing mean annual rainfall for Kano, Katsina and Zaria. The monthly coefficient indicated a significant onset of rainfall in Bida basin is March and lasted till October yearly. The month of August was estimated to have 321.0989mm rainfall per year and the month with the highest rainfall yearly, followed by September, July, June, May, October, April and March in that order. The forecast from both ARIMA $(0,0,0)(0,1,1)_{12}$ and linear regression models considered revealed same pattern of rainfall for 2023. The study will be helpful to our teaming population of farmers in planning their activities and to other who required water for other activities to plan their water harvesting strategies.

Acknowledgements

The authors wish to acknowledge the European Centre for Medium-Range Weather Forecasts (ECMWF) for the free access to the ERA5 Interim (European Reanalysis) grid-based daily rainfall data that was used for this work. No funding agency was involved in the study design, collection, analysis and interpretation of data in the writing of this manuscript.

Competing Interests

Authors have declared that no competing interests exist.

References

Abe, A.; Adeniji, Q. A.; Rabiu, J. A; Adegboyega, O.; Raheem, I. O.; Rasaki, M.G.; Sada, S. M.; & Fidelis, L. (2022). Statistical Analysis and Forecasting of Rainfall Patterns and Trends in Gombe North-Eastern Nigeria. *Iraqi Journal of Physics*, *20*(2), 64-77.

Ekpoh, I. J. (2007). Climate and Society in Northen Nigeria: Rainfall Variability and Farming. *The International Journal Series on Tropical Issues*, *8*(3), 157-162.

Elouissi, A.; Sen, Z.; & Habi, M. (2016). Algerian Rainfall Innovative Trend Analysis and its Implications to Macta Watershed. *Arabian Journal of Geosciences*, 9(303), 1-12.

Emaziye, P. O (2015). The Influences of Temperature and Rainfall on the Yields of Maize, Yam and Cassava among Rural Households in Delta State, Nigeria. *Journal of Biology, Agriculture and Healthcare* 5 (1), 63-69.

Essou, G.R., F. Sabarly, P. Lucas-Picher, F. Brissette, and Poulin A. (2016). Can Precipitation and Temperature from Meteorological Reanalyses Be Used for Hydrological Modelling? *Journal of Hydrometeorology*, 17, 1929–1950.

Obaje N. G, Moumouni A., Goki N. G & Chaanda M. S. (2011). Stratigraphy, Paleogeography and Hydrocarbon Resource Potentials of the Bida Basin in North-Central Nigeria. *Journal of Mining and Geology*, 47(2) 97–114.

Okwokwo O. I, Adetona A. A., Adewumi , T., Sunday, O., Adediran S. (2018). Interpretation of High-resolution Aeromagnetic Data to Determine Sedimentary Thickness over Part of Bida Basin, North Central Nigeria. *Journal of Geology and Mining Research* 10(6), 72-80.

Olayide O. E., Tetteh I. K., Popoola L. (2016). Differential Impacts of Rainfall and Irrigation on Agricultural Production in Nigeria: Any Lessons for Climate-Smart Agriculture? *Agricultural Water Management*, 30-36.

Oti J. O., Kabo-bah A. T., & Ofosu E. (2020). Hydrologic Response to Climate Change in the Densu River Basin in Ghana. *Heliyon* 6 (2020) e04722

Yila K. M., Gboku M. L. S, Lebbie M. S & Kamara L.I (2023). Changes in Rainfall and Temperature and Its Impact on Crop Production in Moyamba District, Southern Sierra *Atmospheric and Climate Sciences*, 13, 19-43.

Appendix I: Forecast from Linear Regression Model for the period 2018 to 2023

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2018	-8.1274249	-59.908750	43.65390	-87.42141	71.16656
Feb 2018	-3.9622898	-55.743615	47.81904	-83.25628	75.33170
Mar 2018	14.0068994	-37.774426	65.78822	-65.28709	93.30089
Apr 2018	47.9785210	-3.802804	99.75985	-31.31547	127.27251
May 2018	129.3663589	77.585033	181.14768	50.07237	208.66035
Jun 2018	197.5425751	145.761250	249.32390	118.24859	276.83656
Jul 2018	250.2144670	198.433142	301.99579	170.92048	329.50846
Aug 2018	312.6466291	260.865304	364.42795	233.35264	391.94062
Sep 2018	275.6774399	223.896114	327.45877	196.38345	354.97143
Oct 2018	120.5274399	68.746114	172.30877	41.23345	199.82143
Nov 2018	-0.2076952	-51.989021	51.57363	-79.50168	79.08629
Dec 2018	-8.4087763	-60.190102	43.37255	-87.70277	70.88521
Jan 2019	-8.6843022	-60.484914	43.11631	-88.00782	70.63922
Feb 2019	-4.5191671	-56.319779	47.28144	-83.84269	74.80436
Mar 2019	13.4500221	-38.350590	65.25063	-65.87350	92.77354
Apr 2019	47.4216437	-4.378968	99.22226	-31.90188	126.74517
May 2019	128.8094816	77.008870	180.61009	49.48596	208.13300
Jun 2019	196.9856978	145.185086	248.78631	117.66218	276.30922
Jul 2019	249.6575897	197.856978	301.45820	170.33407	328.98111
Aug 2019	312.0897519	260.289140	363.89036	232.76623	391.41327
Sep 2019	275.1205627	223.319951	326.92117	195.79704	354.44409
Oct 2019	119.9705627	68.169951	171.77117	40.64704	199.29409
Nov 2019	-0.7645725	-52.565184	51.03604	-80.08810	78.55895
Dec 2019	-8.9656535	-60.766265	42.83496	-88.28918	70.35787
Jan 2020	-9.2411795	-61.062059	42.57970	-88.59574	70.11338
Feb 2020	-5.0760443	-56.896924	46.74484	-84.43060	74.27851
Mar 2020	12.8931449	-38.927735	64.71402	-66.46141	92.24770
Apr 2020	46.8647665	-4.956113	98.68565	-32.48979	126.21933
May 2020	128.2526043	76.431725	180.07348	48.89805	207.60716
Jun 2020	196.4288205	144.607941	248.24970	117.07426	275.78338
Jul 2020	249.1007124	197.279833	300.92159	169.74615	328.45527
Aug 2020	311.5328746	259.711995	363.35375	232.17832	390.88743
Sep 2020	274.5636854	222.742806	326.38457	195.20913	353.91824
Oct 2020	119.4136854	67.592806	171.23457	40.05913	198.76824
Nov 2020	-1.3214497	-53.142329	50.49943	-80.67601	78.03311
Dec 2020	-9.5225308	-61.343410	42.29835	-88.87709	69.83203
Jan 2021	-9.7980567	-61.640184	42.04407	-89.18515	69.58904
Feb 2021	-5.6329216	-57.475049	46.20921	-85.02002	73.75417
Mar 2021	12.3362676	-39.505860	64.17840	-67.05083	91.72336
Apr 2021	46.3078892	-5.534238	98.15002	-33.07921	125.69499
May 2021	127.6957270	75.853600	179.53785	48.30863	207.08282
Jun 2021	195.8719433	144.029816	247.71407	116.48485	275.25904
Jul 2021	248.5438352	196.701708	300.38596	169.15674	327.93093
Aug 2021	310.9759973	259.133870	362.81812	231.58890	390.36309
Sep 2021	274.0068081	222.164681	325.84894	194.61971	353.39390
Oct 2021	118.8568081	67.014681	170.69894	39.46971	198.24390
Nov 2021	-1.8783270	-53.720455	49.96380	-81.26542	77.50877
Dec 2021	-10.0794081	-61.921536	41.76272	-89.46650	69.30769
Jan 2022	-10.3549340	-62.219288	41.50942	-89.77607	69.06620
Feb 2022	-6.1897989	-58.054153	45.67456	-85.61093	73.23133
Mar 2022	11.7793903	-40.084964	63.64374	-67.64174	91.20052
Apr 2022	45.7510119	-6.113342	97.61537	-33.67012	125.17214
May 2022	127.1388498	75.274495	179.00320	47.71772	206.55998
Jun 2022	195.3150660	143.450712	247.17942	115.89393	274.73620

```
Jul 2022    247.9869579 196.122604 299.85131 168.56582 327.40809
Aug 2022    310.4191200 258.554766 362.28347 230.99799 389.84025
Sep 2022    273.4499309 221.585577 325.31429 194.02880 352.87106
Oct 2022    118.2999309  66.435577 170.16429  38.87880 197.72106
Nov 2022     -2.4352043 -54.299559  49.42915 -81.85634  76.98593
Dec 2022    -10.6362854 -62.500640  41.22807 -90.05742  68.78485
Jan 2023    -10.9118113 -62.799370  40.97575 -90.36848  68.54486
Feb 2023     -6.7466761 -58.634235  45.14088 -86.20334  72.70999
Mar 2023     11.2225130 -40.665046  63.11007 -68.23415  90.67918
Apr 2023     45.1941347  -6.693424  97.08169 -34.26253 124.65080
May 2023    126.5819725  74.694414 178.46953  47.12531 206.03864
Jun 2023    194.7581887 142.870630 246.64575 115.30152 274.21486
Jul 2023    247.4300806 195.542522 299.31764 167.97341 326.88675
Aug 2023    309.8622428 257.974684 361.74980 230.40558 389.31891
Sep 2023    272.8930536 221.005495 324.78061 193.43639 352.34972
Oct 2023    117.7430536  65.855495 169.63061  38.28639 197.19972
Nov 2023     -2.9920816 -54.879640  48.89548 -82.44875  76.46459
Dec 2023    -11.1931626 -63.080722  40.69440 -90.64983  68.26350
                        ME     RMSE      MAE MPE MAPE      MASE
ACF1
Training set -7.291601e-16 39.08628 25.31126 NaN  Inf 0.7755994
0.06230028
Test set      9.998653e-01 45.17768 29.91492 Inf  Inf 0.9166671
0.04733514
             Theil's U
Training set        NA
Test set             0
```

Appendix II: Forecast from ARIMA(0,0,0)(0,1,1)12 for the period 2018 to 2023

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2018	1.479427	-50.973688	53.93254	-78.74069	81.69954
Feb 2018	8.222846	-44.230269	60.67596	-71.99727	88.44296
Mar 2018	21.114119	-31.338996	73.56723	-59.10600	101.33424
Apr 2018	51.136507	-1.316608	103.58962	-29.08361	131.35662
May 2018	132.658846	80.205732	185.11196	52.43873	212.87896
Jun 2018	191.408915	138.955800	243.86203	111.18880	271.62903
Jul 2018	233.824683	181.371568	286.27780	153.60457	314.04480
Aug 2018	313.654563	261.201449	366.10768	233.43445	393.87468
Sep 2018	278.952447	226.499332	331.40556	198.73233	359.17256
Oct 2018	129.491562	77.038447	181.94468	49.27145	209.71168
Nov 2018	8.321035	-44.132080	60.77415	-71.89908	88.54115
Dec 2018	1.319639	-51.133476	53.77275	-78.90048	81.53976
Jan 2019	1.479427	-51.338764	54.29762	-79.29903	82.25788
Feb 2019	8.222846	-44.595345	61.04104	-72.55561	89.00130
Mar 2019	21.114119	-31.704072	73.93231	-59.66433	101.89257
Apr 2019	51.136507	-1.681684	103.95470	-29.64195	131.91496
May 2019	132.658846	79.840655	185.47704	51.88039	213.43730
Jun 2019	191.408915	138.590724	244.22711	110.63046	272.18737
Jul 2019	233.824683	181.006492	286.64287	153.04623	314.60314
Aug 2019	313.654563	260.836372	366.47275	232.87611	394.43302
Sep 2019	278.952447	226.134256	331.77064	198.17399	359.73090
Oct 2019	129.491562	76.673371	182.30975	48.71311	210.27001
Nov 2019	8.321035	-44.497156	61.13923	-72.45742	89.09949
Dec 2019	1.319639	-51.498552	54.13783	-79.45881	82.09809
Jan 2020	1.479427	-51.701334	54.66019	-79.85353	82.81238
Feb 2020	8.222846	-44.957915	61.40361	-73.11011	89.55580
Mar 2020	21.114119	-32.066642	74.29488	-60.21884	102.44707
Apr 2020	51.136507	-2.044254	104.31727	-30.19645	132.46946
May 2020	132.658846	79.478085	185.83961	51.32589	213.99180
Jun 2020	191.408915	138.228154	244.58968	110.07596	272.74187
Jul 2020	233.824683	180.643922	287.00544	152.49173	315.15764
Aug 2020	313.654563	260.473802	366.83532	232.32161	394.98752
Sep 2020	278.952447	225.771686	332.13321	197.61949	360.28540
Oct 2020	129.491562	76.310801	182.67232	48.15861	210.82452
Nov 2020	8.321035	-44.859726	61.50180	-73.01192	89.65399
Dec 2020	1.319639	-51.861122	54.50040	-80.01332	82.65259
Jan 2021	1.479427	-52.061449	55.02030	-80.40428	83.36313
Feb 2021	8.222846	-45.318030	61.76372	-73.66086	90.10655
Mar 2021	21.114119	-32.426757	74.65499	-60.76958	102.99782
Apr 2021	51.136507	-2.404369	104.67738	-30.74720	133.02021
May 2021	132.658846	79.117970	186.19972	50.77514	214.54255
Jun 2021	191.408915	137.868039	244.94979	109.52521	273.29262
Jul 2021	233.824683	180.283807	287.36556	151.94098	315.70839
Aug 2021	313.654563	260.113687	367.19544	231.77086	395.53827
Sep 2021	278.952447	225.411571	332.49332	197.06874	360.83615
Oct 2021	129.491562	75.950686	183.03244	47.60786	211.37527
Nov 2021	8.321035	-45.219841	61.86191	-73.56267	90.20474
Dec 2021	1.319639	-52.221237	54.86052	-80.56406	83.20334
Jan 2022	1.479427	-52.419158	55.37801	-80.95135	83.91020
Feb 2022	8.222846	-45.675739	62.12143	-74.20793	90.65362
Mar 2022	21.114119	-32.784466	75.01270	-61.31665	103.54489
Apr 2022	51.136507	-2.762078	105.03509	-31.29427	133.56728
May 2022	132.658846	78.760262	186.55743	50.22807	215.08962
Jun 2022	191.408915	137.510330	245.30750	108.97814	273.83969
Jul 2022	233.824683	179.926098	287.72327	151.39391	316.25546

```
Aug 2022     313.654563 259.755979 367.55315 231.22379 396.08534
Sep 2022     278.952447 225.053862 332.85103 196.52167 361.38322
Oct 2022     129.491562  75.592977 183.39015  47.06079 211.92233
Nov 2022       8.321035 -45.577550  62.21962 -74.10974  90.75181
Dec 2022       1.319639 -52.578946  55.21822 -81.11113  83.75041
Jan 2023       1.479427 -52.774508  55.73336 -81.49481  84.45366
Feb 2023       8.222846 -46.031089  62.47678 -74.75139  91.19708
Mar 2023      21.114119 -33.139816  75.36805 -61.86012 104.08835
Apr 2023      51.136507  -3.117428 105.39044 -31.83773 134.11074
May 2023     132.658846  78.404911 186.91278  49.68461 215.63308
Jun 2023     191.408915 137.154980 245.66285 108.43468 274.38315
Jul 2023     233.824683 179.570748 288.07862 150.85045 316.79892
Aug 2023     313.654563 259.400628 367.90850 230.68033 396.62880
Sep 2023     278.952447 224.698511 333.20638 195.97821 361.92668
Oct 2023     129.491562  75.237627 183.74550  46.51733 212.46580
Nov 2023       8.321035 -45.932901  62.57497 -74.65320  91.29527
Dec 2023       1.319639 -52.934296  55.57357 -81.65459  84.29387
                    ME     RMSE      MAE  MPE MAPE      MASE
ACF1
Training set -2.088870 40.32533 25.61099 -Inf  Inf 0.7847840
0.05063604
Test set     -3.251216 45.31454 29.57244 -Inf  Inf 0.9061727 -
0.02834885
             Theil's U
Training set        NA
Test set             0
```