ACADEMY JOURNAL OF SCIENCE AND ENGINEERING
Available online @
www.academyjsekad.edu.ng
AJSE 17 (2) 2023
Original Research

FORECASTING STREAM FLOW OF KADUNA RIVERAT WUYA GAUGING STATION IN NIGER STATE

Garba H.[1*] and Hamman Y.U.[2]

Abstract

The Box Jenkins methodology for forecasting was used to identify suitable models and forecast future stream flow values for average and maximum discharge of Kaduna River at Wuya gauging station in Niger state and to predict extreme events of flooding. The average and maximum monthly stream flow data used for the analysis were obtained from the Nigerian Hydrological Serves Agency from January 1988 to September 2021 (33 years). Autoregressive Integrated Moving Average ARIMA (0,1,4)(0,1,4) and Autoregressive Integrated Moving Average ARIMA (0,1,3)(0,1,3) models were the most suitable models.. Both models achieved the lowest normalized Bayesian Information Criterion (BIC) which were the criteria for chosen the best model after prognosis. The selected models also have the best Ljung-Box Q statistical significance P Value of (0.11,0.175,0.216,0.539) for ARIMA (0,1,4)(0,1,4) and (0.010,0.1053,0.123,0.247) for (0,1,3)(0,1,3). Diagnostic checks of the models revealed that the models selected have a residual that was white noise. The models used for forecasting had the highest Coefficient of determination (R^2) of 0.795 and 0.813 with the least Mean absolute percentage error (MAPE) of 28.785%, 34.461% respectively for ARIMA (0,1,4) (0,1,4) and ARIMA (0,1,3) (0,1,3). Both model forecast was reasonable having their MAPE between 21% to 50%. The initial analysis indicated that the average stream flow had a mean value of 518.9 m^3/s, standard deviation of 414.5, coefficient of Skewness of 0.81 and coefficient of kurtosis of -0.42 while the maximum stream flow of Kaduna River had mean value of 813.7 m3/s, standard deviation of 687.6, coefficient of Skewness of 0.94 and coefficient of kurtosis of -0.2. The Box Jenkings model's future simulations of the average and maximum discharges through the year 2026 indicates long-term periodicity.

Keywords: Forecasting, Stream Flow, ARIMA model

Corresponding author's e-mail:

OPEN ACCESS
This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

1.0 INTRODUCTION

Hydrology is a branch of science that examines the distribution, occurrence, circulation, and properties or characteristics of water in the environment. Several activies regarding water resources management are dependent on accurate monthly stream flow forecsting such as flood control, reservoirs operation, water supply planning, hydropower generation (Belotti *et al* 2021) Water exists in three states: liquid, solid, and vapor. It travels through the system along many different routes through the atmosphere, the surface of the land, and the subsurface. Water is also temporarily stored in a variety of places, including the soil, wetlands, lakes, flood plains, aquifers, oceans, and the atmosphere. Estimating the amount and quality of water in the various stages and also comprehending the underlying physical and stochastic processes is what hydrology is all about (Wang and Yang, 2014). The science of hydrology also investigates the occurrence, distribution, flow, and properties of the earth's waters as well as how these waters interact with their environment at each stage of the hydrologic cycle. In carrying our statistical modelling of monthly stream flow using time series and artificial neuron network models, (Nabeel *et al 2021) u*sed monthly stream flow for the period January 2000 tto December 2019 by utilizing ARIMA and the non linear authoregressive N,A.R time series models. The predicted Box Jenkins Models was ARIMA (1,1,0 and 0,1,1) while the predicted artificial neural networks for (N.LR)model was (M.L.P.1-3-1). The reults of the study indicates that the traditional Box Jenkins model was more accurate than the N.L.R model. Performing a one step ahead of forecast during the year 2019, the accuracy beetwen the forecasted and recoreded monthly stream flow of both models are as follows Box Jenkins gave root mean squard (RMSE =48.7) and (R^2 = 0.801) while the NAR model gave (RMES 93.40 and R^2= 0.269).

Stream flows are surplus runoff or storm water runoff from precipitation that runs across land but does not percolate or permeate into the subsurface because of either high intensity or because the soil is saturated. Intensity of rainfall events, topography and geology, basin area and slope, length of river or stream, land use (agricultural and urban development), and climate all affect stream flows (Pegram, 2015). Flooding of the stream or river banks happens when there is too much rainfall and the stream flow exceeds the capacity of the stream or river.Stochastic hydrology, according to Pegram (2015), is the statistical subfield of hydrology that deals with the probabilistic modeling of hydrological systems that include random elements.

According to Adnan et al. (2017), accurate streamflow forecasting is crucial for managing water resources. , Adnan et al. (2017) predicted stream flow with two time series models: the Autoregressive Moving Average (ARMA) model and the Autoregressive Integrated Moving Average (ARIMA), using streamflow data from 1974 to 2010. The models were trained using data for the first 28 years, and forecasting was done using data from the most recent 7 years. By contrasting the root mean square error (RMSE), the mean absolute percentage error (MAPE), and the Nash efficiency (NE), the prediction accuracy of both time series models is evaluated. The outcomes showed that the ARIMA model outperforms the ARMA time series models.

One of the most important aspects of several water resources projects in both Sudan and South Sudan is the forecasting of the monthly streamflow for the White Nile River at the Malakal station(Mohamed 2021), the

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

observed flow from this station completely determines how the Jabal al Awliya dam in central Sudan operates. at Malakal station, Stream flow was modeled and projected using seasonal autoregressive integrated moving average (SARIMA) models, which are linear stochastic models. Monthly flow data covering the period from 1970 to 2013 were utilized for the analysis. A close examination of the original series reveals a regular annual pattern.

To find the best strategy for forecasting monthly streamflow time series, Wu and Chau (2010) study various data-driven models. For the investigation process, four sets of data from various areas in China were used. A comparison is made between four models, ARMA (Auto-Regressive Moving Average), ANN (Artificial Neural Networks), KNN (K-Nearest Neighbor), and ANN-PSR (Phase Space Reconstruction-based Artificial Neural Networks). They study concluded that that the KNN model outperforms the other three models, but it only shows marginal superiority to ARMA.

Martins et al., (2011) conducted a research on the monthly stream flow of the River Benue for the 26 years between 1974 and 2000 to analyze and forecast the stream flow of the River Benue, using Autoregressive Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA).Findings from their study reveals that the ARMA model outperformed the ARIMA model, and they suggested using ARMA type models as preliminary models that might serve as the starting point for understanding the dynamics of the streamflow process.

Competing water uses for irrigation, future demand and hydropower generation and preparing for potential impacts of climate change (flood analysis) motivated our concern to forecast stream flow in this study. Therefore harnerssing the potenntials of the basin will help in preparing for accurate flood alert in advance.

2.0 Materials and Methods

2.1Materials

2.1.1 The Study Area

The main tributary of the Niger River in central Nigeria is the Kaduna River. It rises in the Jos plateau in a northwestern direction, south-west of Jos town, and north-east of Kaduna town. When it reaches Mureji in Niger State, it completes its flow to the Niger River by taking a southerly and south westerly direction(Garba et al., 2013). The river is 550 km long. From its source along the western margin of the Jos Plateau, the river flows northwest across the Kaduna plains. Just before it reaches the city of Kaduna, it turns to the southwest, cutting several gorges through rugged terrain between Kaduna and Zungeru. Finally, the river flows south through the broad, level Niger valley, and enters the Niger River opposite Pategi. Major tributaries joiningKaduna River along its course include the Mariga river, the Tubo river, the Sarkin river, the Pawa river,and the Galma river and (Abubakar et al., 2015). The section of the River in Niger State is the primary area of focus.

Average and Maximum Stream flow data from gauging station located at Wuya (latitude 9^{o} 07’ 12’’N and longitude 5^{o} 49’48’’E) which is owned and managed by the Nigerian Hydrological Service Agency NIHSA located in Niger State (Figure 1) was used for this study.

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

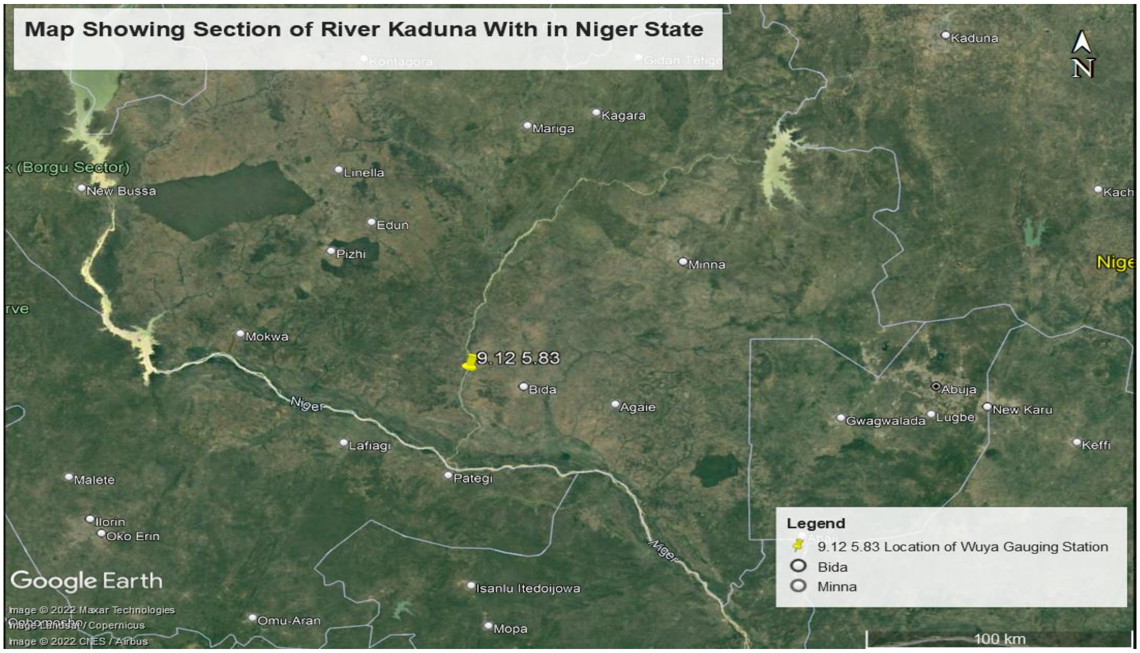

Figure 1: Map Showing Kaduna River and Location of Wuya Gauging Station in Niger State (Source: Google Earth 2022)

2.1.2 Data Collection

The stream flow data required for this work was the average monthly stream flows and peak discharge of Kaduna River which was obtained from the gauging station located at Wuya. The stream flow data from the gauging station was daily stream flow data in (m^3/s). The average and maximum/peak discharge data in (m^3/s)from 1988 to 2021 (33 years) was process and used for this research work. Microsoft Excel was used for this analysis. Also the statistical software used for the stochastic models in this reach work wasMini-Tab and IBM SPSS (Statistical Package for Social Science).

2.1.3 Time Series Analysis of Stream Flow Data

The Box Jenkins Methodology for time series analysis was utilized in this research work as described by Nochai and Nochai (2006), Adhikari and Agrawal (2013), Al-Saati et al. (2021). Figure 2 below is a schematic diagram depicting the flow process used by Box and Jenkins:

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

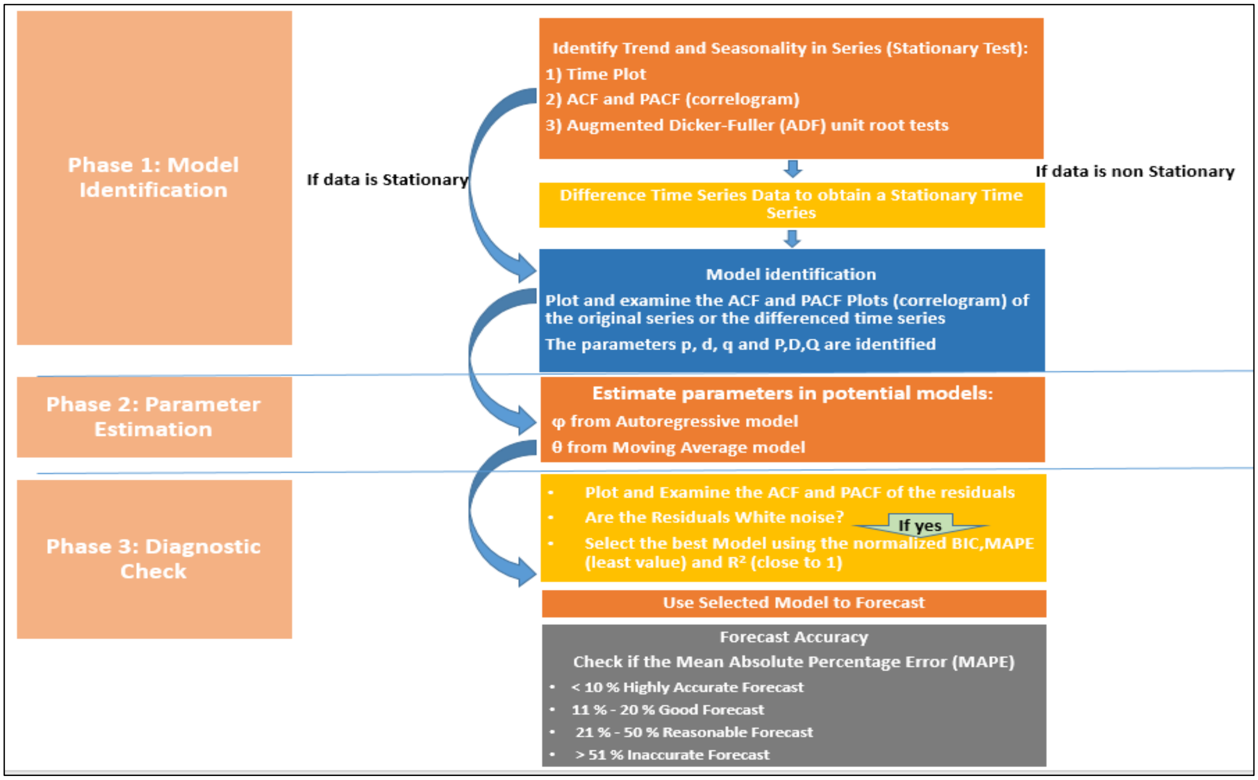

Figure 2: The Box-Jenkins Procedure for Best Model Selection

2.2 Methods

2.2.1 Time Series Plot

Time series plot is a data visualization tools that display data points at consecutive time intervals. each point in the plot corresponds to both time and the measured variable. The stream flow data of Kaduna River was plotted using the time series plot to depict the characteristic of the data either to have a trend or seasonal variation or whether it is already a stationary data.

2.2.2 Model Identification

At this stge the ACF graph was examined to see if the series is stationary or not. At this point, visual observation was used. The time series was be regarded as stationary if a graph of the ACF of the time series values either cuts off or dies down pretty soon. Time series data should be viewed as non-stationary if an ACF graph decays very slowly. If the series is not stationary, differencing would be used to make it stationary. In other words, a number of differences replace the original series. The differenced series is then supplied for an ARMA model. Differencing was carried out until a data plot shows that the series changes about a constant level and the ACF graph either abruptly ends or soon fades away (Nochai and Nochai, 2006).

2.2.3 Model Parameter Estimation

Parameters for a tentative model was estimated and the parameters were selected. The parameters estimated are ϕ from Autoregressive model and θ from Moving Average model.

2.2.4 Model Diagnostic Checks

The model was checked for adequacy by considering the properties of the residuals whether the residuals from an ARIMA model has the normal distribution and should be random. An overall check of model adequacy

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

was done by looking at the Ljung-Box Q statistic. If the p-value associated with the Q statistic is small (p-value < α), the model is considered inadequate. A new or modified model was generated and the analysis repeatedly performed until a satisfactory model is determined.

2.2.5 Forecasting with The Model

Forecasting with the identified Model for one period or several periods into the future with the parameters for a tentative model was done at this stage(Nochai and Nochai, 2006).

2.2.6 Forecasting Accuracy

After building the model, it is necessary to make a one-step-ahead forecast. Numerous calculations are generated and tested in accordance with the forecast accuracy to verify the forecasting accuracy, these includes the coefficient of determination, root mean squared error, mean absolute error, mean absolute percentage error, maximum absolute error, and maximum absolute percentage error, (Nochai and Nochai, 2006, Al-Saati et al., 2021).

3.0 RESULTS AND DISCUSSIONS

3.1 Time Series Plot

A time plot of stream flow versus time as generated for average and maximum stream flow of Kaduna River, a trend line was fitted to the plot. The time plot (Figure 3 and 4) reveals that there is seasonality, trend and elements of non-stationary in the stream flow data.

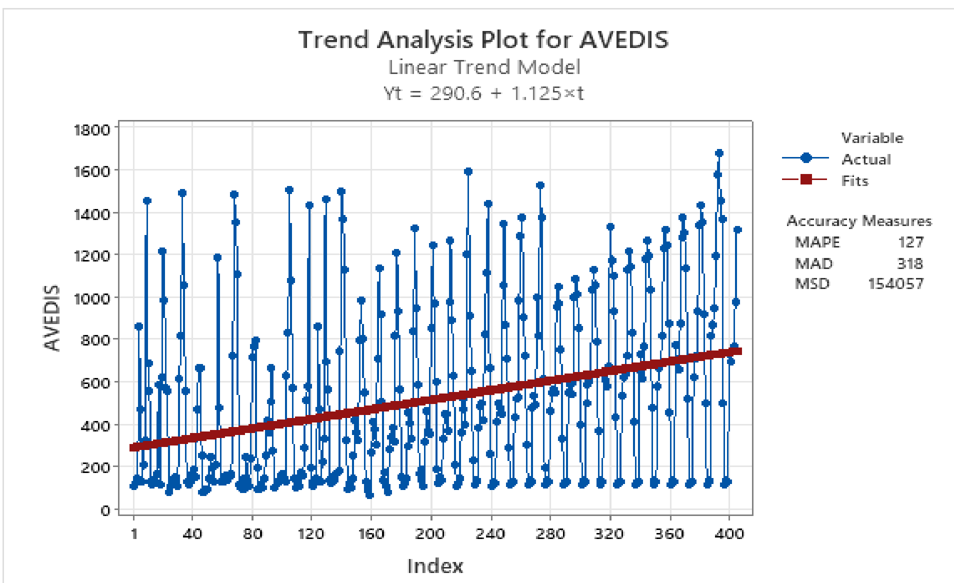

Figure 3: Plot Showing the Trend in Average Stream Flow of Kaduna River

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

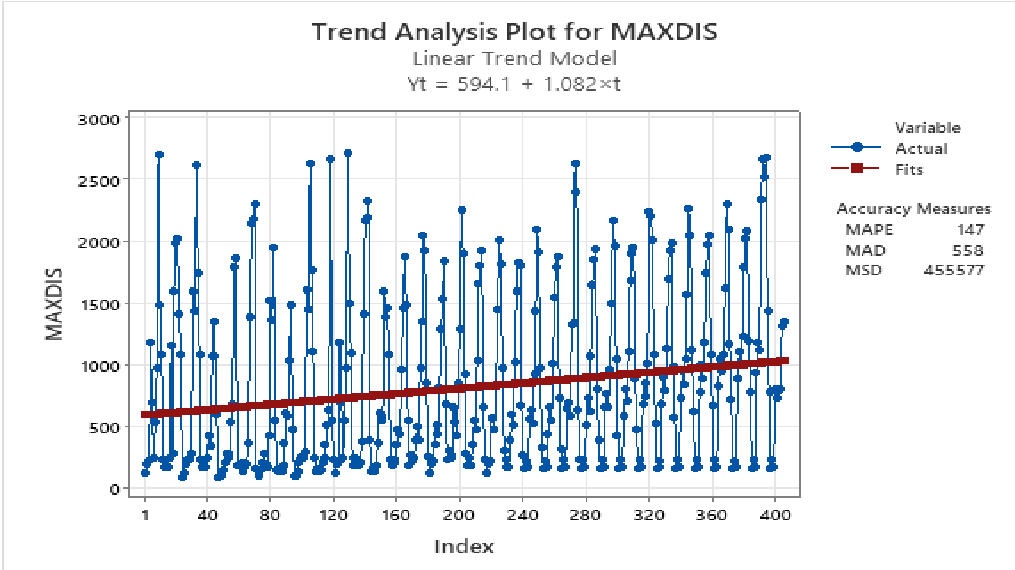

Figure 4. Plot Showing the Trend in Maximum Stream Flow of Kaduna River

3.2 Model Identification

Seasonality, trend and non-stationarity in the time series data for maximum and average monthly stream flow of Kaduna River was identified, hence, there is need to difference both time series to convert the time series from been non stationary to a stationary time series. This was achieved by applying the first order difference d=1 for the non-seasonal component and D=1 for the seasonal components. AR = 0 was selected for both average and maximum stream flow while MA=4 and MA=3 were selected for average and maximum stream flow respectively. The model identified hence were ARIMA (0,1,4)(0,1,4) and ARIMA (0,1,3)(0,1,3)for average and maximum stream flow respectively.

3.3 Parameter Estimation

The parameter estimates for ARIMA (0,1,4)(0,1,4) for the average stream flow discharge of Kaduna River is as presented in Table 1 while the Lag, chi-square, DF and P-values is presented in Table 2.

Table 1 Final Estimates of Parameters for ARIMA (0,1,4)(0,1,4)

Type	Coefficient	SE Coefficient	T-Value	P-Value
MA 1	0.64119	0.00547	117.21	0.000
MA 2	0.1579	0.0508	3.11	0.002
MA 3	0.1329	0.0580	2.29	0.023
MA 4	0.0624	0.0510	1.22	0.222
SMA 12	1.0357	0.0501	20.66	0.000
SMA 24	-0.2458	0.0726	-3.39	0.001
SMA 36	-0.0835	0.0726	-1.15	0.251
SMA 48	0.0962	0.0515	1.87	0.062
Constant	0.0587	0.0301	1.95	0.051

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

Table 2 Modified Box-Pierce (Ljung-Box) Chi-Square Statistic for ARIMA (0,1,4)(0,1,4)

Lag	12	24	36	48
Chi-quare	11.10	19.91	32.43	37.49
DF	3	15	27	39
P-Value	0.011	0.175	0.216	0.539

The parameter estimates for ARIMA (0,1,3)(0,1,3) for the maximum stream flow discharge of Kaduna River is as presented in Table 3 while the Lag, chi-square, DF and P-values is presented in Table 4

Table 3 Final Estimates of Parameters for ARIMA (0,1,3)(0,1,3)

Type	Coefficient	SE Coefficient	T-Value	P-Value
MA 1	0.7415	0.0446	16.62	0.000
MA 2	-0.0022	0.0545	-0.04	0.968
MA 3	0.2505	0.0486	5.16	0.000
SMA 12	1.1372	0.0496	22.94	0.000
SMA 24	-0.1815	0.0752	-2.41	0.016
SMA 36	-0.0687	0.0531	-1.29	0.197
Constant	0.1181	0.0364	3.25	0.001

3.4 Model Diagnostic Checks

Diagnostic check of the residual ACF and PACF for ARIMA (0,1,4)(0,1,4) for average stream flow reveals that all the Lags are within the 5% level of significance ($\pm 2/\sqrt{n}$) which signifies that the residuals are white noise as depicted in Figure 5 and Figure 6 while that of ARIMA (0,13)(0,1,3) for maximum stream flow reveals that all the Lags are within the 5% level of significance ($\pm 2/\sqrt{n}$) except one Lag which is slightly above the line of significance which signifies that the residuals are also white noise as depicted in Figure 7 and Figure 8.

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

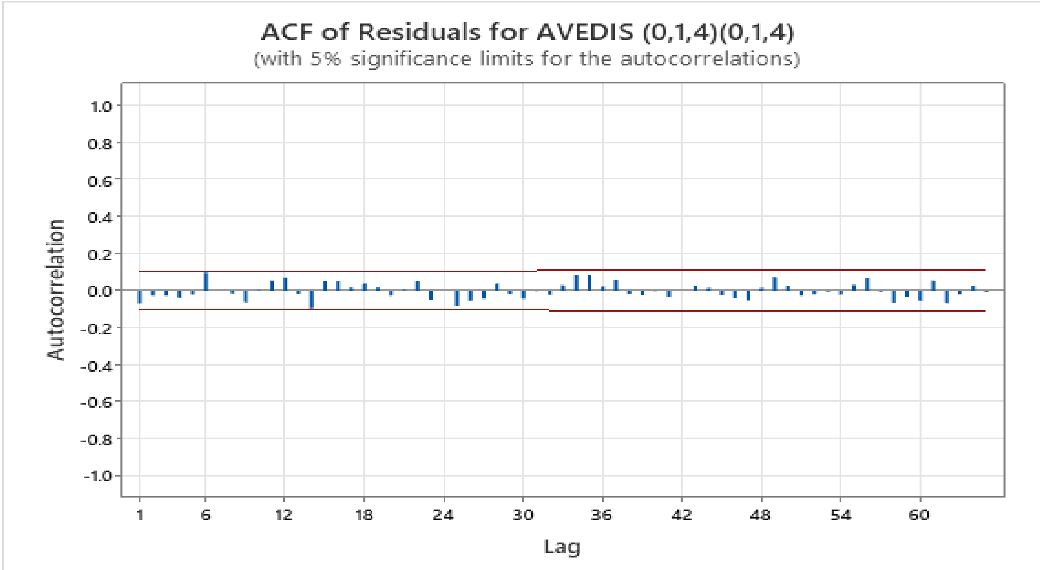

Figure 5 : ACF of Residual ARIMA (0,1,4)(0,1,4) on average stream flow data

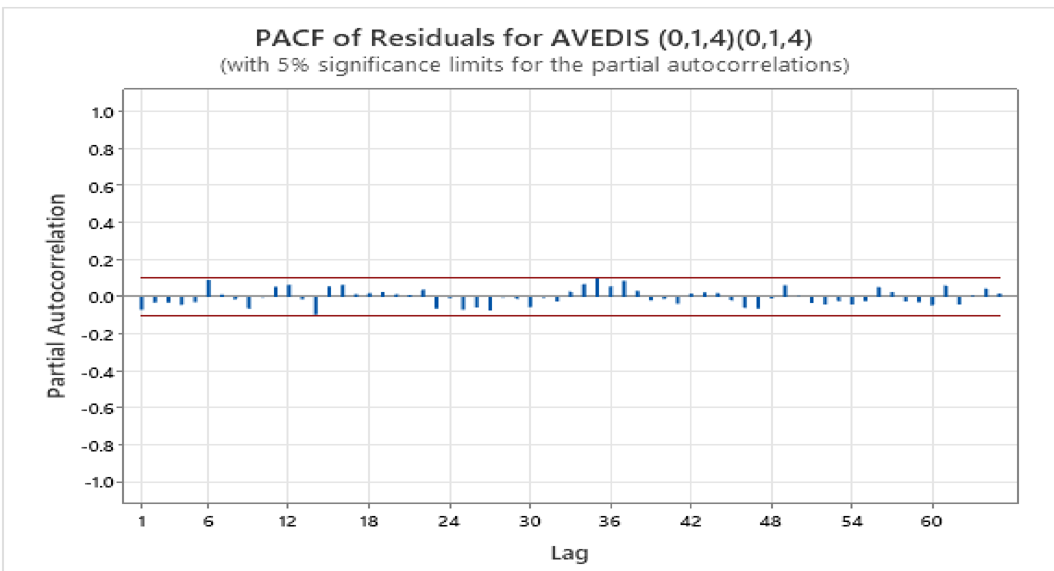

Figure 6: PACF of Residual ARIMA (0,1,4)(0,1,4) on average stream flow data

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

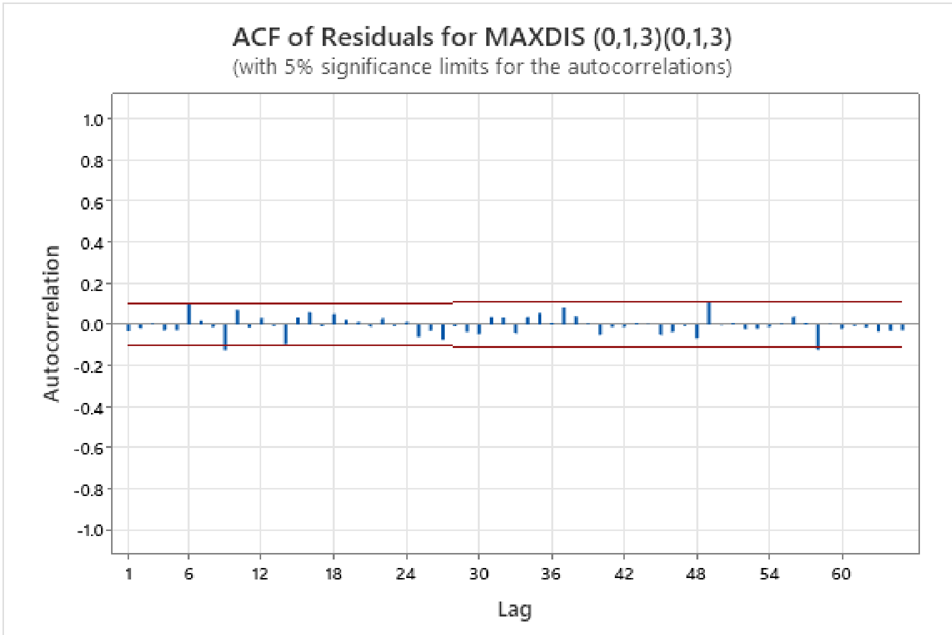

Figure 7: ACF of Residuals of ARIMA (0,1,3)(0,1,3) for maximum stream flow data

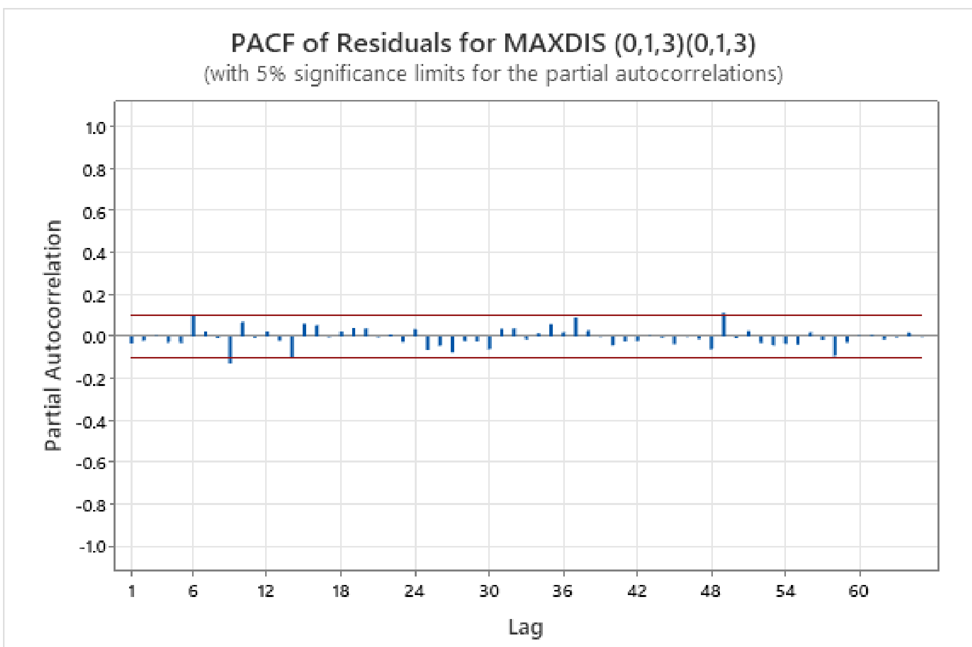

Figure 8: PACF of Residuals of ARIMA (0,1,3)(0,1,3) for maximum stream flow data

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

4.5 Forecasting with The Model and Forecast Accuracy

The identified models for average and maximum stream flow for Kaduna River ARIMA (0,1,4)(0,1,4) and ARIMA (0,1,3)(0,1,3) respectively were used to forecast future values of stream flow for a period of 5years. Figure 9, Figure 10 and Figure 11 are for ARIMA (0,1,4)(0,1,4), while Figures 12, 13 and 14 are for ARIMA (0,1,3)(0,1,3). .Both models have Mean absolute percentage error (MAPE) of 28.785%, 34.461% for ARIMA (0,1,4)(0,1,4) and between 21% to 50%. for ARIMA (0,1,3) (0,1,3).

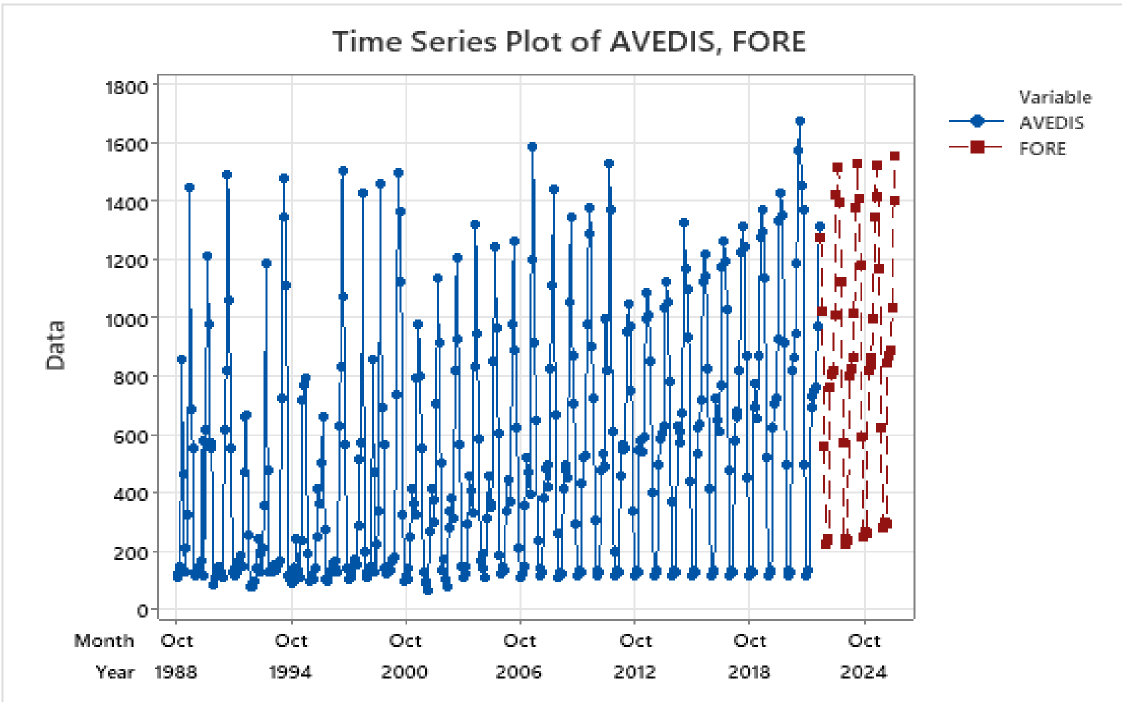

Figure 9: Time Plot of average stream flow data and forecasted ARIMA (0,1,4)(0,1,4)

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

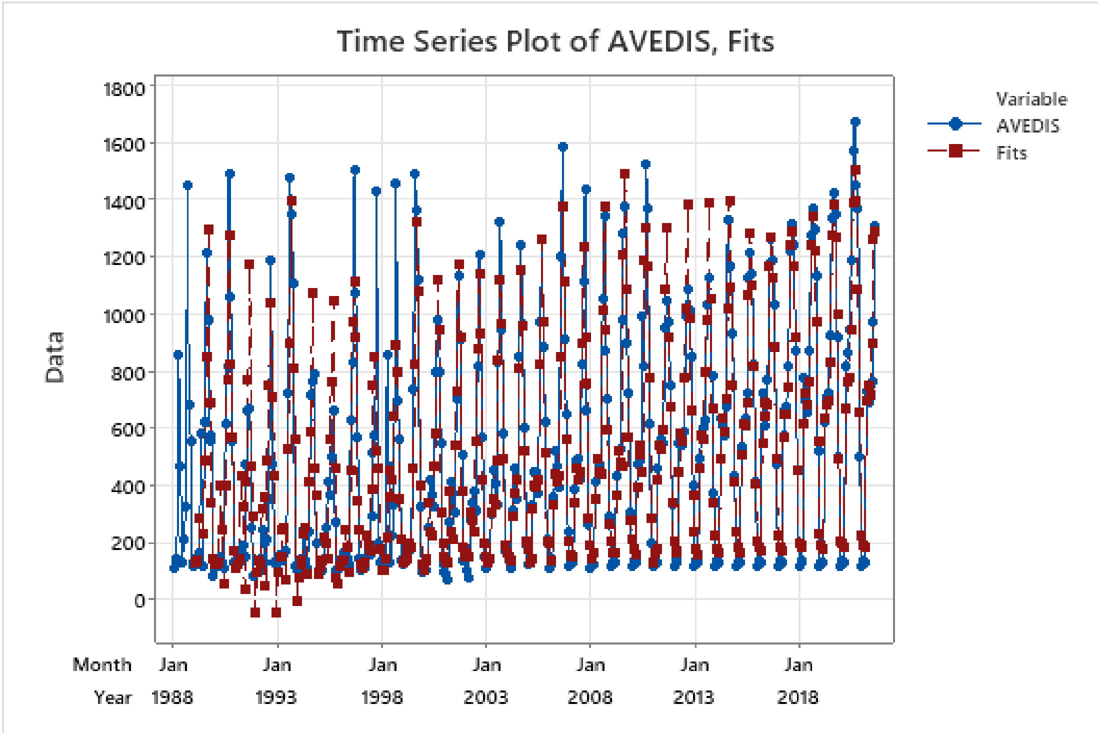

Figure 10: Time Plot of average stream flow data and Model fit valuesARIMA (0,1,4)(0,1,4)

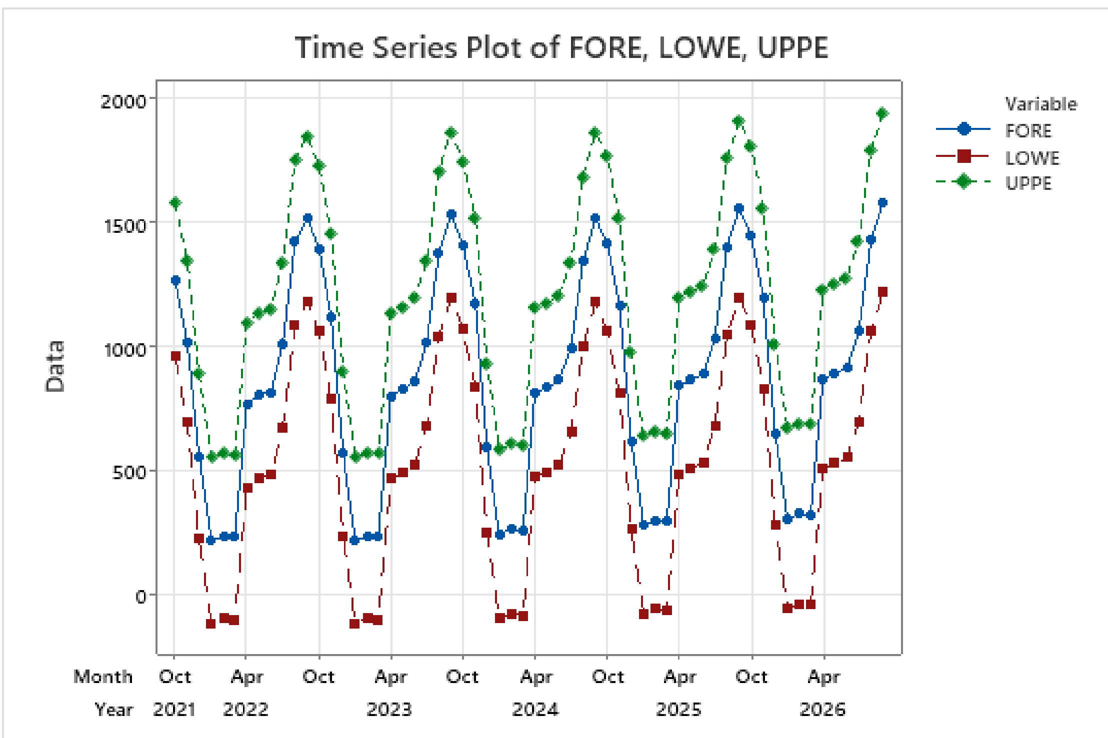

Figure11: Plot of forecasted, lower limit and upper limits of ARIMA (0,1,4)(0,1,4)

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

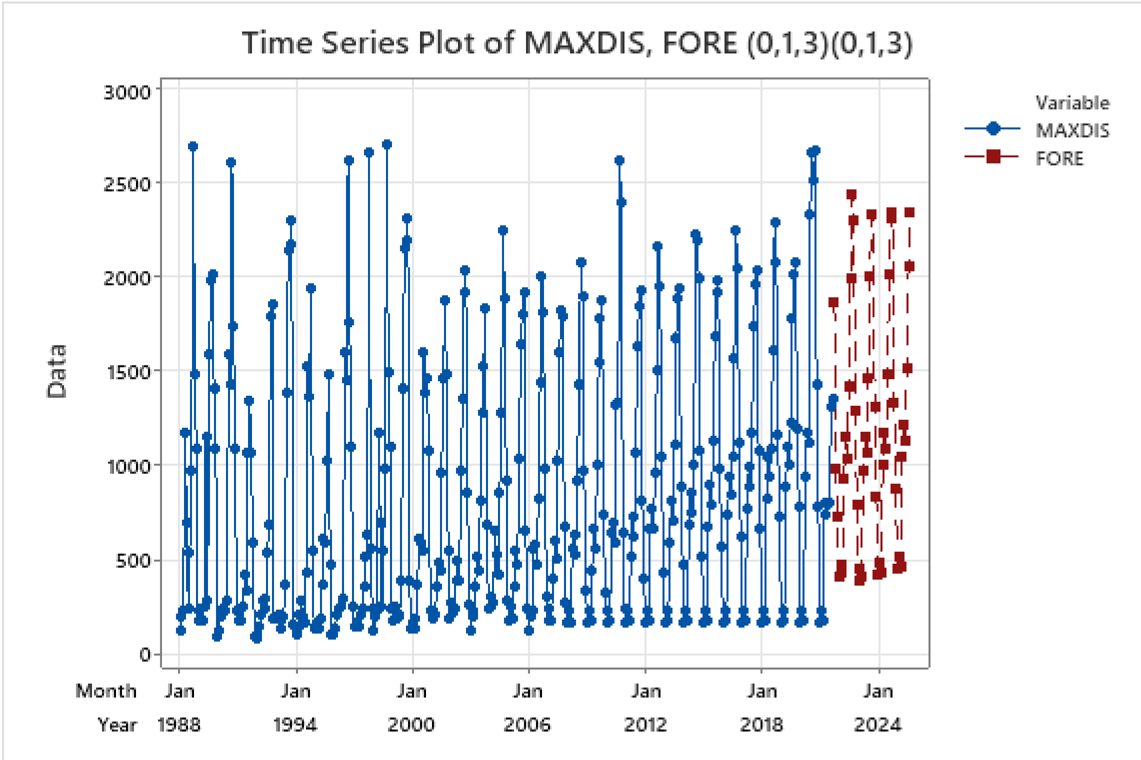

Figure 12: Time Plot of maximum stream flow discharge and forecasted ARIMA (0,1,3)(0,1,3)

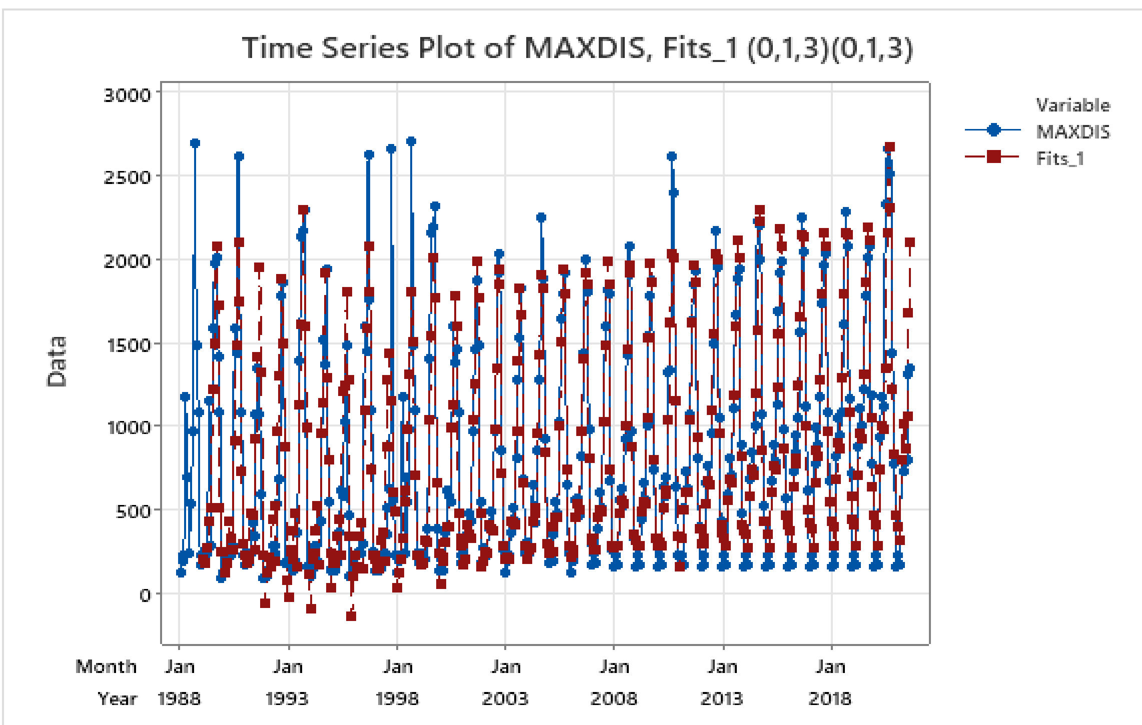

Figure 13:Time Plot of maximum stream flow data and Model fit valuesARIMA (0,1,3)(0,1,3)

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

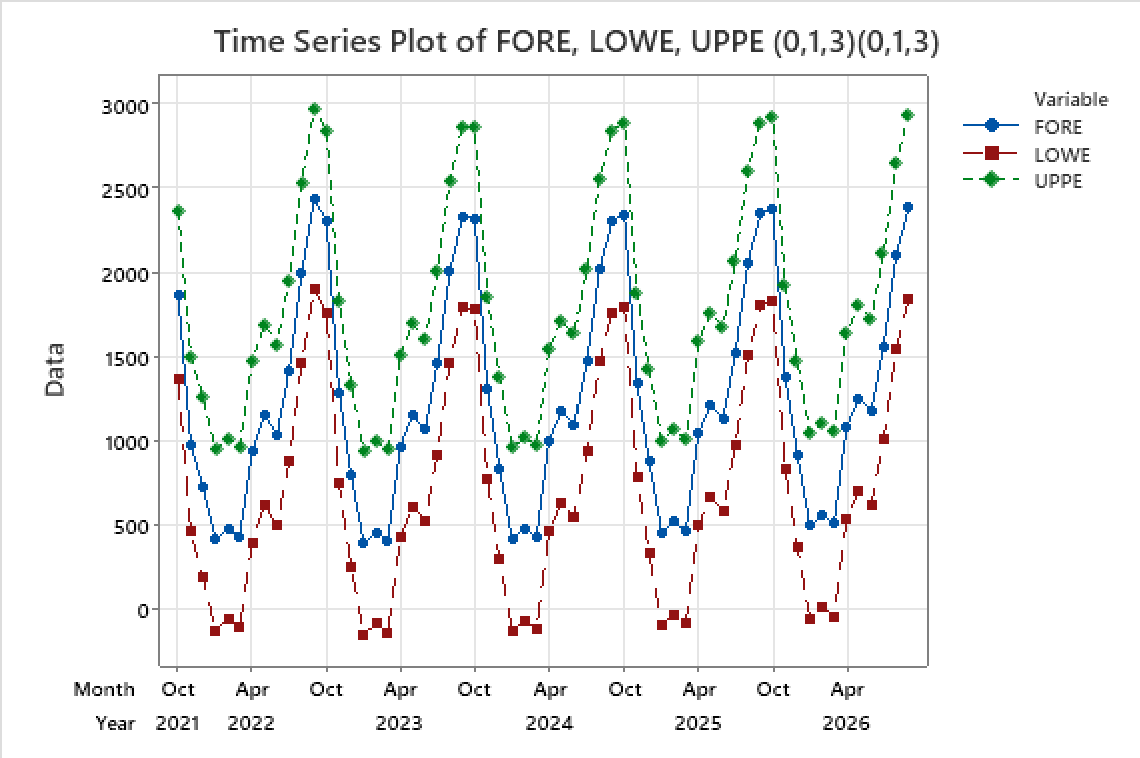

Figure 14:Plot of forecasted, lower limit and upper limits of ARIMA (0,1,3)(0,1,3)

4.0 CONCLUSION

Stream flow of Kaduna River was analyzed in this study. Auto correlation and Partial Auto correlation (correlogram), time plot and trend plot were used to determine the stationarity or otherwise of the data (i.e visual inspection) and all the methods used indicated that both stream flow data are non-stationary. This results demonstrate that both stream flow data needs to be differenced, and the differenced series be used for further analysis.

The initial analysis indicated that the average stream flow had a mean value of 518.9 m3/s, standard deviation of 414.5, coefficient of Skewness of 0.81 and coefficient of kurtosis of -0.42 while the maximum stream flow of Kaduna River had mean value of 813.7 m3/s, standard deviation of 687.6, coefficient of Skewness of 0.94 and coefficient of kurtosis of -0.2.

The Box Jenkins methodology was used to conduct stochastic analysis on the average and maximum stream flow data (from stationarity test to future forecasting). Using the autocorrelation function ACF and partial autocorrelation function PACF of the differenced series of the stream flow, the Autoregressive (p) and Moving Average (q) for the non-seasonal component and the Autoregressive (P) and Moving Average (Q) for the seasonal component of the anticipated model were calculated. The I and (i) for the seasonal and non-seasonal components, respectively, have already been established as being differenced in the other of 1 (i.e first differenced).

After studying the residuals of ACF and PACF and other results from the analysis of

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

stream flow for all the identified models of average and maximum stream flow of Kaduna River, it was found that ARIMA (0,1,4)(0,1,4) and ARIMA (0,1,3)(0,1,3) models were adequate to forecast average and maximum stream flow respectively, having their residuals as white noise.

Finally, the forecasted stream flows for average and maximum stream flow of Kaduna River showed an upward trend and this signifies the more likelihood of flood events within the period of forecast, both model forecast was reasonable having their MAPE between 21% to 50%.

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

REFERENCES

Abubakar, A. J., Olowolafe, E. A., & Obansa, Y. M. (2015). Land Use and Land Cover Change Analysis along Kaduna River Floodplain Using Geospatial Techniques. *IIARD International Journal of Geography and Environmental Management Vol. 1 No. 8 ISSN 2505-8821 Www.Iiardpub.Org*, *1*(8).

Adhikari, R., & Agrawal, R. (2013). An Introductory Study on Time Series Modeling and Forecasting Ratnadip Adhikari R. K. Agrawal. *ArXiv Preprint ArXiv:1302.6613*, *1302.6613*, 1–68.

Adnan, R. M., Yuan, X., Kisi, O., & Curtef, V. (2017). Application of Time Series Models for Streamflow Forecasting. *Organization, Technology and Management in Construction, Vol.9, No.*(1), 2137–2146.

Al-Saati, N. H., Omran, I. I., Salman, A. A., Al-Saati, Z., & Hashim, K. S. (2021). Statistical modeling of monthly streamflow using time series and artificial neural network models: Hindiya barrage as a case study. *Water Practice and Technology*, *16*(2), 681–691. https://doi.org/10.2166/wpt.2021.012

Boletti JM, Mendes JJ, Leme M & Sequema HK (2021); Comparative study of forecsting apparoaches in mothly straemflow series for Brazillian hydr electric power plant using extreme lerning machine. Journal of hydrlogy and hydrme. DOI; 10.2478/johd-2021-0001

Garba, H., Ismail, A., Batagarawa, R. L., Ahmed, S., Ibrahim, A., & Bayang, F. (2013). Climate Change Impact on Sub-Surface Hydrology of Kaduna River Catchment. *Open Journal of Modern Hydrology*, *03*(03), 115–121. https://doi.org/10.4236/ojmh.2013.33015

Martins, Y. O., Ahaneku, I. E., & Sadeeq, M. A. (2011). Parametric Linear Stochastic Modelling of Benue River flow Process. *Open Journal of Marine Science*, *01*(03), 73–81. https://doi.org/10.4236/ojms.2011.13008

Mohamed TM (2021) Forecasting of monthly flow for the White Nile River South Sudan. American journal of water scinece and engineering 2021 7(370) ;103-112;doi 10.116481/ajwse.20210703.12

Nabeel H, Al-Saalt, Islam I, Omran Salman AA, Saalt Z & Hashim KS (2021) Statistical modelling of monthly streamflow using time series and artifical newron networks models; Hindiya Barrage as a case study. Water practice and technology (2021) 16(2) 681-601. 2021

Nochai, R., & Nochai, T. (2006). Arima Model for Forecasting Oil Palm Price. *Proceedings of thea2nd IMT-GT Region Conference on Mathematics, Statstics and Applications, May 2005*, 1-7.

Pegram, G. G. . (2015). Probabilistic Methods and Stochastic Hydrology. In *Hydraulic structures, equipment and water data acquisition systems – Vol. I: Vol. I* (pp. 2014–2015).

Wang, L. K., & Yang, C. T. (2014). Modern water resources engineering. In *Modern Water Resources Engineering* (Issue November). https://doi.org/10.1007/978-1-62703-595-8

Wu, C. L., & Chau, K. W. (2010). Data-driven models for monthly streamflow time series prediction. *Engineering Applications of Artificial Intelligence*, *23*(8), 1350–1367. https://doi.org/10.1016/j.engappai.2010.04.003

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)

Conference on Structural Engineering and Construction Management.

[29] Zaid A. I. O and Travis F. W (2003); "An Examination of the Effect of Target Thickness in the Perforation of Mild Steel Plate by a Flat-ended Projectile". International Journal of Mechanical Sciences. Vol. 16, Issue 6, P. 373 – 380.

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY)