Modelling and Forecasting Standardized Net Allocation to Ondo State, Nigeria

Alabi, Oluwapelumi; Atanlogun, Sunday Kolawole;
Edwin, Oluwatoyin Ayodele
Rufus Giwa Polytechnic, Owo, Ondo State

Abstract

Fluctuations in federal allocations to state and local governments have significant implications, often leading to budget deficits, delayed salaries, stalled development projects, and mounting debt burdens. This study examines the characteristics of monthly net allocation series received by Ondo State from Nigeria's Federal Account Allocation Committee (FAAC) between January 2016 and February 2025. The data were standardized, and unit root and stationarity tests were conducted using KPSS and Augmented Dickey-Fuller tests. The series was found to be stationary at first-order difference. To model the standardized net allocation,

ARIMA (autoregressive integrated moving average) models were developed following the Box-Jenkins methodology. Model performance evaluation revealed relatively low forecast errors across various metrics. Finally, the fitted ARIMA model was applied for out-of-sample forecasting, with predictions from September 2024 to February 2025 compared against actual FAAC allocations.

Keywords: Modelling, forecasting, net allocation, Ondo State, ARIMA model

1. Introduction

Revenue allocation in Nigeria has been a complex and evolving issue since the colonial era. Early commissions such as the Phillipson Commission (1946), Hicks-Phillipson Commission (1951), and Raisman Commission (1958) attempted to establish fair systems for resource sharing. Post-independence, especially during and after the civil war, allocation became highly centralized, with the federal government controlling most revenues. The 1999 Constitution currently governs allocation, mandating the distribution of Federation Account funds as follows: 52.68% to the federal government, 26.72% to states, and 20.60% to local governments (Federal Republic of Nigeria, 1999). This structure aims to enable all levels of government to fulfill their constitutional duties.

Nigeria's heavy reliance on oil revenues, collected centrally and shared through the FAAC, means state governments Ondo State included depend heavily on federal allocations to finance budgets and projects. Created in 1976 from the former Western State, Ondo State is located in the South-West and is rich in petroleum, bitumen, and other natural resources. As an oil-producing state, it also receives the 13% derivation fund. However, despite these advantages, Ondo State remains fiscally dependent on FAAC transfers. Data from the National Bureau of Statistics confirm that most of the state's revenue is derived from federal allocations, raising concerns about its fiscal independence, internally generated revenue (IGR) capacity, and resilience to external shocks.

DOI: https://doi.org/10.5281/zenodo.17251528

Instability in federal allocations has severe implications for state finances, often resulting in deficits, delayed payments, incomplete projects, and rising debt. Eke (2018) highlighted how inconsistent transfers undermine capital project execution, while Adebayo and Yusuf (2020) emphasized that fiscal instability weakens service delivery and public trust. Between 2016 and 2024, Nigeria experienced significant political and economic disruptions including the aftermath of the 2014 oil price crash and the 2020 COVID-19 pandemic that reduced revenues and strained government finances. These shocks compelled states, including Ondo, to adopt fiscal adjustments, borrowing strategies, and revenue reforms. Analyzing federal allocation trends therefore provides vital insights into Ondo State's fiscal sustainability.

This study applies the ARIMA model, developed by Box and Jenkins (1976), to assess allocation dynamics. ARIMA (p, d, q) predicts future values based on past observations, differencing for stationarity, and past forecast errors. Its proven reliability in economic and financial forecasting makes it suitable for analyzing federal allocations to Ondo State.

Time series modeling seeks to create predictive frameworks that generate accurate forecasts from historical data. Time series may be stationary showing consistent statistical properties or non-stationary, requiring transformation for effective analysis (Wei, 2013; Parzen, 1961). Common tools include the Dickey-Fuller test, Fourier Transform, and Hilbert-Huang Transform, the latter being particularly effective for nonlinear and non-stationary data (Huang, 2003).

ARIMA models have been applied across numerous fields, including stock price forecasting (Adebiyi et al., 2014), agricultural yields (Manoj & Madhu, 2014; Padhan, 2012; Hamjah, 2014), GDP projections (Dritsaki, 2015; Zakai, 2016), crude oil price predictions (Abiola & Okafor, 2013), and inflation forecasting in multiple countries (Meyler et al., 1998; Adelekan et al., 2020; Abdulrahman et al., 2018; Jagero et al., 2023). The model has also been used in weather and environmental forecasting (Pinky et al., 2014; Mahmudur et al., 2013). Although machine learning and neural network models have emerged, studies (Serikov et al., 2021; Al-Saati et al., 2021) show that ARIMA often performs better with smaller datasets.

Given the volatility of FAAC allocations to Ondo State, this study develops an ARIMA based forecasting model to evaluate allocation trends, assess consistency, and provide insights for fiscal policy and planning.

2.0 Methodology

This study adopted a quantitative research approach, analyzing monthly FAAC allocations to Ondo State between January 2016 and February 2025, totaling 110 observations. Secondary data were sourced from FAAC reports. The analysis involved descriptive statistics, graphs, and inferential techniques. To enhance forecasting accuracy, the data were standardized to remove correlations among inputs and smooth the distribution. The standardized allocation was used for modeling and given by

$$y_t = \frac{X_i - \bar{X}}{\sigma_X} \qquad (1)$$

2.1 Time Series Model

2.1.1 The Moving Averages (MA)

A moving average is an average of a specific number of time series values around each value of t in the time series, except for the first few and last few terms. It is one of the techniques used for smoothing in time series

DOI: https://doi.org/10.5281/zenodo.17251528

analysis as well as in forecasting and it is only used on a time series that does not have a trend. An example of moving average series with an order q {MA (q)}

$$y_t = C_0 + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} \qquad (2)$$

Where C_0 is a constant, a_t is a white noise series, and $\theta_1, \theta_2, \ldots, \theta_q$ are model parameters (Tsay, 2010).

2.1.2. The Autoregressive Model (AR)

An autoregression refers to a time series model that uses previous observations to predict future observations. An example of AR (p) model with order p:

$$Y_t = \phi_0 + \phi y_{t-1} + \cdots \mp \phi_p y_{t-p} + a_t \qquad (3)$$

Where ϕ_0 is the constant term, ϕ_p are model parameters, and a_t is assumed to be a white noise series. The model is used to explain the weakly stationary stochastic time series and it is a combination of $AR\ (p)$ and $MA\ (q)$ models. An example of $ARMA\ (p, q)$ is given below:

$$Y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \alpha_t - \theta_1 \alpha_{t-1} - \alpha_t$$

$$-\theta_2 \alpha_{t-2} - \cdots - \alpha_t - \theta_q \alpha_{t-q} \qquad (4)$$

An $ARMA$ model combines the ideas of AR and MA models into a compact form so that the number of parameters used is kept small, achieving parsimony in parameterization (Tsay, 2010)

2.1.3 Autoregressive Integrated Moving Average (ARIMA)

The main difference between ARMA and ARIMA model is that there is an integration of differencing part in ARIMA of non-stationary data to ensure the assumption of stationarity is employed. ARIMA model is said to be a unit-root non-stationary because its AR polynomial has a unit-root and a conventional approach for handling unit-root non-stationary is to use differencing (Tsay, 2010). If the differencing $W_t = Y_t - Y_{t-1} = (1 - B)\, Y_t$ or higher-order differencing $W_t = (1 - B)^d Y_t$ of non-stationary time series then we call Y_t an $ARIMA\ (p, d, q)$ process with order $\boldsymbol{p}$ of AR process, $\boldsymbol{d}$ the number of differences made for a series to become stationary and $\boldsymbol{q}$ is the order of a moving average process.

$$\Phi_p(B)(1 - B)^d Y_t = \theta_q(B)\alpha_t \sim ARIMA(p, d, q) \qquad (5)$$

2.1.4 The Box Jenkins Methodology

Box and Jenkins (1970) describe a structured approach for identifying, estimating, validating, and applying autoregressive integrated moving average (ARIMA) models in time series analysis. This technique is recommended when there are at least 30 data points. According to Montgomery et al. (2015), the ARIMA modeling process involves three iterative stages: identifying the model by analyzing past data, estimating the model's unknown parameters, and performing diagnostic checks by examining residuals to assess the model's adequacy.

2.2 Model Identification

Identification of the appropriate ARIMA model requires skills obtained by experience (Box and Jenkins, 1970; Montgomery et al., 2015) postulates the following summary on how to identify the model.

DOI: https://doi.org/10.5281/zenodo.17251528

Model	Autocorrelation	Partial Autocorrelation
ARIMA $(\mathbf{p}, \mathbf{d}, \mathbf{0})$	**Infinite Tails off**	**Finite cut off after p lags**
ARIMA $(\mathbf{0}, \mathbf{d}, \mathbf{q})$	**Finite cut off after q lags**	**Infinite Tails off**
ARIMA $(\mathbf{p}, \mathbf{d}, \mathbf{q})$	**Infinite tail off**	**Infinite tail off**

The parameter p is determined from the partial autocorrelation function (PACF) of stationary data; if the PACF cuts off after a certain number of lags, the last significant lag represents the estimated value of p, while if no cutoff occurs, then $p = 0$ (Box & Jenkins, 1976). Similarly, the parameter q is obtained from the autocorrelation function (ACF) of stationary data; if the ACF cuts off after several lags, the final significant lag corresponds to the estimated value of q (Box & Jenkins, 1976). In an $ARIMA(p, d, q)$ model, the autocorrelation function typically exhibits a combination of exponential decay and damped sine wave patterns beyond the first $q - p$ lags (Box & Jenkins, 1970; Montgomery et al., 2015).

2.3 Parameter Estimation

Montgomery et al. (2015) state that parameter estimation in a tentatively identified model can be carried out using various approaches, including the method of moments, maximum likelihood, and least-squares. Since most ARIMA models are nonlinear, maximum likelihood estimation is often preferred once the values of p, d, and q have been determined. Additionally, backcasting can be applied to obtain estimates of the initial residuals (Box & Jenkins, 1976).

2.4 Diagnostic Checking

Model adequacy is assessed through residual analysis of both AR and MA components to determine whether the fitted model is appropriate. The residuals, or disturbances, should resemble a white noise process (Montgomery et al., 2015). For an adequate model, the residual scatter plot should display a rectangular pattern without any visible trends. Likewise, the residual sample autocorrelation function should show no identifiable structure (Montgomery et al., 2015). Statistical tests such as the approximate chi-square test of model adequacy and the Ljung-Box test may also be employed to confirm adequacy. Once an appropriate model has been established, it can then be applied for forecasting.

2.5 Forecasting

After the model has been validated and evaluated, it becomes suitable for forecasting. It can be applied to predict future values of the time series, and these forecasts can then support informed decision-making. The model is expressed by the following equation:

$$y_{(t)} = c + \emptyset_1 y_{(t-1)} + \emptyset_2 y_{(t-2)} + \cdots + \emptyset_p y_{(t-p)} + \in_{(t)} + \theta_1 \epsilon_{(t-1)} + \theta_2 \epsilon_{(t-2)} + \cdots + \theta_q \epsilon_{(t-q)} \quad (6)$$

Where $y_{(t)}$ is the output at time t, c is the constant term, $\emptyset_1, \emptyset_2, \dots, \emptyset_p$ are the AR coefficients, $\theta_1, \theta_2, \dots, \theta_q$ are the MA coefficients and $\in_{(t)}$ is the error term at time t.

3 Modeling

3.1 Data Description and Modeling

This paper examines the standardized monthly Federal Allocation to Ondo State from January 2016 to February 2025, using data obtained from the Federation Accounts Committee report. Figure 1 presents both the standardized monthly net allocation to the state government and its first difference. As shown in Figure 1(a), the series appears highly irregular and non-stationary. Since

DOI: https://doi.org/10.5281/zenodo.17251528

stationarity is a key assumption, a stationarity test was conducted. If the data is found to be non-stationary, differencing is applied until stationarity is achieved. Accordingly, the first-order difference of the series was taken, based on the results of the KPSS and ADF tests (see Table 1). Figure 1(b) indicates that the series becomes stationary after applying the first-order difference.

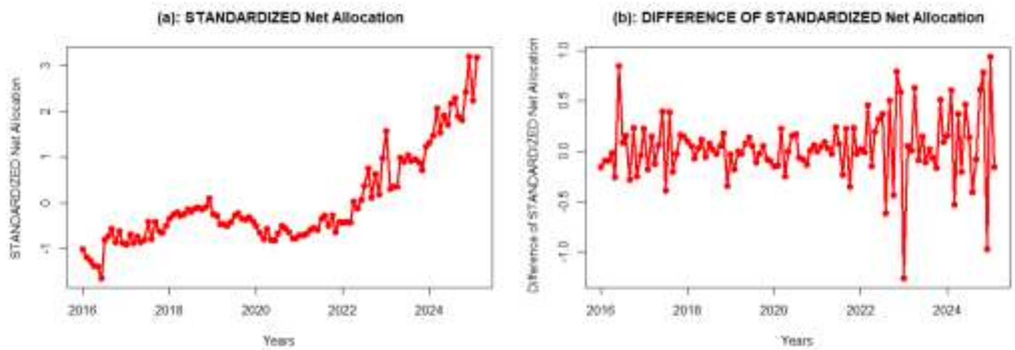


Figure 1: Time plot of Net Allocation to Ondo State Government (NAODSG)

Tests	Values	lag	P-value
(a) KPSS	0.43503	4	0.0621
(b)ADF	-5.7961	4	0.01

p-value smaller than printed p-value[(a)]
p-value greater than printed p-value[(b)]

The KPSS test is employed to determine whether the differenced series is stationary, while the Augmented Dickey-Fuller (ADF) test is used to check for the presence of a unit root. The KPSS test produced a p-value greater than the reported significance level, leading to the acceptance of the null hypothesis that the data is level or trend stationary. This suggests that the differenced series can be considered stationary. In contrast, the ADF test yielded a p-value smaller than the reported significance level, resulting in the rejection of the null hypothesis that the series contains a unit root. Together, these results confirm that the differenced series is stationary and free of unit roots.

For model development and validation, the dataset was split into training and test sets. The training set, covering net allocation data from January 2016 to August 2024, was used to build the model, while the test set, spanning September 2024 to February 2025, was reserved for evaluating the model's accuracy.

DOI: https://doi.org/10.5281/zenodo.17251528

3.1.1Model Identification

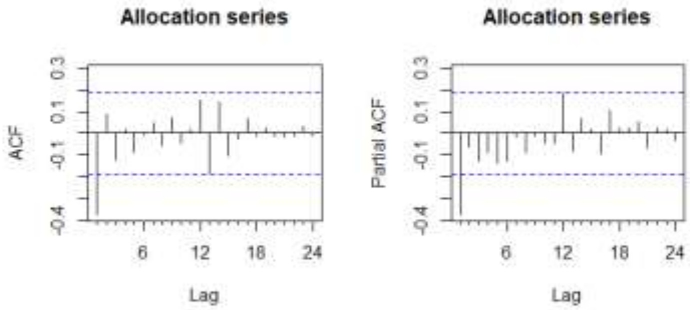


Figure 3**:** ACF and PACF Plots

The autocorrelation function (ACF) shows no clear evidence of slow decay, while the strong negative spike at lag 1 indicates that the series has short memory and can be appropriately modeled with an $MA(1)$ component. Similarly, the partial autocorrelation function (PACF) displays a large negative value at lag 1, with subsequent lags being relatively small and mostly within the confidence bounds, except for a weak spike around lag 12, which may suggest mild seasonality or random noise. Since higher-order lags are insignificant, additional AR terms are unnecessary, implying an $AR(0)$ structure from the PACF plot and an $MA(1)$ component from the ACF plot. To determine the best-fitting model, the $AICc$ values of all candidate models were compared, and the model with the lowest $AICc$ was selected (see Table 2). The final chosen model is $ARIMA(0,1,1)$. Let y_t represent the logarithm of net allocation; then the model can be expressed as:

$$y_t = y_{t-1} + \epsilon_t + \theta_1 \in_{t-1} + \mu$$

Where y_t , ϵ_t and θ_1 are observed value, white noise error term and moving average parameter, respectively.

Table 2: Different ARIMA (p,d,q) fitted models

ARIMA	Drift condition	AICc
(2,1,2)	with drift	25.29889
(0,1,0)	with drift	35.88682
(1,1,0)	with drift	22.47709
(0,1,1)	with drift	19.37092
(0,1,0)		35.13634
(1,1,1)	with drift	21.36151
(0,1,2)	with drift	21.40583
(1,1,2)	with drift	23.51479
(0,1,1)		22.04929

Table 3 presents various ARIMA models, each evaluated using the corrected Akaike Information Criterion ($AICc$), which is essential for model selection as it balances model fit against complexity. Lower AICc values indicate more suitable models that capture the underlying data patterns without introducing unnecessary complexity. Among the alternatives, the $ARIMA(0,1,1)$ model achieves the lowest $AICc$ value (19.371),

DOI: https://doi.org/10.5281/zenodo.17251528

suggesting that it provides the best compromise between accuracy and parsimony. In conclusion, based on the $AICc$ results, the $ARIMA(0,1,1)$ model is identified as the most appropriate choice, offering both strong goodness of fit and simplicity.

3.2 **Estimation**

The final model ARIMA (0,1,1) is estimated by Maximum .ikelihood estimation (MLE) including estimation of the parameters θ_1. The estimated model is

$$y_t = -0.4687 \in_{t-1} + 0.0322$$

$$\text{s.e.} \quad (0.0942) \qquad (0.0136)$$

$\sigma^2 = 0.06766$ $log\ likelihood = -6.56.$ $AIC = 19.13$ $AICc = 19.37$ $BIC = 27.03$

3.3 Ljung and Box Test

The Ljung–Box statistic is calculated as the weighted sum of squares of a sequence of autocorrelations and is used to test whether a time series is simply composed of random values (white noise). The Residual Sum of Squares (RSS), on the other hand, represents the sum of squared residuals and serves as a measure of the discrepancy between observed data and model estimates.

For the ARIMA(0,1,1) model, the Ljung–Box test yields a result of $\chi^2 = 4.8383$ with a p-value of 0.9017. Since the p-value is greater than 0.05, we fail to reject the null hypothesis, indicating that the residuals behave like white noise and can be considered independent and identically distributed (i.i.d.).

Figure 3 presents the residuals of the model, plotted as standardized values over time. The residuals generally fluctuate around zero, suggesting the absence of systematic bias in the forecasts. Furthermore, there is no evident trend or change in variance, which supports the adequacy of the model in capturing the underlying data structure.

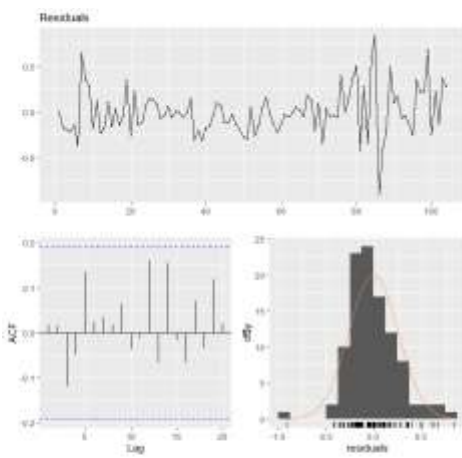


Figure 3: diagnostic plots for residuals from the selected ARIMA (0,0,1) model

3.3 **Evaluation of the Model**

The above ARIMA model is applied to forecast the net allocation to Ondo State Government from September 2024 to February 2025 and compared to the testing set and presented in Table 4

DOI: https://doi.org/10.5281/zenodo.17251528

Table 4: Actual Data against Forecasted Data

Period	Actual	Forecast	Lo 0	Hi 0
September	1.89105	2.200138	2.200138	2.200138
October	1.808568	2.232354	2.232354	2.232354
November	2.423135	2.26457	2.26457	2.26457
December	3.204135	2.296786	2.296786	2.296786
January	2.23231	2.329002	2.329002	2.329002
February	3.17349	2.361218	2.361218	2.361218

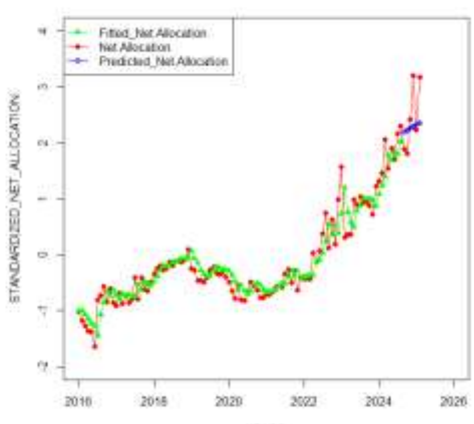
Table 5: Model Evaluation

MODEL	MSE	RMSE	MAE	MAPE
ARIMA(0,1,1)	0.089708	0.0081	0.00531	0.5333209

A comparison was conducted between the actual net allocation and the forecasted values generated by the ARIMA(0,1,1) model for the period spanning September 2024 to February 2025. The results show that the ARIMA model's forecasts closely aligned with the actual net allocation, effectively capturing the underlying patterns and trends in the data, as illustrated in Figure 4.

The model's performance evaluation is summarized in Table 5. The Mean Squared Error (MSE) was 0.089708, indicating the average squared deviation between the actual and forecasted values. The Root Mean Squared Error (RMSE) of 0.0081 measured the standard deviation of forecast errors, while the Mean Absolute Error (MAE) of 0.00531 reflected the average absolute difference between actual and predicted values. Furthermore, the Mean Absolute Percentage Error (MAPE) of 0.5333% highlighted the relative accuracy of the forecasts.

These results demonstrate that the $ARIMA(0,1,1)$ model achieved relatively low forecast errors across multiple evaluation metrics, confirming its adequacy and reliability in modeling the net allocation series.



DOI: https://doi.org/10.5281/zenodo.17251528

Figure 4: Standardized, fitted and predicted net allocation

4.1 **Conclusion**

This study modeled and forecasted FAAC allocations to Ondo State using an $ARIMA(0,1,1)$ model. Findings revealed that the model effectively captured the allocation series' dynamics, with forecasts closely tracking actual outcomes. The model demonstrated low forecast errors, making it a reliable tool for predicting future allocations. Accurate forecasts can help the state government anticipate fiscal inflows and plan expenditure more effectively, reducing risks of budget shortfalls and project disruptions.

4.2 **Recommendations**

1. Continuous Monitoring: Since allocations are influenced by dynamic economic factors, the government should regularly update the forecasting model to account for new conditions.
2. Model Improvement: Although ARIMA performed well, alternative models should be considered periodically to ensure continued accuracy as fiscal conditions evolve.
3. Policy Application: Forecast outcomes should guide fiscal policies, helping the state manage expenditures, borrowing, and development projects in line with expected revenues.
4. Revenue Diversification: Ondo State should strengthen its internally generated revenue base to reduce dependence on volatile FAAC transfers.

References

Abdulrahman, B. M. A., Ahmed, A. Y. A., & Abdellah, A. E. Y. (2018). Forecasting of Sudaninflation rates using ARIMA model. *International Journal of Economics and Financial Issues, 8*(3), 17.

Adebayo, T., & Yusuf, A. (2020). Revenue allocation and economic performance of Nigerian states: A panel data approach. *Journal of Public Administration and Governance, 10*(1), 45–61.

Adebiyi, A. A., Adewuni, A. O., & Ayo, C. K. (2014). Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics.*

Adelekan, O. G., Abiola, O. H., & Constance, A. U. (2020). Modelling and forecasting inflation rate in Nigeria using ARIMA models. *Journal of Economics and Finance, 11*(2), 25–34.

Abiola, A. G., & Okafor, H. O. (2013). Searching for appropriate crude oil price benchmarking method in the Nigerian budgeting process. *Developing Country Studies, 3*(12).

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control* (2nd ed.). San Francisco: Holden-Day.

Dritsaki, C. (2015). Forecasting GDP using ARIMA models: An application to Greece. *International Journal of Economics and Business Administration, 3*(1), 27–40.

Eke, O. A. (2018). Fiscal federalism and revenue allocation in Nigeria: An overview. *Journal of Economics and Public Policy, 13*(2), 45–58.

Hamjah, M. A. (2014). Rice production forecasting in Bangladesh: An application of Box-Jenkins ARIMA model. *Mathematical Theory and Modeling, 4*(4), 1–11.

Huang, N. E., et al. (2003). Applications of Hilbert–Huang transform to non-stationary financial time series analysis. *Applied Stochastic Models in Business and Industry, 19*(3), 245–268.

Jagero, N., Osei, B., & Omondi, G. (2023). ARIMA model validation on inflation data: Evidence from Africa. *Journal of Economic Studies, 50*(2), 188–203.

DOI: https://doi.org/10.5281/zenodo.17251528

Mahmudur, R., Islam, A. H. M. S., Nadvi, S. Y. M., & Rahman, R. M. (2013). Comparative study of ANFIS and ARIMA model for weather forecasting in Dhaka. *IEEE Conference Proceedings*, 1–6.

Manoj, K., & Madhu, A. (2014). An application of time series ARIMA forecasting model for predicting sugarcane production in India. *Faculty of Economic Sciences*, *9*(1), 81–94.

Meyler, A., Kenny, G., & Quinn, T. (1998). Forecasting Irish inflation using ARIMA models. *Central Bank of Ireland Technical Paper Series.*

Okonjo-Iweala, N., & Osafo-Kwaako, P. (2007). Nigeria's economic reforms: Progress and challenges. *Brookings Global Economy and Development Working Paper 6.*

Parzen, E. (1961). An approach to time series analysis. *The Annals of Mathematical Statistics, 32*(4), 951–989.

Padhan, P. C. (2012). Application of ARIMA model for forecasting agricultural productivity in India. *Journal of Agriculture and Social Science, 8*(2), 50–56.

Pinky, S. D., & Tahbilder, H. (2014). Prediction of rainfall using data mining techniques over Assam. *Indian Journal of Computer Science and Engineering*, *5*(2), 1–6.

Serikov, T., Zhetpisbayeva, A., Akhmediyarova, A., Mirzakulova, S., Kismanova, A., Tolegenova, A., & Wójcik, W. (2021). City backbone network traffic forecasting. *International Journal of Electronics and Telecommunications, 67*(3), 319–324.

Tsay, R. S. (2010). *Analysis of financialtime series* (3rd ed.). Hoboken, NJ: Wiley.

Wei, W. W. S. (2013). *Time series analysis: Univariate and multivariate methods* (2nd ed.). Boston: Pearson.

Zakai, M. (2016). A time series modelling on GDP of Pakistan. *Journal of Contemporary Issues in Business Research, 3*(4), 200–210.

DOI: https://doi.org/10.5281/zenodo.17251528