

FORECASTING NIGERIA'S INFLATION RATES (2009–2024) USING SARIMA MODELS

Ifeanyi Gabriel Nwosu

Department of Agricultural and Applied Economics, Rivers State University, Nkpolu-Oroworukwo, Port-Harcourt, Rivers State, Nigeria.

DOI: <https://doi.org/10.5281/zenodo.17368249>

Abstract

The study evaluated the accuracy and performance of SARIMA Models in Forecasting Inflation Rates. The study applied the Box-Jenkins methodology to build an ARIMA model for forecasting Nigeria's monthly inflation rates from November 2009 to October 2024. The results indicated that the ARIMA (3,2,1) (2,0,1) [12] model provided the best fit for predicting monthly inflation rates in Nigeria. This model was then used to forecast inflation from June 2024 to January 2026. The forecasted results are expected to provide policymakers with valuable insights for designing more effective economic and monetary policies, especially to address the forecasted rise in inflation rates in the first quarter of 2026, assuming all other factors remain unchanged.

Key Words: Accuracy, Evaluation, Forecasting, Performance, & Inflation Rates

Introduction

Inflation remains one of the most significant economic challenges faced by Nigeria, influencing a range of macroeconomic factors including purchasing power, economic stability, and investment decisions. As of recent years, Nigeria's inflation rate has experienced considerable fluctuations, driven by factors such as volatile oil prices, currency depreciation, and structural issues within the economy. According to the Central Bank of Nigeria (CBN), inflation has consistently been a concern, with rates reaching double digits in many instances, thereby affecting the standard of living for citizens and complicating the policy-making process. With inflationary pressures on the rise, accurate forecasting becomes vital for effective monetary and fiscal policies. Tuaneh and Essi (2017) reported that the value of the GDP is not very meaningful with consistent inflationary pressure

Several authors, Tuaneh (2018), Tuaneh and Okidim (2019), Tuaneh and Essi (2021), Tuaneh et al (2021), and also Tuaneh and Doodei (2025), have studied inflation rate but in relationship with other variables. also, despite several attempts to predict inflation using various econometric models, previous studies in Nigeria have often encountered limitations, particularly in accounting for both non-seasonal and seasonal patterns inherent in inflation data. Many studies have primarily relied on simpler ARIMA models that overlook the seasonal aspects of inflation, leading to suboptimal forecasts. Furthermore, the lack of consistency and reliability in inflation forecasts has hindered policymakers' ability to make well-informed decisions.

This study seeks to address the existing gap by evaluating the accuracy and performance of SARIMA (Seasonal Autoregressive Integrated Moving Average) models, with a particular focus on the ARIMA (2,2,3)

(2,0,0)[12] model, in forecasting inflation in Nigeria from 2009 to 2024. By incorporating both non-seasonal and seasonal components, SARIMA models offer a more robust approach to capturing the underlying patterns of inflation, including potential annual cycles. This is crucial given the influence of factors such as seasonality in food prices, global commodity shocks, and annual policy adjustments on inflation.

Previous studies on the application of SARIMA models in modelling and forecasting Nigeria's inflation rates (Otu, Osuji, Opara, Mbachu, Iheagwara, 2014, Havi, 2023) have been limited by the inability to effectively capture the seasonal nature of inflation, often leading to inaccurate predictions that fail to inform effective policy decisions. With inflation remaining volatile and unpredictable, there is a clear need for more accurate forecasting models that integrate both trend and seasonal components. Therefore, the problem this study seeks to address is the inadequacy of existing models in providing reliable inflation forecasts that reflect both short-term and long-term dynamics in Nigeria's economy. Therefore, the study will evaluate the effectiveness of the SARIMA model in forecasting inflation in Nigeria, assess the accuracy of the SARIMA model by comparing the error metrics (ME, RMSE, MAE, MPE, MAPE, MASE) with other traditional forecasting models, identify the seasonal and non-seasonal components contributing to inflation dynamics in Nigeria, using the SARIMA model and provide policy recommendations based on the model's forecasting performance and its potential for improving inflation management in Nigeria.

3.0 Methodology

3.1 Data source for the study.

The study used data on inflation rate extracted from the Central Bank of Nigeria (CBN) website www.cbn.ng. The data was extracted from 1st January 1991 to 31st May 2024. The statistical software is r-studio. This is powerful statistical software that allows users to analyze, manage and produce graphical displays of data.

3.2 Model Specification

A model is a simplified system used to simulate certain aspects of the real economy. The method specified for this study is the Box-Jenkins approach (Box and Jenkins, 1976), which incorporates the Autoregressive Integrated Moving Average (ARIMA) model. The ARIMA model seeks to identify patterns in historical data and decomposes it into three main components and they include ; an autoregressive (AR) process, which reflects the memory of past events; an integrated (I) process, which accounts for stabilizing or making the data stationary, thus making it valid for forecasting; and a moving average (MA) process, which models the forecast error (Deebom, Essi & Amos, 2021). The longer the historical data, the more accurate the forecast, as the model learns over time (Out. et al, 2014). These components combine and interact with each other, eventually forming the ARIMA(p,d,q) model. The first component, the AR term, uses the p lags of a time series to improve the forecast. The AR part of ARIMA indicates that the evolving variable of interest is regressed on its own lagged (previous) values. An AR(p) model is expressed in the following form, as shown in equation 1:
$$y_t = \mu + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \dots + \beta_p y_{t-p} + \epsilon_t \quad (3.1)$$

$$= \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i} + \epsilon_t$$

Where;

y_t = The response (dependent) variable being forecasted at time t

y_{t-1} = The lag of the series or the response variable at time lag
(Stimulus)

$\alpha_1, \dots, \alpha_p$ = Are the coefficient of lag that the model estimates

μ = Is the intercept term also estimated by the model

ϵ_t = Error term at time t

This equation demonstrates that the forecasted value of inflation at time t depends on its value in the previous period and a constant. The second component is the integrated stochastic process. A time series is said to be integrated of the first order, $I(1)$, if it must be differenced once to make it stationary. In general, if a time series must be differenced d times to become stationary, it is said to be integrated of order d , denoted as $I(d)$ (Gujarati, 2003).

Similarly, the third component, the $MA(q)$ model, uses the q lags of forecast errors to improve the forecast. The MA part indicates that the regression error is a linear combination of error terms whose values occurred both contemporaneously and at various points in the past. An $MA(q)$ model has the form shown in equation 2.

$$y_t = \beta_0 + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \dots + \beta_q \epsilon_{t-q} + \epsilon_t \quad (3.2)$$

$$= \beta_0 + \sum_{i=1}^q \beta_i \epsilon_{t-i} + \epsilon_t$$

Where: y_t = The response (dependent) variable being forecasted at time t

β_0 = The constant mean of the process

$\beta_1, \beta_2, \dots, \beta_q$ = The coefficient to be estimated

ϵ_t = is the error term at time t

$\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-q}$ = the error in previous time that are incorporated in the response y_t .

This equation indicates that y at time t is equal to a constant plus a moving average of the current and past white noise error terms. However, if no differencing is required to make the series stationary, then an ARMA model is generated with d equal to zero. The autoregressive moving average (ARMA) model refers to a model with p autoregressive terms and q moving average terms. An $ARMA(p, q)$ model is stationary if the series is stationary, as shown in equation 3.

$$y_t = \epsilon_t + \sum_{i=1}^p \alpha_i y_{t-i} + \sum_{j=1}^q \beta_j \epsilon_{t-j} \quad (3.3)$$

To create an ARIMA model, we begin by combining or adding both the autoregressive (AR) process, the moving average (MA) process and the integrated part (I) together as shown in equation

$$(3.4) \quad y_t = \mu + \epsilon_t + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \dots + \beta_q \epsilon_{t-q} \quad (3.4)$$

$$y_t = \mu + \alpha_t + \sum_{i=1}^p \alpha_i y_{t-i} + \sum_{j=1}^q \beta_j \nabla_{t-j} y_t = \nabla_d y_t = (1-B)^d y_t$$

The ARIMA (p,d,q) model can be specified using the backshift operator as:

$$\nabla(B)(1-B)^d Y_t = \nabla(B) \nabla_t \quad (3.5)$$

Similarly, the SARIMA (Seasonal Autoregressive Integrated Moving Average) model is a time series forecasting model that extends the ARIMA model by explicitly modeling seasonality in the data. SARIMA is particularly useful for datasets that show seasonal patterns or trends over time. The general form of the SARIMA model is written as:

$$\nabla_p(\beta^s). (1 - \beta^s)^D Y_t = \theta(\beta^s) \varepsilon_t \quad (3.6) \text{ Where:}$$

Y_t is the time series (inflation Rates) β is the backshift operator $\beta^k Y_t = Y_{t-k}$ s is the length of the seasonal period. D is the number of seasonal differences.

p is the number of seasonal autoregressive terms.

Q is the number of seasonal moving average terms.

ε_t the white noise (error term).

Similarly, the SARIMA model also sometimes referred to as the Multiplicative Seasonal Autoregressive Integrated Moving Average model, is denoted as ARIMA(p,d,q)(P,D,Q) S . The corresponding lag form of the model is:

$$\phi(L)\phi(L^S)(1-L)^d(1-L^S)^D y_t = \theta(L)\theta(L^S)\varepsilon_t \quad (3.7)$$

This model includes the following AR and MA characteristic polynomials in L of order p and q respectively:

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_{p-1} L^{p-1} - \phi_p L^p$$

$$(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_{q-1} L^{q-1} - \theta_q L^q$$

Also, seasonal polynomial functions of order p and q respectively as represented below:

$$(L^S) = 1 - \phi_1 L^S - \phi_2 L^{2S} - \dots - \phi_{P-1} L^{(P-1)S} - \phi_P L^{PS}$$

$$(L^S) = 1 - \theta_1 L^S - \theta_2 L^{2S} - \dots - \theta_{Q-1} L^{(Q-1)S} - \theta_Q L^{QS}$$

Where: $\{y\}$ - the observable time series

$\{\varepsilon_t\}$ -white noise series

p, d, q - order of non-seasonal AR, differencing and non-seasonal MA respectively

P, D, Q - order of seasonal AR, differencing and seasonal MA respectively L -lag operator $L^k y_t = y_{t-k}$

S -seasonal order for example $S=12$ for monthly data

3.4 SARIMA model Estimation Procedures:

1. **Stationarity Check:** Ensure the series is stationary, either by differencing or using transformations.
2. **Seasonality Identification:** Identify if the data exhibits seasonal patterns (e.g., using autocorrelation plots or seasonal decomposition).
3. **Model Selection:** Choose the appropriate values for p, d, q, P, D, Q , and s using techniques such as grid search, ACF/PACF plots, or criteria like AIC/BIC.
4. **Model Fitting:** Fit the SARIMA model to the data using software libraries (e.g., stats models in Python).

5. **Forecasting:** Use the fitted model to make forecasts.

3.3 SARIMA Model Selection

To test whether the series meet the stationarity condition—which denotes time invariant mean, variance, and co-variance—is the first step in creating the SARIMA model. Finding the orders p , q , P , and Q can be aided by visualizing the patterns of the ACF and PACF. These give an idea of the seasonal and non-seasonal lags by using the data on internal correlation between time series observations made at various intervals. At the non-seasonal and seasonal levels, respectively, the ACF and PACF both exhibit spikes and cutoffs at lag k and lag ks . The number of notable spikes indicates the model's order. Shumway and Stoffer (2006) state that Table 1 below illustrates the behavior of ACF and PACF that were taken from AIDOO (2011).

Table 1 Behavior of ACF and PACF for seasonal and Non-seasonal ARMA(p,q)

	Estimator	AR(p)	MA(q)	ARMA(p,q)
Non-seasonal ARMA(p,q)	ACF	Tails off at lag k $K=1,2,3,\dots$	cuts off after lag q	Tails off
	PACF	Cuts off after lag p	Tails off at lags k $k=1,2,3,\dots$	Tails off
		AR(P)s	MA(Q)s	ARMA(P,Q)s
	ACF	Tails off at lag ks $K=1,2,3,\dots$	cuts off after lag Qs	Tails off at ks
pureseasonal ARMA(p,q)	PACF	Cuts off after lag P_s	Tails off at lags ks $k=1,2,3,\dots$	Tails off at ks

Source: Shumway and Stoffer (2006) cited in Aidoo (2016).

The Maximum Likelihood approach is used to estimate the parameters of the various models that the ACF and PACF may produce. The model deemed most suitable is the one with the lowest AIC and BIC selection criterion values. Relative diagnostic checking is the final step in the model selection process; if the model passes these tests, it can be used for forecasting.

Results

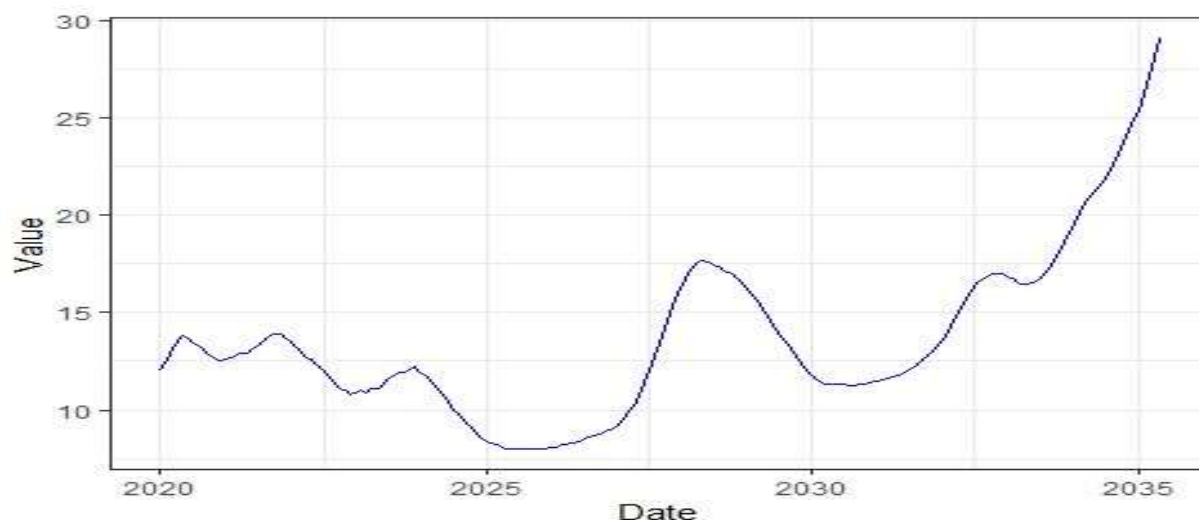


Figure 1: Time plot for the data from 2009 to 2024

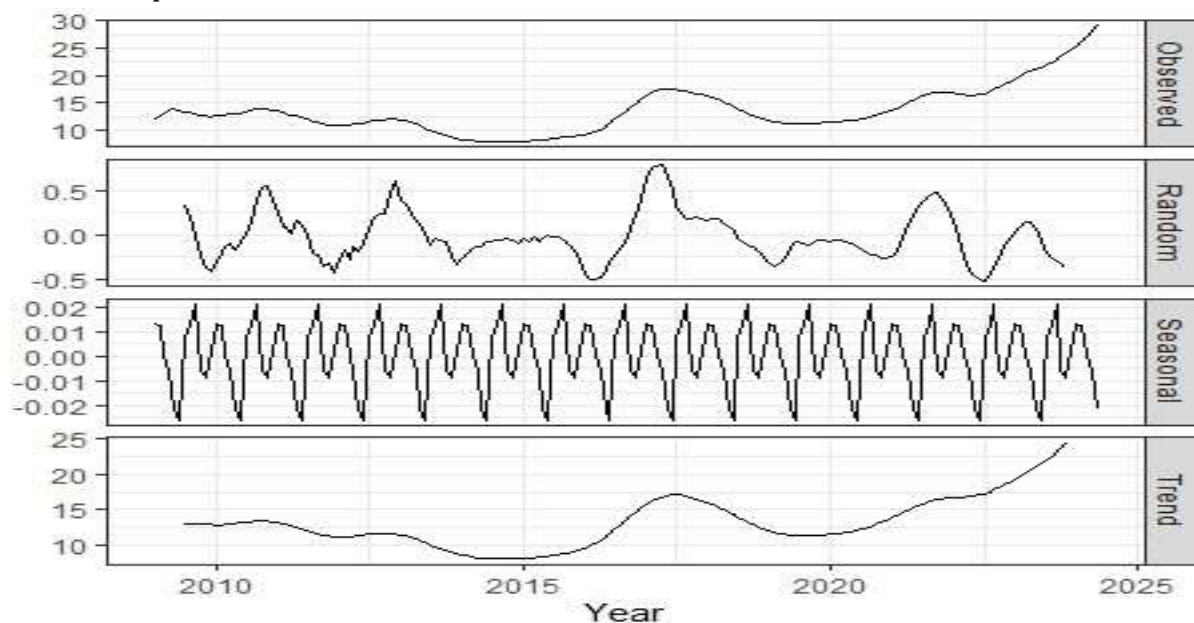


Figure 2: Time plot for Seasonality Check of the Inflation rates from 2009 to 2024

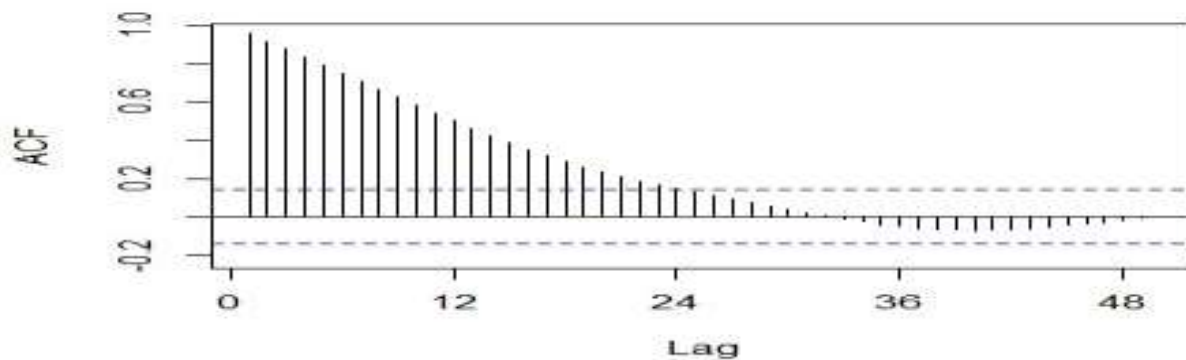


Figure 3: ACF Plots of Original Data on Inflation rates from 2009 to 2024

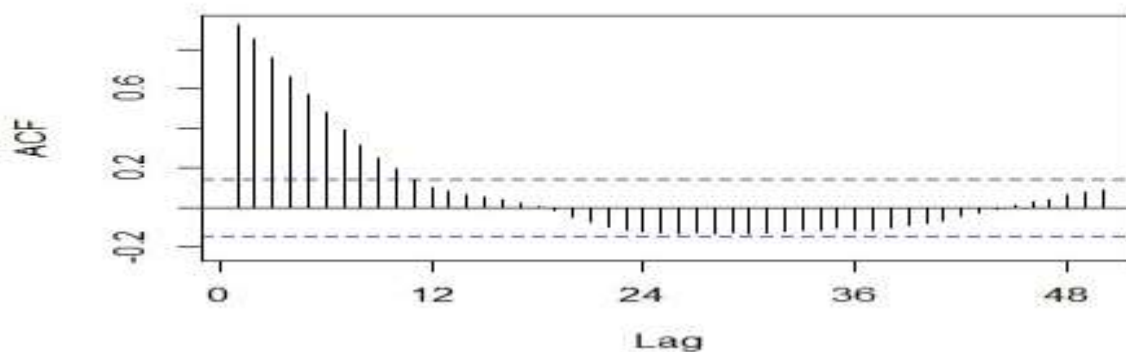


Figure 4: ACF Plots on the differenced Inflation rates from 2009 to 2024

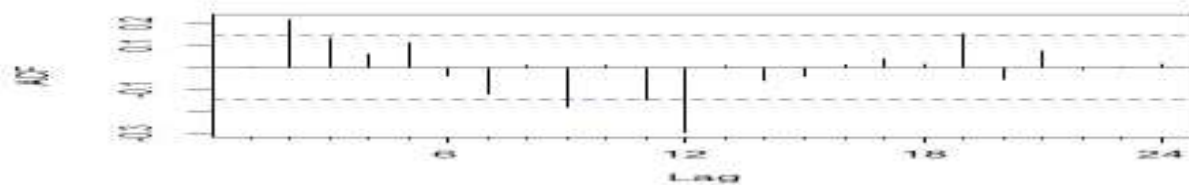


Figure 5: ACF Plots on the second differenced Inflation rates from 2009 to 2024

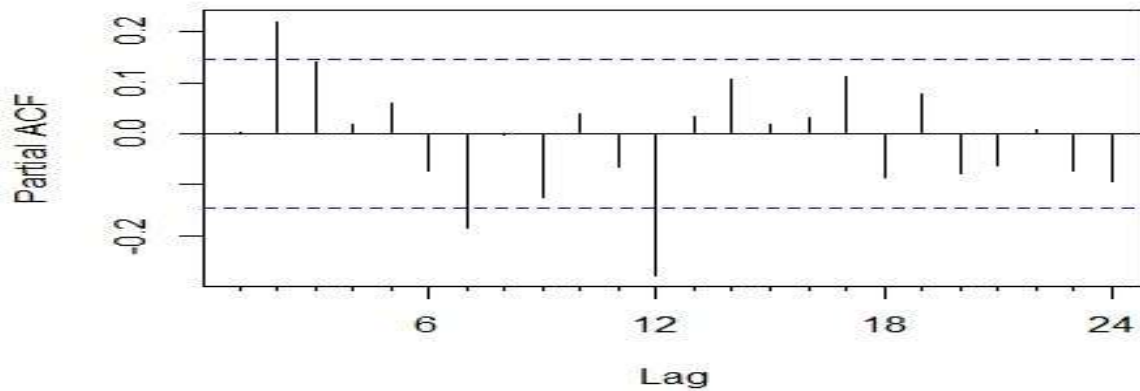


Figure 6: PACF Plots on the second differenced Inflation rates from 2009 to 2024

Table 2: Descriptive Statistics

Variable	Min	1st Qu	Median	Mean	3rd Qu	Max	Variance	std	skewness	Kurtosis
Inflation Rates	8.00	11.27	12.70	13.60	14.23	17.21	17.933	4.235	1.214	4.720

The results in table 2 shows the descriptive statistics. The mean is 13.60, variance is 17.933, standard deviation is 4.235, skewness 1.2139 and kurtosis is 4.720 respectively. This simply the Inflation Rates series is skewed to the right while the kurtosis seems to have flat tail.

Table 3: Augmented Dickey-Fuller Unit Root Test

VARIABLE	Augmented Dickey-Fuller Test						
Level	Lag	First Differ	Lag	Second Differ	Lag	RM order	order K
Inflation Rate	-	0.15163(0.99)	5	-	5	-	1(2)

Table 3 shows the results for the augmented dickey-fuller unit root test. The results show that the series is stationary at second difference with an estimated statistic of -4.6204(0.01)

Table 4. Tentative SARIMA Models

SARIMA Models	Log-Likelihood
ARIMA(2,2,2)(1,0,1)[12]	: -366.8428
ARIMA(0,2,0)	: -302.7963

ARIMA(1,2,0)(1,0,0)[12]	: -346.8572
ARIMA(0,2,1)(0,0,1)[12]	: -318.9219
ARIMA(2,2,2)(0,0,1)[12]	: -331.8437
ARIMA(2,2,2)(1,0,0)[12]	: -363.1251
ARIMA(2,2,2)(2,0,1)[12]	: -371.9795
ARIMA(2,2,2)(2,0,0)[12]	: -371.7642
ARIMA(2,2,2)(2,0,2)[12]	: -375.7874
ARIMA(2,2,2)(1,0,2)[12]	: -376.7413
ARIMA(2,2,2)(0,0,2)[12]	: -329.6811
ARIMA(1,2,2)(1,0,2)[12]	: -371.123
ARIMA(2,2,1)(1,0,2)[12]	: -378.7375
ARIMA(2,2,1)(0,0,2)[12]	: -331.4079
ARIMA(2,2,1)(1,0,1)[12]	: -368.6659
ARIMA(2,2,1)(2,0,2)[12]	: -377.9988
ARIMA(2,2,1)(0,0,1)[12]	: -333.5211
ARIMA(2,2,1)(2,0,1)[12]	: -374.0075
ARIMA(1,2,1)(1,0,2)[12]	: -364.7487
ARIMA(2,2,0)(1,0,2)[12]	: -372.2845
ARIMA(3,2,1)(1,0,2)[12]	: -379.598
ARIMA(3,2,1)(0,0,2)[12]	: -330.0165
ARIMA(3,2,1)(1,0,1)[12]	: -367.5792
ARIMA(3,2,1)(2,0,2)[12]	: -378.9257
ARIMA(3,2,1)(0,0,1)[12]	: -332.1485
ARIMA(3,2,1)(2,0,1)[12]	: -389.1511
ARIMA(3,2,1)(2,0,0)[12]	: Inf
ARIMA(3,2,1)(1,0,0)[12]	: -363.1209
ARIMA(3,2,0)(2,0,1)[12]	: -377.4537
ARIMA(4,2,1)(2,0,1)[12]	: -375.0066
ARIMA(3,2,2)(2,0,1)[12]	: -389.743
ARIMA(3,2,2)(1,0,1)[12]	: -365.771
ARIMA(3,2,2)(2,0,0)[12]	: Inf
ARIMA(3,2,2)(2,0,2)[12]	: Inf
ARIMA(3,2,2)(1,0,0)[12]	: -361.9868
ARIMA(3,2,2)(1,0,2)[12]	: -378.0651
ARIMA(4,2,2)(2,0,1)[12]	: -388.78
ARIMA(3,2,3)(2,0,1)[12]	: Inf
ARIMA(2,2,3)(2,0,1)[12]	: Inf
ARIMA(4,2,3)(2,0,1)[12]	: Inf

Re-fitting the best model(s) without approximations...

ARIMA(3,2,2)(2,0,1)[12] : Inf

ARIMA(3,2,1)(2,0,1)[12] : -338.6423

Best model:

ARIMA(3,2,1)(2,0,1)[12]

Residual Diagnostics

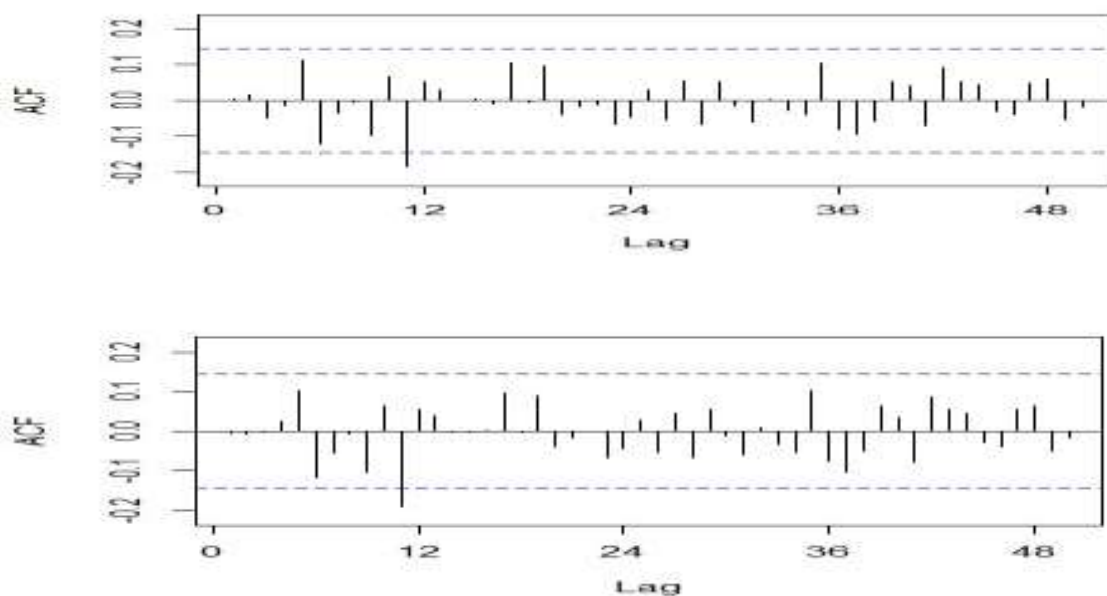


Figure 7: ACF Residual Plot

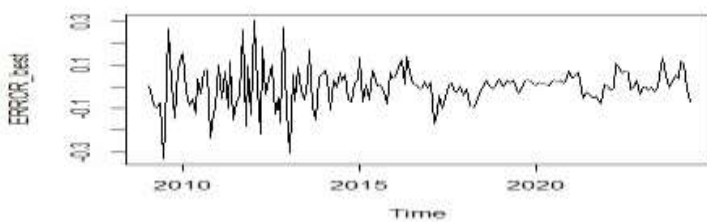


Figure 8: Error of the Best model Plot

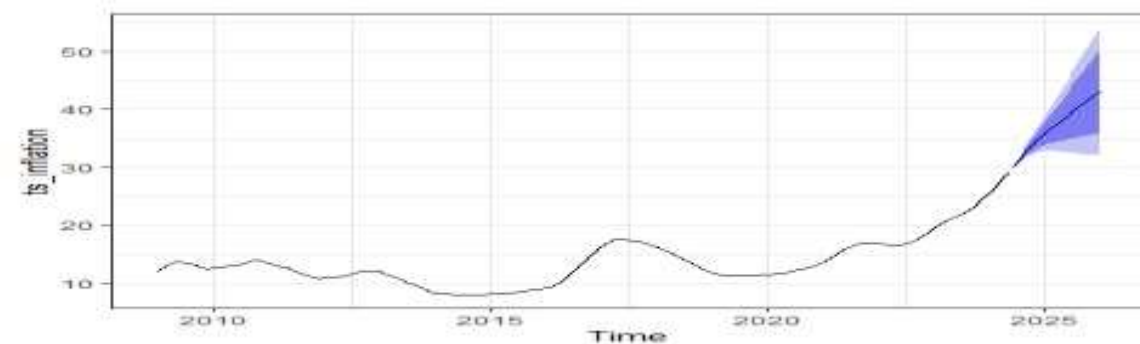


Figure 9: Forecast Plot

The results from re-fitting the best model(s) without approximations show that ARIMA (3,2,2) (2,0,1)[12] has an "Inf" value, which means the fitting process failed or had numerical problems. This may imply overfitting or an issue with the model's parameter limits. In contrast, the ARIMA (3,2,1) (2,0,1)[12] model produced a valid output with a log-likelihood of 177.73 and an AIC of 339.47, making it the best model compared to the others. Thus, this model is chosen as the best one. Additionally, the coefficients of the ARIMA(3,2,1)(2,0,1)[12] model show the connection between prior values of the series (the AR term -0.5440 implies that the first lag has a negative effect on the current value), past errors (the MA term 0.4427 indicates a positive relationship between the first lag of the error term and the series), and seasonal factors (SAR and SMA terms of -0.0156 and -0.4839 show the effect of seasonal components during the first seasonal period). The standard errors of these coefficients imply that the estimates are quite accurate. Nevertheless, the Model Selection indicates that ARIMA (2,2,3)(2,0,0)[12] and ARIMA(3,2,2)(2,0,0)[12] both had similar AIC, AICc, and BIC values, but neither surpassed the ARIMA(3,2,1)(2,0,1)[12] model, making this one the top choice. Furthermore, ARIMA(4,2,1)(2,0,0)[12] had a slightly poorer AIC of -337.58 and displayed higher MAE and RMSE values, thus making it a less attractive option.

For the model diagnostic tests, the Training set error measures include ME, RMSE, MAE, MAPE, ACF1. It was discovered that ARIMA (3,2,1)(2,0,1)[12] yielded the most favorable training set errors with the lowest ME (mean error), RMSE (root mean squared error), MAE (mean absolute error), and MAPE (mean absolute percentage error), indicating it fits the data most closely. Additionally, the ACF1 (first-order autocorrelation) values are low, suggesting that the model's residuals behave well and that the model has effectively captured the time series structure. In comparison, ARIMA (2,2,3)(2,0,0)[12] and ARIMA(3,2,2)(2,0,0)[12] also performed well in terms of error metrics but not as effectively as ARIMA(3,2,1)(2,0,1)[12]. Regarding the convergence of the best model, the log-likelihood value of 177.73 for ARIMA (3,2,1)(2,0,1)[12] confirms that the model explains the data effectively. The relatively low σ^2 of 0.008556 indicates that the residual variance is limited, suggesting a good fit to the inflation data.

Conclusion

The ARIMA models applied to the inflation time series data show that different model types give similar fit and error results. Among the models evaluated ARIMA(2,2,3)(2,0,0)[12], ARIMA(3,2,2)(2,0,0)[12], ARIMA(3,2,1)(2,0,1)[12], and ARIMA(4,2,1)(2,0,0)[12], the ARIMA(3,2,1)(2,0,1)[12] model had the lowest AIC, AIC and BIC values, showing it is the simplest model among those analyzed. Regarding training set error metrics, all models showed similar performance with low values for the mean error (ME), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). The ARIMA (3,2,1) (2,0,1) [12] model had the lowest RMSE and MAE, indicating it offers the most precise fit for the data provided. Moreover, its ACF1 value suggests that the residuals of this model display very low autocorrelation, which further confirms its suitability for the

inflation series. The ARIMA (3,2,1) (2,0,1)[12] model is the best at capturing the key inflation trends with little error and a sensible level of complexity. It can be regarded as the optimal model for predicting inflation in this situation

References

- Aidoo,E(2016) Modelling and Forecasting Inflation Rates in Ghana: An Application of Sarima Models. Thesis for the award of Masters in Applied Statistics, Örebro University School of Business
- Deebom Z, D, Essi , I, D and Amos E(2021), Evaluating Properties and Performance of Long Memory Models from an Emerging Foreign Markets Return Innovations. Asian Journal of Probability and Statistics 11(4): 1-23, 2021; 2582-0230
- Deebom, Z, D & Essi, I, D(2017) , Modeling Price Volatility of Nigerian Crude Oil Markets Using GARCH Model: 1987-2017. International Journal of Applied Science and Mathematical Theory 2489-009X ,3, 4 201
- Havi E.D.K. (2023). Modeling and Forecasting of Inflation in Ghana: SARIMA Approach. ESI Preprints. <https://doi.org/10.19044/esipreprint.8.2023.p512>
- Otu A, O,Osuji G, A., Opara J , Mbachu H, I , Iheagwara A(2014) application of SARIMA models in modelling and forecasting Nigeria's inflation rates
- Tuaneh, G. L. and Essi, I. D. (2017) Simultaneous Equation Modelling and Estimation of Consumption and Investment Functions in Nigeria. International Journal of Economics and Business Management. 3(8) 53-71
- Tuaneh, G. L (2018) Vector autoregressive modelling of the interaction among macroeconomic stability indicators in Nigeria (1981-2016). Asian Journal of Economics, Business and Accounting 9 (4), 1-17
- Tuaneh G. L and Okidim A. I. (2019). On agricultural performance amidst macroeconomic instability in Nigeria; Autoregressive distributed lagged modelling (2010Q1-2017Q4). Asian Journal of Economics, Business and Accounting 10 (2), 1-13
- Tuaneh, G. L., & Wiri, L. (2019). Unrestricted vector autoregressive modelling of the interaction among oil price, exchange rate and inflation rate (1981-2017). Asian Journal of Probability and Statistics. 2(4), 1-19

- Tuaneh, G. L., & Wiri, L. (2021). Markov switching mean vector autoregressive (MSM VAR) modelling of inflation rate and crude oil price interdependence in Nigeria. *Probability Statistics and Econometric Journal*, 3(6), 1-10.
- Tuaneh, G. L., Essi, I. D., & Etuk, E. H. (2021). Markov-Switching Vector Autoregressive (MSVAR) modelling (mean adjusted): Application to macroeconomic data. *Archives of Business Research*, 9(10), 261-274.
- Tuaneh, G. L., Essi, I. D., & Ozigbu, J. (2021). Dynamic linear interdependence between international trade and macroeconomic stability in Nigeria: A vector error correction modelling. *Archives of Business Research*.
- Tuaneh G. L and Doodei Week (2025). Innovative Statistical Modelling of Exchange Rates and Agricultural Sector Performance in Nigeria: The Markov-Switching Vector Autoregressive Modelling. *Asian Journal of Probability and Statistics*. 27(3), 119-133.