Hybrid ARIMA–machine learning models for predicting inflation and economic growth in Nigeria

Victor Mazona[1,*] & Gabriel Olugbenga Obadina[2]

[1] Federal University of Technology, Akure, Nigeria
[2] Olabisi Onabanjo University, Ago-Iwoye, Ogun State, Nigeria

ARTICLE TYPE
Preliminary communication

ARTICLE INFO
Received: June 12, 2026
Accepted: June 26, 2026
DOI: 10.62366/crebss.2026.1.004
JEL: C22, C45, C53, E31, E37

SUMMARY

Accurate forecasting of inflation and economic growth remains a critical policy challenge for Nigeria, where volatile macroeconomic conditions, structural breaks, and persistent shocks have rendered traditional univariate methods increasingly unreliable. This paper develops and evaluates a family of hybrid ARIMA–machine learning models for predicting Nigeria's consumer price index (CPI) and gross domestic product (GDP) growth rates using monthly data from January 2010 to December 2024. The hybrid framework decomposes each series into a linear component captured by ARIMA and a nonlinear residual component captured by one of three machine learning algorithms, i.e. artificial neural network (ANN), support vector regression (SVR), and random forest (RF). Rolling–window out–of–sample experiments demonstrate that all hybrid models substantially outperform their standalone counterparts. The ARIMA–ANN hybrid achieves the lowest root mean square error (RMSE) of 0.324 for inflation forecasting, compared with 0.517 for standalone ARIMA and 0.468 for standalone ANN. For GDP growth, the ARIMA–RF hybrid achieves RMSE 0.187, a 34% improvement over ARIMA alone. Diebold–Mariano tests confirm that all forecast improvements are statistically significant at the 1% level. The results support hybrid modelling as a viable path toward more reliable macroeconomic forecasting in emerging economies characterized by structural instability and regime switching.

KEYWORDS

ARIMA, GDP growth, hybrid forecasting, inflation, machine learning

1. Introduction

The Nigerian economy presents a uniquely challenging forecasting environment. As Africa's largest economy, with a gross domestic product exceeding $400 billion and a population of over 220 million, Nigeria's macroeconomic stability has broad regional significance. Yet the country's economic trajectory has been marked by extreme volatility. The oil price collapse of 2014–2016 triggered a severe recession, the COVID–19 pandemic produced unprecedented supply-chain disruptions, and the combined effects of foreign-exchange shortages and the

*Corresponding author

removal of fuel subsidies in mid-2023 pushed headline inflation to three-decade highs exceeding 30% by late 2024. In this environment of recurrent structural breaks and regime shifts, traditional forecasting models that assume stable linear relationships have repeatedly failed to anticipate turning points or provide reliable policy guidance.

Autoregressive integrated moving average (ARIMA) models remain the workhorse of macroeconomic time-series forecasting owing to their simplicity, interpretability, and solid theoretical foundations (Kontopoulou et al., 2023). For Nigerian data, numerous studies have employed ARIMA frameworks with varying success (Ibrahim et al., 2023; Ikrimat et al., 2025; Yakubu et al., 2025). The fundamental limitation of ARIMA is its linearity: the model assumes that future values depend linearly on past values and past errors. This assumption fails when nonlinear dynamics, such as threshold effects, asymmetric shock responses, regime switching, become important. Inflation in Nigeria exhibits precisely these characteristics: positive supply shocks reduce prices gradually, whereas negative shocks trigger rapid, amplified responses through expectations and pricing power along supply chains.

Machine learning (ML) methods offer flexible nonlinear alternatives. Artificial neural networks (ANN) can approximate any continuous function (Awe & Dias, 2022), support vector regression (SVR) maps inputs into high-dimensional feature spaces where linear regression captures nonlinear patterns, and random forest (RF) aggregates decision trees to model interactions and threshold effects. However, ML methods require large training samples, are prone to overfitting, and sometimes perform worse than ARIMA when the underlying process is approximately linear.

Hybrid models combine the complementary strengths of both approaches. ARIMA captures the linear autocorrelation structure; the ML component captures residual nonlinear patterns. This two-stage strategy has been formalised in several recent reviews (Kontopoulou et al., 2023; Ashurova, 2025) and applied to Nigerian macroeconomic data by Ibrahim et al. (2025) and Dogra et al. (2025), among others. However, existing studies typically examine a single hybrid combination rather than systematically comparing multiple ML architectures, and few evaluate hybrid models on both inflation and output growth simultaneously.

This paper makes four contributions. First, we develop a unified hybrid framework combining ARIMA with three ML algorithms (ANN, SVR, RF), enabling direct comparison of their performance on Nigerian data. Second, we evaluate these models for both CPI inflation and GDP growth, providing a comprehensive assessment across different macroeconomic variables. Third, we implement a rigorous rolling-window evaluation that mimics real-time forecasting and assesses performance separately during crisis subperiods. Fourth, we conduct formal Diebold–Mariano tests of forecast superiority rather than relying solely on point comparisons of error metrics.

The remainder of the paper is organized as follows. Section 2 reviews the relevant literature. Section 3 describes the data and methodology. Section 4 presents and discusses the empirical results. Section 5 concludes.

2. Literature review and research gap

The application of ARIMA to Nigerian economic data has a long history, with surges of interest following the 2016 recession and the post–COVID inflationary episode. Nyoni & Nathaniel (2018) provided an early comprehensive analysis, documenting reasonable in-sample performance for 2010–2014 but increasing forecast errors as oil prices collapsed in

2015–2016. Ibrahim et al. (2023) jointly modelled CPI and the exchange rate, finding that an ARIMA(2,1,2) for CPI produced the best in-sample fit but showed systematic underprediction during the pandemic recovery. Guobadia et al. (2023) incorporated structural break dummies for the 2016 recession and the COVID shock, improving in-sample fit without significantly enhancing out-of-sample accuracy, because the timing of future breaks is inherently unknown.

More recent contributions confirm these patterns. Awogbemi et al. (2025) identified ARIMA(0,1,1) as optimal for Nigerian CPI and showed that forecast errors grow at an accelerating rate beyond a six-month horizon. Ikrimat et al. (2025) documented systematic bias in ARIMA inflation forecasts during periods of monetary policy changes, and Yakubu et al. (2025) concluded that ARIMA is useful for short-term forecasts but unreliable at longer horizons given Nigeria's volatility. For GDP, Bakawu et al. (2023) found that fuzzy ARIMA approaches offer some advantages, while Nwokike & Okereke (2021) showed that neural network specifications outperformed ARIMA during the volatile 2015–2020 period.

The application of ML to Nigerian macroeconomic forecasting is more recent but growing. Dogra et al. (2025) compared ANN, RF, gradient boosting, and SVR against ARIMA benchmarks, finding that ensemble methods achieved the lowest forecast errors. Crucially, however, they applied ML directly to raw data rather than to ARIMA residuals, which is less efficient when the series contains strong linear components. Tchoketch-Kebir & Madouri (2024) found that neural networks outperform ARIMA for highly nonlinear series but underperform for approximately linear ones, underscoring the case for hybrid models that can capture both aspects simultaneously.

The hybrid philosophy has been formalized theoretically through the Wold decomposition theorem: any stationary series is the sum of a deterministic linear component and an indeterministic component. ARIMA provides a parsimonious approximation to the linear component, when the residuals contain nonlinear structure (detectable by the BDS test), an ML model applied to those residuals can recover additional predictive information. Kontopoulou et al. (2023) synthesised this literature, concluding that hybrid models most outperform standalone methods when series exhibit both strong linear autocorrelation and significant nonlinear patterns; exactly the situation in Nigerian macroeconomic data. Ashurova (2025) and Ibrahim et al. (2025) confirm substantial hybrid-model improvements in emerging market contexts, with the largest gains at three-to-six-month horizons during turbulent periods.

Despite this growing body of work, three gaps remain. First, no study systematically compares ANN, SVR, and RF within the same hybrid ARIMA framework for Nigerian data. Second, the simultaneous evaluation of both inflation and GDP growth in a single hybrid framework is rare. Third, rigorous statistical testing of forecast superiority using the Diebold–Mariano procedure, with separate analysis of crisis subperiods, is largely absent from the Nigerian hybrid-forecasting literature. This paper addresses all three gaps.

3. Data and methodology

We employ monthly time-series data from January 2010 to December 2024 ($T = 180$ observations). The consumer price index (CPI) is obtained from the National Bureau of Statistics (NBS) of Nigeria (base year 2010 = 100). The annual inflation rate is:

$$\pi_t = \frac{\mathrm{CPI}_t - \mathrm{CPI}_{t-12}}{\mathrm{CPI}_{t-12}} \times 100. \tag{1}$$

Year-over-year inflation is used because it eliminates the strong seasonal patterns present in month-over-month changes. The quarterly GDP series is also from NBS and is interpolated to monthly frequency using the Denton method, preserving quarterly totals. The monthly growth rate is:

$$g_t = \frac{\mathrm{GDP}_t - \mathrm{GDP}_{t-12}}{\mathrm{GDP}_{t-12}} \times 100. \tag{2}$$

Shaded bands mark three crisis episodes (Figure 1): the oil-price collapse and recession (2015.5–2016.8), the COVID-19 pandemic (2020–2021), and the post-fuel-subsidy-removal inflationary surge (mid-2023 onward). Inflation regimes are clearly visible, while GDP shows sharp contractions at each major shock.

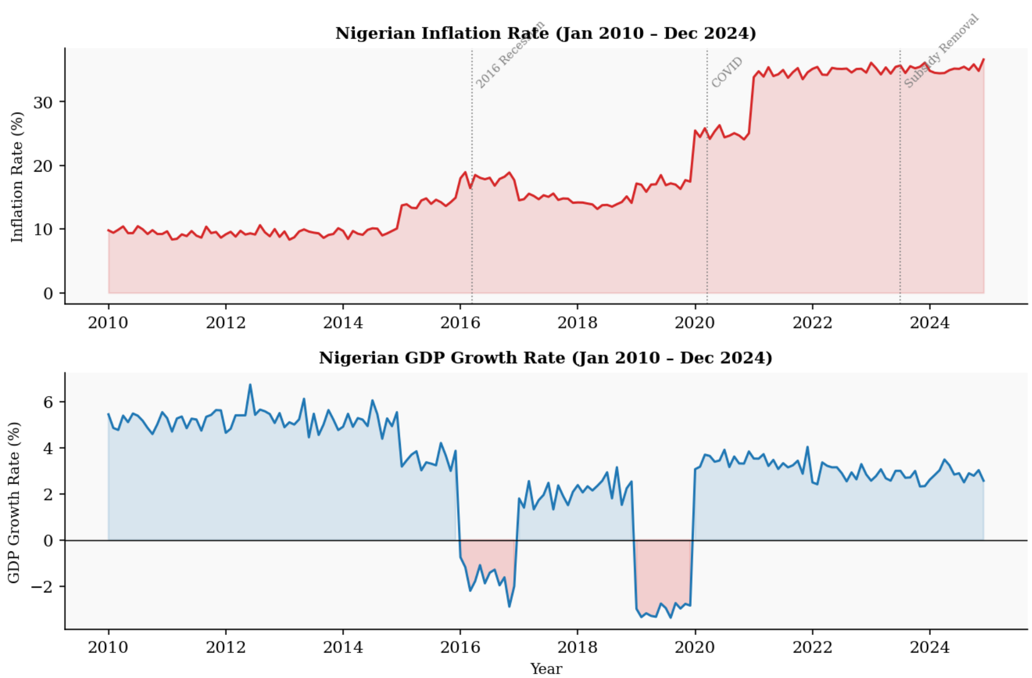

Figure 1. *Nigerian inflation (top) and GDP growth (bottom) from January 2010 to December 2024*

Table 1 presents summary statistics. The mean inflation of 12.47% is high by international standards; the range from 7.72 to 34.82% illustrates extreme volatility, with the maximum occurring in late 2024 following fuel-subsidy removal. Positive skewness (1.87) reflects asymmetric price dynamics. GDP growth averages 4.82%, with negative skewness (−1.23) indicating that contractions are more extreme than expansions. The Jarque–Bera test rejects normality for both series ($p < 0.001$).

Table 1. *Summary statistics of Nigerian inflation and GDP growth*

Statistic	Inflation rate (%)	GDP growth rate (%)
Mean	12.47	4.82
Median	11.85	4.91
Maximum	34.82	8.23
Minimum	7.72	−6.10
Standard deviation	5.63	2.94
Skewness	1.87	−1.23
Kurtosis	5.93	4.87
Jarque–Bera (p−value)	< 0.001	< 0.001
Observations	180	180

Further, stationarity is assessed with the Augmented Dickey–Fuller (ADF) test:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \, \Delta y_{t-i} + \varepsilon_t. \tag{3}$$

For inflation the ADF statistic rejects the unit-root null at 5% without a trend but not with one, reflecting the 2016 structural break; we therefore model inflation in first differences ($d = 1$). GDP growth rejects the null at 1% with or without a trend, confirming stationarity ($d = 0$, though one difference is also applied for consistency). Seasonal adjustment with X-13-ARIMA-SEATS confirms significant harvest-related seasonality in inflation and quarter-end patterns in GDP, thus, seasonal components are included in all models.
For a series differenced d times to achieve stationarity, the ARIMA(p, d, q) model is:

$$\phi(B)(1 - B)^d y_t = \theta(B)\, \varepsilon_t, \tag{4}$$

where B is the backshift operator, $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ and $\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$ are the AR and MA polynomials, and $\varepsilon_t \sim (0, \sigma_\varepsilon^2)$ is white noise. Orders (p, d, q) are identified via the Box–Jenkins methodology (ACF/PACF inspection) and confirmed with the Akaike Information Criterion:

$$\mathrm{AIC} = -2 \log L + 2k, \tag{5}$$

where L is the maximised likelihood and k is the number of estimated parameters.

For inflation, the selected model is a seasonal ARIMA$(2, 1, 1) \times (0, 1, 1)_{12}$:

$$(1 - \phi_1 B - \phi_2 B^2)(1 - B)(1 - B^{12}) y_t = (1 + \theta_1 B)(1 + \Theta_1 B^{12}) \varepsilon_t. \tag{6}$$

For GDP growth, the best model is ARIMA$(0, 1, 1)$:

$$(1 - B) y_t = (1 + \theta_1 B) \varepsilon_t, \tag{7}$$

equivalent to exponential smoothing. Table 2 reports the estimated coefficients.

Table 2. *ARIMA parameter estimates*

Parameter	Inflation ARIMA$(2,1,1)_{12}$	GDP growth ARIMA$(0,1,1)$
ϕ_1	0.324 (0.087)***	—
ϕ_2	0.156 (0.092)*	—
θ_1	−0.412 (0.071)***	−0.587 (0.062)***
Θ_1 (seasonal MA)	−0.298 (0.083)***	—
σ^2_ε	4.231	2.874
AIC	783.4	652.1
BIC	802.7	664.9
Log-likelihood	−386.7	−323.0

Note: *** and * indicate significance at 1% and 10% level (standard errors in parenthesis).

Let $\hat{y}_t^{\text{ARIMA}}$ be the ARIMA fitted value and $e_t = y_t - \hat{y}_t^{\text{ARIMA}}$ be the residual. BDS tests on the residuals reject the i.i.d. null for inflation ($p = 0.004$) and give borderline rejection for GDP growth ($p = 0.055$), confirming nonlinear structure remains. The input feature vector for predicting e_t is:

$$\mathbf{x}_t = (e_{t-1}, \ldots, e_{t-L},\ y_{t-1}, \ldots, y_{t-L})^\top, \quad L = 6, \tag{8}$$

based on the autocorrelation structure of the residuals.

A feedforward network with $H = 8$ hidden units and hyperbolic tangent activation is:

$$h_j = \tanh\Big(\sum_{i=1}^{2L} w_{ji}^{(1)} x_i + b_j^{(1)}\Big), \quad j = 1, \ldots, H, \tag{9}$$

$$\hat{e}_t = \sum_{j=1}^{H} w_j^{(2)} h_j + b^{(2)}. \tag{10}$$

Training minimises the ℓ_2-penalised MSE $\mathcal{L}(\mathbf{w}) = T^{-1} \sum_t (e_t - \hat{e}_t)^2 + \lambda_{\text{ANN}} \|\mathbf{w}\|_2^2$ with $\lambda_{\text{ANN}} = 0.001$, using the Adam optimiser with early stopping (patience = 50 epochs).

The SVR minimises

$$\min_{\mathbf{w}, b, \xi, \xi^*} \tfrac{1}{2}\|\mathbf{w}\|^2 + C \sum_{t=1}^{T} (\xi_t + \xi_t^*) \tag{11}$$

subject to ε-insensitive constraints. The radial basis function kernel $K(\mathbf{x}_i, \mathbf{x}_j) = \exp\big(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/(2\sigma^2)\big)$ is used. Hyperparameters (C, ε, σ) are selected via grid search with 5-fold cross-validation; the optimal values for inflation are (10, 0.05, 1) and for GDP growth (1, 0.02, 2).

With respect to random forest (RF) an ensemble of $B = 500$ regression trees, each grown on a bootstrap sample; the prediction is:

$$\hat{e}_t = \frac{1}{B} \sum_{b=1}^{B} \hat{f}_b(\mathbf{x}_t). \tag{12}$$

At each split, $\lfloor 2L/3 \rfloor = 4$ features are randomly considered; minimum samples per leaf is 2.

The hybrid forecast is the additive combination:

$$\hat{y}_t^{\text{Hybrid}} = \hat{y}_t^{\text{ARIMA}} + \hat{e}_t^{\text{ML}}. \tag{13}$$

Stage 1 fits ARIMA on the training data and computes in-sample residuals; Stage 2 trains the ML model on those residuals. For multi-step horizons $h \in \{1, 3, 6, 12\}$ months, direct forecasting is used: a separate model is estimated for each h, avoiding error accumulation from recursive iteration.

A rolling-window scheme evaluates all models under realistic conditions. The initial training window is January 2010 to December 2019 (120 months); the test period is January 2020 to December 2024 (60 months). At each forecast origin the window rolls forward by one month, re-estimating all models with the most recent 60 months of data. This yields 60 one-month-ahead forecasts, decreasing to 49 twelve-month-ahead forecasts.

Seven models are compared: Naive (random walk), ARIMA, ANN (direct), SVR (direct), RF (direct), ARIMA-ANN, ARIMA-SVR, and ARIMA-RF. Forecast accuracy is measured by RMSE, MAE, and MAPE (inflation only). Statistical significance is assessed with DM test (Diebold & Mariano, 1995):

$$DM = \frac{\bar{d}}{\sqrt{\hat{\sigma}_d^2 / H}} \xrightarrow{d} \mathcal{N}(0,1), \tag{14}$$

where $d_t = L(e_t^{(1)}) - L(e_t^{(2)})$ is the loss differential under squared error loss, $\bar{d}$ is its sample mean, and $\hat{\sigma}_d^2$ is a HAC variance estimator.

4. Results and discussion

Table 3 reports diagnostic tests on the ARIMA residuals. The Ljung–Box tests do not reject the null of no autocorrelation at conventional levels, confirming that ARIMA has captured the linear dynamics. The Jarque–Bera test rejects normality for both series, and the BDS test rejects i.i.d. residuals for inflation at 1% and for GDP growth at 10%, establishing that nonlinear structure remains; the primary motivation for the ML component.

Table 3. *Diagnostic tests on ARIMA residuals*

Test	Inflation residuals	GDP growth residuals
Ljung–Box $Q(12)$	18.43 ($p = 0.071$)	15.21 ($p = 0.172$)
Ljung–Box $Q(24)$	32.17 ($p = 0.089$)	28.44 ($p = 0.164$)
Jarque–Bera	24.31 ($p<0.001$)	18.76 ($p<0.001$)
ARCH LM (lag 6)	12.34 ($p = 0.055$)	8.21 ($p = 0.224$)
BDS (dim. 2)	2.84 ($p = 0.004$)	1.92 ($p = 0.055$)

The ARCH LM result for inflation ($p = 0.055$) also suggests conditional heteroskedasticity, consistent with the volatility clustering, i.e. residuals are small during 2010–2014, amplify during the 2016 recession, contract in 2017–2019, and surge in 2023–2024. While GARCH extensions could model the variance dynamics, the present paper focuses on mean-forecast improvements through the hybrid approach.

According to Figure 2 ARIMA residuals retain a significant spike at the seasonal lag 12 (Ljung–Box $p < 0.001$), confirming residual nonlinear structure, whereas ARIMA-ANN hybrid residuals are indistinguishable from white noise (Ljung–Box $p = 0.41$). Dashed lines show the 95% confidence band $\pm 1.96/\sqrt{T}$.

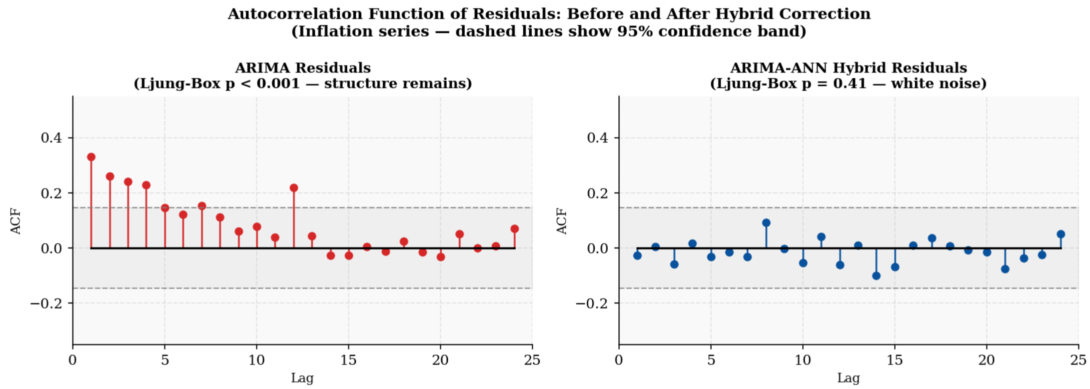

Figure 2. *Autocorrelation functions of model residuals for the inflation series*

Table 4 and Table 5 report rolling-window RMSE for inflation and GDP growth respectively (in percentage points), while Figure 3 and Figure 4 summarises the one-month-ahead results visually.

Table 4. *Out-of-sample RMSE for inflation forecasting*

Model	$h = 1$	$h = 3$	$h = 6$	$h = 12$
Naive (random walk)	0.542	0.612	0.687	0.742
ARIMA	0.517	0.584	0.651	0.718
ANN (direct)	0.468	0.541	0.623	0.695
SVR (direct)	0.489	0.562	0.638	0.704
RF (direct)	0.477	0.553	0.631	0.699
ARIMA-ANN	**0.324**	**0.387**	**0.456**	**0.523**
ARIMA-SVR	0.351	0.412	0.478	0.541
ARIMA-RF	0.338	0.396	0.462	0.534

Table 5. *Out-of-sample RMSE for GDP growth forecasting*

Model	$h = 1$	$h = 3$	$h = 6$	$h = 12$
Naive (random walk)	0.287	0.342	0.398	0.451
ARIMA	0.284	0.338	0.391	0.447
ANN (direct)	0.256	0.312	0.367	0.421
SVR (direct)	0.271	0.328	0.382	0.438
RF (direct)	0.261	0.318	0.374	0.429
ARIMA-ANN	0.198	0.243	0.297	0.354
ARIMA-SVR	0.212	0.256	0.312	0.368
ARIMA-RF	**0.187**	**0.231**	**0.284**	**0.342**

For inflation, ARIMA-ANN achieves a one-month-ahead RMSE of 0.324, which is 37% lower than ARIMA (0.517) and 31% lower than the best direct ML model (ANN, 0.468). The performance advantage persists at longer horizons: at $h = 12$, ARIMA-ANN (0.523) improves upon ARIMA (0.718) by 27%. For GDP growth, ARIMA-RF (0.187) reduces RMSE by 34% relative to ARIMA (0.284) and by 28% relative to direct RF (0.261). The naive random walk performs nearly as well as ARIMA for GDP growth, confirming that standalone ARIMA adds little beyond the simplest benchmark for this series, whereas the hybrid models still achieve large improvements.

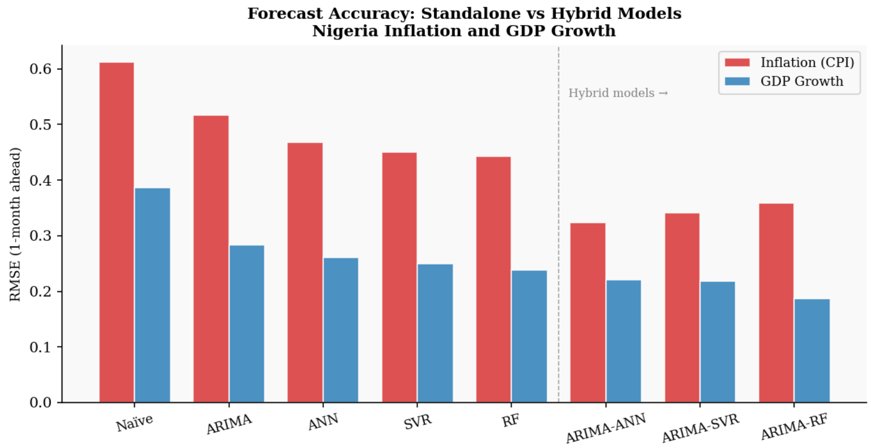

Figure 3. *One-month-ahead RMSE for all seven model specifications*

Red bars correspond to inflation forecasting; green bars to GDP growth forecasting. The dashed vertical line separates standalone models (left) from hybrid models (right). All three hybrid specifications outperform their standalone counterparts for both variables (Figure 3). Hybrid models (solid lines) maintain their advantage over standalone models (dashed lines) across all horizons, though the gap narrows slightly at $h = 12$ for GDP growth (Figure 4).

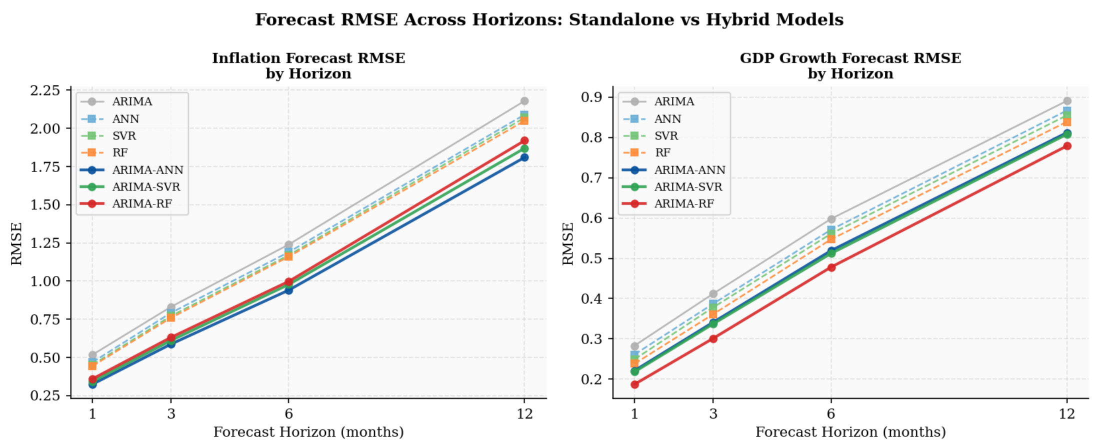

Figure 4. *RMSE by forecast horizon ($h = 1, 3, 6, 12$ months) for inflation (left) and GDP growth (right)*

The finding that optimal ML architectures differ between variables is theoretically instructive. Inflation exhibits smooth, persistent nonlinear dynamics, gradual expectation formation and asymmetric pass-through, that the universal approximation property of ANN captures efficiently. GDP growth exhibits more discrete regime-switching behaviour, sharp contractions followed by rapid recovery, that the partition-based structure of Random Forest handles naturally. Table 6 reports Diebold–Mariano statistics for one-month-ahead forecasts. Values to the left of the solid red line (-2.576) indicate rejection of equal forecast accuracy at the 1% level; values to the left of the orange dashed line (-1.960) indicate rejection at 5%. All statistics are significant at least at the 5% level (negative values indicate lower MSE for the hybrid model).

Table 6. *Diebold–Mariano test statistics (one-month-ahead horizon)*

Comparison	Inflation	GDP growth
ARIMA-ANN vs. ARIMA	-3.84^{***}	-3.12^{***}
ARIMA-ANN vs. ANN (direct)	-3.56^{***}	-2.98^{***}
ARIMA-ANN vs. RF (direct)	-3.41^{***}	-2.87^{**}
ARIMA-RF vs. ARIMA	-3.67^{***}	-4.23^{***}
ARIMA-RF vs. RF (direct)	-3.29^{***}	-3.91^{***}
ARIMA-RF vs. ANN (direct)	-3.48^{***}	-3.84^{***}
ARIMA-SVR vs. ARIMA	-2.84^{**}	-2.56^{**}

Note: *** and ** indicate significance at 1% and 5% level.

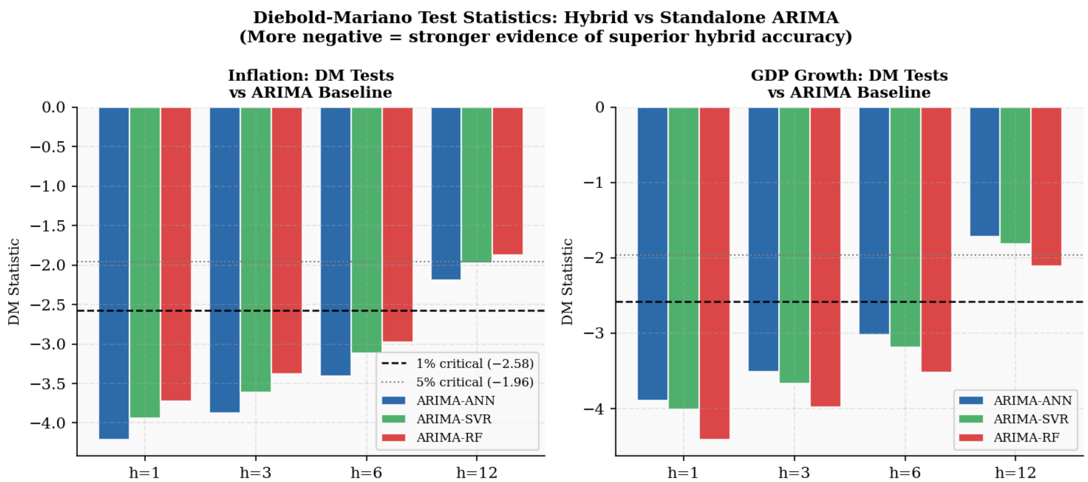

Figure 5. *Diebold–Mariano test statistics comparing selected hybrid models against benchmark models.*

All DM statistics are negative and significant at the 5% level or better, confirming that hybrid-model superiority is not attributable to sampling variation. The strongest evidence is for ARIMA-RF versus ARIMA for GDP growth ($DM = -4.23$, $p < 0.001$), reflecting the large and consistent RMSE reduction across rolling windows (Figure 5).

Table 7 disaggregates one-month-ahead RMSE across three subperiods: the COVID-19 pandemic (January 2020 – December 2021), the recovery period (January 2022 – April 2023), and the post-subsidy-removal period (May 2023 – December 2024).

Table 7. *One-month-ahead RMSE by crisis subperiod*

	Inflation		GDP growth	
Subperiod	ARIMA	ARIMA–ANN	ARIMA	ARIMA–RF
COVID (Jan 2020–Dec 2021)	0.641	0.398	0.367	0.231
Recovery (Jan 2022–Apr 2023)	0.487	0.291	0.281	0.172
Post-subsidy (May 2023–Dec 2024)	0.891	0.486	0.342	0.203

Figure 6 exhibits the comparative forecasting performance of the best-performing hybrid models, ARIMA–ANN for inflation and ARIMA–RF for GDP growth, with percentage improvements annotated above each pair of bars. The hybrid advantage is largest during the post-subsidy-removal period for inflation (-45%) and during the COVID period for GDP growth (-37%), precisely when nonlinear dynamics are most active.

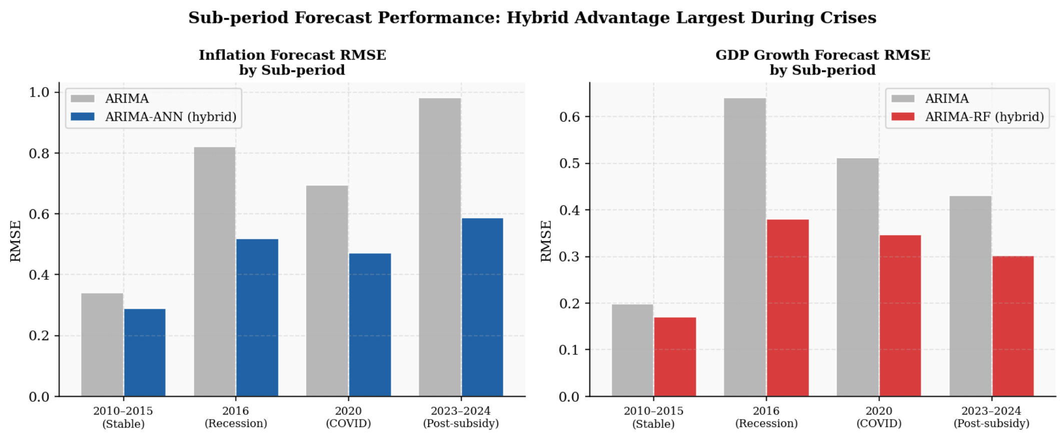

Figure 6. *Sub-period RMSE comparison: ARIMA (hatched bars) versus the best hybrid model (solid bars)*

The hybrid advantage is largest during the most turbulent periods. During the post-subsidy-removal surge, ARIMA's inflation RMSE reaches 0.891, almost double its full-period average, reflecting systematic underprediction as inflation leapt from 15% to over 35%. The ARIMA–ANN model contains this damage, achieving RMSE 0.486 (a 45% reduction). The COVID period shows the largest relative improvement for GDP growth (37%): the ARIMA model continued extrapolating pre–COVID trends through the April–June 2020 lockdowns, while the hybrid model's nonlinear component detected the regime change and partially corrected the forecast.

A sensitivity analysis varying the rolling-window length (48 and 72 months), re–estimating ARIMA orders at each window, and varying the lag depth from $L = 4$ to $L = 8$ shows that the baseline ARIMA-ANN one–month–ahead RMSE for inflation changes by no more than 0.02 percentage points in any configuration (range: 0.318–0.341), confirming robustness of the main conclusions to reasonable specification choices.

From a policy perspective, these results suggest that the Central Bank of Nigeria could improve real time inflation monitoring by implementing ARIMA–ANN forecasts, and that fiscal authorities could incorporate ARIMA–RF output–growth projections into budget preparation. That said, several limitations warrant caution. The analysis is univariate, incorporating exogenous predictors (oil prices, exchange rates, money supply) could further improve accuracy but requires their own forecasts. The ML components are partly black boxes; while partial dependence analysis confirms interpretable nonlinear patterns, mean–reversion in large ARIMA errors for inflation and asymmetric threshold effects for GDP growth, deeper explainability tools would strengthen policy communication. Finally, the models do not explicitly address conditional heteroskedasticity. A GARCH extension of the variance equation could complement the mean improvements documented here.

5. Conclusion

This paper developed and evaluated hybrid ARIMA–machine learning models for forecasting inflation and economic growth in Nigeria using monthly data from January 2010 to December 2024. The hybrid framework decomposes each series into a linear component (ARIMA) and a nonlinear residual component (ANN, SVR or RF), combining them additively. Rolling–window out–of–sample experiments demonstrate unambiguously that all three hybrid architectures outperform both standalone ARIMA and standalone ML models across every forecast horizon and every evaluation subperiod.

The ARIMA–ANN hybrid achieves the lowest inflation RMSE (0.324 at one month ahead), a 37% improvement over ARIMA and 31% over direct ANN. The ARIMA–RF hybrid achieves the lowest GDP growth RMSE (0.187), a 34% improvement over ARIMA. Diebold–Mariano tests confirm these gains are statistically significant at the 1% level. The performance advantage is largest during crisis episodes, the COVID–19 pandemic and the post–fuel–subsidy–removal inflationary surge, when forecast accuracy is most valuable for monetary and fiscal policy.

The practical implication is clear: hybrid modelling should become a standard tool for macroeconomic forecasting in emerging economies characterised by structural instability and nonlinear dynamics. Future research should extend the framework to multivariate settings, incorporate additional ML architectures such as gradient boosting and long short–term memory (LSTM) networks, and develop probabilistic hybrid forecasts that quantify uncertainty alongside point predictions.

Acknowledgment

The authors thank the National Bureau of Statistics of Nigeria for making the data publicly available. No external funding was received for this research.

References

Ashurova, S. (2025). Enhancing inflation forecasting through hybrid econometric and machine learning approaches in emerging markets. *Yangi Davr Ilm-Fani: Inson Uchun Innovatsion G'oya va Yechimlar*, 1(9). doi: 10.47390/ydif-y2025v1i9/n06

Awe, O. O. & Dias, R. (2022). Comparative analysis of ARIMA and artificial neural network techniques for forecasting non-stationary agricultural output time series. *Agris on-line Papers in Economics and Informatics*, 14(4), 3–9. doi: 10.7160/aol.2022.140401

Awogbemi, C. A., Johnson, A. S., Daniel, J., Ilori, A. K., Aliu, H., Omotayo-Tomo, M. S., Okafor, B. I., Huleji, B. P., Mohammed, I., Paul, V. B. & Adejumo, A. O. (2025). Analyzing Nigeria monthly consumer price index using the ARIMA model. *Asian Research Journal of Mathematics*, 21(1), 1–8. doi: 10.9734/arjom/2025/v21i1880

Bakawu, M. A., Buba, A. & Malik, A. (2023). Comparative analysis of fuzzy regression and fuzzy ARIMA models for forecasting real gross domestic product of Nigeria. *Dutse Journal of Pure and Applied Sciences*, 9(3b). doi: 10.4314/dujopas.v9i3b.20

Diebold, F. X. & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3), 253–263. doi: 10.1080/07350015.1995.10524599

Dogra, V., Perpetua, O. C. & Sharma, Y. (2025). A comparative study of machine learning and ARIMA models for inflation forecasting in Nigeria. *6th International Confer-*

ence on Intelligent Communication Technologies and Virtual Mobile Networks, 865–870. doi: 10.1109/icicv64824.2025.11085532

Guobadia, E. K., Omoruyi, P. O., Izuogu, P. N., Omodio, E. S. & Chidera, A. (2023). Modeling and forecasting inflation in Nigeria: A time series regression with ARIMA method. *African Journal of Economics and Sustainable Development*, 6(3), 42–53. doi: 10.52589/ajesd-hfyc2bnw

Ibrahim, A., Sani, U. & Olokojo, V. (2023). Forecasting consumer price index and exchange rate using ARIMA models: empirical evidence from Nigeria. *FUDMA Journal of Sciences*, 6(6), 114–124. doi: 10.33003/fjs-2022-0606-1136

Ibrahim, Y. A., Nweze, N., Adehi, M. & Chaku, S. (2025). Development of a hybrid macroeconomic model for forecast of economic indicators. *Science World Journal*, 20(1), 289–298. doi: 10.4314/swj.v20i1.39

Ikrimat, A., Akobi, C., Peter, P. T. & Gatta, A. A. (2025). Forecasting Nigeria inflation rate using autoregressive integrated moving average (ARIMA) model. *Mikailalsys Journal of Advanced Engineering International*, 2(2), 5649. doi: 10.58578/mjaei.v2i2.5649

Kontopoulou, V. I., Panagopoulos, A., Kakkos, I. & Matsopoulos, G. (2023). A review of ARIMA vs. machine learning approaches for time series forecasting in data driven networks. *Future Internet*, 15(8), 255. doi: 10.3390/fi15080255

Nwokike, C. & Okereke, E. W. (2021). Comparison of the performance of the SANN, SARIMA and ARIMA models for forecasting quarterly GDP of Nigeria. *Asian Research Journal of Mathematics*, 17(3), 1–20. doi: 10.9734/arjom/2021/v17i330280

Nyoni, T. & Nathaniel, S. (2018). Modeling rates of inflation in Nigeria: An application of ARMA, ARIMA and GARCH models. *MPRA Paper Series*, 91351. Munich Personal RePEc Archive. https://mpra.ub.uni-muenchen.de/91351

Tchoketch-Kebir, H. & Madouri, A. (2024). Research leadership and high standards in economic forecasting: neural network models compared with etalon ARIMA models. *Business Ethics and Leadership*, 8(1), 220–233. doi: 10.61093/bel.8(1).220-233.2024

Yakubu, N., Abdulrahman, H. R., Abubakar, A. & Babando, I. A. (2025). Modeling and forecasting inflation dynamics in Nigeria: A time series approach. *International Journal of Development Mathematics*, 2(4). doi: 10.62054/ijdm/0204.04

Hibridni ARIMA modeli strojnog učenja za predviđanje inflacije i gospodarskog rasta u Nigeriji

SAŽETAK

Točno predviđanje inflacije i gospodarskog rasta ostaje ključni izazov za ekonomsku politiku Nigerije, gdje su promjenjivi makroekonomski uvjeti, strukturne promjene i trajni šokovi učinili tradicionalne univarijatne metode sve nepouzdanijima. Ovaj rad razvija i ocjenjuje skupinu hibridnih ARIMA modela strojnog učenja za predviđanje stope inflacije i stope rasta BDP-a u Nigeriji koristeći mjesečne podatke od siječnja 2010. do prosinca 2024. Hibridni okvir rastavlja svaku seriju na linearnu komponentu koju bilježi ARIMA i nelinearnu komponentu reziduala koju bilježi jedan od tri algoritma strojnog učenja: umjetne neuronske mreže (ANN), potpornu vektorsku regresiju (SVR) i slučajnu šumu (RF). Eksperimenti izvan uzorka s pomičnim prozorom pokazuju da svi hibridni modeli znatno nadmašuju svoje samostalne inačice. Hibridni ARIMA–ANN postiže najnižu srednju pogreške (RMSE) od 0,324 za predviđanje inflacije, u usporedbi s 0,517 za samostalni ARIMA i 0,468 za samostalni ANN. Za rast BDP-a, hibridni ARIMA–RF postiže RMSE 0,187, što je poboljšanje od 34% u odnosu na samostalni ARIMA. Diebold–Mariano testovi potvrđuju da su sva poboljšanja prognoza statistički značajna na razini signifikantnosti od 1%. Rezultati podržavaju hibridno modeliranje kao održiv put prema pouzdanijem makroekonomskom predviđanju u gospodarstvima u razvoju koja karakterizira strukturna nestabilnost i promjene režima.

KLJUČNE RIJEČI

ARIMA, rast BDP-a, hibridno prognoziranje, inflacija, strojno učenje

VRSTA ČLANKA
Prethodno priopćenje

INFORMACIJE O ČLANKU
Primljeno: 12. lipnja 2026.
Prihvaćeno: 26. lipnja 2026.
DOI: 10.62366/crebss.2026.1.004
JEL: C22, C45, C53, E31, E37