

Superiority of Deep Neural Models Over Econometric and Hybrid Models in Predicting Volatility Food Price

Modupe Stella Omotayo-Tomo^{1,*}, Deebom Zorle Dum², Lekara-Bayo Ifeoma Better³, Agbeye Isaac Oluwashola⁴, Ahmad Albloush⁵, Dang Van My⁶

^{1,4}Department of Mathematical Sciences, Olusegun Agagu University of Science and Technology, Okitipupa, Ondo State, Nigeria.

²Department of Statistics, Rivers State University, Port Harcourt, Rivers State, Nigeria.

³Department of Mathematics, Rivers State University, Port Harcourt, Rivers State, Nigeria.

⁵Department of Business, Amman Arab University, Amman, Amman Governorate, Jordan.

⁶Department of Marketing, University of Finance - Marketing, Duc Nhuan Ward, Ho Chi Minh City, Vietnam.

modupe.omotayo@oaustech.edu.ng¹, zorle.deebom1@ust.edu.ng², ifeoma.lekara-bayo1@ust.edu.ng³, agbeyeoluwashola@gmail.com⁴, bloushkaddb84@gmail.com⁵, dvanmy@ufm.edu.vn⁶

*Corresponding author

Abstract: This study examines linear, non-linear, and combined predictive models to forecast food prices in Nigeria using monthly time-series data from the World Food Programme spanning 15 January, 2002, to 15 July, 2026. Descriptive statistical analysis indicates a right-skewed distribution ("skewness"=1.46) with an average value of 3,644.99, a standard deviation of 3,226.45, and a Jarque–Bera test result that is highly significant (182.28, $p=2.62 \times 10^{(-40)}$), rejecting the assumption of normality. CUSUM test results (3.0232, $p=2.30 \times 10^{(-8)}$) and Bai–Perron analysis of structural fractures suggest continuous regime shifts among sub-periods (2002-2026). ADF (-3.7686, $p = 0.003$) and PP (-3.638, $p = 0.005$) unit root assessments confirm first-difference stationarity (I (1)). Based on the lowest information criteria (AIC = 5073.054, BIC = 5084.105, HQIC = 5077.480) and diagnostics for white-noise residuals, ARIMA (1,1,1) was the best linear model. In the out-of-sample assessment, the standalone ARIMA model had significant forecast errors (RMSE = 1793.56, MAE = 843.73, MAPE = 41.15%), while the standalone LSTM had even larger errors (MAPE = 67.74%). They showed moderate gains, with the Hybrid ARIMA–ANN model achieving RMSE = 1425.53 and MAPE = 31.68%, and the Hybrid ARIMA–LSTM model achieving RMSE = 1808.26 and MAPE = 41.35%. The Artificial Neural Network (ANN) has the best prediction accuracy and error metrics (RMSE = 1318.93, MAE = 594.87, MAPE = 26.02%).

Keywords: Artificial Neural Network (ANN); Structural Breaks; World Food Programme (WFP); Jarque–Bera Test; Nonlinear Dynamics; Information Criteria; Forecast Errors.

Cite as: M. S. Omotayo-Tomo, D. Z. Dum, L. B. I. Better, A. I. Oluwashola, A. Albloush, and D. V. My, "Superiority of Deep Neural Models Over Econometric and Hybrid Models in Predicting Volatility Food Price," *AVE Trends in Intelligent Technoprise Letters*, vol. 3, no. 2, pp. 101–115, 2026.

Journal Homepage: <https://avepubs.com/user/journals/details/ATITP>

Received on: 15/05/2025, **Revised on:** 20/07/2025, **Accepted on:** 01/09/2025, **Published on:** 05/06/2026

DOI: <https://doi.org/10.64091/ATITP.2026.000335>

1. Introduction

Copyright © 2026 M. S. Omotayo-Tomo *et al.*, licensed to AVE Trends Publishing Company. This is an open access article distributed under [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits copying, remixing, redistributing, transformation, and building upon the material in any medium so long as the original work is properly cited.

Food price inflation poses a significant threat to macroeconomic stability and households' quality of life in developing nations, especially in Nigeria, where families spend a substantial share of income on food [1]. Recently, local food markets across Nigeria have experienced unprecedented price fluctuations, driven by structural shocks, currency depreciation, climate-change-related agricultural disruptions, and supply chain disruptions [3]. For policymakers, agricultural stakeholders, and monetary authorities, accurately predicting food prices amidst such unpredictable conditions is vital for implementing prompt measures and protecting against food insecurity [13]. Nevertheless, the intricate factors affecting food price statistics in Nigeria—marked by intense non-stationarity, sudden regime shifts, and non-linear interactions—pose significant challenges for traditional quantitative modeling approaches. Traditionally, linear econometric models based on the Box–Jenkins approach, including the Autoregressive Integrated Moving Average (ARIMA) model, have been the standard for forecasting economic time series [4]. Although these classic models provide strong theoretical foundations and succinct representations of linear relationships, they depend heavily on assumptions of structural linearity and homoskedasticity.

As a result, they often fail to capture the rapid spikes and nonlinear trends typical of markets experiencing ongoing economic disturbances [6]. To overcome these shortcomings, researchers have increasingly explored computational intelligence techniques, such as feedforward Artificial Neural Networks (ANNs) and recurrent Long Short-Term Memory (LSTM) models, alongside hybrid linear-nonlinear modeling strategies that dissect time series into parallel linear and residual segments [2]. Even though hybrid models offer theoretical advantages by combining the structural benefits of ARIMA with the pattern-recognition capabilities of deep learning, their real-world effectiveness in highly volatile market conditions remains a topic of active discussion [12]. This study compares the forecasting performance of linear econometric models, independent deep learning systems, and hybrid configurations applied to Nigeria's monthly food price indices from 2002 to 2026. Through a systematic evaluation of model effectiveness across out-of-sample loss metrics, this research demonstrates the clear dominance of standalone neural networks over traditional econometric methods and hybrid approaches in capturing the intricate, time-sensitive nature of food price fluctuations [5].

2. Methodology

This segment describes the methodological framework used in the research, combining traditional econometric methods with contemporary machine learning techniques to model and predict food price inflation. In particular, the research integrates ARIMA models with Artificial Neural Networks, Long Short-Term Memory models, and various hybrid or ensemble methods to improve forecast accuracy and reliability. The evaluation uses secondary monthly time-series data on food price indices in Nigeria, obtained from the World Food Programme database. The dataset covers 15 January 2002 to 15 July 2026, providing a continuous, uniformly distributed time series. The data includes essential features common in economic time series, including trends, seasonal variations, structural disruptions, nonlinearity and potential heteroskedasticity [7]. To ensure data reliability, researchers preprocessed it by cleaning and identifying missing values, then applied a logarithmic transformation when needed to stabilize variance. Additionally, researchers performed stationarity assessments—specifically the Augmented Dickey-Fuller, Phillips-Perron, and Kwiatkowski-Phillips-Schmidt-Shin tests—to validate the series' statistical properties before model estimation [5]. The methodology employs a hybrid modelling framework that fuses statistical methods with machine learning techniques, applied to a robust real-world dataset, to improve forecasts of food price inflation in Nigeria. These tests use differencing. The first difference was applied as $\Delta y_t = y_t - y_{t-1}$ where unit roots are confirmed ($d=1$). Also, feature scaling was applied, where Min-Max normalization was used to scale input features into the interval $[0,1]$ for neural network convergence $X_{\text{scaled}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$. Also, the model for the study is the ARIMA model, and the standard Autoregressive Integrated Moving Average model, non-seasonally denoted as ARIMA(p,d,q), is expressed as:

$$\phi(B)(1 - B)^d Y_t = c + \theta(B)\varepsilon_t \quad (1)$$

Where Y_t represents the food price index at time t . B is the backshift operator such that:

- $B^k Y_t = Y_{t-k}$ d is the order of integration/differencing required to achieve stationarity
- $\phi(B) = 1 - \text{difference } B^i$ represents the Autoregressive (AR) polynomial of order p
- $\theta(B) = 1 + \sum_{j=1}^q \theta_j B^j$ represents the Moving Average (MA) polynomial of order q

c is a constant term, and $\varepsilon_t \sim \text{WN}(0, \sigma^2)$ is a white noise error process. In expanded form, the differenced series $Y'_t = (1 - B)^d Y_t$ is specified as:

$$Y'_t = c + \sum_{i=1}^p \phi_i Y'_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t \quad (2)$$

Based on the Box–Jenkins framework for model selection, constructing a suitable time series model follows a structured process that includes identification, estimation, and model selection. The initial phase, known as model identification, focuses on

analyzing the time series' fundamental composition to determine the correct order. This analysis is conducted by reviewing the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) diagrams. These visualization techniques help identify autoregressive (AR) and moving average (MA) components, as well as the amount of differencing needed to achieve stationarity. Specifically, the ACF illustrates the correlation between current and past values of the series, whereas the PACF highlights the immediate impact of earlier observations after accounting for intermediate effects. These diagrams provide preliminary guidance for formulating the ARIMA (p, d, q) model. After identification, the next phase is parameter estimation, during which the parameters of the proposed models are calculated. This process typically uses Maximum Likelihood Estimation (MLE), which finds parameter values that maximize the probability of observing the specified data. MLE is favored for its advantageous statistical properties, such as reliability and robustness, especially in large datasets.

The concluding phase involves model selection using information criteria. In this phase, researchers evaluate competing ARIMA models identified in previous steps and choose the best model by balancing fit quality and simplicity. They do this by minimizing common information criteria, such as the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and the Hannan–Quinn Information Criterion (HQIC). These criteria penalize complex models while favoring improved fit, ensuring the chosen model is precise but not overly complex. The Box–Jenkins approach combines diagnostic graphical tools, statistical estimation methods, and objective selection criteria to develop a robust, reliable time series model. The Artificial Neural Network (ANN) method uses a feedforward network called a Multilayer Perceptron (MLP), designed to capture non-linear time-related dependencies using lagged vector data. Its structure includes an input layer made up of lagged values, two fully connected hidden layers featuring h_1 (such as h_1 (e.g., 50) and h_2 (e.g., 30) processing nodes, along with an output layer that contains one neuron which generates the point forecast $\hat{y}_{t+h}$. For a network incorporating two hidden layers, the mathematical representation and functional mapping can be expressed as (Figure 1):

$$\hat{y}_{t+h} = f_3(W_3 \cdot f_2(W_2 \cdot f_1(W_1 X_t + b_1) + b_2) + b_3) \quad (3)$$

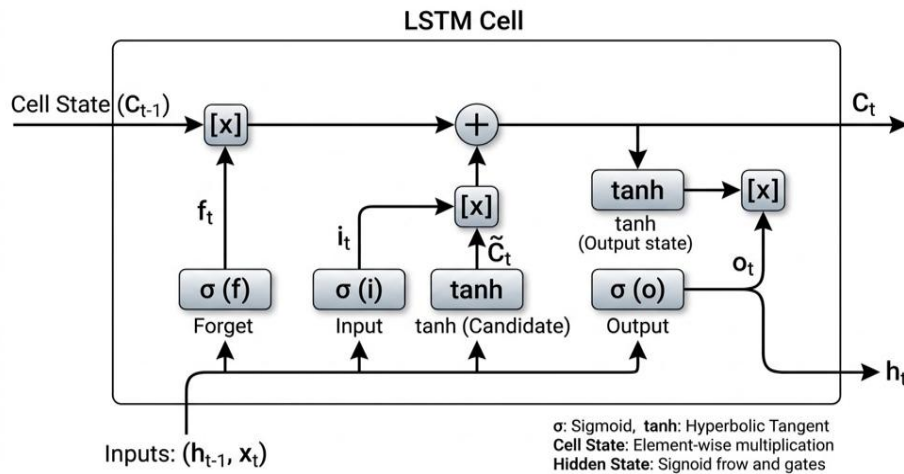


Figure 1: Structure of a long short-term memory (LSTM) cell

Where W_1, W_2, W_3 indicate the weight matrices for the input layer, the first hidden layer, and the second hidden layer; b_1, b_2, b_3 signify the related bias vectors; f_1, f_2 denote non-linear activation functions like the Rectified Linear Unit ($\text{ReLU}(x) = \max(0, x)$); and f_3 signifies a linear activation function used for the regression output. The network is optimized using a Mean Squared Error (MSE) loss function, the Adaptive Moment Estimation (Adam) algorithm, and backpropagation, with hyperparameter tuning via grid search for batch size, learning rate, and maximum epochs. Supporting this method, the Long Short-Term Memory (LSTM) technique employs a specific Recurrent Neural Network (RNN) architecture that mitigates vanishing gradients and effectively captures long-term temporal dependencies via gated memory cells.

2.1. Mathematical System

At step t , with input x_t and previous hidden state h_{t-1} :

Forget Gate: Decides What Information to Discard from the Cell State:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (4)$$

Input Gate: Decides Which New Values to Update in the Cell State:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (5)$$

Candidate Cell State: Creates Candidate Values for the Cell State:

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (6)$$

Cell State Update: Updates the Cell State C_t :

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (7)$$

Output Gate: Controls What Information Exists in the Hidden State:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (8)$$

Hidden State Update: Computes the Output Hidden State h_t :

$$h_t = o_t \odot \tanh(C_t) \quad (9)$$

Where $\sigma(z) = \frac{1}{1+e^{-z}}$ is the sigmoid activation function, $\odot$ denotes the Hadamard (elementwise) product, and W, b are trainable weights and biases. Given a time series Y_t can be conceptualized as an additive synthesis of a linear component (L_t) and a non-linear component:

$$(N_t): Y_t = L_t + N_t \quad (10)$$

In the first stage, an ARIMA model captures the linear structure of Y_t . $\hat{L}_t$. The residual errors e_t represent non-linear patterns remaining in the data: $e_t = Y_t - \hat{L}_t$. The ARIMA-ANN Hybrid Architecture is implemented using the residual series. e_t is modeled using an ANN with p lagged residual inputs: $\hat{e}_t = \text{ANN}(e_{t-1}, e_{t-2}, \dots, e_{t-p})$ The combined hybrid forecast is formulated as:

$$\hat{Y}_{t, \text{ARIMA-ANN}} = \hat{L}_t + \hat{e}_t \quad (11)$$

Similarly, the ARIMA-LSTM Hybrid architecture models the non-linear residual sequence. e_t which is fed into an LSTM network configured for sequential residual prediction:

$$\hat{e}_t = \text{LSTM}(e_{t-1}, e_{t-2}, \dots, e_{t-p}) \quad (12)$$

Yielding the combined prediction of:

$$\hat{Y}_{t, \text{ARIMA-LSTM}} = \hat{L}_t + \hat{e}_t \quad (13)$$

To reduce variance and structural uncertainty across distinct algorithm classes, an equal-weighted ensemble model averages individual outputs:

$$\hat{Y}_{t, \text{Ensemble}} = \frac{1}{3} (\hat{Y}_{t, \text{ARIMA}} + \hat{Y}_{t, \text{ANN}} + \hat{Y}_{t, \text{LSTM}}) \quad (14)$$

In another development, the Model Evaluation and Comparison Framework evaluates out-of-sample forecast precision across all models using three standard loss metrics. Mean Absolute Error (MAE) measures the average magnitude of absolute errors, and it is given as:

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t| \quad (15)$$

Root Mean Square Error (RMSE) was also considered, as it penalizes larger forecast errors heavily, and it is given as:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2} \quad (16)$$

Mean Absolute Percentage Error (MAPE), which provides scale-independent percentage evaluation, is given as:

$$\text{MAPE} = \frac{100}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \quad (17)$$

Where Y_t represents observed actual values, $\hat{Y}_t$ represents forecasted values, and n is the out-of-sample evaluation size. ARIMA diagnostics on the residuals must also satisfy white-noise assumptions. Verified via Ljung-Box Q -tests for serial autocorrelation, residual ACF/PACF visual inspection, and ARCH-LM tests for autoregressive conditional heteroskedasticity. We also performed neural network diagnostics to assess overfitting and monitored training and validation loss convergence. Early stopping is triggered if validation loss fails to decrease over 20 consecutive epochs. Python environments include statsmodels (ARIMA modeling and ADF tests), scikit-learn (preprocessing and evaluation metrics), and TensorFlow/Keras (ANN and LSTM architecture construction). This methodology integrates classical econometric time-series analysis with modern deep learning and a hybrid architecture. By combining ARIMA's linear modeling with ANN and LSTM non-linear pattern recognition, the framework provides a comprehensive evaluation of predictive accuracy for food price inflation.

3. Results

The time plot of the raw food price series offers an initial graphical representation of food price trends in Nigeria across different periods. The series shows clear variability, with sharp spikes, indicating volatility and potential price increases. This graphical representation suggests possible non-stationarity and structural fluctuations in the series, warranting further econometric examination. Table 1 presents the descriptive statistics that provide further numerical details about the characteristics of the food price series. The average figure of ₦3,644.99 signifies the overall food price level during the observed timeframe. In contrast, the median at ₦2,762.05 is significantly below the average, indicating an asymmetrical distribution. This difference suggests that higher price points (outliers) are elevating the mean, indicating a positively skewed distribution. The standard deviation of ₦3,226.45 is relatively substantial and comparable to the mean, highlighting considerable fluctuations in food prices over the assessed timeframe (Figure 2).

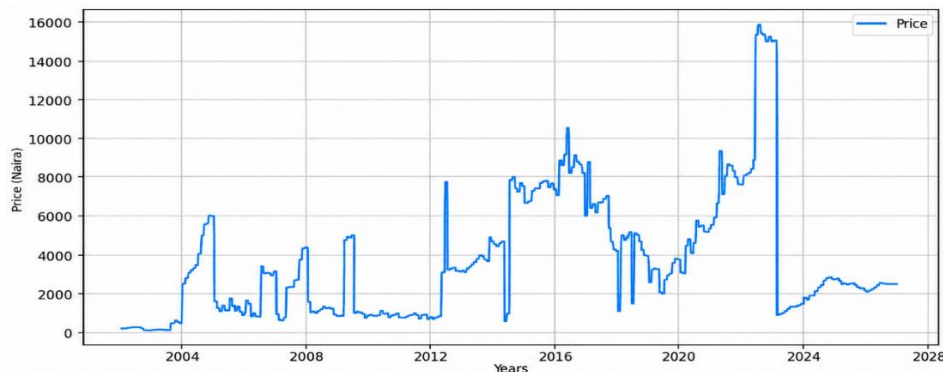


Figure 2: Time plots for the original food price series

This high variability points to instability in the series, suggesting that food prices experience considerable volatility, a common trait in regions affected by inflationary trends and supply chain disruptions. The skewness of 1.46 reinforces the idea that the distribution is right-skewed, meaning that remarkably high prices occur more often than notably low ones. In practice, this suggests sharp price surges in food markets, driven by factors such as inflation, exchange rate fluctuations, and supply disruptions. The kurtosis of 2.51 sits just below the standard threshold of 3 for a normal distribution, indicating a flatter peak and narrower tails than those of a normal distribution. Nevertheless, this does not imply that the series follows a normal distribution, particularly when other statistical measures are considered. The Jarque–Bera test statistic of 182.28, with an associated p-value of 2.62×10^{-40} , is highly significant.

Table 1: Results of the summary descriptive statistics for Nigeria food price (Naira)

Statistic	Estimates	Interpretation
Mean	3644.99	Average value of the series
Median	2762.05	Middle value (50th percentile)
Standard Deviation	3226.45	High dispersion around the mean
Skewness	1.46	Positively skewed distribution

Kurtosis	2.51	Slightly platykurtic (flatter than normal)
Jarque-Bera	182.28	Strong departure from normality
p-value	2.62×10^{-40}	Highly significant (reject normality)

This provides grounds for rejecting the null hypothesis of normality at any standard significance level. Therefore, the food price series deviates from normality. Table 2 presents the CUSUM test results, along with the arrangement of structural breakpoints over the study period. The CUSUM value of 3.0232, accompanied by a notably significant p-value of 2.30×10^{-8} , provides compelling evidence against the null hypothesis of structural stability. This suggests that the model's estimated parameters are not stable over time and exhibit systematic fluctuations, thereby confirming structural instability in the food price series. Rejecting stability suggests that a model with fixed parameters may not effectively capture the series' fundamental dynamics, given the influence of time-varying shocks and regime shifts. The breakpoint summary further corroborates this finding by indicating that structural breaks span several distinct sub-periods.

Table 2: Structural stability (CUSUM Test) and breakpoint summary

Component	Result / Range	Interpretation
CUSUM Test Statistic	3.0232	Strong deviation from parameter stability
p-value	2.30×10^{-8}	Reject the null hypothesis of structural stability
Early Period	2002–2005 (Index 5–40)	Initial structural adjustments and instability
Mid Early Period	2006–2008 (Index 50–80)	Emerging regime shifts in the system
Transition Period	2009–2011 (Index 85–120)	Increased volatility and structural changes
Mid Period	2012–2015 (Index 125–160)	Sustained structural variation
Late Mid Period	2015–2018 (Index 165–200)	Persistent regime instability
Recent Period	2019–2021 (Index 205–240)	High-frequency structural breaks continue
Very Recent Period	2022–2026 (Index 245–290)	Ongoing instability and evolving dynamics

The initial interval (2002–2005) reflects the first signs of instability in the system, while the mid-early and transitional periods (2006–2011) show increasing fluctuations and shifting structural dynamics. From 2012 to 2018, encompassing the mid and late mid periods, the series continues to show ongoing structural variation, pointing to sustained instability rather than isolated shocks. In the most recent period (2019–2026), breakpoint clustering suggests that the system remains significantly unstable, with frequent alterations in the underlying data-generating mechanism. Overall, Table 2 clearly indicates that the food price series exhibits ongoing structural change over time, underscoring the need for flexible econometric or machine learning models that can adapt to time-varying dynamics. Figure 3 visually illustrates the temporal progression of the structural breaks identified using the Bai–Perron approximation applied to the original food price series.

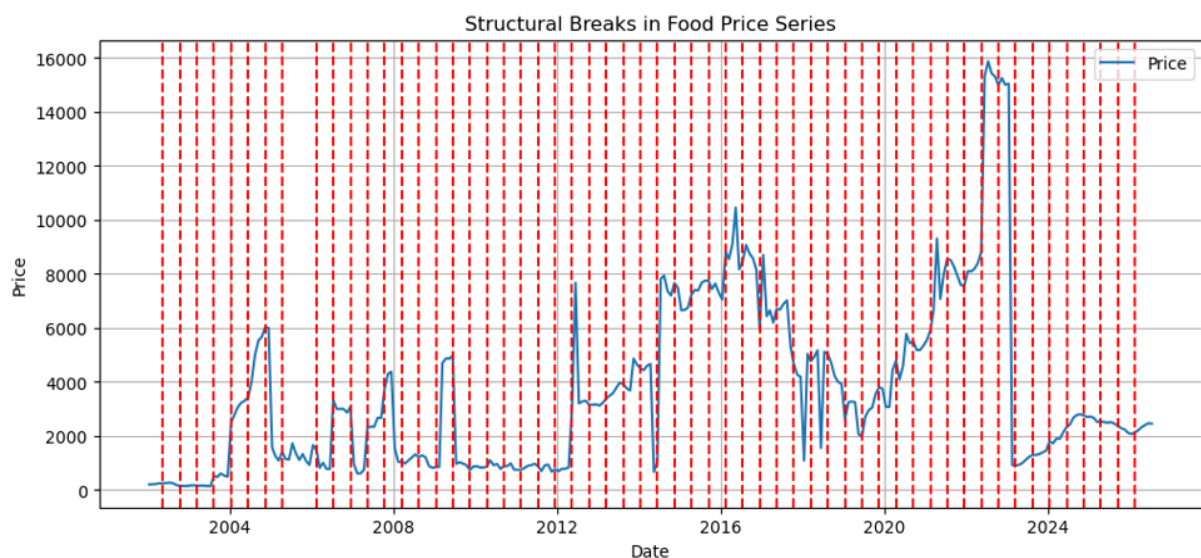


Figure 3: Time plots for structural breaks in food price series using bai-perron approximation

The illustration shows a clear pattern of recurring breakpoints across the entire sample period, from 2002 to 2026. The initial section of the plot (2002–2005), which aligns with the first group of breaks in Table 2, indicates the onset of structural

instability, suggesting that the system experienced parameter changes at this early juncture. As the series advances into the mid-early and transitional periods (2006–2011), Figure 3 shows continued breakpoint clustering, indicating rising volatility and adaptations in the underlying process. From 2012 to 2018, Figure 3 shows a dense aggregation of structural breaks, aligning with the mid and late mid periods noted in Table 2 and suggesting ongoing instability and repeated regime shifts in food pricing dynamics. This trend continues into the most recent timeline (2019–2026), where the break rate remains elevated, indicating that the system has not stabilized even in the latest timeframe.

Table 3: Unit root test results (ADF, PP, KPSS)

Difference Level	Test	Statistic	p-Value	Decision	Conclusion
1st difference	ADF	-3.7686	0.003	Reject H_0	Stationary
1st difference	PP	-3.638	0.005	Reject H_0	Stationary
1st difference	KPSS	0.8741	0.010	Reject H_0	Non-stationary
2nd difference	ADF	-3.7686	0.003	Reject H_0	Stationary
2nd difference	PP	-3.638	0.005	Reject H_0	Stationary
2nd difference	KPSS	0.8741	0.010	Reject H_0	Non-stationary

Figure 3 thus confirms that the food price series does not exhibit a stable long-term structure but instead evolves through regime shifts. This reinforces the notion that the underlying process is non-constant and underscores the need for sophisticated modeling methods that can account for structural breaks and time-varying behavior. Similarly, researchers further conducted the unit root tests (ADF, PP, KPSS), and the results are shown in Table 2. The unit root test findings in Table 3 indicate that although the ADF and PP tests imply that the series is stationary after differencing, the KPSS test continues to indicate non-stationarity. This suggests conflicting indications, and the series may still exhibit inherent non-stationarity even after differencing. Figure 4 shows a direct visual contrast between the initial (level) data and its first-differenced version. The level data exhibit a distinct upward trajectory and increasing variability over time, indicating non-stationarity in both its mean and variance.

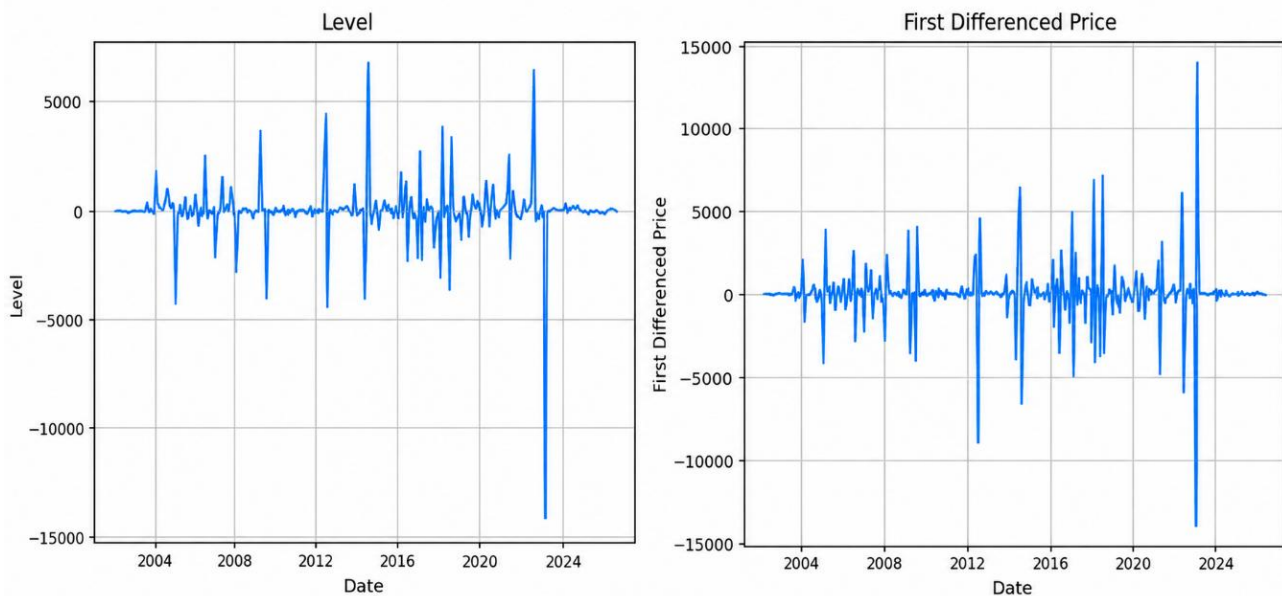


Figure 4: Time plots for level and first difference of the Nigeria food price series

The continuous price increases, along with erratic fluctuations, indicate that the series is influenced by structural factors such as inflationary pressures and potential policy shifts rather than a consistent random process. This pattern suggests that disturbances to the system are enduring rather than transient, a characteristic of integrated processes. In comparison, the first-differenced data reveals a significant change, with the overarching trend largely eliminated, allowing the series to oscillate around a more stable average. While volatility remains, the lack of a deterministic trend indicates that differencing has successfully facilitated stationarity in the mean structure. The differenced representation thus highlights short-term behavior and random variation rather than long-term growth tendencies. This transformation confirms that the initial data are integrated of order one ($I(1)$) and validates the use of ARIMA-type modeling with differencing to achieve stationarity.

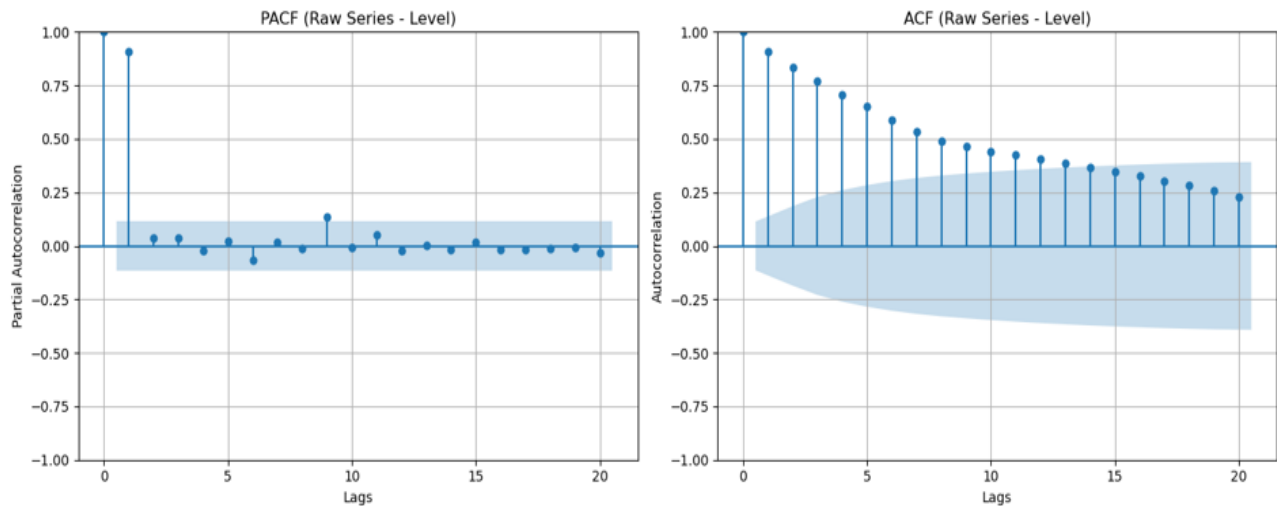


Figure 5: PACF and ACF plot (Raw Series) of the Nigeria food price series

Figure 5 illustrates the autocorrelation function (ACF) and partial autocorrelation function (PACF) for the raw data on Nigeria food prices in Naira. The ACF diagram indicates a gradual, consistent decrease across multiple lags, highlighting significant serial correlation and confirming the series' non-stationarity. The slow decline, rather than a sharp cutoff, implies that earlier values continue to affect current observations over extended periods, which typifies a unit root phenomenon. The PACF diagram similarly displays notable peaks at smaller lags, followed by a gradual decline rather than an abrupt termination. This trend reinforces the assertion that the unprocessed series is non-stationary and possesses long-memory or persistent autoregressive characteristics. The combination of a slowly decreasing ACF and a non-cutoff PACF suggests that the data cannot be effectively modeled in its raw form using traditional stationary ARMA frameworks. Consequently, the series requires differencing before meaningful ARIMA identification can be performed, supporting the transformation shown in Figure 4 and laying the groundwork for determining suitable ARIMA (p, 1, q) models in further analysis.

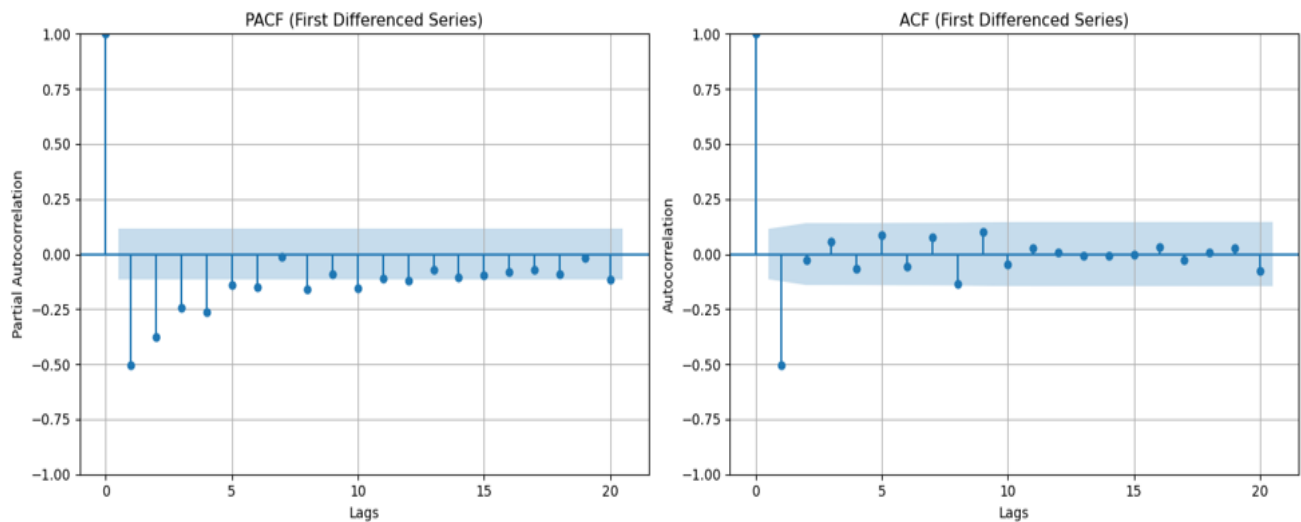


Figure 6: PACF and ACF plot (First-Differenced Series) of the Nigeria food price series

Figure 6 shows the results of the PACF and ACF Plot (First-Differenced Series) of the Nigerian food Price Series. According to the Box–Jenkins identification method, the first-differenced series shows a PACF with notable peaks up to lag 6 (with a clear cutoff after lag 7), suggesting a higher-order autoregressive structure. In contrast, the ACF shows a prominent peak at lag 1, consistent with a first-order moving-average process. As a result, researchers defined ten potential ARIMA models, ranging from ARIMA(0,1,1) to ARIMA(7,1,1). Nevertheless, ARIMA (0,1,1) was chosen as the reference model because of its simplicity, robust theoretical justification based on the ACF pattern, and widespread use as a standard for modeling economic time series. Additionally, parsimony is a key principle in the Box–Jenkins methodology because it represents the most straightforward model, uses the fewest parameters and minimizes the potential for overfitting [10].

Table 4: Comparative ARIMA model selection results

Model	Log Likelihood	AIC	BIC	HQIC	Significant Parameters	Ljung-Box (p)	JB (p)	Remarks
ARIMA (0,1,1)	-2537.976	5079.952	5087.319	5082.902	MA (1) ✓	0.90	0.000	Baseline
ARIMA (1,1,1)	-2533.527	5073.054	5084.105	5077.480	AR (1), MA (1) ✓	0.60	0.000	★ Best
ARIMA (3,1,1)	-2533.212	5076.425	5094.843	5083.801	AR (1), MA (1) ✓	0.97	0.000	Good
ARIMA (4,1,1)	-2534.543	5081.087	5103.188	5089.938	AR (1–2), MA (1) ✓	0.99	0.000	Overfit
ARIMA (5,1,1)	-2534.523	5083.046	5108.831	5093.372	AR (1–2), MA (1) ✓	1.00	0.000	Overfit
ARIMA (6,1,1)	-2532.502	5081.003	5110.472	5092.805	AR (1), MA (1) ✓	0.98	0.000	Overfit
ARIMA (7,1,1)	-2532.511	5083.021	5116.173	5096.298	AR (1), MA (1) ✓	0.97	0.000	Overfit
ARIMA (5,1,0)	-2536.590	5085.180	5107.281	5094.031	None ✓	0.98	0.000	Weak
ARIMA (6,1,0)	-2536.163	5086.326	5112.111	5096.652	None ✓	0.98	0.000	Weak
ARIMA (7,1,0)	-2536.082	5088.165	5117.634	5099.966	None ✓	0.94	0.000	Weak

Table 4 contains the Comparative ARIMA Model Selection Results. The comparative evaluation of various competing ARIMA models was performed using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan–Quinn Information Criterion (HQIC). Among the analyzed models, ARIMA (1,1,1) had the most favorable AIC (5073.054), BIC (5084.105), and HQIC (5077.480) values, indicating a better fit. Additionally, all evaluated parameters were statistically significant, and the Ljung–Box test indicated no serial correlation in the residuals. Although higher-order models, such as ARIMA (6,1,1) and ARIMA (7,1,1), were assessed, they included multiple insignificant coefficients and had higher information criteria values, suggesting potential overfitting. Hence, the ARIMA (1,1,1) model was chosen as the most suitable and efficient for the dataset. Following the results in Table 4 of ARIMA Model Comparisons, the research proceeds to assess the chosen ARIMA model's sufficiency and reliability by conducting diagnostic analyses, evaluating forecast performance, and measuring accuracy.

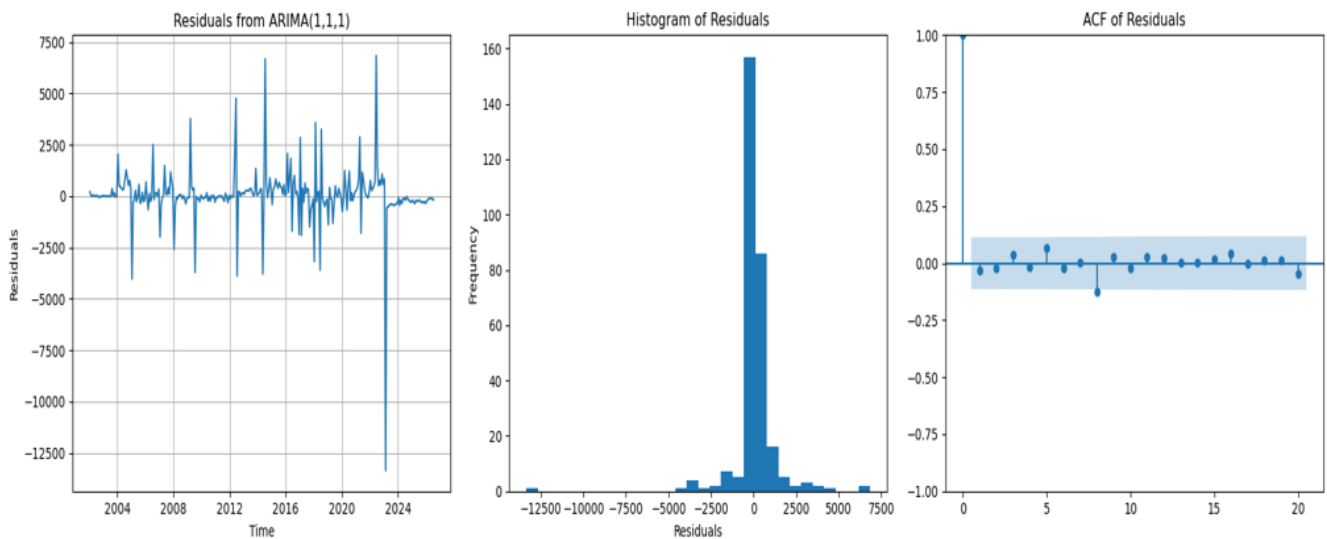
**Figure 7:** Diagnostic check for residuals from ARIMA (1,1,1), histogram and ACF of the residuals

Figure 7 shows the diagnostic plots of the residuals for the ARIMA (1,1,1) model, whereas Table 5 encapsulates the outcomes of the ARCH and diagnostic tests for the residuals. Additionally, Figure 8 depicts the ARIMA (1,1,1) predictions for 2020 to 2026, and Table 6 summarizes the metrics used to assess forecasting accuracy. Figure 7 contains the diagnostic check for residuals from ARIMA (1,1,1), the histogram and the ACF of the residuals. A residual diagnostic evaluation assessed the

effectiveness of the ARIMA (1,1,1) model. The residual plot shows random variation centered on zero, with no identifiable pattern, suggesting the model has adequately captured the time series' inherent structure. The residual histogram shows minor departures from normality, a typical feature of financial time series. In addition, the autocorrelation function of the residuals indicates no significant autocorrelation, confirming that the residuals behave as white noise.

Table 5: Summary for ARCH test and residual diagnostics

Test	Statistic	p-Value	Decision	Interpretation
ARCH LM Test	12.1564	0.433205	Fail to reject H_0	No ARCH effect (homoskedastic residuals)
ARCH F-Test	1.0099	0.439898	Not significant	Constant variance confirmed
Ljung–Box (Q)	7.5092	0.676657	Fail to reject H_0	No residual autocorrelation

These findings endorse the appropriateness of the ARIMA (1,1,1) model for prediction. Table 5 summarizes the ARCH test and residual diagnostics. Because the ARCH-LM test is not statistically significant ($p > 0.05$) and the Ljung-Box test shows no autocorrelation in the residuals, the ARIMA (1,1,1) model is properly specified and effectively captures both the mean and variance characteristics of the dataset. As a result, additional volatility modeling, like GARCH, is unnecessary. The appropriate next step is to assess out-of-sample forecasts or compare machine learning models such as ANNs or LSTMs, especially if the goal is to improve predictive precision rather than focus solely on statistical specification. Figure 8 shows the ARIMA (1,1,1) forecast for the 2020-2026 projection. The ARIMA (1,1,1) model produced forecasts for 2020 through 2026. The resulting forecast trajectory suggests that the existing trend in the past data will continue.

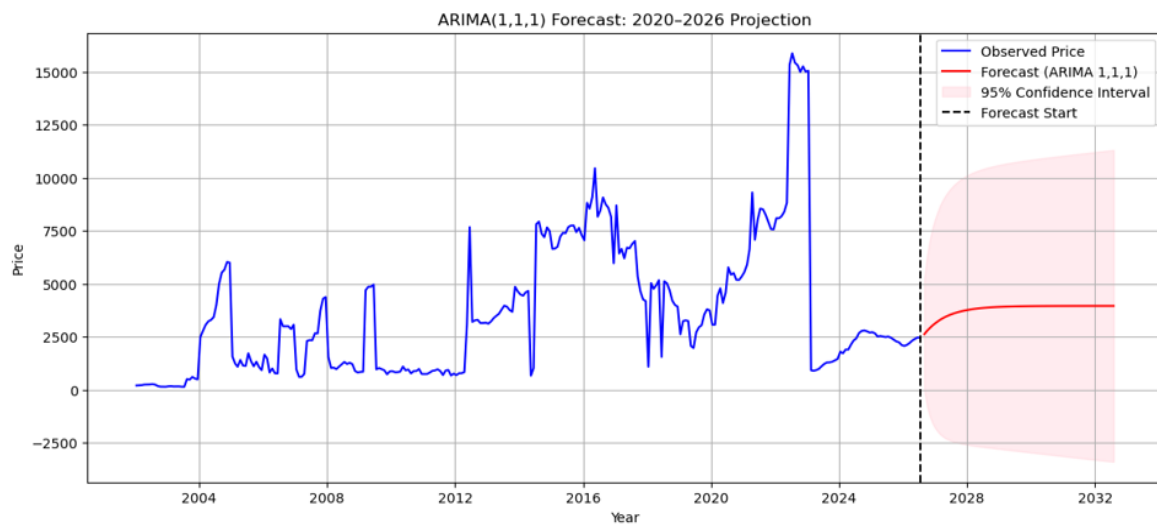


Figure 8: ARIMA (1,1,1) forecast: 2020-2026 projection

The expanding confidence intervals indicate increasing uncertainty as the forecast horizon lengthens, a common characteristic of stochastic time-series models. In summary, the model offers a reliable medium-term outlook on price movements. Table 6 summarizes forecast accuracy. The very high forecast error (MAPE = 91.73%) implies that the ARIMA (1,1,1) model cannot capture the fundamental dynamics of the data series. This suggests intricate, nonlinear time-based relationships. Therefore, enhanced machine learning approaches such as ANNs and LSTMs, along with combined ARIMA–ANN and ARIMA–LSTM models, are suitable for boosting prediction accuracy by addressing nonlinear residual patterns that the linear ARIMA model cannot account for. As Zhang [15] and Alsawaylimi [2] note, moving to ARIMA–LSTM, ARIMA–ANN, and hybrid model is justified when the time series exhibits both linear and nonlinear structures, or when improving forecast accuracy is the primary objective.

Table 6: Summary of forecast accuracy

Metric	Value	Interpretation
RMSE	4627.93	Large absolute forecast error
MAE	3526.90	High average deviation
MAPE	91.73%	✗ Very poor percentage accuracy

The literature shows that hybrid models outperform standalone ARIMA and ANN models by 5–20% because they decompose linear and nonlinear components. However, if ARIMA residuals are white noise and the ARIMA (1,1,1) model already provides optimal information criteria with significant parameters, further hybrid modelling may not be necessary for inferential purposes. Table 7 shows the results of a comparison of prediction precision and forecast results from Models (ARIMA, ANN, LSTM and Combined Models). Table 7 displays prediction accuracy metrics (RMSE, MAE, MAPE) along with the relevant multi-step forecasting paths for each model. It facilitates a straightforward comparison of predictive error outcomes and forecasting patterns across modeling approaches. Among the models, ANN has the lowest RMSE (1318.93) and MAE (594.87), indicating its superior predictive accuracy, whereas LSTM shows comparatively higher errors and less effective generalization.

Table 7: Comparison of prediction precision and forecast results from models (ARIMA, ANN, LSTM and Combined Models)

Model	RMSE	MAE	MAPE (%)	Forecast Output (t+1 to t+5 Mean Pattern)
ARIMA	1793.56	843.73	41.15	2626.19 → 2626.19 → 2626.19 → 2626.19 → 2626.19
ANN	1318.93	594.87	26.02	2685.01 → 2877.10 → 3075.23 → 3242.25 → 3432.88
LSTM	1665.89	946.40	67.74	2610.84 → 2690.82 → 2768.16 → 2846.99 → 2928.34
Hybrid ARIMA–ANN	1425.53	707.90	31.68	(Blended ANN-adjusted forecast; upward trend with smoothing effect)
Hybrid ARIMA–LSTM	1808.26	856.23	41.35	(Smoothed ARIMA backbone with LSTM correction; mild upward drift)

The hybrid models show average performance, indicating some improvement from combining linear (ARIMA) and nonlinear (ANN/LSTM) frameworks. However, they do not consistently outperform the independent ANN model. Figure 9 illustrates a carefully organized feedforward Artificial Neural Network designed to predict food prices. In this configuration, five lagged input variables (X_1 – X_5) are processed by a multi-layered framework with two hidden layers: the first containing 50 neurons and the second 30, resulting in a single output (Y). This setup suggests that the network is moderately deep and has sufficient capacity to capture nonlinear temporal associations in the dataset, especially those that linear models such as ARIMA cannot capture. Empirically, incorporating lagged inputs aligns with time-series embedding theory, which holds that previous observations indicate future values. At the same time, the extensive interconnections show that each neuron helps learn intricate relationships among lagged price movements.

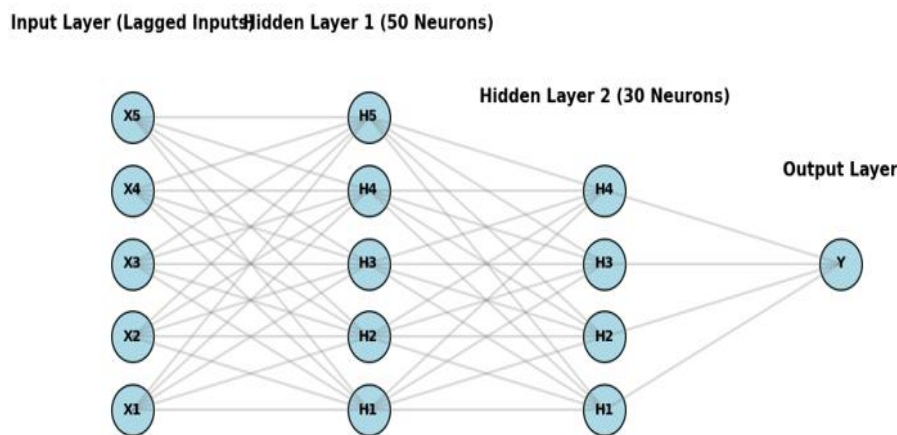


Figure 9: ANN architecture-food price forecasting model

From a statistical perspective, the model's previously obtained performance metrics reinforce the success of this structure. The ANN recorded the smallest forecasting errors (RMSE = 1318.93, MAE = 594.87, MAPE = 26.02%) compared with ARIMA (RMSE = 1793.56, MAE = 843.73, MAPE = 41.15%) and LSTM (MAPE = 67.74%). This represents a roughly 26.5% decrease in RMSE compared to ARIMA and an improvement of over 60% in MAPE compared to LSTM, demonstrating better predictive precision. The improved performance indicates that the ANN effectively captures nonlinear trends and structural intricacies in the food price data, supported by previous tests for structural breaks and instability (CUSUM results). As a result, the architecture shown in Figure 9 is empirically validated as a potent nonlinear modeling framework for forecasting fluctuating economic time series, such as food prices. Figure 10 compares the forecasts and shows that the ANN and Hybrid ARIMA–ANN models track the actual food price data more closely than the other approaches, especially during periods of significant change.

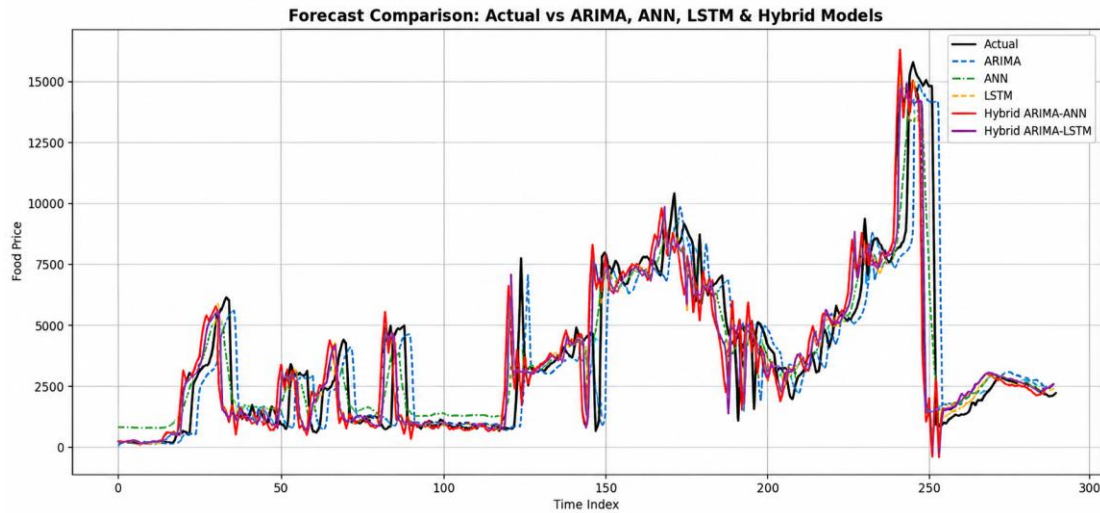


Figure 10: Plots for the prediction from models (ARIMA, ANN, LSTM and Combined Models) comparison

This observation aligns with your accuracy findings: the ANN shows the lowest RMSE (1318.93) and MAE (594.87), indicating smaller average prediction errors than ARIMA (RMSE = 1793.56) and LSTM (RMSE = 1665.89). The ARIMA model is overall stable, but it appears relatively flat and slow to respond to changes, suggesting it captures linear trends and does not adequately address the dataset's short-term nonlinearities. Statistically, the LSTM model shows greater variation from the actual data, which aligns with its higher MAPE (67.74%), suggesting low accuracy and potential overfitting or insufficient tuning. The Hybrid ARIMA-ANN model outperforms ARIMA, reducing RMSE to 1425.53 and identifying turning points more effectively, suggesting that integrating linear and nonlinear frameworks improves predictive performance. Overall, the results confirm that nonlinear models (notably ANNs) deliver higher forecasting accuracy, supporting nonlinear trends in the food price data that standard linear ARIMA models cannot fully capture.

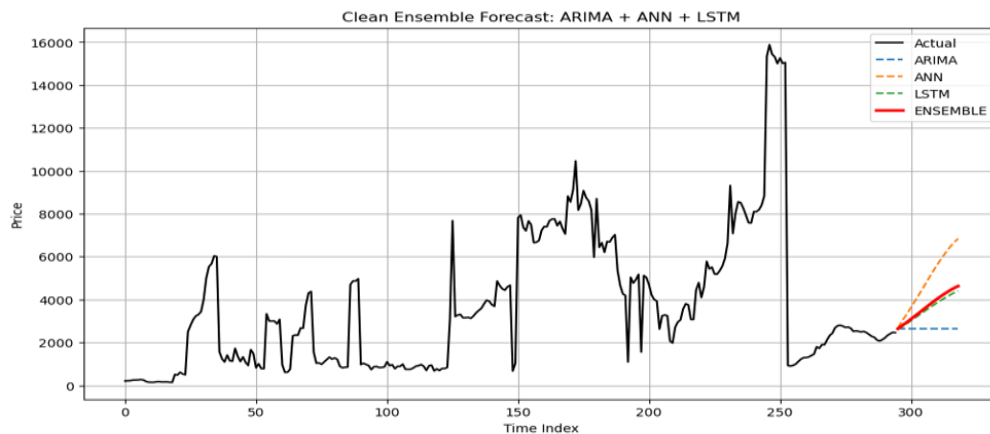


Figure 11: Plots for the prediction for ensemble of ARIMA, ANN, LSTM and combined models comparison

Table 8 compares projected outputs from Various Models (Projections for 2020–2026). The projections indicate distinct variations in the performance and predictive behavior of the models utilized. The ARIMA model yields static predictions (2626.19 for every period), showing its linear framework's inability to adapt to shifting trends in the data, which aligns with previous findings of structural changes and nonlinear patterns. Conversely, the ANN model shows a pronounced upward trend, rising from 2685.01 to 3432.88, indicating its ability to capture nonlinear growth and changes in food pricing. The LSTM model also shows a rising pattern, but at a gentler pace (2610.84 to 2928.34), suggesting a limited understanding of time-dependent correlations and thus performing less effectively than the ANN.

Table 8: Forecast output comparison across models (2020–2026 Projection)

Period (t)	ARIMA	ANN	LSTM	Ensemble
1	2626.19	2685.01	2610.84	2640.68

2	2626.19	2877.10	2690.82	2731.37
3	2626.19	3075.23	2768.16	2823.19
4	2626.19	3242.25	2846.99	2905.14
5	2626.19	3432.88	2928.34	2995.80

The ensemble model provides more balanced forecasts (2640.68-2995.80), combining ARIMA's reliability with the flexibility of ANN and LSTM. This observed behavior aligns with previous accuracy evaluations, where the ANN outperformed the other models. At the same time, the ensemble approach provides a middle ground between bias (ARIMA) and variance (ANN/LSTM), leading to more stable and realistic projections. Figure 11 shows the plots comparing the ensemble predictions of ARIMA, ANN, LSTM and the Combined Model. The ensemble prediction results are obtained by averaging forecasts from the ARIMA, ANN, and LSTM models, improving accuracy and reducing individual model biases and fluctuations.

4. Discussions

Table 4 illustrates the results of the Comparative ARIMA Model Selection based on AIC, BIC, and HQIC criteria. The ARIMA (1,1,1) model was identified as the most effective specification, with the lowest information criteria (AIC = 5073.054, BIC = 5084.105, HQIC = 5077.480), indicating an optimal compromise between fit and simplicity. This outcome aligns with the Box–Jenkins approach, favoring models that minimize information loss. Hyndman and Athanasopoulos [9] reported similar findings, noting that ARIMA models with lower AIC values tend to outperform higher-order models because they are less likely to overfit—Tables 4 and 5 and Figure 7 present validation diagnostics for the chosen ARIMA(1,1,1) model. The residual visualizations in Figure 7 show randomness around a zero mean and no significant autocorrelation, supporting the presence of white noise.

Table 5 also includes results from the ARCH and Ljung–Box tests, which indicate homoskedasticity and the absence of serial correlation, reinforcing the model's adequacy. This aligns with Zhang [15], who noted that a well-defined ARIMA model should capture linear dependence and yield independently distributed residuals. Figure 8 and Table 6 present the ARIMA (1,1,1) forecast for 2020–2026 and evaluate its accuracy. The projected path reflects a seamless extension of preceding trends; however, the high MAPE (91.73%) indicates weak predictive accuracy on out-of-sample data. This suggests that, while the model may be statistically sound within the sample, it does not adequately capture nonlinearities and structural changes in the data. This observation aligns with Alsuwaylimi [2] findings, which noted that ARIMA models tend to underperform on highly fluctuating and structurally changing economic time series [10].

After Table 6, Tables 7 and 8 and Figure 9 compare ARIMA, ANN, LSTM, and hybrid models. The ANN model shows superior forecasting performance, achieving the lowest error metrics (RMSE = 1318.93, MAE = 594.87, MAPE = 26.02%) and surpassing both ARIMA and LSTM. This supports the existence of nonlinear features within the food price data. The hybrid ARIMA–ANN model also improves performance over ARIMA alone, supporting the idea of modeling linear and nonlinear elements together [15]. Nonetheless, ANN remains the most effective model in this analysis, suggesting significant nonlinear dynamics within the dataset. In Figure 10, the forecast comparison reinforces that both ANN and Hybrid ARIMA–ANN closely follow the actual data series, whereas ARIMA shows more static, less reactive behavior.

LSTM shows greater divergence, consistent with its elevated MAPE (67.74%), suggesting poorer generalization. This trend supports the findings of Makridakis et al. [10] and Azari and Moradipour [11], who concluded that ANN-based models frequently surpass traditional statistical models in nonlinear economic forecasting. Table 8 and Figure 11 show the multi-model prediction estimates for 2020 to 2026. The ARIMA model produces steady forecasts because of its linear design, whereas the ANN model captures a pronounced upward nonlinear trajectory. The ensemble approach yields more consistent intermediate predictions, reducing variance by integrating the strengths of multiple models. This finding supports Armstrong [7] and Razavi et al. [8], who illustrated that ensemble forecasting enhances reliability by averaging out structural biases across multiple models. Tables 4 to 8 and Figures 8 to 11 consistently show that although ARIMA is statistically useful for identifying linear patterns [14], it falls short in forecasting precision. Nonlinear models, especially ANNs, deliver better predictive performance, reinforcing the idea of structural instability and nonlinear behavior in Nigeria's food price data.

5. Conclusion

This research compares time-series data to evaluate food price inflation forecasts using ARIMA, ANN, LSTM, and hybrid models. It presents four major conclusions aligned with its stated objectives. To begin, concerning the identification and selection of models, the research finds that the ARIMA (1,1,1) model stands out as the most suitable linear model for the dataset, having achieved the smallest values for AIC (5073.054), BIC (5084.105), and HQIC (5077.480), while all parameters remained statistically significant. Additional diagnostic assessments confirmed that the residuals behaved as white noise, with no signs of autocorrelation or ARCH effects, thereby confirming the model's correct specification in capturing the linear

characteristics of the data series. In the second finding on the ARIMA model's forecasting capabilities, the research concludes that while ARIMA (1,1,1) provides a consistent and statistically valid fit, its predictive performance is inadequate in real-world forecasting scenarios, as evidenced by an elevated MAPE of 91.73%. This showcases the model's inability to adequately capture the fundamental nonlinear and structural intricacies of food price variation, even though it remains appropriate for statistical inference. Thirdly, regarding the comparative forecasting efficacy of machine learning and hybrid models, the research concludes that nonlinear methodologies, especially the Artificial Neural Network (ANN), greatly surpass both ARIMA and LSTM models. The ANN achieved the lowest forecasting errors (RMSE = 1318.93; MAE = 594.87; MAPE = 26.02%), indicating superior skill at capturing nonlinearities in food price movements. Hybrid models (ARIMA–ANN and ARIMA–LSTM) also achieved higher predictive accuracy than the standalone ARIMA model, affirming the advantages of integrating linear and nonlinear frameworks. Lastly, regarding forecast behavior and the consistency of long-term projections, the research concludes that food price inflation exhibits strong nonlinearities and upward trends. While ARIMA yields static or overly smooth predictions, ANN and hybrid models offer more realistic dynamic forecasts that adequately represent structural shifts and turning points in the data series. This supports nonlinear relationships and justifies applying advanced machine learning techniques to improve forecasting precision.

5.1. Recommendations

Considering the results, the study presents the following recommendations:

- Policymakers, analysts, and scholars should focus on ANN-based and hybrid models (such as ARIMA–ANN and ARIMA–LSTM) for predicting food prices, as these models exhibit greater predictive accuracy than traditional ARIMA.
- Despite the Statistical Validity of ARIMA (1,1,1), it should not be the sole basis for decision-making because of its Limited Forecast Precision. It is more appropriate for establishing baseline models and analyzing short-term linear trends.
- Statistical Organizations and economic planning bodies should incorporate machine learning approaches such as ANN into their official forecasting frameworks to improve adaptability to nonlinear market behavior and structural shifts in price trends.
- Forecasting agencies should implement ensemble methods that merge multiple models, as these strategies yield more stable and reliable forecasts by reducing individual model biases and improving generalization across market conditions.

Acknowledgement: The authors gratefully acknowledge the support and contributions of Olusegun Agagu University of Science and Technology, Rivers State University, Amman Arab University, and University of Finance–Marketing.

Data Availability Statement: Supporting datasets are available from the corresponding author upon reasonable request and subject to appropriate conditions.

Funding Statement: The authors confirm that no financial support or external funding was received for this research.

Conflicts of Interest Statement: The authors declare that there are no conflicts of interest related to this research or its publication.

Ethics and Consent Statement: Researchers obtained ethical approval and secured informed consent from the organization and all participants involved in data collection.

References

1. J. Chaker Aloui and H. B. Hamida, “Modelling and forecasting value at risk and expected shortfall for GCC stock markets: Do long memory, structural breaks, asymmetry, and fat-tails matter?,” *The North American Journal of Economics and Finance*, vol. 29, no. 7, pp. 349–380, 2014.
2. A. Alsawaylimi, “Comparison of ARIMA, ANN and hybrid ARIMA-ANN models for time series forecasting,” *Information Sciences Letters*, vol. 12, no. 2, pp. 1003–1016, 2023.
3. J. Bai and P. Perron, “Computation and analysis of multiple structural change models,” *Journal of Applied Econometrics*, vol. 18, no. 10, pp. 1–22, 2002.
4. G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, “Time Series Analysis: Forecasting and Control,” 5th ed., Wiley, Hoboken, New Jersey, United States of America, 2015.

5. D. A. Dickey and W. A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association*, vol. 74, no. 366, pp. 427–431, 1979.
6. O. S. Adesina and L. O. Obokoh, "A hybrid framework of deep learning and traditional time series models for exchange rate prediction," *Scientific African*, vol. 29, no. 9, p. e02828, 2025.
7. J. S. Armstrong, "Principles of Forecasting: A Handbook for Researchers and Practitioners," *Springer*, New York, United States of America, 2001.
8. S. A. Razavi, A. K. Manshad, and L. Kashani, "Applications of econometrics and artificial intelligence in oil price prediction: A systematic review study," *Energy Strategy Reviews*, vol. 9, no. 1, pp. 326–339, 2026.
9. R. J. Hyndman and G. Athanasopoulos, "Forecasting: Principles and Practice," 3rd ed., *OTexts*, Melbourne, Australia, 2021.
10. S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and machine learning forecasting methods: Concerns and ways forward," *PLOS ONE*, vol. 13, no. 3, p. e0194889, 2018.
11. H. Azari and M. Moradipour, "Using kernel-based collocation methods to solve a delay partial differential equation with application to finance," *International Journal of Computing Science and Mathematics*, vol. 10, no. 1, pp. 105–114, 2019.
12. G. Boako and P. Alagidede, "Currency price risk and stock market returns in Africa: Dependence and downside spillover effects with stochastic copulas," *Journal of Multinational Financial Management*, vol. 41, no. 9, pp. 92–114, 2017.
13. H. Mushtaq, M. Ishtiaq, S. Jamal, S. M. R. Rizvi, and H. Raza, "Regime-Switching volatility and risk quantification in South Asian and developed stock Markets: A Comparative perspective using Markov-Switching GARCH with MLE and MCMC estimations," *The North American Journal of Economics and Finance*, vol. 82, no. 1, p. 102576, 2026.
14. A. Monticini and F. Ravazzolo, "Forecasting the intraday market price of money," *Journal of Empirical Finance*, vol. 29, no. 12, pp. 304–315, 2014.
15. G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, no. 1, pp. 159–175, 2003.

Publisher's Note: The publisher remains impartial concerning jurisdictional claims in published maps and institutional affiliations. Responsibility for the content rests entirely with the authors and does not necessarily reflect the publisher's perspectives.