IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 31, No. 10, October, 2026, pp. 3467-3480
Journal homepage: http://iieta.org/journals/isi

Weighted Neutrosophic XGBoost Model for Sparse Annual Economic Time-Series Forecasting: Evidence from Nigeria's Price Level Index

Lawrence O. Obokoh[1], Kayode Samuel Adekeye[2], Olumide Sunday Adesina[1*]

[1] Johannesburg Business School (JBS), University of Johannesburg, Johannesburg 2092, South Africa
[2] Office of the DVC Teaching & Learning, University of the Gambia, Kanifing P.O. Box: 3530, The Gambia

Corresponding Author Email: adesinao@uj.ac.za

Copyright: ©2026 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.18280/isi.311019

Received: 20 May 2026
Revised: 3 August 2026
Accepted: 11 August 2026
Available online: 31 October 2026

Keywords:
neutrosophic sets, XGBoost, adaptive sample weighting, economic time-series forecasting, sparse data, price level index

ABSTRACT

Forecasting economic indicators in developing economies is challenging when observations are limited, and the underlying series exhibits substantial volatility and structural change. This study develops a Weighted Neutrosophic XGBoost (WN-XGBoost) model for sparse annual economic time-series forecasting. The approach operationalizes the truth, indeterminacy, and falsity components of neutrosophic theory as both input features and adaptive sample weights, allowing observation reliability and uncertainty to influence model training. We examine performance using Nigeria's price level index (PLI), measured relative to the United States, with annual observations from 1990 to 2024. After constructing lagged predictors, 32 observations are retained for evaluation through five-fold expanding-window rolling validation, yielding 25 out-of-sample forecasts. The proposed model is compared with random forest (RF) variants, support vector regression (SVR), autoregressive integrated moving average (ARIMA), Prophet, and Neutro long short-term memory (LSTM) using root mean squared error (RMSE), mean absolute error (MAE), and Diebold–Mariano (DM) tests. WN-XGBoost records the lowest errors among the evaluated models, with an RMSE of 10.964 and an MAE of 8.250. However, the DM results show that its performance is not significantly different from several competing models, warranting caution in interpreting its advantage. Recursive forecasts for 2025–2029 indicate an overall increase in the PLI, while the widening prediction intervals point to substantial uncertainty at longer horizons. The findings provide application-specific evidence that neutrosophic information incorporated through adaptive weighting may be useful for forecasting with limited annual economic data.

1. INTRODUCTION

Predicting economic indicators remains difficult in developing economies because of instability, inadequate data, and limited measurement capacity. One of the measures of economic performance is the price level index (PLI). The PLI identifies how purchasing power parity (PPP) compares to market exchange rates when evaluated against the United States dollar as the base (United States = 100). It is essential that the PLI forecast has minimal error because it is a fundamental tool for studying inflation patterns, exchange rate movements, and worldwide economic developments.

Researchers and financial analysts use traditional time-series forecasting methods, such as autoregressive integrated moving average (ARIMA) and exponential smoothing (ES). The traditional methods offer simple modeling solutions, and the results are easy to interpret. The macroeconomic systems, such as PLI, exhibit nonlinear and complex behavior, which traditional models, such as ARIMA and its extensions, have limited capacity to model to capture nonlinear system behavior [1, 2]. Recent studies demonstrate that machine learning techniques have gained popularity because they provide a platform that explains nonlinear connections while handling complicated data sets.

Ensemble methods, which include random forest (RF) and boosting algorithms, are robust in modelling time-series, as shown in several recent studies, including previous studies [3, 4]. The machine learning models are not without limitations; most machine learning models assume that data remains unchanged, yet they do not account for the exceptional constraints that arise from unknown events, undisclosed information, and data deficiencies. Economic datasets present a significant challenge because they experience common fluctuations, measurement errors, and structural changes [5].

Fuzzy logic, intuitionistic fuzzy sets, and uncertainty-aware frameworks provide a solution to the challenges associated with the limitations of machine learning models. Fuzzy logic enables users to represent partial membership, yet its inability to model indeterminate situations results in restricted performance during highly uncertain conditions [6]. An extension of fuzzy logic is the development of neutrosophic set theory, which is a generalization of the fuzzy system [7]. The theory introduces three separate elements that include truth (T), indeterminacy (I), and falsity (F) to demonstrate a complete range of uncertainty. Fuzzy logic considers T + F = 1, whereas in neutrosophic logic, T, I, and F are independent.

Recent studies show that neutrosophic models function effectively in both time-series forecasting decision-making processes and classification operations [8, 9]. Previous methods only used neutrosophic information as additional inputs for conventional learning models, rather than creating genuine learning systems founded on neutrosophic theory. This limitation motivates the development of an uncertainty-aware learning framework that integrates neutrosophic components directly into the model training process through adaptive weighting and uncertainty-guided optimization.

RF and other tree-based models achieve high performance because these models use their capability to decrease variance while they model complex nonlinear patterns [10]. The XGBoost algorithm uses a boosting approach to enhance prediction accuracy through its process of reducing forecast errors [11, 12]. Previous studies [2, 4, 5] showed that ensemble techniques achieve better results than individual models when applied to time-series forecasting tasks.

Following the limitations of the machine learning model to handle uncertainty measurement directly, the neutrosophic logic frameworks combined with neural networks and deep learning techniques provide better prediction results in situations with uncertain and noisy conditions [13, 14]. The neutrosophic and machine learning frameworks are expected to combine their strengths and produce a minimal forecast error, as shown in various studies. Studies [15, 16] have shown that machine learning with boosting-based methods provides strong forecasting abilities for nonlinear economic time-series systems.

On the application of neutrosophic time series models for forecasting purposes, the previous study [8] showed that it resulted in better prediction accuracy; the previous study [9] introduced an inconsistency-based neutrosophic forecasting model. The previous study [17] developed a hybrid model that combined neutrosophic and particle swarm optimization methods to achieve better prediction results. The previous study [18] demonstrated that selecting appropriate parameters serves as a crucial factor for forecasting results in neutrosophic models. Additional studies on neutrosophic frameworks that are combined with neural networks and deep learning methods to achieve better results in environments that contain uncertain and noisy data are found in previous studies [19, 20]. The framework has been applied to various life scenarios, including classification, anomaly detection, and decision-making, because it can manage uncertain situations [21]. The algebraic structure within the complex neutrosophic soft model was investigated by the previous study [22]. The neutrosophic system enables machine learning to work with uncertain data through its specialized representation, while traditional models thrive on large data. This makes neutrosophic learning relevant for economic forecasting characterized by limited data and structural instability.

Recent forecasting research has increasingly examined uncertainty and machine-learning approaches in economic and financial time-series prediction. Dovern [23] develops a measure of macroeconomic expectation uncertainty and demonstrates its relationship with economic volatility, while Luo et al. [24] show the value of machine-learning methods for forecasting financial market volatility. These studies highlight the importance of explicitly considering uncertainty and nonlinear information when developing forecasting models for complex economic time series.

The neutrosophic framework remains underexplored for multi-step-ahead forecasting, including its application in economic time-series forecasting.

In addition, recent forecasting research has increasingly focused on uncertainty-aware learning, Bayesian forecasting, and ensemble methods to improve predictive performance under complex economic conditions [25, 26]. Similarly, neutrosophic approaches have recently been applied to economic forecasting and decision support by explicitly modelling uncertainty and incomplete information [27, 28]. However, limited attention has been given to integrating neutrosophic uncertainty with adaptive ensemble learning for sparse economic time-series forecasting, which is the primary contribution of the present study.

Existing neutrosophic forecasting methods mostly use feature augmentation; a few integrate uncertainty directly into learning optimization, and sparse-data uncertainty-aware forecasting remains underexplored.

Beyond improving forecasting accuracy, the proposed framework contributes to intelligent information systems by integrating data modelling, neutrosophic uncertainty management, adaptive ensemble learning, and predictive analytics into a unified decision-support framework. By transforming uncertain economic observations into reliability-aware learning weights, the framework enhances the ability of information systems to generate robust forecasts for sparse datasets. Such capabilities are particularly relevant for economic planning, policy evaluation, and other decision-support applications where data availability is limited and uncertainty must be explicitly managed.

The main contributions of this study are summarized as follows:

(a) The introduction of a weighted neutrosophic ensemble learning framework for forecasting under sparse and uncertain economic conditions.
(b) The framework uses adaptive sample weighting to integrate neutrosophic T, I, and F components into the boosting objective.
(c) The proposed approach is specifically designed for small macroeconomic datasets where conventional large datasets are not available.
(d) Show that uncertainty-guided weighting helps to stabilize predictive learning performance during sparse-data economic forecasting tasks.

The proposed framework combines neutrosophic elements as flexible weighting components for ensemble models to allow multi-step prediction. The system aims to provide flexibility in managing unpredictable elements while enhancing the accuracy of its predictions. The novelty of the proposed framework lies in its adaptive neutrosophic weighting mechanism, which integrates uncertainty directly into the optimization process rather than using uncertainty only as additional input features.

The remainder of this paper is structured as follows: The methods and model development are presented in Section 2. The results, discussion of results, and limitations of the study are presented in Sections 3, 4, and 5, respectively. Finally, the conclusion and recommendations are in Sections 6 and 7, respectively.

2. METHODS

2.1 Neutrosophic representation

The proposed weighting mechanism was designed for low-

sample environments, which require uncertainty representation to become more important than the ability to access extensive data. The conventional machine learning models in sparse economic datasets face higher instability risks because of their dependency on noisy observations and structural changes, and their restricted training data. The proposed framework uses neutrosophic reliability measures to determine adaptive learning importance, which helps to maintain temporal structures while reducing the effects of unstable observations. The formulation is discussed as follows:

Let the time-series dataset be defined as:

$$\{x_t\}_{t=1}^{n}, x_t \in \mathbb{R} \tag{1}$$

where, x_t represents the PLI at time t. If the objective is to predict the one-step-ahead future value, the framework would be:

$$y_t = x_{t+1} \tag{2}$$

To enable machine learning modeling, the time series is transformed into a supervised learning dataset.

Each observation is transformed into a neutrosophic triplet:

$$x_t \rightarrow (T_t, I_t, F_t) \tag{3}$$

where,

•$T_t \in [0,1]$: degree of truth (reliability)

•$I_t \in [0,1]$: degree of indeterminacy (uncertainty)

•$F_t \in [0,1]$: degree of falsity (deviation)

Let:

$$\delta_t = \mid x_t - x_{t-1} \mid$$

be the temporal variation (δ), which captures the change between consecutive observations. Global statistics are defined as:

$$x_{\max} = \max(x_t)$$

$$x_{\min} = \min(x_t), \mu = \frac{1}{n}\sum_{t=1}^{n} x_t$$

The neutrosophic components are computed as:

T: Stability measure

$$T_t = 1 - \frac{\delta_t}{x_{\max}} \tag{4}$$

Small $\delta_t \Rightarrow T_t \approx 1$(stable data), while large $\delta_t \Rightarrow T_t$

I: Volatility measure

$$I_t = \frac{\delta_t}{x_{\max} - x_{\min}} \tag{5}$$

The indeterminacy captures relative fluctuation. Higher δ_t translates to higher uncertainty

F: Deviation measure

$$F_t = \frac{\mid x_t - \mu \mid}{x_{\max}} \tag{6}$$

Falsity measures deviation from central tendency. Higher deviation = higher falsity. To capture temporal dependencies, lagged variables are introduced:

$$x_{t-1}, x_{t-2}$$

The feature vector at time t is defined as:

$$X_t = [x_t, x_{t-1}, x_{t-2}, T_t, I_t, F_t] \tag{7}$$

The learning task becomes:

$$X_t \rightarrow y_t = x_{t+1}$$

The proposed formulations of *T*, *I*, and *F* are application-specific operationalizations derived from the fundamental principles of neutrosophic theory. Consistent with previous work [7], in the neutrosophic framework, truth represents the reliability of an observation, indeterminacy quantifies the uncertainty associated with temporal changes, and falsity measures the deviation of an observation from the overall data behavior. These measures preserve the independent representation of truth, uncertainty, and falsity while adapting neutrosophic theory to sparse economic time-series forecasting.

2.1.1 Neutrosophic weighting scheme

To incorporate uncertainty into model training, a weighting function is defined:

$$w_t = T_t(1 - I_t)(1 - F_t)$$

High T_t increases importance. High I_t , F_t reduces influence. The weighting scheme ensures that:

$$w_t \in [0,1]$$

and is used as sample weights during training.

Weighted learning objective:

$$\min_{\theta} \sum_{t=1}^{n} w_t \cdot \ell(y_t, \hat{y}_t) \tag{8}$$

The choice of multiplicative weighting formulation is because it simultaneously maximizes reliability (T) while penalizing uncertainty (I) and deviation (F). This formulation ensures that highly reliable observations contribute more strongly to the learning objective, while unstable observations are automatically down-weighted.

The framework employs additional input features to enrich the representation of the economic time series by explicitly capturing data reliability, uncertainty, and deviation. Second, the same components are combined through the adaptive weighting function $w_t = T_t(1 - T_t)(1 - F_t)$, which determines the contribution of each observation during model optimization. While feature integration enables the learning algorithms to recognize uncertainty-related patterns, adaptive weighting directly influences the optimization process by assigning greater importance to reliable observations and reducing the influence of uncertain ones. This distinction allows the study to evaluate whether uncertainty contributes more effectively through feature augmentation or through uncertainty-guided optimization.

The proposed multiplicative formulation was adopted because it jointly captures the interaction among truth,

indeterminacy, and falsity, ensuring that observations receive high weights only when they are both reliable and associated with low uncertainty and deviation. This is not the same as equal weighting or linear weighting schemes.

2.2 Weighted neutrosophic forecasting models

This section is divided into interconnected sections, namely the RF model, the XGBoost model, and the Linear Regression model.

(i) RF model

RF is an ensemble of decision trees:

$$\hat{y}_t = \frac{1}{M}\sum_{m=1}^{M} f_m(X_t) \tag{9}$$

where, M is the number of trees. f_m denotes a decision tree.

2.2.1 Weighted random forest

$$\hat{y}_t = \frac{1}{M}\sum\nolimits_{m=1}^{M} f_m\left(X_t; \mathcal{D}_m^{(w)}\right) \tag{10}$$

where,

$\hat{y}_t$ is the predicted value of $y_t = x_{t+1}$,

X_t is the feature vector at time t,

f_m is the m-th decision tree,

$\mathcal{D}_m^{(w)}$ is the weighted bootstrap sample,

$$p_t = \frac{w_t}{\sum_{i=1}^{n} w_i}$$

Bootstrap samples are drawn according to p_t.

Trees give more importance to High T_t (reliable data), Low I_t, F_t (less uncertainty).

The weighted bootstrap samples are generated according to the neutrosophic reliability weights, allowing observations with higher truth values and lower uncertainty levels to contribute more strongly during ensemble construction.

(ii) XGBoost model

XGBoost is a boosting-based model that minimizes:

$$\mathcal{L} = \sum_{t} \ell(y_t, \hat{y}_t) + \sum_{k} \Omega(f_k) \tag{11}$$

where, ℓ is the loss function and Ω is the regularization term.

Prediction:

$$\hat{y}_t = \sum_{k=1}^{K} f_k(X_t) \tag{12}$$

Neutrosophic-Weighted XGBoost:

$$\mathcal{L} = \sum_{t=1}^{n} w_t \cdot \ell(y_t, \hat{y}_t) + \sum_{k=1}^{K} \Omega(f_k) \tag{13}$$

where,

$$w_t = T_t(1 - I_t)(1 - F_t)$$

and

$$\hat{y}_t = \sum\nolimits_{k=1}^{K} f_k(X_t) \tag{14}$$

where,

K= number of boosting trees,

f_k= regression tree learner,

$\mathcal{F}$= functional space of trees.

Then the weighted objective:

$$\mathcal{L}^{(k)} = \sum\nolimits_{i=1}^{n} w_i \ \ell\left(y_i, \hat{y}_i^{(k-1)} + f_k(x_i)\right) + \Omega(f_k) \tag{15}$$

where,

w_t is the neutrosophic uncertainty weight associated with observation t,

$\ell(\cdot)$ is the loss function,

$\hat{y}_t$ is the prediction from the previous boosting iteration,

(f_k) is the newly added regression tree at boosting stage k,

$\Omega(f_k)$ is the regularization term controlling model complexity.

The weighting mechanism modifies the boosting objective by assigning larger optimization importance to observations exhibiting higher reliability and lower uncertainty.

(iii) Linear Regression

To quantify forecasting uncertainty, 95% prediction intervals were constructed using a residual-based approach. After fitting the forecasting model, the residuals from the training data were used to estimate the residual standard deviation (σ). Assuming approximately normally distributed forecast errors, the lower and upper prediction limits were computed as $\hat{y}_{t+h} \pm 1.96\sigma$, where $\hat{y}_{t+h}$ denotes the point forecast. This approach provides a simple and computationally efficient measure of forecast uncertainty suitable for sparse economic time-series data.

$$\hat{y}_t = \beta_0 + \sum_{i=1}^{p} \beta_i X_{t,i} \tag{16}$$

Table 1. Comparative structure of forecasting configurations

Model Configuration	Neutrosophic Features (T, I, F)	Adaptive Weighting
Baseline RF	No	No
Neutrosophic RF	Yes	No
Weighted RF	Yes	Yes
WN-XGBoost	Yes	Yes

Note: T = truth; I = indeterminacy; F = falsity; RF = random forest; WN-XGBoost = Weighted Neutrosophic XGBoost.

The research study focuses on four forecasting configurations displayed in Table 1. The comparative framework enables the evaluation of uncertainty-aware learning through neutrosophic feature integration, adaptive weighting, and boosting-based optimization under sparse-data forecasting conditions.

To determine the predictive accuracy of the models, metrics such as the mean absolute error (MAE) and the root mean squared error (RMSE) provided in Eqs. (17) and (18), respectively, are used.

$$MAE = \frac{1}{n}\sum | \ y_i - \hat{y}_i \ | \tag{17}$$

$$RMSE = \sqrt{\frac{1}{n}\sum(y_i - \hat{y}_i)^2} \quad (18)$$

To ensure a fair comparison, all forecasting models were evaluated under the same experimental conditions. The models used identical training and testing partitions generated through rolling forward validation, while forecasting performance was assessed using the same evaluation metrics (RMSE and MAE). Hyperparameter settings were selected using moderate literature-supported values and kept fixed across all experiments to minimize overfitting and maintain consistent model comparison.

In addition to the error-based metrics, the Diebold–Mariano (DM) test was used to statistically compare the predictive accuracy of two competing models. The loss differential is defined as:

$$d_t = L(e_{1,t}) - L(e_{2,t})$$

where, $e_{1,t}$ and $e_{2,t}$ are forecast errors from two competing models, and $L(\cdot)$ is a loss function, typically the squared error loss $L(e_t) = e_t^2$.

The DM test statistic is computed using Eq. (19):

$$DM = \frac{\bar{d}}{\sqrt{\frac{\hat{\sigma}_d^2}{T}}} \quad (19)$$

where,

$\bar{d} = \frac{1}{T}\sum_{t=1}^{T} d_t$ is the mean loss differential,

$\hat{\sigma}_d^2$ is the variance of d_t,

T is the number of forecasts.

The following hypotheses apply

H_0: Equal predictive accuracy, $E[d_t] = 0$

H_1: Different predictive accuracy, $E[d_t] \neq 0$

Decision rule

•DM ~ $N(0,1)$ asymptotically

•Reject H_0 if:

$$| DM | > z_{\alpha/2}$$

The null hypothesis agrees with equal predictive accuracy; the DM statistic asymptotically follows a standard normal distribution. A statistically significant DM value indicates that one model provides superior forecasting performance relative to the other. Algorithm 1 provides a detailed procedure for obtaining the results.

Algorithm 1. Neutrosophic-based time-series forecasting procedure

Input:

Time-series data x_t, $t = 1, 2, \ldots, n$

Output:

Predicted values $\hat{x}_{t+1}$, neutrosophic components (T, I, F), and forecast intervals

Step 1: Data preparation

1. Load the dataset and sort observations by time
2. Generate lag variables: ○lag1 $= x_{t-1}$; ○lag2 $= x_{t-2}$
3. Compute change: $\delta_t = | x_t - x_{t-1} |$

Step 2: Neutrosophic transformation

4. Compute global statistics: ○$x_{\max}, x_{\min}, \mu$
5. For each observation, compute: T_t, I_t, and F_t

Step 3: Dataset construction

6. Form feature vector: $X_t = (x_t, x_{t-1}, x_{t-2}, T_t, I_t, F_t)$
7. Define target: $y_t = x_{t+1}$

Step 4: Model training

8. Perform rolling forward validation using TimeSeriesSplit with five expanding validation folds.
9. For each validation fold:
   • Train the model using all observations contained in the expanding training window.
   • Predict the observations in the subsequent testing window (approximately five future observations).
   • Expand the training window by including the testing observations before proceeding to the next fold.
10. Repeat the procedure until the final fold is completed and aggregate all out-of-sample predictions for performance evaluation.

Step 5: Model evaluation

11. Predict on the test set
12. Compute performance metrics: RMSE, and MAE
13. DM significance tests

Step 6: Multi-step forecasting

13. Initialize with last observed values
14. For each future step $h = 1, \ldots, k$:
    •Update lag variables
    •Recompute T, I, F
    •Predict $\hat{x}_{t+h}$ recursively

Step 7: Confidence interval estimation

$$CI_t = \hat{Y}_t \pm 1.96\sigma$$

2.3 Rolling-forward validation strategy

To obtain an out-of-sample evaluation while preserving the temporal ordering of observations, an expanding-window rolling-forward validation strategy was implemented using the TimeSeriesSplit procedure in Scikit-learn. Unlike conventional k-fold cross-validation, this approach maintains chronological order and prevents future observations from being used in model training.

The dataset comprises 35 annual observations (1990–2024). Following the construction of the one-step-ahead target and two lagged predictors, three observations with incomplete lag or target information were removed, leaving 32 usable observations. Five expanding validation folds were subsequently constructed. The first fold used seven observations for initial model training and the next five observations for testing. In each subsequent fold, the training window was expanded by incorporating the preceding testing observations, while the next five observations were reserved for evaluation. This procedure generated 25 out-of-sample forecasts across the five validation folds.

For the machine-learning models incorporating lagged predictors, forecasts within each testing window were generated using the corresponding observed lag values as they

became available. Thus, each prediction represented a one-step-ahead forecast rather than a recursively generated five-step forecast from the beginning of the testing window. This evaluation procedure is distinct from the separate recursive forecasting exercise for 2025–2029, in which each predicted value was used to construct the lagged inputs for the subsequent forecast. The ARIMA model generated forecasts over the testing-window horizon from the corresponding training-window endpoint.

For performance assessment, predictions from the five validation folds were pooled to calculate the overall RMSE and MAE. Therefore, the reported metrics represent aggregate out-of-sample performance rather than separate fold-level results. Given the small number of observations in each testing window, the resulting performance estimates should be interpreted cautiously.

All competing forecasting models, including Baseline RF, Neutrosophic RF, Weighted RF, Neutrosophic support vector regression (SVR), ARIMA, Prophet, Neutro long short-term memory (LSTM), Neutrosophic Linear Regression, and the proposed Weighted Neutrosophic XGBoost (WN-XGBoost), were evaluated using the same validation framework. Using an identical validation strategy for all models ensured a fair comparison of forecasting performance and avoided optimistic bias from random data partitioning.

The expanding-window strategy is particularly suitable for sparse macroeconomic time-series forecasting because it maximises the amount of information available for model estimation at each iteration while preserving the sequential nature of economic observations. Consequently, the results reflect ordered out-of-sample forecasting performance within the available sample; however, the small number of observations in each testing window reduces the precision of the resulting performance estimate.

2.4 Software libraries and computational environment

The algorithm of the proposed neutrosophic uncertainty-guided framework and the competing models was implemented in Python [29], with various computing and machine learning libraries as its core components. The libraries combine strong performance with consistent results, and they function as standard tools for machine learning modelling and time-series prediction. The ‘Pandas’ library [30] provides efficient data manipulation and analysis. The library ‘NumPy’ facilitates numerical computations, such as matrix operations and neutrosophic transformations, including multidimensional arrays at high performance [31].

The Scikit-learn framework served as the main toolkit, which enabled the execution of multiple machine learning algorithms [32]. Scikit-learn provides useful model functions, such as 'RandomForestRegressor' for bagging-based ensemble learning, 'GradientBoostingRegressor' for boosting-based prediction, and 'LinearRegression'. The training windows for rolling forward (rolling-forward) validation were implemented with 'TimeSeriesSplit', which allowed training windows while keeping unseen data for testing purposes to sustain time-based relationships and prevent data leakage.

Another library is Matplotlib [33], used to generate graphs and performance comparison graphs. To achieve reproducibility of the results, a random seed of 42 was used to manage all random activities throughout the ensemble learning procedures and data distribution techniques. The data processing included several stages, such as data preparation, lag production, and neutrosophic component computation (T, I, and F), weighted learning scheme construction, ensemble model development, evaluation, multi-step forecasting, visualization, and analysis of results. Table 2 presents the hyperparameter configuration and experimental settings.

Table 2. Hyperparameter configuration and experimental settings

Component	Parameter	Configuration	Selection Strategy/ Justification
Validation Strategy	Cross-validation	Rolling Forward Validation	It preserves the chronological order of observations and prevents information leakage in forecasting.
	Validation algorithm method	TimeSeriesSplit	For sequential economic time-series where future observations cannot be used during training.
	Training window	Expanding window	The training dataset increases after each iteration, for the model to learn from all previous observations.
	Testing window	Approximately 5 observations per fold	Provides sufficient out-of-sample evaluation while maintaining an adequate training sample for sparse data.
	Data leakage prevention	Future observations excluded from training	To prevent future observations from being excluded from training.
Reproducibility	Random seed	42	Fixed throughout all experiments to ensure reproducibility and consistent model comparison.
Feature Engineering	Lag variables	lag1, lag2	Two lagged observations were selected to capture short-term temporal dependence while avoiding excessive dimensionality given the limited sample size.
	Neutrosophic features	T, I, F	Included to explicitly represent reliability, uncertainty, and deviation in the learning process.
	Feature scaling	MinMaxScaler	Applied to normalize predictor variables and improve convergence of machine learning models, particularly LSTM and SVR.
RF	Number of trees (n_estimators)	100	A commonly recommended value providing stable ensemble performance without high computational cost.
	Maximum tree depth	Default (unrestricted)	Allows trees to adapt naturally while ensemble averaging controls overfitting.
	Bootstrap sampling	Enabled	Standard RF procedure to improve model diversity.
	Feature selection	$\sqrt{p}$ features per split	Default RF strategy to reduce tree correlation and improve generalization.
	Random state	42	Ensures reproducibility of bootstrap sampling and tree construction.

Weighted RF	Sample weights	Neutrosophic adaptive weights	Sample importance determined using the proposed weighting function (w=T(1–I)(1–F)).
	Number of trees	100	Maintained identical to the Baseline RF to ensure fair comparison.
	Bootstrap sampling	Weighted bootstrap	Observations sampled according to neutrosophic reliability weights.
	Random state	42	Ensures reproducibility of weighted sampling.
XGBoost	Number of estimators	100	Provides sufficient boosting iterations while avoiding overfitting in a small dataset.
	Learning rate	0.05	A conservative learning rate selected to improve generalization and stabilize boosting updates under sparse data.
	Maximum tree depth	3	Restricts model complexity to reduce overfitting while capturing nonlinear relationships.
	Objective function	Squared Error Regression	Appropriate for continuous price level index forecasting.
	Subsample ratio	1.0	All available observations were used because of the limited sample size.
	Column sampling	1.0	All predictor variables retained due to the small feature space.
	Regularization	Default L2 regularization	Prevents unnecessary model complexity during boosting.
	Random state	42	Ensures reproducibility across all boosting iterations.
WN-XGBoost	Sample weighting	Adaptive neutrosophic weighting	Observations weighted according to truth, indeterminacy, and falsity values during optimization.
	Learning rate	0.05	Same as standard XGBoost to isolate the contribution of the proposed weighting mechanism.
	Number of estimators	100	Same as baseline XGBoost for fair experimental comparison.
	Maximum depth	3	Maintained constant across experiments to control model complexity.
	Random state	42	Ensures reproducibility.
SVR	Kernel	RBF	Due to its ability to model nonlinear relationships commonly observed in macroeconomic time series.
	Regularization parameter (C)	Default	Default setting retained to avoid extensive tuning under sparse data conditions.
	Gamma	Scale	Standard implementation that automatically adjusts to feature variance.
ARIMA	Model order	(1,1,1)	As a classical benchmark following standard Box–Jenkins modelling for non-stationary economic series.
	Seasonal component	None	Annual data did not exhibit sufficient seasonal observations to justify seasonal modelling.
Prophet	Trend model	Additive	Appropriate for annual macroeconomic data with long-term trend.
	Changepoints	Automatic	Prophet's default changepoint detection retained.
	Yearly seasonality	Disabled	Annual observations do not provide sufficient information for seasonal estimation.
LSTM Network	Hidden units	32	Moderate network size chosen to balance learning capacity and overfitting risk in sparse datasets.
	Activation function	ReLU	Provides computational efficiency and stable gradient propagation.
	Optimizer	Adam	Widely used adaptive optimizer with good convergence properties.
	Learning rate	0.001	Standard value providing stable optimization.
	Epochs	100	Sufficient iterations for convergence while monitoring overfitting.
	Batch size	4	Small batch size selected because of the limited number of observations.
	Loss function	MSE	Standard loss function for continuous forecasting problems.
Evaluation	Performance metrics	RMSE, MAE	Selected to evaluate both average and large forecasting errors.
	Statistical comparison	DM Test	To determine whether forecasting performance differences between competing models are statistically significant.

Note: T = truth; I = indeterminacy; F = falsity; RF = random forest; ARIMA = autoregressive integrated moving average; LSTM = long short-term memory; MSE = mean squared error; RBF = radial basis function; RMSE = root mean squared error; MAE = mean absolute error; SVR = support vector regression; WN-XGBoost = Weighted Neutrosophic XGBoost.

The study used fixed machine learning model parameters during rolling forward validation tests to maintain consistent experimental results by selecting hyperparameters that would maximize prediction accuracy while maintaining low model complexity during forecasting performance with limited data.

Consequently, the study used a fixed random seed of 42 across all stochastic learning phases to enable the replication of forecasting results. Moderate model complexity configurations were adopted to minimize overfitting associated with the use of extreme hyperparameter optimization methods with limited macroeconomic data.

The hyperparameter values were selected based on established recommendations in the machine learning literature and preliminary empirical experimentation while maintaining moderate model complexity. Because the dataset consists of only 35 annual observations, extensive hyperparameter optimization (e.g., grid search or Bayesian optimization) was intentionally avoided to reduce the risk of overfitting. Fixed hyperparameter settings were maintained throughout the rolling forward validation procedure to ensure

fair comparison among competing models and to enhance the reproducibility of the experimental results.

The rationale for selecting the benchmark model. ARIMA(1,1,1) was selected because it is one of the most widely used classical forecasting models for economic time series and provides a standard statistical benchmark. Prophet was selected because it is a modern forecasting framework designed for trend-based time series and has been widely adopted in practical forecasting applications. The benchmark models were selected to represent both traditional statistical forecasting methods and contemporary machine learning-based forecasting approaches.

Given the limited dataset of 35 annual observations, the LSTM architecture was intentionally kept simple by employing a single hidden layer with 32 units, a small batch size, and a moderate learning rate. In addition, rolling forward validation was adopted to evaluate genuine out-of-sample forecasting performance, thereby reducing the risk of overfitting and optimistic performance estimates.

2.5 Data description

To determine the strength of various models in modelling economic data based on neutrosophic methodology, the P PLI data was used. The PLI measures cross-country price level differences by indicating the amount of U.S. dollars required to purchase an equivalent volume of goods and services. The PLI is the ratio of a PPP conversion factor to the corresponding market exchange rate between two countries, expressed relative to a base country that is set equal to 100.

The PLI data were obtained from the World Bank's World Development Indicators database [34].

The data include annual PLI data for Nigeria relative to the United States from the year 1990 to 2024, with 35 observations for the household and non-profit institutions serving households (NPISHs) final consumption expenditure, a univariate time series that exhibits both time-based changes and economic instability. The annual frequency and limited sample size reflect the practical reality of macroeconomic forecasting in many developing economies where high-frequency observations are unavailable, inconsistent, or unreliable.

3. RESULTS

This section presents the empirical findings of the study. The results present three distinct parts, which include descriptive and exploratory statistics, model performance evaluation, and multi-step-ahead forecasting results. The assessment of model performance uses standard metrics, which include RMSE, MAE, and DM significance tests. The study compares ten different models, which include baseline, neutrosophic, weighted models, and other competing models, to determine how uncertainty affects their predictive accuracy. The forecast results are used to analyze the future PLI movement. Table 3 provides the descriptive statistics for the PLI.

The descriptive statistics in Table 3 show a moderate variability for the PLI, with a mean of 48.28. The standard deviation (23.99) indicates a wide variation from the mean value. The minimum and maximum values are 17.97 and 122.82, respectively, showing a wide price range; therefore, the market experienced substantial price variations during the observed period. The skewness is 1.94, indicating that the distribution is positively skewed; therefore, the high PLI values occur infrequently yet attain extreme points at the upper end. The series exhibits periods of sharp price increases in markets.

Table 3. Descriptive statistics for the price level index (PLI)

Statistic	Value
Mean	48.28
Standard Deviation	23.99
Minimum	17.97
25th Percentile	33.54
Median (50%)	40.69
75th Percentile	52.29
Maximum	122.82
Skewness	1.94
Kurtosis	3.54

The kurtosis value of 3.54 indicates that the distribution is leptokurtic, which generates more substantial tail extensions and produces a peak that exceeds the normal distribution. The dataset contains extreme values that match the pattern of economic fluctuations. A robust modeling technique is required to capture the non-normal distribution through its asymmetric patterns and extreme variable changes, and to model uncertainty; hence, the need for a neutrosophic framework. The exploratory plots 1 to 4 further justify the need for neutrosophic methods. Figure 1 provides the temporal evolution of PLI.

The time-series plot in Figure 1 shows that the PLI experiences multiple instances of nonlinear changes and structural breaks throughout the observed period. The economy shows two distinct phases with extreme economic fluctuations, which start with a sharp mid-to-late 1990s increase and end with an abrupt drop. The system shows irregular movement and different trend characteristics, which create uncertainty and instability, thus making it necessary to apply uncertainty-aware neutrosophic modeling methods. Figure 2 presents the distribution of PLI values.

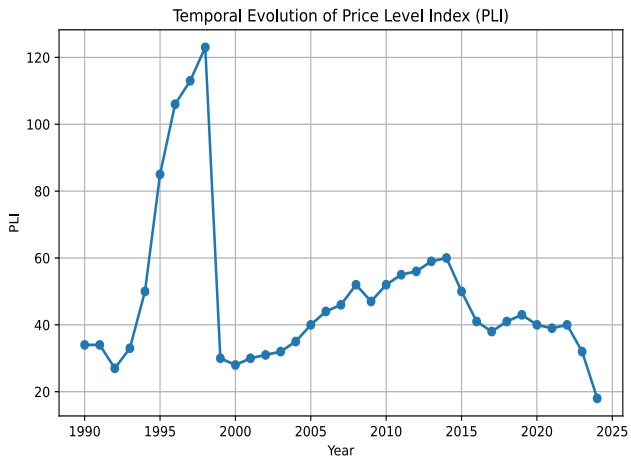

Figure 1. Temporal evolution of price level index (PLI)

The PLI series plot in Figure 2 shows positive skewness with heavy tails and extreme values. Most values exist within common value ranges, but a small number of highly exceptional values create distribution asymmetry. The distribution shows non-normal patterns because of extreme values, which create an uncertain environment that requires

neutrosophic uncertainty-aware modeling methods. Figure 3 provides the rolling mean and volatility of the PLI plot.

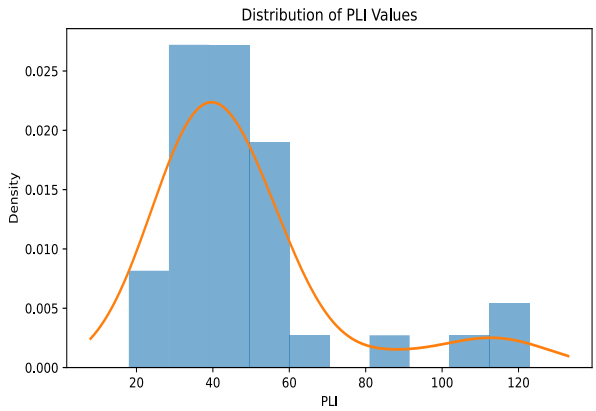

Figure 2. Distribution of price level index (PLI) values

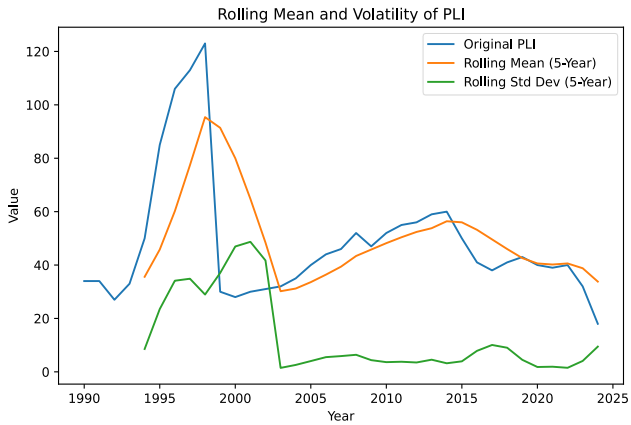

Figure 3. Rolling mean and volatility of price level index (PLI)

The plot in Figure 3 shows rolling statistics that demonstrate that both the average and fluctuating patterns of the PLI series exhibit permanent changes throughout the entire period of observation because the series shows non-stationary characteristics and non-constant volatility patterns. The late 1990s show a significant increase in volatility, which corresponds to two separate economic crises and two different structural changes. The observed pattern of changing variance with unstable trend development further necessitates the use of uncertainty-aware modeling methods, such as neutrosophic analysis.

The delta plot in Figure 4 shows periods of abrupt changes in the PLI series, with extremely large fluctuations occurring during the late 1990s. The dataset shows sharp temporal variations, which indicate economic instability and rising uncertainty. The irregular movement pattern supports the need for neutrosophic methods to effectively capture uncertainty and volatility in forecasting.

The T, I, and F metrics show how data reliability, uncertainty, and deviation from normality behave throughout time. The relationship between T, I, and F was developed throughout the 35-year period. Figure 5 shows the visualization of the T, I, F relationship over the observed period.

Figure 5 provides a graphical display of the relationship between values of T, I, and F. The T component reaches high values that are near 1 almost throughout the time frame, which shows that the neutrosophic data remains reliable and stable throughout the entire period. The I component shows low values for most years but experiences a sudden increase during the late 1990s, indicating a temporary period of increased uncertainty and instability. The assessment period shows low values for the F component, exhibiting a typical pattern interruption that occurs sometimes.

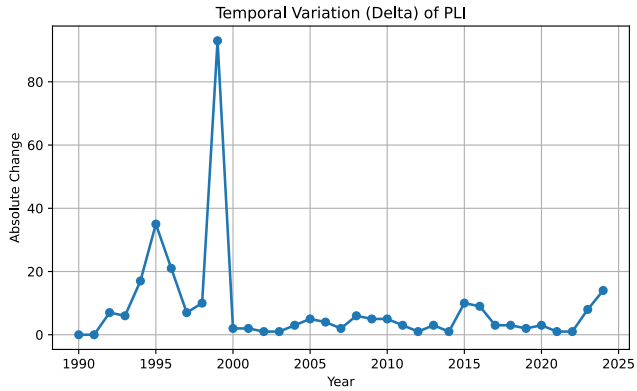

Figure 4. Temporal variation (delta) of price level index (PLI)

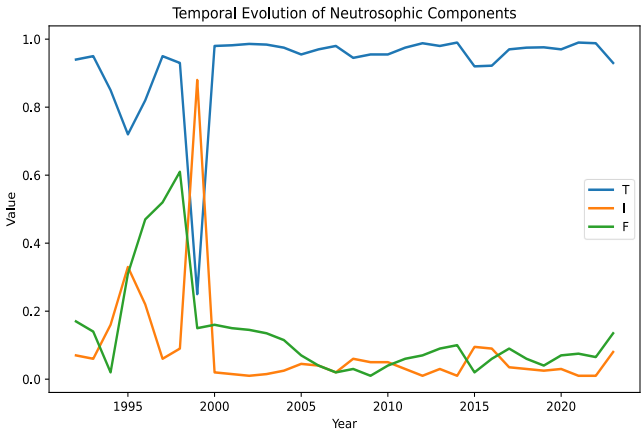

Figure 5. Temporal development of neutrosophic truth (T), indeterminacy (I), and falsity (F) components

The late 1990s saw an economic disruption, which caused both I and F to move suddenly. The two components maintain their low levels after the transition period, while T maintains its high levels, indicating that the data has become more stable. The plot in Figure 5 shows that the neutrosophic time series data remains stable except for uncertainty periods, which proves that neutrosophic modeling effectively represents different time segments.

3.1 Model performance evaluation results

To evaluate the performance of the nine different models discussed in this study, the RMSE and MAE were used; the results are presented in Table 4.

The results in Table 4 demonstrate that XGBoost has a lower forecasting error than the competing models because it produced the lowest RMSE of 10.964 and MAE of 8.250, which indicates that WN-XGBoost demonstrated competitive forecasting performance. The weighted framework demonstrated competitive performance while incorporating uncertainty-aware learning. The Neutro LSTM results would

be interpreted cautiously because the dataset contains only 35 annual observations. Although a deliberately simple architecture was adopted to limit model complexity, the small training sample restricts the stability with which a deep learning model can learn temporal patterns. Accordingly, LSTM is treated primarily as a comparative benchmark rather than a basis for concluding the suitability of deep learning for sparse annual economic forecasting.

Although XGBoost achieved the lowest forecasting errors based on RMSE and MAE, DM significance tests presented in Table 5 are further used to determine whether XGBoost significantly outperformed the competing models at the 5% level of significance.

Neutrosophic is shortened to Neutro. The DM test results provided in Table 5 show that WN-XGBoost produced lower forecasting errors than most competing models during rolling forward validation testing. The WN-XGBoost model showed better performance than Neutrosophic SVR, Baseline RF, Neutro, ARIMA, and Prophet because their *p*-values were less than 0.05. On the other hand, there were no significant differences between XGBoost and Linear Regression, Neutro-RF, Weighted Neutrosophic RF, and Neutro LSTM because their *p*-values were greater than 0.05. The results indicate that WN-XGBoost demonstrated more stable forecasting behavior under uncertain and sparse observational conditions, while Neutrosophic-based RF models and Neutro LSTM models showed similar predictive accuracy in sparse-data economic forecasting. Table 6 presents the WN-XGBoost forecasts for 2025–2029

Table 4. Model predictive performance with rolling forward validation

Rank	Model	RMSE	MAE
1	WN-XGBoost	10.963767	8.250329
2	Neutro LSTM	15.039517	11.877759
3	Neutrosophic RF	15.452448	11.406330
4	Baseline-RF	15.484346	12.930458
5	Weighted Neutrosophic RF	16.362983	12.184133
6	Neutrosophic SVR	26.288776	19.197008
7	ARIMA	48.967549	27.708595
8	Prophet	73.352355	39.327229
9	Neutrosophic Linear Regression	108.145261	38.844659

Note: RF = random forest; SVR = support vector regression; LSTM = long short-term memory; WN-XGBoost = Weighted Neutrosophic XGBoost; ARIMA = autoregressive integrated moving average; RMSE = root mean squared error; MAE = mean absolute error.

Table 5. Diebold–Mariano (DM) significance tests with Weighted Neutrosophic XGBoost (WN-XGBoost) as reference model

Compared Model	DM Statistic	*p*-Value	Significant at 5%
Neutrosophic Linear Regression	−1.235	0.217 > 0.05	No
Neutrosophic SVR	−2.481	0.013 < 0.05	Yes
Baseline RF	−2.503	0.012 < 0.05	Yes
WN-XGBoost	−2.669	0.008 < 0.05	Yes
Neutro RF	−1.387	0.165 > 0.05	No
Weighted Neutrosophic RF	−1.712	0.087 > 0.05	No
ARIMA	−2.527	0.012 < 0.05	Yes
Prophet	−2.397	0.017 < 0.05	Yes
Neutro LSTM	−1.275	0.202 > 0.05	No

Note: RF = random forest; SVR = support vector regression; LSTM = long short-term memory; ARIMA = autoregressive integrated moving average; RMSE = root mean squared error; MAE = mean absolute error.

Table 6. Price level index (PLI) forecasts for 2025–2029 using Weighted Neutrosophic XGBoost (WN-XGBoost)

Year	Forecast	Lower 95% Confidence Interval	Upper 95% Confidence Interval
2025	20.48	7.22	33.74
2026	39.83	26.57	53.09
2027	51.05	37.79	64.31
2028	82.20	68.94	95.46
2029	83.40	70.14	96.66

The comparative model results provide an exploratory assessment of the role of neutrosophic information in the proposed framework. In particular, the comparison between models incorporating neutrosophic components as features and the WN–XGBoost configuration provides an indication of whether integrating these components into the learning process is associated with additional predictive improvement. The results favor the proposed weighted configuration within the Nigerian PLI application considered. However, this comparison should not be interpreted as a comprehensive ablation of all possible weighting functions. Alternative formulations of the weighting function may produce different results and were not exhaustively evaluated in the present study because of the extremely limited number of annual observations. Consequently, the findings provide application-specific support for the proposed weighting strategy rather than establishing its universal superiority over alternative weighting mechanisms.

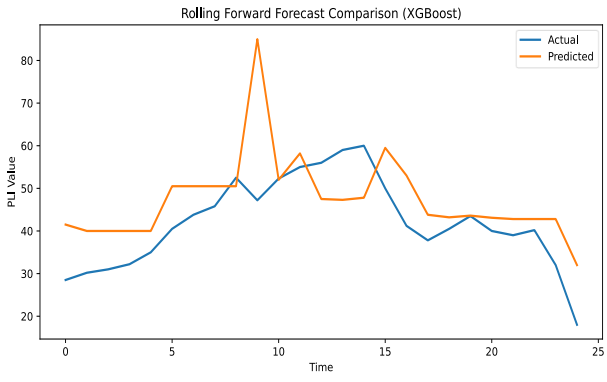

Figure 6. Rolling forward forecast comparison between actual and predicted price level index (PLI) values using Weighted Neutrosophic XGBoost (WN-XGBoost)

Table 6 shows that PLI forecasts rise steadily from 20.48 in 2025 to 83.40 in 2029, which indicates that Nigerian price levels will be below the U.S. threshold, though the price will increase over the years. The increasing relative prices will result in inflationary pressures, which are not advantageous to

international cost competitiveness. The wide confidence intervals demonstrate increasing uncertainty throughout the period, suggesting that researchers should handle long-term predictions with caution. Figure 6 provides the rolling-forward comparison between actual values and predicted values.

Figure 6 presents the rolling-forward forecast comparison between actual and predicted PLI values using WN-XGBoost. The WN-XGBoost model shows satisfactory performance through its ability to capture the PLI time series throughout the entire prediction period. The actual observation data shows an upward and downward movement pattern that the predicted values align with, demonstrating strong forecasting performance through the rolling validation test. The model shows increased prediction accuracy for standard economic conditions, but experiences high predictive values during times of sudden economic fluctuations, which occur when extreme market conditions arise. The predicted series shows only minor deviations from actual data because it stays mostly constant throughout the entire prediction period. The results confirm that the WN-XGBoost framework provides robust forecasting performance for uncertainty-aware economic time-series prediction. Figure 7 presents the correlation heatmap of the forecasting features.

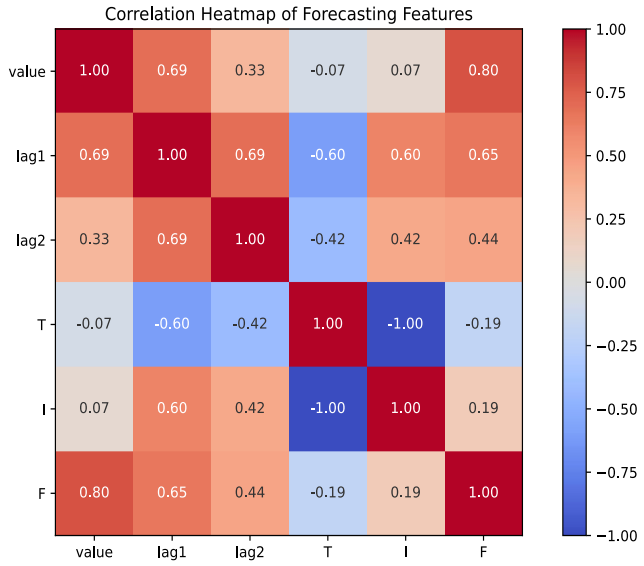

Figure 7. Correlation structure between lagged variables and neutrosophic truth (T), indeterminacy (I), and falsity (F) components

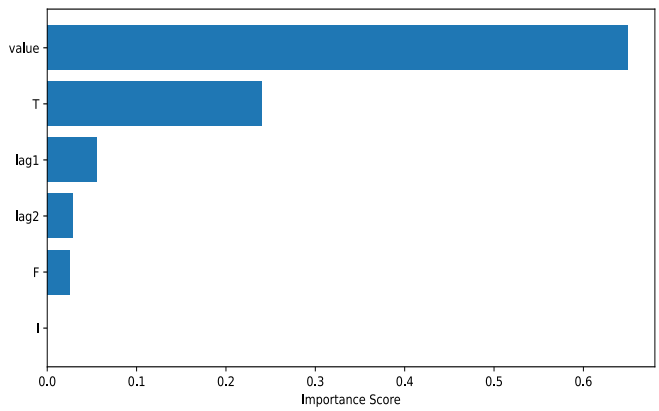

Figure 8. Weighted Neutrosophic XGBoost (WN-XGBoost) feature importance scores for lagged variables and neutrosophic components

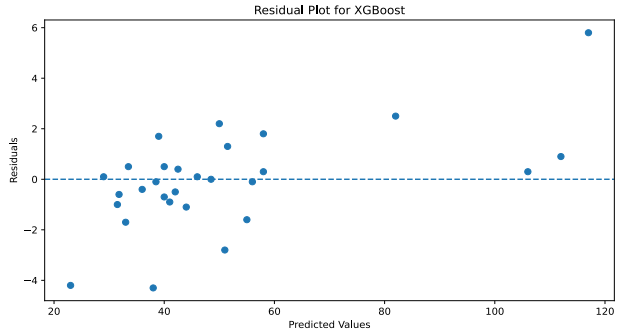

Figure 9. Residual distribution of Weighted Neutrosophic XGBoost (WN-XGBoost) predictions for price level index (PLI) forecasting

The correlation heatmap provided in Figure 7 demonstrates how neutrosophic components successfully capture essential uncertainty patterns while delivering additional insights about the time-series data. The PLI value shows strong positive correlations with lag1 (0.69) and falsity F (0.80) because past data and price deviations from mean values determine future price movements. The time series shows temporal persistence through its lag variables, which suggests strong interdependence. The T component has a strong negative relationship with indeterminacy. The data were sourced from a trusted platform that provides reliable information. The data demonstrate moderate positive relationships with lagged values, which indicates that volatility grows during periods of larger temporal changes. Figure 8 provides the WN-XGBoost feature importance.

The feature importance plot shows that the PLI value serves as the dominant predictor for the future value of the PLI, as shown in Figure 8. The T component of Neutrosophy shows the second-highest importance score because data reliability directly impacts prediction accuracy. The series exhibits temporal dependence because lagged variables (lag1 and lag2) show a moderate impact on forecasting accuracy. The falsity component F shows a minor yet significant effect on results, while the indeterminacy component I has little impact on final predictions. The results show that introducing neutrosophic reliability data enables the model to forecast unpredictable patterns with greater accuracy. Figure 9 displays the WN-XGBoost residual plot.

The Residual distribution of WN-XGBoost predictions for PLI forecasting based on the WN-XGBoost model in Figure 9 shows that the model effectively predicts the overall time series pattern, indicating that prediction errors are distributed closely around the zero-reference line. The residual distribution shows no major forecasting bias because the results display random behavior across all prediction intervals. Ideally, models exhibit reduced forecasting accuracy during economic volatility because larger residuals occur at higher predicted values. The economic dataset demonstrates that sparse-data samples combined with changing economic conditions will increase forecasting errors. Despite the volatility and sparse data, the WN-XGBoost framework maintains steady prediction accuracy because its residual distribution demonstrates consistency despite unpredictable elements and complex system behavior.

4. DISCUSSION

The strength of the weighted neutrosophic model was

assessed along with 8 competing models for their predictive accuracy with metrics such as the RMSE and MAE. The PLI series shows nonlinear fluctuations together with structural changes and extreme variations, which make forecasting more difficult. Multi-step recursive forecasting adds to error accumulation, which results in reduced variance explanation capacity. The forecasting performance was evaluated using RMSE and MAE, appropriate for assessing prediction accuracy under high volatility and sparse-sample economic time-series conditions.

The PLI forecast values show Nigeria's price levels will increase when compared to the United States, the base benchmark (PLI = 100). Nigeria shows short-term ability to maintain a lower-cost status from 2025, when its value starts at 20.48 and continues until 2026. The domestic price increase to 39.83 in 2026 suggests inflation pressures and exchange rate changes may have led to the increase. The 2027 PLI projection provides a forecast value of 51.05, which also does not exceed the base value of 100. The predictions indicate that Nigeria will remain an affordable market, though a more expensive one. The shift holds economic importance because it shows that international price competitiveness will decline while domestic consumers will face higher expenses. The economic pressure will continue to grow to 83.40 in 2029, which confirms the existing upward trend while showing that price levels face continuous upward pressure, making the economy start experiencing extended periods of high relative prices.

The confidence intervals are wide throughout the forecast period, starting with a narrow range that ends in 2025, then reaching their maximum width in 2029. This phenomenon demonstrates that rising uncertainty exists in multi-step forecasting because prediction errors are aggregated over the forecasting period, while volatility effects start to dominate. The short-term forecasts maintain their accuracy; however, the long-term projections need to be seen as general trends instead of exact numerical forecasts. This implies that a fundamental change has occurred in relative price movement, which will have major effects on inflation trends, exchange rate stability, and the country's ability to compete economically.

The Neutrosophic RF model showed good performance, ranking third after WN-XGBoost and WN-LSTM, with an RMSE value of 15.452 and an MAE value of 11.406. The Weighted RF model demonstrated competitive performance with the baseline RF model, with an MAE value of 12.184, and the weighting mechanism demonstrated competitive performance while incorporating uncertainty-aware learning. The DM test shows that the performance of WN-XGBoost is not significantly different from that of WN-RF, whereas it is significantly different from Baseline RF.

The WN-XGBoost model achieved the lowest forecasting errors, smaller than all competing models, with an RMSE of 10.964 and an MAE of 8.250. The results show that neutrosophic uncertainty works better when integrated into the learning process than when treated as a regular input feature. The WN-XGBoost model demonstrates better performance because it excels at detecting nonlinear patterns in sparse time-series datasets. The study shows that neutrosophic uncertainty competitively improves predictive models through adaptive weighting mechanisms that support the learning process instead of direct feature representation and improves predictive performance through adaptive weighting mechanisms. Observations with higher reliability (high T) receive greater influence during model training, whereas observations with higher uncertainty (high I) or deviation (high F) contribute less. This allows the model to prioritize reliable information while reducing the effect of unstable observations.

The neutrosophic components adapt to changing economic conditions. During relatively stable periods, high truth values and low indeterminacy and falsity produce larger sample weights. In contrast, periods of abrupt changes automatically receive lower weights, reducing their influence without discarding valuable observations. This characteristic is particularly beneficial for sparse macroeconomic datasets where removing observations would further reduce the already limited information available.

The superior performance of the WN-XGBoost model demonstrates the effectiveness of this adaptive weighting strategy. Compared with models that only incorporate neutrosophic features, the proposed framework achieved the lowest RMSE and MAE. The results indicate that integrating uncertainty directly into the learning objective provides greater predictive benefits than feature augmentation alone. The weighting mechanism therefore improves learning stability by emphasizing reliable observations throughout the boosting process, resulting in more robust forecasting under sparse and uncertain economic conditions.

The results indicate that incorporating neutrosophic uncertainty into the learning process is associated with improved forecasting performance for the Nigerian PLI series examined in this study. The comparative results provide application-specific support for adaptive weighting, although they do not establish that the proposed weighting formulation is universally superior to alternative weighting strategies.

The proposed WN-XGBoost consistently outperformed the competing models, suggesting that uncertainty-guided optimization can enhance predictive performance when historical observations are limited. However, these findings should be interpreted within the context of the dataset used in this study. Since the analysis was conducted using only the PLI, the observed forecasting improvements may not necessarily generalize to other macroeconomic indicators or more complex economic systems. Economic variables often exhibit distinct dynamics influenced by structural changes, policy interventions, and external shocks, which were beyond the scope of the present investigation. Consequently, the results primarily demonstrate the effectiveness of the proposed methodology rather than establishing broad economic conclusions. Future research should evaluate the framework using multivariate datasets, higher-frequency observations, and cross-country economic indicators to further assess its generalizability and robustness.

From an information-systems perspective, the proposed framework can serve as an uncertainty-aware forecasting component within a decision-support system. It transforms historical observations into neutrosophic measures and produces forecasts with prediction intervals to support economic monitoring, planning, and scenario analysis. However, full system implementation, including data acquisition, visualization, and user-interface components, is beyond the scope of this study.

The key takeaway from this study is that uncertainty-aware learning methods are essential when there are two conditions: when the sample size is small and when the data shows unstable patterns. The system learns to predict outcomes based on two factors: it uses the existing time-based information, and it assesses the trustworthiness of available data. The neutrosophic weighting mechanism functions as an

uncertainty regularization method that enables the model to decrease the impact of unreliable data on forecast outcomes, focusing on reliable learning patterns. The study results demonstrate that direct uncertainty integration into optimization methods enhances learning stability under conditions of limited data access.

A further limitation concerns the small number of observations available in each validation testing window. With approximately five annual observations per fold, individual forecast errors may have a relatively strong influence on fold-specific performance measures. Therefore, the aggregate RMSE and MAE should be interpreted cautiously as sample-specific measures of predictive performance rather than precise estimates of general forecasting ability.

5. CONCLUSION

The study shows that integrating neutrosophic uncertainty as a direct component of ensemble optimization enables stable economic predictions during periods of limited data and unstable market conditions. The results indicate that using uncertainty-based weighting methods can enhance learning stability under conditions where conventional large-data assumptions do not hold. Also, the multi-step forecasting results provide additional insights into the economic implications. The PLI forecast shows that the values fall below the benchmark level, though they reveal a significant upward trend in relative domestic price levels within the five years. The upward trend indicates that inflation rates may rise while international price competitiveness may decrease. The WN-XGBoost showed competitive predictive performance under sparse data through an adaptive weighting method.

The results indicate that uncertainty-aware learning improves forecasting accuracy compared with conventional statistical and machine learning models under sparse-data conditions. However, the findings should be interpreted cautiously because the empirical evaluation was based solely on the PLI. Consequently, the results primarily demonstrate the effectiveness of the proposed forecasting methodology rather than supporting broader economic conclusions. Future research should validate the framework using multivariate, higher-frequency, and cross-country datasets, while exploring alternative uncertainty quantification techniques to further improve forecasting robustness.

From an information systems perspective, the proposed framework demonstrates how uncertainty-aware data modelling and intelligent ensemble learning can strengthen forecasting-based decision-support systems operating under sparse and uncertain data environments.

REFERENCES

[1] Adesina, O.S., Obokoh, L.O. (2025). A hybrid framework of deep learning and traditional time series models for exchange rate prediction. Scientific African, 29: e02818. https://doi.org/10.1016/j.sciaf.2025.e02818

[2] Hyndman, R.J., Athanasopoulos, G. (2018). Forecasting: Principles and Practice, 2nd ed. OTexts, Melbourne, Australia.

[3] Hewamalage, H., Bergmeir, C., Bandara, K. (2022). Global models for time series forecasting: A simulation study. Pattern Recognition, 124: 108441. https://doi.org/10.1016/j.patcog.2021.108441

[4] Hewamalage, H., Bergmeir, C., Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. International Journal of Forecasting, 37(1): 388-427. https://doi.org/10.1016/j.ijforecast.2020.06.008

[5] Chen, T., Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, California, USA, pp. 785-794. https://doi.org/10.1145/2939672.2939785

[6] Dash, P.K., Dash, S., Liew, A.C., Rahman, S. (1995). A fuzzy engineering approach for time series forecasting of electric load. Journal of Intelligent & Fuzzy Systems, 3(4): 261-271. https://doi.org/10.3233/ifs-1995-3402

[7] Smarandache, F. (2013). Neutrosophic measure and neutrosophic integral. Neutrosophic Sets and Systems, 1: 3-7. https://philarchive.org/archive/SMANMA.

[8] Abdel-Basset, M., Chang, V., Mohamed, M., Smarandache, F. (2019). A refined approach for forecasting based on neutrosophic time series. Symmetry, 11(4): 457. https://doi.org/10.3390/sym11040457

[9] Guan, H., Dai, Z., Guan, S., Zhao, A. (2019). A neutrosophic forecasting model for time series based on first-order state and information entropy of high-order fluctuation. Entropy, 21(5): 455. https://doi.org/10.3390/e21050455

[10] Breiman, L. (2001). Random forests. Machine Learning, 45(1): 5-32. https://doi.org/10.1023/a:1010933404324

[11] Friedman, J.H. (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29(5): 1189-1232. https://doi.org/10.1214/aos/1013203451

[12] Ke, G., Meng, Q., Finley, T., et al. (2017). LightGBM: A highly efficient gradient boosting decision tree. In Advances in Neural Information Processing Systems 30 (NIPS 2017), pp. 3146-3154.

[13] Edalatpanah, S.A., Hassani, F.S., Smarandache, F., Sorourkhah, A., Pamucar, D., Cui, B. (2024). A hybrid time series forecasting method based on neutrosophic logic with applications in financial issues. Engineering Applications of Artificial Intelligence, 129: 107531. https://doi.org/10.1016/j.engappai.2023.107531

[14] Lim, B., Zohren, S. (2021). Time-series forecasting with deep learning: A survey. Philosophical Transactions of the Royal Society A, 379(2194): 20200209. https://doi.org/10.1098/rsta.2020.0209

[15] Adesina, O.S., Obokoh, L.O. (2026). Predicting the South African stock market using the hybrid of LSTM-ARFIMA model. In Springer Proceedings in Business and Economics, Cape Town, South Africa, pp. 41-53. https://doi.org/10.1007/978-3-032-13388-5_4

[16] Rusyana, A., Wigena, A.H., Sumertajaya, I.M., Sartono, B. (2024). Unifying variable importance scores from different machine learning models using simulated annealing. Ingénierie des Systèmes d'Information, 29(2): 649-657. https://doi.org/10.18280/isi.290226

[17] Berbiche, N., El Alami, J. (2023). Enhancing anomaly-based intrusion detection systems: A hybrid approach integrating feature selection and Bayesian hyperparameter optimization. Ingénierie des Systèmes d'Information, 28(5): 1177-1195. https://doi.org/10.18280/isi.280506

[18] Ye, J. (2014). Single valued neutrosophic cross-entropy for multicriteria decision making problems. Applied Mathematical Modelling, 38(3): 1170-1175. https://doi.org/10.1016/j.apm.2013.07.020

[19] Sharma, M., Kandasamy, I., Vasantha, W.B. (2021). Comparison of neutrosophic approach to various deep learning models for sentiment analysis. Knowledge-Based Systems, 223: 107058. https://doi.org/10.1016/j.knosys.2021.107058

[20] Rashno, E., Akbari, A., Nasersharif, B. (2022). Uncertainty handling in convolutional neural networks. Neural Computing and Applications, 34(19): 16753-16769. https://doi.org/10.1007/s00521-022-07313-2

[21] Wu, X., Ding, Y., Wang, L., Ding, D., Zhang, H. (2026). A novel neutrosophic transformer for handling imbalanced data. Journal of King Saud University - Computer and Information Sciences, 38: 476. https://doi.org/10.1007/s44443-026-00664-z

[22] Rahoumah, F., Yow, K.S., Nik Long, N.M.A., Gasim, M., Pham, H. (2025). Towards a new approach in complex neutrosophic soft ring theory. Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, 50(6): 1841-1859. https://doi.org/10.1177/18758967251380442

[23] Dovern, J. (2024). Eliciting expectation uncertainty from private households. International Journal of Forecasting, 40(1): 113-123. https://doi.org/10.1016/j.ijforecast.2023.01.002

[24] Luo, J., Klein, T., Walther, T., Ji, Q. (2024). Forecasting realized volatility of crude oil futures prices based on machine learning. Journal of Forecasting, 43(5): 1422-1446. https://doi.org/10.1002/for.3077

[25] Aastveit, K.A., Cross, J.L., van Dijk, H.K. (2022). Quantifying time-varying forecast uncertainty and risk for the real price of oil. Journal of Business & Economic Statistics, 41(2): 523-537. https://doi.org/10.1080/07350015.2022.2039159

[26] Martin, G.M., Frazier, D.T., Maneesoonthorn, W., et al. (2024). Bayesian forecasting in economics and finance: A modern review. International Journal of Forecasting, 40(2): 811-839. https://doi.org/10.1016/j.ijforecast.2023.05.002

[27] Mohmmed, K., Hussin, H., Ali, N., Tayfor, A. (2025). Practical applications of neutrosophic logic in enhancing the accuracy of economic forecasting models and supporting decision-making in banks. International Journal of Neutrosophic Science, 25(4): 282-294. https://doi.org/10.54216/IJNS.250424

[28] Abbas, A. (2026). Quantifying uncertainty in economic growth prediction using the neutrosophic Muth distribution. International Journal of Neutrosophic Science, 27(2): 95-109. https://doi.org/10.54216/IJNS.270209

[29] Python Software Foundation. (2023). Python (Version 3.11.2). https://www.python.org.

[30] McKinney, W. (2010). Data structures for statistical computing in Python. In Proceedings of the Python in Science Conference, Austin, Texas, USA, pp. 56-61. https://doi.org/10.25080/majora-92bf1922-00a

[31] Harris, C.R., Millman, K.J., van der Walt, S.J., et al. (2020). Array programming with NumPy. Nature, 585(7825): 357-362. https://doi.org/10.1038/s41586-020-2649-2

[32] Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12: 2825-2830.

[33] Hunter, J.D. (2007). Matplotlib: A 2D graphics environment. Computing in Science & Engineering, 9(3): 90-95. https://doi.org/10.1109/mcse.2007.55

[34] World Bank. (2024). Price level index (Households and NPISHs Final consumption expenditure). In World Development Indicators, World Bank (WB). https://data360.worldbank.org/en/indicator/WB_WDI_PA_NUS_PRVT_PLI, assessed on June 6, 2026.