Journal of Agriculture and Food Research 24 (2025) 102365

Contents lists available at ScienceDirect

Journal of Agriculture and Food Research

journal homepage: www.sciencedirect.com/journal/journal-of-agriculture-and-food-research

Towards improving farmers livelihood in Nigeria using food price forecasting

Divinefavour Odion [a], Joaquin Gajardo [b,e], Thijs Defraeye [b,d], Thomas Motmans [c], Kanaha Shoji [b], Roberta Evangelista [c], Daniel Onwude [b,*]

[a] Constructor University Bremen, Campus Ring 1, Bremen, 28759, Germany
[b] Empa, Swiss Federal Laboratories for Materials Science and Technology, Laboratory for Biomimetic Membranes and Textiles, Lerchenfeldstrasse 5, St. Gallen, CH-9014, Switzerland
[c] BASE Foundation, Elisabethenstrasse 22, 4051, Basel, Switzerland
[d] Food Quality and Design, Wageningen University & Research, P.O. Box 17, Wageningen, 6700 AA, the Netherlands
[e] Institute of Agricultural Sciences, ETH Zurich, Universitätstrasse 2, Zurich, 8092, Switzerland

A R T I C L E I N F O

Keywords:
Food price forecasting
Market intelligence
Online machine learning
Smallholders
Farmer's income
Food waste
Food loss

A B S T R A C T

Nigeria's agricultural sector represents approximately 25 % of the country's overall GDP and is a major source of employment for its population. This sector is largely driven by smallholder farmers who grow fruits and vegetables on farms under 4 ha. Despite their significant contribution to food production in Nigeria, most smallholder farmers, approximately 70 %, live in poverty, earning less than $1.9 per day. One of the key factors contributing to this situation is a lack of access to market price information. Farmers currently rely only on historical prices observed in local markets to decide on when, what, where and the price to sell their produce. This can lead to suboptimal decisions, resulting in food loss and loss of potential income. To address this challenge, we developed a machine learning online pipeline. It utilizes a Random Forest model trained on historical monthly fresh produce prices and other macroeconomic factors like currency exchange rates for Nigeria, that are regularly scraped from the internet. We deployed our trained model through an open-source mobile application, Coldtivate. Our model accurately predicted market prices for crops such as tomatoes, onions, potatoes, and plantains in various Nigerian states. The prediction success rate of our model varied across the various states in Nigeria. It ranged from 1 % to 20 % in Mean Absolute Percentage Error (MAPE) for predictions up to 8 months ahead. When evaluated on a hold-out test set, it yielded an RMSE of ₦45.16. The average MAPE of our model, when considering state-time-commodity averages, is up to 5 % lower than other baseline models, including the benchmark rolling-average, CatBoost, XGBoost, and SARIMA. By detecting patterns and trends in food prices, farmers can use our tool to make more informed decisions about when and what to sell to optimize profit, thereby improving revenue. Furthermore, our model provides a foundation for future machine learning model development in food price forecasting in agrarian countries.

1. Introduction

In Nigeria, agriculture is a vital sector, contributing to over 24 % of the country's total GDP and employing two-thirds of its labor force comprising mostly of farmers [1]. Most of these farmers are smallholders, estimated to be around 38 million, who grow crops on farms that are less than 4 ha in size [2]. Despite their contribution to food production in the country, over 72 % of these smallholder farmers live below the poverty line of $1.9 per day [3]. One of the key reasons contributing to this situation is the lack of market price intelligence among these farmers. The majority of fruits and vegetables produced by this category of farmers are from the Northern regions and largely sold in the Southern regions where a significant number of food markets are located [4]. At these markets, smallholders, traders, and food vendors sell fresh produce such as tomatoes, onions, bell peppers and so on directly to consumers at diverse price ranges. Prices vary across the country, influenced mostly by farm-market proximity and seasonality. As an example, prices for fresh produce are generally lower in markets closer to farming regions. This is a result of the cost of transporting goods from the farms to the market. Additionally, there is a lower

* Corresponding author.
E-mail addresses: thijs.defraeye@empa.ch (T. Defraeye), daniel.onwude@empa.ch (D. Onwude).

https://doi.org/10.1016/j.jafr.2025.102365
Received 6 July 2025; Received in revised form 29 August 2025; Accepted 14 September 2025
Available online 15 September 2025

demand in these areas compared to larger cities in the south. In these cities, higher demand drives up prices (See Fig. S5). In spite of this interesting dynamics, currently, there are no standard methods or defined approaches for estimating the future market prices of commodities produced by smallholder farmers in Nigeria. Smallholders often resort to reviewing the historical prices of a food observed in local markets in the recent past. This method is equivalent to having a naive model that predicts the future prices of fresh produce using only the prices from the immediate previous time steps. It does not account for external factors and anomalies such as fuel prices, foreign exchange rates, and inflation that may influence the changes in the prices of certain goods. All this makes that smallholder famers would sell their produce at unfavorable prices, leading to loss in revenue due to lack of adequate market price intelligence (access to existing and future market price) [5]. This highlights the need for a more accurate and holistic approach that considers relevant economic influences such as foreign exchange rates. It also needs to include variables such as domestic fuel prices and inflation rates. These considerations will lead to more accurate price predictions.

Several researchers have developed machine learning models with algorithms such as Autoregressive Integrated Moving Average (ARIMA), Long Short-Term Memory (LSTM), Support Vector regressor (SVR), Tree-based algorithms, and Neural networks. These models aim to predict the market prices of avocado, rice, wheat, corn, and so on. Ma et al. [6] presented a commodity price forecasting system to aid small and marginal farmers in India to make informed decisions on when to sell their harvest. The system integrated data from the Indian Ministry of Agriculture and Farmers Welfare's website Agmarket with machine learning algorithms to generate accurate price forecasts. The forecasts were presented in an easy-to-understand format that highlights the key historical pricing data that informs each forecasted trend and seasonal components [7]. They also presented a software-application led solution to help farmers optimize profits. The study compared a range of machine learning and deep learning methods, including Autoregressive Integrated Moving Average (ARIMA), Long Short-Term Memory (LSTM), Support Vector Regression (SVR), Prophet, and Extreme Gradient Boosting (XGBoost), to identify effective forecasting models for agricultural and economic data. While LSTM frequently emerged as the top performer (e.g., achieving a mean squared error of 0.304 in one study), these approaches generally shared two key limitations. First, they neglected critical external predictors such as domestic fuel prices, inflation, and foreign exchange rates, factors that strongly influence fresh produce prices. Second, they overlooked deployment considerations, particularly online or incremental learning strategies for continuous model updating in dynamic market environments.

To the best of our knowledge, few studies have explicitly incorporated such economic variables into agricultural price forecasting. No known study applied online machine learning models to predict the prices of tomato, potato, onion, and plantain in Nigeria up to eight months ahead. This gap is significant because external economic factors capture much of the true variance in price data, and without them, models risk poor generalization. Moreover, despite the growing literature on food and commodity price forecasting, there is little evidence on how these technologies can be deployed at scale to support smallholder farmers or integrated into agricultural development programs. Related work across supply chain forecasting, food price prediction, and stock market modeling further highlights these gaps. For example, Real et al. [8] improved supply chain demand forecasts with nonlinear ML methods, while Juan et al. [9] leveraged weather and sales data to estimate avocado sales. Other contributions include swarm intelligence–based models for food price forecasting [10], hybrid ML–DL frameworks for stock prices [11], and integrated ARIMA–ML models for staple crops [12]. While these approaches achieved strong predictive accuracy in their respective domains, they remain highly data-intensive, computationally costly, or limited by the exclusion of explanatory variables.

Building on these insights, our study takes a distinct approach. We address the challenges of limited historical data, only 75 monthly observations per state–commodity series since 2017, by designing a forecasting framework tailored to small datasets. This scarcity makes complex models like LSTMs or Transformers prone to overfitting, while also limiting their ability to capture temporal dynamics. To overcome this, we propose a robust forecasting tool that integrates economic predictors into a parsimonious model architecture, combined with an online learning framework for continuous retraining. Our solution is deployed through a publicly accessible mobile application, enabling the tracking of vital metrics and offering actionable forecasts to smallholder farmers. In doing so, our work makes three key contributions: (1) Novel integration of economic predictors, where we incorporate inflation, fuel prices, and exchange rates into fresh produce price forecasting. (2) Adaptation to data-scarce environments by designing models suited to short historical series while minimizing overfitting risks. (3) Practical deployment via online learning, where we develop and release a mobile-based forecasting tool that directly supports smallholders in Nigeria. To achieve these, we selected a Random Forest model, which is known to perform robustly under conditions of limited data and noisy predictors [13]. Random Forests do not require strong parametric assumptions, handle missing values and nonlinearities well, and are less sensitive to overfitting when compared to deep networks in small-sample regimes. This choice also reflects the practical realities of the data collation efforts in Nigeria, where improving the data infrastructure for food prices remains an ongoing challenge. Our system accurately predicts market prices for tomato, onion, plantain, irish potato (potato), and sweet potato in each state of Nigeria, up to 8 months into the future. There are 36 states and a Federal Capital Territory in the federal republic of Nigeria. These states are divided into six geopolitical regions: the North-East, North-West, North-Central, South-East, South-West, and South-South. Each of these regions is determined by a combination of geographical, historical, and cultural elements. The North-East region, for example, includes states such as Borno, Adamawa, and Yobe, while the North-West region encompasses states such as Kano, Kaduna, and Sokoto. The North-Central region includes states such as the Federal Capital Territory, Niger, and Kwara, while the South-East region is home to states such as Abia, Anambra, and Enugu. The South-West region comprises of Lagos, Oyo, and Ogun states, while the South-South region is made up of states such as Rivers, Delta, and Akwa Ibom.

We benchmark our model against a simple rolling-average baseline, SARIMA, and other tree-based algorithms including Catboost, and XGboost. Our model is made accessible via an open-source data science mobile application called Coldtivate [14] with an intuitive user interface, providing small-holder farmers with the real-time insights of market prices and trends to make informed decisions and improve their business returns. This research paper elucidates the advantages of a unified, consolidated model capable of discriminating between the 37 states, and precisely predicting forthcoming food prices for a diverse assortment of fresh produce vended in each state. Our research extends the literature by addressing overlooked methodological and practical challenges while offering a deployable solution to strengthen the resilience and livelihoods of smallholder farmers.

1.1. The state of the art of food price forecasting in low- and medium-income countries

Recent studies have increasingly highlighted the importance of robust forecasting tools for agricultural markets in low- and middle-income countries. In such markets, volatility in food prices directly affects household welfare and farmer livelihoods. A 2025 systematic review emphasizes that ensemble methods and deep learning models have become dominant approaches for forecasting staple crops such as rice, wheat, and corn, especially under conditions of sparse and noisy data that characterize many LMIC settings [15]. In parallel, large-scale comparative studies show that hybrid pipelines that combine time series methods with machine learning consistently outperform univariate baselines across commodities and markets [16]. In Nigeria specifically, new empirical work demonstrates that tree-based models such as XGBoost outperform recurrent neural networks when applied to long-horizon food price forecasting, reinforcing the view that ensemble approaches are particularly well suited to volatile African markets [17]. Other Nigerian studies have expanded the predictor set to include macroeconomic variables such as exchange rates and inflation, showing that these exogenous factors significantly improve accuracy over purely historical models [18].

At the global level, interpretability has become an equally important objective. The AGRICAF model integrates explainable machine learning with econometrics to deliver twelve-month forecasts for major commodities while retaining transparency of factor contributions, offering a pathway for operational adoption by non-specialist users [19]. Finally, adaptive food price inflation forecasting frameworks have demonstrated how exogenous drivers and automated model selection can strengthen public information systems and support budget planning, illustrating the policy relevance of methodological improvements [20]. Together, these developments show three clear trends shaping the field: ensemble and hybrid models are most reliable under data volatility, external macroeconomic drivers materially improve predictive performance, and interpretability is essential for farmer and policymaker adoption. These insights provide the foundation for the present study, which adapts tree-based forecasting to the Nigerian context while embedding external economic indicators and prioritizing practical usability for smallholders.

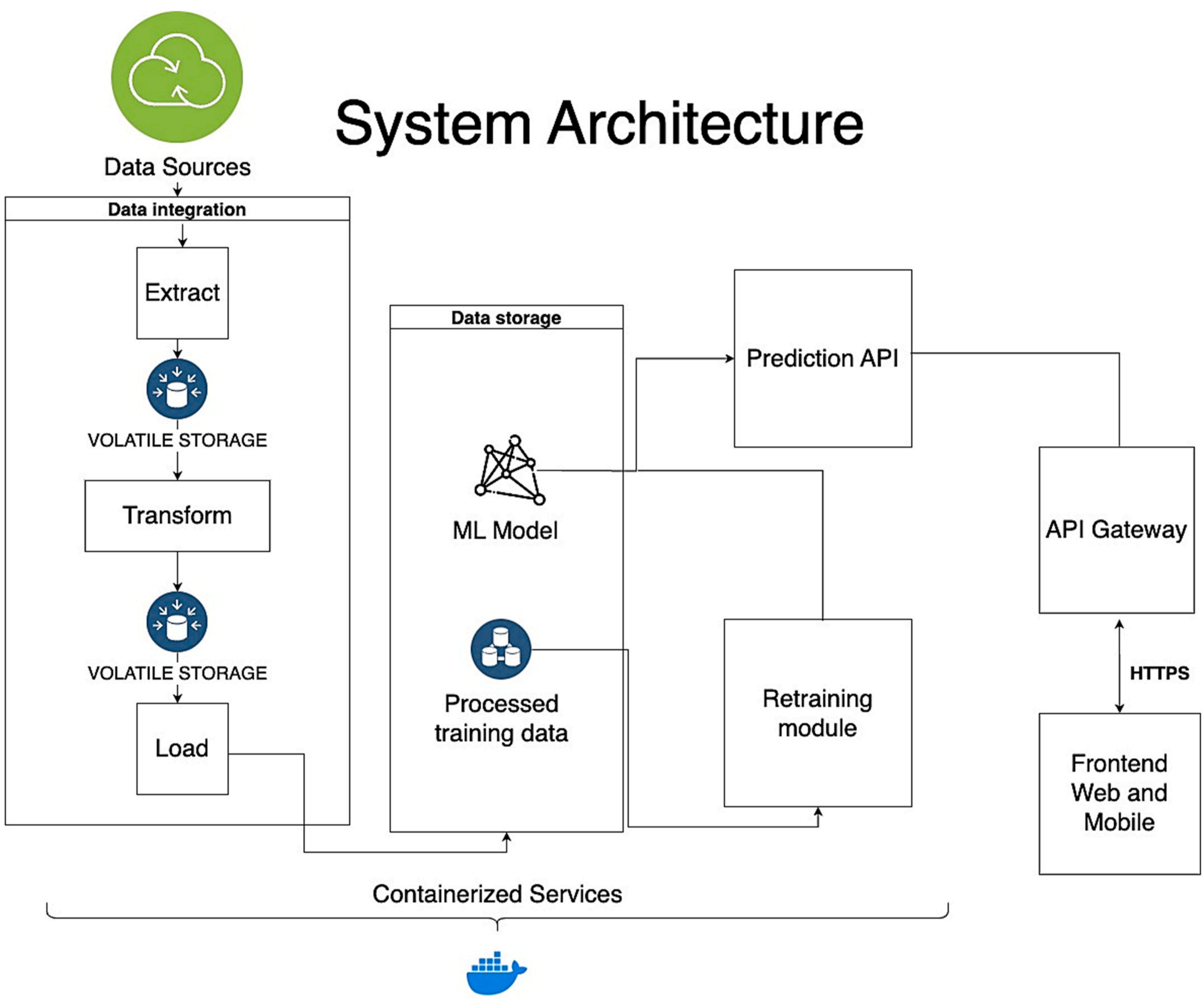

Fig. 1. Flowchart detailing the System Architecture of the food price forecasting tool.

1.2. The proposed methodology

A Random Forest model trained on the average monthly prices of 5 selected commodities from 2017 to 2022 was used to predict the market prices of tomato, onion, plantain, potato (Irish potato) and sweet potato. We selected these commodities because they are some of the more popular fresh produce sold by smallholders in Nigeria. In addition, there quite a number of historical open-source data for these fresh produce relative to others. This model includes 6 other predictors namely, Last Month Price, Last 5 Month Price, Consumer Price Indexes, USD to Naira exchange rates, Crude oil prices and State price roll-ups. A web-scraping pipeline was designed to obtain the food price data and external predictors. Data were obtained from three official repositories: (i) the official e-library site of the National Bureau of Statistics (NBS) [21], which publishes monthly "Selected Food Prices Watch" reports in Excel format; (ii) the official website of the Central Bank of Nigeria (CBN) [22], which provides daily ₦/US$ rates and other macroeconomic indicators in downloadable tables; and (iii) the Yahoo finances website [23]. The model is retrained monthly using this data. An open-source data science-based mobile application called Coldtivate [14], host the model. This mobile app allow users to choose from a list of fresh produce and states. Users can then view the 8-month price forecasts for the chosen food in their selected state.

2. Data and methods

This section discusses details about the characteristics of the training data, findings from the exploratory data analysis carried out on the data, the preprocessing steps to train machine learning models, and feature selection processes. A retraining module is connected to the scraping module and is used to regularly retrain the machine learning model on the most recently scraped data. A flowchart representation of the described methodology is given in Fig. 1.

The data for this study was obtained by scraping three online repositories and processed with the pipeline depicted in Fig. 1. An Extract Transform Load (ETL) architecture was utilized to gather and process the data into a suitable format for machine learning model (re)training. The processed data was subsequently loaded into the Data storage, along with the serialized machine learning model, thus completing the ETL process. The model is retrained on updated data at monthly intervals. A prediction Application Programming Interface (API) was developed to facilitate the querying of the model with prompts containing the target state and food names in JSON format. The model, in turn, provides 8-month forecast related to the query prompts to the mobile application's front-end via the interface. The proposed architecture and API have the potential to enhance the efficiency and timeliness of food price predictions, to the benefit of users.

2.1. Training data and features

The training data covered historical food prices (monthly average) from January 2017 until October 2022. A consolidated label encoded dataset with monthly averaged prices for all 5 commodities in all 37 states is used for training the machine learning models. The processed dataset consists of 9 features which are, Date, State labels, Food labels, Last Month Lagged prices, Last 5 months lagged prices, Consumer Price Indexes, USD to Naira Conversion rates, Crude oil prices and State roll-ups. The features used for training the models are grouped into four categories, which are:

1. Label features ('st_label', 'comm_label'): The label features consisting of state and food labels are categorical features that represent the 37 states and 5 commodities, which have been label encoded.
2. Lagged features (Last_Month_Price', 'Last_5M_Price'): The lagged features consist of 'Last_Month_Price' and 'Last_5M_Price' which represent the food prices shifted one and five times respectively (i.e for a given datapoint in the market price timeseries, the 'Last_Month_Price' feature represents the price of one food commodity in one state in the previous month and the 'Last_5M_Price' represents the price 5 months ago).
3. Exogenous features ('CPI', 'USDtoNaira', 'Crude_prices'): The Exogenous features include the 'Crude_price', 'USDtoNaira', and 'CPI' features. The Consumer Price Index ('CPI') feature consists of the monthly Consumer price indexes for all items classified under "Goods and services" at the country level. The CPIs are calculated with reference to a 2009 base period. The 'USDtoNaira' feature contains the monthly averaged USD to Naira exchange rates consistent with the timespan of the food prices. The 'Crude_prices' feature consists of official monthly Crude oil domestic prices in Naira per barrel.
4. Roll-up features ('state_roll'): The 'state_roll' feature represents the aggregated mean of the prices of all 5 commodities per month for each state.

2.1.1. Linearity of the data

To complement the exploratory data analysis, we carried out formal tests of linearity on each of the 185 state–commodity series. Following standard practice, the autoregressive order was chosen by the Akaike Information Criterion. We implemented a battery of well-established tests, namely the Keenan (1985) quadratic index LM-F test [24], the Luukkonen–Saikkonen–Teräsvirta (1998) smooth transition autoregressive LM test [25], the Ramsey RESET test (Ramsey, 1969) [26], and the Brock–Dechert–Scheinkman (1987) independence test on autoregressive residuals [27]. These tests are designed to capture different classes of nonlinear departures from a linear autoregressive structure.

Aggregate results (Table 1) show that 27 % of the series reject linearity at the 5 % level in at least one of the raw tests. By test, the rejection rates are 16 % for Keenan, 5 % for LSTR, 23 % for RESET, and 0 % for BDS. After applying the Benjamini–Hochberg correction at $q = 0.05$, the overall rejection rate falls sharply to 0.5 %. At the individual level, the RESET test shows 0.5 %, while Keenan, LSTR, and BDS yield 0 %. These results suggest that while pockets of nonlinearity exist, particularly under the Keenan and RESET diagnostics, the evidence becomes negligible after correcting for multiple testing. This proves that many food price dynamics can be approximated by linear models, yet our use of tree-based methods remains appropriate: they capture complex nonlinearities when present without sacrificing accuracy when dynamics are simple. For farmers, traders, and policy makers, this ensures that the forecasts they rely on are both robust and adaptable, providing dependable guidance across diverse market conditions. The full set of results, including the detailed series-level outcomes as well as state-level and commodity-level aggregations are available on request from the authors.

Table 1
Aggregate results of linearity tests across 185 state–commodity series. Entries report the share of series rejecting linearity at the 5 percent level before and after Benjamini–Hochberg (BH) adjustment, for each test in the suite (Keenan, LSTR, RESET, BDS).

N series	Aggregate (raw)	Aggregate (BH)	Keenan (raw)	Keenan (BH)	LSTR (raw)	LSTR (BH)	RESET (raw)	RESET (BH)	BDS (raw)	BDS (BH)
185	27 %	0.5 %	16.2 %	0 %	5.4 %	0 %	23.2 %	0.5 %	0 %	0 %

2.2. Methods

We conducted an Exploratory Data Analysis (EDA) on the training data to determine useful statistical properties for modeling, such as autocorrelation tests, location-based analysis, and multicollinearity tests on the food price time series. This preliminary step was essential for informing our approach to developing an accurate forecasting model. Furthermore, we conducted correlation tests for the exogenous variables to determine those that were worth including in the model to improve prediction accuracy. After concluding the EDA on the data and features, we trained various machine learning models and a statistical model which include CatBoost, Random Forest, XGBoost, and SARIMA. Finally, we compared the 4 models against a naïve rolling-average benchmark model to measure the advantages of machine learning in this problem. To ensure a fair comparison, each model was trained on the same dataset. We performed Grid Search cross-validation to tune each model's hyperparameters for the best fit. Subsequently, the models were evaluated on the hold-out test set using Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE) as the key metrics for determining the most suitable approach. The Random Forest model had the lowest overall Mean Absolute Percentage Error (MAPE) results (See Fig. 6) with an average of approximately 11 % MAPE and 45.16 RMSE for all commodities and states.

2.2.1. Data preprocessing

The data preprocessing consisted of two steps. First, the categorical columns containing the state and food names were converted into numerical features using the LabelEncoder class from the scikit-learn (sklearn) Python library [28]. This was necessary to ensure that the algorithms could effectively process the state names and food names. Then, missing values in the lagged feature columns (consisting of Last_Month_Price and Last_5M_Price) were filled in using linear interpolation. In the training dataset, the first five rows of the lagged price columns for each state-food combination contained missing values due to the absence of data from prior months, as the dataset started from January 2017. For example, 'Last_Month_Price' had one missing value because of a lack of data from December 2016, and 'Last_2M_Price' had two missing values due to the unavailability of data for November and December 2016. To address this issue, the food prices for each state in January 2017 were reverse linearly interpolated and inserted into the missing value columns.

The present study reports on two distinct experimental setups designed to train and evaluate machine learning models for the given task. In the first experiment, a fixed train-test split ratio was utilized to hold out a portion of the data as a single test set. The models were trained on the remaining data and evaluated on the held-out test set. In the second experiment, we employed a temporal cross-validation framework to evaluate the models on multiple test sets. Specifically, the data was split into several non-overlapping temporal folds, and the models were trained and evaluated on each fold iteratively. The final results were obtained by averaging the performance of the models across all folds and presented as a consolidated single metric for easy comparison between models.

2.2.1.1. Train and test split. In the first experiment (single test-set evaluation), the processed dataset was split into two parts for model training. Utilizing the date column of the dataframe and the Python libraries 'pandas' [29] and 'datetime', we partitioned the dataframe into two disjoint subsets. The first subset comprised a training set encompassing the time interval spanning from '2017-01′ to '2022-02,' and the second subset consisted of a test set containing the period between '2022-03′ to '2022-10'. The training set represented approximately 88 % of the entire dataset and the test set represented the remaining 12 % of the latest historical data in the dataframe. The standard practice in conventional machine learning approaches involves dividing the available dataset into train and test sets, where the test set comprises a larger proportion. However, for our forecasting objective, we faced a scarcity of open-source historical food prices. To overcome this limitation, we adjusted the ratio of test data in our study to a lower percentage than what is commonly used in conventional approaches. This split allows for a robust evaluation of the machine learning algorithms, where the model is trained on a significant portion of the data, and the remaining portion is used to test the model's generalization performance. All models were trained and fine-tuned on the training dataset, and then evaluated using MAPE and RMSE on the test dataset.

In the second experiment, we employed a temporal cross-validation methodology to evaluate the generalization performance of the machine learning models over time. To do this, we initially created a training set with monthly averaged food prices from January 2017 to December 2020 and designated the following eight months (January 2021 to August 2021) as the initial test set. We then performed temporal cross-validation by iteratively training all five models on the expanding training set and assessing their predictive accuracy on each successive test set. The Mean Absolute Percentage Error (MAPE) was used as the performance metric throughout this process. To account for the temporal dynamics of the data, we conducted sequential iterations of the experiment by increasing the size of the training set by 1 month per iteration and shifting the test set one month forward per iteration. This process was repeated for 26 train-test splits, resulting in 26 MAPE score sets for each model. To obtain a comprehensive evaluation metric, we computed the final MAPE score for each model as the average of the MAPE across all train-test splits. By leveraging the temporal cross-validation framework, our study provides a rigorous assessment of the models' predictive power over time, which is crucial for practical applications in the food market.

2.2.2. Machine learning modeling

In order to compare different machine learning algorithms for the specific task, we trained a set of 5 models. This included a benchmark 5-month rolling-average model, a SARIMA model, and three machine learning models which include Random Forest, XGboost and CatBoost. The performance of these models was then evaluated on a separate test set to determine which one was the best performing. In this evaluation, we used the Mean absolute percentage error (MAPE) and Root mean square error (RMSE) as the primary metrics to compare the accuracy of the different models. The 5 models included in the evaluation phase are briefly described below.

2.2.2.1. Naive model (benchmark). The naive model approximates the current approach employed by smallholders in Nigerian markets, which stems from utilizing historical price data from recent months as the foundation for their food price forecasts. We use a simple rolling-average algorithm. A single forecast from this model is calculated as the arithmetic average of the last 5 months of prices.

2.2.2.2. Random forest model. The Random Forest model is a popular tree-based model that implements gradient bagging (Fig. S3). It is a modeling approach of training several weak estimators in which a model forecast is computed as the arithmetic average of the predictions of all the weak estimators/trees in the model. We implemented the Random forest model with Python's 'sklearn.ensemble' library [30]. See Fig. S2 in the Supplementary Materials for the schematic diagram of gradient bagging algorithms. The prediction model is as follow:

$$\widehat{Y} = \frac{1}{K}\sum_{k=1}^{K} T_k(X) \quad (1)$$

where T is the individual decision tree in the forest, T_k represent each tree trained on a bootstrap sample of the training data and uses a random subset of features at each split, X is the input feature vector.

2.2.2.3. CatBoost model. Catboost is a novel gradient boosting tree-based algorithm developed by Yandex researchers [31]. Gradient boosting is an ensemble modeling technique that utilizes several weak learners/models in a sequential manner. Gradient boosting starts by fitting a weak learner to the data, then a second learner is fit to the same data with the aim of accurately predicting the cases where the previous learner performed poorly, and this process is repeated for as many learners as specified in the ensemble. The combination of the predictions of all the models in the ensemble is expected to be better than a single model. See Fig. S4 in the Supplementary materials for the schematic diagram of gradient boosting algorithms. Mathematically, CatBoost can be represented as follows:

Given a training dataset with N samples and M features, where each sample is denoted as (x_i, y_i), as x_i is a vector of M features and y_i is the corresponding target variable, CatBoost aims to learn a function $F(x)$ that predicts the target variable y.

$$F(x) = F_o(x) + \sum_{m=1}^{M} \sum_{i=1}^{N} f_m(x_i) \tag{2}$$

where, $F(x)$ represents the overall prediction function that CatBoost aims to learn. It takes an input vector x and predicts the corresponding target variable y. $F_o(x)$ is the initial guess or the baseline prediction. It is often set as the mean of the target variable in the training dataset. This term captures the overall average behavior of the target variable. $\sum_{m=1}^{M}$ represents the summation over the ensemble of trees. M denotes the total number of trees in the ensemble. $\sum_{i=1}^{N}$ represents the summation over the training samples. N denotes the total number of training samples. $f_m(x_i)$ represents the prediction of the m-th tree for the i-th training sample. Each tree in the ensemble contributes to the overall prediction by making its own prediction for each training sample.

2.2.2.4. XGBoost model. XGBoost is a tree-based model that implements gradient boosting (See Fig. S4 in Supplementary materials) designed to optimize computation speeds and machine learning performance. In contrast, CatBoost, another gradient boosting algorithm, focuses on handling categorical features effectively and reducing overfitting through a unique boosting strategy. While both models offer valuable advantages, it is essential to evaluate their performance in the specific context of our study. XGBoost was developed by [32]. The objective function (loss function and regularization) at iteration t that we need to minimize is the following:

$$\mathscr{L}^{(t)} = \sum_{i=1}^{n} l\left(y_i, \widehat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \tag{3}$$

where:

- l is the loss function (e.g., squared error).
- f_t is the tree added at iteration t.
- $\Omega(f_t) =$ is the regularization term (controls complexity)

2.2.2.5. SARIMA model. SARIMA stands for Seasonal Auto-Regressive Integrated Moving Average. SARIMA is an extension of ARIMA for forecasting univariate time series with a seasonal component. The statistical model consists of 2 components, which are trend and seasonality. We utilized Python's 'statsmodel' api library [33] to implement the SARIMA model. The mathematical representation of SARIMA model is:

$$\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^{D_{yt}} = \Theta_Q(B^s)\theta_q(B)\epsilon_t \tag{4}$$

where:

- $\phi_p(B)$: non-seasonal AR polynomial; $\theta_q(B)$: non-seasonal MA polynomial; $\Phi_P(B^s)$: seasonal AR polynomial; $\Theta_Q(B^s)$: seasonal MA polynomial; ϵ_t : white noise; p, d, q: non-seasonal AR, differencing, MA orders; P, D, Q: seasonal AR, differencing, MA orders; s: seasonal period (e.g., 12 for monthly data).

2.2.3. Feature selection

We conducted linear correlation analysis on various covariates that could improve the prediction accuracy of the models. Research showed that some economic factors influence the market price of food items in Nigeria such as the domestic fuel prices, foreign exchange rates, and inflation [34]. The NBS [21], Yahoo Finances [23] and CBN [22] maintain online repositories where historical data on such economic variables can be sourced from in the same manner as the food prices. The data for the monthly Crude oil domestic rates, the monthly consumer price indexes for all items, and monthly conversion rates of the United States Dollar to Nigerian Naira were obtained from these sources. The time granularity and historical timespan for each of these features are the same as that of the food prices, however, their spatial granularity is different. The spatial granularity for the Crude oil prices, consumer price indexes, and USD to Naira conversion rates is set at the country level. To select the most relevant lagged features, we combined statistical correlation analysis with model-based evaluation. As shown in the correlation matrix (Fig. S1), 'Last_Month_Price' exhibited the strongest linear correlation with the target variable (r = 0.95), while 'Last_5M_Price' also showed a high correlation, capturing medium-term temporal dependencies. These initial results guided the construction of candidate lagged features (1-, 3-, and 5-month lags, as well as rolling averages and standard deviations).

We then evaluated these candidates using Random Forest feature importance scores, which measure entropy gain during split decisions. The results (Fig. S2) confirmed that 'Last_Month_Price' and 'Last_5M_Price' consistently yielded the highest entropy gains, significantly outperforming other lagged features. This explains their selection as the most informative predictors. Moreover, from an economic standpoint, the last month's price reflects immediate price persistence, while the 5-month lag captures medium-term seasonal or cyclical variations, making them both statistically robust and contextually meaningful.

2.2.4. Model (re)training

We aim to make our model publicly available [14]. Users can call for future forecasts of the selected commodities in real-time. In order to achieve this, we have implemented an online-learning pipeline that is divided into (1) Scraping pipeline, and (2) Retraining module.

2.2.4.1. Scraping pipeline. The scraping pipeline involves python scripts that gather the most recent food prices, consumer price indexes, crude oil export prices and conversion rates from various sources. The data is extracted from the e-library portal of the National Bureau of Statistics [21], Yahoo Finances website [23], and the official repository of the Central Bank of Nigeria (CBN) [22]. After the data is collected, it is incorporated into a 'master' dataframe (a single consolidated dataset containing historical food prices for the 5 commodities across all 37 states), which is then transformed and encoded as described in Section 2.2.1. Additionally, the scraping module includes a prediction API that allows for making predictions by calling the Random Forest model through the command line interface. The model generates predictions by processing a subset of relevant independent variables. The first month's forecast serves as the basis for the subsequent predictions, which are achieved by shifting the lagged features of the subset one step to the right. This process is repeated for all eight forecasting steps, with each forecast replacing the corresponding feature in the sequence. A visual representation of this process can be seen in Fig. 2, which depicts a flowchart of the prediction procedure.

2.2.4.2. Retraining module. The model is retrained periodically (monthly) using the updated and preprocessed training dataset from the scraping pipeline. The retraining is executed in three consecutive steps:

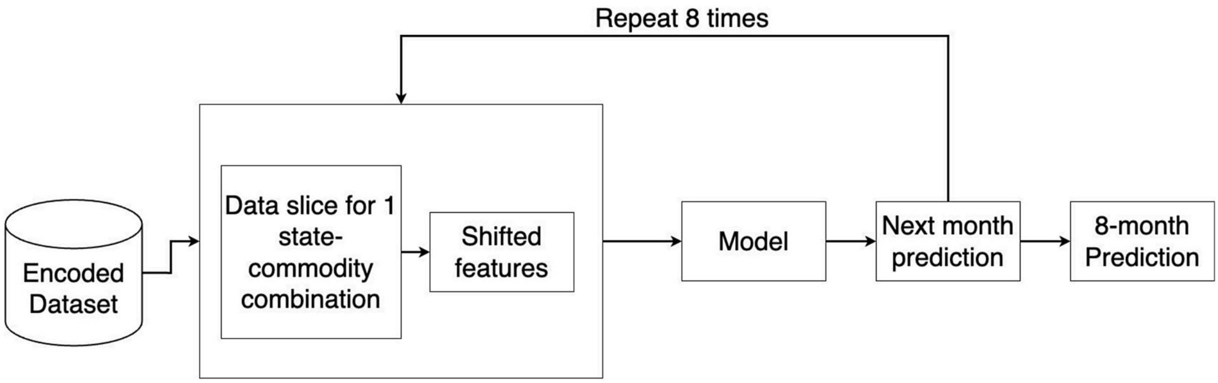

Fig. 2. Flowchart representation of a single prediction call via the prediction module.

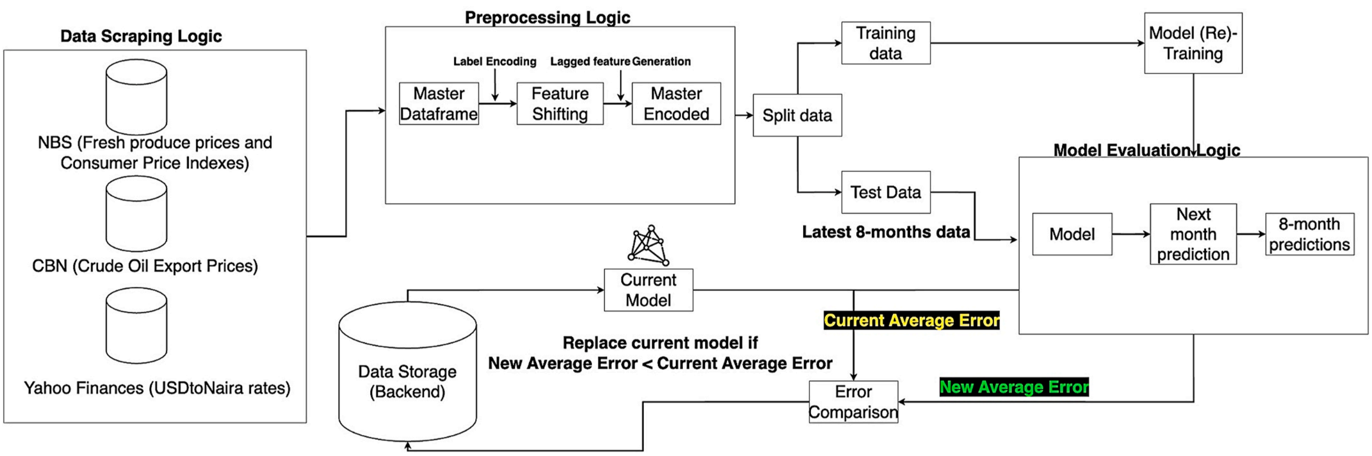

Fig. 3. Flowchart representation of (Re)Training process.

1. A new model is trained using the updated and consolidated encoded data from the latest scraping attempt.
2. The new and current saved models are evaluated on a single hold-out test set consisting of the latest scraped prices using MAPE.
3. If the new model performs better, it replaces the old model in the backend. If the old model performs better, it remains in place and is re-evaluated during the next retraining iteration. To visualize the order of execution of the scraping and retraining modules, see Fig. 3.

2.2.5. *Key performance metrics*

We have chosen two key performance metrics to assess our model's performance on the test set: the Mean Absolute Percentage Error (MAPE) and the Root Mean Square Error (RMSE). MAPE is one of the most popular metrics for assessing the performance of machine learning forecasting models. This is the case in [35] where MAPE is proven to be quite suitable for this application. RMSE measures the average difference between the predicted and actual values. When compared to the MAPE metric, RMSE is less sensitive to extreme values or outliers in the data but on the other hand is also harder to interpret. Therefore, in this paper we use MAPE as the primary metric for evaluation of the models and RMSE as a secondary metric.

The formulas are illustrated as follows:

$$MAPE = \frac{1}{n} \times \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right| \times 100 \tag{5}$$

$$RMSE = \sqrt{\left[\frac{\Sigma(F_t - A_t)2}{n}\right]} \tag{6}$$

n = The number of times the summation occurred
A_t = Actual value
F_t = Forecast value

3. Results and discussion

3.1. *Historical food prices*

Here, we present line plots depicting the food prices for selected states, namely Lagos, Kano, and Plateau, in Fig. 4. Kano and Plateau are two of the largest producers of fresh produce in Nigeria, while Lagos has the largest consumer market for fresh produce in Nigeria. This analytical approach enabled us to ascertain the historical trends and seasonal attributes of the food prices. Through this preliminary analysis, we could evaluate the necessity of employing machine learning techniques to tackle the problem at hand. In the event that the food prices exhibit minimal fluctuations over time, there may be no pressing need to utilize sophisticated forecasting methodologies to predict future prices. Conversely, if the observations demonstrate considerable temporal variance in the historical prices, it would justify the implementation of more complex forecasting models.

Fig. 4 reveals a considerable variability in the prices of the commodities within this time frame. For instance, the price of tomato in Lagos state increased from slightly above ₦100 ($0.22) per kg in January 2017 to approximately ₦400 ($0.86) per kg in October 2022. While a discernible positive trend is observed in the time series, seasonality components are less evident in the line plots, as is typical of such visualizations. All the historical food price plots exhibit similar

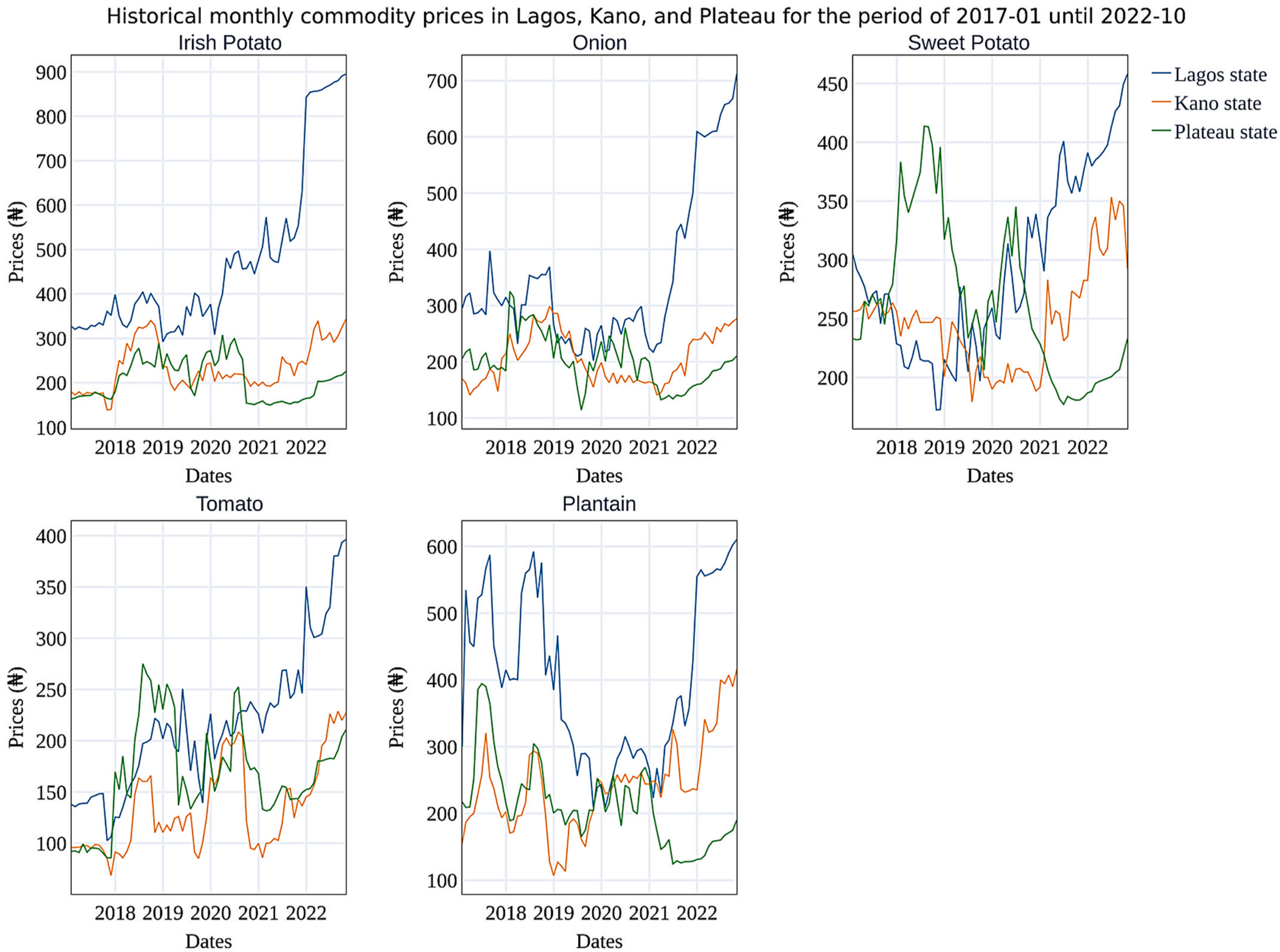

Fig. 4. Line plots of the historical monthly prices of potato, onion, sweet potato, tomato, and plantain in Lagos, Kano, and Plateau states within the timeframe of January 2017 until October 2022. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

patterns of large variances and standard deviations from the mean price within the timeframe (2017-01 until 2022-10) that are far greater than 100 %. In contrast, the case of plantain prices in Plateau state is notable for an indeterminate trend component, featuring several positive and negative peaks within the given timeframe. Given the substantial temporal fluctuations in the historical food prices, using a non-linear model such as the machine learning model proposed in this research paper is justified for accurate forecasting of future prices.

3.2. Model comparison

This section provides an overview of the results obtained from training various algorithms on the dataset and evaluating their performance using a hold-out test set. As outlined in Section 2.2.2, we experimented with 5 different models, each with its own set of hyperparameters. In this section, we compare the results of each model when evaluated on the same hold-out test sets. We present the results of evaluating the models on a single test set and those from temporal cross-validation (multiple test sets).

3.2.1. Single hold-out test set evaluation using MAPE

Fig. 5 shows the consolidated Mean Absolute Percentage Error (MAPE) for the 5 models, averaged over all 5 commodities and all 37 states. We utilized monthly averaged food price data, spanning January 2017 to February 2022, for model training. The models were then evaluated on a hold-out test set, consisting of March to October 2022 data. This set encompassed five commodities across 37 states, assessing predictive efficacy for price movements. From the aggregated MAPE line plot (Fig. 5), it is evident that the best performer is the Random Forest model, with XGBoost and Catboost close behind. One possible explanation for this performance is that Random Forest models tend to handle noisy and small datasets better [36]. Unsurprisingly, the SARIMA model performs worse on average than the basic benchmark model, which predicts the last 5-month rolling average prices. This inferior performance of SARIMA stems from two main limitations. It assumes stationarity, which is often violated in Nigerian food price data marked by shocks and regional heterogeneity. In addition, being univariate, it cannot incorporate key exogenous drivers such as exchange rates or CPI, which reduces its ability to capture the dynamics present in the data. The benchmark model, on the other hand, achieves a MAPE below 20 % for most forecast steps (excluding the 6th and 8th months). This observation does not undermine the importance of machine learning in this problem but emphasizes the need for further analysis. The results are considered myopic because they are based on a single test set at a certain time and do not account for the entire time range of data. To really visualize the advantage of machine learning over the baseline, we explore temporal cross-validation as discussed in section 3.2.3.

We visualized the distribution of the 8-month average forecast Mean Absolute Percentage Error (MAPE) for each model using box plots (Fig. 6). As shown in Fig. 6, the box plots clearly illustrate the

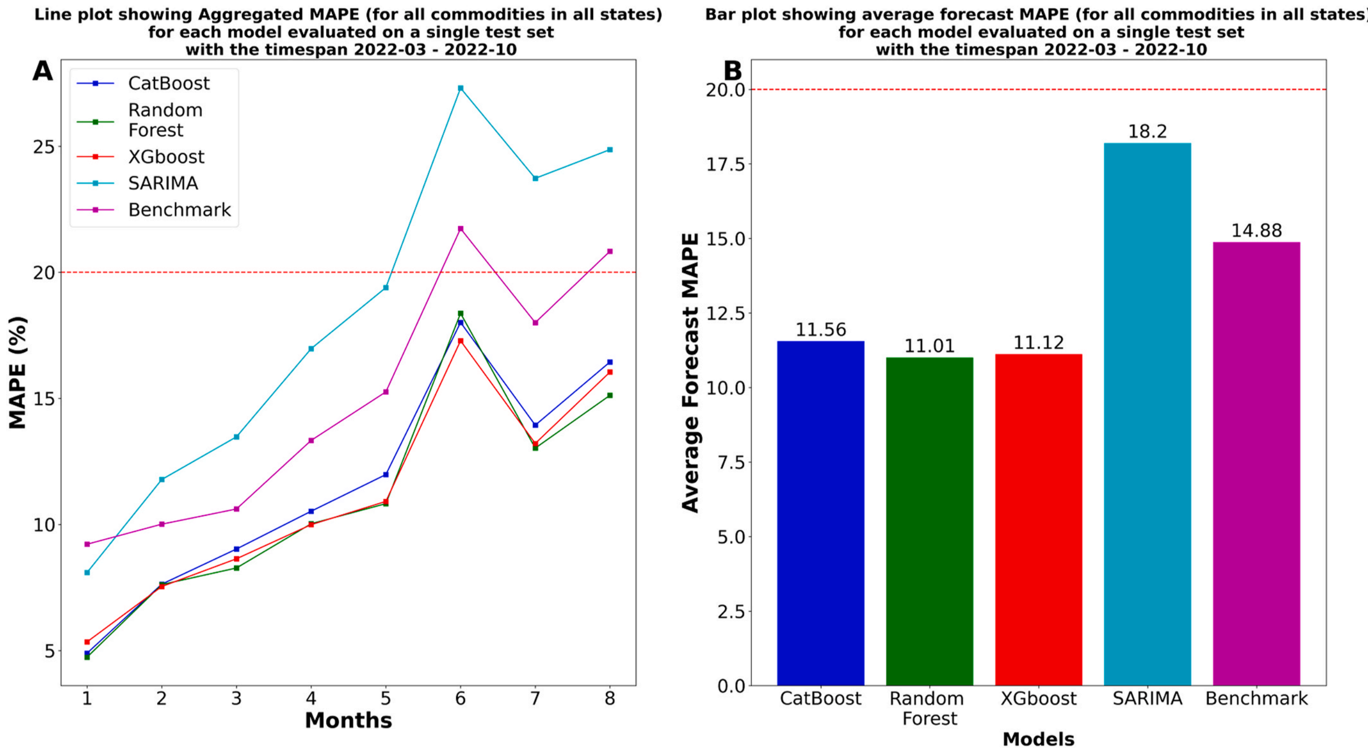

Fig. 5. (A) Line plot showing Aggregated MAPE (for all 5 commodities in all 37 states) for each model evaluated on a single test set with the timespan 2022-03 - 2022-10 (B) Bar plot showing the average forecast MAPE for each model evaluated on a single test set with the timespan 2022-03 - 2022-10. The red dotted line represents the 20 % acceptable error limit. The bars from left to right, correspond to the CatBoost, Random Forest, SARIMA, XGBoost, and Benchmark rolling-average models, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

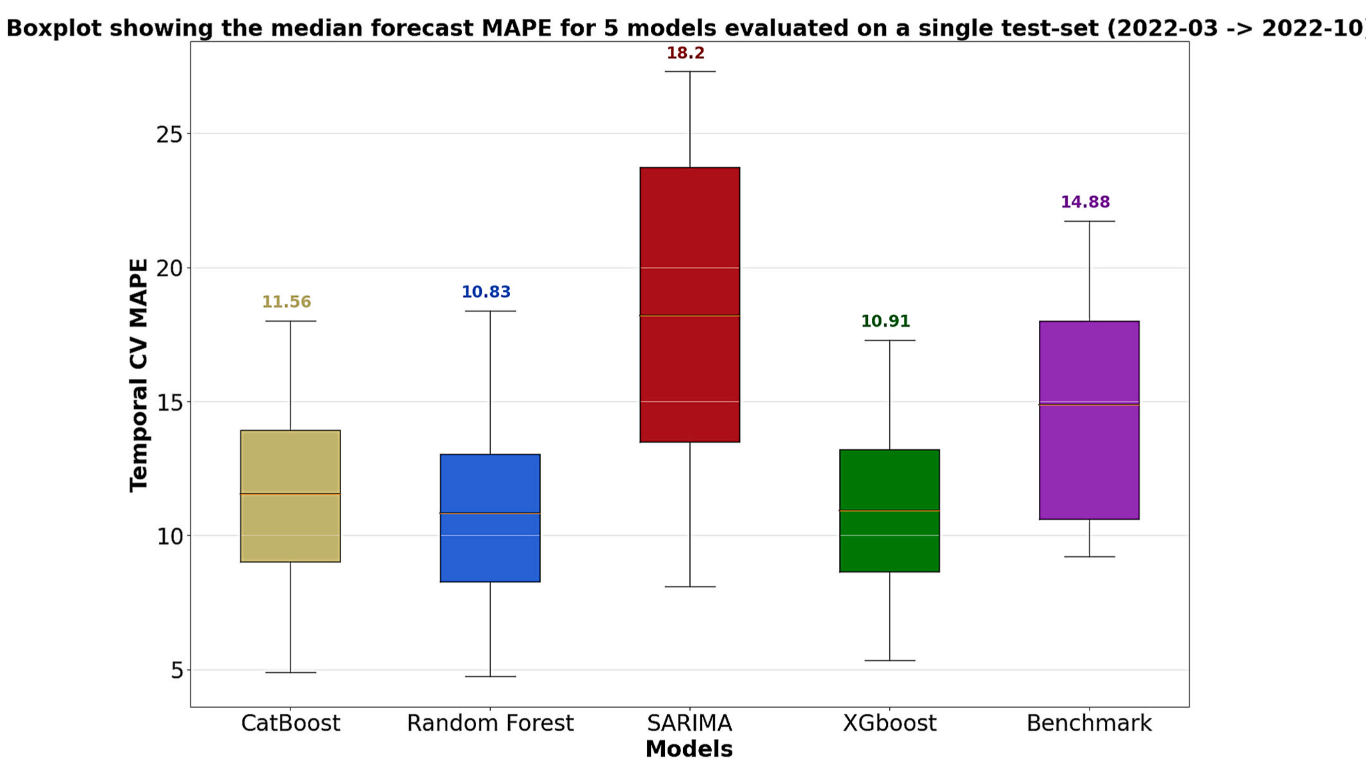

Fig. 6. Box-plot illustrating the aggregated MAPE (for all commodities in all states) for each model evaluated on a single test set spanning from 2022-03 to 2022-10. Each box represents the interquartile range (IQR) with whiskers as quantiles, and the middle line denoting the median average forecast MAPE. The numbers at the top correspond to the median values of the boxes. The boxes, from left to right, correspond to the CatBoost, Random Forest, SARIMA, XGBoost, and Benchmark rolling-average models, respectively.

Table 2
Performance Evaluation of 5 Models evaluated on a single test set with the timespan 2022-03 - 2022-10 – Mean Absolute Percentage Error (MAPE) and Average Root Mean Squared Error (RMSE) Values For All Commodities Across All States.

MODELS	MAPE (%)	RMSE (Naira)
Benchmark	14.88	60.84
SARIMA	18.20	70.96
Random Forest	11.01	45.16
XGBoost	11.12	50.17
CatBoost	11.56	45.48

distribution of the average forecast MAPE for each model. The box in the plot represents the interquartile range (IQR) of the errors, with the line inside the box representing the median MAPE. The whiskers extending from the box show the range of the MAPE, excluding any outliers. The results indicate a superior accuracy of the Random Forest model in predicting the target variable. This was particularly notable in the early time steps, specifically months 1 and 2, as illustrated by the line plot in Fig. 5. However, it demonstrated less consistency than the XGBoost model. We prioritize the forecast accuracy for the earlier months because these are the most important for driving informed decision-making, especially time-constrained ones.

3.2.2. Single hold-out test set evaluation using RMSE

The Root Mean Squared Errors (RMSEs) of the five models were computed in a single test-set evaluation and presented in Table 2. RMSE represents the square root of the average of the squared differences between predicted and actual values. The average RMSE values displayed in Table 2 were computed by averaging the RMSEs of model forecasts for all commodities in all states, as done for the MAPE metric. The results indicate that decision tree-based algorithms performed similarly, with the Random Forest model having the lowest RMSE value of ₦45.16. The Catboost model had an average RMSE of ₦45.48, making it a close second, while XGBoost performed relatively worse with an average RMSE of ₦50.17. In contrast, the SARIMA and the rolling-average benchmark models produced less promising results. The SARIMA model had an average RMSE of ₦70.96, while the rolling-average benchmark had an RMSE of ₦60.84. The results further indicate that Random Forest outperformed other models such as XGBoost, SARIMA, and the rolling-average benchmark suggesting a more predictive power when it comes to predicting food values for smallholder farmers in Nigeria. By accurately predicting food values, farmers can anticipate market trends, adjust their production strategies, and make informed decisions regarding when to sell their produce. This information can aid in optimizing crop selection, timing of harvest, pricing strategies, and resource allocation.

3.2.3. Temporal cross-validation

Prior to evaluating the performance of the five proposed models using Temporal Cross-Validation (CV), it is imperative to conduct a thorough analysis of the historical food prices time series. This would enable us to visually inspect the price variances within the CV duration and establish a reasonable threshold for assessing the efficacy of the models. To this end, Fig. 7 illustrates the line plots of the food prices averaged across all states (37) for the duration spanning from January 2020 to September 2022. Additionally, the corresponding standard deviations from the mean price during this period are depicted. Notably, for the Irish potato, onion, and plantain commodities, the standard deviation ranges approximately ±70 %. Conversely, the sweet potato and tomato commodities exhibit lower standard deviations at approximately ±40 %. Aside from the inherent unpredictability of price fluctuations, substantial variances can also be attributed to the considerable differences between states in the Northern regions, where the majority of food is produced, and states in the Southern regions, where most of the food is consumed. Further details on this topic can be found in Section 5 of the Supplementary material.

The flowchart of the temporal/time-series cross-validation performed in this study is shown in Fig. 8. The results of the cross-validation experiment for the 5 models are presented in the supplementary material (Table S1–S10). The total number of test sets used in the cross-validation is 26, and the consolidated results for each food class were averaged over 26. In simple terms, the temporal CV aims to showcase how well the models perform over time by utilizing rolling test sets. Instead of focusing on a single 8-month window, the evaluation expands to 33 months, concurrently increasing the training data. For smallholders, the model exhibiting the least average forecast MAPE from temporal CV which in our case is the Random Forest (See Table S2 and S6 in supplementary materials) will likely be the most adaptive. It would best accommodate the price data's temporal variance. Consequently, it

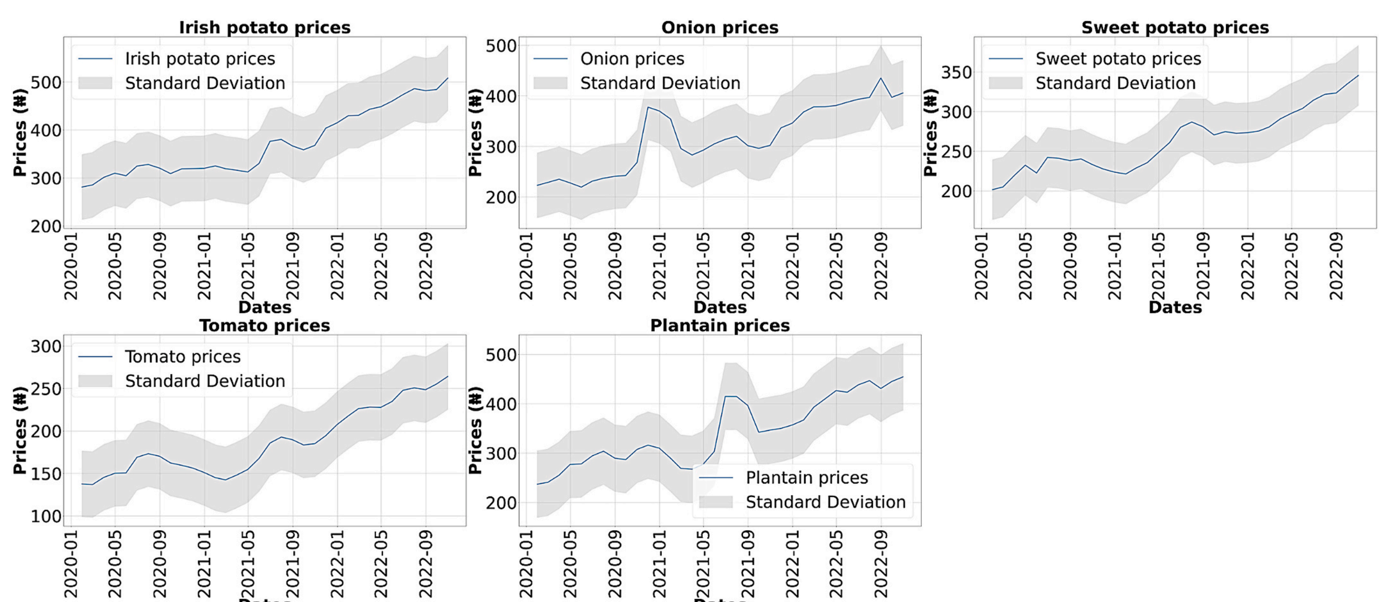

Fig. 7. Line plot showing the monthly historical food prices and the standard deviations from the mean (per food) within the time frame of 2020-01 until 2022-09. The data is averaged over all 37 states.

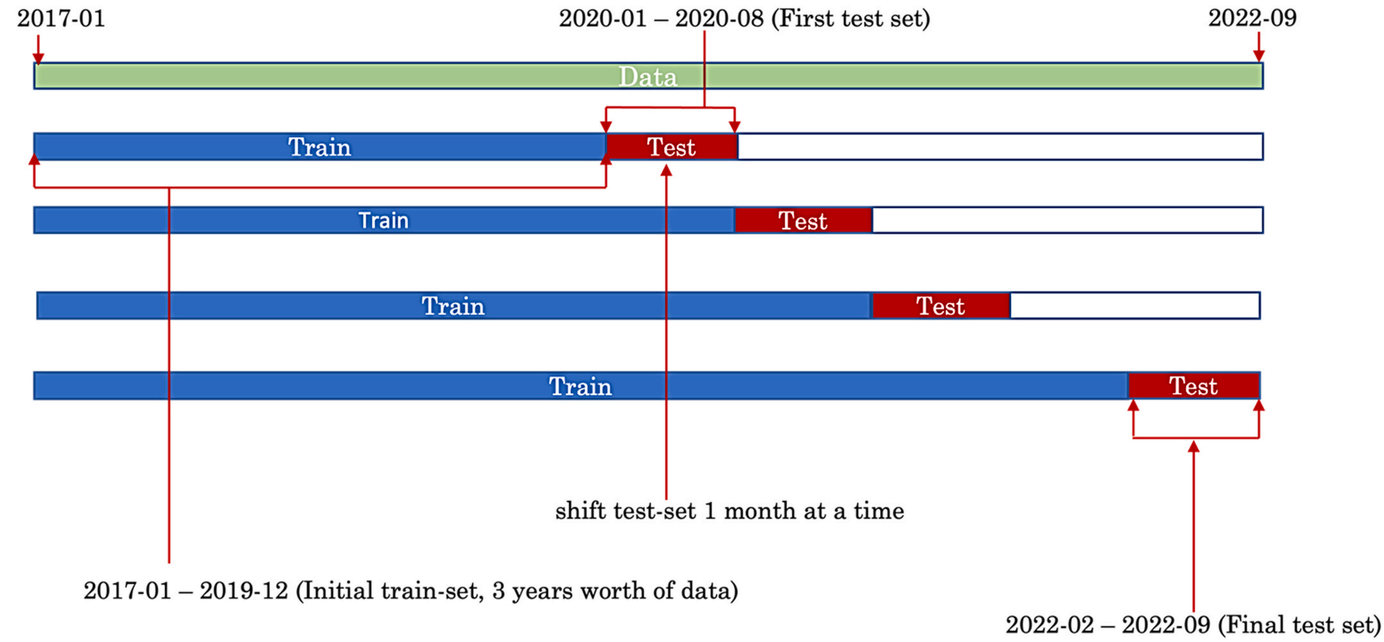

Fig. 8. Graphical representation of temporal cross-validation.

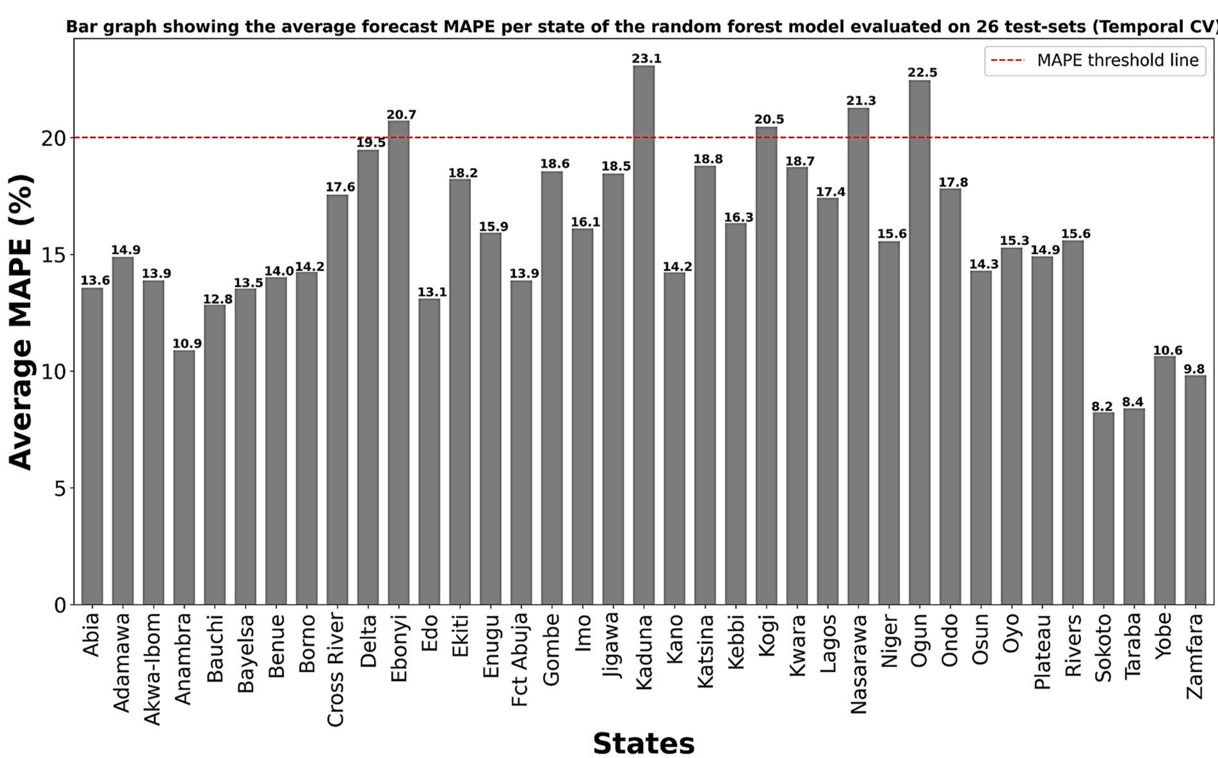

Fig. 9. Bar graph depicting the mean forecast MAPE per state, averaged over 26 test sets, for the random forest model using a temporal cross-validation approach. The red dotted line represents the 20 % acceptable error limit. States with a MAPE exceeding this threshold include Kaduna, Nasarawa, Ebonyi, Kogi, and Ogun. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

would demonstrate the greatest predictive capability, remaining robust as new data are integrated via the pipeline (section 2.2.4.1). Maintaining a model that efficiently generalizes to new data ensures the continuous provision of quality market intelligence. This means that smallholders are empowered with information to make informed decisions even in the future to come.

Leveraging the insights garnered from the aforementioned plots and recent research in the field of time series forecasting, we established a threshold of 20 % Mean Absolute Percentage Error (MAPE) for the evaluation of the five models under consideration. Furthermore, we considered this 20 % threshold acceptable for our specific use case, targeting the improvement of smallholder farmer income in Nigeria through food price forecasting.

3.2.3.1. State-wise temporal cross validation. In this section, the results of temporal cross-validation Mean Absolute Percentage Error (MAPE) for the Random Forest model are presented on a state-by-state basis. Fig. 9 displays the 8-month average forecast MAPE computed across the CV test-sets (26) for all commodities in the form of bar plots. In addition to the consolidated results (see Table S1–S10 in the supplementary materials), our finding further indicates that the random forest model outperforms the benchmark and the other models on average across all states. However, it is also evident that the simpler 5-month rolling average model performs similarly well to the machine learning models, as seen in the consolidated results tables (S1–S10). This result is due to two main reasons: first, the training data is not sufficient for the machine learning models to reach their maximum potential, leading to similar performance as simpler models such as the Benchmark. Second, the forecasting problem at hand is highly linear, making it easier for linear models to achieve comparable performance. While these are the primary explanations supported by our analysis, further investigation may reveal additional contributing factors. By leveraging online machine learning and model retraining techniques, we assert that the advantages of machine learning models will become increasingly apparent as the training dataset grows larger.

The analysis further reveals that the model did not perform so well in 5 states, namely Kaduna, Nasarawa, Ebonyi, Kogi, and Ogun, where the 8-month average forecast MAPE slightly exceeds 20 %. This represents 13.5 % of the total states in Nigeria and provides an opportunity for model optimization. Several factors may contribute to these marginal higher errors: stronger price volatility in urban consumption centers, inconsistencies in the National Bureau of Statistics data for some commodities, and structural differences between production-oriented and consumption-oriented markets. Addressing these limitations may require incorporating additional predictors such as transportation costs, local weather shocks, and market-level demand indicators.

Nonetheless, the performance of our model gave satisfactory results and can therefore be a useful tool for smallholder farmers to increase income due to better market intelligence. Table S1 of the supplementary materials also contain details of the performance of the random forest model in different states.

3.2.3.2. Commodity-wise temporal cross validation. Fig. 10 shows the cross-validation MAPE results for the Random Forest model, displayed per food commodity. The 8-month average forecast MAPE, computed from the CV test-sets across all states, is represented by each bar. The result based on the commodity-based cross-validation indicate that the Random Forest model performs well for all commodities, with the highest accuracy in Irish potato and Plantain. All food commodity classes have an average MAPE below 20 %, which is considered good performance. This demonstrates that the model's forecasts typically deviate from actual prices by an average of 20 %. This performance metric, superior to methods heavily dependent on historical prices, offers smallholders substantial market intelligence. As a result, they are empowered to make decisions about their fresh produce that are more grounded in data.

Table 3
Results of Diebold–Mariano tests comparing model forecast accuracy against the five-month rolling average benchmark. Counts of significant wins, losses, and non-significant outcomes are reported for each model and error metric (MAE, MAPE, MSE).

Model	Metric	Wins_vs_Baseline	Losses_vs_Baseline	Non-significant
CatBoost	MAE	23	6	156
	MAPE	25	7	153
	MSE	13	4	168
XGBoost	MAE	9	4	172
	MAPE	12	5	168
	MSE	5	3	177
Random Forest	MAE	21	5	159
	MAPE	25	6	154
	MSE	10	5	170
Sarima	MAE	6	10	169
	MAPE	5	11	169
	MSE	2	3	180

3.2.4. *Statistical significance of model predictions*

To complement the error based evaluation, we applied Diebold Mariano tests [37] to compare forecast accuracy across models for MAE, MAPE, and MSE at the eight month horizon. The value of alpha was set at 0.05 as it is the standard and widely accepted level in statistical testing [38]. Tests were run pairwise and then aggregated by commodity and by state. Table 3 in the section below details the results of the statistical tests.

Against the five month rolling average benchmark, the tree based models show clear and statistically meaningful gains. CatBoost records

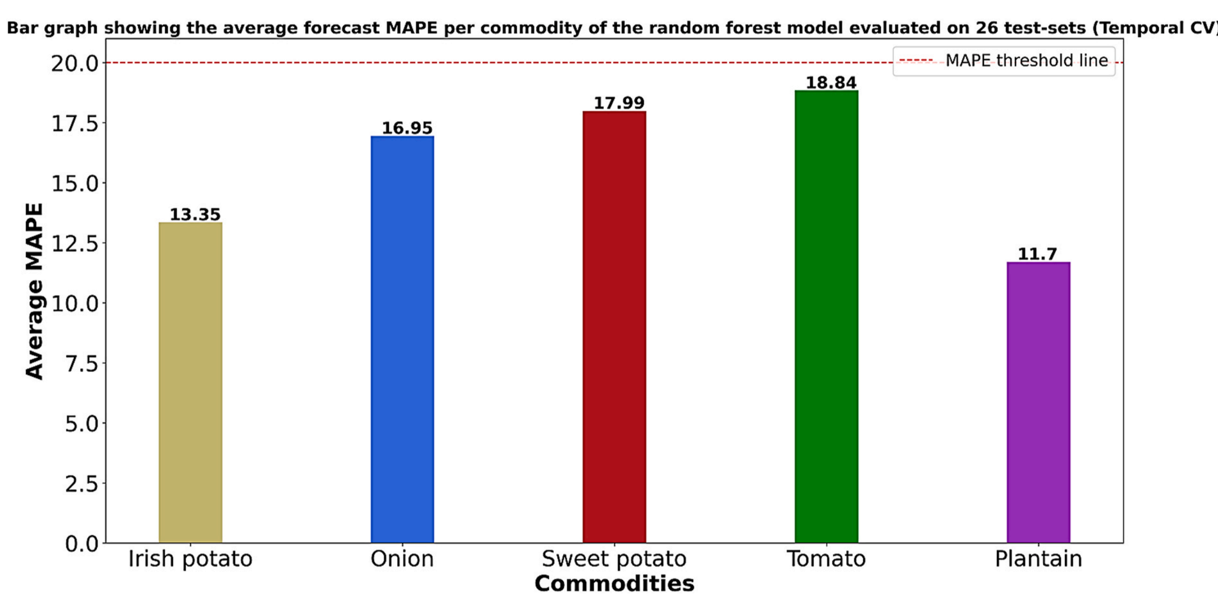

Fig. 10. Bar graph showing the 8-month average forecast MAPE per food commodity for the random forest model evaluated over 26 test-sets in a temporal cross-validation approach. The red dotted line signifies the acceptable error limit. Average forecast MAPE for all food commodity classes are below the desired threshold of 20 %. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

61 significant wins versus 17 losses across the three metrics. Random Forest records 56 wins versus 16 losses. XGBoost records 26 wins versus 12 losses. SARIMA shows only 13 wins versus 24 losses. These counts confirm that the tree based group outperforms the benchmark with statistical support, while SARIMA does not. Between the tree based models there is no systematic dominance across commodities or states. CatBoost has a small edge in the number of significant wins over Random Forest, but both deliver consistent and significant improvements over the benchmark. This is aligned with the error based results in Sections 3.2.2 and 3.3.3 and supports our choice to rely on a tree-based approach in production.

We retain Random Forest as the primary model due to its strong accuracy at the early horizons that matter most for decision making, its stable behaviour across states, and its lower sensitivity to tuning in our retraining pipeline. Full disaggregated results, including the commodity-specific and state-specific aggregates and the series-level tables, are not included here to maintain brevity but are available from the authors upon request.

3.2.5. *Validation of random forest model*

Here, we compare the Random Forest model's real-time price predictions with actual data from August 2022 to March 2023. A specific state is chosen for each commodity to demonstrate the model's effectiveness.

Fig. 11 presents line plots depicting the price trends of irish potato, onion, sweet potato, tomato, and plantain in various Nigerian states. Our analysis reveals the superiority of the random forest model in forecasting the price trends of these commodities as seen in Figs. 5 and 6. In the case of Irish potato prices in Niger state, the RF model offers highly accurate price predictions with minor deviations. The model exhibits an overestimation bias in the initial four months, while later predictions slightly underestimate the actual prices. Notwithstanding these deviations, it's notable that the average forecast error remains within the standard deviation of the observed prices, demonstrating the model's reliability. Analogously, the model demonstrates a comparable performance when forecasting Onion prices in Kwara state. Under-confidence is detected in the predictions for the third and fourth months but the model accurately adapts to the rising trend seen in the final months. Regarding the Sweet potatoes market in Taraba state, the RF model demonstrates commendable accuracy, with a notable exception in the final two months where it tends to overestimate prices. This indicates a potential limitation of the model in accounting for sudden market changes or rare events. A similar limitation is observed in predicting Plantain prices in Enugu state, where the model fails to account for a sudden drop in prices towards the beginning of September. The model, nevertheless, manages to reconcile this forecast error over the subsequent six months, achieving accurate predictions. Finally, when forecasting Tomato prices in Plateau state, the model accurately captures the overall positive trend, albeit with a slight underestimation bias during the fourth, fifth, and sixth months. A slight overestimation is noticed for the last two months, reinforcing the need for constant monitoring and adjustment of the model parameters to optimize its performance.

In summary, the current model boasts of a great forecasting accuracy for a majority of the states in Nigeria. This means that smallholder farmers can be confident of the forecast provided by this tool and make smart decisions to reduce loss and scale their profit margins.

4. Model implementation in open-source mobile app

The deployment of our machine learning (ML) model presents a compelling use case for supporting smallholder farmers in making informed decisions. Specifically, the model can help farmers identify the optimal regions to sell their harvest, plan planting and harvest seasons to maximize profits during sales, and make other crucial decisions related to crop management. These benefits extend beyond smallholders to encompass larger organizations such as government bodies, NGOs, and research institutions. The model leverages historical food price data and provides accurate forecasts of future prices several months ahead. This information is invaluable to larger organizations when making decisions and policies that impact smallholders positively.

To provide users with access to the model's insights, we have developed a data science-based mobile application called Coldtivate [14]. The app is open-source (https://gitlab.com/b1866/coldtivate/mobile-app-react-native) and available on both Android and iOS. Within the Coldtivate app, users will be able to select fresh produce and states (See Fig. 12) within Nigeria to generate monthly forecasts of average fresh produce prices up to eight months in advance. Note that the app is also available for stakeholders worldwide. The app's interface includes a graph that displays both historical and forecasted prices, allowing users to visualize recent price trends and make informed decisions accordingly. As an example Fig. 12B show the market price of tomato in Lagos state from April 2023 to August 2023. Similarly, the farmer could decide from Fig. 12C what state and month they would target to sell their produce and make the most profit. With such a powerful tool at their disposal, they can optimize their revenue and at the same time reduce food loss by knowing when and where to sell even before harvest, which is one of the main drivers of food loss at the farm.

4.1. *Challenges*

In this section, we discuss the challenges faced during the development of the fresh produce market price forecasting tool. We also provide recommendations on how to resolve them for further development or applications of the price forecasting tool.

1. Inadequate Training Data: The available open-source data for training the machine learning algorithms is limited. Currently, only monthly averaged prices from January 2017 to March 2023 are available, which results in a mere 75 data points per food commodity per state. To address this challenge, it is necessary to have more historical food commodity prices available, with a higher frequency (e.g., daily or weekly). This can be done through private-public partnership of government organizations, NGOs, policy makers, and research organizations in Nigeria to gather relevant and trustworthy statistical data and make it open-source. By doing so, the methodology outlined in this paper can be successfully reproduced with improved results. The positive implications of gathering relevant data is not limited to fresh produce price forecasting, they can be converted to other data-driven solutions for smallholders, such as interactive web maps hosting relevant geospatial insights [39], smart supply-chains with decentralized cooling and refrigerated transport solutions, etc., as recommended in [4].
2. Poorly Structured Open-source Repositories: Another challenge faced in the development of the food price forecasting tool was the poor structure of the NBS e-library site [21]. The food price and Consumer Price Index (CPI) data is uploaded on the e-library site as Microsoft Excel sheets, which presents a challenge for the web scraping module to efficiently and accurately update the training data. Changes to the structure of the Excel sheets, such as dimensions, column names, sheet numbers, or indexes, can result in incorrect data being scraped. To mitigate this risk, the scraping module has been tested to remain robust in the event of some changes. However, it is not possible to anticipate all possible modifications to the Excel sheets, leaving the possibility of future failures if the scraping module is not regularly maintained and updated in tandem with the website. A solution to this challenge would be for the National Bureau of Statistics (NBS) to improve the web-scraping friendliness of their website by using dynamic HTML tables instead of traditional Microsoft Excel worksheets to store open-source data.
3. Inconsistent Data Quality: While developing our food price forecasting tool, we recognized areas for improvement related to data quality within the NBS dataset. Occasionally, inconsistencies and

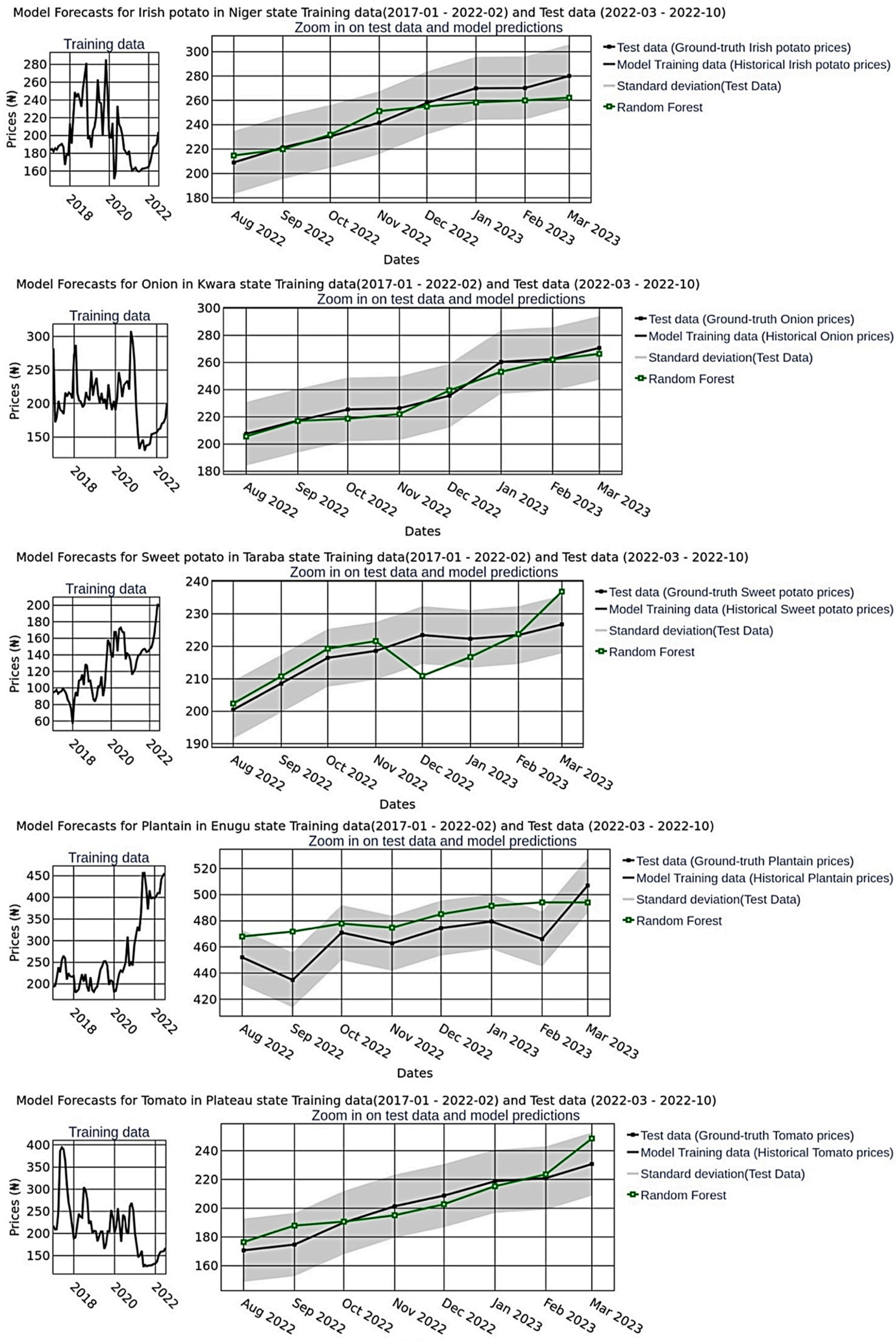

Fig. 11. Verification of Model's forecasts for Irish potato, onion, sweet potato, tomato, and plantain prices in various nigerian states.

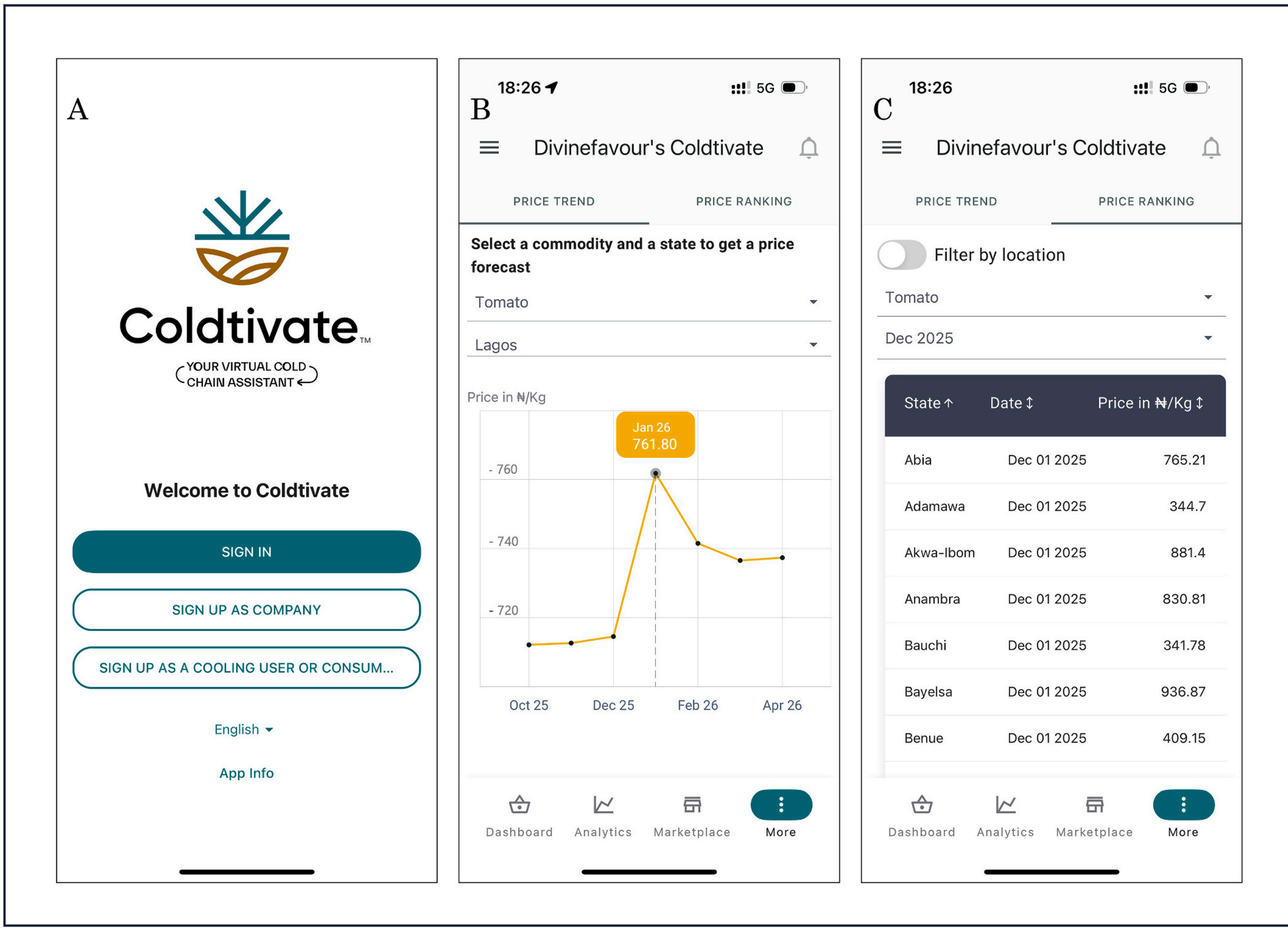

Fig. 12. Initially designed sketches for a machine learning model prediction graph in Coldtivate. The Price trend section displays the historical prices and forecasts for a user-selected food and state. The price ranking section displays the prices of a user-selected food for all states per month, arranged in ascending or descending order based on user preference.

gaps were present in the food prices, which can potentially influence the accuracy of model predictions. We view these inconsistencies not as flaws, but as opportunities for enhancement. By advocating for stringent data quality control processes, we aim to ensure the usage of only high-quality, complete data for training and testing our model, thus bolstering the model's reliability. We were able to make significant strides in reducing this challenge using data imputation techniques, such as linear interpolation.

4. Limited Features for Model Training: The lack of availability of additional features, such as geopolitical events and demand and supply levels, limits the ability of the model to capture all the factors affecting food prices. This could be a reason for the less accurate predictions and lower overall model performance in some Nigerian states. To address this challenge, additional variables/features that capture the economical influencers of food prices in Nigerian markets should be made openly available, collected and incorporated into the model training, as long as they are relevant and have a significant impact on food prices.

While the challenges discussed may affect the potential impact of the ML forecasting tool in Coldtivate, it's important to underscore the tool's existing value. Our methodology, even at its current stage, provides adequate prediction accuracy, offering valuable market price insights and trends for 85 % of states (see Fig. 9) and all major produced commodities in Nigeria. This translates to substantial support for smallholder farmers in their decision-making process. Furthermore, by providing high-quality historical price data with a higher frequency at weekly or daily levels, the methodology presented in this paper can be scaled and applied to achieve better results. As a result, the overall impact of ML forecasting tools on the lives of smallholder farmers can be increased significantly.

4.2. Future extensions

While our current study focuses on static forecasting of state-level food prices, several extensions are possible. Future work could integrate stochastic approaches to explicitly model uncertainty in commodity markets, where shocks from climate variability or macroeconomic volatility can alter price dynamics [20]. In particular, incorporating prediction intervals would be a valuable addition, as it would provide farmers and policymakers with not only point forecasts but also measures of confidence around these forecasts, enhancing decision-making under uncertainty. A further direction is the inclusion of network-based models that capture interdependencies between regional markets and transport corridors, providing a richer picture of price transmission [17]. Dynamic optimization methods could also be employed to support decision-making on storage, transport, or sales timing, linking price forecasts with operational choices in the supply chain. Finally, extensions that incorporate human and institutional factors, such as trader behavior, market regulations, and cooperative bargaining, would bring the forecasting pipeline closer to real-world complexity and policy relevance [4]. These pathways highlight the potential of combining machine learning with modern mathematical tools to improve resilience and profitability in smallholder agriculture.

5. Conclusions

This paper presents a new application of machine learning in the field of agriculture and the food supply chain, specifically for fresh produce price forecasting in Nigeria. A robust system has been implemented to accurately predict real-time food prices across different states, using an online-learning approach. This approach periodically scrapes open-source repositories for historical food prices and external factors such as inflation, foreign exchange rates, and domestic fuel prices. It also retrains a machine learning model with the latest data to enhance forecast accuracy. Several machine learning and statistical models were evaluated using the MAPE and RMSE performance metrics, and the best performing model, Random Forest, was selected. The analysis showed that a non-linear model trained on historical prices of fresh produce, supported by economic influencers as predictors, can accurately forecast future fresh produce prices. Our key findings are:

- Compared to SARIMA, CatBoost, XGboost, and a 5-month rolling average benchmark, Random Forest model accurately predicts 8 month average market prices of tomato, onion, plantain, and potato across 37 states in Nigeria. The model achieved an average Mean Absolute Percentage Error (MAPE) of <20 %. This error is much less than the typical price variation throughout the year and between different years.
- The criteria with the largest influence on the model performance is the historical market prices followed by USD to Naira exchange rates. The historical market prices accounted for approximately 92 % of the model performance while exchange rates coupled with other exogenous factors contributed to the model accuracy by 7–8 %.
- Non-linear, tree-based machine learning algorithms are better suited for fresh produce market price prediction for a volatile value chain like Nigeria and provide forecasts that are at least 5–7 % more accurate when compared to linear models such as SARIMA and rolling-average (benchmark). Tree-based models combined with external economical variables also ensure a model more robust to outliers.
- Our model performed better in fresh produce producing states like Kano, Plateau, and Sokoto with less volatility compared to states where they are largely consumed like Lagos.

We have integrated this model into a mobile application called Coldtivate [14]. The purpose of Coldtivate is to help reduce postharvest losses in the food supply chain and increase smallholder farmer's income. Smallholders can use Coldtivate to access valuable insights on maximizing the shelf life of their fresh produce and market intelligence, including historical food prices and price forecasts. Policy makers, government organizations, NGOs, and other stakeholders can also use the features offered by Coldtivate including the fresh produce knowledge hub, fruit and vegetable digital twins and price forecasting models to make informed decisions to mitigate postharvest losses and increase smallholder income.

CRediT authorship contribution statement

Divinefavour Odion: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Joaquin Gajardo:** Writing – review & editing, Methodology, Supervision. **Thijs Defraeye:** Writing – review & editing, Project administration, Methodology, Funding acquisition. **Thomas Motmans:** Writing – review & editing, Funding acquisition. **Kanaha Shoji:** Writing – review & editing, Methodology. **Roberta Evangelista:** Writing – review & editing, Project administration, Methodology, Funding acquisition. **Daniel Onwude:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was partially funded by the project "Scaling up Your Virtual Cold Chain Assistant" commissioned by the German Federal Ministry for Economic Cooperation and Development and being implemented by BASE and Empa on behalf of the German Agency for International Cooperation (GIZ). The funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jafr.2025.102365.

Data availability

Data will be made available on request.

References

[1] Nigeria at a glance | FAO in Nigeria | food and agriculture organization of the united nations." Accessed: February. 8, 2023. [Online]. Available: https://www.fao.org/nigeria/fao-in-nigeria/nigeria-at-a-glance/en/.
[2] V. Ricciardi, N. Ramankutty, Z. Mehrabi, L. Jarvis, B. Chookolingo, How much of the world's food do smallholders produce? Global Food Secur. 17 (Jun. 2018) 64–72, https://doi.org/10.1016/J.GFS.2018.05.002.
[3] Doris Dokua Sasu, "Nigeria: share of global poverty 2016-2023 | Statista," share of global population living in extreme poverty in Nigeria 2016-2023. https://www.statista.com/statistics/1287840/share-of-global-population-living-in-extreme-poverty-in-nigeria/. (Accessed 13 February 2023).
[4] D. Onwude, et al., Bottlenecks in Nigeria's fresh food supply chain: what is the way forward? Trends Food Sci. Technol. 137 (Jul. 2023) 55–62, https://doi.org/10.1016/J.TIFS.2023.04.022.
[5] F.N. N, J. C, H.A. E, J.A. E, C.N. A, Ict for agriculture knowledge management in Nigeria: lessons and strategies for improvement, Science Journal of Agricultural Research and Management 2012 (2012), https://doi.org/10.7237/SJARM/192.
[6] W. Ma, K. Nowocin, N. Marathe, G.H. Chen, An interpretable produce price forecasting system for small and marginal farmers in India using collaborative filtering and adaptive nearest neighbors, in: ACM International Conference Proceeding Series, Association for Computing Machinery, Jan. 2019, https://doi.org/10.1145/3287098.3287100.
[7] C.Z. Yuan, S.K. Ling, Long short-term memory model based agriculture commodity price prediction application, in: ACM International Conference Proceeding Series, Association for Computing Machinery, Aug. 2020, pp. 43–49, https://doi.org/10.1145/3417473.3417481.
[8] R. Carbonneau, K. Laframboise, R. Vahidov, Application of machine learning techniques for supply chain demand forecasting, Eur. J. Oper. Res. 184 (3) (Feb. 2008) 1140–1154, https://doi.org/10.1016/J.EJOR.2006.12.004.
[9] J. Rincon-Patino, E. Lasso, J.C. Corrales, Estimating avocado sales using machine learning algorithms and weather data, Sustainability 10 (10) (Sep. 2018) 3498, https://doi.org/10.3390/SU10103498, 2018*, Vol.* 10*, Page* 3498.
[10] Z. Mustaffa, Y. Yusof, S.S. Kamaruddin, Enhanced artificial bee colony for training least squares support vector machines in commodity price forecasting, J. Comput. Sci. 5 (2) (Mar. 2014) 196–205, https://doi.org/10.1016/J.JOCS.2013.11.004.
[11] S. Mehtab, J. Sen, A time series analysis-based stock price prediction using machine learning and deep learning models, Int. J. Bus. Forecast. Market. Intell. 6 (4) (Apr. 2020) 272, https://doi.org/10.1504/IJBFMI.2020.115691.
[12] Y.E. Shao, J.T. Dai, Integrated feature selection of ARIMA with computational intelligence approaches for food crop price prediction, Complexity 2018 (2018), https://doi.org/10.1155/2018/1910520.
[13] L. Breiman, Random forests, Mach. Learn. 45 (1) (Oct. 2001) 5–32, https://doi.org/10.1023/A:1010933404324.
[14] Coldtivate app. https://yourvcca.org/coldtivate/. (Accessed 18 September 2025).
[15] A. Theofilou, S.A. Nastis, A. Michailidis, T. Bournaris, K. Mattas, Predicting prices of staple crops using machine learning: a systematic review of studies on wheat, corn, and rice, Sustainability 17 (12) (Jun. 2025) 5456, https://doi.org/10.3390/SU17125456, 2025*, Vol.* 17*,* Page 5456.

[16] R.L. Manogna, V. Dharmaji, S. Sarang, Enhancing agricultural commodity price forecasting with deep learning, Sci. Rep. 15 (1) (Dec. 2025) 1–24, https://doi.org/10.1038/S41598-025-05103-Z. SUBJMETA=4066,4068,704,843,844; KWRD=ENERGY+ECONOMICS,ENVIRONMENTAL+ECONOMICS.
[17] M. Adamu Abubakar, B.E. Okpe, M. Audu Nirsal, Predicting Food Price Trends in Nigeria Using Advanced Machine Learning Techniques: LSTM and Xgboost, 2025, https://doi.org/10.47772/IJRISS.
[18] H.O. Ajasa, O.B. Akanbi, Predicting food prices in Nigeria using machine learning: symbolic regression, International Journal of Research and Innovation in Applied Science 10 (6) (2025) 979–995 [Online]. Available: https://ideas.repec.org/a/bjf/journl/v10y2025i6p979-995.html. (Accessed 24 August 2025).
[19] R. Zelingher, Democratising agricultural commodity price forecasting: the AGRICAF approach. https://arxiv.org/pdf/2410.20363, Oct. 2024. (Accessed 24 August 2025).
[20] M.J. MacLachlan, M.K. Adjemian, X. Etienne, M. Sweitzer, R. Volpe, W. Zeng, Adaptive food price forecasting improves public information in times of rapid economic change, Nat. Commun. 16 (1) (Dec. 2025) 6282, https://doi.org/10.1038/S41467-025-61660-X.
[21] Home | National Bureau of Statistics." Accessed: December. 29, 2022. [Online]. Available: https://nigerianstat.gov.ng/.
[22] Central bank of Nigeria :: crude oil price. https://www.cbn.gov.ng/rates/crudeoil.asp. (Accessed 29 December 2022).
[23] USD/NGN (NGN=X) live rate, chart & news - Yahoo finance. https://finance.yahoo.com/quote/NGN%3DX/history?period1=1632164022&period2=1663700022&interval=1mo&filter=history&frequency=1mo&includeAdjustedClose=true&guccounter=2. (Accessed 29 December 2022).
[24] D.M. Keenan, A tukey nonadditivity-type test for time series nonlinearity, Biometrika 72 (1) (Apr. 1985) 39–44, https://doi.org/10.1093/BIOMET/72.1.39.
[25] R. Luukkonen, P. Saikkonen, T. Terasvirta, Testing linearity against smooth transition autoregressive models, Biometrika 75 (3) (Sep. 1988) 491, https://doi.org/10.2307/2336599.
[26] Tests for specification errors in classical linear least-squares regression analysis on JSTOR. https://www.jstor.org/stable/2984219. (Accessed 26 August 2025).
[27] C.F. Baum, S. Hurn, K. Lindsay, The BDS test of independence, STATA J. 21 (2) (2021) 279–294, https://doi.org/10.1177/1536867X211025796.
[28] F. Pedregosa Fabianpedregosa, et al., Scikit-learn: machine learning in python gaël varoquaux bertrand thirion vincent dubourg alexandre passos PEDREGOSA, VAROQUAUX, GRAMFORT ET AL. Matthieu perrot, J. Mach. Learn. Res. 12 (2011) 2825–2830 [Online]. Available: http://scikit-learn.sourceforge.net. (Accessed 3 March 2023).
[29] J. Reback, et al., Pandas-Dev/Pandas: Pandas 1.2.2, Feb. 2021, https://doi.org/10.5281/ZENODO.4524629.
[30] sklearn.ensemble.RandomForestClassifier — scikit-learn 1.2.1 documentation. https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html. (Accessed 16 February 2023).
[31] A.V. Dorogush, V. Ershov, A. Gulin, Catboost: Gradient Boosting with Categorical Features Support, Oct. 2018, https://doi.org/10.48550/arxiv.1810.11363.
[32] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system. https://github.com/dmlc/xgboost. (Accessed 16 February 2023).
[33] Introduction — statsmodels. https://www.statsmodels.org/stable/index.html. (Accessed 16 February 2023).
[34] I.O. Fasanya, F. Olawepo, Determinants of food price volatility in Nigeria, Agric. Tropica Subtropica 51 (4) (Dec. 2018) 165–174, https://doi.org/10.2478/ats-2018-0019.
[35] H. Hewamalage, K. Ackermann, C. Bergmeir, Forecast evaluation for data scientists: common pitfalls and best practices. http://arxiv.org/abs/2203.10716, Mar. 2022.
[36] R. Díaz-Uriarte, S. Alvarez de Andrés, Gene selection and classification of microarray data using random forest, BMC Bioinf. 7 (1) (Jan. 2006) 1–13, https://doi.org/10.1186/1471-2105-7-3/FIGURES/1.
[37] F.X. Diebold, R.S. Mariano, Comparing predictive accuracy, Source: J. Bus. Econ. Stat. 20 (1) (2002) 134–144.
[38] R.A. Fisher, Statistical Methods for Research Workers, Oliver and Boyd, Edinburgh, Scotland. - References - Scientific Research Publishing, 1925 [Online]. Available: https://www.scirp.org/reference/referencespapers?referenceid=2056938. (Accessed 26 August 2025).
[39] D. Odion, et al., A GIS-based interactive map enabling data-driven decision-making in Nigeria's food supply chain, MethodsX 10 (Jan. 2023) 102047, https://doi.org/10.1016/J.MEX.2023.102047.