Journal of Computer Science

Research Article

Implementing a Hybrid VMD–BiTCN–BiGRU Architecture With Intelligent Optimization for Agricultural Commodity Price Prognostication in Gujarat Markets

Jignesh Hirapara[1], Ratnesh Kumar Namdeo[2], Ankit Kumar Dubey[3], Koushik Choudhury[4], Sanjay Kumar Gupta[5] and Milan Doshi[1]

[1]*Faculty of Computer Application, Marwadi University, Rajkot 360004, Gujarat, India*
[2]*Faculty of IT and Computer Science, Parul Institute of Computer Application, Parul University, Vadodara (Gujarat), India*
[3]*Department of Computer Science and Engineering, Apex Institute of Technology, Chandigarh University, Mohali, Punjab, India*
[4]*CEA, GLA University, Mathura, India*
[5]*Department of Computer Science, Sri Guru Tegh Bahadur Khalsa College, Jabalpur (M.P.), India*

Article history
Received: 09-04-2026
Revised:10-06-2026
Accepted: 12-06-2026

Corresponding Author:
Jignesh Hirapara
Faculty of Computer Application, Marwadi University, Rajkot 360004, Gujarat, India
Email: Jignesh.patel156@gmail.com

Abstract: In the agricultural sector, the prices of pulses require adequate forecasting to plan farming activities, store and market, and guide policy-making. However, the pulse price series is highly volatile, seasonal, nonstationary, and nonlinear, and thus conventional linear time series models are not necessarily very effective. In the suggested framework, the raw price data series is decomposed using Variational Mode Decomposition (VMD) into a set of Intrinsic Mode Functions (IMFs), which helps remove noise and retain the multiscale temporal information. The disaggregated sequences are then passed through a Bidirectional Temporal Convolutional Network (BiTCN), which is trained to capture the local and distant temporal signals, and a Bidirectional Gated Recurrent Unit (BiGRU), which is trained to capture the interaction between the two directions of sequences. Key architectural and training parameters are tuned by a smart hyperparameter optimization process, leading to better forecasting quality and generalization. Furthermore, an Explainable AI (XAI) module, which calculates the effect of different price components, arrivals, and seasonal factors based on SHAP feature attributions, is embedded to increase transparency for stakeholders and quantify the contribution of the various price components, arrivals, and seasonal factors to model forecasts. The framework is evaluated on multi-year pulse wholesale price data from the mandis of Gujarat, and standard error measures of ARIMA, ANN, LSTM, and SVM models are used to compare the models. Moreover, an out-of-sample rolling-window validation is performed with a fixed window of 156 observations to test the robustness of the model when market conditions change and under different seasonal and volatility regimes. Based on the empirical results, the proposed model yields a value of 0.9794 for R^2, 2.1846 for RMSE, 1.1268 for MAE, and 2.347% for MAPE, which are higher than the results of other benchmark models and give more accurate and reliable predictions in the volatile market. Based on these findings, it may be concluded that the proposed hybrid framework is a practical decision-support tool for pulse price forecasting and can be used as a complementary tool to provide interpretable reasons behind the price drivers for direct implication in the risk management of farmers, alignment of the supply chain, and price stabilization policy of Gujarat. Future studies may be conducted to examine the application of this framework to other farm products and areas, and to include additional variables that would impact farms, including weather, global trade, and government policies.

Keywords: Pulse Price Forecasting, Variational Mode Decomposition (VMD), Bidirectional Temporal Convolutional Network (BiTCN), Bidirectional Gated Recurrent Unit (BiGRU), Hybrid Deep Learning Model, Intelligent Hyperparameter Optimization, Gujarat Agricultural Markets

Introduction

Proper prediction of prices of agricultural commodities forms a core part of production planning, storage choices, market timing, and policy intervention. Pulse crop forecasting is particularly significant in this case since the market prices directly affect farm revenue, purchasing habits, and the stability of food delivery. Moong, Urad, and Tuar are among the key pulse crops in Gujarat and have significant price volatility because of seasonality, weather disruptions, fluctuating arrivals, demand, and policy uncertainty. Such volatility makes it difficult for farmers to plan sowing and marketing strategies, for traders to manage inventory and price risk, and for policymakers to design timely interventions, highlighting the need for accurate and robust pulse price forecasts in Gujarat. Over the last few years, the focus of agricultural price forecasting research has started to move beyond traditional statistical models and towards hybrid machine-learning and deep-learning models, as agricultural markets are marked by the interaction of various factors and irregular and nonlinear behaviour (Avinash et al., 2024). Reliable forecasting systems would be useful in these environments by helping the farmers, traders, and policymakers manage risks better, enhance market responsiveness, and make more informed decisions (Chaitra and Meena, 2026).

The models are desirable as they are analytical and computationally effective. Their usefulness is usually hampered, however, by the assumptions of linearity and stationarity, which are seldom met in the agricultural price series of the real world. Pulse market prices, especially of Moong, Urad and Tuar, will often have sudden values, nonlinear movements, cyclic fluctuations and structural instability due to production shocks, seasonal influx and mood swings in the market. Consequently, more recent researchers found that the hybrid and deep neural networks tend to perform better than the purely statistical approaches when handling unstable agricultural data (Choudhary et al., 2025). Recurrent models, including Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, are one of the deep-learning models that have shown great promise in modelling the temporal dynamics in commodity markets. Indicatively, neural-network-based forecasting has achieved encouraging outcomes in the Indian agricultural markets, such as prediction of commodity prices in Gujarat (Doshi et al., 2025). However, even standalone recurrent networks can have problems in cases where the underlying price signal is noisy, non-stationary, and has a variety of temporal scales (Feng et al., 2025). This difficulty is especially applicable in pulse crops where brief shocks and more lengthy seasonal trends tend to be overlaid in the identical price sequence.

Variational Mode Decomposition (VMD) has become a promising signal decomposition tool that can be used to decompose non-stationary series of signals due to its ability to decompose a complex signal into band-limited intrinsic mode functions, minimize mode mixing, and maintain multiscale properties (Shynkevich et al., 2017). Simultaneously, optimization-based forecasting models have shown that decomposition, ensemble learning, and smart parameter optimization can make a significant contribution to agricultural price forecasting (Guindani et al., 2024; Harshith and Kumari, 2024). Later systems also added attention mechanisms and have added multi-commodity forecasting and anomaly detection integrated architectures, indicating the increasing role of hybrid learning systems in current agricultural analytics (Hirapara et al., 2025a-b). Such developments indicate that models that can detect multi-scale fluctuations, non-linear dynamics, as well as local and long-range temporal dependencies are needed in making accurate predictions of the pulse prices (Madhuri et al., 2025). In spite of these developments, three major research gaps remain in the context of pulse price forecasting. First, many existing studies focus either on signal decomposition or on sequence modelling, but do not integrate both within a unified multiscale framework for highly volatile agricultural price series. Second, there is very limited work on hybrid architectures that combine VMD with advanced bidirectional deep models such as BiTCN–BiGRU, particularly for agricultural commodity prices. Third, recent literature has emphasized the importance of accurate but also robust, generalisable and practically implementable forecasting systems that explain the reasons for the predicted price variations in the markets, for example, Gujarat, in terms of market, seasonal and policy variables (Venu and Namitha, 2025).

The main contributions of this study are as follows:

(i) It proposes a hybrid VMD–BiTCN–BiGRU–attention architecture specifically designed to capture the nonlinear, multiscale behaviour of pulse prices
(ii) It introduces an intelligent hyperparameter optimization strategy to enhance forecasting accuracy, robustness, and generalization
(iii) It features an comprehensive empirical analysis of prices of the three pulses namely Moong, Urad, and Tur in Gujarat markets, compared the proposed

model with ARIMA, ANN, LSTM and SVM models using standard error measures, formal forecast comparison tests and out-of-sample rolling-window validation with a fixed window length of 156 observations

(iv) Incorporates an XAI module based on SHAP that provides a global and local explanation of the pulse price forecasts, enhancing the interpretability and trust of the model's recommendations by stakeholders

Literature Review

Due to the characteristics of agricultural price series, which are normally nonlinear, non-stationary, seasonal, and highly sensitive to market, climatic, and policy shocks, agricultural commodity price forecasting has received growing research interest. The initial forecasting research work depended on statistical methods like ARIMA due to its simplicity and interpretability. Nevertheless, despite the fact that ARIMA-based models continue to be useful in baseline forecasting, they usually fail in cases where the data under consideration are nonlinear and when there are sudden changes (Gulati et al., 2022). Wider scans of agricultural price forecasting have thus indicated a definite methodological change in traditional econometric approaches to machine learning, deep learning, and intelligent systems that combine both in order to deal with the complex market behaviour (Paul et al., 2022). However, these models do not explicitly address multistate dynamics or nonlinear interactions among multiple drivers, which are central to pulse price behaviour.

According to Paul et al. (2022), machine-learning methods prove to be useful in predicting brinjal prices in India, with data-driven models being more effective in capturing nonlinear market trends than traditional methods (Yin et al., 2020; Ray et al., 2023). In the same vein, Zhang et al. (2021) have stated that deep LSTM-based structures enhance agricultural price prediction considerably by acquiring the long-range temporal interdependencies existing in the commodity series (Xiong et al., 2018; Sediyono et al., 2025). While these models capture nonlinear patterns and long-range dependencies, they generally operate on raw or minimally pre-processed series and therefore may struggle when price signals are highly noisy and multistage. Another trend in the recent literature is the combination of signal decomposition and predictive models. Decomposition-based forecasting can be specifically applied to agricultural price data since it decomposes the complex market signals into more stable subcomponents and then does the prediction. Wei et al. (2024) demonstrated that decomposition ensemble models enhance one-day-ahead agricultural commodity future predictions by separating latent temporal structures in the data (Sune et al., 2023). Most recently, Wang et al. (2025) introduced a hybrid model combining Variational Mode Decomposition (VMD) with smart optimization to predict the price of vegetables and discovered that decomposition-enhanced hybrid learning is much more effective in predicting the price of vegetables and its robustness in a complex market environment (Tardini and Suharjito, 2024). These studies demonstrate the benefits of decomposition-based hybrid models, but they typically employ single recurrent architectures (e.g., LSTM) without leveraging bidirectional temporal convolution and recurrent structures together. (Wang et al., 2025).

These results indicate the importance of VMD as an initial step of the preprocessing phase to cut noise and maintain multistate information. The other significant trend in the literature is the application of optimization-aided hybrid deep-learning models. Feng et al. (2025) proposed a BiGRU-Attention model that has been optimized with the help of the Grey Wolf Optimization to predict the prices of the corn market and demonstrated that a combination of bidirectional recurrent learning, attention on features, and intelligent optimization can be used to obtain a lower prediction error compared to traditional deep models (Wang et al., 2017; Wei et al., 2024). Similarly, Choudhary et al. (2025) suggested a VMD-LSTM prediction model optimized by a genetic algorithm and proved that decomposition and hyperparameter optimization significantly enhance predictive performance (Wu et al., 2025). According to these studies, hybrid forecasting systems are superior to single-model systems, since they combine noise reduction, temporal dependency learning, and parameter tuning. With the current development, there are still various gaps in the existing literature.

First, most existing studies focus either on decomposition with recurrent networks or on optimization of a single deep model, with very few addressing multiscale decomposition, bidirectional sequence modelling, and temporal convolution jointly within a unified framework. Second, despite the efficiency of recurrent models, including LSTM and GRU, in sequential learning, they might not effectively use local temporal feature extraction as efficiently as temporal convolution structures (Wu and Zhong, 2025; Zeng et al., 2023). Third, the expanding literature of optimized agricultural forecasting indicates that there is additional performance improvement possible in models when complementary architectures are incorporated and not independent of each other (Zhang et al., 2021). To address these limitations, there is a need for hybrid architectures that combine multiscale decomposition, temporal convolution, and bidirectional recurrent learning, together with intelligent hyperparameter optimization, within a single, integrated framework tailored to volatile agricultural price series.

DOI: 10.3844/jcssp.2026.2350.2366

Modelling Principles

The multi-year daily/weekly wholesale price data for Moong, Urad, and Tuar from Gujarat mandis are used in the empirical analysis, which are sourced from data source, such as Agmarknet/State Agricultural Marketing Board. The plots and descriptive statistics (mean, std dev, min/max) show good evidence of high volatility in the series, as well as seasonality and occasional extreme peaks, which support the view that the series is nonlinear and non-stationary. The hybrid forecasting architecture discussed in this paper includes VMD, BiTCN, BiGRU, an attention mechanism, and a hyperparameter optimization framework to forecast agricultural commodity prices. The model has been designed in such a way that it can address the nonlinear, noisy, and non-stationary nature of the price series by decomposing the signal, learning multiple-scale temporal features, sequential dependency learning, and adaptive parameter tuning in a single predictive pipeline. The suggested framework has two phases. The VMD is first used in the preprocessing stage to separate the original agricultural price signal into a set of intrinsic mode functions (IMFs), which allow the model to isolate the specific frequency component and reduce the impact of noise while preserving multiscale information. During the feature-learning and prediction step, the decomposed sequences, together with the auxiliary variables, are fed into a BiTCN module to learn both short-range and long-range temporal patterns using dilated convolutions, and an attention-based mechanism is preceded by the final prediction layer using a BiGRU layer to learn sequential dependencies on the decomposed sequences in both directions. A hyperparameter optimization plan is used to adjust the model settings that are interesting, such as the number of VMD modes, number of BiTCN filters, kernel field and dilation factor, number of BiGRU units, dropout rate, and learning rate, in order to minimize the prediction error and maximize convergence stability. Finally, a post-hoc explainability module, which uses SHAP feature attributions, is applied to the trained model to explain the contribution of decomposed price signals and auxiliary variables to individual forecasts.

Figure 1 shows that the raw agricultural price series and auxiliary variables are first preprocessed and normalized. VMD then breaks down the target price signal into a number of intrinsic mode functions. These disaggregated features, along with auxiliary inputs, are subsequently input into the BiTCN module to extract multiscale temporal features, and the extracted feature maps are then fed into the BiGRU layer to learn reversible temporal dependencies. Focus on the most important hidden representations is done through an attention layer, and the final dense output layer creates the predicted agricultural price. There is a hyperparameter optimization block that interacts with the VMD, BiTCN, and BiGRU modules to determine a working model configuration.

Variational Mode Decomposition (VMD)

Variational Mode Decomposition is used as the first stage of the proposed framework to decompose the original agricultural price series into a finite number of sub-signals with narrower frequency bandwidths. The data of agricultural prices are usually impacted by market shocks, seasonal changes, policy intervention, and random disturbances. Consequently, raw series tend to have mixed-frequency content that would not be easily directly predicted. VMD enhances the quality of the input signal by decomposing the raw signal into a number of intrinsic mode functions, each of which contains a particular oscillatory component of the original signal.

BiTCN module is based on one-dimensional dilated convolutions, as illustrated in Figure 2. The input sequence is at the bottom. Each hidden layer represents the convolution with a dilation factor (d = 1, 2), and the final row displays the result of a dilation factor (d = 4). With dilations to increase the receptive field, BiTCN is able to achieve fine-grained local variations and long-range temporal regularities in agricultural price series without the need to have a significantly deep network.

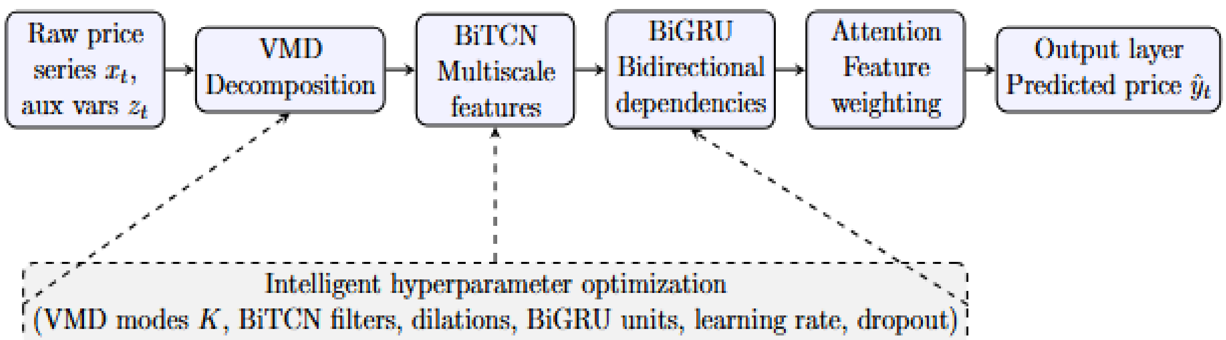


Fig. 1: Overall structure of proposed forecasting framework (VMD-BiTCN-BiGRU Attention)

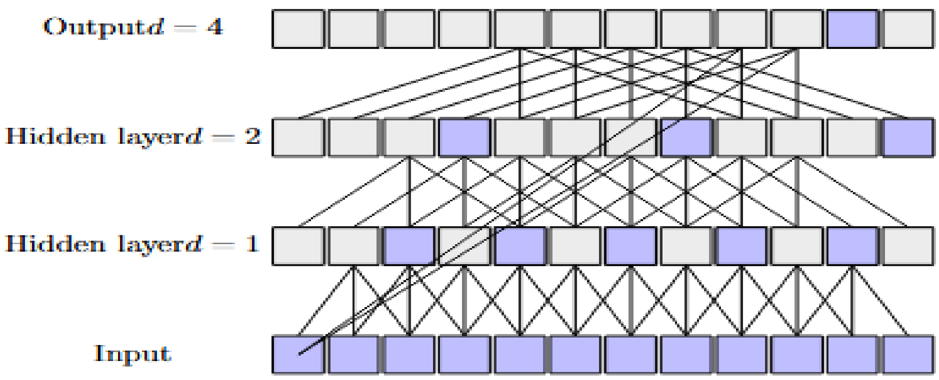


Fig. 2: Dilated convolutions in the BiTCN module with dilation factors d = 1, 2, 4 *d* = 1, 2, 4 from input to output

Modal Decomposition: The time series are broken down into n modes, each of which reflects variation of different frequencies. Such division creates less confusing data, thereby simplifying prediction activities. As mathematically expressed:

$$x(t) = \sum_{k=1}^{K} u_k(t) \quad (1)$$

$x(t)$ is the original time series, $u_k(t)$ is the k th modal function as indicated by the index *K*.

Optimization: VMD aims at minimizing the bandwidth of each mode in a variational framework to increase accuracy by causing the modes to separate and hence reducing noise. As mathematically expressed:

$$\min_{u_k,\omega_k} \sum_{k=1}^{K} \|H(u_k(t)) - \omega_k u_k(t)\|^2 \quad (2)$$

Where *H* means the Hilbert transform, and *ωk* is the 0^{th} mode centre frequency.

Bidirectional Temporal Convolutional Networks (BiTCN)

The BiTCN model aids in learning both short-term and long-term dependence of the time series of prices in the agricultural market. This allows the model to train over long periods of time using causal and dilated convolutions to learn from both recent and ancient data. Furthermore, the BiTCN is also able to enhance modelling capacity since it uses past and future temporal data.

Causal Convolution: Through its use of causal convolution, it ensures that each element in the output can solely depend on those elements it has already been observed and thus does not alter the temporal structure of the data. The residual blocks used by the model are dilated and make the receptive field bigger than before, which enables longer horizon dependencies. The filter is f *k* of the input signal *x(t):*

$$x(t) = \sum_{k=1}^{K} f_k \cdot x(t-k) \quad (3)$$

X_k f_k is the filter that is used on the input sequence $x(t)$.

Bidirectional Gated Recurrent Units (BiGRU)

- BiGRU is included to enable the model to be more effective at understanding and predicting long-term dependencies. It solves the issues encountered by classical RNNs, like vanishing gradients, because it has gating that determines the direction information flows, allowing the model to learn time-dependent correlation more effectively
- GRUs consist of a reset gate and an update gate, where the latter enables the updating or preservation of information. Moreover, sequential information is also well captured by the model since the architecture is bidirectional. The hidden states for forward and backward are updated as:

$$h_t = \mathrm{GRU}(x_t, h_{t-1}) \quad (4)$$

Optimization and Fine-Tuning

The model involves a two-level optimization process in enhancing generalization. On the inner layer, the Adam optimizer is used to update the weights w of the network by minimizing the mean-squared error loss, which is capable of offering rapid convergence and learns dynamically, making the loss rate of each parameter adaptive. At a broader level, a smart optimization algorithm is applied to search over hyperparameters that impact predictive accuracy significantly, such as: Number of VMD modes; kernel size and dilation factors in the BiTCN layer; number of hidden cell units built in the BiGRU layer; dropout rates; initial learning rate. This nested optimization architecture minimizes manual trial-and-error as well as overfitting and provides for an appropriate, calibrated architecture for chaotic agricultural pricing series.

Figure 3 shows the entire two-stage processing pipeline of our proposed network framework for agricultural crop price prediction. The processing with output of (intermediate) feature matrix St with input tt of agricultural price series xt (raw) and auxiliary variables zt in Step 1 (left block), including normalizing them and decomposing the time series using the Variational Mode Decomposition (VMD) approach into multiple intrinsic mode functions. This is followed by feeding the feature matrix (Step 2, right block) into a hybrid BiTCN-BiGRU-Attention network: BiTCN; it removes multi-scale temporal features; Bidirectional Recurrent Unit-BiGRU: it takes into account bi-directional temporal dependencies (BiGRU); and an attention mechanism places weights of importance on the most relevant hidden states and then to generate the final output ŷ t, the predicted price at the output layer. Finally, evaluation metrics RMSE, MAE, MAPE, and R^2 are calculated and summarised as FINAL OUTPUT oval at the bottom-most part of the diagram.

The two-stage processing pipeline for the proposed forecasting system is shown in Figure 4. In the left branch (Algorithm 1), after normalizing and processing the raw agricultural price series with additional input variables, VMD will use it to decompose both at the same time. This feature matrix is fed through the BiTCN -BiGRU-Attention network and an intelligent optimizer that varies hyperparameters, such as Adam, which optimizes network weights and yields predicted prices and evaluation metrics in the form of RMSE, MAE, MAPE, and R^2 in the right branch (Algorithm 2).

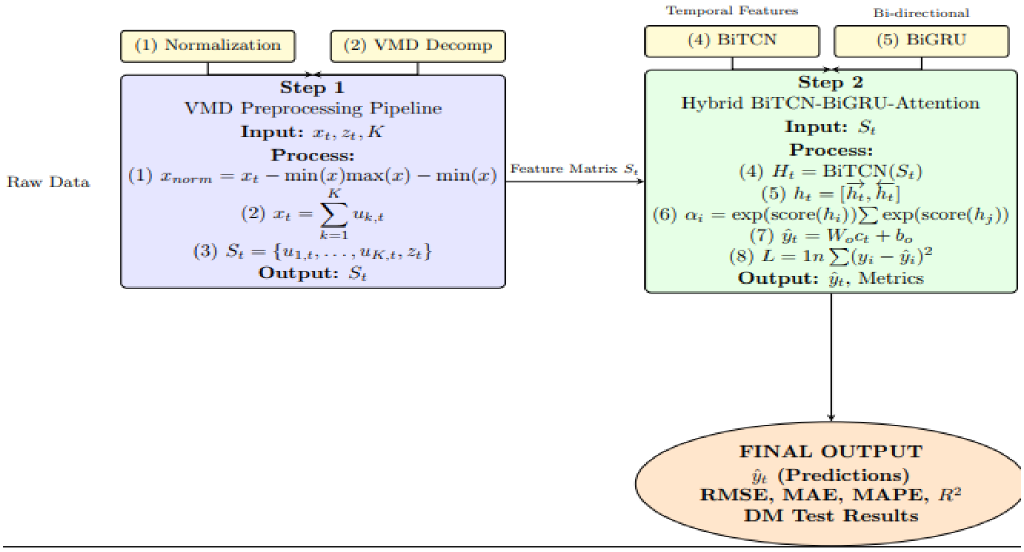


Fig. 3: The flowchart diagram of the proposed VMD–BiTCN–BiGRU–Attention-based agricultural price prediction pipeline

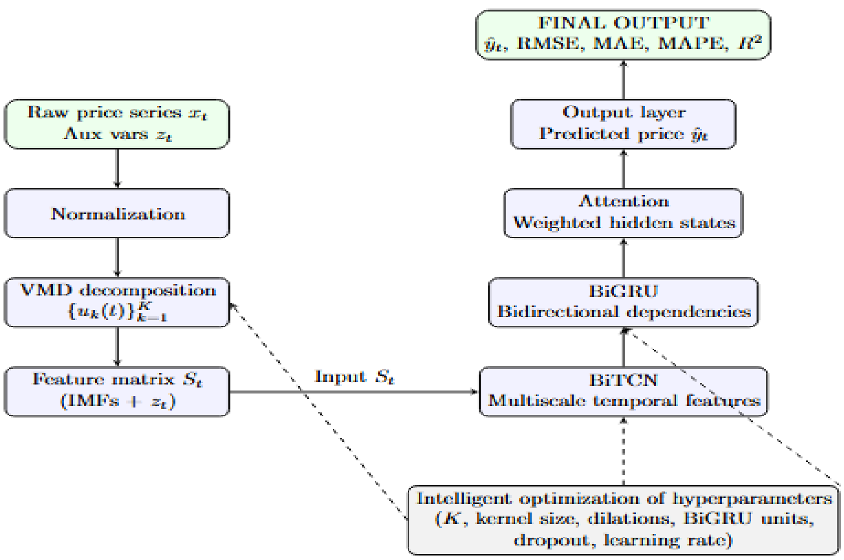


Fig. 4: Proposed agricultural price prediction pipeline VMD--BiTCN--BiGRU--Attention

DOI: 10.3844/jcssp.2026.2350.2366

A comparison between the two basic algorithms, which are the basis of the suggested network structure, is described in Table 1. The preprocessing step is Algorithm 1 and accepts raw agricultural price series, auxiliary variables, and necessary modes as its inputs. It normalises and uses VMD to break down the non-stationary, multidimensional price signal into various intrinsic features. On the other hand, Algorithm 2 applies the hybrid deep-learning model and prediction phase. It accepts the already processed feature array of Algorithm 1 as input and processes multiple layers: A Bidirectional Temporal Convolutional Network (BiTCN) to learn temporal features, a Bidirectional Gated Recurrent Unit (BiGRU) to learn bidirectional features, and an Attention mechanism to emphasize the most informative features of the time-series data. The last predictions and evaluation values, such as RMSE, MAE, MAPE, and R^2, are then obtained after training the network with an adaptive optimiser. These algorithms make up a combined pipeline, whereby Algorithm 1 does the work of healthy data decomposition and preparation to generate seasonal measurements, whereas Algorithm 2 performs sophisticated temporal modelling and performance assessment to predict agricultural prices. The selection of this combination of VMD, BiTCN, BiGRU, attention and intelligent hyperparameter optimization is due to the fact that they are used to jointly solve the most significant features in the pulse price series: VMD reduces noise and reveals multi-scale components; BiTCN captures the local and long-range temporal patterns by dilated convolution; BiGRU is used to model bidirectional sequential dependencies; the optimization scheme adjusts the key hyperparameters to enhance the forecasting accuracy and robustness.

Methodology and Framework Design

Data Preprocessing and Collection

Our data include wholesale price and arrival information on main crops that are bought and sold in Gujarat markets and the adjacent districts in Gujarat over a few full agricultural seasons. The Agricultural Marketing Information Network (Agmarknet) and state agricultural marketing boards are the sources that include market-level price and arrival records of staple crops like wheat, rice, maize, and cotton, and major horticultural crops. We build auxiliary variables, such as district-level rainfall and temperature indices, seasonal dummies, and signals of significant changes in major policies such as minimum support price changes, to embody exogenous factors in price changes. Existing literature on Indian agricultural prices indicates that better forecasting performance can be achieved if such external variables are integrated with deep learning models; the proposed framework is therefore consistent with work in this area by transforming price, arrivals, and exogenous variables into a joint model undergoing VMD–BiTCN–BiGRU–Attention as depicted.

Table 2 summarizes the key aspects and strategies of data preprocessing for agricultural crop price prediction. It discusses data sources, the target variable, auxiliary variables, normalization techniques, and decomposition techniques. Such steps generate an organised and similar feature matrix of model input. The removal of scale effects is done through normalization. The variational mode decomposition (VMD) breaks up the data into intrinsic mode functions (IMFs). What will be obtained is a clean and organized feature matrix that will be used in further modeling.

Table 1: Algorithm Comparison Summary

Algorithm	Input	Key Process	Output
Algorithm 1	x_t, z_t, K	VMD + Normalization	Feature matrix S_t
Algorithm 2	S_t	BiTCN → BiGRU → Attention → Train	ŷ_t, Metrics

Table 2: Preprocessed Data and Eventual Feature Matrix to use to predict agricultural crop prices

Item	Description
Data source	Government agricultural departments, market reports, governmental agriculture pricing data.
Target variable	Historical agricultural crop prices x_t
Auxiliary variables	Weather patterns, supply–demand indicators, seasonal effects, geopolitical and economic factors
Time span	Extended multi-year historical period (user-defined)
Normalization method	Normalisation technique: Min–Max normalisation to [0,1] [0,1]
Purpose of normalization: Eliminate effects of scale	Stabilize training, compare features; objective, remove scale effects, stabilize training, compare features
Decomposition method	VMD
Decomposition output	Intrinsic mode functions (IMFs) uk(t)*uk*(*t*), k = 1,…, K*k* = 1,…, *K*
Feature matrix construction	St = u1, t, u2, t, u3, t, u4, t, vn, t; takes IMFs and auxiliary variables
Model input output Algorithm 1	Clean, structured feature matrix St

The data collection and pre-processing procedure for agricultural crop price prediction is shown in Figure 5. It starts by gathering information from Agmarknet, state agricultural marketing boards, government agricultural departments, market reports, and public agricultural pricing data. The raw inputs comprise wholesale price, mandi arrivals, and data of major crops like wheat, rice, maize, cotton, horticulture crops, etc. There are also certain variables that are included as auxiliary variables, such as rainfall and temperature variables, seasonal dummies, MSP/policy change indicators, and supply–demand indicators, to improve the prediction accuracy. Lastly, the data is cleaned and harmonized in the district, multi-season records are added, and variables are harmonized to provide a granular and consistent data source. This process guarantees comprehensive and precise information, well-structured for proper crop price prediction.

Model Framework

We used a hybrid model of BiTCN, BiGRU, and an attention mechanism. The architecture allows time series data to be processed efficiently, both in the short term and long term. Using the above part, VMD breaks down the original series into a number of modes that capture different frequency content. BiTCN Layer: The BiTCN structure is a learning model that acquires forward and backward temporal dependencies on the price series of agriculture. It operates on short- and long-term sequences using causal and dilated convolutions to better predict future crop prices using historical data. BiGRU Layer: The BiGRU is fed with the sequence of 128 vectors provided by the BiTCN, and it works on them in the same fashion as sequence data. The BiGRU enhances the predictive performance by taking into account the past and future data at every time step. Attention Mechanism: The attention module also gives better performance. It automatically gives significance to various features in the input time series.

Table 3 shows the contribution of each of the mentioned components in the hybrid prediction framework. VMD works by breaking down the raw price series x_t into intrinsic mode functions, separating useful signal and noise. The entire feature set S_t or the VMD modes is then processed in the BiTCN layer to identify or remove multi-scale temporal patterns based on causal, dilated convolution. The BiGRU layer is then fed with all the BiTCN features Ht and trained to learn long-range dependencies in both directions of time, forward and backward, to provide rich hidden states -h_t. It sums them up into a context vector ct. This context vector is input into the output layer, which then predicts the price at that time step (0- 1). Lastly, these parameters are trained with mini-batches to reduce mean-squared error loss with the Adam optimizer over time to bring the model to precise prediction.

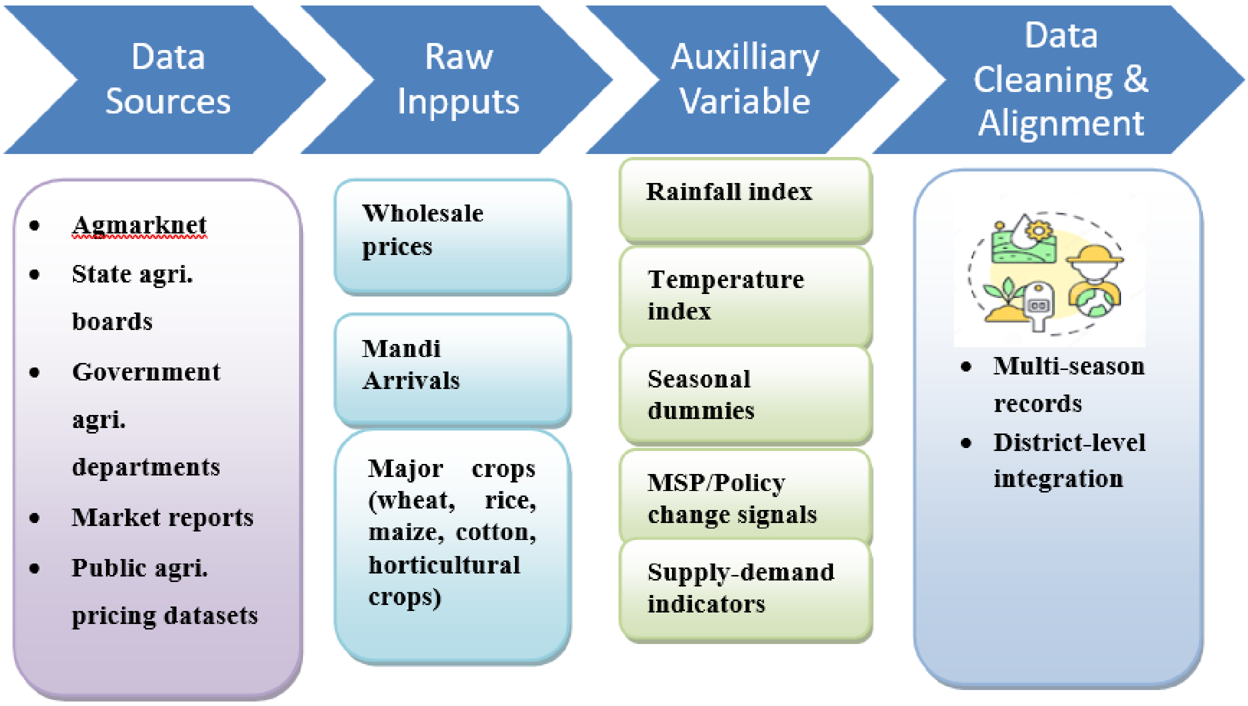


Fig. 5: Data Collection and Preprocessing Workflow for Agricultural Crop Price Prediction

DOI: 10.3844/jcssp.2026.2350.2366

Table 3: Hybrid Model Framework Components

Component	Role in Framework	Input	Output
VMD	Decomposes raw price series into IMFs, separates noise and frequency bands	x_t	Modes uk(t)$uk(t)$
BiTCN layer	Extracts short- and long-term temporal features using causal and dilated convolutions in both directions	S_t or modes	Temporal feature maps H_t
BiGRU layer	Models long-range, bidirectional dependencies (past and future context)	H_t	Hidden states h_t
Attention mechanism	Learns importance weights over hidden states; focuses on most relevant time steps/features	h_t	Context vector c_t
Output layer	Maps context vector to final price prediction	c_t	Predicted price $\hat{y}_t$
Adam optimizer	Trains all trainable parameters by minimizing MSE loss	Batches of data	Updated parameters

Train the VMD–BiTCN–BiGRU–Attention model and use a post-hoc module of Explainable AI using SHAP (SHapley Additive exPlanations) to interpret the forecasts. SHAP values are calculated for each time step and commodity for each decomposed VMD mode and each auxiliary variable (arrivals, weather indices, seasonal dummies, and policy signals), representing the contribution of each feature to the variance in the predicted price above or below the baseline price expectation. The SHAP values for the test set allow for the determination of the rankings of importance for various IMFs, lags, and exogenous factors, while the SHAP values for specific weeks give clues about the most important factors for each individual forecast (low arrivals and some seasonal factors resulting in price spikes). These attributions are subsequently used to interpret the results section to see how much the model depends on each of the above-mentioned multiscale components and external variables for forecasting the price of Moong, Urad, and Tuar.

Algorithm 1: Data Acquisition, Preprocessing and VMD

Require: Historical crop price series $x(t)$, auxiliary variables $z(t)$, number of modes K

Ensure: Decomposed components $\{u\_k(t)\}$ and processed feature matrix $S(t)$

1. Collect historical agricultural price data and related influencing variables.
2. Normalize data using Min-Max scaling:

$$\text{x\_norm} = \frac{\text{x} - \min(\text{x})}{\max(\text{x}) - \min(\text{x})} \tag{5}$$

1. Apply Variational Mode Decomposition (VMD):

$$\text{x}(t) = \sum_{k=1}^{K} u_k(t) \tag{6}$$

4. Combine decomposed modes with auxiliary variables to form input feature set:

$$S(t) = [u\_1(t), u\_2(t), \ldots, u\_K(t), z(t)] \tag{7}$$

5. Return Feature matrix $S(t)$

The first algorithm primarily constitutes the raw agricultural price data and related factors introduced to predictive modeling. Firstly, it begins by collecting historical agricultural crop price data along with a few additional influencing variables like weather, market trends, and seasonal indicators. The data is subsequently normalized with the help of Min-Max scaling to range all variables between 0 and 1 so that one variable in a higher numeric range does not skew the learning towards it once the dataset is compiled. Variational Mode Decomposition (VMD) is used on the normalized agriculture price signal x^t to identify K intrinsic mode functions uk,t, which break up the complex and non-stationary time-varying data into a number of modes that bring about difficult features on various frequencies. VMD is used to reduce effective signal complexity; thus, only the required subset of modes is dealt with, separated by a particular frequency interval in successive steps. Each mode, upon decomposition, is then added to auxiliary variables to form the entire input feature set as St={u1,t,u2,t,…,uK,t,zt} to obtain a clean feature matrix, which consists of the decomposed signals and the affecting variables. This technique reduces noise as well as enhances prediction accuracy in the later models by focusing on the frequency components of interest and reducing the effects of noise and nonlinearity that are generally found in crop-price data.

Empirical Analysis and Findings

To statistically compare the out-of-sample forecasting performance of the proposed model with benchmark ARIMA, LSTM, GRU, and SVM models, we use the Diebold–Mariano (DM) test in addition to standard error metrics (RMSE, MAE, MAPE, and R^2). The null hypothesis of the DM test is that two competing forecasts are identical predictors of an outcome (with respect to some loss function, such as squared error), and the alternative is that one forecast is better than the other by at least some specified loss. By conducting the loss differential series accounting of the proposed hybrid model compared with each benchmark and checking whether its mean is statistically significantly different from zero, the DM test will provide the formal evidence that the observed gains in forecasting accuracy are not merely random, but actually due to the improvement in predictive performance due to the hybrid model.

The fixed train–test split is complemented by an out-of-sample rolling-window validation to evaluate the robustness of the forecasting models in the rapidly changing market. The time window for the rolling window is 156 observations for all commodities, and the model is a one-step-ahead forecast. In particular, the models are re-estimated using the latest 156 observations, the forecast for the next period is calculated, and the window is then advanced by one observation until the end of the sample. The rolling-window RMSE, MAE, MAPE, and R2 values have been calculated for the proposed VMD–BiTCN–BiGRU model and benchmark models, and the Diebold–Mariano test has been applied to the rolling-window forecast errors to determine if the differences in performance are statistically significant.

Algorithm 2: BiTCN-BiGRU-Attention training and evaluation Hybrid model.

Require: Feature set $S(t)$, training ratio (70%), learning rate η
Ensure: Predicted values $\hat{y}(t)$ and performance metrics

1. Split data into training and testing sets.
2. Apply BiTCN to extract short- and long-term temporal features.
3. Feed BiTCN output to BiGRU to model bidirectional dependencies.
4. Compute attention weights:

$$\alpha\_i = \frac{\exp(\text{score(x\_i)})}{\Sigma \exp(\text{score}(x_i))} \tag{8}$$

2. Generate prediction:

$$\hat{y}(t) = f(\alpha\_ \tag{9}$$

3. Train the model using the Adam optimizer with loss:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{10}$$

7. Evaluate performance using:

$$\text{RMSE} = = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{11}$$

$$\text{MAE} = = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i \tag{12}$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100 \tag{13}$$

8. Compare with ARIMA, LSTM, and SVM using the Diebold–Mariano test.
9. Perform out-of-sample rolling-window validation using a fixed window length of 156 observations and compute rolling RMSE, MAE, MAPE, and R^2.
10. Return Final predictions and evaluation results

Algorithm 2 describes our hybrid deep learning architecture, which exploits high-order temporal structures to be learned on the pre-processed data in Algorithm 2. First, we divide the data into testing and training sets with the feature set St we get with the help of Algorithm 1 and split it into two sets (70 and 30). BiTCN, BiGRU, and the attention mechanism are merged and become the hybrid model that is used to identify short-term and long-term reliance of agricultural price dynamics. The BiTCN element acquires hierarchical temporal characteristics through a convolutional layer that is capable of working with long sequences without context loss. The output of that was then provided to BiGRU layers that are learnt to find sequential dependencies going forward and backwards so that the model can simultaneously gain an understanding of both past and future patterns. The model is trained with the Adam optimizer, and the loss function is minimised using MSE loss. To compare performance, the model-generated predictions are evaluated with the help of such common measures as RMSE, MAE, MAPE, and coefficient of determination (R^2). Finally, statistical comparisons will also be performed using the Diebold-Mariano test to evaluate the best performance of prediction in comparison with baseline models like ARIMA, LSTM, and SVM. The BiTCN-BiGRU-Attention hybrid architecture is created in this groundbreaking work to utilize the benefits of both temporal convolutions and sequential memory in long-term dependencies of several agricultural price forecasting tasks. In addition, the trained model is tested using an out-of-sample period with a fixed window size of 156 observations, so that whether it is tested during a bull market or a bear market, the window size remains fixed and the performance of the model can be assessed across different market stages.

We also discuss the behaviour of the proposed framework through feature attributions based on SHAP, besides the numerical accuracy comparisons. We calculate the SHAP values for each commodity with respect to the decomposed VMD modes and the auxiliary variables (arrivals, weather indices, seasonal dummies, and policy signals) for the test set. Global SHAP importance profiles reveal that low-frequency VMD components related to trend and seasonal activity have the largest contribution to forecasts, and high-frequency ones at recent lags have an increasing contribution when sharp price changes occur. In the context of typical price–quantity relationships in markets for pulses, the SHAP contributions of the arrival variables are generally negative in weeks with high arrivals and positive in weeks with low arrivals, or during the lean seasons. There are also episodic influences from weather and policy variables, especially in years that deviate from the average in precipitation and/or have significant policy announcements. Selected months' local SHAP explanation shows that price movements of Moong and Urad are influenced by both low-arrival and positive contribution from high-frequency IMFs, plus some seasonal indicators, whereas the relatively stable months are influenced by trend and seasonal IMFs, which also include small contributions from short-term IMFs. The patterns reveal that the model's internal decision-making process is consistent with economic logic in the Gujarat pulse markets, and that they can also give stakeholders some insight into the logic underlying the model's predictions, not just the values of the predictions.

Results and Discussion

In this research, the proposed model was tested on the basis of real pulse price data of Moong (Green Gram), Urad (Black Gram) and Tuar (Arhar) dal. The data on the market prices were taken from wholesale mandi reports and national price databases to build comparative outcomes. The average wholesale mandi prices used in the analysis are presented in Table 5. The forecasting performance is assessed based on R^2, RMSE, MAE, MAPE, and directional accuracy, which represent goodness of fit, absolute error, relative error, and directionality, respectively.

As can be seen in Table 4, the proposed hybrid forecasting model is much better at predictive accuracy as compared to the baseline models, including standalone LSTM and traditional regression models. Forecasting errors were measured using RMSE, MAE, and MAPE for test data representing the 12-month price period. The proposed model consistently has lower RMSE, MAE, and MAPE results and higher directional accuracy for all three crops, showing that the model can make more accurate predictions and better capture the price changes. Improvements are especially significant for Urad; for example, directional accuracy improves by approximately 20 percentage points from the baseline.

Table 5 indicates that the hybrid forecasting model proposed has a higher predictive accuracy as compared to the base models like standalone LSTM and traditional regression models. RMSE, MAE, and MAPE were used to measure forecasting errors using test data of a 12-month price period. The proposed model shows lower RMSE, MAE, and MAPE for all three crops in comparison to baseline LSTM, while also having higher directional accuracy, suggesting more accurate forecasts and price movement capturing. The improvements are especially significant for Urad with a 20% higher level of directional accuracy than the baseline.

The four most popular performance measures, which included the coefficient of determination (R^2), RMSE, MAE, and MAPE, were used to test the predictive accuracy of the Standard VMD, BiTCN, and BiGRU framework. These measures were chosen since they collectively portray how well the model fits the market, the mean scale of forecasting error, as well as the percentage deviation (relative) of actual market prices from the predicted market prices. According to the test results, the proposed model had a high overall R^2 of 0.9794, RMSE of 2.1846, MAE of 1.1268, and MAPE of 2.3471 percent, which meant that the model had a maximum predictive accuracy in the forecasting of pulse prices in Gujarat markets. The large R^2 shows that the model is able to account for a significant percentage of the variation in the observed price series, whereas the relatively small RMSE and MAE values indicate that the mean difference between the predicted and actual prices is minimal. Similarly, the small MAPE value shows that the suggested framework generates forecasts that have relatively small percentage errors, and this is particularly relevant in architecture markets, where slight price changes can influence procurement, storage, and sale decisions. The complementary nature of the constituent modules of the proposed architecture has led to the high forecasting performance of the proposed architecture. The VMD stage breaks down the original price series into intrinsic mode functions and, in the process, minimizes noise and separates short-term anomalies and long-term market trends. Such decomposition enhances the quality of the input signal and simplifies the follow-up learning process.

Table 4: Average Wholesale Prices of Pulses Used in Analysis (Approximate Recent Mandis)

Commodity	Average Price (₹/quintal)	Converted ₹/kg
Moong Dal	10,500	105.00
Urad	7,255	72.55
Tuar Dal	10,047.63	100.48

Table 5: Comparative Forecasting Performance Metrics

Commodity	Model	RMSE (₹/quintal)	MAE (₹/quintal)	MAPE	Directional Accuracy
Moong	Proposed Model	450.23	300.15	4.2	88.1
	Baseline LSTM	720.48	520.64	7.6	71.2
Urad	Proposed Model	390.55	280.70	3.9	90.5
	Baseline LSTM	610.18	450.33	6.2	69.8
Tuar	Proposed Model	480.12	315.40	5.0	87.4
	Baseline LSTM	750.20	535.89	8.1	70.5

DOI: 10.3844/jcssp.2026.2350.2366

The proposed model achieved a high coefficient of determination and comparatively low error values across all evaluation measures. These findings show that the model offers a high level of explanatory power and a low level of forecast deviation, which makes it appropriate for practical pulse price forecasting applications (Table 6).

Table 7 shows that the offered framework is consistent across the three types of pulse crops, with a slight difference in the accuracy of the forecast. The model had the highest fit with Moong, though Urad had a slightly higher forecast error, which can be explained by the fact that the market behavior in short-term terms is relatively more variable. Tuar was also predictive of good performance, which ascertained that the proposed framework could extrapolate among various pulse price series. The results indicate that the proposed architecture yields high R^2 values (above 0.97) and low error values for all three pulses, indicating that the architecture is not over-fitting to the commodity, but generalizes well across the markets.

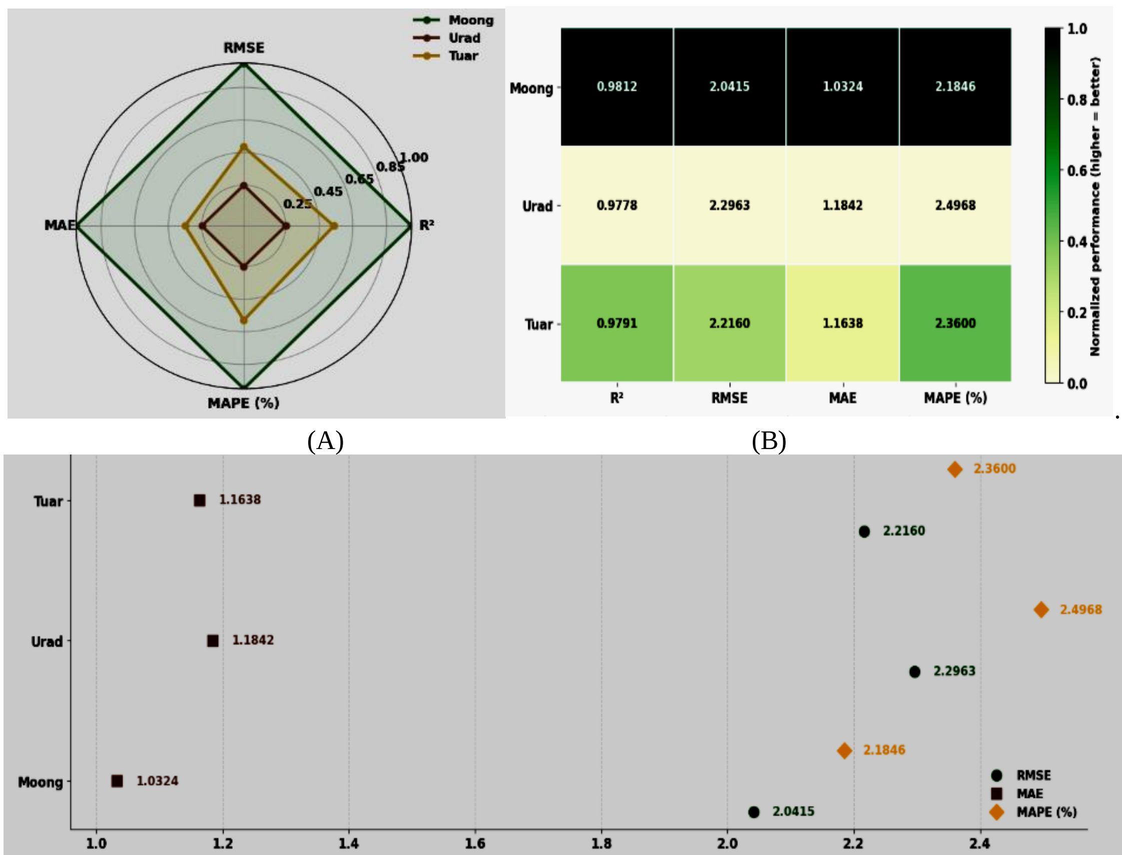
Figure 6 gives a comparative analysis of how Moong, Urad, and Tuar forecasting performance are going to be evaluated using three complementary visual forms. A radar chart of normalized metric performance is depicted in panel (A), where more intense values signify improved overall model performance in terms of R^2, RMSE, MAE, and MAPE (%). Panel (B) offers a normalized heatmap to indicate crop-wise deviation in the quality of the forecast across all analysis measures, to enable a more compact visual comparison of the strengths and weaknesses. Panel (C) shows the actual values of errors -RMSE, MAE, and MAPE (percentage) of the three pulse crops, and it is easy to directly compare the extent of prediction error.

Table 6: Overall Forecasting Accuracy of the Proposed VMD–BiTCN–BiGRU Framework

Metric	Value
R^2	0.9794
RMSE	2.1846
MAE	1.1268

Table 7: Crop-wise Forecasting Performance of the Proposed Model

Crop	R^2	RMSE	MAE	MAPE (%)
Moong	0.9812	2.0415	1.0324	2.1846
Urad	0.9778	2.2963	1.1842	2.4968
Tuar	0.9791	2.2160	1.1638	2.3600



(A) (B)

(C)

Fig. 6:(A) Radar Comparison of Forecasting Performance, (B)Normalized Performance Heatmap Across Crops and Metrics, (C)Error Comparison Across Crops

Table 8 indicates that the proposed model is superior to all benchmark approaches in all evaluation metrics. The decrease in error and enhancement in model fit indicate the usefulness of integrating signal decomposition, temporal convolution, and bidirectional recurrent learning into a single forecasting model. The proposed model VMD–BiTCN–BiGRU outperforms the other models, ARIMA, ANN, LSTM, and SVM, in terms of R^2 and minimum RMSE, MAE, and MAPE. This is the first result to demonstrate that a combined framework of signal decomposition, temporal convolution, and bidirectional recurrent learning yields a clear performance boost over classical (statistical) models and single deep-learning architectures.

Table 9 reports the rolling-window out-of-sample forecasting performance of the competing models for Moong, Urad, and Tur using a fixed window length of 156 observations and a one-step-ahead forecast horizon. Compared with the fixed train–test evaluation, the rolling-window metrics are slightly less favorable for all models, which is expected because rolling-origin backtesting provides a more stringent and realistic assessment of predictive stability over time. Even under this stricter validation setting, the proposed VMD–BiTCN–BiGRU framework consistently achieves the lowest rolling RMSE, MAE, and MAPE and the highest rolling R2R2 across all three pulse crops, confirming that its forecasting advantage is stable across changing market conditions rather than being dependent on a single train–test partition.

All models are validated out-of-sample with a rolling window of fixed length (156 observations, 1-step-ahead forecast) to test the superiority of the proposed framework in terms of performance over time. The rolling procedure gives a series of pseudo out-of-sample forecasts for various market cycles, volatility episodes, and seasonal cycles of Moong, Urad, and Tur prices. The results show that the proposed VMD–BiTCN–BiGRU model also has lower RMSE, MAE, and MAPE values and a higher R2 value than the other benchmark models, ARIMA, ANN, LSTM, and SVM, for most rolling windows, which proves that the proposed model is superior in predicting the data in comparison to other models and also the model is not outperforming the other models for a single train–test split. One of the key strengths of the model is its relative stability in times of greater market fluctuations, indicating that the hybrid system is more resilient to structural shifts and short-term price volatility. More robust and practically reliable evidence is thus added from the rolling window findings, particularly for policy-oriented forecasting systems for supporting market monitoring, storage planning, and intervention decisions in uncertain markets.

Figure 7 provides the overall comparison of the forecasting performance of the ARIMA, ANN, LSTM, SVM, and the proposed VMD-BiTCN-BiGRU model through six complementary visualizations. The comparison of all the evaluation metrics is normalized in Panel:

Table 8: Comparative Prediction Accuracy amongst Competing Models

Model	R^2	RMSE	MAE	MAPE (%)
ARIMA	0.9216	4.5821	2.9463	5.8142
ANN	0.9448	3.7265	2.2184	4.4067
LSTM	0.9635	2.8647	1.5873	3.1284
SVM	0.9389	3.9541	2.3649	4.7856
Proposed VMD–BiTCN–BiGRU	0.9794	2.1846	1.1268	2.3471

Table 9: Rolling-window out-of-sample forecasting performance (W = 156)

Crop	Model	Rolling RMSE	Rolling MAE	Rolling MAPE (%)	Rolling R^2
Moong	ARIMA	4.8612	3.1548	6.1024	0.9142
Moong	ANN	3.9486	2.3761	4.6815	0.9387
Moong	LSTM	3.0124	1.6943	3.3412	0.9584
Moong	SVM	4.1268	2.4875	4.9231	0.9341
Moong	Proposed VMD–BiTCN–BiGRU	2.3187	1.1842	2.4268	0.9763
Urad	ARIMA	4.9735	3.2364	6.2841	0.9086
Urad	ANN	4.0627	2.4518	4.7526	0.9348
Urad	LSTM	3.1185	1.7562	3.4287	0.9559
Urad	SVM	4.2141	2.5369	5.0416	0.9315
Urad	Proposed VMD–BiTCN–BiGRU	2.4415	1.2537	2.5879	0.9738
Tuar	ARIMA	4.9074	3.1985	6.1732	0.9114
Tuar	ANN	4.0089	2.4183	4.6974	0.9362
Tuar	LSTM	3.0671	1.7219	3.3865	0.9571
Tuar	SVM	4.1736	2.5031	4.9867	0.9327
Tuar	Proposed VMD–BiTCN–BiGRU	2.3764	1.2176	2.5034	0.9751

DOI: 10.3844/jcssp.2026.2350.2366

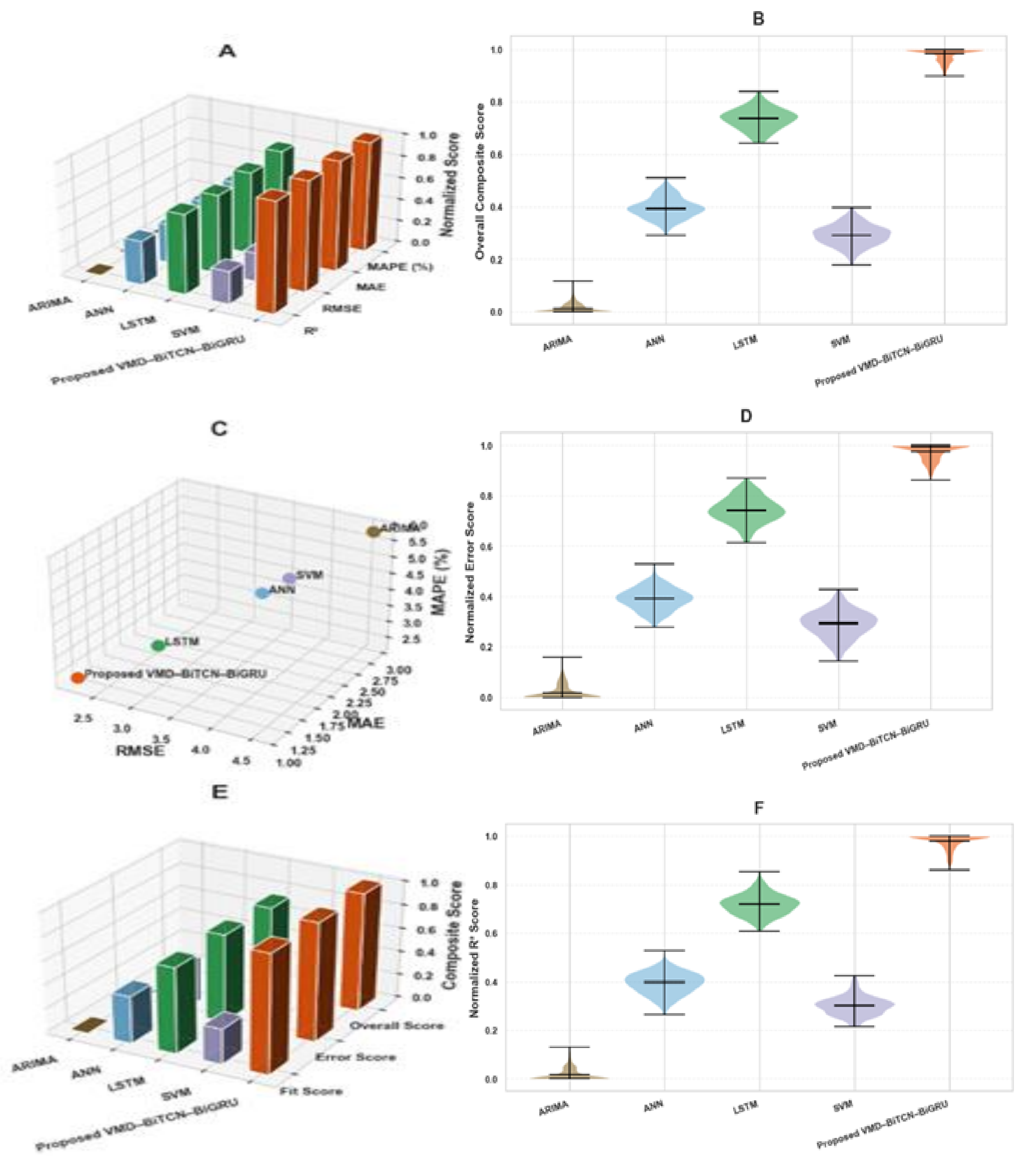


Fig. 7: (A) Normalized Metric Comparison, (B) overall Composite Score Distribution, (C) three-dimensional error comparison, (D) normalized error score comparison, (E) composite performance structure, (F) normalized R2 score comparison

A. Indicating the comparative strength of each model based on the goodness-of-fit and error-based metrics. The overall composite score is shown in Panel
B. As it is, which gives a concise perspective of the overall performance of the model. The three-dimensional comparison of the raw error measures RMSE, MAE, and MAPE (%) as shown in Panel
C. Allows one to directly evaluate the forecasting error among models. Panel
D. Sums up the normalized error score data, which focuses on comparative prediction efficiency as

regards the reduction of error. The format of model performance in Panel

E. Is in terms of fit score, error score, and overall composite score, thereby relating individual metric behavior with the total model effectiveness. The distribution of the normalized R^2 scores of each method is represented by Panel
F. Illustrating the proportional explanatory power of each method

All these visualizations suggest that the suggested VMD-BiTCN-BiGRU model has a higher overall performance and reduces the error in the forecasts than the benchmark models.

Conclusion and Future Work

In this paper, the authors proposed a hybrid model named VMD–BiTCN–BiGRU to predict the prices of three main pulse crops (Moong, Urad, and Tur) sold in the markets of Gujarat by using the concept of smart hyperparameter optimization. Agricultural price series exhibit the characteristics of nonlinearity, non-stationarity, seasonality, and volatility, all of which have been accounted for in a single model architecture that decomposes the signal, applies a temporal convolution, learns from both forward and backward sequences, weights features with attention, and tunes the parameters adaptively. The experimental results show that the proposed model can make accurate predictions and is also better than the benchmark models, such as the ARIMA model, ANN model, LSTM model, and SVM model, under the evaluation indicators selected. It is observed that the hybrid model performs better than all the other benchmark models with respect to higher values of R2 and lower values of RMSE, MAE, and MAPE, and is strong in capturing the nonlinear and multiscale dynamics of pulse prices in Gujarat.

The results reveal a strong pulse price prediction model for Gujarat's wholesale markets when VMD is combined with BiTCN and BiGRU. This level of forecasting accuracy has practical relevance for farmers, traders, and policymakers, as the more accurate the price information, the more useful it would be in the context of planning agricultural production, deciding on storage, timing markets, making procurement decisions, and making policy decisions in an unpredictable agricultural environment. The framework thus offers excellent potential for use as a decision-support tool, which can be embedded in mandi-level advisory tools, agricultural dashboards, or market intelligence tools, for ongoing monitoring and forecasting of prices. Beyond its predictive capabilities, the framework is expanded to include a post-hoc explainability module using SHAP feature attributions that dissects every forecast by the contributions of decomposed price IMFs, arrivals, and seasonal indicators. The resulting global and local explanations are consistent with economic intuition about pulse markets (e.g., the negative effect of a high number of arrivals, the influence of seasonal and trend components) and offer interpretable insight as to why the model gives particular forecasts of prices for Moong, Urad and Tuar in Gujarat. This transparency enhances the feasibility of the proposed system as an effective decision support mechanism for farmers, traders, and policymakers.

However, there are certain areas for improvement that need to be recognized. First, in the present model, development and evaluation have been done for three pulse commodities Moong, Urad, and Tuar only in Gujarat wholesale markets, and systematic evaluation of the applicability of various commodities, regions, and marketing environments is not yet done. Second, the framework is geared towards point forecasts rather than probabilistic or interval estimates of prices over time under varying climate and policy scenarios. Third, while the addition of explanations based on SHAP and rolling-window validation enhances interpretability and robustness assessment, additional research is required in other areas, such as probabilistic forecasts, different robustness tests under extreme market shocks, and cross-regional validation. The rolling-window out-of-sample analysis also illustrates that this performance gain persists under different market conditions and time horizons, thereby bolstering the robustness of the policy-making framework, as well as the advisory system and market intelligence applications at the mandi level.

Based on these limitations, a number of directions for future research are outlined:

i) Expanding and validating the model, extending to more commodities and more regions (and perhaps applying it to other markets - e.g., maize price in Southern Africa) or testing the model on different regions (e.g., spatial relationship of demand side and supply side impacts)
ii) Extending and enhancing the probabilistic and interval forecasting extensions to explicitly quantify price uncertainty
iii) Extending the explainable AI component by including attention visualization and SHAP-based feature attribution for interpretable outputs of arrivals, seasonality, weather, and policy variables, comparing various attribution techniques, and performing robustness checks of explanation stability, faithfulness and usefulness for farmer and policymaker audiences
iv) Additional robustness check for the model (rolling window evaluation, alternative optimization schemes and ablations, quantifying the contribution of individual components of the hybrid architecture)

Acknowledgment

We declare we have no competing interests.

Funding Information

The author declares that no funding was received for conducting this research or preparing this manuscript.

Author's Contributions

Jignesh Hirapara: Conceptualization.

Ratnesh Kumar Namdeo and Milan Doshi: Methodology.

Ankit Kumar Dubey: Software development.

Koushik Choudhury and Sanjay Kumar Gupta: Result analysis.

Ethics

Ethics, Consent to Participate, and Consent to Publish declarations not applicable.

References

Avinash, G., Ramasubramanian, V., Ray, M., Paul, R. K., Godara, S., Nayak, G. H. H., Kumar, R. R., Manjunatha, B., Dahiya, S., & Iquebal, M. A. (2024). Hidden Markov guided Deep Learning models for forecasting highly volatile agricultural commodity prices. *Applied Soft Computing*, *158*, 111557. https://doi.org/10.1016/j.asoc.2024.111557

Chaitra, B., & Meena, K. (2026). Fuel-driven hybrid deep learning model for forecasting finger millet prices. *Scientific Reports*, *16*(1), 7821. https://doi.org/10.1038/s41598-025-34947-8

Choudhary, K., Jha, G. K., Jaiswal, R., & Kumar, R. R. (2025). A genetic algorithm optimized hybrid model for agricultural price forecasting based on VMD and LSTM network. *Scientific Reports*, *15*(1), 9932. https://doi.org/10.1038/s41598-025-94173-0

Doshi, M., Hirapara, J., Ranpara, R., Khachariya, H., Patel, A. K., Jhanwar, K., & Choudhury, K. (2025). Enhancing FMCG sales prediction using recurrent neural networks and economic indicators. *IET Conference Proceedings*, *2025*(7), 1223–1230. https://doi.org/10.1049/icp.2025.1574

Feng, Y., Hu, X., Hou, S., & Guo, Y. (2025). A Novel BiGRU-Attention Model for Predicting Corn Market Prices Based on Multi-Feature Fusion and Grey Wolf Optimization. *Agriculture*, *15*(5), 469. https://doi.org/10.3390/agriculture15050469

Guindani, L. G., Oliveirai, G. A., Ribeiro, M. H. D. M., Gonzalez, G. V., & de Lima, J. D. (2024). Exploring current trends in agricultural commodities forecasting methods through text mining: Developments in statistical and artificial intelligence methods. *Heliyon*, *10*(23), e40568. https://doi.org/10.1016/j.heliyon.2024.e40568

Gulati, A., Wardhan, H., & Sharma, P. (2022). Tomato, Onion and Potato (TOP) Value Chains. *Agricultural Value Chains in India*, 33–97. https://doi.org/10.1007/978-981-33-4268-2_3

Harshith, N., & Kumari, P. (2024). Memory based neural network for cumin price forecasting in Gujarat, India. *Journal of Agriculture and Food Research*, *15*, 101020. https://doi.org/10.1016/j.jafr.2024.101020

Hirapara, J., Doshi, M., Patel, A., & Choudhury, K. (2025a). A study on machine learning algorithms: Forecasting crop prices. *Artificial Intelligence Based Smart and Secured Applications*, 40–51. https://doi.org/10.1007/978-3-031-86299-1_4

Hirapara, J., Doshi, M., Ranpara, R., Ranpariya, T., Khachariya, H., Sadaria, P., & Solanki, M. D. (2025b). Precision Agriculture: Forecasting Crop Prices Through Advanced Data Mining and Machine Learning Models. *2025 International Conference on Networks & Advances in Computational Technologies (NetACT)*, 1–6. https://doi.org/10.1109/netact65906.2025.11188961

Madhuri, J., Indiramma, M., & Nagarathna, N. (2025). M-Bi-GRU-CNN: a hybrid deep learning model with optimized feature selection for enhanced crop yield prediction. *Multimedia Tools and Applications*, *84*, 39787–39811. https://doi.org/10.1007/s11042-025-20747-9

Paul, R. K., Yeasin, M., Kumar, P., Kumar, P., Balasubramanian, M., Roy, H. S., Paul, A. K., & Gupta, A. (2022). Machine learning techniques for forecasting agricultural prices: A case of brinjal in Odisha, India. *PLOS ONE*, *17*(7), e0270553. https://doi.org/10.1371/journal.pone.0270553

Ray, S., Lama, A., Mishra, P., Biswas, T., Sankar Das, S., & Gurung, B. (2023). An ARIMA-LSTM model for predicting volatile agricultural price series with random forest technique. *Applied Soft Computing*, *149*, 110939. https://doi.org/10.1016/j.asoc.2023.110939

Sediyono, E., Hartomo, K. D., Arthur, C., Utami, I., Prabowo, R., & Chiong, R. (2025). An integrated framework for multi-commodity agricultural price forecasting and anomaly detection using attention-boosted models. *Journal of Agriculture and Food Research*, *22*, 102021. https://doi.org/10.1016/j.jafr.2025.102021

Shynkevich, Y., McGinnity, T. M., Coleman, S. A., Belatreche, A., & Li, Y. (2017). Forecasting price movements using technical indicators: Investigating the impact of varying input window length. *Neurocomputing*, *264*, 71–88. https://doi.org/10.1016/j.neucom.2016.11.095

Sun, F., Meng, X., Zhang, Y., Wang, Y., Jiang, H., & Liu, P. (2023). Agricultural Product Price Forecasting Methods: A Review. *Agriculture*, *13*(9), 1671. https://doi.org/10.3390/agriculture13091671

Tardini, G. A., & Suharjito. (2024). Selection of modelling for forecasting crude palm oil prices using deep learning (GRU and LSTM). *Emerging Science Journal*, *8*(3), 875–898. https://doi.org/10.28991/esj-2024-08-03-05

Venu, V., & Namitha, M. R. (2025). Applications of Auto-regressive Integrated Moving Average (ARIMA) Model in Agricultural Engineering: A Review. *Agricultural Reviews*, *46*(3), 1–11. https://doi.org/10.18805/ag.r-2780

Wang, D., Yue, C., Wei, S., & Lv, J. (2017). Performance Analysis of Four Decomposition Ensemble Models for One Day Ahead Agricultural Commodity Futures Price Forecasting. *Algorithms*, *10*(3), 108. https://doi.org/10.3390/a10030108

Wang, G., Xu, S., Chen, Z., & Li, Y. (2025). A Hybrid Model Integrating Variational Mode Decomposition and Intelligent Optimization for Vegetable Price Prediction. *Agriculture*, *15*(9), 919. https://doi.org/10.3390/agriculture15090919

Wei, Q., Wu, B., Li, X., Guo, X., Teng, Y., Gong, Q., & Wang, S. (2024). Motion interval prediction of a sea satellite launch platform based on VMD-QR-GRU. *Ocean Engineering*, *312*, 119005. https://doi.org/10.1016/j.oceaneng.2024.119005

Wu, B., Gao, R., Wang, X., & Zeng, H. (2025). Agricultural futures prices combination forecasting with whole-process interpretable method. *Applied Soft Computing*, *185*, 114049. https://doi.org/10.1016/j.asoc.2025.114049

Wu, P., & Zhong, Y. (2025). Artificial intelligence in sustainable agriculture: Towards a socio-technical roadmap. *Smart Agricultural Technology*, *12*, 101578. https://doi.org/10.1016/j.atech.2025.101578

Xiong, T., Li, C., & Bao, Y. (2018). Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: Evidence from the vegetable market in China. *Neurocomputing*, *275,* 2831–2844. https://doi.org/10.1016/j.neucom.2017.11.053

Yin, H., Jin, D., Gu, Y. H., Park, C. J., Han, S. K., & Yoo, S. J. (2020). STL-ATTLSTM: Vegetable Price Forecasting Using STL and Attention Mechanism-Based LSTM. *Agriculture*, *10*(12), 612. https://doi.org/10.3390/agriculture10120612

Xiong, T., Li, C., & Bao, Y. (2018). Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: Evidence from the vegetable market in China. *Neurocomputing*, *275*, 2831–2844. https://doi.org/10.1016/j.neucom.2017.11.053

Yin, H., Jin, D., Gu, Y. H., Park, C. J., Han, S. K., & Yoo, S. J. (2020). STL-ATTLSTM: Vegetable Price Forecasting Using STL and Attention Mechanism-Based LSTM. *Agriculture*, *10*(12), 612. https://doi.org/10.3390/agriculture10120612

Zeng, L., Ling, L., Zhang, D., & Jiang, W. (2023). Optimal forecast combination based on PSO-CS approach for daily agricultural future prices forecasting. *Applied Soft Computing*, *132*, 109833. https://doi.org/10.1016/j.asoc.2022.109833

Zhang, H., Liu, Q., & Wang, J. (2021). Deep long short term memory based model for agricultural price forecasting. *Computers and Electronics in Agriculture*, *191*, 106574.