https://doi.org/10.56093/JISAS.V80I3.4
Available online at https://isas.live/newsite
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 80(3) 2026 427–437

Hybrid Decomposition-Based Deep Learning Models for Agricultural Commodity Price Forecasting: Wavelet and EMD Approaches for Indian Cardamom Markets

Martin Joseph and Imran Khan

Sher-e-Kashmir University of Agricultural Sciences and Technology-K, Srinagar, J&K, India

Received 11 May 2026; Revised 05 September 2026; Accepted 07 September 2026

SUMMARY

Accurate short-term price forecasting for high-value agricultural commodities is critical for farmer income stability, trader planning, and policy formulation. Cardamom (*Elettaria cardamomum*), India's premier spice by export value, exhibits strongly nonlinear, nonstationary, and seasonally complex price dynamics that challenge conventional econometric models. This study develops and evaluates a comprehensive family of hybrid decomposition-based deep learning models combining Discrete Wavelet Transform (DWT) and Empirical Mode Decomposition (EMD) with Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks for weekly cardamom price forecasting across three major Kerala markets: Kattapana, Nedumkandam, and Thodupuzha. These four hybrid architectures are benchmarked against standalone LSTM, GRU, ARIMA, and SARIMA models across five evaluation metrics: RMSE, MAE, MAPE, R-squared, and Directional Accuracy (DA)- the percentage of weeks in which the model correctly predicted the direction of week-on-week price change. Results demonstrate that wavelet-based hybrids outperform all competing models across all markets. In Kattapana, the Wavelet-GRU model achieved the strongest out-of-sample forecast accuracy (RMSE = 35.82, MAPE = 2%), together with the highest goodness-of-fit (R-squared (0.98) and directional accuracy (DA = 86.49%) recorded among all sixteen models evaluated, while Wavelet-LSTM and Wavelet-GRU attained MAPE of 5–6% in Nedumkandam and Thodupuzha. EMD-based hybrids ranked second overall, with EMD-GRU achieving RMSE = 94.83 and MAPE = 4% in Kattapana, confirming the benefit of adaptive signal decomposition ahead of sequence learning; however, this advantage was market-dependent and did not hold in Nedumkandam, where EMD hybrids underperformed the standalone LSTM/GRU baselines. The study provides a systematic, multi-market evaluation of hybrid decomposition-neural architectures for spice price forecasting-adomain underserved in the existing literature and offers actionable guidance for farmers, traders, and agricultural market regulators.

Keywords: Cardamom price forecasting; Wavelet transform; Empirical mode decomposition; LSTM; GRU; hybrid models; agricultural commodity markets; time-series forecasting; deep learning; Kerala spice markets.

1. INTRODUCTION

India holds an unrivalled position in global spice trade, accounting for approximately 75% of the world's spice production by variety. Among its diverse spice portfolio, small cardamom (*Elettaria cardamomum*), revered as the "Queen of Spices," commands exceptional economic significance. Cultivated almost exclusively along the Western Ghats-an ecologically sensitive region known historically as the "Cardamom Hills" - Indian cardamom is traded through regulated electronic weekly auctions managed by the Spices Board of India under the Ministry of Commerce and Industry. In 2023–2024, India exported 6,168 tonnes of small cardamom valued at approximately ₹999.6 crore, underlining its macroeconomic importance for smallholder farmers and export earnings alike.

Despite institutional mechanisms for price discovery, cardamom markets exhibit persistent and severe price volatility. Weekly modal prices in the three principal Kerala trading centres - Kattapana, Nedumkandam, and Thodupuzha - in Idukki district have ranged from ₹120 to ₹3,900 per kilogram over the study period, with coefficients of variation exceeding 46–57%. This volatility originates from a confluence

Corresponding author: Martin Joseph
E-mail address: martinjoseph1822@gmail.com

of factors: (i) strong biannual production seasonality driven by the monsoon-dependent crop cycle; (ii) global supply shocks, particularly Guatemala's dominance as the world's largest producer since 2000 and its documented 44% production decline due to thrips infestation; (iii) festival-driven domestic demand surges; and (iv) speculative behaviour in electronic auction markets. The resulting price series are nonstationary, nonlinear, and characterised by frequent abrupt spikes properties that fundamentally limit the performance of classical linear time-series models.

Traditional econometric approaches such as Autoregressive Integrated Moving Average (ARIMA) and its seasonal extension (SARIMA) have been widely applied in agricultural price forecasting (Box and Jenkins, 1970; Paul *et al.* , 2010; Reddy, 2018) due to their interpretability and computational simplicity. However, their core assumptions of linearity and stationarity render them ill-suited for commodity markets with complex multi-scale dynamics. The emergence of deep learning architectures particularly Long Short-Term Memory (LSTM; Hochreiter and Schmidhuber, 1997) and Gated Recurrent Unit (GRU; Cho *et al.* , 2014) networks has substantially advanced time-series forecasting by enabling models to learn long-range nonlinear dependencies without explicit stationarity requirements. Comparative studies consistently demonstrate LSTM and GRU's superiority over ARIMA for agricultural and financial time series (Sima *et al.* , 2018; Sirisha *et al.* , 2022; Peter *et al.* , 2020).

Hybrid models that combine signal decomposition techniques with predictive models represent a further generation of improvement. Decomposition-based pre-processing disentangles multi-scale price components before forecasting, providing the downstream model with cleaner, more structured input. Two complementary decomposition paradigms have attracted substantial research attention. Discrete Wavelet Transform (DWT) decomposes a signal into time-frequency localised components via orthogonal filter banks, preserving both the timing and scale of price fluctuationsa critical advantage for non-stationary series. Empirical Mode Decomposition (EMD), introduced by Huang *et al.* (1998), is an adaptive, data-driven method that decomposes a signal into Intrinsic Mode Functions (IMFs) without requiring a pre-defined basis, making it particularly suited to series with complex, evolving oscillatory structures. Both paradigms have shown promise when combined with LSTM and GRU in financial and commodity price applications (Kriechbaumer *et al.* , 2014; Liang *et al.* , 2019; Qi *et al.* , 2023), yet their comparative performance for Indian spice price forecasting has not been evaluated.

A review of existing literature reveals critical gaps. First, very few studies have specifically targeted cardamom price forecasting using machine learning or hybrid methods. Second, no published study has systematically compared the full family of Wavelet–neural and EMD–neural hybrid architectures for any Indian spice commodity. Third, multi-market comparative evaluation necessary to establish robustness across heterogeneous price dynamics is largely absent from the spice forecasting literature. This study addresses all three gaps.

The specific objectives of this study are: (1) to develop and train Wavelet-LSTM, Wavelet-GRU, EMD-LSTM, and EMD-GRU hybrid models for weekly cardamom price forecasting across three major Kerala auction markets; (2) to benchmark these hybrid models against standalone LSTM, GRU, ARIMA, and SARIMA baselines using a comprehensive suite of evaluation metrics; and (3) to provide interpretable, market-specific insights that inform practical decision-making by farmers, traders, and policymakers in the Indian cardamom value chain.

2. LITERATURE REVIEW

2.1 Classical Time-Series Models in Agricultural Price Forecasting

The ARIMA framework, formalised by Box and Jenkins (1970), remains a widely employed baseline in agricultural price studies. Its ability to capture linear autocorrelation structures through autoregressive and moving average components, combined with differencing to induce stationarity, made it the method of choice for decades. Paul *et al.* (2010) applied ARIMA to model inland fish production in India. Reddy (2018) employed SARIMA for tomato price forecasting, demonstrating the value of explicit seasonal parameterisation. Despite this utility, comparative studies confirm ARIMA's fundamental limitation: its

linear structure systematically underestimates peak deviations in volatile commodity markets, producing large errors precisely at the price extremes most consequential for market participants (Mitra *et al.* , 2017; Kulkarni *et al.* , 2020).

2.2 Deep Learning Approaches

LSTM networks, introduced by Hochreiter and Schmidhuber (1997), overcome the vanishing gradient problem of standard RNNs through a gated cell-state mechanism comprising forget, input, and output gates. This architecture enables the retention of long-range temporal dependencies, making it particularly effective for price series exhibiting persistent trends and seasonality. Sima *et al.* (2018) reported that LSTM reduced forecasting error by 84–87% relative to ARIMA on economic time series. Sagheer *et al.* (2019) applied deep LSTM architectures to oilfield production forecasting with superior results. Sirisha *et al.* (2022) reported LSTM forecasting accuracy of 97.01% versus 94.38% for SARIMA in a profit-prediction benchmark. GRU, proposed by Cho *et al.* (2014), simplifies the LSTM architecture by merging cell and hidden states into a single vector and employing only two gates (update and reset), reducing parameter count while maintaining comparable learning capacity. Peter *et al.* (2020) reported GRU outperforming LSTM on cryptocurrency price data with lower MAPE, attributed to GRU's computational efficiency preventing overfitting on shorter datasets.

2.3 Wavelet-Based Hybrid Models

Wavelet Transform provides a multi-resolution decomposition framework that localises both frequency content and temporal occurrence of signal features a critical property for non-stationary series where the frequency composition changes over time. Kriechbaumer *et al.* (2014) demonstrated forecast improvements of $50–126 per tonne for metal prices using wavelet-ARIMA frameworks. Shafaei *et al.* (2016) applied wavelet-GRU to Chinese energy consumption forecasting. Liang *et al.* (2019) showed that wavelet denoising of financial time series substantially improved LSTM stock-price prediction accuracy. Branco *et al.* (2022) validated Wavelet-LSTM for electrical fault forecasting, achieving robust performance under noise. Across these domains, the consistent finding is that wavelet pre-processing enhances deep learning model performance by providing stationarised, multi-scale input components that isolate trend and seasonal information from high-frequency noise.

2.4 EMD-Based Hybrid Models

EMD, introduced by Huang *et al.* (1998), adaptively decomposes a signal into a finite set of IMFs, each representing a locally symmetric oscillatory mode with a meaningful instantaneous frequency. Unlike the fixed basis functions of wavelet or Fourier methods, EMD requires no a priori assumptions about the signal's structure. Awajan *et al.* (2019) reviewed EMD applications in time-series forecasting and identified its superiority for signals with rapidly varying amplitude and frequency. Sarah and Shabri (2014) developed EMD-ARIMA hybrids for rice price forecasting. Qi *et al.* (2023) proposed GRU-CEEMDAN-Wavelet ensembles for stock index forecasting, demonstrating that adaptive decomposition outperforms fixed-basis alternatives for financial series. A persistent limitation of EMD is mode mixing, where a single IMF may contain oscillations of different time scales, reducing decomposition interpretability and potentially degrading downstream model performance relative to wavelet approaches.

2.5 Research Gap

Despite the extensive literature on hybrid forecasting, cardamom price forecasting using decomposition-based deep learning remains largely unexplored. The few existing Indian spice price studies rely primarily on ARIMA-class models (Nikhil, 2008; Reeja, 2011) and do not employ modern deep learning architectures. No study has simultaneously compared Wavelet-neural and EMD-neural hybrid families across multiple heterogeneous spice auction markets. This study fills this gap by providing the first systematic, multi-market, multi-metric evaluation of these architectures for cardamom price forecasting.

3. MATERIALS AND METHODS

3.1 Study Area and Data

Weekly modal cardamom prices (Rs/Quintal, converted to Rs/kg for analysis) for three markets Kattapana, Nedumkandam, and Thodupuzhawere collected from AGMARKNET (Agricultural

Marketing Information Network, Directorate of Marketing and Inspection, Government of India). The three series cover different observation windows reflecting the availability of auction records at each market: Kattapana comprises 543 weekly observations (03 November 2013 – 24 March 2024); Nedumkandam comprises 603 weekly observations (16 January 2011 – 27 August 2024); and Thodupuzha comprises 722 weekly observations (02 January 2011 – 27 October 2024). Weeks in which no auction was recorded (e.g., market holidays) were treated as missing and estimated using Newton's forward and backward interpolation, consistent with the standard practice for auction-based agriculture price series. These three markets differ meaningfully in their price volatility profiles (Table 1), providing a rigorous multi-environment test of model robustness.

Table 1a. Data source and observation windows for the three cardamom auction markets

Market	Observation Period	No. of Weekly Observations
Kattapana	03 Nov 2013 – 24 Mar 2024	543
Nedumkandam	16 Jan 2011 – 27 Aug 2024	603
Thodupuzha	02 Jan 2011 – 27 Oct 2024	722

Note: Source — AGMARKNET (Agricultural Marketing Information Network), Directorate of Marketing and Inspection, Government of India.

Table 1. Descriptive statistics of weekly modal cardamom prices (Rs/Kg) for three major auction markets.

Statistic	Kattapana	Nedumkandam	Thodupuzha
Minimum (Rs/Kg)	500.00	500.00	120.00
Mean (Rs/Kg)	1203.55	1059.55	1064.74
Maximum (Rs/Kg)	3900.00	2500.00	3000.00
Range (Rs/Kg)	3400	2000	2880
Std. Deviation	698.02	487.93	576.91
CV (%)	57	46	54
Std. Error	29.95	19.87	21.47

Note: CV = Coefficient of Variation. Standard error calculated as $Std.Dev./\sqrt{n}$.

Kattapana exhibits the highest price variability (CV = 57%, range = Rs. 3,400), while Nedumkandam is the most stable (CV = 46%). The time series of all three markets showed pronounced price spikes around 2019–2020 followed by increased volatility, a pattern reflecting both domestic demand shocks and disruptions to global cardamom supply chains during this period.

3.2 Data Pre-processing and Stationarity Testing

Prior to modelling, stationarity of each price series was assessed using two complementary formal tests: the Augmented Dickey-Fuller (ADF) test, whose null hypothesis is the presence of a unit root (non-stationarity), and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test, whose null hypothesis is stationarity. Deploying both tests mitigates the risk of incorrect inference from either test alone. Test results are presented in Table 2.

Table 2. Augmented Dickey-Fuller (ADF) and KPSS unit root test results after first differencing

Market	ADF Statistic	ADF p-value	KPSS Statistic	KPSS p-value
Kattapana	-3.54	0.0069	0.0714	≥ 0.10
Nedumkandam	-4.50	0.0002	0.1515	≥ 0.10
Thodupuzha	-4.97	0.00003	0.1613	≥ 0.10

Note: ADF null hypothesis — unit root (non-stationary). KPSS null hypothesis — stationarity. A KPSS $p-value \geq 0.10$ *indicates non-rejection of stationarity at the 10% significance level.*

All three markets achieved stationarity after first differencing, as confirmed by significant ADF test statistics (all $p<0.01$) and non-significant KPSS statistics. Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots of the original and first-differenced series confirmed the appropriateness of d = 1 differencing and informed the identification of ARIMA/SARIMA orders. For deep learning models, all price series were normalised to the range [0, 1] using min-max scaling prior to training, and inverse-transformed for evaluation in original price units.

3.3 Baseline Models

ARIMA(p, d, q) models were identified following the Box-Jenkins methodology: candidate orders were selected from ACF/PACF analysis and the best specification chosen by minimising AIC and BIC. The best ARIMA models were ARIMA (2,1,2) for Kattapana and ARIMA(3,1,1) for both Nedumkandam and Thodupuzha. SARIMA (p,d,q) $(P,D,Q)_s$ models incorporated a 26-week seasonal period identified from spectral analysis and confirmed by ACF at seasonal lags, reflecting the semi-annual cardamom production and marketing cycle. Standalone LSTM and GRU models were trained with a consistent two-layer architecture: a

first layer with 256 units (LSTM/GRU, tanh activation, return sequences = True), dropout (30%), a second layer with 224 units, dropout (10%), and a dense output layer with linear activation. The Adam optimiser was employed throughout.

3.4 Wavelet Decomposition and Hybrid Architecture

Discrete Wavelet Transform (DWT) was applied to each price series using the Daubechies 4 (D4) mother wavelet at decomposition level 3. The D4 wavelet was selected for its compact support and orthogonality, properties that minimise boundary artefacts and preserve signal energy across decomposition levels, and for its established performance in financial and commodity signal analysis (Sifuzzaman *et al.* , 2009). Decomposition level 3 was chosen because, for a weekly series, it separates the signal into detail components spanning approximately 2, 4, and 8-week cycles (D1–D3) together with a longer-term approximation (A3) a scale range that aligns with the short- to medium-term auction and seasonal dynamics relevant to cardamom pricing, while retaining sufficient effective sample length within each component for stable network training; deeper decomposition levels were not pursued given the length of the available series. The decomposition yields one approximation sub-series (A3), capturing long-term trend, and three detail sub-series (D1, D2, D3), capturing progressively higher-frequency fluctuations. Wavelet pre-processing denoises the original series by reconstructing it from the approximation and selected detail components, removing high-frequency noise while preserving economically meaningful price dynamics. This single denoised, reconstructed series rather than the individual sub-series was used as input to the LSTM and GRU networks, yielding the Wavelet-LSTM and Wavelet-GRU hybrid models. Network architecture, optimiser, and training procedure were identical to the standalone baselines described in Section 3.3.

3.5 Empirical Mode Decomposition and Hybrid Architecture

Empirical Mode Decomposition (EMD) was applied to each original price series, yielding a set of Intrinsic Mode Functions (IMFs) and a residual trend. The number of IMFs extracted varied by market 4 for Kattapana, 3 for Nedumkandam, and 5 for Thodupuzha reflecting their differing signal complexities. Each IMF was independently modelled with a dedicated single-layer GRU network (224 units, tanh activation, trained with the Adam optimiser); final price forecasts were obtained by summing the individual IMF forecasts and the residual. The 224-unit single-layer configuration mirrors the second layer of the standalone baseline architecture (Section 3.3), preserving comparability of model capacity across the hybrid families; a single layer was judged adequate given that individual IMFs typically exhibit simpler, near-monotonic dynamics than the original composite price series. For the same reason of cross-model comparability, identical hyperparameters were applied to every IMF within a market rather than tuning each component separately. We recognise that IMFs can differ substantially in their statistical characteristics (e.g., dominant frequency and variance), and that component-specific hyperparameter tuning could further improve EMD hybrid performance; this is noted explicitly as a limitation in Section 5.3. This decompose-forecast-reconstruct strategy enables the network to specialise in the temporal dynamics of each frequency component.

3.6 Model Evaluation Metrics

All models were evaluated on a held-out test set comprising the final portion of each market's price series. Five complementary metrics were computed. Root Mean Squared Error (RMSE) penalises large deviations and is expressed in original price units. Mean Absolute Error (MAE) provides a robust average absolute deviation, also in original price units. Mean Absolute Percentage Error (MAPE) is a scale-free relative error that enables cross-market comparison. The coefficient of determination (R-squared) is a goodness-of-fit measure rather than an error metric; it is defined as R-squared= $1 - (SS_{res} / SS_{tot})$, where SS_{res} is the sum of squared prediction residuals and SS_{tot} is the total sum of squares of the observed values about their mean. Because R-squared benchmarks a model's error against that of a naive mean predictor, it takes negative values whenever a model performs worse than simply forecasting the sample mean as occurs for the linear ARIMA/SARIMA baselines on several of the volatile series examined here (Sections 4.3–4.4). Directional Accuracy (DA) is the percentage of time steps for which the model correctly predicted the direction (increase or decrease) of the week-on-week price change, **calculated as** $DA = (1/N)\Sigma D_t \times 100$, **where** $D_t = 1$ **if** the sign of the predicted change equals the sign of the observed change and 0 otherwise,

following the directional-accuracy formulation used in prior deep-learning forecasting studies (e.g., Peter *et al.*, 2020). DA has direct operational relevance for market participants making hold-or-sell decisions. Together, the magnitude-based metrics (RMSE, MAE, MAPE), the goodness-of-fit measure (R-squared), and the directional metric (DA) provide a comprehensive, multi-dimensional assessment of forecast quality.

4. RESULTS AND DISCUSSION

4.1 Wavelet Hybrid Model: Training Configuration and Convergence

Table 3 summarises the training configuration and convergence behaviour of Wavelet-LSTM and Wavelet-GRU models across the three markets; these are diagnostic training-process indicators rather than out-of-sample performance measures, which are reported separately in Sections 4.3–4.5. Validation losses and MAE values reported are in normalised (scaled) units. The Wavelet-GRU model for Kattapana reached the lowest validation loss among all hybrid architectures (0.000019, MAE = 0.0031), converging in only 38 epochs indicating that wavelet pre-processing provides an exceptionally clean, structured input that enables rapid and stable GRU learning. The Wavelet-LSTM for Kattapana also demonstrated near-zero validation loss (0.0000308, MAE = 0.0040), requiring somewhat more epochs (58) consistent with LSTM's larger parameter count. For Thodupuzha the most volatile market Wavelet-LSTM required 88 epochs to converge, reflecting the greater complexity of the underlying signal, yet still achieved a low validation loss (0.00015). These training diagnostics are consistent with, though not directly predictive of, the out-of-sample performance reported in subsequent sections.

Table 3. Training configuration and convergence diagnostics of Wavelet-LSTM and Wavelet-GRU models for three cardamom markets

Model	Market	Val. Loss	Epochs	MAE (scaled)
Wavelet-LSTM	Kattapana	0.0000308	58	0.0040
Wavelet-LSTM	Nedumkandam	0.0013	52	0.0276
Wavelet-LSTM	Thodupuzha	0.00015	88	0.0089
Wavelet-GRU	Kattapana	0.000019	38	0.0031
Wavelet-GRU	Nedumkandam	0.0010	40	0.0246
Wavelet-GRU	Thodupuzha	0.000041	82	0.0047

Note: Validation loss and MAE are in normalised (min-max scaled) units, not comparable across markets. D4 wavelet at level 3 used for all markets.

4.2 EMD Hybrid Model: Training Configuration and Convergence

Table 4 presents the training configuration and convergence behaviour of EMD-LSTM and EMD-GRU models; as in Section 4.1, these figures describe the training process rather than out-of-sample forecast accuracy. The number of IMFs extracted varied by market (4, 3, and 5 for Kattapana, Nedumkandam, and Thodupuzha, respectively), reflecting the differing spectral complexity of each price series. EMD-GRU for Kattapana achieved the lowest average MSE per IMF (0.0004), indicating precise IMF-level fitting. The Thodupuzha EMD-LSTM required 200 training epochs, substantially more than any other model, owing to the combination of 5 IMFs and the highly erratic price behaviour of this market. Nevertheless, the average MSE per IMF remained low (0.0014), indicating reliable convergence even under challenging conditions.

Table 4. Training configuration and convergence diagnostics of EMD-LSTM and EMD-GRU hybrid models for three cardamom markets

Model	Market	No. IMFs	Epochs	Avg. MSE/ IMF
EMD-LSTM	Kattapana	4	54	0.0006
EMD-LSTM	Nedumkandam	3	47	0.0005
EMD-LSTM	Thodupuzha	5	200	0.0014
EMD-GRU	Kattapana	4	127	0.0004
EMD-GRU	Nedumkandam	3	66	0.0006
EMD-GRU	Thodupuzha	5	40	0.0016

Note: Average MSE per IMF is in normalised units, retained here (rather than RMSE) because it is computed and averaged across IMFs of differing scale during training; all out-of-sample performance metrics in Sections 4.3–4.5 are reported as RMSE. IMF counts: Kattapana = 4, Nedumkandam = 3, Thodupuzha = 5.

4.3 Out-of-Sample Forecast Performance: Kattapana Market

Table 5 presents the out-of-sample forecast performance for all eight models in the Kattapana market. The results establish a clear performance hierarchy. ARIMA and SARIMA produced similarly poor RMSE values (415.28 and 415.56) with MAPE near 19% and negative R-squared(–0.79 and –0.80); a negative R-squared indicates that these linear models performed worse than a naive mean predictor for this volatile series (see Section 3.6 for

the interpretation of negative R-squared). Standalone GRU (RMSE = 133.14, MAPE = 5%) substantially outperformed LSTM (RMSE = 222.24, MAPE = 10%), consistent with GRU's efficiency advantage on relatively short training datasets.

EMD-based hybrids provided a further step-change improvement over the standalone networks in this market. EMD-GRU achieved RMSE = 94.83, MAPE = 4%, and R-squared (0.90), while EMD-LSTM reached RMSE = 153.93, MAPE = 6%, R-squared(0.75). The wavelet-based hybrids dominated all other models. Wavelet-GRU attained the best performance across every metric: RMSE = 35.82, MAE = 21.70, MAPE = 2%, R-squared (0.98), and DA = 86.49% the highest directional accuracy among all sixteen models tested across all markets. Wavelet-LSTM achieved RMSE = 101.68, MAPE = 4%, R-squared(0.89). The Wavelet-GRU represents a 91.4% reduction in RMSE relative to ARIMA, underscoring the impact of combining multi-resolution wavelet decomposition with recurrent deep learning.

Table 5. Out-of-sample forecast performance metrics for all models in the Kattapana market. Green shading = Wavelet hybrids (Tier 1); Blue shading = EMD hybrids (Tier 2)

Model	RMSE	MAE	MAPE (%)	R^2	DA (%)
ARIMA	415.28	281.21	19.0	–0.79	51.85
SARIMA	415.56	281.53	19.0	–0.80	53.09
LSTM	222.24	142.60	10.0	0.48	54.05
GRU	133.14	72.11	5.0	0.81	63.51
EMD-LSTM	153.93	91.72	6.0	0.75	59.46
EMD-GRU	94.83	56.70	4.0	0.90	63.51
Wavelet-LSTM	101.68	61.27	4.0	0.89	63.51
Wavelet-GRU	35.82	21.70	2.0	0.98	86.49

Note: RMSE and MAE in Rs/Kg. MAPE and DA in %. R^2 is dimensionless. Bold formatting per metric omitted for print consistency; Wavelet-GRU achieves the best value on every metric.

4.4 Out-of-Sample Forecast Performance: Nedumkandam Market

Table 6 presents Nedumkandam results. This market displayed the most erratic price dynamics, with ARIMA and SARIMA producing very poor forecasts (RMSE 628–645, MAPE 59–60%, R-squared –4.3 to –4.6; see Section 3.6 for the interpretation of such strongly negative R-squared values, which arise here because the linear models' squared errors substantially exceed the total variance of the observed series). Standalone LSTM and GRU improved substantially (RMSE 176–188, MAPE 6–8%), but directional accuracy remained very low (25.30%), indicating difficulty capturing short-term price reversals. In this market, the EMD-based hybrids did not outperform the standalone deep learning baselines: EMD-LSTM (RMSE = 191.21) was marginally worse than standalone LSTM (188.42), and EMD-GRU (RMSE = 198.11) was clearly worse than standalone GRU (176.17). This indicates that, for the highly irregular Nedumkandam signal, EMD decomposition did not provide exploitable additional structure and may instead have introduced noise into the IMF-level inputs, plausibly through mode mixing (see Section 2.4).

Wavelet-LSTM achieved the best performance in Nedumkandam: RMSE = 108.17, MAE = 58.33, MAPE = 5%, R-squared (0.80), DA = 40.96%, outperforming Wavelet-GRU (RMSE = 115.37, MAPE = 6%) one of the few markets where LSTM slightly exceeded GRU. The superiority of the wavelet hybrids over both the EMD hybrids and the standalone networks in this most volatile market is particularly notable: the structured, fixed-basis multi-resolution representation provided by the D4 wavelet appears more robust to abrupt price discontinuities than EMD's data-driven IMF decomposition, which is more susceptible to mode mixing under highly irregular price conditions.

Table 6. Out-of-sample forecast performance metrics for all models in the Nedumkandam market

Model	RMSE	MAE	MAPE (%)	R^2	DA (%)
ARIMA	628.03	587.12	59.0	–4.32	25.56
SARIMA	645.25	604.45	60.0	–4.62	25.56
LSTM	188.42	99.70	8.0	0.40	25.30
GRU	176.17	75.55	6.0	0.47	25.30
EMD-LSTM	191.21	132.51	12.0	0.38	32.53
EMD-GRU	198.11	137.24	13.0	0.34	34.94
Wavelet-LSTM	108.17	58.33	5.0	0.80	40.96
Wavelet-GRU	115.37	67.98	6.0	0.77	40.96

Note: Colour coding as in Table 5. Wavelet-LSTM is the top performer in this market; EMD hybrids underperform the standalone LSTM/GRU baselines here.

4.5 Out-of-Sample Forecast Performance: Thodupuzha Market

Table 7 presents results for Thodupuzha, characterised by intermediate volatility (CV = 54%) but the widest price range after Kattapana. ARIMA and SARIMA produced RMSE exceeding 809–812 with MAPE approximately38% and R-squared (approximately −2.29) structurally inadequate fits. Standalone GRU (RMSE = 222.51, MAPE = 8%) again outperformed LSTM (RMSE = 326.67, MAPE = 12%). EMD-LSTM and EMD-GRU both achieved RMSE approximately 159–166 and MAPE = 7%, a meaningful improvement over the standalone networks and, unlike in Nedumkandam, confirming that decomposition-assisted modelling benefits forecasting in this market.

Wavelet-LSTM and Wavelet-GRU performed comparably in Thodupuzha, achieving RMSE of 128.39 and 129.05 respectively, both with MAPE = 6% and R-squared (0.90). While the absolute improvement over the EMD hybrids is modest here (approximately19%), the consistent advantage of wavelet over EMD hybrids across all three markets establishes the robustness of the wavelet–deep learning paradigm. Notably, Wavelet-GRU achieved the highest DA in Thodupuzha (37.62%) among all eight models, an operationally important result for market participants seeking directional signals.

Table 7. Out-of-sample forecast performance metrics for all models in the Thodupuzha market

Model	RMSE	MAE	MAPE (%)	R^2	DA (%)
ARIMA	811.76	686.50	38.0	–2.29	33.33
SARIMA	809.85	685.71	38.0	–2.29	33.33
LSTM	326.67	218.47	12.0	0.40	19.80
GRU	222.51	132.79	8.0	0.72	17.82
EMD-LSTM	159.76	112.12	7.0	0.85	30.69
EMD-GRU	165.84	115.01	7.0	0.84	34.65
Wavelet-LSTM	128.39	94.80	6.0	0.90	36.63
Wavelet-GRU	129.05	93.74	6.0	0.90	37.62

Note: Colour coding as in Table 5. Wavelet-LSTM and Wavelet-GRU perform comparably in this market.

4.6 Cross-Market Performance Synthesis

Examining performance across all three markets together reveals a consistent hierarchy. Wavelet-based hybrids dominate in all markets on RMSE, MAPE, and R-squared. Wavelet-GRU is the single best model in Kattapana and Thodupuzha; Wavelet-LSTM leads in Nedumkandam. EMD-based hybrids rank second in Kattapana and Thodupuzha, where they outperform the standalone LSTM and GRU baselines, but this advantage does not generalise to Nedumkandam, where EMD hybrids trail the standalone networks (Section 4.4) indicating that the benefit of EMD decomposition in this study is market-dependent rather than universal. The MAPE values achieved by the top two tiers in the markets where decomposition helps (2–7% across Kattapana and Thodupuzha) fall well within the operational tolerance of agricultural commodity trading, where 5–10% price uncertainty is typical in real market conditions. The directional accuracy of Wavelet-GRU (86.49% in Kattapana) is particularly striking: correctly predicting the direction of weekly price change 86% of the time is a practically valuable signal for both short-position traders and farmers deciding whether to sell or hold stock at auction.

The consistent superiority of wavelet over EMD decomposition in this study can be attributed to two structural properties. First, the Daubechies 4 wavelet provides a mathematically orthogonal, fixed-basis decomposition that produces stable, reproducible sub-series components across training and test windows an advantage when the same decomposition parameters must generalise to unseen data. Second, DWT's multi-resolution property separates the price signal into frequency bands aligned with the biological and market seasonality of cardamom (26-week semi-annual cycle, annual trend), enabling the GRU and LSTM models to specialise in the dynamics of each frequency component without interference. EMD's data-driven basis, while more flexible in principle, is more susceptible to mode mixing in series with sharp amplitude changes, which are frequent in cardamom markets; this structural instability likely explains why EMD hybrids only reliably outperformed the standalone networks in the two comparatively fewer erratic markets (Kattapana and Thodupuzha) and not in Nedumkandam.

5. DISCUSSION

5.1 Implications for Agricultural Price Forecasting

This study contributes to a growing body of evidence that hybrid decomposition-neural architectures

represent a structurally superior approach to volatile agricultural commodity price forecasting compared to both classical econometric models and standalone deep learning. The MAPE values achieved by Wavelet-GRU (2% in Kattapana) and Wavelet-LSTM (5% in Nedumkandam) are substantially lower than those reported in comparable studies for other agricultural commodities: for example, Mitra *et al.* (2017) reported MAPE of 8–12% for the best ARIMA-ANN hybrid for potato prices, and Reddy (2018) obtained MAPE values of 6–9% for tomato prices using SARIMA. The present results are also consistent with Kriechbaumer *et al.* (2014), who demonstrated that Wavelet-ARIMA outperformed pure ARIMA for metal prices, and extend their finding to a deep learning context. The EMD-based hybrids provide a complementary, if less consistent, line of evidence for the same broader conclusion: EMD-GRU's MAPE of 4% in Kattapana and the EMD hybrid's MAPE of 7% in Thodupuzha still improve markedly on the standalone LSTM/GRU baselines in those two markets, reinforcing that decomposition assisted forecasting, in general, provides a structural benefit over forecasting directly from the raw price series even where, as in Nedumkandam, the specific choice of decomposition method matters a great deal to whether that benefit is realised.

The three markets display meaningfully different volatility profiles Kattapana most volatile, Nedumkandam least, Thodupuzha intermediate and the models' relative performance varies accordingly. Wavelet-GRU's exceptional DA of 86.49% in the most volatile market (Kattapana) suggests that D4 wavelet decomposition is particularly effective at isolating the recurring seasonal and trend components of a highly spiky series, enabling the GRU to learn cleaner directional signals; EMD-GRU's more modest but still respectable DA of 63.51% in the same market indicates that adaptive IMF-based decomposition captures part, but not all, of this structure. In the least volatile market (Nedumkandam), where directional accuracy is universally low across all models (25–41%) and where the EMD hybrids underperform the standalone networks (Section 4.4), the challenge lies not in noise that a decomposition can remove but in the genuine unpredictability of short-term price reversals driven by local, auction-specific factors that neither wavelet nor EMD pre-processing can resolve.

5.2 Practical Implications

The Wavelet-GRU and Wavelet-LSTM models developed in this study have direct operational applications along the cardamom value chain. Smallholder farmers, who typically sell at weekly Spices Board auctions, could use weekly price forecasts with 2–6% MAPE to make informed hold-or-sell decisions, reducing the revenue loss associated with selling during predicted price troughs. Cardamom traders and processors could optimise procurement schedules and inventory financing based on directional forecasts. Where wavelet-based forecasts are unavailablefor instance, in streaming or low-latency settings where the reconstructed series cannot be recomputed as new observations arrive the EMD-GRU and EMD-LSTM models offer a viable, if less accurate, fallback in the markets where they add value over standalone networks (Kattapana and Thodupuzha), though the Nedumkandam results caution against relying on EMD hybrids in markets with comparably erratic short-term dynamics. The Spices Board of India and state agricultural marketing departments could integrate rolling weekly forecasts from these models into market intelligence bulletins, reducing speculative uncertainty and improving price transparency. At the policy level, predictive signals from these models could inform timely interventions such as export permit adjustments or buffer stock releases to moderate extreme price spikes that disproportionately harm smallholder welfare.

5.3 Limitations

Several limitations of this study should be acknowledged. First, the hybrid models are univariate they use only historical price data and do not incorporate exogenous drivers of cardamom prices such as weekly rainfall indices, Guatemala export volumes, domestic festival calendars, or cost-of-production data. Incorporating such variables within a multivariate Wavelet-LSTM or Wavelet-GRU framework represents an important direction for future work. Second, while the Daubechies 4 wavelet and decomposition level 3 were selected based on established practice in commodity price decomposition literature and the rationale given in Section 3.4, a systematic hyperparameter search over wavelet families and decomposition levels may yield further performance improvements. Third, EMD's susceptibility to mode mixing particularly

evident in the volatile Nedumkandam market, where EMD hybrids underperformed the standalone baselines suggests that Ensemble EMD (EEMD) or Complete EEMD with Adaptive Noise (CEEMDAN), which mitigate mode mixing by ensemble averaging, should be evaluated in future studies. Fourth, as noted in Section 3.5, identical hyperparameters (architecture, unit count, optimiser) were applied to every IMF within a market for comparability across model families; component-specific hyperparameter tuning, informed by each IMF's individual statistical characteristics, may further improve EMD hybrid performance and should be explored in future work. Fifth, directional accuracy, while practically important, remains moderate in Nedumkandam and Thodupuzha (37–41%), indicating that turning-point prediction remains an open challenge even for the best models tested.

6. CONCLUSION

This study developed and evaluated a comprehensive family of hybrid decomposition-deep learning models for weekly cardamom price forecasting across three major Kerala auction markets, addressing a significant gap in the literature on Indian spice commodity price prediction. Among the eight models tested per market, wavelet-based hybrids (Wavelet-GRU and Wavelet-LSTM) consistently outperformed all alternatives, achieving MAPE values of 2–6% across markets and R-squared values of 0.77–0.98. EMD-based hybrids (EMD-GRU and EMD-LSTM) provided reliable second-tier performance in Kattapana and Thodupuzha, with MAPE of 4–7% and meaningful improvements over standalone deep learning models, but did not outperform the standalone networks in the more erratic Nedumkandam market.

The principal conclusions are: (i) Wavelet pre-processing using the Daubechies 4 filter at level 3 substantially enhances LSTM and GRU forecasting accuracy for nonstationary agricultural commodity prices, by isolating multi-scale trend and seasonal components and reducing the noise burden on the recurrent network. (ii) EMD-based decomposition provides a viable, market-dependent alternative when wavelet pre-processing is not available, with clear benefits in Kattapana and Thodupuzha but no benefit in Nedumkandam, underscoring that its performance advantage over standalone deep learning is not universal. (iii) GRU models consistently match or exceed LSTM performance with fewer parameters and shorter training times, supporting the practical case for GRU as the preferred recurrent architecture in agricultural price forecasting applications. (iv) The Wavelet-GRU model, with RMSE = 35.82 and directional accuracy = 86.49% in Kattapana, provides actionable weekly forecasts that can directly inform hold-or-sell decisions for farmers, procurement planning for traders, and market intelligence reporting by regulatory bodies.

Future research should extend these frameworks in three directions: (a) incorporation of exogenous variables weather indices, international spice trade data, demand-side indicators within a multivariate hybrid architecture; (b) application of CEEMDAN or Variational Mode Decomposition (VMD) to further address EMD mode mixing, together with component-specific hyperparameter tuning for EMD hybrids (Section 5.3); and (c) evaluation of Transformer-based architectures, including the Temporal Fusion Transformer and PatchTST, against the Wavelet-GRU benchmark established here. The models and methodology developed in this study are transferable to other high-value spice commodities pepper, vanilla, and turmeric offering a scalable contribution to evidence-based price risk management in India's spice economy.

7. DATA AND CODE AVAILABILITY STATEMENT

The weekly modal cardamom price data used in this study (Kattapana: 03 Nov 2013–24 Mar 2024, n = 543; Nedumkandam: 16 Jan 2011–27 Aug 2024, n = 603; Thodupuzha: 02 Jan 2011–27 Oct 2024, n = 722) are publicly available from AGMARKNET (https://agmarknet.gov.in), Directorate of Marketing and Inspection, Government of India. The processed dataset, trained model weights, and source code used for data pre-processing, wavelet/EMD decomposition, and model training and evaluation for one representative market are available at https://github.com/martin1822/Time-series-forecasting-of-cardamom-using-machine-learning-models; the same scripts can be readily adapted to the other two markets by modifying the data paths and market filter. Supplementary materials, including full ADF/KPSS test outputs and per-IMF training logs, are available from the corresponding author upon reasonable request.

DECLARATIONS

- **Conflict of Interest**

  The author declares that he has no conflict of interest.

- **Data Availability**

  Not Applicable.

- **Ethical Approval**

  Ethical approval was not required for this study because it did not involve human participants, animals, or collection of identifiable personal information.

- **Clinical Trial Number**

  Not Applicable. This study did not involve a clinical trial.

- **Consent To Publish**

  Not applicable. This study did not involve the recruitment or participation of human subjects.

- **Consent To Participate**

  This study did not involve the recruitment or participation of human subjects.

- **Funding**

  The author received no specific funding or financial support for this study.

  The authors confirm that the paper is not under review or in publication in full or partial in any other Journal.

REFERENCES

Awajan, A.M., Ismail, M.T. and Al Wadi, S. (2019). A review on empirical mode decomposition in forecasting time series. Italian Journal of Pure and Applied Mathematics, **42**, 301–323.

Box, G. and Jenkins, G. (1970). Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.

Branco, N.W., Cavalca, M.S.M., Stefenon, S.F. and Leithardt, V.R.Q. (2022). Wavelet LSTM for fault forecasting in electrical power grids. MDPI Sensors, **22**, 8323.

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. and Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv:1406.1078.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. Neural Computation, **9**(8), 1735–1780.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N., Tung, C.C. and Liu, H.H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London A, **454**, 903–995.

Kriechbaumer, T., Angus, A., Parsons, D.J. and Rivas Casado, M. (2014). An improved wavelet-ARIMA approach for forecasting metal prices. Resources Policy, **41**, 20–29.

Kulkarni, A., Misra, I. and Sharma, P. (2020). Time series data analysis for stock market prediction using machine learning techniques. International Journal of Advanced Trends in Computer Science and Engineering, **9**(3).

Liang, Z., Zhang, J., Chen, Y. and Zhang, C. (2019). Wavelet denoising combined with deep learning for stock price prediction. Neurocomputing, **331**, 81–89.

Mitra, S., Paul, R.K. and Das, M.K. (2017). ARIMA-GARCH-ANN hybrid for potato price forecasting. Journal of the Indian Society of Agricultural Statistics, **71**(2), 121–130.

Nikhil, P. (2008). Arecanut marketing and prices under economic liberalisation in Karnataka. Unpublished study, Karnataka State Agricultural Marketing Board.

Paul, R.K., Yadav, R., Rai, A. and Rana, S. (2010). Forecasting of inland fish production in India using ARIMA model. Indian Journal of Fisheries, **57**(4), 65–70.

Peter, M., Vogt, L. and Kempf, M. (2020). ARIMA vs. deep learning models for time series forecasting of Bitcoin. Proceedings of the 2nd International Conference on Finance and Economics.

Qi, X., Li, Q. and Zhang, Y. (2023). GRU-CEEMDAN-Wavelet for stock index forecasting. Expert Systems with Applications, **224**, 119987.

Reddy, M.A. (2018). SARIMA model for tomato price forecasting in major Indian states. International Journal of Current Microbiology and Applied Sciences, **7**(5), 1234–1242.

Reeja, P. (2011). Forecasting the price of natural rubber at Kottayam and Bangkok markets using ARIMA and ANN models. Unpublished thesis, Kerala Agricultural University.

Sagheer, A. and Kotb, M. (2019). Time series forecasting of petroleum production using deep LSTM recurrent networks. Neuro computing, **323**, 203–213.

Sarah Abadan, S. and Shabri, A. (2014). EMD-ARIMA hybrid model for rice price forecasting. Applied Mathematical Sciences, **8**(42), 2071–2080.

Shafaei, M., Adamowski, J., Fakheri-Fard, A., Dinpashoh, Y. and Adamowski, K. (2016). A wavelet-ARIMA-ANN hybrid model for precipitation forecasting. Journal of Water and Land Development, **28**, 27–36.

Sifuzzaman, M., Islam, M. and Ali, M. (2009). Application of wavelet transform and its advantages compared to Fourier transform. Journal of Physical Sciences, **13**, 121–134.

Sima, S. and Namin, A.S. (2018). Forecasting economic and financial time series: ARIMA vs. LSTM. Proceedings of the 17th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 1394–1401.

Sirisha, U.M., Belavagi, M.C. and Attigeri, G. (2022). Profit prediction using ARIMA, SARIMA and LSTM models in time series forecasting: a comparison. IEEE Access, 10.

Spices Board India. (2024). Export statistics: small cardamom 2023–2024. Spices Board of India, Kochi. Available at: www.spicesboard.gov.in.