PLANT SCIENCE TODAY
ISSN 2348-1900 (online)
Vol 11(sp4): 01–09
https://doi.org/10.14719/pst.5849

HORIZON
e-Publishing Group
HePG

RESEARCH ARTICLE

Heterogeneous autoregression inspired neural network framework for predicting groundnut price volatility in India

Krupesh S[1], Kalpana M[2*], Venkatesa Palanichamy N[2], Shivakumar K M[3], Vijayabhama M[1]

[1]Department of Physical Sciences and Information Technology, Tamil Nadu Agricultural University, Coimbatore 641 003, Tamil Nadu, India
[2]Office of the Dean, Agricultural College and Research Institute, Tamil Nadu Agricultural University, Coimbatore 641 003, Tamil Nadu, India
[3]Department of Agricultural Economics, CARDS, Tamil Nadu Agricultural University, Coimbatore 641 003, Tamil Nadu, India

*Email: kalpanam@tnau.ac.in

OPEN ACCESS

ARTICLE HISTORY
Received: 21 October 2024
Accepted: 24 November 2024
Available online
Version 1.0 : 31 December 2024
Version 2.0 : 23 January 2025

Check for updates

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

CITE THIS ARTICLE
Krupesh S, Kalpana M, Venkatesa PN, Shivakumar KM, Vijayabhama M. Neural network framework for predicting groundnut price volatility in india. Plant Science Today.2024;11(sp4):01-09. https://doi.org/10.14719/pst.5849

Abstract

Groundnut, predominantly cultivated as a rainfed crop, is highly susceptible to significant price volatility. This study aimed to investigate and enhance the performance of traditional models for forecasting the realized volatility of groundnut price returns across five Indian states (Tamil Nadu, Telangana, Karnataka, Maharashtra, and Gujarat) by evaluating traditional models and neural network-based frameworks. Using groundnut price returns data spanning fourteen years and six months (01 January 2010 to 30 June 2024), weekly realized volatility was computed. The predictive behaviour of Heterogeneous Autoregression (HAR)-based neural network frameworks was evaluated. Neural networks were assessed using time series cross-validation, and model metrics were employed to generate Model Confidence Sets (MCS). These sets were ranked based on model inclusion. The Extended Cochran–Armitage test was applied to identify and compare the best-performing models. Subsequently, model forecasts were tested and compared using the two-sided Diebold-Mariano test, and Model Confidence Sets were generated to evaluate predictive performance. For this unconventional weekly realized volatility forecast, the HAR (1,6,12) framework emerged as the most effective. Notably, the implementation of Convolutional Neural Network (CNNs) combined with RNNs, such as Conv1D-GRU and Conv1D-LSTM, demonstrated superior and consistent predictive performance across all states. Among standalone neural networks, GRU performed on par with CNN-based RNNs. These findings highlight the potential of CNN and GRU models as effective and accurate methods for forecasting agricultural price volatility.

Keywords

convolutional neural network; Conv1D-GRU; Conv1D-LSTM; groundnut; realized volatility; heterogeneous autoregression; recurrent neural network;

Introduction

Accurately forecasting fluctuations in agricultural prices is a critical issue within the global food system. Understanding price volatility is essential for anticipating price movements in agricultural commodities. Addressing price volatility is crucial for safeguarding farmers 'livelihoods and ensuring economic stability. Fluctuations in agricultural markets can significantly impact farmers' income, leaving them vulnerable to financial instability and food insecurity (1,2). Price volatility disrupts farmers' investment and production decisions, impending their ability to cover basic costs and plan for future seasons (3). Mitigating this volatility through effective forecasting and market interventions can enhance predictability and resilience, thereby fostering sustainable agricultural practices and ensuring rural economic stability (4). Furthermore, stabilizing prices is vital for food security, as prices volatility often leads to increased consumer costs, affecting the affordability of essential goods.

Recent advancements have demonstrated that a reliable and consistent benchmark econometric model for addressing this issue is the Heterogeneous Autoregressive Realized Volatility (HAR-RV) model. This study utilizes low-frequency data, specifically daily groundnut price, and explores various logical combinations of time horizons to enhance interpretability and forecasting performance. Lyócsa et al. (5), attempted to use low-frequency data for realized volatility forecasting and concluded that such data could be as effective as high-frequency data when the latter is not available or difficult to obtain. While studies on realized volatility forecasting are relatively limited, research on agricultural commodities is even scarcer. For instance, Liu employed LSTM models, while Bucci and Christensen et al. utilized Feed forward neural networks (FNNs) and LSTM in their respective studies on realized volatility (6–8). These studies proved that neural networks outperform traditional models. However, it is important to note that all existing literature predominately focuses on FNNs and LSTM. Notable works by Bucci, Sahiner et al., and Diane and Brijlal have successfully applied Artificial Neural Networks to forecast realized volatility (9–11).

Despite these advancements, there remains a limited body of research that integrates the HAR-RV framework with neural networks to predict the realized volatility of commodities. This gap is even more pronounced in studies focused on the agricultural sector.

Therefore, our primary objectives of this investigation are to:

(a) assess whether the HAR-RV framework is effective for volatility studies using low-frequency data, such as daily prices of agricultural crops, and

(b) explore alternative neural network architectures, beyond ANN and LSTM, that can be integrated with the HAR-RV framework to improve predictive accuracy.

Materials and Methods

This study examines price returns of groundnut, a crop known for its inherent volatility, in India. Groundnut, often referred to as the ‘King of Oil seeds’ in India, holds significant importance in the country’s agricultural economy. The study aims to enhance forecasting models for agricultural price volatility by leveraging Artificial Neural Networks, Convolutional Neural Networks, and Recurrent Neural Networks.

This research focuses on forecasting the realized volatility of groundnut price returns across five states of India (Gujarat, Maharashtra, Karnataka, Telangana, and Tamil Nadu). These states were selected based on the triennial cultivation area and production levels, which are predominantly concentrated in the southern region of the country.

Data collection

The dataset consists of daily modal prices aggregated from the Agricultural Marketing Information Network (AGMARKNET, https://agmarknet.gov.in/) across five key states in India — Gujarat, Maharashtra, Karnataka, Telangana, and Tamil Nadu — over a span of 14 years and 6 months (01 January 2010 to 30 June 2024). These states were specifically chosen due to their significant contributions to groundnut production and the availability of comprehensive data.

Gujarat leads in both cultivation area, accounting for 35% of the total area, and production, contributing 43% of the total groundnut yield in the country. Tamil Nadu contributes to around 7% of the total area and 10% of total production; Karnataka contributes 11% of the cultivated area and 6% of production. Maharashtra and Telangana cover 5% and 2% of the total cultivated area, respectively, and account for 4% and 3% of total production, respectively, as reported by the Directorate of Economics and Statistics, Ministry of Agricultural and Farmers Welfare, Government of India.

To address the issue of missing data entries, the median imputation technique was employed, ensuring the consistency and reliability of the dataset for analysis. For analysing price volatility, the rate of returns at time was calculated using the formula:

$$r_{i,t} = \ln\left(p_{t,i}/p_{t,i-1}\right)*100$$

Here, $p_{t,i}$ and $p_{t,i-1}$ represents the groundnut price on the i^{th} day of the t^{th} week. This approach allowed for the computation of weekly realized volatility based on the collected modal price data, providing a robust foundation for further predictive modelling and analysis.

Empirical Analysis and Volatility Measurement

The groundnut median prices in Maharashtra exhibited the highest variability, followed closely by Tamil Nadu. In contrast, the median prices in Gujarat, Karnataka, and Telangana demonstrated a significant degree of stability. A pronounced disparity in the range of median prices was observed in Maharashtra and Tamil Nadu compared to the other states, which can be attributed to substantial price volatility in these regions.

Analysing of groundnut pricing trends and returns across the five states revealed a consistent upward trend in prices over time in all states. However, price fluctuations displayed a uniform pattern across the states, except for Maharashtra. Between 2015 and 2018, groundnut prices in Maharashtra followed an upward trajectory and then stabilized, with minimal volatility, as reflected in the price returns.

In Tamil Nadu and Karnataka, groundnut price returns showed a substantial positive spike, followed by

Telangana, which exhibited more frequent but less pronounced spikes. The price returns in Maharashtra and Gujarat displayed notable peaks in both directions, with these fluctuations being more pronounced in Maharashtra.

Realized Volatility

Traditionally, volatility studies and forecasts are conducted using the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) approach, which includes Univariate GARCH models and Multivariate GARCH models such as VECH, Baba-Engel-Kraft-Kroner (BEKK), Constant Correlation models (CCC), and Dynamic Correlation (DCC) models (12). But, in this study, a more relevant measure of observable volatility, realized volatility is used to measure the volatility. The weekly realized volatility of groundnut price returns at time *t* was calculated using the formula:

$$RV_t = \sqrt{\frac{1}{N_t}\sum_{i=1}^{N_t}(r_{t,i} - \mu_t)^2}$$

where $N_t = 7$ is the length of the period *t*.

The realized volatility of groundnut price returns and their kernel density is shown in Fig. 1. The distribution patterns of realized volatility in Karnataka, Telangana, and Tamil Nadu are notably similar. Among these, Karnataka exhibited some notable high points in volatility compared to the others. Maharashtra followed these states, demonstrating relatively lower volatility. As evident from the returns, groundnut price volatility in Maharashtra remained extremely low between 2015 to 2018.

In contrast, Gujarat exhibited the least price volatility and has maintained a consistent level of stability throughout the period.

Heterogeneous Autoregression

The Heterogeneous Autoregression (HAR) model is a widely used tool for forecasting realized price volatility in financial time series data. It is particularly effective in recognizing how volatility is influenced by varying time frames. The model is built on the premise that market participants operate on diverse time scales, which collectively shape the dynamics of market volatility.

The HAR model leverages historical volatility data to predict future volatility levels. It operates under the hypothesis that the market consists of traders with varying investment horizons and distinct information sets, contributing to a complex and layered pattern of volatility over time.

The predictive power of the HAR model and its adaptations has been extensively studied. For instance, Pappas et al. developed a novel volatility forecasting model that integrates market realized variances and semi variances into the HAR framework, achieving notable improvements in forecasting accuracy (13). Similarly, Baek & Park introduced the sparse vector HAR (VHAR) model, which captures the dynamics of multinational stock volatility while improving forecasting performance (14). Clements and Preve studied explored the effects of various estimators, transformations, and forecasting schemes on the HAR model, concluding that simpler solutions often outperform standard HAR forecasts (15). Zhu et al. extended HAR-type models for the Shanghai Stock Exchange Composite by explicitly accounting for time-varying coefficients, and they investigated the HAR model's predictive accuracy using various estimators, transformations, and forecasting schemes (16). They also examined the impact of replacing high-frequency data with low-frequency data in the HAR model.

In this study, we try to employ the Heterogeneous Autoregression (HAR) model given in Eqn. (2) and Eqn. (3) as the foundational framework.

$$RV_w \sim \sum_{N(\delta)} RV_w^{(\delta)}, \delta \in (1,2,4,6,12)$$

where,

$$RV_w^{(\delta)} \sim \frac{1}{\delta}\sum_{\delta} RV_{w-\delta}, \delta \in (1,2,4,6,12)$$

This model differs from the conventional HAR (1,5,22) model typically employed with high-frequency data. Special emphasis was placed on incorporating the influence of autoregressive terms to effectively capture temporal effects and improve the precision of realized volatility forecasts. The linear partial correlation coefficients of the features revealed a consistent pattern across all the states analysed. After rigorous trials, the HAR (1,6,12) model was identified as the most suitable framework for modelling weekly realized volatility, demonstrating superior performance in capturing the underlying dynamics.

Neural Network Models

Architecture: The proposed Neural Network Architectures are presented in Fig. 2. The tested feed forward networks start with an input layer of shape structured as *batch size * time steps * no.of features*.

Network Depth and Width: The depth and breadth of the network, activation functions, optimizers, and number of epochs were selected methodically based on the outcomes of hyper-parameter tuning conducted with KerasTuner. The models were designed to be sufficiently complex, with an extensive number of nodes, to approximate any continuous function, as agreed upon by Cybenkot's Universal Approximation theorem (17).

During the tuning process, single, double, and triple layers were subjected to testing. ANN performed optimally with a single hidden layer, while LSTM and GRU required two layers each. CNN achieved effective results with a single Conv1D layer. The Conv1D-LSTM and Conv1D-GRU models included a single Conv1D layer followed by their respective RNN layer. The number of neurons per layer was tested from the set $\{2^i, i \in 1,2,\dots,8\}$.

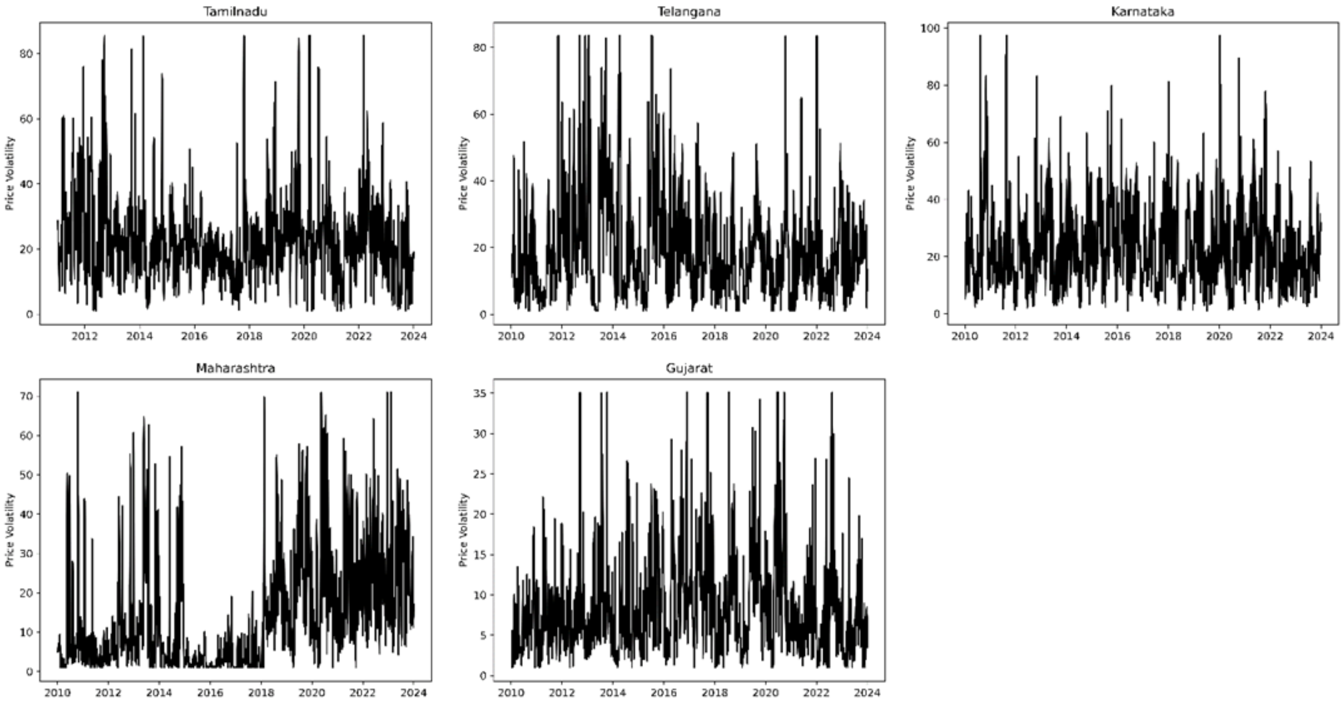

Fig. 1. Realized volatility of groundnut price returns.

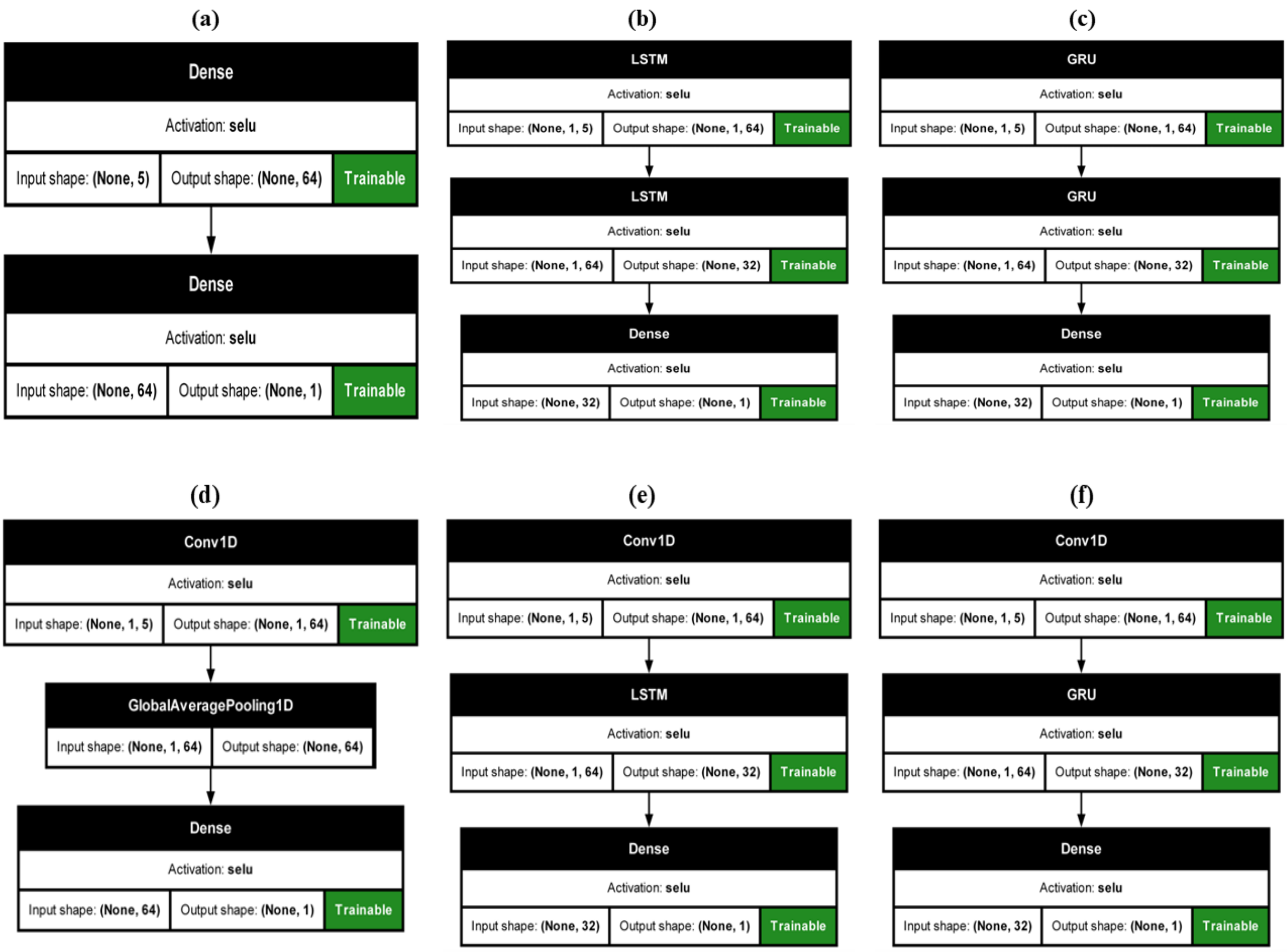

Fig. 2. Model Architectures: (a) Artificial Neural Network (ANN); (b) Long-Short Term Memory (LSTM); (c) Gated Recurrent Unit (GRU); (d) Convolutional Neural Network(Conv1D); (e) Conv1D-LSTM; (f) Conv1D-GRU.

Initialization and Activation: The Scaled Exponential Linear Unit (SELU) activation function outperformed other non-linear activation functions, yielding superior predictive results. The LeCun-Normal kernel initializer was found to complement SELU effectively, enhancing predictive accuracy while mitigating vanishing and exploding gradient issues. This configuration was consistently applied across the networks.

For LSTM and GRU layers, incorporating sigmoid and tanh functions as recurrent activation functions significantly improved prediction accuracy. Initially, the ReLU activation functions in conjunction with HeNormal initialization was employed to restrict predictions to positive values. Batch Normalization layers were subsequently added to enhance the models' training efficiency and generalization capacity.

Despite the initial success of these methods, SELU emerged as the superior choice for the dataset across all states, even though it occasionally produced predictions in the negative range. Unlike functions that feature zero gradients for all inputs over specific ranges, SELU ensures active neuron participation throughout training, allowing continuous learning.

Optimization: Three optimizers, Adaptive Moment Estimation (Adam), Stochastic Gradient Descent (SGD), and Root Mean Square Propagation (RMSprop) were tested. Among these SGD was the one that performed the best and was used to optimize the networks' weights, despite its slow convergence rate. Various learning rate schedules were tested before finalizing the optimizer, with SGD proving efficient in minimizing loss.

Model Training: Based on the nature of the time series data from different states, the data was partitioned into training, validation, and testing sets. The training set was used to train the model, while the validation set was used to monitor the training process to prevent overfitting issues. Early stopping callback were used as a precautionary measure against overfitting.

The model was trained by optimizing Mean Squared Error (MSE) loss function, since it seemed to have a better significance in moderating the model. In order to assess the accuracy of the predictions the following metrics were used: Mean Absolute Error, Mean Squared Error, Root Mean Squared Error and Pearson R^2 .

Model Performance Evaluation

Cross Validation: In order to verify the predictive performance of the models, *k*-fold cross validation was performed. Each time series dataset was split into three training sets and corresponding testing sets, with a two-period gap between them to prevent information leakage between the sets. This was repeated five times to avoid any bias in the analysis.

For each of the six models across the five-time series, Mean Absolute Error and R-squared values were computed for the 5x3 validation sets. The MAE and R^2 values were averaged across the splits in each of the five runs. Model confidence sets were then generated, ranking the models based on their order of inclusion in the set.

To assess the association between the observed ranks and the models, the Extended Cochran–Armitage test was performed. This test builds on the traditional Cochran–Armitage trend test, which evaluates the relationship between an ordinal predictor and a binary outcome. The extended version accommodates a categorical dependent variable with more than two levels, making it suitable for this analysis.

Subsequently, pairwise ordinal independence tests were conducted to differentiate high-performing models from lower-performing ones based on their ranks. This step provided a clearer distinction in model performance, identifying those with superior predictive capabilities.

Model Testing: After cross validation, the models were tested to confirm their predictive accuracy as indicated by the validation results. The test set was employed for this evaluation. To compare the predictive performance of different neural networks, a two-sided Diebold-Mariano (DM) test was applied.

The DM test statistically compares the forecasting accuracy of two competing models. It calculates the difference in loss values (e.g., prediction errors) for each time period and tests whether the mean of these differences is statistically different from zero. Under the null hypothesis, the mean difference in forecast errors equals zero, indicating that neither model outperforms the other significantly.

This approach ensures robust evaluation of model performance, highlighting the models with superior predictive accuracy in both validation and testing phases.

Results and Discussion

The HAR-RV framework for predicting realized price volatility is usually used where high frequency data are available, i.e., intra-day data points collected over months or years, is available. In such cases, daily, weekly, and monthly realized volatilities are commonly used as features. However, in this study, we used daily price returns of groundnut, with only one data point per day. As a result, weekly realized volatility was utilized as the measure of volatility, and feasible features for the prediction process were tested accordingly. To identify suitable features, linear partial correlation coefficients were analyzed for weekly realized volatility across different time periods $t \in 1,2,3,\ldots,24$. It was observed that no features corresponding to time periods greater than t=12, exhibited significant correlation with weekly realized volatility. Therefore, the weekly realized volatility for time periods t = 1,2,4,6,12, which demonstrated significant correlations, were selected as features for the neural network models.

Following the initial configuration, the models were trained, and cross validated. From Table 1, it is evident that the Conv1D-LSTM model consistently ranked first or second, followed by the GRU and Conv1D-GRU models. The Linear Regression model consistently ranked last throughout the cross-validation process, while the LSTM model exhibited the poorest performance in most instances.

Upon generating the Model Confidence Set, an Extended Cochran–Armitage test was performed on the tabulated ranks to determine whether there was an association between the models and their cross-validation rankings. The hypotheses were formulated as follows:

***H*₀:**There is no association between the models and their rank.

***H*₁:**There is an association between the models and their ranks.

Test statistic: χ^2- 289.01; df=5;α=0.05; p-value<2.2*e*-16

This proves that there is strong evidence to reject the null hypothesis(H_0).

A pairwise comparison for ordinal independence was subsequently performed. The results indicated no significant performance differences between theConv1D-LSTM and Conv1D-GRU (p - value = 0.051). Similarly, the comparison showed no significant difference between the performance of GRU and Conv1D-LSTM (p - value = 0.331) and the performance of GRU and Conv1D-GRU (p - value = 0.194). All other pairwise comparisons showed strong evidence to reject the null hypothesis, confirming significant differences in model performance.

From Table 1 and the test results, the following can be observed:

(1) Linear Regression model was outperformed by all neural networks across all instances.

(2) Conv1D – LSTM and Conv1D – GRU models consistently demonstrated the best performances across all states.

(3) Among standalone networks like ANN, LSTM and GRU, GRU recorded consistent performance than other models and was comparable, in some cases, to the CNN-based LSTM and GRU models.

These findings suggest that a simple recurrent neural network such as GRU can be a reliable choice for forecasting the realized volatility of groundnut prices. Moreover, integrating CNN with RNN models significantly enhanced the predictive performance, particularly for LSTM, due to the superior feature extraction capabilities of the Conv1D layer. Although Conv1D-GRU exhibited incremental improvements, further research is required to fully understand its potential.

Interestingly, all models performed well in Gujarat, the state with the lowest price volatility, highlighting that neural networks can be effective under low-volatility conditions. Despite the lower loss values in Gujarat, there appeared to be a trade-off between goodness-of-fit and loss. In contrast, R^2 values in Maharashtra, a state with high volatility, were notably higher than those in Gujarat. Models performed better in states with high price volatility (Tamil Nadu, Maharashtra, and Karnataka) than in states with low volatility.

The models were then tested on unseen data, with the results tabulated in Table 2. The forecasts of the models were then compared using the two-sided Diebold Mariano test, which calculates a t-statistic for each comparison. A heat map illustrates the significance of the test statistic (Fig. 3). The null hypothesis (H_0) posits that the forecasts of the models being compared do not differ. A lower p - value rejects the null hypothesis. The testing results aligned with cross-validation outcomes, with more pronounced reflections of model behaviour. All neural networks outperformed the linear HAR-RV model in terms of lower errors. However, in Gujarat, the differences between HAR-RV and other neural networks were not statistically significant, likely due to the state's consistently low-price volatility. GRU produced promising forecasts for realized volatility, while Conv1D-GRU and Conv1D-LSTM performed comparably across all states. Notably, LSTM failed to outperform ANN or CNN, consistent with cross-validation findings. Standalone ANN and CNN models struggled to generalize time-dependent patterns as effectively as RNNs, but CNN integration with RNNs demonstrated significant improvements, warranting further investigation.

This study exclusively examined groundnut price volatility across five major producing states using realized volatility over various time periods, without incorporating potentially influential exogenous factors such as weather, policy changes, or global market trends. While the findings offer an initial understanding of volatility structures and may generalize to other regions, future research could benefit from incorporating a broader range of exogenous variables and extending the analysis to other crops.

Conclusion

Modelling groundnut price returns volatility using the HAR-RV framework based on daily price data is a substantial advancement in the field of agriculture. Our study demonstrates that the HAR (1,6,12) framework is a reliable approach for modelling and predicting realized volatility in groundnut prices. This framework implies that weekly realized volatility is influenced by the volatility of the preceding week, the previous month, and the past three months.

Table 1. Ranks obtained through Model Confidence Set generation during the cross validation process

Modals \ Ranks	1	2	3	4	5	6	7
LR[a]	0	0	0	0	0	0	75
ANN	0	12	0	15	36	12	0
LSTM	0	0	0	1	12	62	0
GRU	25	14	31	5	0	0	0
Conv1D	0	2	14	42	16	1	0
Conv1D-LSTM	28	32	6	2	7	0	0
Conv1D-GRU	22	15	24	10	4	0	0

[a]Linear Regression model was omitted for further analysis as the model was outperformed by all other models invariably.

Table 2. Out - of - bag sample validation of neural network models for different states

State	Model	MAE	MSE	RMSE	R^2
Tamil Nadu	**LR**	7.192	50.332	7.095	0.416
	ANN	4.802	28.358	5.325	0.489
	LSTM	5.785	36.164	6.014	0.400
	GRU	2.898	15.284	3.909	0.850
	Conv1D	967	27.601	5.254	0.501
	Conv1D-LSTM	2.708	14.158	3.763	0.861
	Conv1D-GRU	**2.599**	**13.944**	**3.734**	**0.863**
Telangana	**LR**	5.875	46.689	6.833	0.423
	ANN	4.770	33.929	5.825	0.448
	LSTM	7.020	52.576	7.251	0.317
	GRU	**2.693**	18.880	4.345	0.781
	Conv1D	3.835	26.919	5.188	0.568
	Conv1D-LSTM	2.730	**18.618**	**4.315**	**0.784**
	Conv1D-GRU	2.759	19.372	4.401	0.775
Karnataka	**LR**	6.911	53.274	7.299	0.401
	ANN	4.589	25.556	5.055	0.486
	LSTM	6.654	42.120	6.490	0.334
	GRU	2.462	13.387	3.659	0.871
	Conv1D	4.401	24.588	4.959	0.475
	Conv1D-LSTM	**2.338**	12.789	3.576	**0.877**
	Conv1D-GRU	2.440	**12.716**	**3.566**	**0.877**
Maharashtra	**LR**	7.420	70.889	8.420	0.375
	ANN	4.347	41.432	6.437	0.631
	LSTM	5.680	53.290	7.300	0.521
	GRU	**3.234**	**28.817**	**5.368**	**0.836**
	Conv1D	5.291	48.630	6.974	0.515
	Conv1D-LSTM	3.318	29.887	5.467	0.830
	Conv1D-GRU	3.292	29.608	5.441	0.832
Gujarat	**LR**	1.324	4.915	2.217	0.781
	ANN	1.174	4.012	2.003	0.821
	LSTM	1.348	5.034	2.244	0.776
	GRU	**1.188**	**3.644**	**1.909**	**0.838**
	Conv1D	1.241	4.026	2.006	0.821
	Conv1D-LSTM	1.237	4.293	2.072	0.809
	Conv1D-GRU	1.190	4.104	2.026	0.817

To enhance these models, the incorporation of neural networks proved to be highly effective. Among the trained and tested models—ANN, LSTM, GRU, CNN, CNN-LSTM, and CNN-GRU—the integration of LSTM and GRU with CNN emerged as particularly powerful for predicting weekly realized volatility, showing promising potential for further exploration. The simple GRU model also displayed competitive performance, largely due to its straightforward architecture compared to LSTM.

Neural network models demonstrated superior fit in high-volatility price scenarios. Conversely, in low-volatility conditions, the trade-off between goodness-of-fit and model loss was minimal.

This study focused on analysing the price volatility of groundnut crops from five major producing states. While these findings provide insights that may generalize to other regions, the study exclusively relied on realized volatility and did not account for potentially influential exogenous factors such as weather, policy changes, or global market trends.

Future research could expand on these results by applying the models to additional agricultural commodities across diverse locations and incorporating significant exogenous

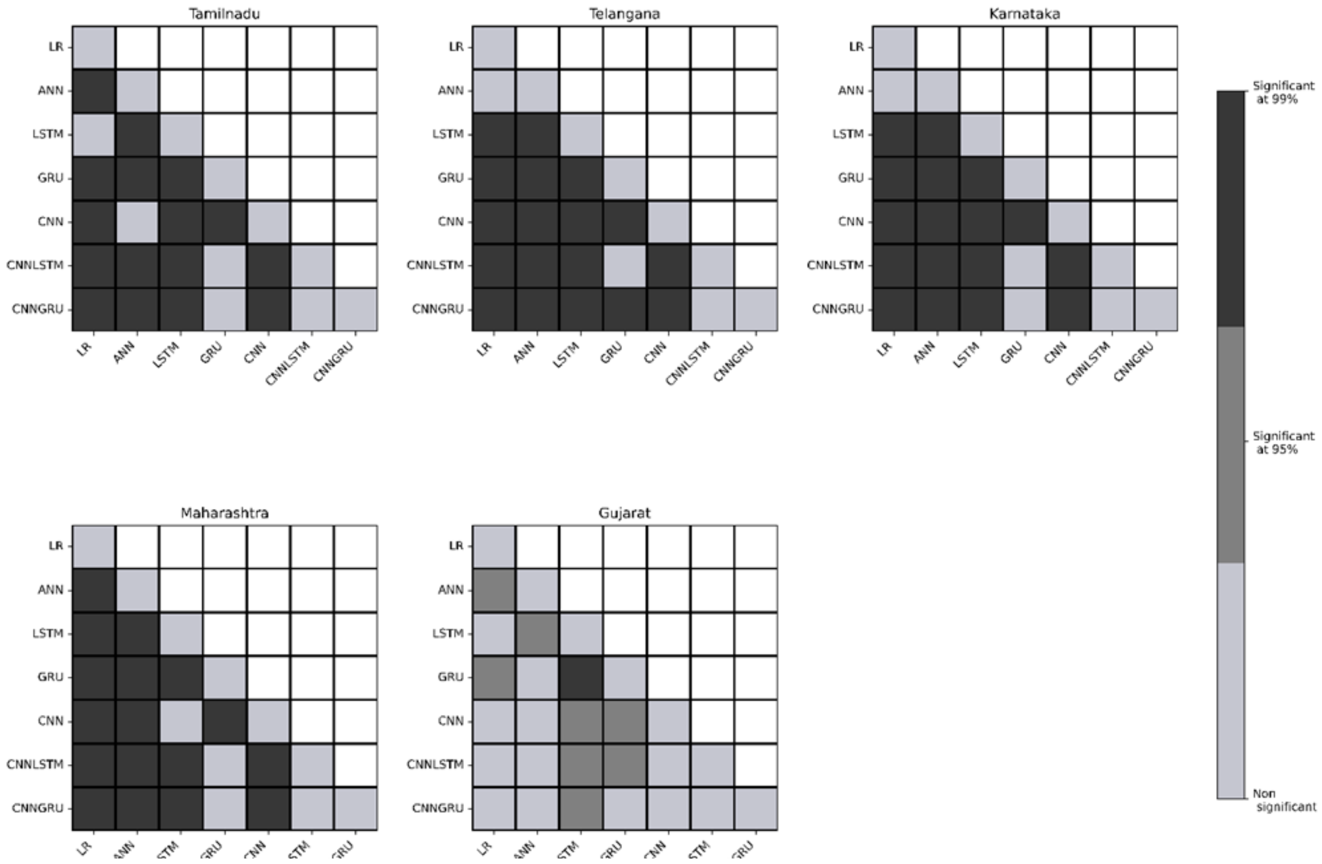

Fig. 3. Significance of the statistic obtained from multiple comparison of forecasts by Diebold Mariano Test.

variables. This would enable the development of more robust and effective non-linear modelling tools for agricultural price volatility.

Limitations

This study exclusively focused on groundnut crop prices from five major producing states. The volatility structures in lesser-producing states were not analyzed. Additionally, the proposed models were tested solely using realized volatility across different time periods, without incorporating any exogenous variables.

Data Availability

The dataset used in this article is available upon request to the corresponding author. Researchers and other users interested in accessing the data can contact the corresponding author directly for further details.

Authors' Contributions

SK and MK conceived and presented the idea; SK performed the analysis and prepared the manuscript. KMS verified the analytical methods; NVP and MV provided critical comments and approved the manuscript. MK supervised the project.

Compliance with ethical standards

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

Ethical issues: None

References

1. Bellemare MF. Rising food prices, food price volatility and social unrest. Amer J Agri Eco. 2015 Jan;97(1):1-21. https://doi.org/10.1093/ajae/aau038
2. Binswanger-Mkhize HP. Is there too much hype about index-based agricultural insurance?. Journal of Development Studies. 2012 Feb 1;48(2):187-200. https://doi.org/10.1080/00220388.2011.625411
3. Myers RJ, Jayne TS. Multiple‐regime spatial price transmission with an application to maize markets in southern Africa. Amer J Agri Eco. 2012 Jan;94(1):174-88. https://doi.org/10.1093/ajae/aar123
4. Minot N. Food price volatility in sub-Saharan Africa: Has it really increased?. Food Policy. 2014 Apr 1;45:45-56. https://doi.org/10.1016/j.foodpol.2013.12.008
5. Lyócsa Š, Plíhal T, Výrost T. FX market volatility modelling: Can we use low-frequency data?. Finance Research Letters. 2021 May 1;40:101776. https://doi.org/10.1016/j.frl.2020.101776
6. Liu Y. Novel volatility forecasting using deep learning–long short term memory recurrent neural networks. Expert Systems with Applications. 2019 Oct 15;132:99-109. https://doi.org/10.1016/j.eswa.2019.04.038
7. Bucci A. Realized volatility forecasting with neural networks. Journal of Financial Econometrics. 2020;18(3):502-31. https://doi.org/10.1093/jjfinec/nbaa008

8. Christensen K, Siggaard M, Veliyev B. A machine learning approach to volatility forecasting. Journal of Financial Econometrics. 2023 Dec 15;21(5):1680-727. https://doi.org/10.1093/jjfinec/nbac020
9. Bucci A. Cholesky–ANN models for predicting multivariate realized volatility. Journal of Forecasting. 2020 Sep;39(6):865-76. https://doi.org/10.1002/for.2664
10. Sahiner M, McMillan DG, Kambouroudis D. Do artificial neural networks provide improved volatility forecasts: Evidence from Asian markets. Journal of Economics and Finance. 2023 Sep;47(3):723-62. https://doi.org/10.1007/s12197-023-09629-8
11. Diane L, Brijlal P. Forecasting stock market realized volatility using random forest and artificial neural network in South Africa. Int J Eco and Fin Issues. 2024 Mar 18;14(2):5-14. https://doi.org/10.32479/ijefi.15431
12. Bucci A. Forecasting realized volatility: A review. Journal of Advanced Studies in Finance (JASF). 2017;8(16):94-138. Available from: https://mpra.ub.uni-muenchen.de/id/eprint/83232
13. Hizmeri R, Izzeldin M, Nolte I, Pappas V. A generalized heterogeneous autoregressive model using market information. Quantitative Finance. 2022 Aug 3;22(8):1513-34. https://doi.org/10.1080/14697688.2022.2076606
14. Baek C, Park M. Sparse vector heterogeneous autoregressive modeling for realized volatility. J Korean Stat Soc. 2021 Jun;50(2):495-510. https://doi.org/10.1007/s42952-020-00090-5
15. Clements A, Preve DP. A practical guide to harnessing the har volatility model. Journal of Banking and Finance. 2021 Dec 1;133:106285. https://doi.org/10.1016/j.jbankfin.2021.106285
16. Zhu Q, You M, Wu S. Forecasting volatility with time-varying coefficient regressions. Discrete Dynamics in Nature and Society. 2020;2020(1):3151473. https://doi.org/10.1155/2020/3151473
17. Cybenko G. Approximation by superpositions of a sigmoidal function. Math Control Signals Syst. 1989 Dec;2(4):303-14. https://doi.org/10.1007/BF02551274