



A Random Forest-Convolutional Neural Network Deep Learning Model for Predicting the Wholesale Price Index of Potato in India

Soumik Ray¹ · Tufleuddin Biswas¹ · Walid Emam² · Shikha Yadav³ · Priyanka Lal⁴ · Pradeep Mishra⁵



Received: 2 April 2024 / Accepted: 27 April 2024

© The Author(s), under exclusive licence to European Association for Potato Research 2024

Abstract

The wholesale price index (WPI) is a crucial economic indicator that provides insights into the pricing dynamics of different goods within a country, especially potato commodities. In this study, we tried to build a hybrid machine learning model technique for predicting the volatile price index of potato. We introduced the Random Forest-Convolutional Neural Network (RF-CNN) model to predict agricultural volatility price index commodities. Traditional statistical time series models (Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroscedasticity (GARCH)) were also investigated for comparison with machine learning models (Support Vector Machine (SVM), Long Short-Term Memory (LSTM), and Convolutional Neural Network (CNN)). Because the data set was volatile, the GARCH model outperformed the ARIMA model which had a lower goodness of fit value. The performance of the SVM model was comparable to that of the statistical models. However, after utilizing an input lag selection strategy based on autocorrelation function (ACF) and RF, the machine learning models outperformed the statistical models. We used LSTM and CNN models with the appropriate input lag feature assessed by ACF and RF. Our findings indicate that the RF-CNN model beats the other models in terms of error accuracy, with improvements of 67% for root mean square error, 95% for mean absolute percentage error, 63% for mean absolute error and mean absolute squared error on the training set, and more than 90% on the testing set for all goodness of fit. Based on the error accuracy, the RF-CNN model can be utilized to better predict the potato price index in the long term. We hope our study will benefit stakeholders and policymakers by providing a realistic potato price forecast. Furthermore, our study contributes to the growing corpus of research on machine learning models for time series.

Keywords Convolutional Neural Network · Deep learning · Long Short-Term Memory · Potato wholesale price index · Time series analysis

Extended author information available on the last page of the article

Introduction

The wholesale price index (WPI) serves as a critical economic indicator, providing insights into the pricing dynamics of various commodities within an economy (Dhakre and Bhattacharya 2016). Amongst these commodities, potatoes are a staple food crop with significant economic implications, particularly in the Indian agricultural landscape (Lin et al. 2023). Gebrechristos and Chen (2018) reported a substantial holding in potatoes' agricultural and dietary value, which earned the moniker 'The King of Vegetables' within the Solanaceae family. Understanding the behaviour of the wholesale price index of potatoes is crucial for policymakers, agricultural stakeholders, and consumers alike. The potato market is projected to witness substantial growth, with an estimated market size of USD 115.74 billion in 2024, expected to reach USD 137.46 billion by 2029, reflecting a Compound Annual Growth Rate (CAGR) of 3.5% over the forecast period (2024–2029) (source: <https://www.mordorintelligence.com/industry-reports/potato-market/market-size>). Its resilience to pests, diseases, and diverse climatic conditions has positioned the potato as a leading crop in developing nations, offering enhanced food security compared to other food crops (Zaheer and Akhtar 2016). Şahinli (2020) demonstrated that the Autoregressive Integrated Moving Average (ARIMA) method achieves good consumer potato price forecasting accuracy according to the mean absolute percentage error (MAPE), the root mean square error (RMSE), and the mean absolute deviation (MAD). Sahu et al. (2024) employed ARIMA to forecast potatoes' future area, production, and productivity in India and the major growing states. Additionally, they utilized Markov Chain analysis to analyse the export behaviour of potatoes. Mishra et al. (2024a) worked on potato production data by ARIMA and ETS (Error-Trend-Seasonality) in predicting potato production to assess the efficacy of these models, and the results showed that the ETS model consistently performed better in predicting potato production for the testing data set than the ARIMA model when considering the chosen countries. Mishra et al. (2023) and Yadav et al. (2024) studied potato forecasting using machine learning and the traditional ARIMA approach. Shankar et al. (2024) reported time series models including ARIMA, Artificial Neural Networks (ANN), ARIMA-Generalized Autoregressive Conditional Heteroscedasticity (GARCH), ARIMA-ANN, Ensemble Empirical Mode Decomposition (EEMD)-ARIMA, and EEMD-ANN. They got varying best fit models as per the varying locations; that is the EEMD-ANN model was the best fit for Chandigarh, Delhi, and Shimla markets, whilst the EEMD-ARIMA model was the best fit for Dehradun. In their 2022 study, Adudotla et al. forecasted potato arrivals and prices in Agra, the largest potato-producing district in Uttar Pradesh, employing ARIMA, Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM) models. Upon comparing the Root Mean Square Error (RMSE) values, ANN exhibited the lowest RMSE value, indicating its superior performance for this dataset. Özden (2023) compared the performance of Convolutional Neural Network (CNN), LSTM, and random forest models on multivariate, multi-step prediction tasks using daily price and trade volume data of onion, potato, and

garlic in Turkey. Whilst all models yielded promising results, the CNN model emerged as the most effective in prediction accuracy. Ray et al. (2023) investigated an improved hybrid ARIMA-Long Short-Term Memory (LSTM) model based on the random forest lag selection criterion. They showed that the hybrid model applied to volatile monthly price indices of pulses.

The concept of hybrid models, combining different algorithms or techniques from statistical and machine learning domains, is becoming increasingly popular in machine learning. This approach draws upon the long-standing practice in time series literature, where researchers have utilized hybridization to leverage the strengths of various models for more efficient forecasting (Gulay et al. 2024). Hybrid models merge two or more distinct network architectures to harness the strengths of each type and offer researchers the opportunity to capitalize on the unique capabilities of each constituent model, thereby enhancing the accuracy and effectiveness of forecasting tasks (Abbasimehr et al. 2024). This approach allows for a synergistic integration of statistical methods with machine learning techniques, enabling more robust and adaptable solutions for analysing time series data (Chen et al. 2023; Kumar and Yadav 2023). By integrating various network structures or blending deep learning with statistical models, hybrid approaches offer several advantages, including enhanced prediction accuracy, improved interpretability, and more efficient management of data constraints (Salman et al. 2024). Wang et al. (2018) used ARIMA and the hybrid ARIMA-ANN models to estimate Ecological Footprint and ecological capacity in China. It demonstrated that the hybrid ARIMA-ANN performed better than other models. Júnior et al. (2019) devised a hybrid approach to determine the optimal combination for integrating estimates from linear and nonlinear models, and the hybrid system outperformed single and existing hybrid models in the literature, demonstrating superior performance.

This research comprehensively explores the complex dynamics inherent in the wholesale price index of potatoes in India. Employing a multifaceted approach, we integrate statistical analysis and cutting-edge machine-learning techniques. This includes traditional methods such as ARIMA and GARCH, alongside advanced deep learning architectures such as Support Vector Machine (SVM), CNN, and LSTM networks. By combining these diverse methodologies, our study aims to provide a holistic understanding of the factors influencing potato WPI fluctuations. We seek to uncover hidden patterns and relationships within the data through rigorous analysis and modelling. This enables more accurate forecasting and informed decision-making for stakeholders to anticipate price movements, mitigate risks, and optimize decision-making processes across the agricultural supply chain. Through the integration of statistical analysis and machine learning methodologies, coupled with a focus on forecasting, this research provides actionable insights that can drive sustainable agricultural practices, facilitate market stability, and ensure food security in India. The complete framework of our research is represented in Fig. 1. The rest of the manuscript has been separated into “Materials and Methodology” section, “Results and Discussion” section, followed by the “Conclusion” section.

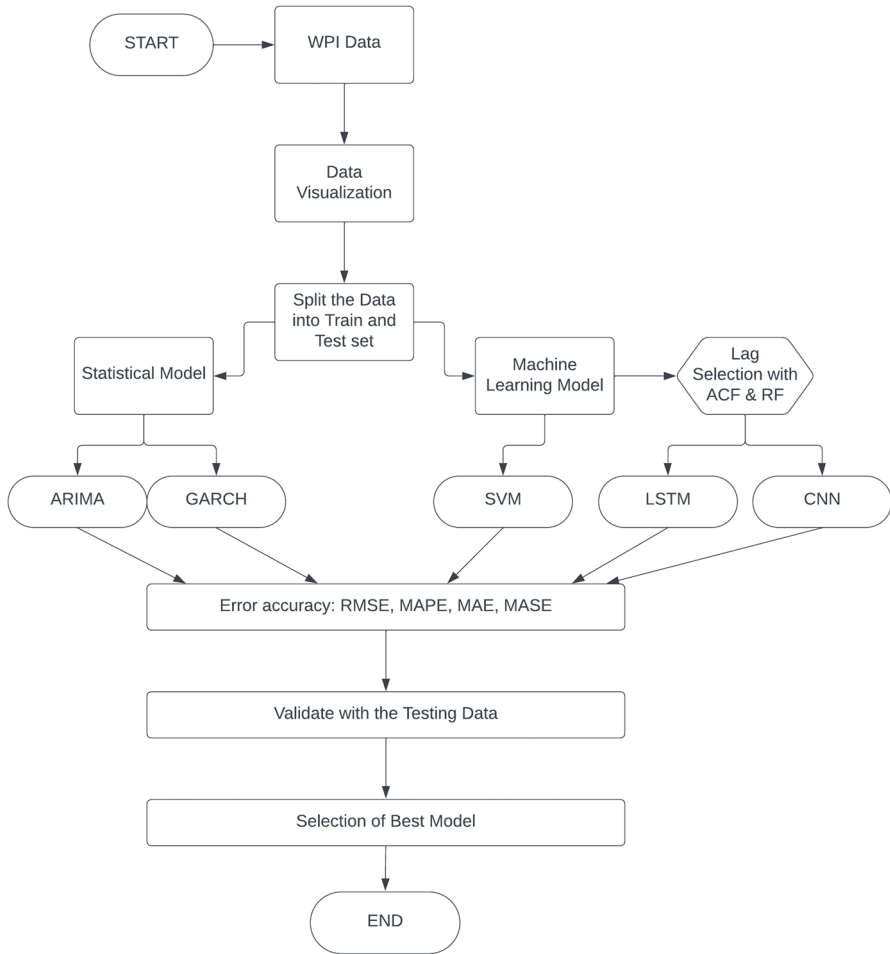


Fig. 1 Framework of the research methodology

Materials and Methods

This section provides a brief description of the methodology used to forecast the Whole Price Index of potato. In this study, the Whole Price Index (WPI) data was collected from the Economic Adviser, Ministry of Commerce, Government of India, from Jan 2005 to Jan 2024. The study uses different forecasting models like Autoregressive Integrated Moving Average (ARIMA), Generalized Autoregressive Conditional Heteroscedasticity (GARCH), Support Vector Machine (SVM), Long Short-Term Memory (LSTM), and Convolutional Neural Network (CNN) (Mishra et al. 2021; Mishra et al. 2024b). Each model was applied to the same training dataset to assess their effectiveness in predicting future values. For forecasting the data, different steps were followed as discussed in this section (Mishra et al. 2024a). The autocorrelation function (ACF)

and random forest (RF) technique are utilized to estimate the input lag selection for the deep learning model (Ray et al. 2023).

Data Collection and Processing

The dataset used in this study comprises the wholesale price index of potato from Jan 2005 to Jan 2024, on which different models were used. Data processing was conducted on the data collected for any missing value interpolation and extrapolation. This section briefly describes the forecasting models used in the study. The reason behind employing different methods is that combining multiple methods may improve prediction performance rather than any individual method (Bento et al. 2021). This study primarily focuses on utilizing the data set to forecast using statistical and machine learning tools, thereby merging the predictions of these methods (Jagait et al. 2021). The data collected after suitable treatment was considered fit for forecasting models. Following the data processing, data analysis was conducted first by splitting data into training and testing data sets. The training set was used for training the model, whereas the test data set was used for testing the model (Mishra et al. 2024b). This process helps us to train the models on the data set and test accuracy on the testing set. There were 229 data points, out of which 217 data points were used as a training set and 12 data points were used as a testing set.

This study aims to provide insights into the comparative performance of ARIMA, GARCH, SVM, LSTM, and CNN models with a hybrid approach of lag selections for time series forecasting, ultimately aiding researchers in selecting the most appropriate model for their specific forecasting needs. A brief description of each modeling technique has been discussed below.

Autoregressive Integrated Moving Average

ARIMA is a time series forecasting model that mostly captures linear dependencies between observations by modelling the relationship between an observation and a number of lagged observations. The ARIMA model is specified by its order parameters (p, d, q) , where p represents the autoregressive order, d represents the degree of differencing, and q represents the moving average order (Rahman et al. 2022). It is a widely used model for prediction that can handle nonstationary data series. The model can be formulated as

$$Y_t = c + \phi_1 y_{dt-1} + \phi_p y_{dt-p} + \dots + \theta_1 e_{t-1} + \theta_q e_{t-q} + e_t \quad (1)$$

where e is an error term and c is a constant.

Generalized Autoregressive Conditional Heteroscedasticity

GARCH is a model commonly used to capture volatility clustering and time-varying volatility in time series. The GARCH model extends the traditional ARIMA model by incorporating conditional heteroscedasticity. GARCH models are specified by

their order parameters (p , q) for the autoregressive and moving average components of the conditional variance equation. The fact that the conventional time series models had constant standard deviation (Black and Scholes 1973), conditional heteroscedasticity is known for accuracy in the results. Generalizations of ARCH have been developed to overcome the drawbacks and weaknesses of earlier models.

Support Vector Machine

Support Vector Machine (SVM) is a type of supervised learning technique that is commonly employed for both classification and regression applications. SVM can be employed in time series forecasting to address univariate and multivariate forecasting. Support Vector Machines (SVM) utilize a hyperplane in a feature space with several dimensions to separate the target values effectively. In time series forecasting, the SVM parameters, including the kernel type, regularization parameter, and kernel parameters, are optimized using grid search or cross-validation approaches. Support Vector Machine (SVM) is a machine learning technique used for tasks including classification and regression, and it requires labelled data for training. SVM can be employed in time series forecasting to address univariate and multivariate forecasting functions by considering the forecasting problem as a regression problem. Support Vector Machine (SVM) is a machine learning algorithm that identifies a hyperplane in a feature space with several dimensions. This hyperplane is chosen to segregate the target values effectively.

Long Short-Term Memory

LSTM is a specific form of recurrent neural network (RNN) designed to model sequential input and capture long-term dependencies. LSTM networks are suitable for time series forecasting because they have the capacity to retain information across long sequences. The LSTM model is trained using the training dataset, and hyperparameters such as the number of LSTM layers, the number of units in each layer, and the learning rate are optimized using grid search or random search approach.

Convolutional Neural Network

CNN is a deep learning method primarily used for image recognition but has also shown effectiveness in sequence modelling tasks. In time series forecasting, CNN can be applied directly to the raw input sequence or preprocessed data extracted from the sequence. For time series forecasting, CNN typically involves convolutional layers followed by pooling and fully connected layers. Hyperparameters such as kernel size, number of filters, and pooling size are optimized during the training process. This paper uses the features of the time series of WPI of potato to take advantage of the CNN network.

Proposed Technique

The most challenging approach is to select an appropriate input feature lag for any machine or deep learning model. So, our proposed methods followed the estimating input feature lag with ACF and RF methods. We independently estimated the input lag using the ACF and RF method, then tried to build the LSTM and CNN model with a hybrid approach.

Model Evaluation

The performance of each model was assessed on the testing dataset using suitable evaluation metrics measures, including Mean Absolute Error (MAE), Mean Absolute Squared Error (MASE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). In addition, graphical representations such as time series plots, residual plots, and Radar and Taylor plots were used to evaluate the accuracy of the model to depict the underlying patterns and dynamics in the data. Statistical tests, such as the Ljung-Box test, can be used to ascertain if any model exhibits a significantly superior performance compared to the others.

The analysis was done in the Python platform with some special packages, i.e. `pmdarima` (<https://pypi.org/project/pmdarima/>), `arch` (<https://pypi.org/project/arch/>), and `TesnsorFlow` (<https://keras.io>).

Results and Discussion

Basic Statistics

The performance metrics mean, standard deviation, skewness, kurtosis, and outlier test were applied to the potato wholesale price index (WPI) (Table 1). The dataset recorded an average of 170.510 with a variance of 64.227, indicating the presence of an outlier in November 2020, as tested by Grubbs' test statistic (4.992; $Z_{\text{tab}}=3.645$). The boxplot diagram in Fig. 2 further demonstrates the presence of outliers. The skewness and kurtosis values showed that the distribution was asymmetrical and leptokurtic. Based on the monthly plot, the maximum potato price is expected to rise in November (Fig. 2).

Table 1 Statistical performance of potato price index series

| | Mean | SD | Skewness | Kurtosis |
|-----|---------------------------|-----------------|----------|----------|
| WPI | 170.510 | 64.227 | 1.4131 | 3.048 |
| | Grubbs' test for outliers | | | |
| | 4.992 (test statistic) | 491.1; Nov 2020 | | |

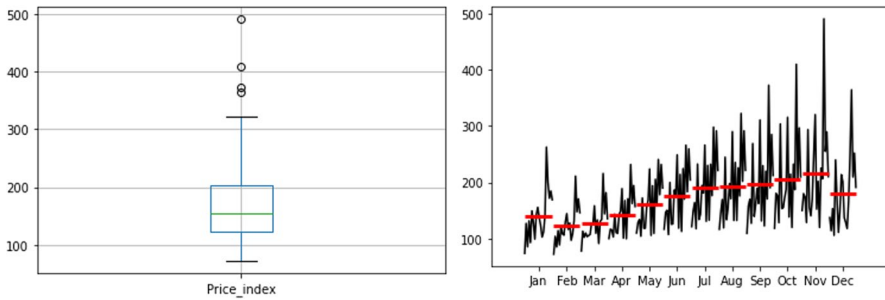


Fig. 2 Boxplot and monthly plot of potato WPI

Model Development

Statistical Model

The main purpose of this research is to find the best time series model with their prediction accuracy of the potato price index data. We employed statistical models such as ARIMA and GARCH based on the assumption criteria and machine learning models such as SVM, LSTM, CNN, hybrid RF-LSTM, and RF-CNN. The models were chosen using the technique described in the “Materials and Methods” section and compared for accuracy in predicting the data series. The goal is to select the most accurate and appropriate model for the analysed data series. As the analysis was carried out in Python, `autoarima` package (<https://pypi.org/project/pmdarima/>) was utilized and estimated that ARIMA(2,1,2) model was the best based on lower Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Also, the Augmented Dickey-Fuller (ADF) test (Dickey and Fuller 1979) satisfied the data being stationary at 1st difference with a p -value of less than 0.05 (test statistic -6.182 at lags=19). So, ARIMA(2,1,2) model was applied to the training data and estimated the parameter and standard error (Table 2). The Ljung-Box test confirmed that residual autocorrelation at different lags was zero, since the p -value

Table 2 Statistical model parameter and residual test

| | Parameter estimation | | | Ljung-Box test | | |
|---------------|----------------------|-----------|-------|----------------|-----------|------------|
| | Coefficient | Estimates | SE | Residuals | Statistic | p -Value |
| ARIMA (2,1,2) | ar1 | 1.205 | 0.113 | Lag 5 | 0.730235 | 0.981268 |
| | ar2 | -0.3652 | 0.11 | Lag 10 | 4.250527 | 0.935338 |
| | ma1 | -0.6985 | 4.557 | Lag 15 | 21.611454 | 0.118404 |
| | ma2 | -0.3015 | 1.389 | | | |
| GARCH (1, 1) | mu | 0.5828 | 1.651 | Lag 5 | 0.475188 | 0.993004 |
| | omega | 3.4192 | 1.396 | Lag 10 | 4.029097 | 0.946025 |
| | alpha[1] | 0.7774 | 0.229 | Lag 15 | 19.889211 | 0.176224 |
| | beta[1] | 0.4431 | 0.227 | | | |

was greater than 0.05. The diagnostic checking in Fig. 3 confirmed the presence of volatility in the residuals ($Q-Q$ plot, Residual plot), implying that the GARCH model had to be employed for enhanced prediction.

Before applying the GARCH, it was necessary to check the presence of volatility of residuals using the ARCH-LM test. The test confirmed that the volatility was observed at different lags up to 20 with p -value less than 0.01 (test statistic 32.588). As a result, the GARCH model was applied to the trained residual data generated from ARIMA. GARCH (1,1) was chosen as the best model due to its lower AIC and BIC. The Ljung-Box test likewise demonstrated residual normality, with a p -value larger than 0.05 (Table 2). We also provide the conditional volatility and standardized residual plots from the best fitted GARCH model (Fig. 4). Higher conditional variance was reported when volatility behaviour was detected in the data. This variance volatility prediction confirmed that the GARCH model was superior to the mean model ARIMA, as we compared it with lower RMSE, MAPE, MAE, and MASE in both training and testing sets (Tables 5 and 6).

Machine Learning Model

After gathering information on prediction behavior in statistical models, we attempted to use certain prominent deep learning models such as SVM, LSTM, and CNN. Before implementing these models, the nonlinearity diagnostic check should be performed. As shown in Table 3, the BDS test (Brock et al. 1996) with two embedding dimensions was found to be significant (p -value less than 0.01),

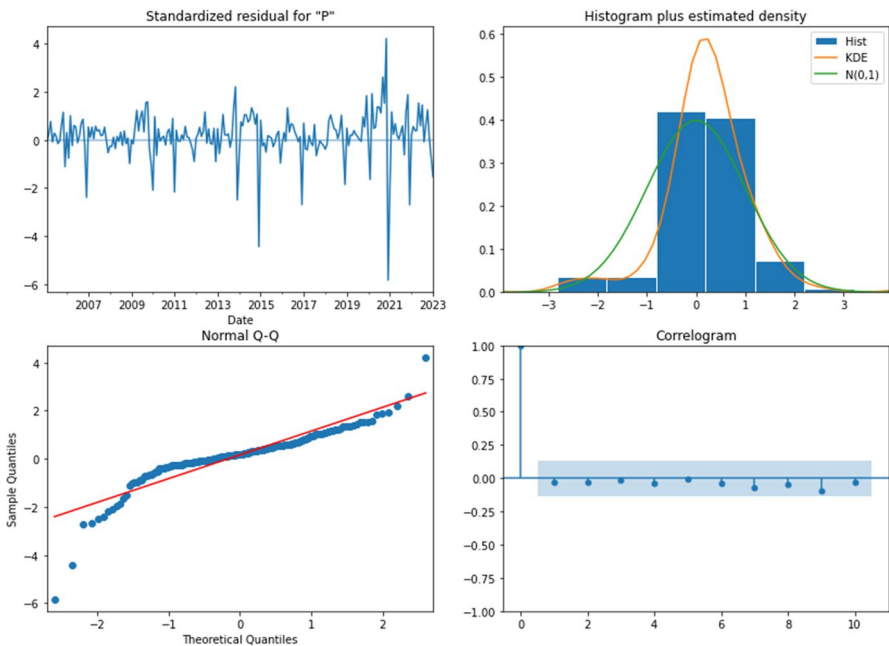


Fig. 3 Diagnostic check for ARIMA

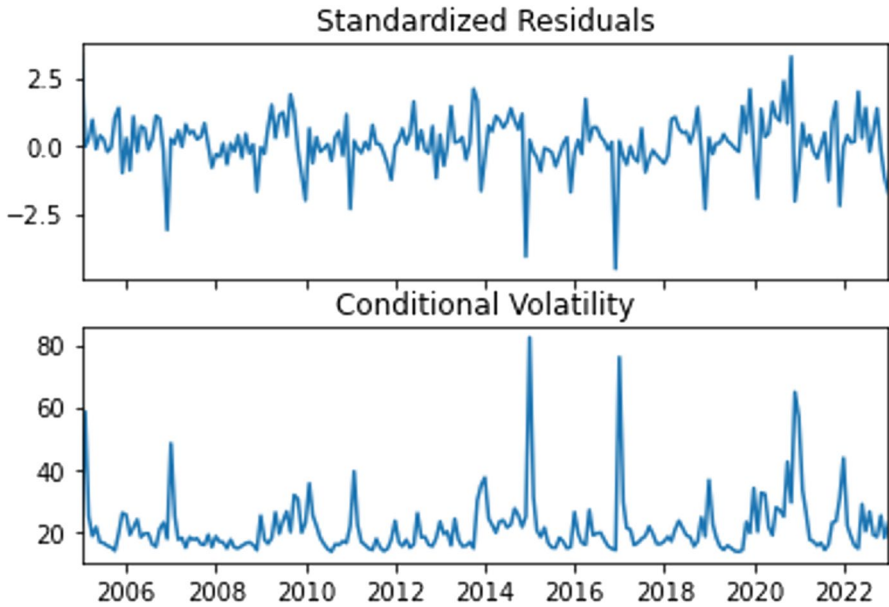


Fig. 4 Conditional volatility of variance

Table 3 BDS test for nonlinearity

| Epsilon | Embedding dimension | Statistic | <i>p</i> -Value |
|---------|---------------------|-----------|-----------------|
| 1 | 2 | 55.6346 | <0.01 |
| | 3 | - 13.9323 | <0.01 |
| 2 | 2 | 72.5543 | <0.01 |
| | 3 | 212.525 | <0.01 |
| 3 | 2 | 64.0861 | <0.01 |
| | 3 | 197.9581 | <0.01 |
| 4 | 2 | 70.5501 | <0.01 |
| | 3 | 190.5247 | <0.01 |

indicating that the data set is not independent and identically distributed (*i.i.d*). The potato WPI data set shown below can then be used to estimate and predict using machine or deep learning models.

By confirming the nonlinearity of the data, the Support Vector Machine (SVM) algorithm was employed with 100 regularization parameters (*C*), 0.1 gamma with 0.1 epsilon parameter (Table 4) Lama et al. (2024). The SVM model was quite similar to the statistical models (ARIMA and GARCH), followed by the goodness of fit (Tables 5 and 6). So, we applied two deep learning algorithms by adopting the input lag selection criteria (Ray et al. 2023). Autocorrelation function (ACF) plot and random forest (RF) technique were utilized to estimate the appropriate input lag for deep learning models LSTM and CNN. As reported in Fig. 5, the maximum significant lag

Table 4 Machine learning model parameter

| | Lag selection criterion | | Neuron | Dense | Total param | Epochs |
|-----|-------------------------|-----|--------|-------|-------------|--------|
| | ACF | RF | | | | |
| | LSTM | 4 | | | | |
| CNN | | | | | 301 | |
| SVM | C | 100 | | | | |
| | Gamma | 0.1 | | | | |
| | Epsilon | 0.1 | | | | |

Table 5 Measure of error accuracy in train data

| Model | RMSE | MAPE | MAE | MASE |
|---------------|-----------|----------|----------|----------|
| ARIMA (2,1,2) | 23.443251 | 8.85187 | 15.20347 | 0.821624 |
| GARCH (1, 1) | 22.3412 | 8.8503 | 15.0043 | 0.80187 |
| SVM | 23.9801 | 7.7146 | 19.6732 | 0.9971 |
| ACF-LSTM | 29.426986 | 1.31526 | 22.19953 | 1.1997 |
| ACF-CNN | 28.021202 | 1.645126 | 22.74275 | 1.22906 |
| RF-LSTM | 7.793024 | 0.45379 | 6.554748 | 0.35423 |
| RF-CNN | 7.50902 | 0.364997 | 5.55503 | 0.300204 |

Heat mapping with high to low (Red to Green)

Table 6 Measure of error accuracy in test data

| Model | RMSE | MAPE | MAE | MASE |
|---------------|-----------|----------|-----------|----------|
| ARIMA (2,1,2) | 42.79223 | 1.740842 | 35.71723 | 1.93022 |
| GARCH (1, 1) | 31.70423 | 1.50388 | 28.94211 | 1.56408 |
| SVM | 41.100908 | 1.84072 | 31.50057 | 1.9871 |
| ACF-LSTM | 19.85913 | 0.86462 | 17.33601 | 0.93687 |
| ACF-CNN | 13.8701 | 0.703345 | 11.67085 | 0.630715 |
| RF-LSTM | 4.32044 | 0.20969 | 3.96496 | 0.21427 |
| RF-CNN | 3.285906 | 0.16029 | 3.0212809 | 0.16327 |

Heat mapping with high to low (Red to Green)

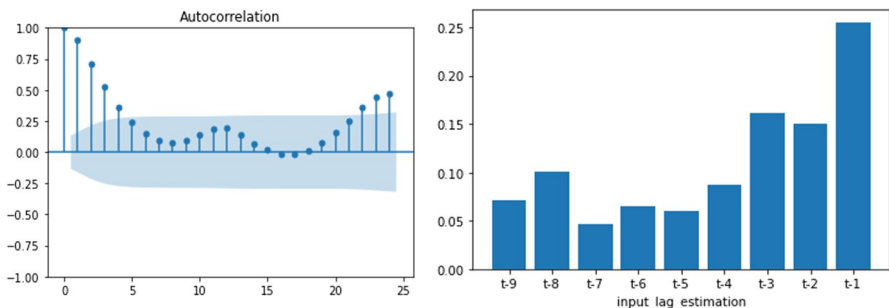


Fig. 5 Input lag estimation with ACF and RF

was observed up to four lags in the correlogram of ACF, whilst the maximum correlation was obtained in the first lag in RF classifications. So, we employed LSTM and CNN twice with input features 4 and 1. The hypertuning parameter follows 100

neurons, one dense, activation function ‘ReLU’ (Rectified Linear Unit); ‘Adam’ optimizer; and ‘mse’ loss function with 25 epochs (Table 4). As the analysis platform was Python, ‘MinMaxScaler’ was first utilized to normalize the potato WPI data series. A ‘TimeseriesGenerator’ from the ‘Keras’ package was used to modify the data series to oversee neural network learning. After developing the model, both LSTM and CNN, we incorporate the run test of residuals. The test satisfied that the residuals are random as the p -value is greater than 0.05. This confirmed that the proposed model satisfied all the assumptions of the randomness of residuals, indicating the model to be a good fit for potato WPI data.

Discussion and Comparison

The research objective focused on comparing statistical and machine learning models. We also implemented the input lag feature technique with ACF and RF for better accuracy (Ray et al. 2023). Finally, we compared all the selected models based on

Table 7 The measure of forecasting accuracy of different models

| Date | Actual | ARIMA(2,1,2) | GARCH (1,1) | SVM | ACF-LSTM |
|----------|--------|--------------|-------------|---------|----------|
| Feb-23 | 146.0 | 146.785 | 178.677 | 241.556 | 151.117 |
| Mar-23 | 135.2 | 136.271 | 156.591 | 180.954 | 145.475 |
| April-23 | 157.0 | 137.412 | 154.769 | 144.044 | 156.913 |
| May-23 | 190.0 | 142.627 | 161.471 | 133.490 | 174.606 |
| Jun-23 | 204.3 | 148.494 | 166.313 | 154.695 | 182.405 |
| Jul-23 | 221.3 | 153.659 | 170.571 | 186.008 | 191.733 |
| Aug-23 | 221.4 | 157.741 | 175.685 | 199.251 | 191.788 |
| Sept-23 | 212.2 | 160.772 | 178.248 | 214.715 | 186.725 |
| Oct-23 | 209.7 | 162.935 | 179.129 | 214.805 | 185.355 |
| Nov-23 | 210.3 | 164.433 | 178.998 | 206.476 | 185.683 |
| Dec-23 | 191.1 | 165.450 | 178.531 | 204.196 | 175.203 |
| Jan-24 | 169.1 | 166.127 | 188.752 | 204.744 | 163.347 |
| | Actual | ACF-CNN | RF-LSTM | RF-CNN | |
| Feb-23 | 146.0 | 168.871 | 149.664 | 148.922 | |
| Mar-23 | 135.2 | 160.893 | 140.514 | 139.226 | |
| April-23 | 157.0 | 177.004 | 159.047 | 158.797 | |
| May-23 | 190.0 | 201.401 | 187.579 | 188.424 | |
| Jun-23 | 204.3 | 211.973 | 200.123 | 201.262 | |
| Jul-23 | 221.3 | 224.542 | 215.176 | 216.524 | |
| Aug-23 | 221.4 | 224.616 | 215.265 | 216.614 | |
| Sept-23 | 212.2 | 217.814 | 207.099 | 208.355 | |
| Oct-23 | 209.7 | 215.965 | 204.888 | 206.110 | |
| Nov-23 | 210.3 | 216.409 | 205.418 | 206.649 | |
| Dec-23 | 191.1 | 202.214 | 188.540 | 189.412 | |
| Jan-24 | 169.1 | 185.949 | 169.442 | 169.660 | |

the minimum RMSE, MAPE, MAE, and MASE values in both training and testing sets (Tables 5 and 6). Table 7 and Fig. 7 represent all the selected model predictions for the 12 ahead point forecast. A total of seven models were compared based on their predicting accuracy. As reported in Tables 5 and 6, the RF-CNN model has the lowest RMSE, MAPE, MAE, and MASE followed by the RF-LSTM model for both training and testing sets. ACF-LSTM and ACF-CNN models were not found to be superior in the training set. This could imply that ACF failed to estimate the lag input feature for this volatile data series. The Radar plot (Fig. 6) for training and testing sets provides enough evidence to identify the best model amongst all applied models using the statistical measure for goodness of fit. The Taylor diagrams were used to analyse the spatial design of projected and actual WPI magnitudes across all models (Fig. 6). These diagram simulation plots provide visually appealing model performance evaluations using statistical indicators like standard deviation (SD) and correlation coefficient (r). The diagram shows similar conclusions from the training and testing evaluation statistic as with the Radar plots. As indicated in the plot, RF-CNN model prediction in the last 12 ahead points have less variation with the actual observation with a high positive correlation, confirming the appropriate model for the potato WPI data series (Fig. 7) (Mishra et al. 2024c).

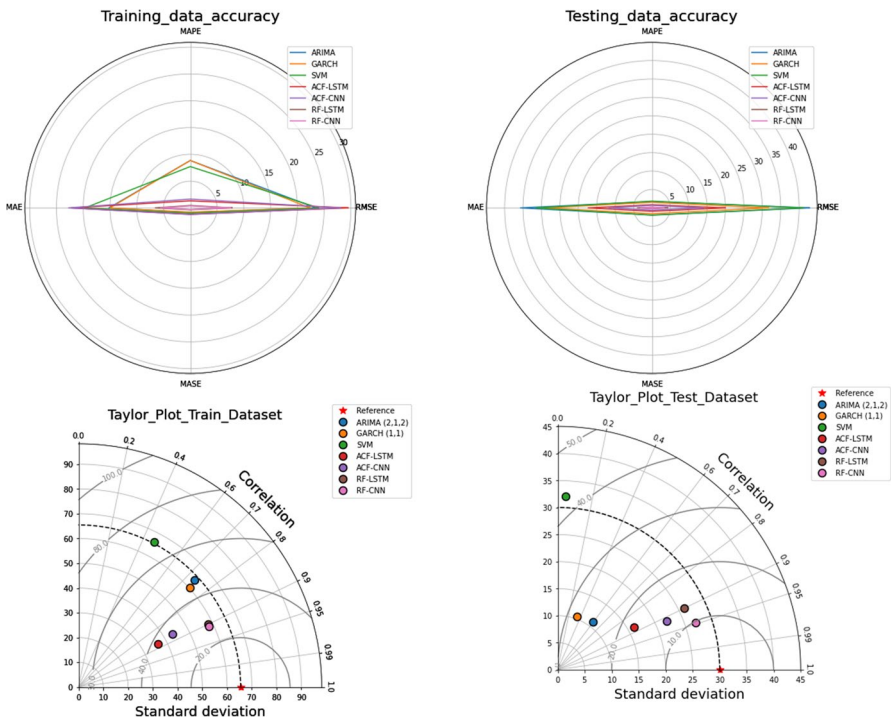


Fig. 6 Accuracy testing with Radar plot and Taylor plot

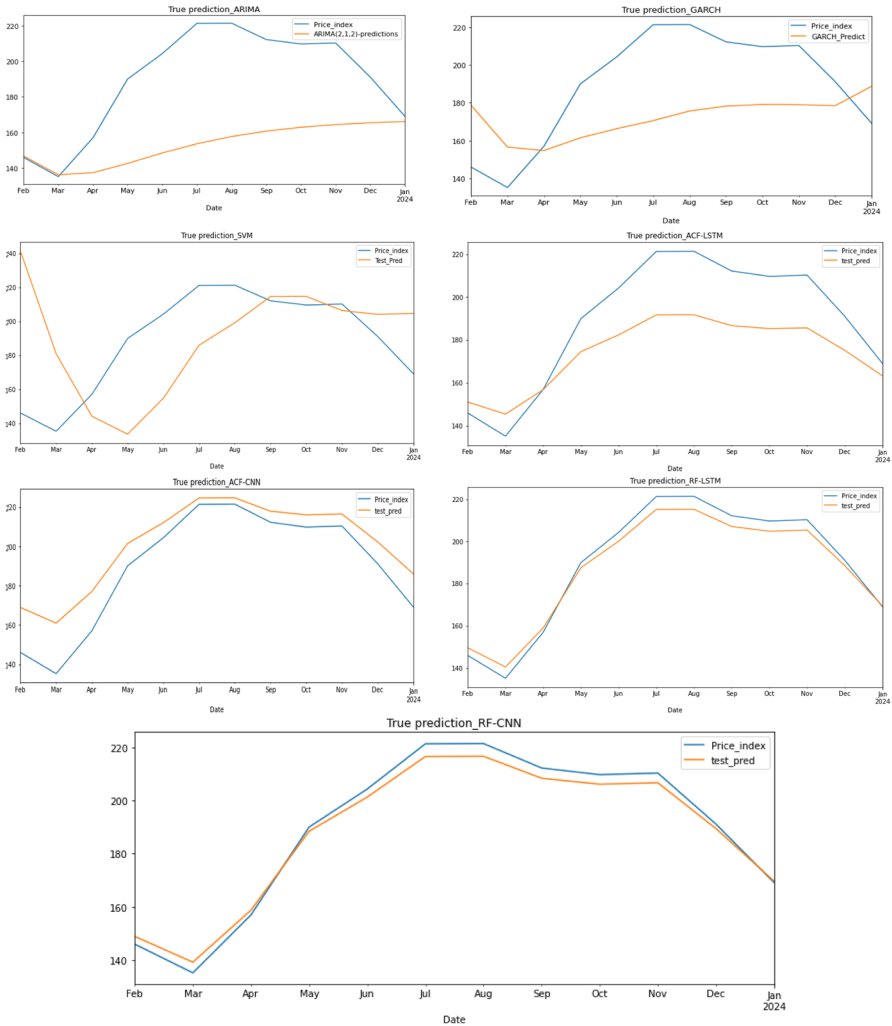


Fig. 7 Test prediction based on different models

Conclusion

In this research, we have introduced a hybrid technique of the RF-CNN model for predicting agricultural volatile price index commodities. Potato WPI data was used for building such model. The traditional statistical time series model (ARIMA and GARCH model) was also studied to compare it with the machine learning model (SVM, LSTM, and CNN model). As the data set followed a volatile nature, the GARCH model performed better than the ARIMA model based on a lower goodness of fit value. The performance of the SVM model is *at par* with the statistical models. However after adopting the input lag selecting method using ACF and RF,

the machine learning models performed well enough than the statistical models. We applied the LSTM and CNN model using the selected input lag feature estimated by ACF and RF. Our findings suggest that the RF-CNN model outperforms the other models in terms of error accuracy, with improvements of 67% for RMSE, 95% for MAPE, 63% for MAE and MASE for the training set, and more than 90% for the testing set for all goodness of fit.

This manuscript highlights the significance of accurate lag length selection for deep learning models (LSTM and CNN) for time series forecasting. We revised a hybrid technique of RF-CNN to predict accurate volatile series. We anticipate our study will aid stakeholders and policymakers by giving reliable potato price forecasts. Also, our study adds to the expanding body of literature on machine learning models for time series.

Funding The study was funded by Researchers Supporting Project number (RSPD2024R749), King Saud University, Riyadh, Saudi Arabia.

Data Availability On reasonable request, the corresponding author will provide data supporting the study's results. The raw data cannot be made public for reasons of confidentiality and privacy. However, researchers who satisfy the requirements for access to confidential data can be given access to aggregated and anonymized data as well as the statistical analysis codes. To request access to the data, interested researchers can get in touch with the corresponding author at WhatsApp+919560073489 or pradeep-jnkvv@gmail.com.

Declarations

Ethics Approval and Consent to Participate Not applicable.

Conflict of Interest The authors declare no competing interests.

References

- Abbasimehr H, Behboodi A, Bahrini A (2024) A novel hybrid model to forecast seasonal and chaotic time series. *Expert Syst Appl* 239:122461
- Adudotla SS, Bobba P, Pathan Z, Kata T, Sobin CC, Jahfar (2022) A method for price prediction of potato using deep learning techniques. In: International conference on intelligent vision and computing. Springer Nature Switzerland, Cham. pp 619–629
- Bento PMR, Pombo JAN, Calado MRA, Mariano SJPS (2021) Stacking ensemble methodology using deep learning and ARIMA models for short-term load forecasting. *Energies* 14:7378. <https://doi.org/10.3390/en14217378>
- Black F, Scholes M (1973) The pricing of options and corporate liabilities. *J Pol Econ* 81(3):637
- Brock WA, Scheinkman JA, Dechert WD, LeBaron B (1996) A test for independence based on the correlation dimension. *Economet Rev* 15:197–235
- Chen L, Wu T, Wang Z, Lin X, Cai Y (2023) A novel hybrid BPNN model based on adaptive evolutionary Artificial Bee Colony Algorithm for water quality index prediction. *Ecol Ind* 146:109882
- Dhakre DS, Bhattacharya D (2016) Price behaviour of potato in agra market—a statistical analysis. *Indian Res J Ext Educ* 14(2):12–15
- Dickey DA, Fuller WA (1979) Distribution of the estimators for autoregressive time series with a unit root. *J Am Stat Assoc* 74:427–431. <https://doi.org/10.2307/2286348>
- Gebrechistos HY, Chen W (2018) Utilization of potato peel as eco-friendly products: a review. *Food Sci Nutr* 6(6):1352–1356

- Gulay E, Sen M, Akgun OB (2024) Forecasting electricity production from various energy sources in Türkiye: a predictive analysis of time series, deep learning, and hybrid models. *Energy* 286:129566
- Jagait RK, Fekri MN, Grolinger K, Mir S (2021) Load forecasting under concept drift: online ensemble learning with recurrent neural network and ARIMA. *IEEE Access* 9:98992–99008
- Júnior DSDOS, de Oliveira JF, de Mattos Neto PS (2019) An intelligent hybridization of ARIMA with machine learning models for time series forecasting. *Knowl-Based Syst* 175:72–86
- Kumar B, Yadav N (2023) A novel hybrid model combining β SARMA and LSTM for time series forecasting. *Appl Soft Comput* 134:110019
- Lama A, Ray S, Biswas T et al (2024) Python code for modeling ARIMA-LSTM architecture with random forest algorithm. *Softw Impacts*. <https://doi.org/10.1016/j.simpa.2024.100650>
- Lin Y, Li S, Li B, Li G, Jin L, Liu J (2023) Methodological evolution of potato yield prediction: a comprehensive review. *Front Plant Sci* 14:1214006
- Mishra P, Yonar A, Yonar H, Kumari B, Abotaleb M, Das SS, Patil SG (2021) State of the art in total pulse production in major states of India using ARIMA techniques. *Curr Res Food Sci* 1(4):800–806
- Mishra P, Al Khatib AMG, Lal P, Anwar A, Nganvongpanit K, Abotaleb M, Punyapornwithaya V (2023) An overview of pulses production in India: retrospect and prospects of the future food with an application of hybrid models. *Natl Acad Sci Lett* 1–8. <https://doi.org/10.1007/s40009-023-01267-2>
- Mishra P, Alhussan AA, Khafaga DS, Lal P, Ray S, Abotaleb M et al (2024a) Forecasting production of potato for a sustainable future: global market analysis. *Potato Res*. <https://doi.org/10.1007/s11540-024-09717-0>
- Mishra P, Al khatib AMG, Alshaib BM, Kuamri B, Tiwari S, Singh AP, Yadav S, Sharma D, Kumari P (2024b) Forecasting potato production in major South Asian countries: a comparative study of machine learning and time series models. *Potato Res*. <https://doi.org/10.1007/s11540-023-09683-z>
- Mishra P, Al Khatib AMG, Yadav S, Ray S, Lama A, Kumari B, Yadav R (2024c) Modeling and forecasting rainfall patterns in India: a time series analysis with XGBoost algorithm. *Environ Earth Sci* 83(6):1–15. <https://doi.org/10.1007/s12665-024-11481-w>
- Özden C (2023) Comparative analysis of CNN, LSTM and random forest for multivariate agricultural price forecasting. *Black Sea J Agric* 6(4):422–426
- Rahman UH, Ray S, Mohammad A, Al G, Lal P, Mishra P et al (2022) State of art of SARIMA model in second wave on COVID-19 in India. *Int J Agricult Stat* 18(1):141–152
- Ray S, Lama A, Mishra P, Biswas T, Das SS, Gurung B (2023) An ARIMA-LSTM model for predicting volatile agricultural price series with random forest technique. *Appl Soft Comput* 149:110939
- Şahinli MA (2020) Potato price forecasting with Holt-Winters and ARIMA methods: a case study. *Am J Potato Res* 97(4):336–346. <https://doi.org/10.21203/rs.3.rs-4011255/v1>
- Sahu PK, Das M, Sarkar B, VS A, Dey S, Narasimhaiah L, Mishra P, Tiwari RK, Raghav YS (2024) Potato production in India: a critical appraisal on sustainability forecasting price and export behaviour. *Potato Res*. <https://doi.org/10.1007/s11540-023-09682-0>
- Salman D, Direkoglu C, Kusaf M, Fahrioglu M (2024) Hybrid deep learning models for time series forecasting of solar power. *Neural Comput Appl* 1–18. <https://doi.org/10.1007/s00521-024-09558-5>
- Shankar SV, Chandel A, Gupta RK, Sharma S, Chand H, Aravinthkumar A, Ananthkrishnan S (2024) Comparative study on key time series models for exploring the agricultural price volatility in potato prices
- Wang Z, Yang L, Yin J, Zhang B (2018) Assessment and prediction of environmental sustainability in China based on a modified ecological footprint model. *Resour Conserv Recycl* 132:301–313
- Yadav S, Al khatib AMG, Alshaib BM, Ranjan S, Kumari B, Alkader NA, Mishra P, Kapoor P (2024) Decoding potato power: A global forecast of production with machine learning and state-of-the-art techniques. *Potato Res*. <https://doi.org/10.1007/s11540-024-09705-4>
- Zaheer K, Akhtar MH (2016) Potato production, usage, and nutrition—a review. *Crit Rev Food Sci Nutr* 56(5):711–721

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Soumik Ray¹  · Tufleuddin Biswas¹ · Walid Emam² · Shikha Yadav³ ·
Priyanka Lal⁴ · Pradeep Mishra⁵ 

✉ Pradeep Mishra
pradeepjnkvv@gmail.com

- ¹ Department of Agricultural Economics and Statistics, Centurion University of Technology and Management, Paralakhemundi, Odisha, India
- ² Department of Statistics and Operations Research, Faculty of Science, King Saud University, P.O. Box 2455, 11451 Riyadh, Saudi Arabia
- ³ Department of Geography, Miranda House, University of Delhi, New Delhi 11007, India
- ⁴ Department of Agricultural Economics and Extension, School of Agriculture, Lovely Professional University, Punjab, India
- ⁵ College of Agriculture, Jawaharlal Nehru Krishi Vishwa Vidyalaya (JNKVV), Rewa 486001, India