International Journal of Horticulture, Agriculture and Food Science (IJHAF)
ISSN: 2456-8635
[Vol-10, Issue-4, Oct-Dec, 2026]
Issue DOI: https://dx.doi.org/10.22161/ijhaf.10.4
Peer-Reviewed Journal



Evaluation of suitable forecasting models for major vegetable crops price prediction in Madurai district-A Statistical Approach

K. Prabakaran

Professor (Agrl. Statistics), Department of Agricultural Economics, Agricultural College and Research Institute, Madurai – 625 104, India
ORCID ID: 0000-0001-8901-0907

Received: 09 Oct 2026; Received in revised form: 11 Nov 2026; Accepted: 16 Nov 2026; Available online: 23 Nov 2026
©2026 The Author(s). Published by AI Publications. This is an open-access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/)

Abstract— *This study examines the growth trends and dispersion patterns of brinjal, bhendi, tomato, and onion prices in the Madurai district, while also evaluating the suitability of various price prediction models. Secondary price data for these selected crops were utilized in the analysis. The results revealed that prices for all four crops exhibited positive growth rates and high dispersion patterns during the study period. The forecasting performance of ARIMA, ANN, SVR, and LSTM models was evaluated using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The evaluation indicates that the SVR model performed better for bhendi, brinjal, and onion, yielding lower error values across all three metrics. In contrast, the LSTM model demonstrated the best performance for tomato price prediction, registering the lowest MAE and RMSE values. Overall, SVR emerged as the most consistent model for forecasting vegetable prices, while LSTM proved superior specifically for tomato price prediction.*

Keywords— ***Vegetables, Trend, Dispersion pattern, Forecast models***

I. INTRODUCTION

Crop performance projections are critical for agricultural logistics and trade policies. India is the world's second-largest producer of vegetables. Tamil Nadu is a leading horticulture state in the country, dedicating about 3.34 lakh hectares to vegetable production with an annual output exceeding 82 lakh metric tonnes. Tapioca, onion, tomato, brinjal, moringa (drumstick), and lady's finger (okra) account for over 75% of the total area under vegetable cultivation. Specifically, Brinjal (eggplant), bhendi (okra), onion and tomato are among the major vegetable crops produced in the Madurai district of Tamil Nadu. The highest concentration of vegetable farming in this occurs in blocks such as Usilampatti, Sedapatti, and Chellampatti.

Horticultural production is influenced by several factors, including cultivated area, productivity, weather conditions, irrigation availability, input use, technology, pest and disease incidence and market conditions. Consequently, the production performance of a crop may change considerably from year to year, even when its long-term direction remains stable. Understanding the historical patterns of area, production and productivity is therefore important for assessing whether changes in output are primarily associated with the expansion of cultivated area or improvements in productivity.

Along with production performance, price behaviour is an important consideration in horticultural crops. Most horticultural commodities are perishable and, therefore, cannot be stored for long periods without appropriate infrastructure. Farmers may consequently face pressure to sell their produce soon after harvest. When market prices fluctuate considerably, the timing of harvesting and marketing becomes a crucial factor influencing farm returns. Price uncertainty can also affect decisions regarding crop selection, area allocation and the quantity of inputs used in cultivation.

The importance of district-level analysis arises from the fact that agricultural performance can vary substantially even within the same state. State-level

averages may mask differences in cropping patterns, resource availability, climatic conditions and market access that exist between districts. A district-specific analysis can therefore provide more relevant information for local production planning and market intelligence. The present study focuses on Madurai district to examine both long-term changes in production characteristics and short-term movements in retail prices.

This integrated approach is particularly relevant because a crop with favourable production growth may still experience considerable price instability. Similarly, high historical price dispersion does not necessarily imply poor short-term forecastability. Examining these characteristics separately and then bringing them together can provide a clearer understanding of the production and market environment of selected horticultural crops in Madurai district.

Objectives of the study:

- To assess the selected vegetable crops (Onion, Tomato, Brinjal and Bhendi) price trend in Madurai district.
- To study the selected vegetable crops price dispersion pattern in Madurai district.
- To evaluate the suitable price prediction model for selected vegetable crops in Madurai district.

II. METHODLOGY

In the Madurai district, onion, tomato, brinjal, and bhendi are the major cultivation crops, exhibiting high production volumes among the various vegetable crops. Therefore, these four commodities were purposively selected for this study. Secondary daily price data for the selected crops were collected from major markets within the Madurai district—specifically Mattuthavani market, Paravai market, and Uzhavar Santhai—covering the period from 2019 to 2024. To ensure consistency across all forecasting approaches, the raw daily price observations were transformed into a uniform, model-ready format through a systematic pre-processing pipeline before applying the ARIMA, SVR, ANN, and LSTM models. This pipeline included the aggregation of daily prices into a weekly frequency, data cleaning and preparation, and the partitioning of each time series into distinct training and testing datasets for model development and evaluation.

For the SVR, ANN and LSTM models, the same weekly price series was additionally transformed into a supervised learning format, where each observation comprised a set of lagged price values as predictor variables and the corresponding future price as the target variable. Creating both representations from the same verified weekly series ensured that all forecasting models were developed using an identical underlying price history for each crop, with differences arising only from the manner in which the data were organised for model input.

To furnish the supervised-learning models with predictors capturing the recent history of each price series, a set of lagged variables was generated from the weekly series of each crop. Specifically, for a given week t, lagged price values were constructed at lags of 3, 6, 9, 12 and 24 weeks, that is, as P_{t-3}, P_{t-6}, P_{t-9}, P_{t-12} and P_{t-24}, where P_t denotes the observed weekly price in week t. These specific lags were selected to represent, within a single feature set, price behaviour over a short recent window of approximately one to three months (lags 3, 6 and 9), a medium window of roughly one quarter (lag 12) and a longer window of approximately six months (lag 24). This multi-window approach allows the supervised learning models to capture both recent momentum and longer seasonal or cyclical patterns that typically characterise horticultural commodity prices.

Because the machine learning and deep learning models compared in this study, namely SVR, ANN and LSTM are sensitive to the numerical scale of their input variables and because prices of the four crops examined differ substantially in their typical magnitude, the lagged predictors and the target price series generated were rescaled to a common range prior to model fitting, using min-max normalization. For a given series, the min-max scaled value corresponding to an original observation x_t is computed as

$$x_t' = (x_t - x_{min}) / (x_{max} - x_{min})$$

where x_t is the original, unscaled value of the price or lagged predictor in week t; x_{min} and x_{max} are, respectively, the minimum and maximum values taken by that variable over the estimation and x_t' is the resulting scaled value, which lies in the closed interval [0, 1]. Min-max scaling was preferred over alternative scaling methods, such as standardisation to zero mean and unit variance, because it bounds every rescaled observation to a known, fixed range, which is particularly advantageous for the artificial neural network and long short-term memory models, whose internal activation functions are most effective over a bounded input range.

For multi-step forecasting, the fitted models were applied recursively to generate the twelve-week-ahead forecasts. In this approach, the forecast for the first future week was generated using the available lagged observations and the resulting forecast was subsequently used as an input for generating the forecast for the next week. The recursive forecasting framework can be expressed as

$$\hat{P}_{t+h} = f\left(P^*_{t+h-l_1}, P^*_{t+h-l_2}, \ldots, P^*_{t+h-l_k}\right),\ h = 1,$$

where $\hat{P}_{t+h}$ denotes the forecast price h weeks ahead, $f(\cdot)$ represents the fitted forecasting model, $l_1, \ldots, l_k$ are the selected lag values and P^* denotes either an observed price or a previously forecast price, depending on whether the corresponding lag falls within the observed or forecast period. This recursive procedure was applied consistently to the SVR, SVM, ANN and LSTM models for generating the twelve-week-ahead forecasts.

Training and Testing Data Preparation

To enable out-of-sample evaluation of forecasting performance, the model-ready weekly series for each crop was partitioned chronologically into training and testing datasets. The training dataset comprised the earlier observations and was used for model identification, parameter estimation and network training for the ANN and LSTM models, while the testing dataset consisted of the final 12 weeks, corresponding to the forecasting horizon. Chronological partitioning was adopted to preserve the temporal sequence of the data and to prevent information from future observations from influencing model training. The same training-testing split was applied across all crops and forecasting models to ensure a consistent and unbiased comparison of forecasting accuracy.

Forecast accuracy assessment

The forecasting performance of the five models was evaluated using the testing dataset of each crop's weekly price series, which was excluded from all stages of model training and hyperparameter optimisation. Model performance was assessed using four complementary accuracy measures: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and the Coefficient of Determination (R^2).

III. RESULTS AND DISCUSSION

The weekly retail price observations of onion, tomato, brinjal, and bhendi for the period from 2014 to 2019 were utilized in this study to examine the compound growth rate (CGR) and the Cuddy-Della Valle Index (CDVI), as well as to evaluate the suitability of various forecasting models for the selected crops.

I. Assessment of selected vegetable crops price trend and dispersion pattern in Madurai district

The compound growth rate (CGR)and Cuddy Della Valle Index(CDVI) were computed for the selected vegetable crops and are presented in Table 1.

Table 1: Compound growth rate and Cuddy Della Valle Index for the selected vegetable crops price in Madurai district

Crop Details	CAGR	P Value	CDVI(%)	Dispersion Pattern
Onion Price	36.85**	0.007	38.16	High
Tomato Price	36.82**	0.003	75.12	High
Brinjal Price	36.77**	0.008	38.32	High
Bhendi Price	36.78**	0.005	32.68	High

The results in Table 1 reveal that the prices of all four selected crops onion, tomato, brinjal, and bhendi exhibited high, positive, and statistically significant growth rates. Based on the Cuddy-Della Valle Index (CDVI), price instability ranged from 32.68% for bhendi and 38.16% for onion to 38.32% for brinjal. The highest price instability was recorded for tomato, yielding a CDVI of 75.12%. These findings demonstrate a high price dispersion pattern across all selected crops during the study period. The historical price trends for these crops are further illustrated in Figure 1.

II. Evaluation of suitable price prediction model for selected vegetable crops in Madurai district

To identify the most suitable forecasting model in this study, the predictive performance of ARIMA, ANN, SVR, and LSTM models was evaluated and compared. Following model identification, parameter estimation, and diagnostic checking, the forecasting accuracy of each model was evaluated using three standard metrics: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE). Particular emphasis was placed on MAPE, as it provides a scale-free measure that allows for reliable comparison across different crops with distinct price magnitudes. These evaluation results are presented in Table 2.

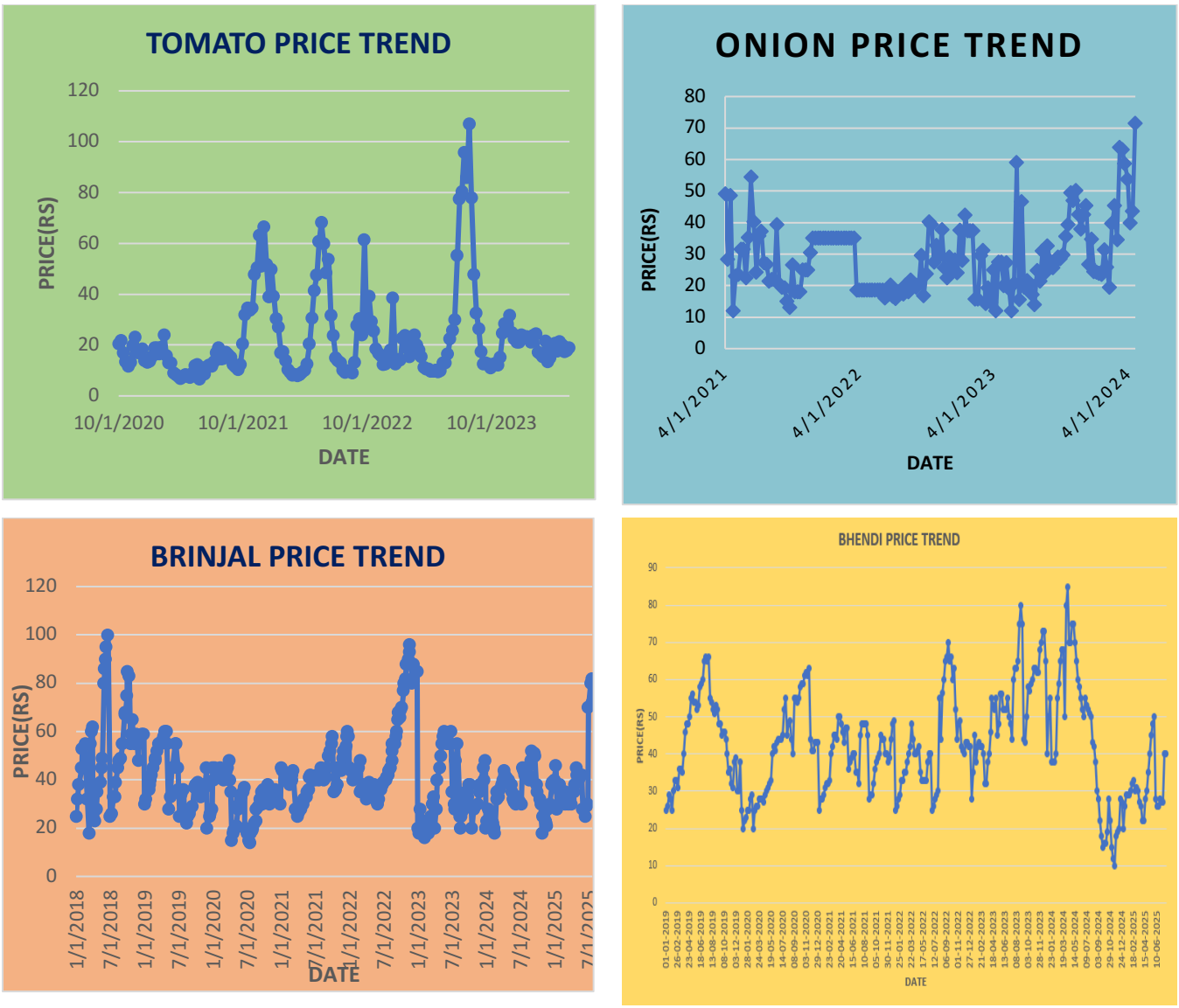

Fig.1: Trend in Tomato, Onion, Brinjal and Bhendi crops prices in Madurai district

Table 2: Evaluation of suitable forecasting models for the selected vegetable crops price in Madurai district

Crops	Model	RMSE	MAE	MAPE
Bhendi	**ARIMA**	21.18	18.76	0.75
	ANN	9.6	6.16	0.15
	SVR	6.84	4.51	**0.15**
	LSTM	19.77	17.56	0.63
Brinjal	**ARIMA**	12.01	8.79	0.26
	ANN	6.35	4.31	0.14
	SVR	6.27	4.3	**0.14**
	LSTM	11.27	9.35	0.33

Onion	**ARIMA**	9.96	8.10	0.49
	ANN	22.64	11.81	0.30
	SVR	20.94	12.71	**0.27**
	LSTM	29.32	17.30	0.40
Tomato	**ARIMA**	18.12	13.72	0.29
	ANN	12.14	7.78	0.29
	SVR	12.68	9.64	0.26
	LSTM	10.47	7.95	**0.25**

The empirical results demonstrate that for bhendi, the SVR model achieved the lowest error rates, with an RMSE of 6.84, an MAE of 4.51, and a MAPE of 0.15, indicating that SVR is the most suitable model for forecasting bhendi prices. Similarly, for brinjal, the SVR model yielded the lowest error metrics (RMSE = 6.27, MAE = 4.30, and MAPE = 0.14), confirming its selection for brinjal price prediction in the Madurai district. A consistent pattern emerged for onion price prediction, where the SVR model again minimized the error criteria, registering an RMSE of 20.94, an MAE of 12.71, and a MAPE of 0.27. Consequently, SVR was identified as the optimal model for onion price forecasting.

In contrast, for tomato price prediction, the LSTM model demonstrated superior performance, registering the lowest overall errors with an RMSE of 10.47, an MAE of 7.95, and a MAPE of 0.25. Based on these metrics, the LSTM model was selected as the most appropriate framework for tomato price prediction. Overall, SVR emerged as the most consistent and robust model for forecasting onion, brinjal, and bhendi prices, whereas LSTM proved superior specifically for capturing the highly volatile price dynamics of tomatoes.

Twelve weeks forecast using the selected models

Based on the comparative evaluation, Support Vector Regression (SVR) was selected as the final forecasting model for onion, bhendi, and brinjal prices, while the Long Short-Term Memory (LSTM) model was selected for tomato price prediction. The twelve-week-ahead point forecasts generated by these models are presented in Table 3.

Table 3: Twelve weeks ahead selected vegetable crops price forecasted values in Madurai district

Week	Forecast date	Onion	Tomato	Forecast date	Bhendi	Brinjal
1	23.05.2024	25.79	23.38	29.07.2025	50.42	31.75
2	30.05.2024	25.80	20.99	05.08.2025	49.84	33.32
3	06.06.2024	26.13	23.99	12.08.2025	50.05	36.50
4	13.06.2024	26.20	23.99	19.08.2025	50.01	39.73
5	20.06.2024	26.34	23.63	26.08.2025	46.65	38.00
6	27.06.2024	26.71	24.20	02.09.2025	45.16	34.88
7	04.07.2024	27.05	24.38	09.09.2025	47.83	32.05
8	11.07.2024	27.07	24.06	16.09.2025	48.07	32.72
9	18.07.2024	27.08	23.56	23.09.2025	50.86	33.98
10	25.07.2024	27.09	23.17	30.09.2025	51.11	34.32
11	01.08.2024	27.12	23.40	07.10.2025	46.83	32.65
12	08.08.2024	27.37	24.34	14.10.2025	44.69	36.55

The results in Table 3 reveal distinct short-term price movements among the selected vegetable crops. The forecasted retail price for onion exhibits a steady upward trend over the twelve-week horizon. Conversely, the price

forecasts for the other three crops—tomato, bhendi, and brinjal—show notable instability across the twelve-week period, characterized by fluctuations with price increases in certain weeks followed by declines in others.

IV. CONCLUSION

This study demonstrates that all four selected vegetable crops—onion, tomato, brinjal, and bhendi—in the Madurai district exhibit positive growth trends alongside high price dispersion patterns. Through a rigorous comparative evaluation, Support Vector Regression (SVR) was identified as the optimal forecasting model for onion, bhendi, and brinjal price prediction. In contrast, the Long Short-Term Memory (LSTM) model proved superior for capturing the specific price dynamics of tomatoes. The twelve-week-ahead forecasts indicate persistent price instability for tomato, brinjal, and bhendi, characterized by frequent fluctuations, while onion prices are projected to follow a steady upward trajectory. Ultimately, the robust forecasting frameworks identified in this study provide accurate market price intelligence that can help farmers, consumers, wholesalers, and policymakers make well-informed production, marketing, and trade decisions.

CONFLICT OF INTEREST

I declare **no conflict of interest**, particularly regarding any financial or personal relationships with other people or organizations that could inappropriately influence or bias this work.

AUTHOR FUNDING

There is no fund involved and it was carried out as a university sub project. That project outcome submitted as a research article.

AUTHOR CONTRIBUTION

The entire research work carried out individually. (Statement of the problem, Data collection, Analyse the data, Interpretation, Report and Article publication)

ARTICLE TYPE

Original Research Article

REFERENCES

[1] Areef M, Rajeswari S, Vani N and Mohan Naidu G,2020, “Forecasting of Onion Prices in Bangalore Market: An Application of Time Series Models” , Indian Journal of Agricultural Economics, Vol.75, No.2, April-June 2020,pp217-227.

[2] Atharva Ratnaparkhe, Somya Kothari ,Siddharth Bhatia, Bimal Kumar Mawandiya and Bharatkumar Modi, 2022, “Price and Production Forecasting Model of Okra – A Case Study” Proceedings of the 2nd Indian International Conference on Industrial Engineering and Operations Management Warangal, Telangana, India, August 16-18, 2022.

[3] Mohit Kumar, Rohit Kumar Sharma, Upasana D Bhopala and Bhoomi Suthar, 2024, “Forecasting prices of onion in major wholesale markets of Gujarat”, International Journal of Statistics and Applied Mathematics 2024; SP-9(1): 22-29.

[4] Pal Vikash, Darji VB and Chaudhari Raju, 2021, “Price forecasting of Brinjal: A statistical evaluation”, The Pharma Innovation Journal 2021; SP-10(12): 2105-2110.

[5] Tamilselvi C, Md Yeasin , Ranjit Kumar Paul and Amrit Kumar Paul,2024, “Can Denoising Enhance Prediction Accuracy of Learning Models? A Case of Wavelet Decomposition Approach”, Forecasting 2024, 6, 81–99.