



International Journal of Research in Agronomy

E-ISSN: 2618-0618
P-ISSN: 2618-060X
© Agronomy
NAAS Rating (2025): 5.20
www.agronomyjournals.com
2025; 8(10): 944-946
Received: 07-08-2025
Accepted: 10-09-2025

Anjali Gilhare
M.Sc. Scholar, Department of
Agricultural Statistics and Social
Sciences (L.), Indira Gandhi Krishi
Vishwavidyalaya, Raipur,
Chhattisgarh, India

ML Lakhera
Professor, Department of
Agricultural Statistics and Social
Sciences (L.), Indira Gandhi Krishi
Vishwavidyalaya, Raipur,
Chhattisgarh, India

Yamuna Kashyap
M.Sc. Scholar, Department of
Agricultural Statistics and Social
Sciences (L.), Indira Gandhi Krishi
Vishwavidyalaya, Raipur,
Chhattisgarh, India

Sweta Ramole
Professor, Department of
Agricultural Statistics and Social
Sciences (L.), Indira Gandhi Krishi
Vishwavidyalaya, Raipur,
Chhattisgarh, India

Kuldeep Tandan
M.Sc. Scholar, Department of
Agricultural Statistics and Social
Sciences (L.), Indira Gandhi Krishi
Vishwavidyalaya, Raipur,
Chhattisgarh, India

Corresponding Author:
Anjali Gilhare
M.Sc. Scholar, Department of
Agricultural Statistics and Social
Sciences (L.), Indira Gandhi Krishi
Vishwavidyalaya, Raipur,
Chhattisgarh, India

Price prediction of pigeon pea (Arhar) using statistical and machine learning models in Rajnandgaon District of Chhattisgarh

Anjali Gilhare, ML Lakhera, Yamuna Kashyap, Sweta Ramole and Kuldeep Tandan

DOI: <https://www.doi.org/10.33545/2618060X.2025.v8.i10m.4089>

Abstract

The present study aimed to forecast the monthly modal prices of pigeon pea (Arhar) in the Rajnandgaon district of Chhattisgarh using statistical and machine learning models. Twenty years of secondary data (2003-2023) were utilized from the Agmarknet portal, Directorate of Economics and Statistics, and meteorological sources. The study compared the predictive performance of the traditional Autoregressive Integrated Moving Average (ARIMA) model and the Support Vector Regression (SVR) model. Descriptive statistics revealed high volatility and seasonal variation in Arhar prices. The ARIMA model captured linear trends effectively but struggled with abrupt fluctuations. In contrast, the SVR model achieved superior forecasting accuracy with an R^2 value of 0.9460 compared to 0.4469 for ARIMA, and significantly lower MAE and RMSE values (42.4 and 35.67 respectively). Among SVR kernels, the Linear and Radial Basis Function (RBF) kernels performed best. These findings highlight SVR's effectiveness in handling nonlinear price dynamics, offering a reliable forecasting framework to aid policymakers, traders, and farmers in informed decision-making and price risk management.

Keywords: Arhar, price prediction, ARIMA, SVR, Kernels

1. Introduction

Pigeon pea (*Cajanus cajan* L.), commonly known as Arhar or Tur, is a major pulse crop cultivated across India, valued for its nutritional content and soil-enriching ability. It plays a vital role in ensuring food and nutritional security, especially for the vegetarian population. The Rajnandgaon district of Chhattisgarh, with favorable agro-climatic conditions, is a leading region in pulse cultivation. However, the prices of pigeon pea are highly volatile due to climatic variability, market fluctuations, and supply-demand mismatches, which cause uncertainty in farmers' income. Accurate price forecasting serves as an essential tool for strategic market planning, crop selection, and policy formulation.

Traditional time-series models such as ARIMA have been widely used for price forecasting but are limited in capturing nonlinear and complex market patterns. Machine learning models like Support Vector Regression (SVR) offer a data-driven alternative capable of modeling nonlinear dependencies and improving predictive accuracy. Therefore, this study focuses on comparing ARIMA and SVR models for forecasting the monthly prices of pigeon pea in the Rajnandgaon district.

Mishra *et al.* (2024) ^[5] reported machine learning techniques, including XGBoost, to forecast rainfall patterns across different seasons in India. Comparative analysis shows XGBoost outperforms traditional models like state space, SVM, ANN, and Random Forest in capturing complex nonlinear rainfall trends. The findings provide valuable insights for water resource management, agriculture, and power generation, highlighting the effectiveness of advanced machine learning in meteorological forecasting.

Sahu *et al.* (2020) ^[6] The study aimed to forecast paddy prices using a univariate Seasonal ARIMA (SARIMA) model based on monthly price data. The SARIMA (2,1,2) (1,1,1) model was identified as the most suitable for price prediction. Forecasts were generated for a 13-month

period from March 2020 to March 2021. Model selection was based on the highest R^2 value and the lowest Mean Absolute Percentage Error (MAPE), ensuring reliable forecasting performance.

Jadhav *et al.* (2017) ^[2] The study aimed to forecast farm prices of Paddy, Ragi, and Maize in Karnataka using time series data (2002-2016) through the ARIMA model. Forecast accuracy was evaluated using MSE, MAPE, and Theil's U coefficient. The results showed that ARIMA effectively predicted prices for 2020, with low error values confirming its reliability for price forecasting.

Meesad, P. *et al.* (2013) ^[4] This study applies Support Vector Regression (SVR) with various windowing operators as preprocessing techniques to predict stock prices and trends for ACI Ltd. on the Dhaka Stock Exchange. The results show that the Win SVR models effectively capture market patterns, producing accurate forecasts when compared with actual prices.

2. Materials and Methods

The study will cover a span of 20 years, from 2003-2023. it will focus on major pulse crop pigeon pea (Arhar). Secondary data required for the study, including information on minimum price, maximum price and modal price of selected crops for the year 2003-2023 from sources (<https://agmarknet.gov.in/>), e-NAM, state mandi boards, (<https://desagri.gov.in/>), Department of Meteorology, I.G.K.V, Raipur etc.

All the analytical works would be done by using an open-source software called “R: The R Project for Statistical Computing”, which is fantastic for graphics and programming and is available at <https://www.R-project.org>.

2.1 ARIMA (Autoregressive Integrated Moving Average process) Model

The ARIMA (Autoregressive Integrated Moving Average) model is a powerful tool for analysing non-stationary time series data. By incorporating past values and stochastic errors, ARIMA (p, d, q) effectively captures patterns and provides accurate forecasts.

2.2 Support Vector Regression (SVR): SVR is a supervised machine learning technique used for regression tasks. It is based on Statistical Learning Theory and is effective in handling small samples, nonlinearity, and high-dimensional data.

SVR model that fits the best possible line (or curve) within a margin. It works finds a hyperplane that minimizes error within a given tolerance. The key objective is to maximize the margin, which is the distance between the hyperplane and the nearest

data points, called support vectors. This approach enhances the model's generalization and predictive accuracy.

Table 1: Different types of kernels

S. No.	Kernel Type	Equation
1.	Linear	$k(x,y) = x^T y + c$
2.	Polynomial	$k(x,y) = (ax^T y + c)^2$
3.	Radial Basis Function (RBF)	$k(x,y) = \exp(-\frac{\ x-y\ ^2}{2\sigma^2})$

2.3 Model Evaluation Metrics

Coefficient of Determination (R ²)	$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$
Mean Absolute Error (MAE)	$MAE = \frac{1}{n} \sum_{i=1}^n y_i - \hat{y}_i $
Mean Squared Error (MSE)	$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$

3. Results and Discussion

3.1 Traditional methods and Machine learning techniques

This study utilizes traditional statistical and machine learning methods to predict crop prices using historical data from the central plains of Chhattisgarh. The main objective is to compare traditional statistical price predicting methods with machine learning techniques.

3.1.1 Arhar prices

Figure 1 illustrates the ARIMA model's actual versus predicted modal prices for Arhar during 2023, revealing notable discrepancies, especially during sharp price movements at the year's end where the model underestimates peaks. The ARIMA predictions tend to lag behind actual values, struggling with price volatility and abrupt transitions, which results in higher prediction errors and lower correlation with real trends.

In contrast, Figure 2 demonstrates the SVR model's (linear kernel) superior alignment with actual Arhar prices over the years, accurately capturing both trends and major price fluctuations. The SVR model closely tracks rapid increases and declines, evidencing its capacity to handle complex, nonlinear market dynamics and producing lower errors, thereby confirming its suitability for agricultural price forecasting tasks.

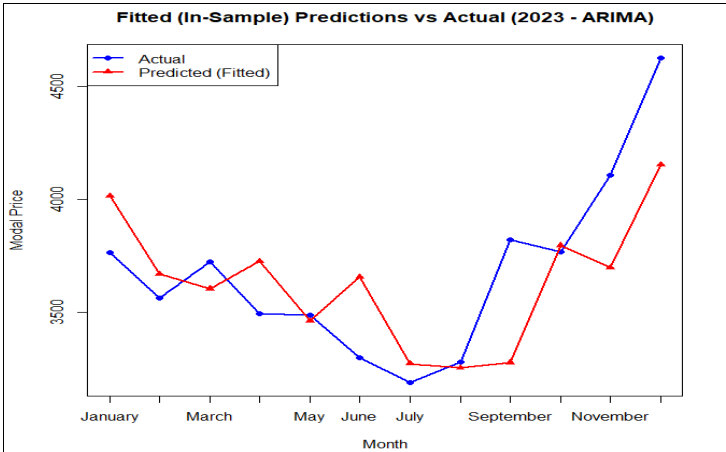


Fig 1: Comparison of actual and predicted Arhar price for ARIMA model

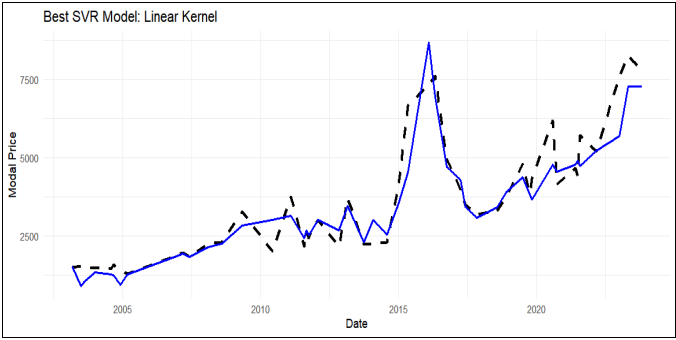


Fig 2: Comparison of actual and predicted Arhar price for SVR model

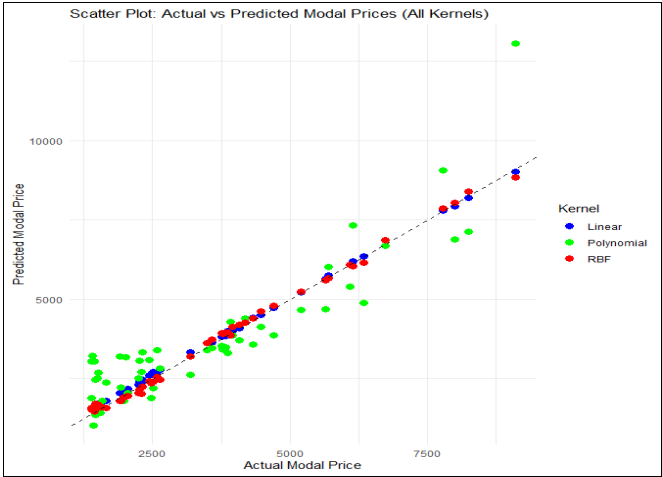


Fig 4: Comparison of all kernels for arhar SVR method

3.1.2 Evaluation metrics

Table 2 presents a comparative analysis of the ARIMA and SVR models on price prediction for Arhar crop. The SVR model significantly outperforms the ARIMA model for Arhar crop, demonstrating much lower MAE and RMSE values, which indicate higher accuracy and less prediction error. For instance, in Arhar, SVR’s MAE and RMSE (42.4 and 35.67) are substantially lower than ARIMA’s (222 and 284.25). Additionally, SVR shows higher R-Square values indicating a better fit and predictability than ARIMA for Arhar crop. This comparison underscores the superiority of SVR for accurate and reliable crop price forecasting in this study. The result is akin to those of Mohanty *et al.* (2023), suggested that the decision tree regressor is found to be the best model, for predicting crop price, over others.

Table 2: Comparison of evaluation metrics for both models:

METRICS	ARHAR CROP	
	ARIMA Method	SVR Method
MAE	222	42.4
RMSE	284.25	35.67
R-Square	0.4469	0.9460

3.2 Comparison of all kernels

The comparison of all SVR kernels for Arhar price predicting (Fig. 3 and Fig. 4) shows that both linear and RBF (radial) kernels closely follow the actual monthly price trend, with minimal deviation, while the polynomial kernel exhibits greater fluctuation and consistent over/underestimation during price peaks and troughs. The scatter plot further confirms that predictions from linear and RBF kernels are tightly clustered along the ideal reference line, indicating high accuracy, whereas the polynomial kernel points display more dispersion, reflecting less reliable performance.

Conclusion

The comparative analysis between ARIMA and SVR models demonstrated that the Support Vector Regression model provides significantly higher forecasting accuracy for pigeon pea prices. SVR effectively captures nonlinear price variations and seasonal effects that ARIMA fails to model. The SVR model with a linear kernel emerged as the best fit, offering an R² value of 0.9460, highlighting its robustness for agricultural price forecasting. The comparative results confirm that both Linear and RBF kernels outperform the Polynomial kernel in terms of forecast accuracy, stability, and generalization capability. Among them, the Linear kernel emerged as the most balanced and computationally efficient option, making it the preferred model for pigeon pea (Arhar) price prediction in Rajnandgaon district.

References

1. Gujarati DN. Basic Econometrics. New York: McGraw-Hill; 2004.
2. Jadhav V, Chinnappa RB, Gaddi GM. Application of ARIMA model for forecasting agricultural prices. Vol. 19. 2017:981-92.
3. Makridakis S, Wheelwright SC. Forecasting: Methods and Applications. New York: John Wiley and Sons; 1978.
4. Meesad P, Rasel RI. Predicting stock market price using support vector regression. In: 2013 International Conference on Informatics, Electronics and Vision (ICIEV); 2013 May. p. 1-6. IEEE.
5. Mishra P, Al Khatib AMG, Yadav S, Lama A, Kumari B, Sharma D, *et al.* Modeling and forecasting rainfall patterns in India: a time series analysis with XGBoost algorithm. Environ Earth Sci. 2024;83:163.
6. Sahu CR, Lakhera ML, Walke SS. Forecasting paddy prices in Ambikapur market of Chhattisgarh state using SARIMA approach. Int J Curr Microbiol Appl Sci. 2020;9(4):1366-72.

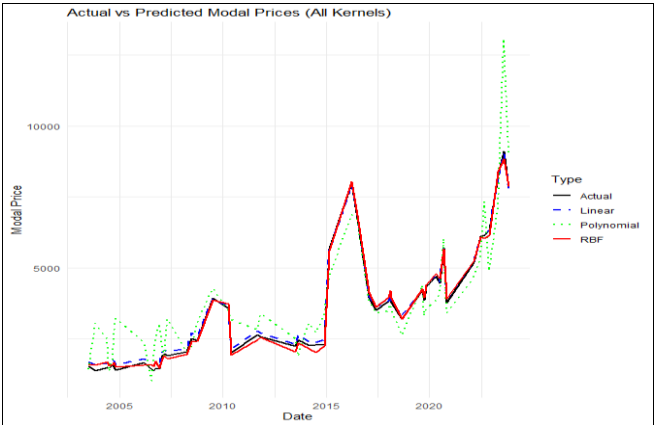


Fig 3: Comparison of all kernels for arhar SVR method