

Performance Assessment of Statistical, Machine Learning and Hybrid Forecasting Approaches for Pearl Millet Production in the Jammu Division

Divyam Sharma, Manish Sharma and Nishant Jasrotia

Division of Statistics and Computer Science, SKUAST-Jammu

Received: 16 July 2026; Revised: 19 August 2026; Accepted: 23 August 2026

Abstract

Pearl millet (*Pennisetum glaucum*) is an important cereal crop contributing to food and nutritional security, necessitating accurate production forecasting for effective agricultural planning. The present study evaluated and compared the forecasting performance of statistical and machine learning-based time series models for pearl millet production in the Jammu division using secondary data for the period 1966–2024. Simple Exponential Smoothing (SES), Double Exponential Smoothing (DES), Autoregressive Integrated Moving Average (ARIMA), Nonlinear Support Vector Regression (NLSVR), and Hybrid ARIMA-SVR models were developed and evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), the Diebold–Mariano test, and the Chi-square goodness-of-fit test. The Brock–Dechert–Scheinkman (BDS) test rejected the null hypothesis of independent and identically distributed residuals, indicating significant nonlinear dependence in the ARIMA residuals and supporting the use of machine learning models. Among all models, NLSVR recorded the lowest testing RMSE (0.42) and MAE (0.31) and adequately fitted the historical data and hence is the most suitable model for forecasting pearl millet production in the Jammu division.

Key words: Pearl millet; Time series forecasting; ARIMA; NLSVR; Hybrid ARIMA-SVR; Machine learning.

AMS Subject Classifications: 62D05, 91B82

1. Introduction

Agriculture is the backbone of the Indian economy and according to the economic survey 2024-2025, it contributes approximately 17 to 18% to the country's Gross Value Added (GVA) and providing livelihoods to about 46.1% of the country's workforce, thereby making a significant contribution to national economic growth, rural development, and food security (Mehmood, 2025). India possesses diverse agro-climatic conditions, which facilitate the cultivation of a wide variety of crops. Since independence, the agricultural sector has

made remarkable progress in terms of area, production, and productivity of several crops. Among cereal crops, pearl millet (*Pennisetum glaucum*), commonly known as bajra, ranks sixth globally and fourth after rice, wheat, and maize in India (Rani *et al.*, 2024). It is one of the major food crops cultivated in the arid and semi-arid regions of the country and is valued for its drought tolerance and nutritional quality. Pearl millet is believed to have originated in Africa and later spread to India. It is widely cultivated in countries such as India, Nigeria, Sudan, and Saudi Arabia and is commonly known as bulrush millet, bajra, or cattail millet. The grains are rich in nutrients, containing 9 to 15 per cent protein, about 5 per cent fat, and 2 to 7 per cent minerals, besides being good sources of vitamins A and B, thiamine, and riboflavin (Pal *et al.*, 1996).

Pearl millet is mainly cultivated during the kharif season in India. However, summer cultivation under irrigated conditions generally provides higher yields and better grain quality than kharif cultivation (Reddy *et al.*, 2013). Summer cultivation also experiences lower disease incidence due to the disease escape mechanism. Nevertheless, kharif pearl millet occupies a larger area because rainfall substantially reduces irrigation costs. Despite its importance, prolonged dry spells and adverse climatic conditions often affect crop productivity. Keeping in view the importance of millet, the United Nations General Assembly at its 75th session in March 2021, at the behest of the Government of India, declared 2023 the International Year of Millets (IYM 2023), with the aim to create awareness about health and nutritional benefits of millets as also to increase production and consumption of millets. During 2024–25, pearl millet in the Jammu division recorded a production of 8.69 thousand tonnes (Directorate of Economics and Statistics, Jammu and Kashmir, 2025). Although bajra is the principal millet crop grown in the Union Territory of Jammu and Kashmir, its productivity remains relatively low, necessitating the adoption of improved cultivars and appropriate package of practices to ensure sustainable cultivation.

Considering the economic, nutritional, and climatic significance of pearl millet, it is essential to study its production behaviour over time and assess future prospects. Time series analysis provides an effective framework for understanding historical trends and generating reliable forecasts. In this context, various statistical and machine learning approaches such as exponential smoothing techniques Simple Exponential Smoothing (SES) and Double Exponential Smoothing (DES), Autoregressive Integrated Moving Average (ARIMA) models, Nonlinear Support Vector Regression (NLSVR), and hybrid ARIMA–SVR models have been widely adopted to capture both linear and nonlinear patterns in agricultural time series data. Exponential smoothing methods are simple yet powerful tools for short-term forecasting, while ARIMA models are effective in modelling linear dependencies in time series data. However, agricultural data often exhibit nonlinear behaviour due to climatic variability and socio-economic factors. To address this limitation, machine learning techniques such as NLSVR are employed. Furthermore, hybrid models like ARIMA–SVR combine the strengths of both statistical and machine learning approaches, leading to improved forecasting accuracy. Therefore, the present study was undertaken to evaluate and compare the forecasting performance of exponential smoothing, ARIMA, NLSVR, and hybrid ARIMA–SVR models for pearl millet production in the Jammu division and to identify the most appropriate model for accurate future forecasting.

2. Material and methods

Data: The study was carried out for the Jammu region. The secondary data on pearl millet production and area (in hectares) for a period of 59 years (1966-2024) were collected from various authentic sources, including the Directorate of Economics and Statistics, Jammu and Kashmir, and the Indian Institute of Millets Research.

Statistical time series models:

(i) Exponential smoothing: Exponential smoothing is a widely used time series forecasting technique that generates forecasts by assigning exponentially decreasing weights to past observations, thereby placing greater emphasis on more recent values. Depending on the characteristics of the time series, Simple Exponential Smoothing (SES) and Holt's Linear Trend (Double Exponential Smoothing) methods may be employed. SES is appropriate for series exhibiting no systematic trend or seasonality, whereas Holt's method extends SES by incorporating a linear trend component (Holt, 1957). The SES forecasting equation is given by

$$F_{t+1} = F_t + \alpha(Y_t - F_t)$$

where α is the smoothing constant, taking values between 0 and 1, and F_t is the forecast for Y_t .

For Holt's linear trend method, the level and trend components are estimated as

$$\text{Level equation: } l_t = \alpha Y_t + (1 - \alpha)(l_{t-1} + b_{t-1})$$

$$\text{Trend equation: } b_t = \gamma(l_t - l_{t-1}) + (1 - \gamma)b_{t-1}$$

$$\text{Forecast equations: } Y_{t+1} = l_t + hb_t$$

where, l_t denotes an estimate of the level of the series at time t , b_t denotes an estimate of the trend (slope) of the series at time t , α is the smoothing parameter for the level and γ is the smoothing parameter for the trend which range from 0 to 1. The last estimated level plus h times the last estimated trend value is the h -step-ahead forecast. Forecasts thus represent a linear function of h .

(ii) Auto regressive integrated moving average: ARIMA (Auto regressive integrated moving average) models are widely used for modelling and forecasting univariate time series data, particularly when non-stationary series can be transformed into stationary series through differencing. An ARIMA model is denoted as ARIMA (p, d, q) , where p represents the auto regressive order, d denotes the degree of differencing required to achieve stationarity, and q represents the moving average order. The general form of an ARIMA model is expressed as:

$$Y_t = \sum_{i=1}^p \beta_i Y_{t-i} + e_t + \sum_{j=1}^q \phi_j e_{t-j}$$

where Y_t is the observed value at time t , α is the intercept, β_i are the autoregressive coefficients, ϕ_j are the moving average coefficients, and e_t is the random error term (Box and Jenkins, 1976).

Machine learning time series models

(i) Non linear support vector regression (NLSVR) model: NLSVR primarily

converts the initial input space into a feature space with higher dimensions. Subsequently, a linear regression model is built within it, effectively representing non-linear regression in the original space. In the context of a dataset represented by $Z = \{X_i Y_i\}_{i=1}^N$, where the input vector X_i belongs to the n -dimensional real space, Y_i represents the scalar output, and N denotes the size of Z , the general equation of NLSVR can be expressed as follows:

$$f(x) = W^T \phi(x) + b$$

where w represents the weight vector, $\phi(x)$ stands for the non-linear mapping function, b signifies the bias term, and superscript T indicates the transpose operation. The data is used to estimate the coefficients w and b through the minimization of a regularized risk function:

$$R(\theta) = \frac{1}{2} \|W\|^2 + C \left[\frac{1}{N} \sum_{i=1}^N L_E(Y_i, f(X_i)) \right].$$

The above equation comprises two elements: the first is a regularized component represented as half times the norm of w squared, while the second component is referred to as the empirical error, denoted as $\frac{1}{N} \sum_{i=1}^N L_E(Y_i, f(X_i))$. The regularized risk function effectively balances the optimization of both of these components simultaneously, thus preventing both underfitting and overfitting issues of the model.

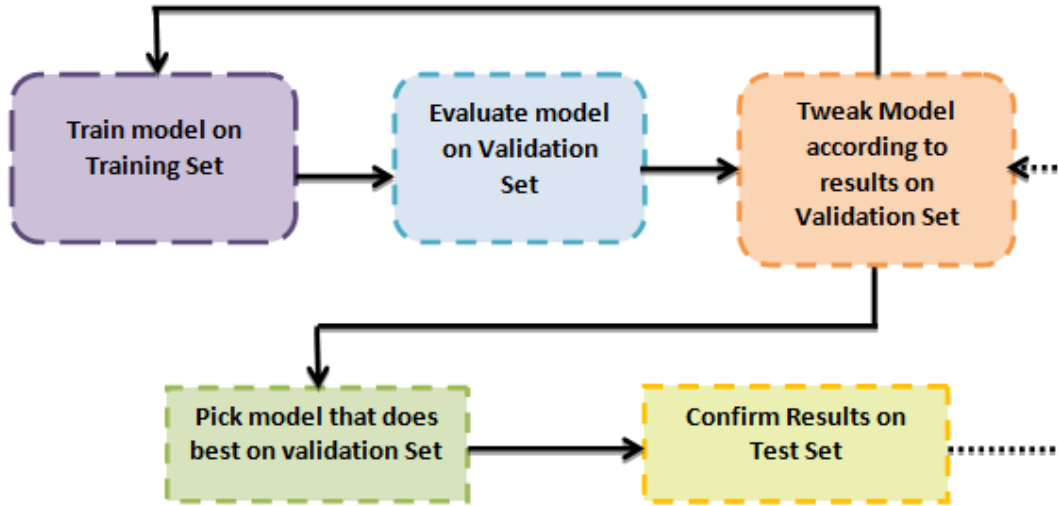


Figure 1: Standard machine learning workflow using training, validation, and test datasets

(ii) **Hybrid models:** Hybrid time series models combine statistical and machine learning techniques to improve forecasting accuracy by capturing both linear and nonlinear patterns in the data. In this study, an ARIMA–SVR hybrid model was used for forecasting pearl millet production in the Jammu region. First, an ARIMA model is fitted to the original series to capture the linear structure, and residuals are obtained as:

$$e_t = Y_t - \hat{Y}_t^{(ARIMA)}.$$

Since residuals may contain nonlinear information, Support Vector Regression (SVR) is applied to model them. The final hybrid forecast is generated by combining both components:

$$\hat{Y}_t^{(Hybrid)} = \hat{Y}_t^{(ARIMA)} + \hat{e}_t^{(SVR)}.$$

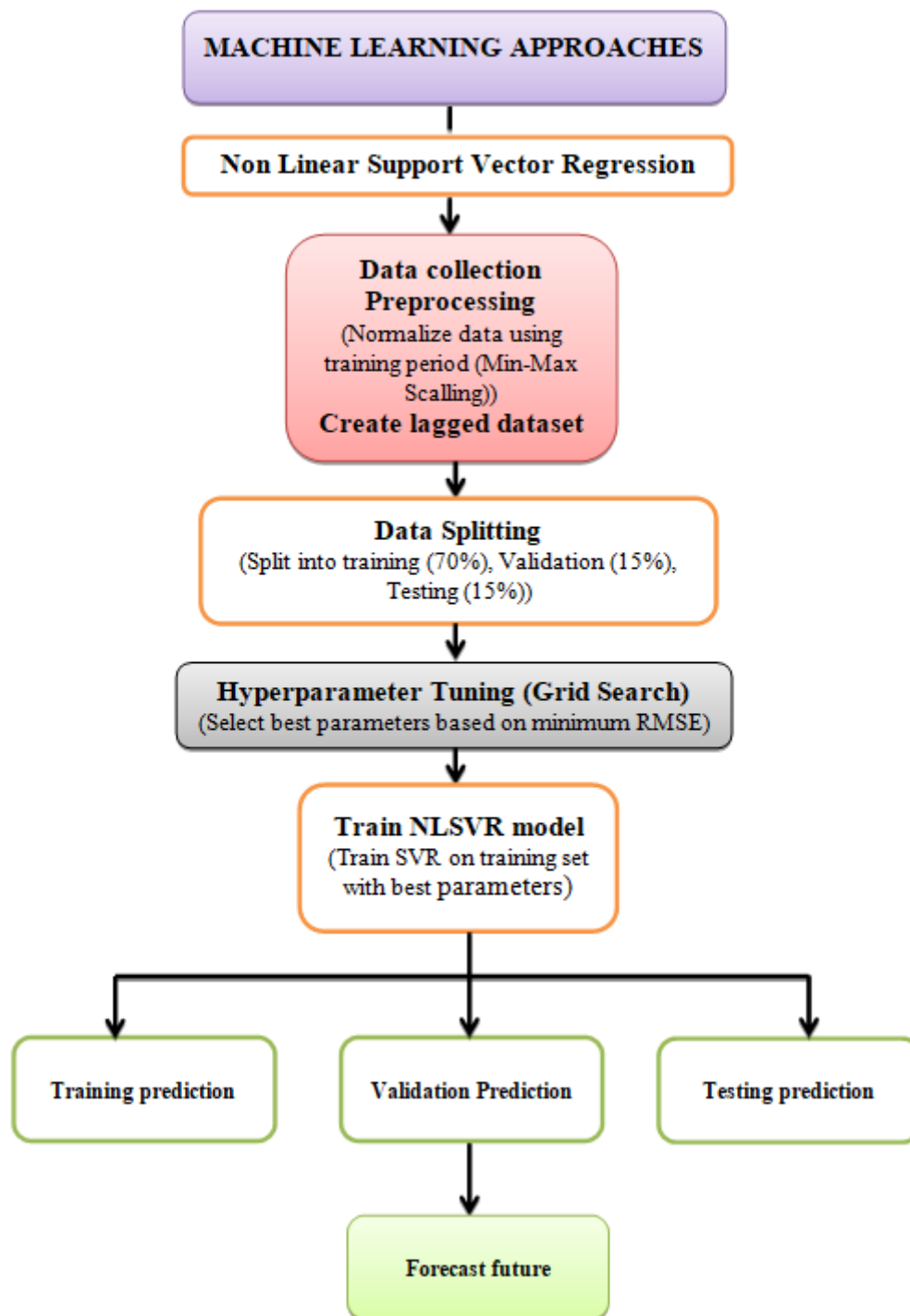


Figure 2: Standard workflow of the NLSVR-based machine learning approach

Train test validation split

The dataset was divided into three subsets: training, validation, and test datasets. The training set (70% of the data) was used to develop and fit the predictive model by learning underlying patterns from the data. Model performance was subsequently evaluated using the validation set (15% of the data), which facilitated hyperparameter tuning and optimization through iterative adjustments. Based on validation results, the best-performing model was selected for further analysis. Finally, the selected model was evaluated on the independent test set (15% of the data), which was not involved in model development, to obtain an unbiased estimate of predictive accuracy, robustness, and generalization capability on unseen data.

Model adequacy

To evaluate model performance, RMSE and MAE were computed for all forecasting models.. Lower values of these measures indicated better model performance and prediction accuracy. The RMSE and MAE is given by the formulas:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (Y_t - \hat{Y}_t)^2}{(n-2)}}$$

where, Y_t is the Actual value, $\hat{Y}_t$ is the Predicted value and n is the total number of observations

$$MAE = \frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t|$$

where, n is the number of error variables and $|Y_t - \hat{Y}_t|$ is the absolute error.

3. Results and discussion

The descriptive statistics of pearl millet production during 1966-2024 revealed a mean production of 8.15 thousand tonnes, with values ranging from 3.70 to 13.59 thousand tonnes. The standard deviation (2.22 thousand tonnes) and coefficient of variation (27.27%) indicated moderate variability in production over the study period. The skewness (0.16) and kurtosis (2.31) values suggested an approximately symmetric and slightly platykurtic distribution. The Durbin–Watson statistic (DW) is 1.10 indicated the presence of positive autocorrelation in the series, justifying the application of time-series models for capturing temporal dependence and generating reliable forecasts.

Table 1: Descriptive statistics of pearl millet production from 1966-2024

Variable	N	Mean	Max	Min	SD	CV (%)	Skewness	Kurtosis	Durbin–Watson test
Production (000tonnes)	59	8.15	13.59	3.7	2.22	27.27	0.16	2.31	1.1

As shown in figure 3, pearl millet production did not remain constant from 1966 to 2024. Production increased in some years and decreased in others, showing noticeable fluctuations over time. These ups and downs indicate that production is influenced by various

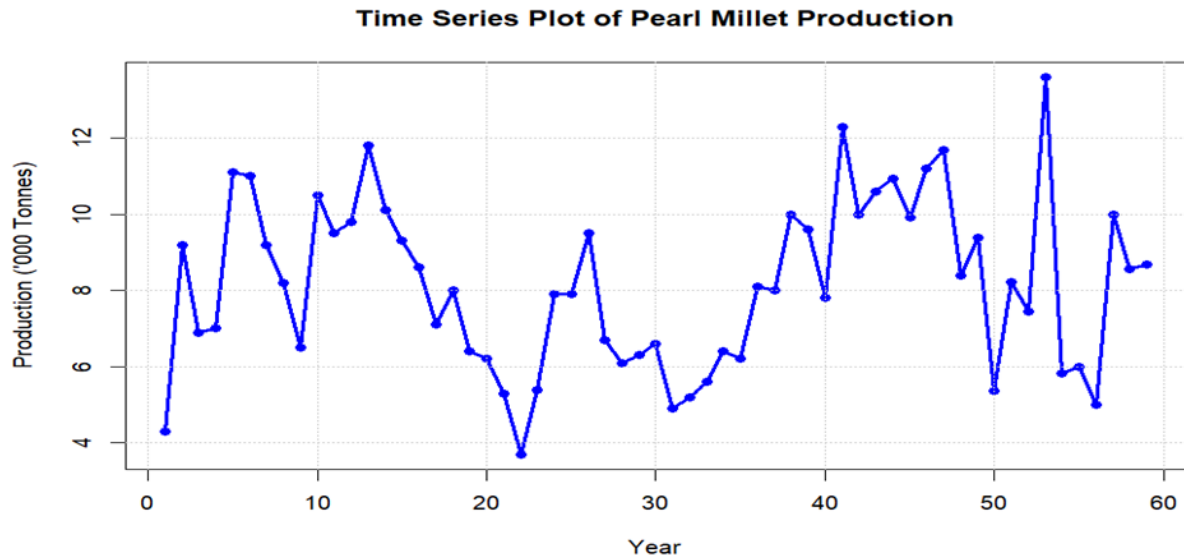


Figure 3: Graphical representation showing fluctuations in data of pearl millet crop for the period 1966-2024

factors and follows a time-dependent pattern, making it suitable for forecasting using time-series models.

Table 2: Simple exponential smoothing and double exponential smoothing parameters

Models	α	β
Simple exponential smoothing	0.5	-
Double exponential smoothing	0.1	0.6

Simple Exponential Smoothing (SES) and Double Exponential Smoothing (DES) models were employed to forecast pearl millet production. The smoothing parameters were optimized through a grid-search procedure using the training and validation datasets. The optimal parameter for SES was $\alpha = 0.5$, indicating that equal weight was assigned to recent and past observations. For DES, the optimal values were $\alpha = 0.1$ and $\beta = 0.6$, suggesting that the model emphasized long-term historical information while effectively capturing trend variations in the data. These parameter values were selected based on the minimum RMSE and MAE.

As shown in figure 4, the ACF plot exhibits a gradual decline in autocorrelation across lags, indicating the presence of temporal dependence in pearl millet production data. The PACF plot shows significant spikes at the initial lags, suggesting that autoregressive components may be present and that the series is suitable for time-series modelling.

The ADF test results in Table 3 showed that the original pearl millet production series was non-stationary, whereas the first-differenced series became stationary.

As shown in Table 4, four ARIMA models were evaluated using RMSE and MAE

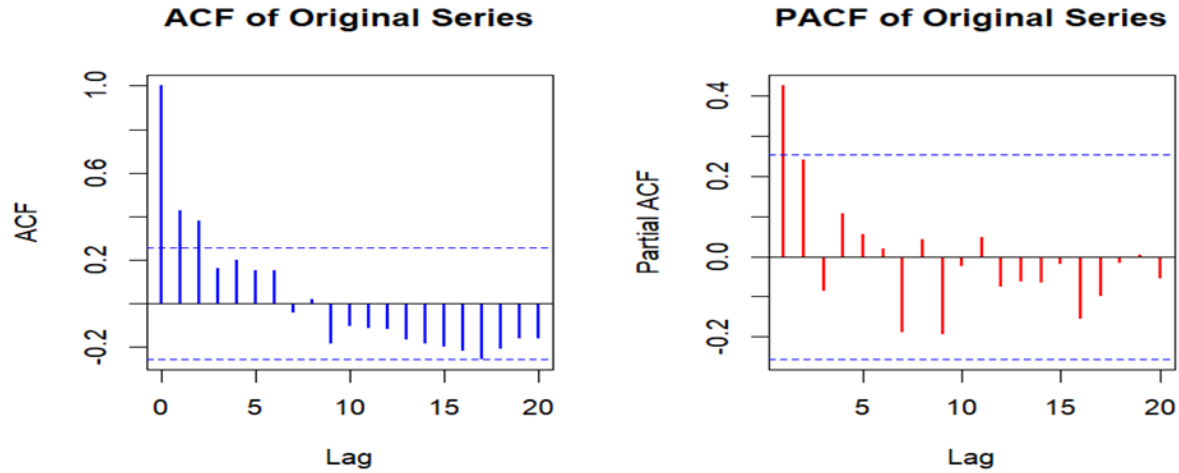


Figure 4: Graphical representation of ACF and PACF of original series for production in J&K

Table 3: Stationarity test results for pearl millet production under J&K through ADF test

	Original series			First differenced series		
	Statistics	Lag	p-value	Statistics	Lag	p-value
ADF Test	-2.03	3	0.55	-4.04	3	0.01

values across training, validation, and testing datasets. Among the candidate models, ARIMA(1,1,1) exhibited the lowest validation RMSE (1.15) and MAE (0.91), indicating superior predictive performance. Therefore, ARIMA(1,1,1) was selected as the most suitable model for forecasting pearl millet production in Jammu and Kashmir.

The parameter estimates and diagnostic statistics reported in Table 5 demonstrated that the model parameters were significant and that the residuals behaved as white noise. Therefore, ARIMA(1,1,1) was identified as the most suitable model for forecasting pearl millet production in Jammu and Kashmir.

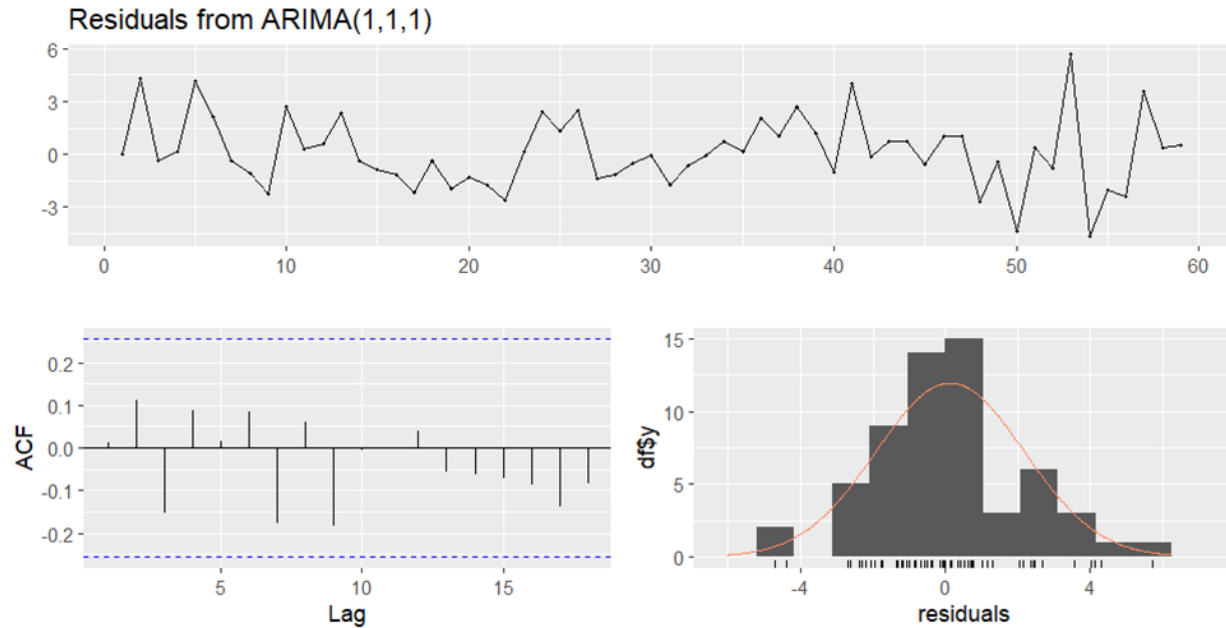
As shown in figure 5, the residuals of the ARIMA(1,1,1) model were randomly scat-

Table 4: ARIMA models for pearl millet production under J&K during the study period

Models	Training		Validation		Testing	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
ARIMA(0,1,1)	1.82	1.43	1.38	1.07	4.15	3.83
ARIMA(1,1,0)	1.81	1.42	1.18	0.97	3.91	3.57
ARIMA(1,1,1)	1.8	1.41	1.15	0.91	2.96	2.56
ARIMA(1,1,2)	1.81	1.49	1.19	0.91	3.93	3.56
ARIMA(2,1,2)	1.5	1.14	2.11	1.89	4.98	4.58
ARIMA(2,1,1)	1.75	1.42	1.2	0.98	4	3.64

Table 5: Estimates of parameters of ARIMA (1, 1, 1) for production

Model	Ljung-Box test		Parameters	Estimates	SE	p-value
ARIMA(1,1,1)	Statistics	p-value	AR1	0.12	0.61	0.00**
	6.86	0.73	MA1	-0.46	0.56	0.01**

**Figure 5: Residual time series plot of ARIMA (1, 1, 1) for production in J&K**

tered around zero with no visible trend or systematic pattern. Although the ARIMA model adequately captured the linear pattern in the data, statistical models are developed under certain assumptions, particularly that the residuals should not contain any systematic information. Therefore, the nonlinearity of the residuals needs to be tested, which can be done using the Brock–Dechert–Scheinkman (BDS) test.

Table 6: Results of the BDS test for testing non-linearity

	Epsilon	Embedding dimension=2		Embedding dimension=3	
		Statistics	p-value	Statistics	p-value
BDS Test	0.89	2.44	0.01	-0.08	0.93
	1.78	-1.22	0.22	-2.22	0.02
	2.68	-2.63	0	-2.74	0

The BDS test was applied to the residuals of the fitted ARIMA model to examine the presence of nonlinear dependence in the pearl millet production series. As presented in Table 6, the BDS test revealed significant statistics ($p < 0.05$) for most epsilon values under embedding dimensions 2 and 3, indicating the presence of nonlinear dependence in the ARIMA residuals. These results suggest that the ARIMA model alone could not completely capture the underlying data structure, thereby justifying the use of machine learning models to capture the remaining nonlinear patterns. Therefore, machine learning models were also used, and their performance was compared with the statistical models at a later stage. The

first machine learning model employed was Nonlinear Support Vector Regression (NLSVR). For model development, the pearl millet production series was transformed into a supervised learning dataset using five lagged observations as input variables and divided into training (70%), validation (15%), and testing (15%) datasets. The lagged dataset, spanning from 1971 to 2024, was divided chronologically into training (1971 - 2006), validation (2007-2015), and testing (2016-2024) subsets to preserve temporal order and ensure a realistic evaluation of forecasting performance on future unseen data. To avoid information leakage, the data were normalized using the minimum and maximum values of the training dataset only.

Table 7: Optimal hyper-parameter values of the fitted NLSVR models

Specifications	NLSVR	ARIMA(1,1,1) residual series
No. of input lags	5	5
C	100	1
γ	0.05	0.01
ϵ	0.001	0.001

Before fitting the NLSVR model, the optimal hyperparameters were identified through a grid search procedure based on minimum validation RMSE. This NLSVR model was developed using five lagged observations, and a Radial Basis Function (RBF) kernel was used, as shown in Table 7. The optimal hyperparameters ($C = 100$, $\gamma = 0.05$, $\epsilon = 0.001$) were selected through grid search based on the minimum validation RMSE, enabling the model to effectively capture nonlinear patterns in pearl millet production data. Similarly, the Hybrid ARIMA(1,1,1)-SVR model, the residuals obtained from the fitted ARIMA(1,1,1) model were first transformed into a supervised learning dataset by considering five lagged residual observations as input variables. The dataset was subsequently divided into training (70%), validation (15%), and testing (15%) subsets, and normalization was performed using the training dataset to avoid information leakage. Before fitting the Hybrid model, the optimal hyperparameters were identified through a grid search procedure based on minimum validation RMSE. The optimal parameter combination ($C = 1$, $\gamma = 0.01$, $\epsilon = 0.001$) was identified through grid search, allowing the model to capture residual nonlinearities and enhance forecasting accuracy. Similar findings were reported by Rathod *et al.* (2018), who developed an NLSVR model for oilseed production using an RBF kernel and determined the optimal values of $C = 8.19$, $\gamma = 3.06$, and $\epsilon = 0.15$ through cross-validation. Their study demonstrated that systematic hyperparameter optimization considerably improved forecasting accuracy. Likewise, Pandit *et al.* (2023) reported that tuning the SVR hyperparameters before modelling the ARIMA residuals enhanced the predictive capability of hybrid forecasting models by effectively capturing nonlinear components that remained unexplained by the linear ARIMA model. Thus, the findings of the present investigation further confirm that careful hyperparameter optimization is an essential step for obtaining reliable forecasts using NLSVR and hybrid ARIMA-SVR models.

After developing the statistical and machine learning models, their forecasting performance was first compared using the Diebold–Mariano (DM) test based on the actual and predicted values. As presented in Table 8, the Diebold–Mariano test indicates that there is no statistically significant difference in forecasting accuracy among ARIMA, NLSVR, and Hybrid ARIMA–SVR models (all p-values > 0.05). Since the DM test did not identify a statistically superior model, the final model selection was based on forecast accuracy measures.

Table 8: Diebold–Mariano test results based on testing actual and predicted values

Model	DM Statistics	p-value	Accuracy Differences Between Models
Hybrid vs ARIMA	-0.5	0.31	Not Significant
Hybrid vs NLSVR	-0.73	0.23	Not Significant
NLSVR vs ARIMA	0.52	0.69	Not Significant

Table 9: Comparison of different time series models on the basis of RMSE and MAE

Models	Training			Validation			Testing		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
SES	1.85	1.47	18.38	1	0.84	8.55	3.01	2.6	39.76
DES	6.37	4.86	64.9	1.37	1.29	12.83	4.94	4.62	70.8
ARIMA(1,1,1)	1.8	1.41	17.96	1.15	0.91	9.65	2.96	2.56	38.94
NLSVR	0.09	0.03	4.07	0.19	1.37	11.41	0.42	0.31	36.22
ARIMA(1,1,1)-SVR	1.56	1.22	15.98	1.13	0.9	9.48	2.89	2.49	37.52

Accordingly, the performance of the SES, DES, ARIMA(1,1,1), NLSVR, and Hybrid ARIMA(1,1,1)-SVR models were evaluated using the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) across the training, validation, and testing datasets, as presented in Table 9. Among the models evaluated, the NLSVR model recorded the lowest errors across training, validation, and testing datasets, indicating the highest forecasting accuracy. The Hybrid ARIMA (1,1,1)-SVR model improved upon ARIMA, while DES showed the poorest performance. Therefore, based on the minimum RMSE and MAE values obtained on the independent testing dataset, the NLSVR model was identified as the most appropriate model for forecasting pearl millet production.

Table 10: Forecasted pearl millet production using the NLSVR and SES model

Year	NLSVR	Lower Confidence Limit (95%)	Upper Confidence Limit (95%)
2026 – 2027	9.27	7.57	10.96
2027 – 2028	9.95	8.25	11.65
2028 – 2029	9.49	7.79	11.19
2029 – 2030	9.19	7.49	10.89
2030 – 2031	9.28	7.58	10.98

The NLSVR model predicts slight fluctuations in pearl millet production during 2026–2030, with forecasted values ranging from 9.19 to 9.95 thousand tonnes. Overall, production is expected to remain relatively stable, showing no significant long-term increasing or decreasing trend as presented in Table 10. Moreover, it is observed that the forecasted values through the proposed model, *i.e.*, NLSVR, with respect to the future are varying. This variation is also observed through the periodogram of the spectral analysis, as well as the ACF spikes, which show a positive trend for the initial lags and followed by negative values for several lags as shown in Figure 4.

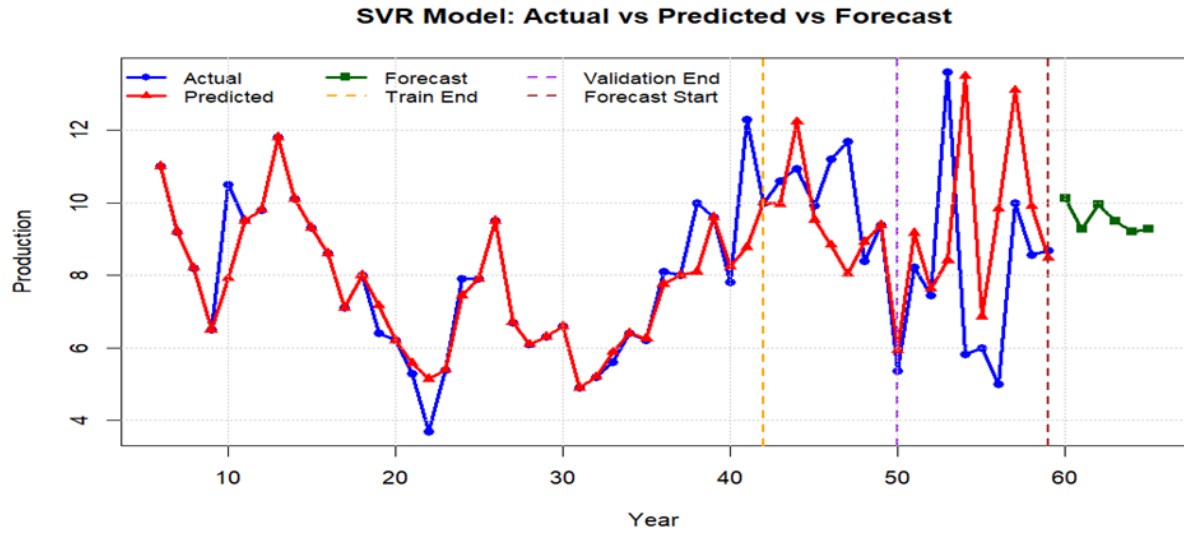


Figure 6: Graphical representation of forecasted value of pearl millet production through NLSVR

As illustrated in figure 6, the NLSVR model showed satisfactory predictive performance, with predicted values closely aligning with observed production during the training and validation phases. Despite minor deviations in the testing period, the model effectively captured the overall temporal trend and indicated relatively stable future production.

Table 11: Actual and predicted pearl millet production in the testing dataset and forecast for 2025

Year	Actual	NLSVR	Hybrid ARIMA-SVR	ARIMA	SES	DES
2015	5.36	5.94	9.07	9.75	9.59	11.59
2016	8.21	9.16	9.41	7.8	9.59	11.71
2017	7.44	7.62	9.81	8.25	9.59	11.84
2018	13.59	8.4	9.28	7.88	9.59	11.97
2019	5.82	13.48	9.43	10.51	9.59	12.1
2020	6	6.85	9.34	8.06	9.59	12.23
2021	5	9.82	9.32	7.36	9.59	12.36
2022	10	13.09	9.41	6.44	9.59	12.48
2023	8.56	9.89	9.37	8.2	9.59	12.61
2024	8.69	8.46	9.39	8.17	9.59	12.74
2025	-	10.13	9.38	8.36	8.49	5.98
χ^2	Statistics		df	Critical value (0.05, 53 df)		p-value
(value for NLSVR)	16.92		53	70.99		0.99

Table 11 presents the actual and predicted pearl millet production values in the testing dataset obtained using the different models, along with the forecast for 2025. A comparison of the predicted values with the observed production indicated that the NLSVR model exhibited the least deviation from the actual production values in the testing dataset, particularly for the most recent observation (2024), showed better predictive performance. Since the NLSVR

model demonstrated the highest forecasting accuracy based on the lowest RMSE, MAE and MAPE values on the testing dataset, it was identified as the best-performing model. To further assess the adequacy of the best-performing model, a Chi-square goodness-of-fit test was conducted for the NLSVR model. The null hypothesis (H_0) stated that there was no significant difference between the observed and NLSVR-predicted pearl millet production values, whereas the alternative hypothesis (H_1) stated that a significant difference existed between them. The test yielded a χ^2 statistic of 16.92 with 53 degrees of freedom. Since the calculated χ^2 value was substantially lower than the critical value (70.99) at the 5 per cent level of significance, and the corresponding p-value (0.99) exceeded 0.05, the null hypothesis was not rejected. These results indicate that there was no statistically significant difference between the observed and NLSVR-predicted production values. Therefore, the NLSVR model provided an adequate fit to the historical pearl millet production data and may be considered a reliable model for forecasting future production.

4. Conclusion

The present study compared the forecasting performance of statistical, machine learning, and hybrid time series models for pearl millet production in the Jammu division. The forecasting models were compared based on their predictive performance using forecast accuracy measures, including RMSE and MAE. Among the evaluated models, the NLSVR model achieved the highest forecasting accuracy and demonstrated an adequate fit to the historical data. The model forecasts indicate that pearl millet production is likely to remain relatively stable during 2026–2030, ranging from 9.19 to 9.95 thousand tonnes. Therefore, the NLSVR model can be considered a reliable tool for production forecasting and can support informed agricultural planning and policy decisions.

Acknowledgements

The authors are grateful to the Reviewer for the valuable comments on the original version of this manuscript.

Conflict of interest

The authors do not have any financial or non-financial conflict of interest to declare for the research work included in this article.

References

- Box, G. E. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*.
- Department of Agriculture & Farmers Welfare. (2023). *Statewise Millet Production*. Ministry of Agriculture & Farmers Welfare, Government of India.
- Directorate of Economics and Statistics, Jammu and Kashmir. (2025). *Digest of Statistics*. <https://ecostatjk.nic.in/>
- Directorate of Millets Development. (2024). *Indian Millets Production Statistics*. Ministry of Agriculture and Farmers Welfare, Government of India. <https://www.milletstats.com/>
- Government of India. *Economic Survey 2024–25*. New Delhi: Ministry of Finance, Department of Economic Affairs; 2025.

- Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages (onr memorandum no. 52). *International Journal of Forecasting*, **20**, 5–10.
- Holt, C. C. (1957). *Forecasting Trends and Seasonals by Exponentially Weighted Averages*. Carnegie Institute of Technology, Pittsburgh Office of Naval Research memorandum no. 52.
- ICAR-All India Coordinated Research Project on Pearl Millet. *About Pearl Millet*. Jodhpur (Rajasthan), India.
- Jasrotia, N., Sharma, M., Bhat, A., Mahajan, S., and Gupta, S. (2024). Forecasting models for the production of walnut in Jammu and Kashmir: A comparative study. In *Special Proceedings of the 26th Annual Conference of the Society of Statistics, Computer and Applications* (pp. 51–60). Society of Statistics, Computer and Applications. ISBN 978-81-950383-5-0.
- Mahajan, S., Sharma, M., and Gupta, A. (2020). ARIMA modelling for forecasting of rice production: A case study of India. *Agricultural Science Digest-A Research Journal*, **40**, 404–407.
- Mehmood, T. (2025). Agriculture as the backbone of India's economy: An analysis of contribution, barriers and policy reforms. *International Journal of Finance, Management and Economics*, **8**, 1134–1140. <https://doi.org/10.33545/26179210.2025.v8.i2.660>
- Pal, M., Deka, J., and Rai, R. K. (1996). *Fundamentals of Cereals Crop Production*. Tata McGraw Hill Publishing Co. Ltd., New Delhi.
- Pandit, P., Sagar, A., Ghose, B., Dey, P., Paul, M., Alqadhi, S., and Abdo, H. G. (2023). Hybrid time series models with exogenous variable for improved yield forecasting of major Rabi crops in India. *Scientific Reports*, **13**, 22240.
- Rani, R., Khandelwal, V., Jhunjhadia, S., Ambawat, S., Bhanwariya, S., Singh, A., and Kumar, V. (2024). Quantifying genetic diversity based on morphological and molecular analysis in pearl millet (*Pennisetum glaucum* [L.] R. Br.) inbred lines. *Annals of Arid Zone*, **63**, 105–110.
- Rathod, S., Singh, K. N., Patil, S. G., Naik, R. H., Ray, M., and Meena, V. S. (2018). Modeling and forecasting of oilseed production of India through artificial intelligence techniques. *Indian Journal of Agricultural Sciences*, **88**, 22–27.
- Reddy, A., Yadav, O. P., and Singh, I. P. (2013). Prospects for kharif (rainy season) and summer pearl millet in Western India. Working Paper, International Crops Research Institute for Semi-Arid Tropics, Andhra Pradesh, India.
- Risnawati, I., Afendi, F. M., and Sumertajaya, I. M. (2026). Performance evaluation of seasonal ARIMA-SVR and seasonal ARIMAX-SVR hybrid methods on forecasting paddy production. *BAREKENG: Journal of Mathematics and Its Applications*, **20**, 367–380.
- Sharma, D. (2024). *Forecasting Models for Production of Pearl Millet (Bajra) and its Scenario in Jammu Region*. [Unpublished M.Sc. Thesis, SKUAST JAMMU].
- Sharma, M., Jasrotia, N., Kumar, B., Bhat, A., and Mahajan, S. (2018). Modeling of monthly arrival of rohu fish using ARIMA in Jammu region of J&K state. *Journal of Animal Research*, **8**, 259–262.