International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0008
e ISSN: 2584-2854
Volume: 03
Issue: 01 January 2025
Page No: 35-39

Comparing MLR and SVR in Evaluating the Impacts of Climate Change on Sugarcane Production in Saharanpur District

Akanksha Sharma[1], Dr. Charu Saraf [2]
[1]Research Scholar, Department of Botany, Shri Venkateshwar University, Gajraula, Uttar Pradesh, India.
[2]Assistant Professor, School of Applied Sciences, Shri Venkateshwar University, Gajraula, Uttar Pradesh, India.
Email ID: *sharmasharmasa89@gmail.com[1], charuagg.9923@gmail.com[2]*

Abstract
Climate change poses significant challenges to agriculture, with impacts on crop yield and food security. In Saharanpur District, Uttar Pradesh, sugarcane—a key economic crop—is particularly vulnerable to variations in climatic factors such as temperature, rainfall, and humidity. This study evaluates the predictive capabilities of Multiple Linear Regression (MLR) and Support Vector Regression (SVR) models in assessing the impacts of climate variability on sugarcane production. The analysis incorporates sugarcane yield data and seven climatic variables: maximum and minimum temperatures, annual rainfall, humidity, solar radiation, wind speed, and soil moisture. The study compares the performance of MLR and SVR based on metrics such as R^2 and Mean Squared Error (MSE), using 10-fold cross-validation for robustness. Results indicate that while MLR captures linear relationships effectively, SVR, with its ability to model non-linear interactions, outperforms MLR, achieving an R^2 of 0.87 on testing data compared to 0.74 for MLR. These findings highlight SVR's superior predictive accuracy and its potential as a robust tool for agricultural planning in climate-sensitive regions. The study underscores the importance of adopting advanced modeling techniques to mitigate the risks posed by climate variability on critical crops like sugarcane.
Keywords: *Multiple Linear Regression (MLR) and Support Vector Regression (SVR)*

1. Introduction
Climate change has emerged as a critical challenge for global agriculture, significantly affecting crop production and food security [1]. In regions like Saharanpur, where sugarcane is a key economic crop, variations in temperature, rainfall, and other climatic factors can disrupt yield patterns, posing risks to farmers and local economies [2]. Understanding these impacts and developing predictive models is essential for implementing adaptive strategies to mitigate climate-induced risks [3]. Traditional modeling approaches like Multiple Linear Regression (MLR) have been widely used to analyze the linear relationships between climatic variables and crop yields [4]; however, the complexity of climate dynamics often necessitates advanced techniques like Support Vector Regression (SVR) to capture non-linear interactions effectively [5]. This study compares the efficacy of MLR and SVR in evaluating the impacts of climate variability on sugarcane production, providing insights into the suitability of these methods for agricultural planning in climate-sensitive regions.

2. Literature Review
The impact of climate change on agriculture has been extensively studied, with significant emphasis on understanding its effects on crop yields and developing predictive models [6]. Sugarcane, a climate-sensitive crop, is particularly vulnerable to variations in temperature, rainfall, and humidity, which can drastically influence its growth and yield [7]. Traditional statistical methods such as Multiple Linear Regression (MLR) have been used to model the relationship between climatic variables and agricultural outputs due to their simplicity and interpretability [8]. However, MLR often struggles to account for the non-linear and complex interactions inherent in climate-agriculture dynamics [9]. Recent advancements in machine learning, such as Support Vector Regression (SVR), offer robust alternatives by effectively modeling non-linear relationships and

handling high-dimensional data [10]. Studies have shown that SVR outperforms linear models in predicting crop yields under varying climatic conditions, making it a promising tool for agricultural planning [11], [12]. This growing body of research underscores the need for comparative studies to identify the most effective modeling techniques for specific crops and regions, such as sugarcane production in Saharanpur District.

3. Research Methodology

3.1. Variables of The Study

The dataset for this study includes sugarcane yield data along with several climate variables specific to the Meerut District. The dependent variable is Sugarcane Yield (SY), measured in quintals. The independent variables considered are key climate factors: Maximum Temperature (MAT) in degrees Celsius, Minimum Temperature (MIT) in degrees Celsius, Annual Rainfall (AR) in centimeters, Humidity (HU) in percentage, Solar Radiation (SR) in megajoules per square meter per day, Wind Speed (WS) in meters per second, and Soil Moisture (SM) in percentage.

3.2. Model Development

The dataset included sugarcane yield as the dependent variable and seven climatic factors (maximum temperature, minimum temperature, annual rainfall, humidity, solar radiation, wind speed, and soil moisture) as independent variables, divided into 85% training and 15% testing sets. Multiple Linear Regression (MLR) was used to model linear relationships, with Ridge and Lasso regressions explored to reduce overfitting. Support Vector Regression (SVR) employed an RBF kernel to capture non-linear relationships, with hyperparameters such as regularization, epsilon, and kernel type optimized using grid search. Both models were evaluated using Mean Squared Error (MSE) and R2, with 10-fold cross-validation ensuring robust performance. The comparison highlighted the strengths and suitability of each model for predicting sugarcane yield under varying climatic conditions.

4. Results and Discussion

4.1. Descriptive Statistics

Table 1 presents the descriptive statistics for the climatic variables and sugarcane yield in Saharanpur District, based on 500 observations. The mean maximum temperature (MAT) was 30.21°C with a standard deviation of 3.65°C, indicating moderate variability. Minimum temperature (MIT) averaged 15.47°C with a slightly higher standard deviation of 4.40°C. Annual rainfall (AR) showed significant variability, with a mean of 115.36 cm and a standard deviation of 41.80 cm, reflecting the region's diverse precipitation patterns. Humidity (HU) had a mean value of 60.77% and a standard deviation of 12.47%, while solar radiation (SR) averaged 15.17 MJ/m²/day. Wind speed (WS) and soil moisture (SM) had mean values of 2.52 m/s and 14.83%, respectively, with standard deviations indicating moderate variation. The skewness and kurtosis values for all variables were close to zero, suggesting that the data distributions were approximately normal. The sugarcane yield (SY) had a mean of 2751.76 quintals with a standard deviation of 300.91 quintals, indicating consistent production levels across the dataset.

Table 1 Descriptive Statistics of Climate Variables in Saharanpur District (N=500)

	MAT	MIT	AR	HU	SR	WS	SM	SY
Mean	30.21	15.47	115.36	60.77	15.17	2.52	14.83	2751.76
Std. Dev.	3.65	4.40	41.80	12.47	3.76	1.14	6.02	300.91
Kurtosis	-0.70	-0.53	-1.14	-1.07	-0.47	-1.14	-1.21	-0.25
Skewness	-0.10	-0.12	-0.03	-0.06	-0.09	0.02	0.02	0.06
Minimum	20.87	3.86	32.52	32.79	3.78	0.41	3.95	1939.47
Maximum	39.71	27.61	192.59	85.71	24.71	4.61	26.95	3699.57

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0008

e ISSN: 2584-2854
Volume: 03
Issue: 01 January 2025
Page No: 35-39

4.2. Multiple-Linear Regression (MLR)

The Multiple Linear Regression model was developed to evaluate the linear relationships between the climatic variables and sugarcane yield. Table 2 outlines the hyperparameters considered and the optimized values determined through model tuning. A standard linear regression was selected over Ridge and Lasso regressions, as it provided the best fit without overfitting.

Table 2 **MLR Model Parameters**

Algorithm	Hyper parameter	Ranges	Best Optimized Value
Multiple Linear Regression (MLR)	Type of Regression	[Linear, Ridge, Lasso]	Linear
	Alpha	[0.01, 0.1, 1, 10, 100]	-
	Random State	[1,2,3….100]	46
	Test Size	[0.15, 0.2….0.3]	0.15

The performance metrics for the MLR model are presented in Table 3. The model achieved an average accuracy of 77.20% ± 6.43% based on 10-fold cross-validation. The training phase resulted in an R2 value of 0.80, indicating that 80% of the variability in sugarcane yield was explained by the model. In the testing phase, the R2 value slightly decreased to 0.74, demonstrating good generalization to unseen data.

Table 3 **MLR Performance Metrics**

Accuracy (%) (CV=10)	Training		Testing		Equation
	MSE	R^2	MSE	R^2	
77.20 ± 6.43	0.00	0.80	0.00	0.74	((0.2*MAT) + (-0.06*MIT) + (-0.14*AR) + (0.06*HU) + (0.51*SR) + (-0.19*WS) + (0.37*SM))

Figure 1 compares the predicted versus actual sugarcane yields for both the training and testing datasets. The plots illustrate that the predicted values closely align with the actual values, particularly in the training dataset, indicating a strong fit. In the testing dataset, the predictions remain consistent, although with a slight dispersion, which is expected when applying the model to new data.

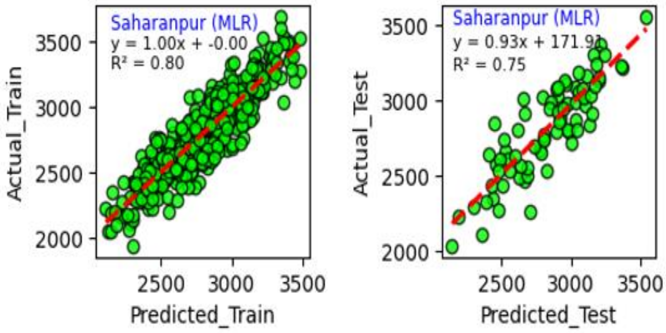

Figure 1 **MLR Model Training & Testing - Comparison of Predicted vs. Actual Sugarcane Yield**

4.3. Support Vector Regression (SVR)

Support Vector Regression (SVR) was employed to model the non-linear relationships between climatic factors and sugarcane yield. Table 4 details the hyperparameters optimized during model development. The SVR model was tuned to achieve the best balance between accuracy and computational efficiency. A linear kernel was selected over the RBF kernel for simplicity and comparable performance, with the regularization parameter C set at 10 and epsilon (ϵ\epsilonϵ) at 0.01. The random state and test size were consistent with those used in the MLR model to ensure comparability.

Table 4 **SVR Model Parameters**

Algorithm	Hyper parameter	Ranges	Best Optimized Values
Support Vector Regression (SVR)	C	[10,20,30…100]	10
	Epsilon	[0.01, 0.001, 0.0001]	0.01
	Kernel	[linear, rbf]	Linear
	Random State	[1,2,3….100]	46
	Test Size	[0.15, 0.2….0.3]	0.15

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0008
e ISSN: 2584-2854
Volume: 03
Issue: 01 January 2025
Page No: 35-39

Table 5 summarizes the performance of the SVR model, demonstrating its robustness and high accuracy in predicting sugarcane yields. For Saharanpur District, the SVR model achieved an R2 value of 0.78 for the training set and 0.87 for the testing set, outperforming the MLR model in both phases. The Mean Squared Error (MSE) values were consistently low, reinforcing the model's reliability.

Table 5 SVR Performance Metrics

Support Vector Regression (SVR)					
Site	Accuracy (%) (CV=10)	Training		Testing	
		MSE	R^2	MSE	R^2
Saharanpur	77.10 ± 3.85	0.00	0.78	0.00	0.87

Figure 2 illustrates the predicted versus actual sugarcane yields for the training and testing datasets in Saharanpur District. The plots highlight the SVR model's strong predictive performance, with the predicted values closely aligning with the actual values. The testing dataset plot shows an R2 of 0.87, indicating the model's excellent ability to generalize to unseen data. The minimal dispersion of points around the regression line underscores SVR's robustness in capturing complex interactions between climatic variables and crop yield.

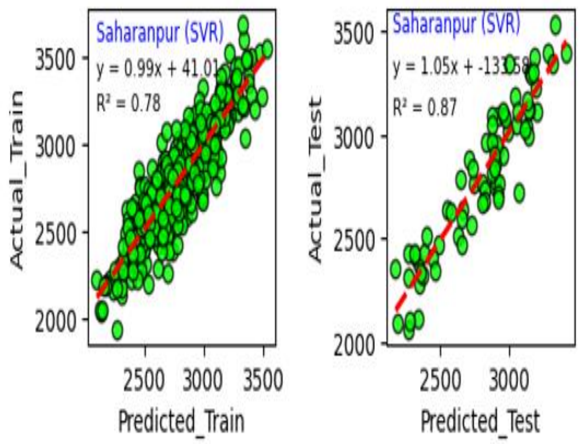

Figure 2 SVR Model Training & Testing - Comparison of Predicted vs. Actual Sugarcane Yield

Conclusion

This study assessed the impacts of climate variability on sugarcane production in Saharanpur District by comparing the performance of Multiple Linear Regression (MLR) and Support Vector Regression (SVR) models. The findings revealed that while MLR effectively captured linear relationships between climatic variables and sugarcane yield, its predictive accuracy was limited by its inability to account for complex, non-linear interactions inherent in climate dynamics. In contrast, SVR demonstrated superior performance, achieving an R^2 of 0.87 on the testing dataset, compared to 0.74 for MLR. The results underscore the importance of adopting advanced machine learning techniques like SVR for agricultural modeling in climate-sensitive regions. The SVR model's ability to capture non-linear relationships makes it a robust tool for predicting sugarcane yields and can aid in designing adaptive strategies to mitigate the risks posed by climate variability. These findings have significant implications for agricultural planning and policymaking. By integrating advanced predictive models, stakeholders can make informed decisions to enhance crop resilience and ensure sustainable agricultural practices. Future research should explore the application of these models to other crops and regions, incorporate additional climatic and soil parameters, and leverage larger datasets to further enhance model robustness and generalizability.

References

[1]. A. N. N. Technique, "Forecasting Sugar Cane Yield in the Eastern Area of Thailand with ANN Technique," Aust. J. Basic Appl. Sci., vol. 13, no. 1, pp. 112–116, 2019, doi: 10.22587/ajbas.2019.13.1.14.

[2]. T. Wang et al., "Applications of UAS in Crop Biomass Monitoring: A Review," Front. Plant Sci., vol. 12, no. April, pp. 1–16, 2021, doi: 10.3389/fpls.2021.616689.

[3]. H. You, M. Zhou, J. Zhang, W. Peng, and C. Sun, "Sugarcane nitrogen nutrition estimation with digital images and machine learning methods," Sci. Rep., vol. 13, no. 1, pp. 1–12, 2023, doi: 10.1038/s41598-023-42190-2.

[4]. M. K. Linnenluecke, N. Nucifora, and N.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0008
e ISSN: 2584-2854
Volume: 03
Issue: 01 January 2025
Page No: 35-39

[5]. Thompson, "Implications of climate change for the sugarcane industry," Wiley Interdiscip. Rev. Clim. Chang., vol. 9, no. 1, 2018, doi: 10.1002/wcc.498.
[6]. V. Venugopal, R. Shanmugam, and L. P. Kamalakkannan, "Heat-health vulnerabilities in the climate change context-comparing risk profiles between indoor and outdoor workers in developing country settings," Environ. Res. Lett., vol. 16, no. 8, 2021, doi: 10.1088/1748-9326/ac1469.
[7]. S. N. Khan et al., "County-level corn yield prediction using supervised machine learning," European Journal of Remote Sensing, vol. 56, no. 1. 2023. doi: 10.1080/22797254.2023.2253985.
[8]. S. Emami, H. Dehghanisanij, M. Achite, N. Al-Ansari, and N. T. T. Linh, "Application of ANFIS, ELM, and ANN models to assess water productivity indicators based on agronomic techniques in the Lake Urmia Basin," Appl. Water Sci., vol. 13, no. 2, pp. 1–15, 2023, doi: 10.1007/s13201-022-01851-9.
[9]. Y. Shahare et al., "A Comprehensive Analysis of Machine Learning-Based Assessment and Prediction of Soil Enzyme Activity," Agric., vol. 13, no. 7, 2023, doi: 10.3390/agriculture13071323.
[10]. M. G. Mallikarjuna, S. C. Nayaka, and T. Kaul, Next-Generation Plant Breeding Approaches for Stress Resilience in Cereal Crops. 2022. doi: 10.1007/978-981-19-1445-4.
[11]. L. Benos, A. C. Tagarakis, G. Dolias, R. Berruto, D. Kateris, and D. Bochtis, "Machine learning in agriculture: A comprehensive updated review," Sensors, vol. 21, no. 11, pp. 1–55, 2021, doi: 10.3390/s21113758.
[12]. H. Xu et al., "Machine learning approaches can reduce environmental data requirements for regional yield potential simulation," Eur. J. Agron., vol. 129, no. June, 2021, doi: 10.1016/j.eja.2021.126335.
[13]. M. N. Thimmegowda et al., "Weather-Based Statistical and Neural Network Tools for Forecasting Rice Yields in Major Growing Districts of Karnataka," Agronomy, vol. 13, no. 3, pp. 1–20, 2023, doi: 10.3390/agronomy13030704.