Environment and Ecology 43 (1A) : 210—221, January—March 2025
Article DOI: https://doi.org/10.60151/envec/NSRV4438
ISSN 0970-0420

Modelling and Forecasting of Sugarcane (*Saccharum officinarum* L.) Area, Production and Yield in Eastern Uttar Pradesh using ARIMA: A Data-Driven Approach to Sustainability

Harshit Mishra, Supriya, R.R. Kushwaha, Manish Kumar, Sandeep Gautam

Received 11 December 2024, Accepted 11 February 2025, Published on 27 February 2025

ABSTRACT

Sugarcane is a vital commercial crop in Eastern Uttar Pradesh, contributing significantly to both the agricultural and industrial sectors. This study utilizes the ARIMA model to forecast the area, production, and yield of sugarcane in the region, aiming to address the challenges posed by yield variability due to factors such as climate change and market fluctuations. Historical data from 1960 to 2022, covering 15 districts, was analyzed using time series methods. The ARIMA (p,d,q) model, with its autoregressive and moving average components, was employed to project trends up to 2035. The average area under sugarcane cultivation over this period was 343.46 thousand hectares, with a compound annual growth rate (CAGR) of 0.69%. The production reached an average of 12,633.05 thousand tonnes, while the yield averaged 44.81 tonnes per hectare. The ARIMA model demonstrated strong predictive capability, as evidenced by performance metrics such as RMSE and R^2, validating its suitability for agricultural forecasting. The findings highlight the need for data-driven planning to enhance sugarcane productivity and economic sustainability, providing valuable insights for policymakers, farmers, and industry stakeholders in mitigating risks and optimizing resources.

Keywords Area, ARIMA, Forecasting, Instability, Production, Sugarcane, Sustainability, Yield.

INTRODUCTION

Sugarcane serves as a critical agricultural commodity globally, with Brazil leading the production at an astounding 752.9 million tons in 2024. This immense output, predominantly sourced from the South-Central region, highlights dominance of Brazil in the sugarcane market. Following Brazil, India ranks second with a production of 405.4 million tons, underscoring its significant role in the global sugarcane industry. Thailand, holding the third position, contributes 131 million tons, while China follows closely with 110

Harshit Mishra[1], Supriya[2]*, R.R. Kushwaha[3], Sandeep Gautam[5]
[1,5]PhD Research Scholar, [2]Associate Professor, [3]Professor
Department of Agricultural Economics, Acharya Narendra Deva University of Agriculture and Technology, Kumarganj, Ayodhya (UP) 224229, India

Manish Kumar[4]
Assistant Professor, Department of Agricultural Statistics, Acharya Narendra Deva University of Agriculture and Technology, Kumarganj, Ayodhya (UP) 224229, India

Email: drsupriya@nduat.org
*Corresponding author

million tons (Mehmood *et al.* 2024). The combined production of these top four nations constitutes a substantial portion of the world's total sugarcane output. Additionally, the trend towards increased ethanol production from molasses in Brazil reflects changing market dynamics influenced by rising automobile ownership (Verma *et al.* 2021).

Sugarcane is one of the most important commercial crops in India, contributing significantly to the economy, especially in states like Uttar Pradesh, Maharashtra and Karnataka. Its cultivation not only provides raw material for the sugar industry but also supports other sectors through its by-products, such as molasses, bagasse and press mud, which are used in alcohol production, energy generation, and as fertilizers, respectively (Hooda *et al.* 2020, Kumar 2022). As of 2023, India stands as the second-largest sugarcane producer globally, next to Brazil, with an annual production of approximately 405 million tons from an area of around 5.1 million hectares (Tyagi *et al.* 2023). The state of Uttar Pradesh alone accounts for nearly 43% of the country's sugarcane production, demonstrating its pivotal role in this sector (Supriya *et al.* 2024).

In Eastern Uttar Pradesh, sugarcane is a critical crop due to favorable agro-climatic conditions, with districts like Gorakhpur, Deoria and Kushinagar leading in its cultivation. However, despite its importance, the yield in these regions has shown significant variability over the years, impacted by factors such as erratic rainfall patterns, soil fertility issues, and fluctuating market prices (Padhan 2012). According to data from the Directorate of Economics and Statistics, the average yield of sugarcane in Eastern Uttar Pradesh was around 70 tonnes per hectare in 2020, which is lower than the national average of 75 tonnes per hectare (Mishra and Thakur 2021). Such variations necessitate accurate forecasting and modelling to ensure better resource allocation, policy planning, and market stability for the benefit of farmers and industries alike.

The time series analysis, particularly using the Autoregressive Integrated Moving Average (ARIMA) model, has emerged as a best tool for forecasting agricultural output. The ARIMA model has been widely used in agricultural research to predict crop yields, production areas, and productivity (Saini *et al.* 2023). For example, a study by Sahu and Mishra (2014) applied ARIMA to forecast the production of irrigated crop such as maize, while Vishwajith *et al.* (2014) forecasted pulses production in India up to 2020 using a similar approach. For sugarcane, Yassen *et al.* (2005) utilized ARIMA to model and forecast yield in Pakistan, demonstrating the model's efficacy in predicting future trends based on historical data. In the context of Eastern Uttar Pradesh, modelling and forecasting sugarcane production, area and yield are crucial for ensuring food security, stabilizing farmer incomes, and supporting the overall agro-industrial economy of the region.

This study aims to apply the ARIMA model to predict the future behavior of sugarcane area, production, and yield in Eastern Uttar Pradesh. Accurate forecasts will help policymakers, farmers, and stakeholders make informed decisions regarding crop management, resource allocation and industrial planning, thereby enhancing the sustainability of sugarcane cultivation in this vital region.

MATERIALS AND METHODS

This study focused on predicting the area, production, and yield of sugarcane in Eastern Uttar Pradesh using historical data from selected districts (Fig. 1). A total of 15 districts, including Allahabad, Azamgarh, Bahraich, Ballia, Basti, Deoria, Faizabad, Ghazipur, Gonda, Gorakhpur, Jaunpur, Mirzapur, Pratapgarh, Sultanpur and Varanasi, were chosen for the analysis due to the availability of reliable and consistent data. The remaining 13 districts were excluded from the study due to data unavailability, and these districts were Ambedkar Nagar, Amethi, Balrampur, Chandauli, Kaushambi, Kushinagar, Maharajganj, Mau, Sant Kabir Nagar, Sant Ravidas Nagar, Shravasti, Siddharth Nagar and Sonbhadra.

The dataset, which spanned from 1960 to 2022 (Table 1), was gathered from credible government sources such as the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Sankhyakiya Patrika, and the Directorate of Economics & Statistics. This data underwent descriptive

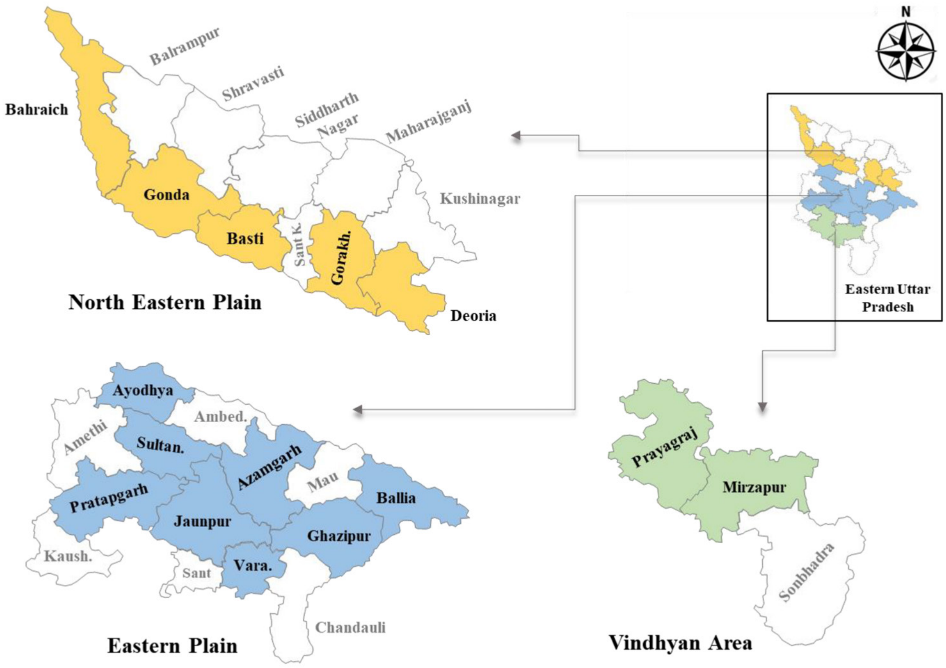

Fig. 1. Highlighted districts (yellow from North East Plain, blue from Eastern Plain, and green from Vindhyan region) in Eastern UP selected for area, production, and yield forecasting of sugarcane (Source: Author's compilation).

statistical analysis, including key measures such as maximum and minimum values, mean, skewness, and kurtosis. These measures helped in identifying trends in the area, production, and yield of sugarcane over time, as previously demonstrated by Kathayat and Dixit (2021), Singh (2024), Mishra *et al*. (2023) and Bharati and Singh (2019).

To model and forecast the sugarcane trends, several trend models were applied. These models included:

Linear model $Y_t = b_0 + (b_1 t)$

Quadratic model $Y_t = b_0 + (b_1 t) + (b_2 t^2)$

Compound model $\ln(Y_t) = \ln(b_0) + t \ln(b_1)$

Cubic model $Y_t = b_0 + (b_t t) + (b_2 t^2) + (b_3 t^3)$

Exponential model $Y_t = b_0 e(b_t t)$ or $\ln(Y_t) = \ln(b_0) + (b_1 t)$

Logarithmic model $Y_t = b_0 + b_1 \ln(t)$

Growth model $\ln(Y_t) = b_0 + b_1 t Y_t$

Instability in the data were analyzed using the Cuddy-Della Valle instability index (1978), a method extensively used in studies such as those by Prabakaran *et al.* (2013), Dharmaraja *et al.* (2020), and Srivastava *et al.* (2022). The index enabled the measurement of variability in the sugarcane data over the study period. The index is calculated as follows:

Where CV defined as:

$$CV_t = (CV) \times \sqrt{(1-R^2)} \text{ ....} \quad (1)$$

$$CV = \frac{\sigma}{\bar{X}} \times 100$$

Here, σ represents the standard deviation and $\bar{X}$ is the mean of the series.

Furthermore, Autoregressive Integrated Moving Average (ARIMA) models were employed to analyze the time series data. These models incorporated autoregressive (AR) and moving average (MA) components, with differencing techniques applied to stabilize the series (Srivastava *et al.* 2023). The ARIMA

Table 1. Area, production, and yield of sugarcane in Eastern UP. Source: ICRISAT, and Directorate of Economics & Statistics, 2024.

Year	Area ('000' ha)	Production '000' tonne	Yield (Ton/ha)	Year	Area ('000' ha)	Production '000' tonne	Yield (Ton/ha)
1960	281.00	4729.50	25.400	1992	329.41	16236.30	46.081
1961	286.10	5362.60	26.204	1993	316.98	15852.30	47.375
1962	284.20	6556.70	24.647	1994	323.19	17017.30	48.036
1963	306.70	7254.70	28.087	1995	333.76	16817.00	47.972
1964	300.40	7583.90	29.015	1996	331.68	17282.70	48.347
1965	312.70	9155.63	29.725	1997	325.84	9435.95	48.456
1966	304.30	9511.85	28.381	1998	320.91	8114.29	44.058
1967	283.40	10796.00	37.192	1999	334.16	9085.54	46.105
1968	305.60	13172.00	43.451	2000	353.52	9011.93	47.421
1969	331.00	14575.00	42.686	2001	378.67	10721.36	48.666
1970	319.60	12291.00	37.451	2002	393.24	10438.73	47.476
1971	317.90	10446.00	32.605	2003	384.62	9417.12	43.850
1972	334.30	13342.00	37.347	2004	350.32	9790.11	49.032
1973	326.60	10870.00	32.238	2005	380.28	10937.12	47.727
1974	315.10	11702.00	35.168	2006	397.90	12042.77	47.519
1975	312.80	11636.00	35.268	2007	403.29	10431.00	49.309
1976	304.00	13035.00	41.343	2008	391.96	9690.59	46.412
1977	307.30	12282.00	38.545	2009	381.28	10431.15	47.539
1978	315.80	11322.00	34.386	2010	393.78	10984.27	49.748
1979	286.70	9664.00	30.144	2011	396.87	12893.00	53.931
1980	271.90	10866.00	38.178	2012	408.05	13370.72	56.428
1981	310.80	13916.00	43.619	2013	435.68	13895.87	59.104
1982	349.10	13935.00	38.143	2014	410.64	14194.45	59.949
1983	326.20	13773.00	41.290	2015	402.06	14372.26	60.635
1984	288.90	10570.00	35.222	2016	394.89	14390.23	65.464
1985	280.50	12668.00	42.281	2017	400.33	15663.37	63.335
1986	299.00	13278.00	44.260	2018	370.30	16701.55	65.569
1987	308.50	14369.00	43.660	2019	404.19	17240.09	68.757
1988	313.20	13909.00	40.889	2020	441.17	22643.02	64.648
1989	316.80	15186.00	44.981	2021	434.78	21616.09	61.242
1990	330.49	16193.70	45.317	2022	433.40	23698.75	67.660
1991	349.87	17513.80	48.314				

model, expressed as ARIMA (p,d,q), allowed for the determination of autoregressive terms, degrees of differencing, and lagged forecast errors. The model was developed using data from 1960 to 2009, validated with data from 2010 to 2022, and used for projections extending to 2035. Several studies, including those by Choudhury *et al*. (2017), Khaemba *et al*. (2021) and Kumar *et al*. (2024), have demonstrated the effectiveness of this approach.

The performance and adequacy of the models were evaluated using several criteria: Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), R-squared (R^2), and Akaike's Information Criterion (AIC). Each of these metrics provided insights into the accuracy and reliability of the models. For instance, RMSE and MAPE offered measures of error magnitude, while R^2 assessed the goodness-of-fit. AIC helped determine the best model by accounting for both goodness-of-fit and model complexity (Kumar *et al.* 2023). Diagnostic checks on residuals were carried out using autocorrelation function (ACF) and partial autocorrelation function (PACF) plots to ensure that the residuals exhibited the characteristics of white noise, confirming the suitability of the models for forecasting. This diagnostic approach has been validated in various studies, including those by Ahmar *et al.* (2023), Umesh *et al*. (2023) and Kulkarni *et al.* (2018).

RESULTS AND DISCUSSION

Performance of sugarcane in eastern UP

Table 2 presents a comprehensive analysis of the performance of sugarcane in Eastern Uttar Pradesh over the period from 1960 to 2022. The table provides important statistics on the area under cultivation, production, and yield of sugarcane, segmented into three distinct periods.

Table 2 shows the average area under sugarcane cultivation over the entire period (1960–2022) was 343.46 thousand hectares. During Period I (1960–1980), the average area was 305.11 thousand hectares, which increased to 322.14 thousand hectares in Period II (1981–2000), and further rose to 399.44 thousand hectares in Period III (2001–2022). The increase in area from Period I to Period III is notable, with the CAGR indicating a positive trend of 0.69% across the whole period. However, the CAGR was negative in Period I (–0.16%) and marginally positive in Period II (0.65%), indicating variability in growth rates. The skewness of the area data was positive in Period III (0.17), suggesting a slight rightward skew, which implies that larger values are more frequent

Table 2. Performance of sugarcane area, production and yield in Eastern Uttar Pradesh during 1960-2022.

Particulars	Whole period (1960-2022)	Period I (1960-1980)	Period II (1981-2000)	Period III (2001-2022)
	Area ('000' ha)			
Mean	343.46	305.11	322.14	399.44
Standard error	5.82	3.78	4.24	4.76
Standard deviation	46.18	17.34	18.96	22.33
Kurtosis	-0.90	-0.74	0.22	0.27
Skewness	0.52	-0.24	-0.42	0.17
Minimum	271.90	271.90	280.50	350.32
Maximum	441.17	334.30	353.52	441.17
CAGR	0.69%	-0.16%	0.65%	0.62%
	Production ('000' tonnes)			
Mean	12633.05	10293.04	13708.24	13889.26
Standard error	471.27	588.10	672.55	903.98
Standard deviation	3740.60	2695.03	3007.74	4240.03
Kurtosis	1.00	-0.34	-0.85	0.56
Skewness	0.62	-0.62	-0.59	1.17
Minimum	4729.50	4729.50	8114.28	9417.12
Maximum	23698.70	14575.00	17513.80	23698.70
CAGR	2.59%	4.04%	-2.15%	3.67%
	Yield (ton/ha)			
Mean	44.81	33.69	44.59	55.64
Standard error	1.38	1.24	0.81	1.74
Standard deviation	10.97	5.67	3.62	8.15
Kurtosis	-0.28	-1.04	1.01	-1.58
Skewness	0.30	0.06	-1.11	0.19
Minimum	24.65	24.65	35.22	43.85
Maximum	68.76	43.45	48.46	68.76
CAGR	1.57%	1.96%	0.42%	1.51%

Table 3. Trends in area, production and yield of sugarcane in Eastern Uttar Pradesh.

	Model summary and parameter estimates					
	Area ('000' ha)					
	Model summary			Parameter estimates		
Equation	R^2	Significance	Constant (b)	b_1	b_2	b_3
Linear	0.759	0.00	273.21	2.20	-	-
Exponential	0.781	0.00	0.00	0.01	-	-
Logarithmic	0.508	0.00	227.24	36.43	-	-
Quadratic	**0.861**	**0.00**	**276.32**	**7.33**	**-0.51**	**0.01**
Cubic	0.829	0.00	305.09	-1.07	0.07	-
Compound	0.781	0.00	0.00	0.01	-	-
Growth	0.781	0.00	0.00	0.01	-	-
	Production ('000' tonnes)					
Linear	0.280	0.00	9180.20	107.90	-	-
Exponential	0.281	0.00	0.00	0.01	-	-
Logarithmic	0.331	0.00	5035.40	2381.30	-	-
Quadratic	**0.716**	**0.00**	**5808.80**	**373.12**	**17.42**	**-1.06**
Cubic	0.657	0.00	2648.80	1295.40	-46.24	-
Compound	0.281	0.00	0.00	0.01	-	-
Growth	0.281	0.00	0.00	0.01	-	-
	Yield (ton/ha)					
Linear	0.853	0.00	27.12	0.55	-	-
Exponential	0.866	0.00	0.00	0.01	-	-
Logarithmic	0.697	0.00	12.48	10.14	-	-
Quadratic	**0.896**	**0.00**	**24.34**	**1.26**	**-0.03**	**0.00**
Cubic	0.896	0.00	24.10	1.33	-0.04	-
Compound	0.866	0.00	0.00	0.01	-	-
Growth	0.866	0.00	0.00	0.01	-	-

in recent years. The kurtosis values indicate that the distribution of the area under cultivation has been relatively platykurtic (negative kurtosis) or close to normal, especially in the earlier periods.

The average production of sugarcane over the entire period was 12,633.05 thousand tonnes. In Period I, the average production was 10,293.04 thousand tonnes, which rose significantly to 13,708.24 thousand tonnes in Period II, and further increased to 13,889.26 thousand tonnes in Period III. The CAGR for production was 2.59% across the whole period, with a peak in Period I at 4.04% and a decline in Period II to -2.15%. The production growth in Period III was robust at 3.67%. The high standard deviation in production (4,240.03 tonnes in Period III) compared to earlier periods suggests increased variability in production levels. The skewness value of 1.17 in Period III indicates a positive skew, reflecting a tendency towards higher production values in recent years. The kurtosis values vary from platykurtic to leptokurtic across periods, indicating differing degrees of distribution shape.

The average yield of sugarcane over the whole period was 44.81 tonnes per hectare. In Period I, the average yield was 33.69 tonnes per hectare, which improved to 44.59 tonnes per hectare in Period II, and increased further to 55.64 tonnes per hectare in Period III. The yield data exhibits an increasing trend with a CAGR of 1.57% for the entire period. Period I saw a growth rate of 1.96%, which slowed to 0.42% in Period II but picked up to 1.51% in Period III. The yield's standard deviation was highest in Period I (5.67 tonnes/ha), reflecting greater variability in early years. The skewness of yield was relatively stable across periods, with a positive skew in Period III (0.19) suggesting a tendency for higher yields. The kurtosis values indicate that yield distributions have varied from platykurtic to leptokurtic, highlighting changes in the distribution of yields over.

Trends in area, production and yield

The trend analysis of sugarcane in Eastern Uttar Pradesh, as presented in Table 3 with graphical representations in Figs. 2–4, offers significant overview the dynamics of area, production and yield over time.

For the area under sugarcane cultivation, the quadratic model stands out as the most appropriate, evidenced by the highest R^2 value of 0.861, indicating a strong correlation between the area and time. The model's statistical significance at the 0.00 level further reinforces its reliability. The constant (b) is estimated at 276.32, with coefficients b_1, b_2 and b_3 being 7.33, -0.51 and 0.01, respectively. These coefficients suggest an initial increase in the area under sugarcane, followed by a deceleration, reflecting a complex, non-linear relationship. This pattern may be attributed to a variety of factors including changes in agricultural practices, policy shifts, or economic constraints that limit further expansion after an initial growth phase. The cubic model, while also significant with an R^2 of 0.829, does not outperform the quadratic model in terms of explaining the variations in the area under sugarcane cultivation. Other models such as the linear (R^2 = 0.759), exponential (R^2 = 0.781), and compound (R^2 = 0.781) exhibit lower R^2 values, indicating a less precise fit. The logarithmic model, with an R^2 of 0.508, shows the weakest correlation, emphasizing the superiority of the quadratic model in capturing the area trends.

In analyzing the production of sugarcane, the quadratic model again provides the best fit with an R^2 value of 0.716, highlighting its ability to capture the trend over time. The significance level is maintained at 0.00, ensuring the robustness of the model. The constant (b) is recorded at 5808.80, with coefficients b_1, b_2 and b_3 at 373.12, 17.42 and –1.06, respectively. These figures suggest an overall positive trend in production, although the negative b_3 coefficient implies that the rate of increase may taper off or even decline slightly in the long term. This pattern could be a reflection of fluctuating environmental conditions, variations in input availability, or shifts in market demand impacting production levels. Other models, including the cubic model (R^2 = 0.657), also show significant results but do not match the explanatory power of the quadratic model. The linear model, with an R^2 of 0.280, and the exponential model (R^2 = 0.281) demonstrate a weaker correlation with production trends, indicating that these models may not fully capture the complexity of sugarcane production in the region.

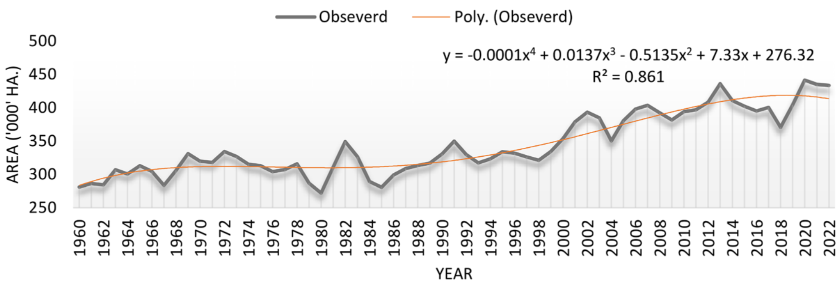

Fig. 2. Observed and expected trends of area under sugarcane in Eastern Uttar Pradesh.

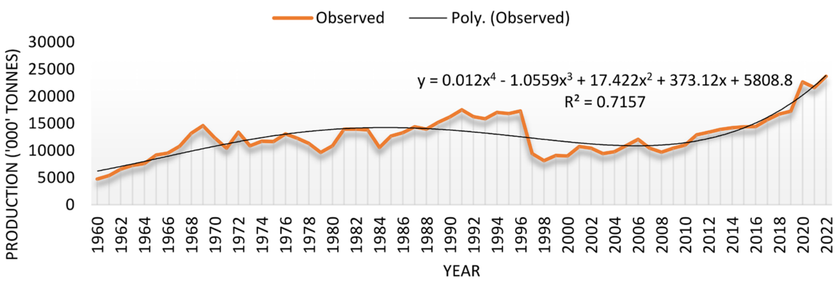

Fig. 3. Observed and expected trends of production under sugarcane in Eastern Uttar Pradesh.

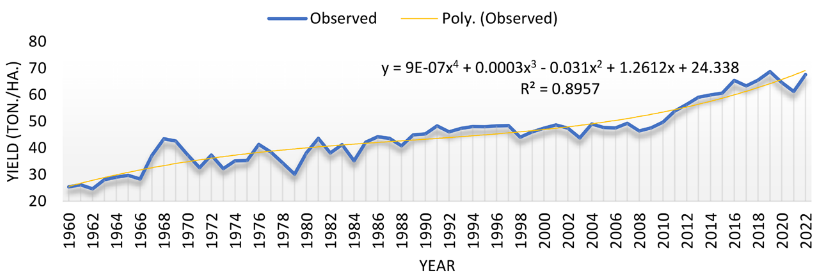

Fig. 4. Observed and expected trends of yield under sugarcane in Eastern Uttar Pradesh.

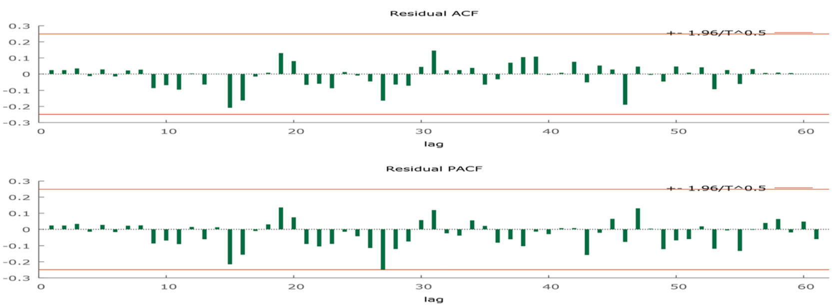

Fig. 5. ACF and PACF graphs of residuals for the best fitted models of area under sugarcane in Eastern Uttar Pradesh.

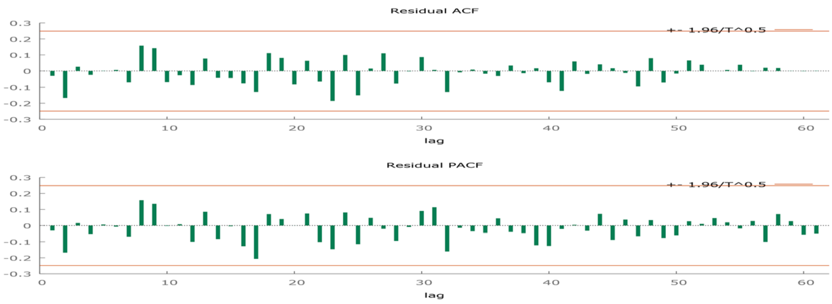

Fig. 6. ACF and PACF graphs of residuals for the best fitted models of production under sugarcane in Eastern Uttar Pradesh.

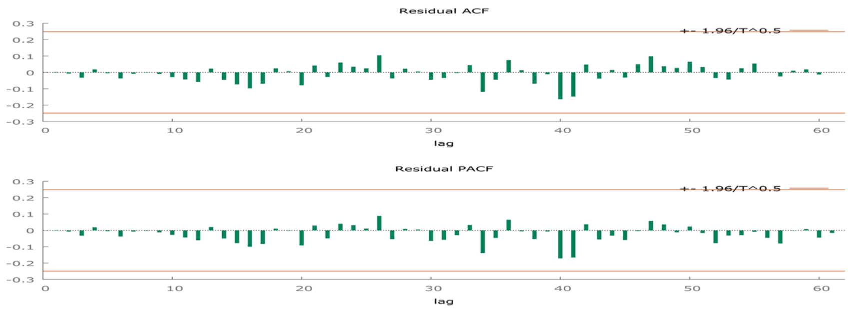

Fig. 7. ACF and PACF graphs of residuals for the best fitted models of yield under sugarcane in Eastern Uttar Pradesh.

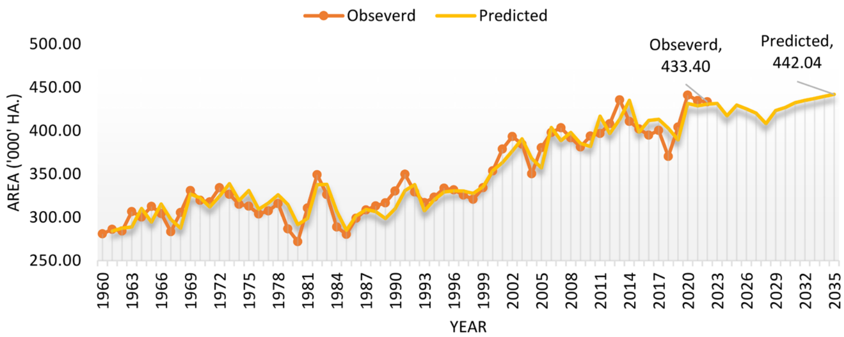

Fig. 8. Observed and forecasting of area of sugarcane in Eastern Uttar Pradesh.

The yield analysis of sugarcane reveals that the quadratic model is the most suitable, with an R^2 value of 0.896, the highest among all models tested. The constant (b) is estimated at 24.34, with coefficients b_1, b_2 and b_3 at 1.26, -0.03 and 0.00, respectively. The positive b_1 coefficient indicates an increasing yield trend, though the small negative b_2 suggests a potential stabilization or slight decline in growth rate over time. This trend could be influenced by advancements in agricultural techniques, better seed varieties, or improvements in soil management practices. The cubic model, which also has an R^2 of 0.896, closely matches the quadratic model in terms of fit, but the quadratic model is preferred due to its slightly better explanation of the yield trend. The linear (R^2 = 0.853), exponential (R^2 = 0.866), and compound (R^2 = 0.866) models, while significant, do not offer as comprehensive an understanding of yield trends as the quadratic model.

Instability analysis

The analysis of sugarcane instability in Eastern Uttar Pradesh, as presented in Table 4, highlighted signif-

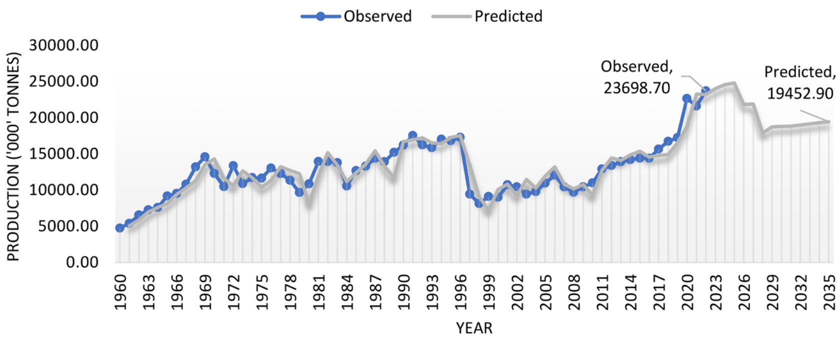

Fig. 9. Observed and forecasting of production of sugarcane in Eastern Uttar Pradesh.

icant variations in the area, production, and yield across different periods. The R^2 value for the area under sugarcane cultivation during the whole period was 0.759, indicating moderate stability. However, this stability was not uniform across the sub-periods. Period I exhibited a low R^2 of 0.036, reflecting high instability in the area. Period II showed a slight improvement with an R^2 of 0.223, while Period III further improved to 0.407, suggesting a gradual stabilization over time. The CV for the entire period was 13.446%, with lower values in the sub-periods, indicating reduced variability in the area over time. The CV_t decreased consistently from 6.652 in the whole period to 4.411 in Period III, indicating increased area stability.

Table 4. Instability in area, production and yield of sugarcane in Eastern Uttar Pradesh.

Statistics	Whole period	Period I	Period II	Period III
	Area ('000' ha)			
R^2	0.759	0.036	0.223	0.407
CV	13.446	5.683	5.886	5.589
CV_t	6.652	5.723	5.330	4.411
	Production ('000' tonnes)			
R^2	0.280	0.447	0.056	0.797
CV	29.610	26.183	21.941	30.527
CV_t	25.337	19.970	21.903	14.077
	Yield (ton/ha)			
R^2	0.853	0.333	0.527	0.854
CV	24.482	16.831	8.113	14.647
CV_t	9.456	14.103	5.730	5.743

Production instability was more pronounced, as evidenced by a low R^2 of 0.280 for the whole period. Period I had a moderate R^2 of 0.447, while Period II showed a significant drop to 0.056, indicating high instability. Period III displayed a substantial improvement with an R^2 of 0.797, reflecting better stabilization. The CV for production was highest in Period III at 30.527%, which was consistent with the overall pattern of volatility. The CV_t also showed fluctuations, with the highest stability observed in Period III at 14.077.

Yield instability revealed a more consistent pattern, with a high R^2 of 0.853 for the whole period, suggesting strong stability. However, this stability fluctuated across periods, with R^2 values of 0.333 in Period I, 0.527 in Period II, and 0.854 in Period III. The CV and CV_t followed a similar trend, indicating that yield stability was more controlled in the later periods, particularly in Period II and III.

Table 5. Best fitted ARIMA model for area, production and yield under sugarcane in Eastern Uttar Pradesh.

	ARIMA model	R^2	RMSE	MAPE	MAE	AIC
Area	ARIMA (1,1,10)	0.915	13.29	3.236	11.066	532.932
Production	ARIMA (0,1,9)	0.850	1416.10	8.964	1087.50	1114.548
Yield	ARIMA (1,1,10)	0.927	2.891	5.289	2.258	339.355

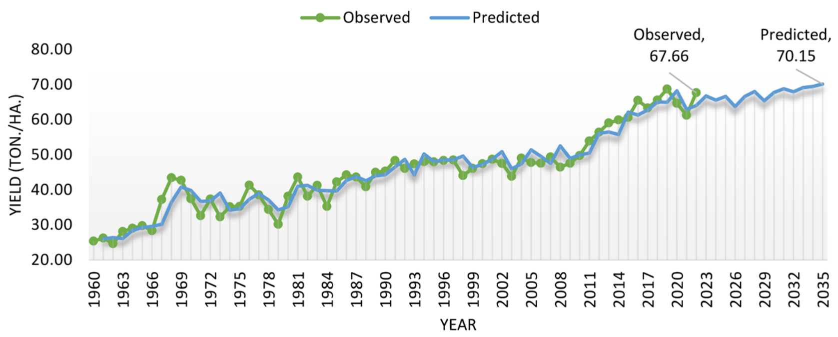

Fig. 10. Observed and forecasting of yield of sugarcane in Eastern Uttar Pradesh.

Modelling and forecasting

The modelling and forecasting of sugarcane production in Eastern Uttar Pradesh were performed using various ARIMA models. The ACF and PACF plots of the first difference for the area under sugarcane, illustrated in Figs. 5–7, suggested optimal tentative values for the parameters p and q as 1 and 10, respectively.

The ARIMA (1,1,10) model was found to deliver superior predictive performance for the area under sugarcane, demonstrated by the highest R^2 value of 0.915 and the lowest RMSE of 13.290 and MAE of 11.066 (Table 5).

For sugarcane production, the ARIMA (0,1,9) model yielded the most accurate results, with an

Table 6. Model validation and forecasting of area, production and yield under sugarcane in Eastern Uttar Pradesh.

Year	Area ('000' ha)		Production ('000' tonnes)		Yield (ton/ha)	
	Observed	Predicted	Observed	Predicted	Observed	Predicted
2010	393.780	381.872	10984.300	9297.330	49.748	49.943
2011	396.870	416.625	12893.000	12693.500	53.931	50.476
2012	408.050	396.648	13370.700	14421.900	56.428	56.093
2013	435.680	413.220	13895.900	14077.500	59.104	56.507
2014	410.640	435.112	14194.500	14832.500	59.949	55.734
2015	402.060	398.529	14372.300	15340.800	60.635	62.088
2016	394.890	411.801	14390.200	14531.800	65.464	61.318
2017	400.330	413.155	15663.400	14761.900	63.335	62.742
2018	370.299	402.765	16701.500	14914.100	65.569	65.057
2019	404.189	389.371	17240.100	16519.300	68.757	64.949
2020	441.174	431.349	22643.000	18770.900	64.648	68.194
2021	434.776	428.655	21616.100	23272.500	61.242	62.932
2022	433.397	430.570	23698.700	23188.600	67.660	64.052
2023	-	431.543	-	24019.700	-	66.724
2024	-	417.353	-	24568.900	-	65.555
2025	-	429.467	-	24783.600	-	66.610
2026	-	425.105	-	21847.700	-	63.716
2027	-	420.232	-	21863.500	-	66.616
2028	-	408.318	-	17870.900	-	67.996
2029	-	423.285	-	18731.500	-	65.381
2030	-	426.920	-	18798.800	-	67.733
2031	-	432.304	-	18832.600	-	68.836
2032	-	435.153	-	18987.700	-	67.967
2033	-	437.326	-	19142.700	-	69.132
2034	-	439.717	-	19297.800	-	69.473
2035	-	442.038	-	19452.900	-	70.148

R^2 of 0.850 and RMSE of 1416.100, confirming its suitability for forecasting. Additionally, the ARIMA (1,1,10) model was most effective for yield prediction, achieving the highest R^2 value of 0.927, accompanied by the lowest RMSE of 2.891 and MAE of 2.258.

In 2010, the observed area for sugarcane cultivation in Eastern Uttar Pradesh was 393.78 thousand hectares, with production at 10,984.30 thousand tonnes and a yield of 49.748 tonnes per hectare (Table 6). The model predicted a slightly lower area of 381.872 thousand hectares and significantly lower production of 9,297.33 thousand tonnes, though the yield was closely aligned at 49.943 tonnes per hectare. By 2035, forecasts suggest that the area will reach 442.038 thousand hectares, with production increasing substantially to 19,452.90 thousand tonnes and yield rising to 70.148 tonnes per hectare. Figs. 8 – 10 illustrate the observed and predicted data from 1960 to 2022, highlighting forecast values until 2035. These projections reflect an anticipated positive trend in sugarcane production, likely driven by advancements in agricultural practices and technology, thus contributing to improved livelihood status in the region.

CONCLUSION

Sugarcane cultivation in Eastern Uttar Pradesh shows a dynamic yet somewhat fluctuating trend in area, production and yield, as revealed by the ARIMA model's predictions. The model forecasted a slightly reduced area of 381.872 thousand hectares and a markedly lower production of 9,297.330 thousand tonnes, while the yield closely aligned with historical trends at 49.943 tonnes per hectare in 2010. However, projections for 2035 suggest substantial growth, with the area expanding to 442.038 thousand hectares, production rising significantly to 19,452.90 thousand tonnes and yield reaching 70.148 tonnes per hectare. This forecast shows a strong upward trend in sugarcane output, driven by better agronomic practices, technological advancements and favorable agro-climatic conditions. These findings emphasize the importance of strategic planning in addressing future challenges such as resource allocation and market stability, ensuring sustained agricultural productivity. This approach ensures long-term sustainability for the sugarcane industry in Eastern Uttar Pradesh.

REFERENCES

Ahmar AS, Singh PK, Ruliana R, Pandey AK, Gupta S (2023) Comparison of Arima, Sutte ARIMA and holt-winters, and NNAR models to predict food grain in India. *Forecasting* 5 : 138—152. https://doi.org/10.3390/forecast5010006.

Bharati RC, Singh AK (2019) Predicting rice production using autoregressive integrated moving average model. *Journal of Agric Search* 6 : 205—210. https://doi.org/10.21921/jas.v6i04.16905.

Choudhury N, Saurav S, Kumar RR, Budhlakoti N (2017) Modelling and forecasting of total area, irrigated area, production and productivity of important cereal crops in India towards food security. *International Journal of Current Microbiology and Applied Sciences* 6 : 2591—2600. https://doi.org/10.20546/ijcmas.2017.610.305.

Cuddy JDA, Della Valle PAD (1978) Measuring the instability of time series data. *Oxford Bulletin of Economics and Statistics* 40 : 79—85. https://doi.org/10.1111/j.1468-0084.1978.mp40001006.x

Dharmaraja S, Jain V, Anjoy P, Chandra H (2020) Empirical analysis for crop yield forecasting in India. *Agricultural Research* 9 : 132—138. https://doi.org/10.1007/s40003-019-00413-x

Directorate of Economics & Statistics (2024) Agricultural statistics at a glance. Ministry of Agriculture & Farmers Welfare, Government of India.

Hooda E, Verma U, Hooda BK (2020) Arima and State-Space models for sugarcane (*Saccharum officinarum*) yield forecasting in Northern agro-climatic zone of Haryana. *Journal of Applied and Natural Science* 12 : 53—58. https://doi.org/10.31018/jans.v12i1.2229

ICRISAT (2024) Annual report. International Crops Research Institute for the Semi-Arid Tropics.

Kathayat B, Dixit AK (2021) Paddy Price Forecasting in India Using Arima Model.

Khaemba PF, Muiruri PW, Kibutu TN (2021) Trend analysis in sugarcane growth in Mumias sugar belt, Western Kenya, for the period 1985—2015. *Interdisciplinary Journal of Rural and Community Studies* 3 : 31—40. https://doi.org/10.51986/ijrcs-2021.vol3.01.04.

Kulkarni S, Mandal SN, Sharma GS, Mundada MR, Meeradevi (2018) Predictive analysis to improve crop yield using a neural network model. In: International conference on advances in computing, communications and informatics (ICACCI). IEEE Publications, pp 74—79. https:/ doi.org/10.1109/ICACCI.2018.8554851.

Kumar L (2022) Forecasting of area, production and productivity of wheat in Madhya Pradesh and its instability. *Indian Journal of Economics and Development* 18 : 3—3.

Kumar M, Prakash G, Rana SK (2024) Statistical modeling for analysis of growth and trend pattern of wheat production in selected states of India. *Asian Journal of Research in Crop Science* 9 : 66—75. https://doi.org/10.9734/ajrcs/2024/v9i1246.

Kumar N, Kushwaha RR, Meena NR, Mishra H, Yadav APS (2023) A study on costs and returns of paddy cultivation in Ambedkar Nagar district of Uttar Pradesh. *International Journal of Statistics and Applied Mathematics* 8 : 107—111.

Mehmood Q, Raza A, Abro AA, Shaheen N, Riaz M (2024) Development of linear, nonlinear and hybrid models for forecasting sugarcane yield. *Sarhad Journal of Agriculture* 40 : 754—759. https://doi.org/10.17582/journal.sja/2024/40.3.754.759.

Mishra D, Singh KK, Mishra H, Srivastava AB (2023) Resource use efficiency (RUE) of lentil cultivation in Sultanpur District of Uttar Pradesh. *Environment and Ecology* 41 : 1209—1216.

Mishra M, Thakur RK (2021) Comparative study of non-linear growth models for area and production of sugarcane crop in Uttar Pradesh. *International Journal of Agricultural and Statistical Sciences,* pp 17.

Padhan PC (2012) Application of Arima model for forecasting agricultural productivity in India. *Journal of Agriculture and Social Sciences* 8 : 50—56.

Prabakaran K, Sivapragasam C, Jeevapriya C, Narmatha A (2013) Forecasting cultivated areas and production of wheat in India using Arima model. *Golden Research Thoughts* 3 : 1—7. https://doi.org/10.9780/2231-5063/332013/2831.

Sahu PK, Mishra P (2014) Instability scenario of maize in India and forecasting using ARIMA model. *International Journal of Agricultural and Statistical Sciences* 10 : 425—435.

Saini P, Nagpal B, Garg P, Kumar S (2023) Evaluation of remote sensing and meteorological parameters for yield prediction of sugarcane (*Saccharum officinarum* L.) Crop. *Brazilian Archives of Biology and Technology* 66 : e23220781. https://doi.org/10.1590/1678-4324-2023220781.

Singh Y (2024) Production and productivity trends of major foodgrain crops in Uttar Pradesh. *Management Journal for Advanced Research* 4 : 149—154.

Srivastava AB, Kushwaha RR, Yadav S, Verma SK, Mishra P (2022) Source of growth for wheat in Uttar Pradesh: Decomposition analysis. *Indian Journal of Economics and Development* 18 : 976—980.

Srivastava AB, Singh KK, Supriya SK, Verma SK, Mishra H, Ahmad DR (2023) Production and export dynamics of wheat in India. *International Journal of Statistics and Applied Mathematics* 8 : 206—209. https://doi.org/10.22271/maths.2023.v8.i3Sc.1023.

Supriya B, Srivastava AB, Kumari B, Yadav S, Shrivastava A, Neerugatti MP (2024) Comparative analysis of sugarcane production for South East Asia Region. *Sugar Tech Srivastava*, *AB: Kumari, Turkey* 26 : 264—273. https://doi.org/10.1007/s12355-023-01346-0.

Tyagi S, Chandra S, Tyagi G (2023) Climate change and its impact on sugarcane production and future forecast in India. A comparison study of univariate and multivariate time series models. *Sugar Tech* 25 : 1061—1069. https://doi.org/10.1007/s12355-023-01271-2.

Umesh KB, Hamsa KR, Vijayalaxmi DK (2023) Trends and prospects of crop production in Karnataka. In: Agricultural Sector in India. Routledge: India, pp 89—106.

Verma AK, Garg PK, Hari Prasad KS, Dadhwal VK, Dubey SK, Kumar A (2021) Sugarcane yield forecasting model based on weather parameters. *Sugar Tech* 23 : 158—166. https://doi.org/10.1007/s12355-020-00900-4.

Vishwajith KP, Dhekele BS, Sahu PK, Mishra P, Norman MD (2014) Time series modelling and forecasting of pulses production in India. *Journal of Crop Weed* 10 : 147—150.

Yassen M, Zakria M, Shahzad ID, Khan MI, Javed MA (2005) Modelling and forecasting the sugarcane yield of Pakistan. *International Journal of Agricultural Biology* 7 : 180—183.