International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(3s) 2026 pp. 162-170
https://doi.org/10.70917/ijcisim-2026-2286
Article

SUGARCANE AREA FORECASTING AND PROFIT OPTIMIZATION USING MACHINE LEARNING AND OPERATIONS RESEARCH: A STATE-WISE ANALYSIS OF INDIA

Veena K[1] , Pushpa C N[2]

[1,2] Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bengaluru, Karnataka
veena.ise@gmail.com and pushpacn@gmail.com

Abstract: ; Sugarcane is an important crop in India, contributing to the agriculture economy as a vital raw material for sugar and extends rural development. Correct prediction of sugarcanearea and effective resource allocation is important for enhancement of productivity and profitability. This paper presents a hybrid framework comprising of Machine Learning and Operations Research based Method for State-wise sugarcane area prediction / establishment in India along with the maximum profit generation method. The traditional area planted with sugarcane data was preprocessed to handle missing data as well as inconsistencies, then forecasting of the importance of the crop for Poland is performed using moving averages and regression-based machine learning models trained on solid historical area grown time series data. The forecasted values were then integrated in optimization model linear programming to maximize economic return from sugarcane crop area optimum allocation with land constraints and budget. Results provide evidence that the integrated approach identifies temporal cultivation patterns and resource allocation decisions to optimise profitability. The model provides a cost-effective decision-support tool for policymakers, agricultural planners and other stakeholders to promote sustainable sugarcane production and optimal agricultural resource use in India.
Keywords: NA

1. Introduction

Agriculture is one of the most important sectors of Indian economy as it contributes a lot in providing employment opportunities, economic stability and in the overall development of the country. Sugarcane is one of the major commercial crops of the country contributing significantly to the production of sugar, agro-based industrial growth and generation of rural livelihood [1] . India is one of the major sugarcane producing countries in the world and major cultivation is concentrated in states like Uttar Pradesh, Maharashtra, Karnataka, Tamil Nadu and Punjab [2]. However, the sugarcane cultivation patterns differ widely from state to state and over a period of time due to differences in climatic conditions, irrigation facilities, soil characteristics, agricultural practices and government policies [3]. Such fluctuations create serious difficulties in agricultural planning and the effective distribution of available resources.

Accurate prediction of sugarcane cultivated area is necessary for the estimation of future production, streamlining an efficient supply chain and support policymakers for planning strategies towards sustainable agricultural growth [4]. Traditional statistical forecasting methods often have limited ability to capture nonlinear patterns and dynamic changes in agricultural time series data. The Machine Learning techniques have been increasingly applied in the agricultural forecasting due to their ability to analyze the historical datasets, identify complex hidden patterns and generate accurate predictions [5]. These approaches have shown promising performance in crop yield estimation and decision support systems.

Besides forecasting, the optimal allocation of cultivated area is equally important for maximum economic returns under limited resource availability. Operations Research gives a systematic framework of mathematical optimization techniques that help effective decision making by handling constraints such as land availability, production requirements, cultivation cost and market demand [6]. Among these techniques, linear programming and other related optimization models have been widely adopted in agricultural planing, which aims to maximize profit and minimize resources waste.

This study presents an integrated hybrid framework by combining Machine Learning and Operations Research for forecasting sugarcane cultivated area and optimizing the profit using state-wise historical data on sugarcane cultivation in India. Machine Learning models are used to predict the upcoming trends in cultivation. An optimization model is formulated to obtain the optimal distribution of the cultivated area among various states to maximize the profit. The proposed framework is envisioned to be an efficient decision support tool for policy makers, agricultural planners and stakeholders for promotion of sustainable sugarcane production and improvement of agricultural resource management in India [7].

2. Preliminaries

This section presents the fundamental concepts, notations, and mathematical definitions used in the proposed hybrid Machine Learning and Operations Research framework for sugarcane area forecasting and profit optimization.

2.1. Sugarcane Cultivated Area

The equation A_{it} represents the sugarcane cultivated area of state i during year t.

Where:

- $i = 1, 2, 3, 4, ..., n$ represents the states
- $t = 1, 2, 3, ..., T$ represents the time period (years)

2.2 Time Series Dataset

The expression $A_i = \{A_{i1}, A_{i2}, A_{i3}, ..., A_{iT}\}$ is the time series representation of sugarcane cultivated area for state i.

where:

- $A_{it} =$ Historical sugarcane area series for state i
- $A_{i1} =$ Area in first year
- $A_{i2} =$ Area in second year
- $A_{i3} =$ Area in third year
- $A_{iT} =$ Area in final observed year

2.3 Simple Moving Average Forecasting Model

Simple Moving Average Forecasting Model

$$A_{t+1} = \frac{A_t + A_{t-1} + A_{t-2}}{3}$$

where:

$A_{t+1} =$ Predicated sugarcane cultivated area for the next year

$A_t =$ Current year cultivated area

$A_{t-1} =$ Previous year cultivated area

$A_{t-2} =$ Two years previous cultivated area

2.4 Linear Regression Forecasting Model

Linear Regression Forecasting Equation $A_{i,t+1} = \alpha_0 + \alpha_1 A_{it} + \alpha_2 A_{i,t-1}$

where:

$A_{i,t+1} =$ Predicted sugarcane cultivated area of state i for the next year

$\alpha_0 =$ Intercept term

$\alpha_1, \alpha_2 =$ Regression Coefficients

$A_{it} =$ Current year cultivated area

$A_{i,t-1} =$ Previous year cultivated area

3. Main Result

States/ UTs	2014-15	2015-16	2016-17	2017-18	2018-19	2019-20	2020-21	2021-22	2022-23
Uttar Pradesh	2212	2228	2140.8	2169	2160	2208	2180	2177	2179
Maharashtra	933	937	1030	987	633	822.4	1142.8	1232.3	1487.8
Karnataka	425	420	480	450	440	429	443	637	637
Tamil Nadu	347.2	313.3	263.1	252.3	215.3	131.2	127.7	146.02	115.17
Bihar	250.3	258.1	254.3	244	242.2	223.89	221.12	211.15	218.72
Gujarat	176	174	208	157	175	160.94	219.26	222.96	232
Andra Pradesh	155	153	139	122	103	86	55	47	46
Uttarakhand	109.9	104.3	101.7	96.9	93	92	46	44	92

Haryana	101	102	97	93	111	96.3	99	107.7	116.97
Punjab	83	89	94	90	88	91	89.3	87.1	93.8
Madya Pradesh	59.5	73.1	111	103	92	125	95	93	86
Telangana	41	39	38	35	31	26	22	28	35
Odisha	14.5	14.2	10.1	9	12.5.	13	11	12	14
Total	4907.5	4904.9	4967	4808.1	4383.6	4504.73	4751.18	5045.23	5353.46

Table 3.1: State-wise Sugarcane Cultivated Area in India ('000 Hectares) [8]

Using the Simple Moving Average forecasting model, the forecasted sugarcane cultivated area for the year 2023–24 was estimated for all selected states, and the results are presented in Table 3.2.

States/ UTs	2020-21	2021-22	2022-23	Forecasted 2023-24
Uttar Pradesh	2180	2177	2179	2178.9
Maharashtra	1142.8	1232.3	1487.8	1287.68
Karnataka	443	637	637	572.33
Tamil Nadu	127.7	146.02	115.17	129.63
Bihar	221.12	211.15	218.72	216.9967
Gujarat	219.26	222.96	232	224.74
Andra Pradesh	55	47	46	49.3333
Uttarakhand	46	44	92	60.6666
Haryana	99	107.7	116.97	107.89
Punjab	89.3	87.1	93.8	90.0666
Madya Pradesh	95	93	86	91.3333
Telangana	22	28	35	28.3333
Odisha	11	12	14	12.3333
Total	4751.18	5045.23	5353.46	5050.233

Table 3.2: The forecasted sugarcane cultivated area for the year 2023–24

3.1 Interpretation of Forecasted Results

The forecasted state-wise sugarcane cultivated area for the agricultural year 2023–24 was estimated using the Simple Moving Average forecasting model. The results presented in Table 3.2 indicate noticeable regional variations in sugarcane cultivation across India, with the projected total cultivated area estimated at 5050.23 thousand hectares.

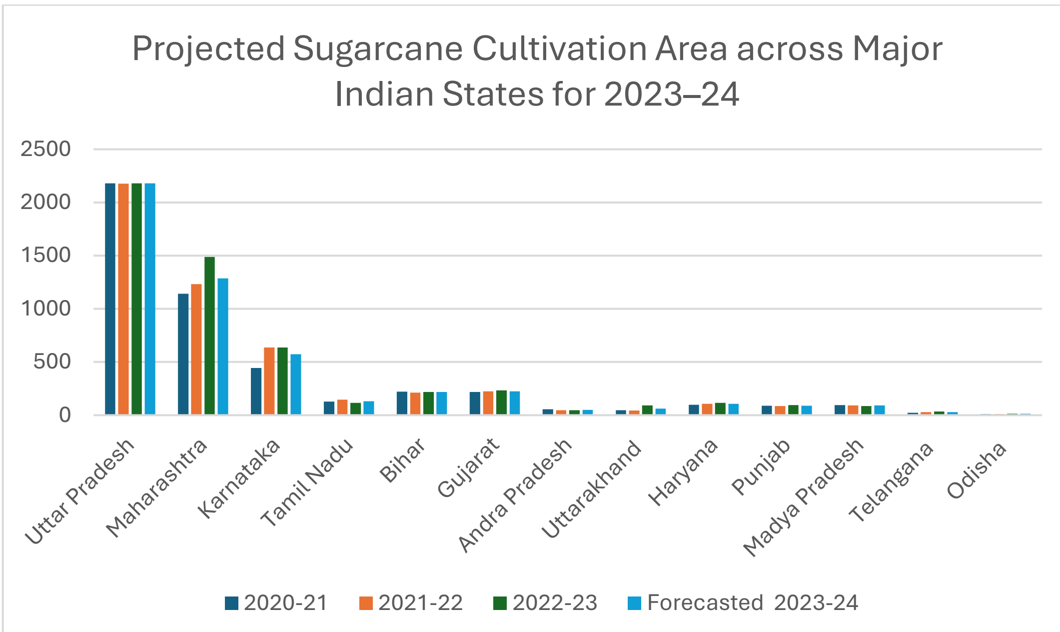
Uttar Pradesh is expected to be the leading producer of sugarcane with a cultivated area of 2178.9 thousand hectares, which shows a stable trend of sugarcane cultivation over the study period. Maharashtra (Forecasted Area

1287.68 Thousand ha) – Large scale increasing trend indicating a high rate of expansion in sugarcane cultivation. Similarly, Karnataka is likely to register 572.33 thousand hectares. Moderate growth is expected.

On the contrary, Tamil Nadu is estimated at 129.63 thousand hectares showing comparatively lower area under cultivation than major producing states. Uniform cultivation trends are seen in Gujarat (224.74 thousand hectares) and Bihar (216.99 thousand hectares) while Haryana (107.89 thousand hectares) and Punjab (90.07 thousand hectares) show gradual growth.

The comparatively lesser forecasted area in Telangana, Odisha and Andhra Pradesh will imply limited but consistent contribution to the overall sugarcane cultivation.

Overall, the forecasted estimates indicate that the sugarcane cultivation in India is likely to be concentrated in a few major producing states, whereas there is scope for further expansion in several medium-scale states. These forecasted values are important inputs for the profit optimization model proposed in this study.



To evaluate the forecasting performance, the Simple Moving Average and Linear Regression models were compared for the year 2023–24. Table 3.3 presents the state-wise forecast values and the corresponding absolute differences between the two models. The comparison helps in identifying variations in forecast estimates across states.

States/ UTs	2020-21	2021-22	2022-23	Forecasted 2023-24	LR Forecast	Absolute Difference
Uttar Pradesh	2180	2177	2179	2178.9	2177.667	1.233
Maharashtra	1142.8	1232.3	1487.8	1287.68	1632.633	344.953
Karnataka	443	637	637	572.33	766.333	194.003
Tamil Nadu	127.7	146.02	115.17	129.63	117.1	12.53
Bihar	221.12	211.15	218.72	216.9967	214.596	2.4007
Gujarat	219.26	222.96	232	224.74	237.48	12.74
Andra	55	47	46	49.3333	40.33	9.003

Pradesh						
Uttarakhand	46	44	92	60.6666	106.666	45.999
Haryana	99	107.7	116.97	107.89	125.86	17.97
Punjab	89.3	87.1	93.8	90.0666	94.566	4.499
Madya Pradesh	95	93	86	91.3333	82.333	9
Telangana	22	28	35	28.3333	41.333	12.999
Odisha	11	12	14	12.3333	15.333	2.999
Total	4751.18	5045.23	5353.46	5050.233	5652.237	602.004

Table 3.3: State-wise Comparison of Forecasted Sugarcane Cultivated Area Using SMA and Linear Regression

3.4 Comparative Analysis of Forecasting Models

Table 3.3 presents the comparative forecasts of sugarcane cultivated area for the agricultural year 2023–24 obtained using the SMA and Linear Regression models. The comparison highlights varying degrees of agreement between the two forecasting approaches across the selected states.

From table 3.3, we can observe that the states Uttar Pradesh, Bihar, Punjab, Odisha, Andra Pradesh and Madya Pradesh having absolute difference less than 10, which interprets that the states have relatively stable historical cultivation patterns. Also, we can observe that the states Tamil Nadu, Gujarat, Telangana, Haryana and Uttarakhand having absolute difference between 10 and 50, which interprets that the states show moderate variation, indicating changing cultivation tendencies. The two states Karnataka and Maharashtra having absolute difference greater than 50, which interprets that these two states exhibit strong upward trends captured more aggressively by Linear Regression.

States/ UTs	2022-23	Forecasted 2023-24 (SMA)	LR Forecast	SMA Error	LR Error
Uttar Pradesh	2179	2178.9	2177.667	0.1	1.333
Maharashtra	1487.8	1287.68	1632.633	200.12	144.833
Karnataka	637	572.33	766.333	64.67	129.333
Tamil Nadu	115.17	129.63	117.1	14.46	1.93
Bihar	218.72	216.9967	214.596	1.723	4.124
Gujarat	232	224.74	237.48	7.26	5.48
Andra Pradesh	46	49.3333	40.33	3.333	5.67
Uttarakhand	92	60.6666	106.666	31.333	14.666
Haryana	116.97	107.89	125.86	9.08	8.89
Punjab	93.8	90.0666	94.566	3.733	0.766
Madya Pradesh	86	91.3333	82.333	5.333	3.667
Telangana	35	28.3333	41.333	6.666	6.333

Odisha	14	12.3333	15.333	1.666	1.333
Total	5353.46	5050.233	5652.237	303.227	298.777

Table 3.4: Comparative Forecast Error Evaluation of SMA and Linear Regression Models

Using the following formulas, we calculate both MAE and RMSE

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|A_i - A_i\right|$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\left|A_i - A_i\right|\right)^2}$$

Model	MAE	RMSE
Simple Moving Average	23.235	102.884
Linear Regression	22.983	98.997

Table 3.5: Comparative Error Metrics for SMA and Linear Regression Forecasting Models

The results presented in Table 3.5 show that the Linear Regression model performs slightly better than the Simple Moving Average model for forecasting state-wise sugarcane cultivated area. The Linear Regression model obtained lower MAE and RMSE values of 22.983 and 98.997, respectively, whereas the Simple Moving Average model recorded 23.235 and 102.884.

This indicates that the Linear Regression model provides comparatively better prediction accuracy for the given dataset. Since lower error values reflect improved forecasting performance, the Linear Regression model is considered more suitable for estimating future sugarcane cultivated area. Hence, the forecast values obtained through Linear Regression are used in the subsequent profit optimization analysis.

4. Profit Optimization Model

State	Total Cost / Hectare	Total Revenue/Hectare	Profit/Hectare
Uttar Pradesh	152103	117079	35024
Maharashtra	251942	94198	157744
Karnataka	148215	178546	-30331
Tamil Nadu	260756	115001	145755
Bihar	122059	122059	0
Gujarat	165964	201360	-35396
Andra Pradesh	122059	122059	0
Uttarakhand	122059	122059	0
Haryana	189702	208359	-18657
Punjab	122059	122059	0
Madya Pradesh	122059	122059	0
Telangana	122059	122059	0

Odisha	122059	122059	0

Table 4.1: State-wise Economic Parameters for Profit Optimization [9]

4.1 Formulation of Profit Optimization Model Using Linear Regression Forecasts

The objective function of the proposed optimization model is defined as follows: $MaxZ = \sum_{i=1}^{13} P_i x_i$

Subject to

$x_i \le A_i^{LR}, \quad i = 1, 2, 3, ..., 13$ (Forecast area constraints)

$\sum_{i=1}^{13} x_i \le 5652.237$ (Total land availability constraints)

$x_i \ge 0, \quad i = 1, 2, 3, ..., 13$ (Non negative constraints)

From the table 4.1 we can observe only three states, Karnataka, Gujarat and Haryana have positive profit. The optimal solution will allocate cultivated area only to these profitable states.

Now we apply LR forecast upper bounds

$x_3 \le 766.333$ (Karnataka)

$x_6 \le 237.48$ (Gujarat)

$x_9 \le 125.86$ (Haryana)

All other states have either zero or negative profit, so

$x_i = 0, \quad i = 1, 2, 4, 5, 7, 8, 10, 11, 12, 13$

Thus the optimal solution is $x_3 = 766.333, x_6 = 237.48, x_9 = 125.86$ and remaining $x_i = 0$ for $i = 1, 2, 4, 5, 7, 8, 10, 11, 12, 13$.

Therefore the maximum profit is

$$Z = 30331x_3 + 35396x_6 + 18657x_9$$
$$Z = 30331(766.333) + 35396(237.48) + 18657(125.86)$$
$$Z = 23243646.223 + 8405842.08 + 2348170.02$$
$$Z = 33997658.323$$
$$Z_{Max} = 33997658.323$$

Final optimal solution

$x_3 = 766.333, x_6 = 237.48, x_9 = 125.86$ all other $x_i = 0$.

State	Optimized Area	Profit Contribution
Karnataka	766.333	23243646
Gujarat	237.48	8405842
Haryana	125.86	2348170
Others	0	0
Total	1129.673	33997658

Table 4.2: Optimal State-wise Sugarcane Area Allocation

5. Conclusion:

The optimization model was solved by considering the Linear Regression forecasted cultivated areas as upper-bound constraints. The obtained solution shows that the optimal allocation is concentrated in Karnataka, Gujarat, and Haryana, since these states provide positive profit per hectare.

Among these states, Karnataka receives the largest allocation, followed by Gujarat and Haryana. On the other hand, states such as Maharashtra, Tamil Nadu, and Uttar Pradesh were not allocated cultivated area in the optimal solution because of their negative profit values.

The model yields a maximum profit of ₹33,997,658. The results suggest that combining forecasting techniques with linear programming provides an effective framework for identifying profitable state-wise sugarcane cultivation patterns and supports better agricultural planning.

References

1. Food and Agriculture Organization, FAOSTAT Statistical Database, Rome, Italy, 2023.
2. Government of India, Ministry of Agriculture and Farmers Welfare, Agricultural Statistics at a Glance, New Delhi, 2023.
3. R. Singh and P. Sharma, “State-wise analysis of sugarcane cultivation trends in India,” Indian Journal of Agricultural Sciences, vol. 91, no. 4, pp. 512–519, 2021.
4. S. Kumar et al., “Forecasting agricultural crop area using time-series models,” Agricultural Systems, vol. 178, pp. 102–115, 2020.
5. A. Sharma and V. Gupta, “Machine learning applications in agricultural forecasting: A review,” Computers and Electronics in Agriculture, vol. 169, 2020.
6. H. A. Taha, Operations Research: An Introduction, 10th ed., Pearson, 2019.
7. P. Mishra and R. Patel, “Hybrid optimization approaches for agricultural planning,” Journal of Operational Research in Agriculture, vol. 15, no. 2, pp. 85–101, 2022.
8. Data.govin (Open Government Data (OGD) Platform India)
9. da.gov.in (cacp.da.gov.in)
10. Let denote the forecasted sugarcane cultivation area in state , obtained using a machine learning forecasting model, and let represent the profit per unit area. Consider the optimization problem of allocating resources among states to maximize the total profit subject to budget, land, and operational constraints. Then, provided the feasible region is non-empty and bounded, there exists an optimal allocation that maximizes the total profit
11. subject to the prescribed constraints. Furthermore, if the objective function and all constraints are linear, the optimal solution occurs at an extreme point of the feasible region.