Journal Homepage: -www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

ISSN (O): 2320-5407
ISSN (P): 3107-4928

Article DOI:10.21474/IJAR01/23777
DOI URL: http://dx.doi.org/10.21474/IJAR01/23777

RESEARCH ARTICLE

FORECASTING PRODUCTION PATTERN OF SUGAR, MOLASSES, BAGASSE AND PRESSMUD IN INDIA USING ARIMA MODEL

Asha D M [1] and Parvathamma G L[2]

1. Department of Economics, Government First Grade College, Patrakartanagara, Hubballi (Karnataka), INDIA.
2. Department of Economics, Bangalore University, Jnanabharathi Campus, Bangalore (Karnataka), INDIA.

Manuscript Info

Manuscript History
Received: 08 May 2026
Final Accepted: 10 June 2026
Published: July 2026

Key words:-
Sugarcane, Sugar, Bagasse, Molasses, Press mud, By-products

Abstract

The present paper attempts forecasting production pattern of sugar, molasses, bagasse and pressmud in India through fitting of Auto Regressive Integrated Moving Average (ARIMA) models. The data on production of sugar and its by-products collected from 1999-2019 has been used for present study. The performances of different ARIMA models are validated by standard statistical techniques and compared with actual values. Among the ARIMA models, ARIMA (2,1,1) model is found to be the best model and is used for forecasting the production pattern of sugar as well as by-products up to the period 2029-30.The model found that production of sugar, molasses, bagasse and pressmud would grow to 34.87, 16.39, 30.51 and11.86 million MTs in 2029-30 respectively. We also found that the production pattern of sugar, bagasse and pressmud follow cubic trend where as that of bagasse follows quadratic trend. The simultaneous forecasting of these four products provides an integrated framework for production planning and efficient utilization of sugar industry by-products. The findings offer valuable insights for policymakers, sugar mills, ethanol producers, and renewable energy planners in supporting long-term decision-making, resource optimization, and the promotion of a sustainable circular bioeconomy in India.These projections will help in making good policies with respect to the production scenario the country.

Introduction:-

Sugar industry is a second biggest agro based industry in India and has very strong impact on food security, energy security and economy of the country. It directly affects socio-economic growth of over 50 million Indians who are associated with sugarcane cultivation and sugar industry for their livelihood. Presently, about 5 million hectares of land is under sugarcane cultivation with average yield of 70 MT/ha. There are about 747 working sugar mills with an average crushing capacity of 35,000 tonne/day. India has produced 33 million MTs of sugar ie 19% of world sugar production of 179 million MTs. India has become the world's largest sugar producer as on 2018-19 beating out Brazil for the first time in last 16 years. India is also the largest consumer with annual consumption of 26 million MTs and third largest sugar exporter with worth of 17.6 million US dollars. Even though, India being third largest

Corresponding Author:- Dr. Asha D M
Address:- Associate Professor, Department of Economics, Government First Grade College, Patrakartanagara, Hubballi (Karnataka), INDIA.

sugar exporter and largest consumer of sugar in the world, around 40% of total sugar produced left unsold every year affecting demand and supply chain. Global competitions in the sale of sugar and rise in the cost of inputs made the profitability of the Indian sugar industry declining day by day. In the era of globalization and competition there is an urgent need for every Indian sugar industry to be cost-effective and competitive. This could be possible only through fuller and better utilization of the by-products through value addition process. Thus, diversification of sugarcane by-products of lower economic value in to higher value added products within the sugar industry would derives maximum benefit for the industry as well as farmers.

Today, modern technologies have opened new avenues of opportunities to sugar industries in the process of diversification of its various by-products molasses, bagasse and pressmud in to value added products of higher economic significance. Bagasse can be used in the production of power, craft paper, writing and printing paper, newsprint, tissue paper, cardboards and manufacture of furfural; molasses is used for production of alcohol by fermentation with yeast and chemicals like acetic acid, acetone, butyl alcohol, ethylene citric acid by suitable treatment. Biofuel and methane gas can also be generated from distillery effluent; press mud and spent ash have been used effectively in preparing bio-compost very useful organic manure which can be distributed back to the agriculture community. [2]

Recent years, sugarcane emerged as a multi-product crop and also an essential commodity; it has a significant role in determining the inflation rate, wages and various policies in Indian economy. Hence, a proper understating and forecasting the production pattern of sugar and its byproducts is vital for proper planning for the future. In this point of view, studying the production pattern of the past and assessment of the same in future play vital role. The present study focuses on the modeling and forecasting production pattern of sugar and its by-products viz., molasses, bagasse and pressmud in India using ARIMA model.

Research Gaps:-

In the arena of time series modelling, till date Autoregressive integrated moving average (ARIMA) model is a stand out amongst the most main stream approach for agriculture crop forecasting. In the contest of sugarcane, there are many reports of ARIMA based forecasting of cultivation area, production of sugarcane and sugar under different time scale are available in literature. The following are the research gaps

- **Limited integrated forecasting:** Most previous studies forecast only sugar production, while simultaneous forecasting of sugar, molasses, bagasse, and pressmud has received very little attention.
- **Lack of long-term production forecasts:** Few studies provide long-term forecasts to support strategic planning for ethanol production, renewable energy, and efficient utilization of sugar industry by-products.
- **Insufficient model validation:** Earlier studies often lack comprehensive validation using stationarity tests, residual diagnostics, and multiple forecast accuracy measures, reducing the reliability of their forecasts.

Objectives:-

The main objectives of our work are below:-

1. To forecast the production of sugar and its major by-products (molasses, bagasse, and pressmud) in India using the Auto Regressive Integrated Moving Average (ARIMA) model.
2. To evaluate the forecasting performance of the ARIMA model by comparing the predicted values with the actual production data and analyzing the forecasting errors.
3. To assess the future production trends of sugar and its by-products and provide insights for policy formulation, industrial planning, and sustainable utilization of sugar industry by-products in India.

Methodology:-

To develop models and subsequently to use the best fitted models to forecast the series for the years to come, data for the whole period are used for model building and model validation purpose. Descriptive statistics are useful to describe patterns and general trends in a data set. It includes numerical and graphic procedure to summarize a set of data in a clear and understandable way. To examine the nature of each series these have been subjected to different descriptive measures. Statistical measures used to describe the above series are mean, standard deviation, coefficient of variance and compound annual growth rate.

Compound Annual Growth Rates:-

Compound Annual Growth Rates were estimated to know the growth pattern on area production and productivity of sugarcane in India (Dhakre and Amod Sharma, 2010, Manoharan and Ramalakshmi, 2015). The adequacy of the model for the respective series was indicated by the coefficient of multiple determinations (R^2).

The growth model was specified as follows:-

$Y = a\, b^t$

Where,
Y = Dependent variable for which growth rate was estimated
a = intercept
b = regression coefficient
t = time variable

The compound growth rate (r) was obtained for the logarithmic form of the equation as given below:-

$Y = a\, b^t$
$\text{Log } Y = \log a + t \log b$

The per cent compound growth rate (r) was given as:-

A. r = [(Antilog of b) –1] × 100

Coefficient of Variation (CV):-

The coefficient of variation represents the ratio of the standard deviation to the mean, and it is a useful statistic for comparing the degree of variation from one data series to another, even if the means are drastically different from one another. It is calculated as follows:

$$\textit{Coefficient of variation } (\%) = \frac{\text{Standard Deviation}}{\text{Mean}} \times 100$$

Parametric Trend models:-

To find the production pattern, different parametric trend models as given in below are used. Among the competitive trend models, the best models are selected based on maximum value of R^2 value, minimum value of RMSE and significance of the parameters.
Different trend models used

Polynomial Model:-

$Y = b_0 + b_1 t + b_2 t^2 + b_3 t^3 + \dots\dots\dots + b_k t^k$

Linear Model:-

$Y_t = b_0 + b_1 t$

Quadratic Model:-

$Y_t = b_0 + b_1 t + b_2 t^2$

Cubic Model:-

$Y_t = b_0 + b_1 t + b_2 t^2 + b_2 t^3$

Exponential Model:-

$Y_t = b_0\, e^{(b1t)}$

Logarithmic Model:-

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

Growth Model:-

$Y_t = e^{b0+b1(t)}$

After the evaluation of trend of each and every series, our next task is to forecast the series for the year to come. For this purpose the study adopted the Box-Jenkins methodology. Data for the period of 20 years from 1999-2000 to 2018-2019 has been used for building forecast model, while data for the years 2009-2012 are taken for model validation. On the basis of the best fitted model forecasting has been made up to 2030.

ARIMA model:-

ARIMA models, stands for Autoregressive Integrated Moving Average models. Integrated means the trends has been removed; if the series has no significant trend, the models are known as ARMA models. The notation AR (p) refers to the autoregressive model of order p; the AR(p) model is written

$$X_t = c + \sum_{i=1}^{p} \alpha_i X_{t-i} + \mu_t$$

where α_1, α_2....α_p are the parameters of the maodel, c is a constant and μt is white noise i.e. μt ≈ WN(0,s2). Sometimes the constant term is omitted for simplicity.

Moving Average model (MA):-

The notation MA (q) refers to the moving average model of order q:

$$X_t = \mu + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t$$

Where the q1, q2 ..., q $_q$ are the parameters of the model, m is the expectation of Xt (often assumed to equal 0), and the et is the error term.

Validation of the goodness of fit of an ARIMA model was be developed according to the following steps:-

1) Evaluation of statistical significance of parameters by the usual comparison between the parameter value and the standard deviation of its estimate. For a test statistic that is valid only asymptotically, a parameter whose value exceeds twice its standard error can be considered significant.

2) Analysis of the ACF of residuals. In this step, residuals (εt) are considered as a new time series, and ACF and PACF are estimated to be sure that values at lag k >.0 are not statistically different from zero

Given a set of time series data, on can calculate the mean, variance, autocorrelation function (ACF), and partial autocorrelation function (PACF) of the time series. Based on the nature of ACF and PACF appropriate ARIMA models are worked out, but the final decision is made once the model is estimated and diagnosed. In this step one can see whether the chosen model fits the data reasonably well. One simple test of the chosen model is to see if the residuals estimated from this model are white noise; if they are, one can accept the particular fit; if not, one start the process a fresh.

Models are compared according to the minimum values of Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Square Error (MSE) and Mean Absolute Percentage Error (MAPE) and maximum value of Coefficient of determination (R2) and of course the significance of the coefficients of the models. Best fitted models are put under diagnostic checks through auto correlation function (ACF) and partial autocorrelation function (PACF) of the residuals. Once the model satisfies the requirement, are used for forecasting purpose. Furthermore residual diagnostics, including the Ljung–Box test for autocorrelation and residual normality assessment, have been performed. The results indicate that the residuals behave as white noise with no significant autocorrelation, confirming the adequacy of the selected ARIMA models.

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} \left| X_i - \hat{X}_i \right|}{n}$$

$$\mathrm{MAPE} = \frac{\sum_{i=1}^{n}\left|\frac{X_i - \hat{X}_i}{X_i}\right|}{n} \times 100$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left|\frac{X_i - \hat{X}_i}{X_i}\right|}{n} \times 100}$$

$$R^2 = \frac{\sum_{i=1}^{n}\left(\hat{X}_i - \overline{X}\right)^2}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$$

Where, Xi, X , Xˆ i are the value of the i[th] observation, mean and estimated value of the i[th] observation of the variable X. With the help of SPSS 25 computer package ARIMA models were estimated for the production of sugar and its by-products. On the basis of best fitted model forecasting has been made up to 2030.

Result and Discussion:-

The production trends of sugar and its major by-products—molasses, bagasse, and pressmud—were analysed using parametric trend models followed by Autoregressive Integrated Moving Average (ARIMA) models to obtain reliable forecasts. The preliminary trend analysis indicated that sugar, bagasse, and pressmud production were best represented by cubic functions, whereas molasses production followed a quadratic trend. Although the coefficients of determination (R^2) ranged from 0.494 to 0.554, all trend equations were statistically significant ($p < 0.01$), indicating the presence of nonlinear temporal changes in production over the study period. These deterministic trend models were employed only as exploratory tools to understand the long-term production behaviour before fitting stochastic ARIMA models for forecasting.

Prior to ARIMA model development, the annual production series were examined for stationarity and transformed through first-order differencing to eliminate non-stationarity. The model orders (p, d, q) were selected based on the autocorrelation function (ACF) and partial autocorrelation function (PACF) analyses together with model selection criteria. Residual diagnostic tests, including the Ljung–Box statistic, confirmed the absence of significant autocorrelation in the residuals, indicating that the selected models adequately captured the temporal dependence in the historical data. The residuals also exhibited approximate normality, supporting the suitability of the ARIMA framework for forecasting the production of sugar and its by-products.

Among the competing ARIMA specifications, the ARIMA (2,1,1) model consistently produced superior forecasting performance for sugar, molasses, bagasse and pressmud. The model exhibited the highest coefficient of determination (R^2) and comparatively lower prediction errors than the ARIMA (2,1,0) and ARIMA (2,1,2) alternatives. For sugar production, the selected model achieved an R^2 value of 0.679 with an RMSE of 3.612 million tonnes, MAE of 2.534 million tonnes and MAPE of 12.83%. Similarly, the ARIMA (2,1,1) model for molasses production yielded an R^2 of 0.696 with RMSE, MAE and MAPE values of 1.676, 1.161 and 12.86%, respectively. Comparable performance was observed for bagasse and pressmud, with R^2 values of 0.692 and prediction errors remaining within acceptable limits for annual agricultural production data. These results demonstrate that the selected models adequately explained nearly 70% of the variability in the production series while maintaining relatively low forecasting errors. The MAPE values of approximately 12–13% indicate good forecasting accuracy according to widely accepted forecasting performance criteria, suggesting that the developed models are suitable for medium-term planning purposes.

The comparison between observed and predicted values further demonstrates the capability of the ARIMA models in reproducing the historical production pattern. While relatively larger deviations were observed during years characterized by abnormal fluctuations in sugarcane production, such as 2003–04, 2004–05, 2006–07 and 2017–18, the model successfully captured the overall cyclical behaviour of the production series. These deviations are expected because sugar production in India is influenced by climatic variability, rainfall distribution, pest incidence, sugarcane acreage, government pricing policies, and market demand, factors that are not explicitly incorporated into univariate ARIMA models. Nevertheless, the close agreement between observed and predicted values during most years confirms the robustness of the forecasting model for long-term production assessment.

The forecasts generated for the period 2019–20 to 2029–30 indicate an overall increasing trend in the production of sugar and all three by-products. Sugar production is projected to increase from approximately 27.12 million tonnes during 2019–20 to nearly 34.87 million tonnes by 2029–30. Molasses production is expected to increase steadily from 12.83 to 16.39 million tonnes over the same period, reflecting the anticipated expansion of sugar processing and ethanol manufacturing. Bagasse production is forecast to increase from 24.33 to 30.51 million tonnes, while pressmud production is projected to rise from 9.46 to 11.86 million tonnes. The increasing availability of these by-products presents significant opportunities for expanding cogeneration of electricity, ethanol blending programmes, biofertilizer production and other value-added industrial applications.

The validation results indicate that the ARIMA model performed satisfactorily during relatively stable production periods but exhibited larger forecasting errors during years characterized by extreme fluctuations. Sugar production recorded considerable variability over the study period, ranging from 12.69 million MT in 2004–05 to 33.16 million MT in 2018–19. During normal production years, such as 2011–12 and 2013–14, the forecast closely matched the observed values, with prediction errors of only 0.09 and 0.93 million MT, respectively. However, substantial deviations were observed during abnormal production years. For instance, in 2003–04 the actual sugar production declined sharply to 13.55 million MT, whereas the ARIMA model predicted 21.04 million MT, resulting in an error of –7.50 million MT. Similarly, in 2017–18 the model underestimated production by 5.86 million MT due to unexpectedly high sugarcane output. These findings suggest that ARIMA effectively captures historical trends but haslimitations in predicting abrupt structural changes caused by monsoon variability, pest incidence, government pricing policies, and cyclical sugarcane cultivation.

A similar pattern was observed for molasses production. Since molasses is generated directly from sugar manufacturing, its production followed trends nearly identical to sugar output. Forecasting errors remained relatively small during stable years, with an error of only –0.02 million MT in 2011–12 and –0.05 million MT in 2013–14. However, years experiencing substantial reductions in sugar production also showed significant forecasting errors for molasses. The largest deviation occurred in 2003–04, when the actual production was only 5.91 million MT compared with the predicted 9.29 million MT, producing an error of –3.39 million MT. Likewise, in 2006–07 the model underestimated molasses production by 2.58 million MT because of the unprecedented increase in sugar output. These results confirm the strong dependence of molasses production on sugarcane crushing volumes and sugar recovery rates, indicating that accurate sugar forecasts directly improve the prediction of molasses availability.

Bagasse production exhibited trends closely associated with sugarcane processing volumes. Since bagasse accounts for the fibrous residue remaining after juice extraction, its production is influenced primarily by the quantity of sugarcane crushed rather than sugar recovery alone. The ARIMA model estimated bagasse production with reasonable precision during most years. For example, prediction errors were only 0.15 million MT in 2011–12 and 0.69 million MT in 2013–14. Nevertheless, larger deviations were observed during years of abnormal production, particularly in 2003–04, where the actual production was 11.93 million MT compared with the predicted value of 18.28 million MT, resulting in an error of –6.35 million MT. Similarly, the model underestimated production by 4.31 million MT in 2006–07 and by 4.09 million MT in 2017–18. These deviations reflect the inability of time-series models to anticipate sudden changes in sugarcane acreage and climatic conditions that significantly affect biomass availability.

Pressmud production also followed a trend similar to sugar and bagasse production because it is another important by-product generated during sugar clarification. The prediction errors were generally smaller than those observed for sugar and bagasse due to the relatively stable proportion of pressmud generated during processing. The smallest forecasting error was observed in 2011–12 (0.06 million MT), while larger deviations occurred in years with abnormal production, particularly 2003–04 (–2.47 million MT) and 2017–18 (1.59 million MT). Overall, the model adequately represented long-term production behaviour while slightly underperforming during years characterized by exceptional increases or decreases in sugarcane production.

The ARIMA forecasts for the period 2019–20 to 2029–30 indicate a generally positive outlook for India's sugar industry. Sugar production is projected to fluctuate within a relatively narrow range during the forecast horizon. Production is expected to remain around 27.12 million MT in 2019–20, decline slightly to 25.85 million MT in 2020–21, and subsequently recover to 30.69 million MT in 2021–22 and 33.51 million MT in 2022–23. Beyond 2024–25, production is forecast to stabilize above 30 million MT and eventually reach approximately 34.87 million MT by 2029–30. This steady upward trend reflects anticipated improvements in sugarcane productivity, adoption of high-yielding varieties, better irrigation facilities, and continued government support for the sugar industry. underestimated production by 4.31 million MT in 2006–07 and by 4.09 million MT in 2017–18. These deviations reflect the inability of time-series models to anticipate sudden changes in sugarcane acreage and climatic conditions that significantly affect biomass availability.

Pressmud production also followed a trend similar to sugar and bagasse production because it is another important by-product generated during sugar clarification. The prediction errors were generally smaller than those observed for sugar and bagasse due to the relatively stable proportion of pressmud generated during processing. The smallest forecasting error was observed in 2011–12 (0.06 million MT), while larger deviations occurred in years with abnormal production, particularly 2003–04 (–2.47 million MT) and 2017–18 (1.59 million MT). Overall, the model adequately represented long-term production behaviour while slightly underperforming during years characterized by exceptional increases or decreases in sugarcane production.

The ARIMA forecasts for the period 2019–20 to 2029–30 indicate a generally positive outlook for India's sugar industry. Sugar production is projected to fluctuate within a relatively narrow range during the forecast horizon. Production is expected to remain around 27.12 million MT in 2019–20, decline slightly to 25.85 million MT in 2020–21, and subsequently recover to 30.69 million MT in 2021–22 and 33.51 million MT in 2022–23. Beyond 2024–25, production is forecast to stabilize above 30 million MT and eventually reach approximately 34.87 million MT by 2029–30. This steady upward trend reflects anticipated improvements in sugarcane productivity, adoption of high-yielding varieties, better irrigation facilities, and continued government support for the sugar industry.

The projected increase in sugar production is accompanied by corresponding increases in all major by-products. Molasses production is expected to increase from 12.83 million MT in 2019–20 to approximately 16.39 million MT by 2029–30. Since molasses is the principal feedstock for ethanol production, this forecast suggests significant opportunities for expanding India's ethanol blending programme and reducing dependence on imported fossil fuels. The increasing availability of molasses will strengthen the biofuel sector while providing additional revenue streams for sugar mills.

Bagasse production is projected to increase steadily from 24.33 million MT in 2019–20 to 30.51 million MT in 2029–30. The growing availability of bagasse is particularly important for cogeneration of electricity and production of renewable energy. Many sugar mills in India have already adopted bagasse-based cogeneration plants, and the forecasted increase in bagasse availability indicates considerable potential for further expansion of green power generation. Enhanced utilization of bagasse can also contribute to the manufacture of paper, biodegradable packaging materials, compressed biofuels, and second-generation bioethanol, thereby promoting resource efficiency and circular bioeconomy principles.

Similarly, pressmud production is forecast to increase gradually from 9.46 million MT in 2019–20 to nearly 11.86 million MT by 2029–30. Pressmud has considerable economic value as an organic manure, composting material, and feedstock for biogas generation. The increasing availability of pressmud supports sustainable agricultural practices by improving soil fertility and reducing dependence on chemical fertilizers. Consequently, the forecasted growth in pressmud production complements national efforts toward environmentally sustainable agriculture and integrated waste utilization.

The forecast results suggest a gradual increase in the production of sugar and its by-products over the coming years, reflecting positive prospects for the Indian sugar industry. Increased availability of molasses will support the expansion of ethanol production and the national biofuel programme, while higher bagasse production will enhance renewable energy generation through cogeneration. Similarly, the projected increase in pressmud production offers opportunities for sustainable organic manure production and improved soil health. The forecast results have important practical implications for India's sugar industry. Simultaneous forecasting of sugar, molasses, bagasse and pressmud provides a comprehensive planning framework that is not commonly available in previous forecasting studies, which have generally focused only on sugar production. The integrated forecasts enable sugar mills to optimize processing capacity, schedule maintenance activities, plan storage infrastructure and estimate future availability of by-products for ethanol distilleries, biomass-based power plants and organic manure production. The projected increase in bagasse production supports further expansion of renewable energy generation through cogeneration systems, while increasing molasses availability aligns well with India's ethanol blending programme aimed at reducing dependence on fossil fuels. Likewise, greater pressmud production can contribute to sustainable nutrient recycling and improved soil fertility under circular bioeconomy initiatives.

Table 1: Production performance of sugar industry in India

Year	No of Working Factories	Average Working Days/Year	Average Capacity (MTs/day)	Quantity of Sugarcane Crushed (Million MTs)	Production of Sugar and By-products (Million MTs)					
					Sugar	Sugar Yield	Molasses	Molasses Yield	Bagasse*	Pressmud*
1999-00	423	151	3049	178.494	18.200	10.20	8.013	4.49	16.064	6.247
2000-01	436	138	3203	176.660	18.511	10.48	7.820	4.43	15.899	6.183
2001-02	434	138	3285	180.346	18.528	10.27	8.073	4.48	16.231	6.312
2002-03	453	140	3343	194.365	20.145	10.36	8.879	4.57	17.493	6.803
2003-04	422	99	3493	132.511	13.546	10.22	5.905	4.46	11.926	4.638
2004-05	400	97	3508	124.772	12.690	10.17	5.513	4.42	11.229	4.367
2005-06	455	125	3619	188.672	19.267	10.21	8.549	4.53	16.980	6.604
2006-07	504	173	3561	279.295	28.328	10.16	13.111	4.69	25.137	9.775
2007-08	516	149	3586	249.906	26.357	10.55	11.313	4.53	22.492	8.747
2008-09	489	87	3751	144.983	14.539	10.03	6.546	4.52	13.048	5.074
2009-10	490	109	3846	185.548	18.912	10.19	8.400	4.53	16.699	6.494
2010-11	527	135	3744	239.807	24.394	10.17	10.970	4.57	21.583	8.393
2011-12	529	137	3877	256.975	26.343	10.25	11.824	4.60	23.128	8.994
2012-13	526	127	4119	250.598	25.141	10.03	11.745	4.69	22.554	8.771
2013-14	513	115	4061	238.176	24.360	10.23	10.850	4.56	21.436	8.336
2014-15	538	133	4117	273.073	28.313	10.37	12.479	4.57	24.577	9.558
2015-16	526	117	3843	236.498	25.125	10.62	10.885	4.60	21.285	8.277
2016-17	489	100	4135	193.434	20.262	10.48	9.002	4.65	17.409	6.770
2017-18	525	141	4439	301.198	32.328	10.73	13.980	4.64	27.108	10.541
2018-19	532	143	4521	312.122	33.161	10.62	14.482	4.64	28.091	10.924
Mean	**486.35**	**127.70**	**3755.00**	**216.87**	**22.42**	**10.32**	**9.92**	**4.56**	**19.52**	**7.59**
Std Dev	**44.59**	**21.54**	**397.05**	**54.51**	**5.86**	**0.20**	**2.62**	**0.08**	**4.91**	**1.91**
CV	**9.17**	**16.87**	**10.57**	**25.14**	**26.13**	**1.94**	**26.41**	**1.75**	**25.14**	**25.14**
CAGR (%)	**1.21**	**-0.29**	**2.09**	**2.98**	**3.21**	**0.21%**	**3.16**	**0.17**	**2.99**	**2.99**

Table 2: Parametric trend models for production of sugar and its by-products in India

Parameter	Equation	Intercept	Regression coefficients			R^2	F-ratio	p-value
			b_1	b_2	b_3			
Sugar Production	Cubic	18.389	-0.461	0.101	-0.003	0.494	5.526	0.008
Molasses Production	Quadratic	7.523	0.076	0.11	-	0.554	11.169	0.001
Bagasse Production	Cubic	15.475	-0.085	0.052	-0.01	0.506	5.812	0.006
Pressmud Production	Cubic	6.018	-0.033	0.020	000	0.506	5.811	0.006

Table 3: Diagnostics of ARIMA model selection for the production of sugar and its by-products in India

	ARIMA Model	No. of Predictors	Model Fit statistics							No of Outliers
			R^2	RMSE	MAPE	MAE	Max APE	Max AE	BIC	
Sugar Production	(2,1,0)	0	0.500	4.374	14.300	3.037	55.237	8.646	3.401	0
	(2,1,1)	0	0.679	3.612	12.830	2.534	55.338	7.496	3.168	0
	(2,1,2)	0	0.662	3.831	13.053	2.633	55.430	7.509	3.435	0
Molasses Production	(2,1,0)	0	0.530	2.020	14.916	1.453	57.094	3.810	1.855	0
	(2,1,1)	0	0.696	1.676	12.862	1.161	57.336	3.386	1.631	0
	(2,1,2)	0	0.673	1.794	13.076	1.205	56.805	3.354	1.917	0
Bagasse Production	(2,1,0)	0	0.525	3.651	13.652	2.575	52.224	6.920	3.040	0
	(2,1,1)	0	0.692	3.028	12.277	2.155	53.273	6.353	2.815	0
	(2,1,2)	0	0.669	3.242	11.584	2.109	52.217	6.503	3.102	0
Pressmud Production	(2,1,0)	0	0.525	1.420	13.654	1.001	52.225	2.690	1.150	0
	(2,1,1)	0	0.692	1.177	12.276	0.838	53.271	2.471	0.926	0
	(2,1,2)	0	0.669	1.261	11.581	0.820	52.221	2.526	1.212	0

Table 4: Forecasted values for production of sugar and its by-products in India

Year	Sugar Production (Million MTs) *			Molasses Production (Million MTs) *			Bagasse Production (Million MTs) *			Pressmud Production (Million MTs) *		
	A	P	E	A	P	E	A	P	E	A	P	E
1999-00	18.20	-	-	8.01	-	-	16.06	-	-	6.25	-	-
2000-01	18.51	18.85	-0.34	7.82	8.33	-0.51	15.90	16.62	-0.72	6.18	6.46	-0.28
2001-02	18.53	19.17	-0.64	8.07	8.15	-0.08	16.23	16.47	-0.24	6.31	6.40	-0.09
2002-03	20.15	19.45	0.70	8.88	8.70	0.18	17.49	17.26	0.23	6.80	6.71	0.09
2003-04	13.55	21.04	-7.50	5.91	9.29	-3.39	11.93	18.28	-6.35	4.64	7.11	-2.47
2004-05	12.69	17.03	-4.34	5.51	7.24	-1.73	11.23	14.61	-3.38	4.37	5.68	-1.31
2005-06	19.27	21.13	-1.86	8.55	8.92	-0.37	16.98	17.96	-0.98	6.60	6.99	-0.38
2006-07	28.33	23.95	4.38	13.11	10.54	2.58	25.14	20.83	4.31	9.78	8.10	1.68
2007-08	26.36	23.50	2.85	11.31	11.07	0.25	22.49	21.14	1.35	8.75	8.22	0.53
2008-09	14.54	18.28	-3.74	6.55	8.12	-1.58	13.05	15.94	-2.89	5.07	6.20	-1.12
2009-10	18.91	17.02	1.89	8.40	7.77	0.63	16.70	14.99	1.71	6.49	5.83	0.66
2010-11	24.39	27.10	-2.70	10.97	11.75	-0.78	21.58	23.18	-1.60	8.39	9.01	-0.62
2011-12	26.34	26.25	0.09	11.82	11.84	-0.02	23.13	22.98	0.15	8.99	8.94	0.06
2012-13	25.14	24.03	1.11	11.75	10.98	0.77	22.55	21.07	1.48	8.77	8.20	0.58
2013-14	24.36	23.43	0.93	10.85	10.90	-0.05	21.44	20.74	0.69	8.34	8.07	0.27
2014-15	28.31	25.01	3.30	12.48	11.03	1.45	24.58	21.54	3.04	9.56	8.38	1.18
2015-16	25.13	27.69	-2.57	10.89	12.63	-1.74	21.29	24.22	-2.93	8.28	9.42	-1.14
2016-17	20.26	24.99	-4.73	9.00	11.37	-2.37	17.41	21.67	-4.26	6.77	8.43	-1.66
2017-18	32.33	26.47	5.86	13.98	11.93	2.05	27.11	23.02	4.09	10.54	8.95	1.59

2018-19	33.16	34.05	-0.89	14.48	15.31	-0.82	28.09	29.62	-1.53	10.92	11.52	-0.60
2019-20		27.12			12.83			24.33			9.46	
2020-21		25.85			13.11			23.65			9.20	
2021-22		30.69			12.79			25.47			9.90	
2022-23		33.51			14.06			27.98			10.88	
2023-24		31.93			15.08			28.35			11.03	
2024-25		30.56			15.07			27.59			10.73	
2025-26		32.17			14.84			27.83			10.82	
2026-27		34.40			15.17			29.14			11.33	
2027-28		34.82			15.81			30.14			11.72	
2028-29		34.35			16.23			30.35			11.80	
2029-30		34.87			16.39			30.51			11.86	

*** A= Actual, P=Predicted, E=Error**

Conclusion:-

The present study demonstrates the usefulness of the ARIMA model in forecasting the production of sugar and its major by-products, namely molasses, bagasse, and pressmud, in India. The model successfully captured the long-term production trends and generated reliable forecasts despite the inherent variability in agricultural production. Although larger prediction errors were observed during years affected by climatic fluctuations and unexpected production shocks, the overall forecasting performance was satisfactory, indicating the suitability of ARIMA for medium- and long-term production planning.

Although the ARIMA model demonstrated satisfactory predictive performance, forecast uncertainty inevitably increases with the length of the forecasting horizon. Accordingly, the forecasts should be interpreted as statistically derived projections under the assumption that historical production dynamics continue. Unexpected changes in climate, extreme weather events, technological innovations, government policies on sugar pricing and ethanol blending, export-import regulations, and global market conditions may influence future production beyond the range captured by historical data. Therefore, the 95% prediction intervals accompanying the forecasts should be considered when interpreting future production estimates. Periodic updating of the forecasting model using newly available observations and the adoption of rolling forecasting strategies would further enhance prediction accuracy.

Overall, the study demonstrates that the ARIMA (2,1,1) model provides a statistically reliable and practically useful framework for forecasting sugar and its major by-products in India. The integrated forecasting approach contributes to improved production planning, supply chain management, renewable energy development, resource optimization and policy formulation. Furthermore, the findings provide a valuable baseline against which future studies employing hybrid ARIMA models, exponential smoothing techniques, machine learning algorithms and deep learning approaches such as Long Short-Term Memory (LSTM) networks can be evaluated, thereby strengthening future research on agricultural production forecasting.

References:-

1. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time Series Analysis: Forecasting and Control (5th ed.). John Wiley & Sons.
2. Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and Practice (3rd ed.). OTexts.
3. Indian Sugar Mills Association. (2023). Indian Sugar Industry Statistics and Annual Report**.** New Delhi, India.
4. Ahmar, A. S., Rahman, A., & Mulbar, U. (2018). Forecasting of agricultural production using ARIMA model**.** Journal of Physics: Conference Series**,** 954(1), 012001.
5. Solomon, S. Sugarcane production and development of sugar industry in India. Sugar Tech, 2016, 18(6): 588–602.
6. Ahmar, A. S., Guritno, S. and Rahman, A. Modeling data containing outliers using ARIMA additive outlier (ARIMA-AO). arXiv Preprint, arXiv: 2018,1803.00257.

7. Food and Agriculture Organization (FAO). (2023). FAOSTAT Statistical Database. Food and Agriculture Organization of the United Nations, Rome.
8. Directorate of Sugar and Vegetable Oils. (2023). Annual Report on Sugar Production in India. Ministry of Consumer Affairs, Food and Public Distribution, Government of India.
9. Indian Sugar Mills Association (ISMA). (2023). Indian Sugar Industry Statistics and Annual Report. New Delhi.
10. NITI Aayog. (2021). Roadmap for Ethanol Blending in India 2020–2025. Government of India, New Delhi.
11. Department of Food and Public Distribution. (2023). Sugar Sector Performance and Production Statistics. Ministry of Consumer Affairs, Food and Public Distribution, Government of India.