Journal of Sugarcane Research (2014) 4 (2) : 64 - 69

RESEARCH ARTICLE

FORECASTING PRODUCTION OF SUGARCANE IN INDIA USING ARIMA MODELS

K. Prakash* and B. Muniyandi

Abstract

This study aims at forecasting sugarcane production in India, based on secondary data during the years from 1970-71 to 2012-13 (43 years). The study considered Autoregressive Integrated Moving Average (ARIMA) model. The selection of ARIMA (2, 1, 0) the best Model is found suitable for sugarcane production. The total cropped production can be increased in future but there has been some fluctuations in the country. The ARIMA model will be useful to project the sugarcane growth in the future period.

Key words: ARIMA Model, sugarcane, production, forecasting, agricultural economics.

Introduction

India is the second largest producer of sugarcane and cane sugar, after Brazil (Hari et al. 2013). Sugarcane, the second most important cash crop in India, is cultivated as a monocrop in large tracts under the supervision of sugar industry. The spatial and temporal continuity of the crop accords the status of what can be termed as semi-perennial habitat (Srikanth et al. 2013). The crop occupies a signiûcant position among the commercial crops in India. According to estimates of the National Commission on Agriculture (1976) and estimations by various agencies, the population of the country is expected to swell to 1.5 billion by 2030 AD (Yadhav 2011) at the present compound growth rate of 1.6 per cent per annum. It is estimated that the per capita consumption is likely to increase and may go up to 35 kg (both white sugar and *gur*) by 2030 AD. At this rate of consumption and expected rise in population, the country may require nearly 52 million tonnes of sweeteners by 2030 AD. With decreasing trends in *gur* and *khandsari* production, the demand for white sugar is likely to increase to 33 million tonnes by 2030 AD. The emerging energy need of ethanol for blending in petrol will require additional sugarcane over and above the cane requirement to produce 33 million tonnes of sugar. In order to meet the growing demand of sugar and energy by 2030 AD in India, around 520 million tonnes of sugarcane with a recovery of 10.75 per cent will be required (312 million tonnes of cane exclusively for white sugar and additional 78 million tonnes of cane exclusively for ethanol production) (www.iisr.nic.in).

There is close association between crop production and prices. An unexpected decrease in production reduces marketable surplus and income of the farmers and leads to price rise. A glut in production can lead to a slump in prices and has adverse effect on farmers' income. Impact on price of an essential commodity has a signiûcant role in determining the inflation rate, wages, salaries and various policies in an economy. In case of commercial crops like sugarcane, production level affects raw material cost of user industries and their competitive advantages

K. Prakash* and B. Muniyandi
Department of Economics, School of Economics, Bharathiar University, Coimbatore – 641 046 Tamil Nadu, India.
*email : prakashkaruva@gmail.com

in the market (Suresh and Krishna Priya 2011). A proper forecast of its production is very important in an economic system. There has been a recent upswing of interest using large datasets for macroeconomic forecasting. An increasing number of time series data describing the state of the economy are available that could be used for forecasting. Also, computational power to handle large amounts of data has steadily risen over time. Thus researchers now attempt to improve their forecasting models by exploiting a broader information base (Klaus Wohlrabe and Teresa Buchen 2014). In the present study, sugarcane production of India has been forecasted using Auto Regressive Integrated Moving Average (ARIMA) models.

Two main theories have been shaping the thinking in economic forecasting. The traditional optimality theory on economic forecasting has grounds in two key assumptions, that the model is a good representation of the economy and that the structure of the economy remains unchanged also in the future. Given these assumptions, several theorems can be proved. In particular, it can be shown that the best in-sample model produces the best forecasts. However, the empirical evidence has undermined the relevance of the two assumptions of the traditional theory together with all the related theorems. This is how the new theory on forecasting emerged, relying on the assumptions that the models are simplified representations and incorrect in many ways, and that economies are subject to sudden shifts (Klara Stovicek, 2007). The objectives of the study were to suggest appropriate ARIMA model for the generation of forecasting production of sugarcane in India and to make ten years forecasts with appropriate prediction interval and to generate forecasts of production of sugarcane in India by using appropriate ARIMA models.

Materials and methods

ARIMA model

The Auto Regressive Integrated Moving Average **(ARIMA)** model is a generalization of an autoregressive moving average (ARMA) model. These models are fitted to time series data either to better understand the data or to predict future points in the series. The existing study applies Box-Jenkins (1970) forecasting model popularly known as ARIMA model. The ARIMA is an extrapolation method, which requires historical time series data of underlying variable and generally ARIMA model is used in macro level data analysis. The model in specific and general forms may be expressed as follows. Let Y_t is a discrete time series variable which takes different values over a period of time. The corresponding AR (p) model of Y_t series, which is the generalizations of autoregressive model, can be expressed as:

$$AR(p): Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon t \quad (1)$$

Where, Y_t is the response variables at time t, Y_{t-1}, Y_{t-2}……Y_{t-p} are the respective variables at different time with lags; $\phi_0, \phi_1, \ldots, \phi_p$ are the coefficients and εt is the error factor. Similarly, the MA (q) model which is again the generalizations of moving average model may be specified as:

MA (q): $$Y_t = \mu_t + \varepsilon_t + \delta_1 \varepsilon_{t-1} + \cdots + \delta_q \varepsilon_{t-q} + vt \quad (2)$$

Where, μ_t is the constant mean of the series, $\delta_1 \ldots \delta_q$ are the coefficients of the estimated error term, is the error term. Combining both the models yields what is called as ARIMA models, which has the general form as:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t + \delta_1 \varepsilon_{t-1} + \cdots + \delta_q \varepsilon_{t-q} + vt$$

If Y_t is stationary at level or I (0) or at first difference I (1) determines the order of integration, which is called as ARIMA model. To identify the order of *p* and *q* the ACF and PACF are applied.

Data

For the present study, production (000'tonnes) data were obtained from secondary sources like *Cooperative Sugar* for the periods of 1970-71 to 2012-13 (43 years). Statistical tools were employed to assess the growth performance of Sugarcane in India as large numbers of data are required for ARIMA model. Model Statistics tools were employed to analyze the data with reference to selected objectives of the study.

Results and discussion

The ARIMA model was formulated after transforming the variable under forecasting in to stationary series. The stationary series was the set of values that varied over time around a constant mean and constant variance. This model is a common method to check the stationary nature of data. Fig. 1 reveals that the data used were non-stationary. Hence, at the first stage non-stationary data in mean was corrected through differencing of the data for examining the production and time variable (Y_t).

Since Y_t was stationary in mean, the next step was to identify the values of p and q. For this, the autocorrelation and partial autocorrelation coefficients (ACF and PACF) of various orders of Y_t were computed and presented in Table 1 and Fig. 2.

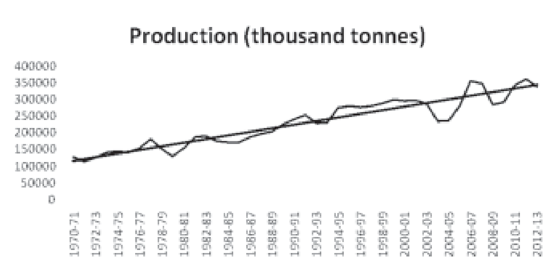

Fig. 1. Time plot of sugarcane production in India

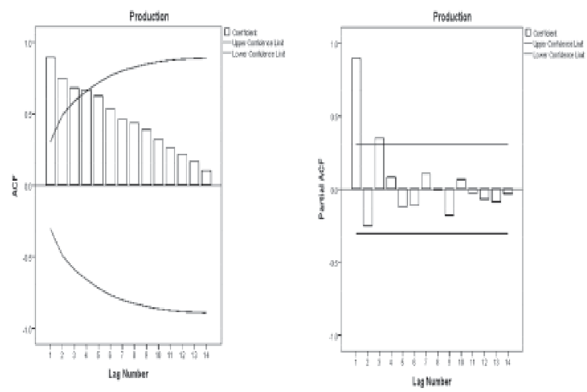

Fig. 2. ACF and PACF of differenced data

The order of an ARIMA model is usually denoted by the notation ARIMA (p,d,q), where p is the order of the autoregressive part, d is the order of the differencing and q is the order of the moving-average process.

The ARIMA model was discussed with values differenced once (d=1) and the model which had the minimum normalized Bayesian information criterion (BIC) was chosen. The BIC value was used to determine the autoregressive order to estimate the error series. The estimation of parameters for sugarcane production was estimated in Best Fitted Model. Based on four ARIMA models and the corresponding normalized BIC values the value of normalized BIC of the chosen ARIMA was 19.947.

Estimation of parameters for sugarcane, production of best fitted models

Model estimation

The second step was the estimation of model parameters using SPSS.20 version and the results estimated are presented in Tables 2 and 3. R^2 value was 0.94. Hence, the most suitable model for sugarcane production was ARIMA (2,1,0), as this model had the lowest normalized BIC value, good R^2 and better model fit statics (RMASE and

Table 1. ACF and PACF of sugarcane production

Lag	Auto Correlation		Box-Ljung Statistics			Partial Auto Correlation	
	Value	DF	SIG	Value	DIF	Partial Autocorrelation	Std. Error
1	0.894	0.152	36.854	1	0.000	0.894	0.152
2	0.750	0.246	63.397	2	0.000	-0.250	0.152
3	0.681	0.294	85.804	3	0.000	0.350	0.152
4	0.666	0.329	107.836	4	0.000	0.084	0.152
5	0.626	0.359	127.777	5	0.000	-0.121	0.152
6	0.535	0.383	142.769	6	0.000	-0.111	0.152
7	0.464	0.400	154.362	7	0.000	0.109	0.152
8	0.436	0.413	164.880	8	0.000	-0.008	0.152
9	0.390	0.423	173.537	9	0.000	-0.179	0.152
10	0.323	0.432	179.661	10	0.000	0.067	0.152
11	0.262	0.437	183.798	11	0.000	-0.030	0.152
12	0.218	0.441	186.777	12	0.000	-0.070	0.152
13	0.170	0.443	188.641	13	0.000	-0.087	0.152
14	0.102	0.445	189.328	14	0.000	-0.032	0.152

MAPE). It is justified that the selection of ARIMA (2,1,0) was the best model to represent the data generating process very precisely.

Table 2. Estimated ARIMA Model of sugarcane production

Constant	Estimate	SE	t	Sig
	4657.915	4352.057	1.070	.291

Diagnostic checking

In this model it is proved that the verification was concerned with checking the residuals of the model to see if they contained any systematic pattern which still could be removed to improve the chosen ARIMA, which has been done through examining the autocorrelations and partial autocorrelations of the residuals of various orders. For this purpose, various autocorrelations up to 10 lags and were computed (Table 4, Fig. 3) and the same along with their significance was tested by Box-Ljung statistic. The results indicated that none of these autocorrelations was significantly different from zero at any reasonable level. The selected ARIMA model was a suitable model for forecasting sugarcane production in India.

Table 3. Estimated ARIMA Model Fit Statistics

Stationary R-squared	R-squared	RMSE	MAPE	Normalized BIC
527	.941	17949.224	5.053	19.947

Table 4. Residual of ACF and PACF of sugarcane production

	ACF		PACF	
Lag	**Mean**	**SE**	**Mean**	**SE**
1	-0.101	0.154	-0.101	0.154
2	-0.069	0.156	-0.080	0.154
3	-0.110	0.157	-0.127	0.154
4	-0.198	0.158	-0.238	0.154
5	0.193	0.164	0.127	0.154
6	-0.191	0.170	-0.228	0.154
7	0.085	0.175	0.022	0.154
8	0.066	0.176	0.033	0.154
9	-0.185	0.176	-0.181	0.154
10	0.100	0.181	-0.009	0.154

Forecasts of sugarcane

The ten year forecast of sugarcane production was estimated by using the best model of ARIMA and (Table 5, Fig. 4). It has predicted that sugarcane production would increase from 326.626 million tonnes in the year 2013-14 to 392.919 million tonnes in 2022-23.

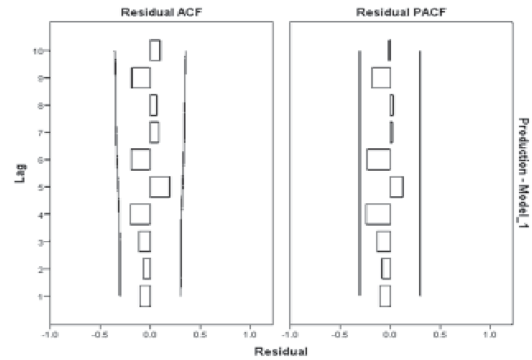

Fig. 3. ACF and PACF plot of residuals

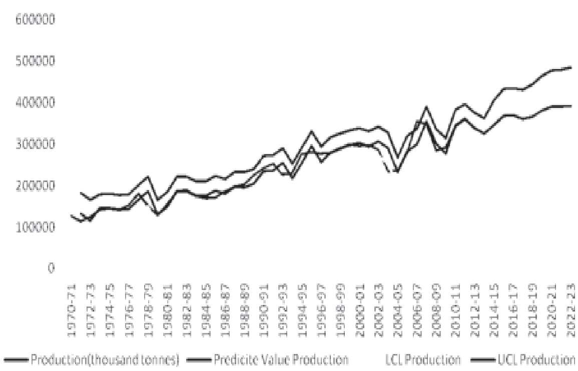

Fig. 4. Actual and estimate of sugarcane production

Conclusion

The study concludes that the total sugarcane production can be increased in the future, though

Table 5. Forecast for the production of sugarcane in India (thousand tonnes)

S.No.	Year	Predicted	LCL	UCL
1	2013-14	326626	290598	362654
2	2014-15	346248	286823	405672
3	2015-16	368952	304406	433497
4	2016-17	369973	304602	435344
5	2017-18	362078	293300	430857
6	2018-19	366817	289489	444146
7	2019-20	381848	298207	465490
8	2020-21	391154	304944	477364
9	2021-22	391402	302920	479884
10	2022-23	392919	300286	485552

there can be some fluctuations in the production of sugarcane in India. The projection of ARIMA model shows that sugarcane will play a vital role in improving the sugar and by-products production in the coming years in the country.

References

Hari K, Reginold Jebitta S, Sivaraman K (2013) Production and characterization of sugarcane juice powder J Sugarcane Res 3(1) 20-34.

Klara Stovicek (2007) Forecasting with ARMA Models the case of Slovenian inflation, Prikazi in analyze Ljubljana, 14 (1), 23-55.

Klaus Wohlrabe, Teresa Buchen (2014) Assessing the macroeconomic forecasting performance of boosting: Evidence for the United States, the Euro Area and Germany J Forecasting0, 33, 231–242.

Srikanth J, Sivaraman K, Kurup NK, Chandrasekhar SD, Sundara B, Rakkiyappan P, Hari K, Ramesh Sundar A, Sankaranarayanan C (2013) Pest scenario in long-term organic and conventional sugarcane production systems J Sugarcane Res 3(1) 47-61.

Suresh KK, Krishna Priya SR, (2011) Forecasting Sugarcane yield of Tamil Nadu using ARIMA models Sugar Tech 13(1) 23–26.

Yadhav RL (2011) Vision 2030, Indian Institute of Sugarcane Research, Lucknow, pp.1-28. Source: www.iisr.nic.in.