Economic Affairs, Vol. **69**(03), pp. 1199-1205, September 2024
DOI: 10.46852/0424-2513.4.2024.2

AESSRA

RESEARCH PAPER

Statistical Modeling and Trend Analysis of Sugarcane Yield in the States of India

Manish Kumar*, Shiv Kumar Rana and S. Vivek Menon

Department of Agricultural Statistics, Acharya Narendra Deva University of Agriculture & Technology, Ayodhya, India

*Corresponding author: manishstats88@gmail.com (**ORCID ID:** 0000-0002-4989-1612)

Received: 05-06-2024 **Revised:** 04-08-2024 **Accepted:** 22-08-2024

ABSTRACT

In this paper, a time series data on sugarcane yield in some selected states of India is considered to determine the growth and trend patterns by fitting well-known statistical models, viz. linear model and exponential model. The trend values have been obtained by fitting the respective models. Moreover, the "goodness of fit" of the respective models has been tested by using Chi-square test statistic. It is demonstrated from the study that both the models are valid for analyzing the growth and trend patterns of the sugarcane yield. Furthermore, the coefficients of determination (R^2), for the fitted models, are also computed to judge their suitability for exploring the trend patterns of sugarcane yield in the concerned states of India.

HIGHLIGHTS

- The growth and trend patterns of sugarcane yield are analyzed for the selected states of India.
- The present study is based on secondary time series data pertaining to the period 2010-2011 to 2019-2020.
- The study revealed that the linear and exponential models are suitable for examining the trend pattern of sugarcane yield in the concerned states of India.

Keywords: Time series modeling, Linear model, Exponential model, Chi-square test, Coefficient of determination

Sugarcane (*Saccharum officinarum*), a member of the Poaceae family, is one of the significant commercial crops, which is grown in large-scale in various tropical and sub-tropical states of India. Sugarcane is a rich source of vital nutrients such as potassium, calcium, magnesium, and other multi vitamins and amino acids. It is mainly used for the production of sugar, jaggery, ethanol, and other products. The health benefits of jaggery have been described by Rao and Singh (2021). Sugar industry is one of the largest agro-based industries which provides lot of employment opportunities to skilled and semi-skilled workers, and hence contributes in the economic growth of the nation. Roy and Chandra (2019) described some useful technologies for enhancing the income of sugarcane farmers, and suggested policy interventions for the same.

In India, Uttar Pradesh is the leading state in the production of sugarcane (178.42 million tons), followed by Maharashtra (64.67 million tons), Karnataka (31.60 million tons), Bihar (12.74 million tons), Tamil Nadu (12.66 million tons), and others including Andhra Pradesh (6.77 million tons). In India, the sugarcane yield was 77.89 tons per hectare during the year 2019-20 [Source: Directorate of Economics & Statistics, DAC&FW, Govt. of India].

Time series analysis of agricultural crops is of utmost importance for examining the long-term existing trends, and in forecasting of future trends for the

How to cite this article: Kumar, M., Rana, S.K. and Menon, S.V. (2024). Statistical Modeling and Trend Analysis of Sugarcane Yield in the States of India. *Econ. Aff.*, **69**(03): 1199-1205.

Source of Support: None; **Conflict of Interest:** None

observations on area, production, and productivity of the selected crops. In time series analysis, an adequate statistical model is fitted to the time series data in order to obtain precise predicted trend values. The various models used for the analysis are: linear model, exponential model, logistic model, and so on. The predicted values obtained on fitting the desired models provide a basis for inventory management and policy formulation for the crops under consideration.

In the past as well as in recent times, statistical analysis of sugarcane production have been made by various scientists and researchers, for instance, Suresh and Priya (2011) utilized univariate Auto Regressive Integrates Moving Average (ARIMA) models for forecasting the sugarcane area, production and productivity of Tamilnadu state of India. Sanjeev *et al.* (2015) fitted ARIMA models to the time series data on sugarcane yield for various districts of Haryana, and suggested the relevance of ARIMA models for short-term forecast. Arti and Rai (2017) utilized the secondary time series data on sugarcane (*Saccharum officinarum*) in Uttar Pradesh and revealed the percentage change and compound annual growth rate (CAGR) of sugarcane in terms of area, production and yield. Rahman and Bee (2019) assessed the trend and pattern of sugarcane production in Shahjahanpur district of Uttar Pradesh in India. Lal *et al.* (2020) analyzed the trend and growth rates of sugarcane production in the districts of western Uttar Pradesh in India. Mishra *et al.* (2021) utilized ARIMA model for forecasting of sugarcane production in various states of India. Singh *et al.* (2021a) analyzed the trend and change point detection of sugarcane production in India by applying several well-known non-parametric methods, viz. Pettitt's test, Standard normal homogeneity test and Buishand's range test. Singh *et al.* (2021b) estimated spatial compound growth rates to determine the growth pattern and instability in the area, production and productivity of sugarcane in major sugarcane growing states of India.

The time series analysis of agricultural crops, other than sugarcane, has been dealt by various authors, for instance, Rajarathinam *et al.* (2010) investigated the area, production and productivity trends, and growth rates of tobacco (*Nicotiana tabacum*) crop based on parametric and non-parametric regression models. Ramteke *et al.* (2015) examined the changes in soybean scenario in Madhya Pradesh with respect to the area sown, production and productivity, taking into consideration the changes in climatic variables like rainfall and temperature during the period of study. Dharmaraja *et al.* (2020) carried out an empirical analysis for predicting bajra crop yield in Alwar district of Rajasthan by incorporating linear regression and other time series models. Saha *et al.* (2021) analyzed the growth rates of area, production, and yield of tea in Bangladesh before independence (1947–1970) and after independence (1972–2018) by using exponential growth model. Lal and Rohtas (2022) examined compound annual growth rates (CGRs) for exploring growth patterns in area, production and productivity of selected spices and crops in Haryana state of India.

The objective of the present study is to examine the growth and trend patterns of sugarcane yield in selected states of India. For the present analysis, a secondary time series data is used, and trend values have been obtained by fitting linear and exponential models to the concerned data. Moreover, the validity and suitability of the respective models have been tested by using Chi-square test statistics and coefficients of determination (R^2).

DATA AND METHODOLOGY

Source of Data

For the purpose of present study, a secondary time series data on sugarcane yield for selected states of India is considered, pertaining to the period 2010 to 2019. The time series data is obtained from the records of Agricultural Statistics at a Glance, Directorate of Economics & Statistics, Govt. of India, and Indian Institute of Sugarcane Research, Lucknow, India.

Terminologies and Notations

In India, there are 28 states, which exhibit various growth patterns of sugarcane yield, viz. (i) increasing growth pattern, (ii) decreasing growth pattern, (iii) constant growth pattern, and (iv) inconsistent growth pattern. For the present analysis, we have considered three states: Uttar Pradesh (S1), Maharashtra (S2), and Andhra Pradesh (S3). In these states, we observe the various above mentioned cases of growth pattern.

Fitting of Statistical Models to the Data

In order to analyze the trend and growth patterns of sugarcane yield in the states S1, S2, and S3, we compute the trend values by fitting linear and exponential models to the concerned time series data on sugarcane yield as follows:

(A) Linear Model:

$$y_t = a + bt \qquad \ldots(1)$$

where y_t denotes the time series value at time t. The values of constants 'a' and 'b' are obtained on solving the following normal equations, which are obtained by using the principle of Least Squares.

$$\sum y_t = na + b\sum t \qquad \ldots(2)$$

$$\sum ty_t = a\sum t + b\sum t^2 \qquad \ldots(3)$$

(B) Exponential Model:

$$y_t = ab^t \qquad \ldots(4)$$

Taking log on both sides of (4), we have

$$Y_t = A + Bt \qquad \ldots(5)$$

where $Y_t = log\ y_t$, $A = log\ a$ and $B = log\ b$.

The normal equations for estimating 'A' and 'B' are as follows:

$$\sum Y_t = nA + B\sum t \qquad \ldots(6)$$

$$\sum tY_t = A\sum t + B\sum t^2 \qquad \ldots(7)$$

Finally, on solving (6) and (7), the values of 'a' and 'b' are obtained as follows:

$$a = antilog\,(A),\ b = antilog\,(B)$$

Data Analysis and Results

The secondary time series data on the sugarcane yield in the states S1, S2 and S3, is depicted in Table 1. Also, the trend values on fitting linear and exponential models to the data pertaining to states S1, S2 and S3 are computed and presented in Tables 2, 3 and 4, respectively. Moreover, the model equations for linear and exponential trends in the respective states are depicted in Table 5.

Table 1: Time series data on sugarcane yield in selected states of India

Year	*Yield (in Tons / Hectare) for the States		
	S1	S2	S3
2010	56.73	84.87	77.94
2011	59.58	84.87	81.79
2012	59.88	74.66	79.42
2013	60.45	82.07	80.13
2014	62.15	82.23	71.85
2015	67.03	74.65	76.66
2016	64.89	82.52	76.02
2017	79.24	92.00	78.68
2018	80.81	77.20	79.36
2019	80.81	78.65	78.72

***Source:** *As per the records of Directorate of Economics & Statistics, New Delhi, and IISR, Lucknow.*

Table 2: Trend values for linear and exponential models in state S1

Year (t)	Yield (O_i)	Trend Values	
		Linear Model (L_i)	Exponential Model (E_i)
2010	56.73	54.11	23.02
2011	59.59	57.02	24.01
2012	59.88	59.93	25.04
2013	60.45	62.84	26.11
2014	62.15	65.75	27.23
2015	67.03	68.66	28.40
2016	64.89	71.58	29.62
2017	79.24	74.49	30.89
2018	80.81	77.10	32.21
2019	80.81	80.31	33.60

Table 3: Trend values for linear and exponential models in state S2

Year (t)	Yield (O_i)	Trend Values	
		Linear Model (L_i)	Exponential Model (E_i)
2010	84.87	82.030	82.073
2011	84.87	81.853	81.882
2012	74.66	81.676	81.692

2013	82.07	81.499	81.502
2014	82.23	81.322	81.312
2015	74.65	81.145	81.123
2016	82.52	80.968	80.935
2017	92.00	80.791	80.746
2018	77.20	80.614	80.559
2019	78.65	80.437	80.371

Table 4: Trend values for linear and exponential models in state S3

Year (t)	Yield (O_i)	Trend Values	
		Linear Model (L_i)	Exponential Model (E_i)
2010	77.94	79.72	78.58
2011	81.79	79.59	78.45
2012	79.42	79.46	78.33
2013	80.13	79.34	78.20
2014	71.85	79.21	78.08
2015	76.66	79.10	77.96
2016	76.02	78.95	77.83
2017	78.68	78.82	77.71
2018	79.36	78.70	77.58
2019	78.72	78.57	77.46

Table 5: Model equations for linear and exponential trends in selected states of India

States	Linear Model	Exponential Model
S1	$y_t = 2.911t - 5797$	$y_t = (65.18002)\ (1.043378)^t$
S2	$y_t = -0.177t + 437.8$	$y_t = (81.31225)\ (0.997675)^t$
S3	$y_t = -0.128t + 337$	$y_t = (78.08111)\ (0.99841)^t$

In Tables 2, 3 and 4, the term 'O_i' denotes the observed value of sugarcane yield (in Tons per Hectare) for the i-th year ($i = 2010, 2011, \ldots, 2019$), '$L_i$' denotes the linear trend value of sugarcane yield for the i-th year, and 'E_i' denotes the exponential trend value of sugarcane yield for the i-th year.

In order to illustrate the relative influence of linear and exponential trend values on the observed values of sugarcane yield for the states S1, S2, and S3, the graphical plots are obtained and demonstrated in Figs. 1 to 6.

Moreover, for testing the suitability of the linear and exponential models, we have obtained the coefficients of determination (R^2) for the states S1, S2, and S3, by using the formula:

$$R^2 = 1 - \frac{\sum\left(Y - \hat{Y}\right)^2}{\sum\left(Y - \bar{Y}\right)^2} \qquad \ldots(8)$$

where Y denotes the observed value of sugarcane yield, $\bar{Y}$ is the mean value of the variable Y, and $\hat{Y}$ is the trend value of the variable Y obtained on fitting the respective statistical model (such as linear model or exponential model as the case may be) to Y.

The values of coefficients of determination (R^2), for the concerned states, by fitting the linear and exponential models, are obtained and presented in Table 6.

Table 6: Values of R^2 for selected states by fitting linear and exponential models

States	Linear Model	Exponential Model
S1	0.859	0.879000
S2	0.010	0.009230
S3	0.020	0.020284

From Table 6, it is revealed that:

(i) In each of the states S1, S2, and S3, the values of R^2 are nearly the same for both the linear and exponential models. Hence, both the models are equally suitable for describing the trend and growth patterns of sugarcane yield in the concerned states.

(ii) In state S1, we observe that $R^2 > 0.5$, which implies that both the linear and exponential models are best fitted, and hence are suitable for exploring the rising growth pattern of sugarcane yield in the concerned state.

(iii) In states S2 and S3, the values of R^2 are infinitesimally small, due to the reason that the variation of yield (Y) from the mean value ($\bar{Y}$) is nearly the same as the variation of yield from the trend value $\hat{Y}$ (i.e., linear trend value and exponential trend value, as the case may be). Symbolically, for states S2 and S3, we have $\sum\left(Y - \hat{Y}\right)^2 \cong \sum\left(Y - \bar{Y}\right)^2$. Moreover, in state S2, we observe a slight decline in growth pattern of sugarcane yield, whereas in state S3, we observe an inconsistent growth pattern of sugarcane yield.

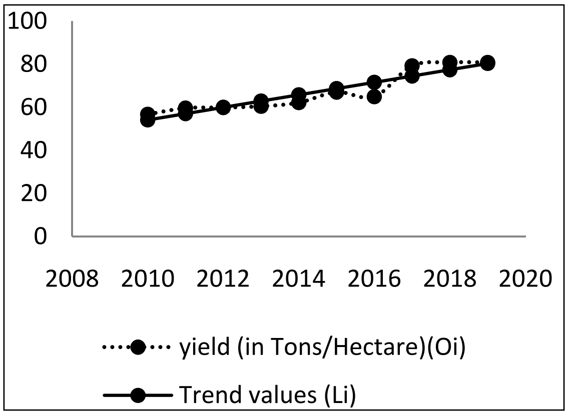

Fig. 1: Trend values for Linear Model in state S1

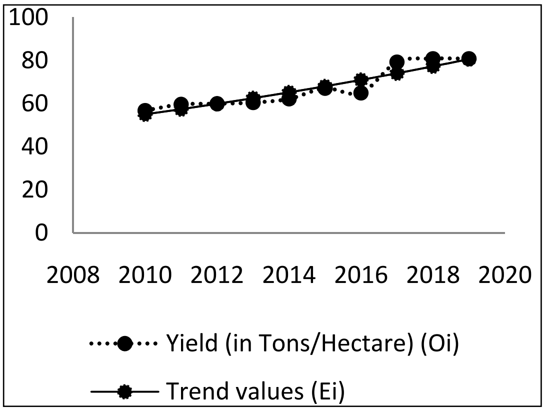

Fig. 2: Trend values for Exponential Model in state S1

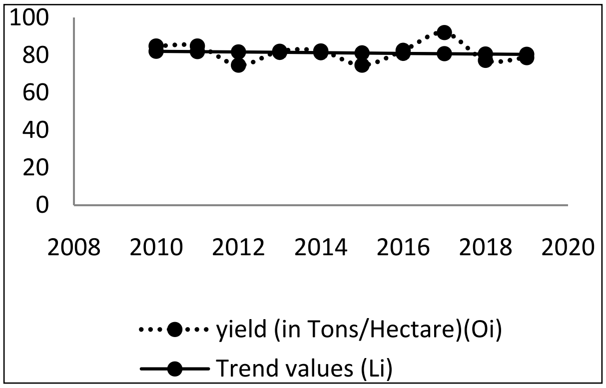

Fig. 3: Trend values for Linear Model in state S2

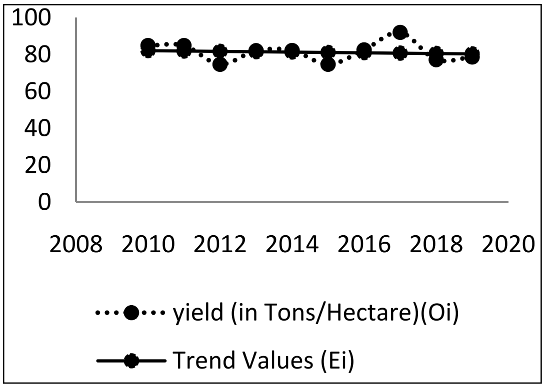

Fig. 4: Trend values for Exponential Model in state S2

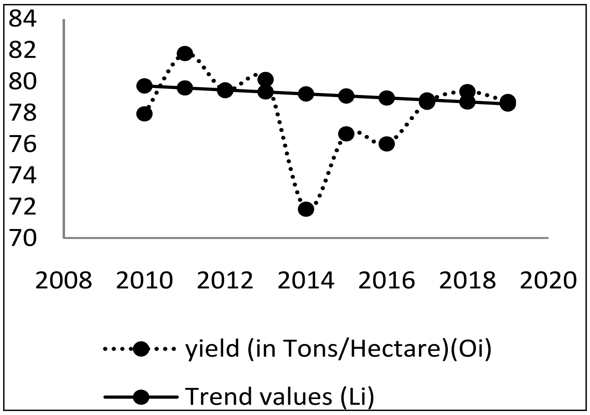

Fig. 5: Trend values for Linear Model in state S3

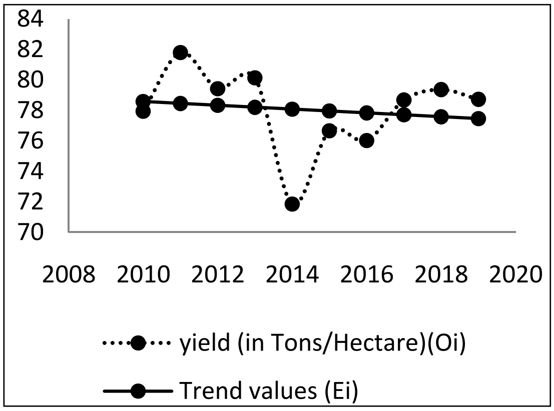

Fig. 6: Trend values for Exponential Model in state S3

Formulation of Hypotheses

We test the following null hypotheses:

H_{0A}: Linear model fits the given data on the sugarcane yield.

H_{0B}: Exponential model fits the given data on the sugarcane yield.

against the following respective alternative hypotheses:

H_{1A}: Linear model does not fit the given data on the sugarcane yield.

H_{1B}: Exponential model does not fit the given data on the sugarcane yield.

The above-mentioned hypotheses are tested for the concerned states S1, S2, and S3 of India.

Hypotheses Testing and Validation

The chi-square values for the fitted linear and exponential models (i.e., χ_L^2 and χ_E^2) in the concerned states S1, S2, and S3 of India, have been computed, and the findings are explored in Table 7. The chi-square values by fitting the concerned models have been obtained using the below mentioned formulae:

$$\chi_L^2 = \sum_{i=1}^{n} \frac{(O_i - L_i)^2}{L_i} = \sum_{i=1}^{10} \frac{(O_i - L_i)^2}{L_i},$$

$$\chi_E^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = \sum_{i=1}^{10} \frac{(O_i - E_i)^2}{E_i},$$

where the terms 'O_i', 'L_i' and 'E_i' have been utilized from the Tables 2, 3, and 4, for the respective states S1, S2, and S3 of India.

Table 7: Values of χ^2 for selected states by fitting linear and exponential models

States	Chi-square values	
	Linear Model (χ_L^2)	Exponential Model (χ_E^2)
S1	1.6507	1.4084
S2	3.1150	3.1170
S3	0.9810	0.8437

The tabulated values of chi-square at 1% and 5% levels of significance with 9 degrees of freedom are given, respectively, by

$$\chi^2_{0.01,9} = 21.67 \text{ and } \chi^2_{0.05,9} = 16.92$$

From Table 7, we observe the following results:

(i) $\chi_{L\,(Si)}^2 < \chi^2_{0.01,9}$ and $\chi_{L\,(Si)}^2 < \chi^2_{0.05,9}$ $(i = 1,2,3)$

(ii) $\chi_{E\,(Si)}^2 < \chi^2_{0.01,9}$ and $\chi_{E\,(Si)}^2 < \chi^2_{0.05,9}$ $(i = 1,2,3)$

Hence, on the basis of above results, we accept the null hypotheses H_{0A} and H_{0B}, and conclude that the linear model as well as the exponential model fit the given time series data on sugarcane yield for the respective states S1, S2, and S3 of India.

DISCUSSION & CONCLUSION

The present paper deals with the time series analysis of sugarcane yield in selected states of India. The growth and trend patterns of sugarcane yield have been examined by fitting statistical models, viz. linear model and exponential model, to the concerned time series data for selected states. It has been observed that the linear and exponential trend values are nearly the same as the observed values for the states S1 (Uttar Pradesh) and S2 (Maharashtra), whereas due to inconsistent growth pattern of sugarcane yield in state S3 (Andhra Pradesh), we observe a slight deviation in the trend values (i.e., linear and exponential) from the observed values. Hence, both the linear and exponential models are appropriate for predicting the future trends of sugarcane yield in the concerned states of India.

It has been observed from the empirical results that the state S1 exhibits increasing growth pattern of sugarcane yield, while the state S2 exhibits a slightly decreasing growth pattern. Moreover, in state S3, we observe an inconsistent growth pattern characterized by a slight decline in sugarcane yield in the year 2019 as compared to the year 2018. In recent years, the sugarcane yield in the state S1 remains the highest as compared to the other selected states S2 and S3 of India.

In order to test the "goodness of fit" of the linear and exponential models for the states S1, S2, and S3, the chi-square test statistic values (i.e., χ_L^2 and χ_E^2) have been computed for the respective states. These values are then compared with the tabulated values of chi-square at 1% and 5% levels of significance. It has been observed that both the models fit the given time series data on sugarcane yield for the concerned states.

Moreover, on the basis of values of coefficient of determination (R^2), we conclude that both the linear and exponential models are suitable for analyzing and exploring the trend patterns of sugarcane yield in the concerned states of India.

The present study could be enhanced further by considering the scenario of sugarcane yield in the other states of India. Moreover, on considering the benefits and usefulness of sugarcane, the potential farmers could be encouraged for its cultivation.

REFERENCES

Arti and Rai, C.K. 2017. Growth rate of area, production and productivity of sugarcane crop in Uttar Pradesh. *Research Journal of Agricultural Sciences*, **8**(2): 423-425.

Dharmaraja, S., Jain, V., Anjoy, P. and Chandra, H. 2020. Empirical analysis for crop yield forecasting in India. *Agricultural Research*, **9**(1): 132-138.

Lal, B., Maurya, K.K., Gupta, R. P., Kumar, M. and Kumar, S. 2020. Statistical analysis of growth pattern of sugarcane production in districts of western Uttar Pradesh. *International Journal of Current Microbiology and Applied Sciences*, **11**: 2331-2346.

Lal, C. and Rohtas 2022. The trends of area, production and productivity of selected spices and traditional crops in Haryana. *Economic Affairs*, **67**(01): 19-24.

Mishra, P., Al Khatib, A.M.G., Sardar, I., Mohammed, J., Karakaya, K., Dash, A., Ray, M., Narsimhaiah, L. and Dubey, A. 2021. Modeling and forecasting of sugarcane production in India. *Sugar Tech.*, **23**(6): 1317-1324.

Rahman, F. and Bee, N. 2019. Trends and pattern of sugarcane production in Shahjahanpur district, Uttar Pradesh: A geographical analysis. *Economic Affairs*, **64**(3): 537-545.

Rajarathinam, A., Parmar, R.S. and Vaishnav, P.R. 2010. Estimating models for area, production and productivity trends of Tobacco (*Nicotiana tabacum*) crop for Anand region of Gujarat state, India. *Journal of Applied Sciences*, **10**(20): 2419-2425.

Ramteke, R., Gupta, G.K. and Singh, D.V. 2015. Growth and yield responses of soybean to climate change. *Agricultural Research*, **4**(3): 319-323.

Rao, G.P. and Singh, P. 2022. Value addition and fortification in non-centrifugal sugar (jaggery): A potential source of functional and nutraceutical foods. *Sugar Tech.*, **24**(2): 387–396.

Roy, M.M. and Chandra, A. 2019. Potential role of appropriate technological interventions in enhancing income of sugarcane farmers in subtropical India. *Agricultural Research*, **8**(4): 531-539.

Saha, J.K., Adnan, K.M.M., Sarker, S.A. and Bunerjee, S. 2021. Analysis of growth trends in area, production and yield of tea in Bangladesh. *Journal of Agriculture and Food Research*, **4**: 1-7.

Sanjeev, Verma, U. and Tonk, M. 2015. Time-series modeling for sugarcane yield estimation in Haryana. *International Journal of Applied Mathematics & Statistical Sciences*, **4**(6): 53-62.

Singh, B., Arya, C.K., Snehdeep and Sharma, M.K. 2021a. Statistical analysis for trend and change point detection of sugarcane production in India. *Journal of Crop and Weed*, **17**(2): 170-176.

Singh, B., Sharma, P., Arya, C.K., Singh, A. and Kumar, J. 2021b. Spatial growth and instability analysis of area, production and yield of sugarcane in India. *Economic Affairs*, **66**(2): 245-252.

Suresh, K.K. and Krishna Priya, S.R. 2011. Forecasting sugarcane yield of Tamilnadu using ARIMA models. *Sugar Tech.*, **13**(1): 23-26.