American Journal of Pharmacy an pharmacology
1(1):6-13, August 2019

Research Paper

Forecasting of Sugar Production Using Regression and Smoothing Analysis

[1]C.Mounika, [2]G.Sreedhar, [3]J.Lakshmidevi.

[1, 2, 3] Lecturer, Dept. of Statistics, C.S.S.R and S.R.R.M Degree & PG College, Kamalapuram, Kadapa -516289.

To cite this article: C.Mounika, G.Sreedhar and J.Lakshmi Devi. Forecasting of sugar production using regression and smoothing analysis: International Journal of Advanced Research in Engineering, 1(1):6-13, May – June 2019.

Email: mounikachemikala@gmail.com

Received: 17th May 2019.| Revised: 12th June 2019.| Accepted: 24th June 2019.

Abstract: Forecasting is the prediction of the variable based on known past values of it or other related variables. After estimation of parameters of the regression model common use of regression is for prediction or forecasting. In the present study the simple linear forecasting method considered to be the optimum technique for forecasting. To test this, we have collected data on sugar production from S.V. Sugar Factory Ltd., Tirupathi and applied different forecasting techniques like Simple Linear Regression Method, Simple Non Linear Regression Method, Simple Exponential Smoothing Method and Holt's Linear Exponential Smoothing Method. The results of these techniques are then compared by using the criteria of Mean Absolute Error (MAE) measure to arrive at the final conclusion.

Keywords: Forecasting, Prediction, Simple Linear Regression Method, Simple Non- Linear Regression Method, Simple Exponential Smoothing Method, Holt's Linear Exponential Smoothing Method.

1. Introduction

Forecasting is the process of making predictions of the future based on past and present data and most commonly by analysis of trends. A common place example might be estimation of some variable of interest at some specified future date. Prediction is a similar, but more general term. Both might refer to formal statistical methods employing time series, cross- sectional or longitudinal data, or alternatively to less formal judgmental methods. Usage can differ between areas of application: for example, in hydrology the terms "forecast" and "forecasting" are sometimes reserved for estimates of values at certain specific future times, while the term "prediction" is used for more general estimates, such as the number of times floods will occur over a long period.

Risk and uncertainty are central to forecasting and prediction; it is generally considered good practice to indicate the degree of uncertainty attaching to forecasts. In any case, the data must be up to date in order for the forecast to be as accurate as possible. In some cases the data used to predict the independent variable is itself forecasted. Forecasting might seem purely hypothetical but companies can utilize forecasting tools to provide as much accuracy as possible. Care must be taken in

- Selection of the method of forecasting for production planning.
- Determining the time period to study.
- Choosing reports on previous activities to help with projecting future production.
- Picking market trends to apply to the forecast.

2. Production forecasting

Production forecasting means to estimate the future demand for goods and services. It also estimates the resources which are required to produce those goods and services. These resources include human resources, financial and material resources. So, production forecasting means to estimate the 6M's of management. The production manager first estimates the future market or demand for the companies' goods and services. Then he estimates the Men (human resources), Money (financial resources), Materials, Machines and Methods, which will be required to produce those goods and services. Production forecasting estimates the future technological developments. It estimates the customer's needs and preferences along with competitors' strategy in the future. So, production forecasting is an estimation of a wide range of future events, which affect the production of the organization.

First production manager studies all the past and present events. Then he makes estimations about the future. So, most of the production forecasts are made for existing goods and services. However, some new products will be introduced into the market in an upcoming future. Forecasts for these new products are called predictions. Production forecasting requires a lot of skill, experience and judgement of the production manager. He must use many statistical techniques and tools to make his forecasts accurate.

Production forecasting is a combination of objective calculations and subjective judgements. That is, it involves systematic collecting and analysing past and present data. This is done objectively with the help of statistical techniques and tools. Production forecasting also involves subjective judgement of the production manager. That is, he has to use his intuition and judgement for forecasting

The success or failure of an organization depends upon the accuracy of its production forecasts. All the plans and strategies of the organization are based on the production forecasts. Production forecasts help to reduce the business risks.Managers in a manufacturing environment often must forecast the amount of inventory and supplies needed to meet demands. The use of forecasting assumes that past trends will continue with little variance into the future. In some industries, production trends may vary by season. However, for forecasting to be effective, even these shifting trends must remain consistent over time. There are a variety of forecasting methods manufacturers may use in production planning. Produces, types of statistical data available for forecasting, degree of accuracy required for your forecasting methods, forecasting for a specific period, product life cycle, etc., are some of the important aspects and factors required for effective production forecasting and planning.

Forecasting is the prediction of the variable based on known past values of it or other related variables. After estimation of parameters of the regression model common use of regression is for prediction or forecasting. In the present study the Simple Linear forecasting method considered to be the optimum technique for forecasting. To test this, we have collected data on sugar production from S.V. Sugar Factory Ltd., Tirupathi and applied different forecasting techniques like Simple Linear Regression Method, Simple Non Linear Regression Method, Simple Exponential Smoothing Method and Holt's Linear Exponential Smoothing Method. The results of these techniques are then compared by using the criteria of Mean Absolute Error (MAE) measure to arrive at the final conclusion.

2.1. Need of the study

Every business entity wants to predict the future to be successful business. Time is more important and precious aspect of the modern management. Future depends on the forecasting techniques and models and these models are provided by the statistical analytical models and techniques. Many statistical techniques are used for predicting and forecasting the future strategies. Hence, the empirical analysis is chosen for the project study.

2.2. Objectives

The objectives of the empirical analysis and forecasting and prediction of the sugar production are:

- To forecast and predict the variables based on the past values.
- To make empirical analysis using the suitable models particularly the regression model.
- To predict the sugar production applying different forecasting techniques such as Regression and Exponential Smoothing Methods.
- To compare the results of the techniques and to arrive at the final conclusion.
- To draw forecast sugar production of Sri Venkateshwara cooperative Sugar Factory Ltd., Tirupathi with necessary inferences.

2.3. Data collection

Collection of the data is purely based on the secondary data. The data that are collected by someone else for a purpose other than the researcher's current project and has already undergone the statistical analysis is called as the Secondary Data. The secondary data is collected from the records of the S.V. Sugar Factory Ltd., Tirupathi particularly from the Production Reports, Sales Report and Financial Statements.

2.4. Scope of study

The empirical study is spread the time span of 19 years that is from 1998-1999 to 2016-2017 for making effective and productive analysis and forecasting of future sugar production.

2.5. Limitations of study

All forecasting is subject to error and change. Forecasting tools are utilised to provide focus for the future but the forecasts may not suit the change as market trends and other company activity changes.

3. Forecasting of sugar production

The objective is to draw forecast using different types of forecasting techniques and compare them to infer which is best suitable technique for the data collected on sugar production of Sri Venkateshwara cooperative Sugar Factory Ltd., Tirupathi. The data is shown in table 1.

Table-1: Sugar Production of Sri Venkateshwara Co1operative Sugar Factory Ltd. Tirupathi.

S. No.	Time period	Sugar production (in qtls)
1	1998-1999	144.1110
2	1999-2000	177.5020
3	2000-2001	123.9560
4	2001-2002	161.3180
5	2002-2003	173.1080
6	2003-2004	263.7700
7	2004-2005	216.7500
8	2005-2006	108.7750
9	2006-2007	215.4050
10	2007-2008	267.3650
11	2008-2009	250.1750
12	2009-2010	137.8660
13	2010-2011	214.1120
14	2011-2012	248.9231
15	2012-2013	227.3124
16	2013-2014	222.3040
17	2014-2015	155.2345

18	2015-2016	214.6587
19	2016-2017	220.4560

3.1. Forecasting yearly output of sugar production using simple linear regression method

Consider a simple linear relationship between the value of production (Y) and time variable (t) as

$Y = a + bt + \epsilon$ (1)

If 'n' observation are measured, then the ordinary least squares estimates of a and b are given by

$$\hat{b} = \frac{\sum Yt - \frac{(\sum Y)(\sum t)}{n}}{\sum t^2 - \frac{(\sum t)^2}{n}} \quad (2)$$

And $\hat{a} = \bar{Y} - \hat{b}\bar{t}$b (3)

Where $\bar{Y} = \frac{\sum Y}{n}$ and $\bar{t} = \frac{\sum t}{n}$ then the estimated relationship is given by $\hat{Y} = \hat{a} + \hat{b}t$

From the results of analysis, it is observed that the estimated two variable linear forecasting model of sugar production is given by

$\hat{Y} = 164.2703 + 3.2735t$ (4)

Using this estimated model we can forecast the value of production for the study periods and the forecasted values are shown in table-2.

Table- 2: Forecast of Yearly Output of Sugar Using Simple Linear Regression Method

S. No.	Time period	Observed production (in qtls)	Forecasted values (in qtls)
1	1998-1999	144.1110	167.5438
2	1999-2000	177.5020	170.8173
3	2000-2001	123.9560	174.0908
4	2001-2002	161.3180	177.3643
5	2002-2003	173.1080	180.6378
6	2003-2004	263.7700	183.9113
7	2004-2005	216.7500	187.1848
8	2005-2006	108.7750	190.4583
9	2006-2007	215.4050	193.7319
10	2007-2008	267.3650	197.0054
11	2008-2009	250.1750	200.2789
12	2009-2010	137.8660	203.5524
13	2010-2011	214.1120	206.8259
14	2011-2012	248.9231	210.0994
15	2012-2013	227.3124	213.3729
16	2013-2014	222.3040	216.6464
17	2014-2015	155.2345	219.9199
18	2015-2016	214.6587	223.1934
19	2016-2017	220.4560	226.4669

Analysis of Errors:

- Mean Error (ME) = 0.0001
- Mean Absolute Error (MAE) = 34.0784
- Mean Absolute Percentage Error (MAPE) = 19.6734
- Standard Deviation of Error (unbiased) = 44.5031
- Mean Square Error (MSE) = 11876.291

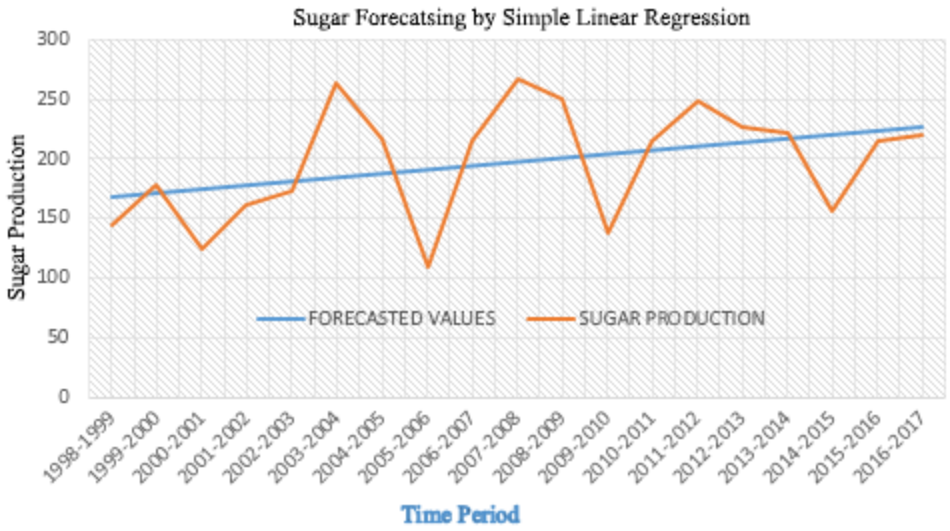

Figure-1: Sugar Forecasting by Simple Linear Regression

3.2. Forecasting of yearly output of sugar using simple non-linear regression method

Consider an exponential relationship between the value production (y) and time variable t as

$y = ae^{bt}e^{\epsilon}$ (5)

Taking logarithms on both sides, we get $\log y = \log a + bt + \epsilon$

Here, logarithms are taken with base 'e', rewriting an above equation as $Y = A + Bt + \epsilon$

Where $A = \log a$ and $B = b$

The ordinary least squares estimates of a and b are given by $\hat{B} = \frac{\sum Yt - \frac{(\sum Y)(\sum t)}{n}}{\sum t^2 - \frac{(\sum t)^2}{n}}$

And $\hat{A} = \bar{Y} - \hat{b}t$

Where $\bar{Y} = \frac{\sum Y}{n}$ and $\bar{t} = \frac{\sum t}{n}$ then the estimated relationship is given by

$\hat{Y} = \hat{A} + \hat{B}t$ (or) $\hat{y} = \hat{a}e^{\hat{b}t}$ (6)

Here, $\hat{a}$ = antilog ($\hat{A}$) and $\hat{B} = \hat{b}$

Using this estimated model the value of production can be estimated.

From the results of analysis it is observed that the estimated two variable non - linear forecasting model of sugar production is given by

$\hat{y} = (158.6074)\, e^{0.008048t}$ (from equation 6)

Table- 3 Forecast of Yearly Output of Sugar Using Simple Non – Linear Regression Method

S. No.	Time period	production	(i precasted value (in qtls)
1	1998-1999	144.1110	159.8890
2	1999-2000	177.5020	161.1810
3	2000-2001	123.9560	162.4835
4	2001-2002	161.3180	163.7964
5	2002-2003	173.1080	165.1200
6	2003-2004	263.7700	166.4542
7	2004-2005	216.7500	167.7993
8	2005-2006	108.7750	169.1552
9	2006-2007	215.4050	170.5221
10	2007-2008	267.3650	171.9000
11	2008-2009	250.1750	173.2890
12	2009-2010	137.8660	174.6893
13	2010-2011	214.1120	176.1009
14	2011-2012	248.9231	177.5238
15	2012-2013	227.3124	178.9583
16	2013-2014	222.3040	180.4044
17	2014-2015	155.2345	181.8622
18	2015-2016	214.6587	183.3317
19	2016-2017	220.4560	184.8131

Analysis of Errors

- Mean Error (ME) = 24.9383
- Mean Absolute Error (MAE) = 43.9504
- Mean Absolute Percentage Error (MAPE) = 22.1483
- Standard Deviation of Error (unbiased) = 45.8004
- Mean Square Error (MSE) = 2609.19

The analysis is graphically represented in figure 2.

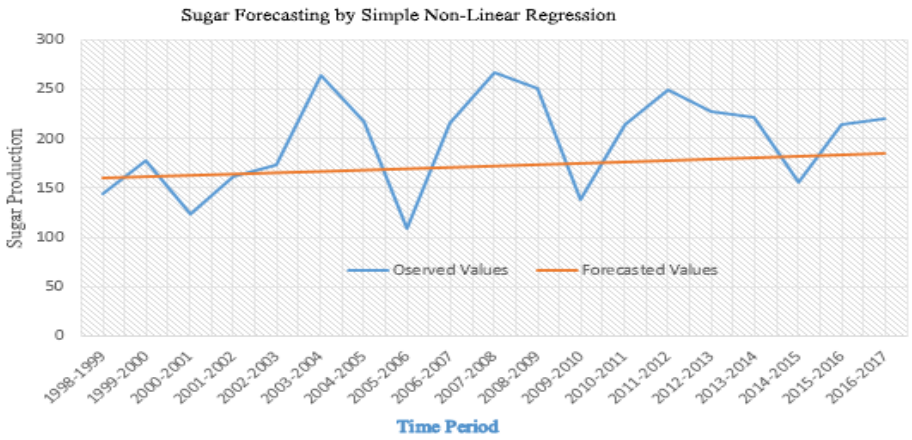

Figure- 2: Sugar Forecasting by Simple Non - Linear Regression

3.3. Forecasting of yearly output of sugar production using simple exponential smoothing

The equation for forecasting in this method is given by

$$F_{t+1} = \alpha X_t + (1- \alpha)F_t \quad (7)$$

This single exponential smoothing requires little storage and few computations. For example, to get the single exponential smoothing forecasting system started we need F_1

$$F_2 = \alpha X_1 + (1- \alpha)F_1$$

Since, the value of F_1 is not known, the first observed value (X_1) is used as the first forecast.

($F_1 = X_1$) and then proceeding using the equation.

$$F_{t+1} = \alpha X_t + (1- \alpha)F_t$$

By trial and error method, it was noticed that for present data the appropriate value of smoothing coefficient is 0.1. Table-4 shows the exponential smoothing results for $\alpha = 0.1, 0.5$ and 0.9.

Table -4: Forecast of Yearly Output of Sugar Using Simple Exponential Smoothing

S. No.	Time period	rved values	(in	Exponential smoothed values A=0.1 (in Qtls)	A=0.5 (in Qtls)	A=0.9 (in Qtls)
1	1998-1999	144.1110		---	---	---
2	1999-2000	177.5020		144.1110	144.1110	144.1110
3	2000-2001	123.9560		147.4501	160.8065	174.1629
4	2001-2002	161.3180		145.1007	142.3813	128.9767
5	2002-2003	173.1080		146.7224	151.8496	158.0839
6	2003-2004	263.7700		149.3610	162.4788	171.6056
7	2004-2005	216.7500		160.8019	213.1244	254.5536
8	2005-2006	108.7750		166.3967	214.9372	220.5304
9	2006-2007	215.4050		160.6345	161.8561	119.9505
10	2007-2008	267.3650		166.1116	188.6306	205.8596
11	2008-2009	250.1750		176.2369	227.9978	261.2145
12	2009-2010	137.8660		183.6307	239.0864	251.2789
13	2010-2011	214.1120		179.0543	188.4762	149.2073
14	2011-2012	248.9231		182.5600	201.2941	207.6215
15	2012-2013	227.3124		189.1963	225.1086	244.7929
16	2013-2014	222.3040		193.0079	226.2105	229.0605
17	2014-2015	155.2345		195.9375	224.2572	222.9796
18	2015-2016	214.6587		191.8672	189.7459	162.0090
19	2016-2017	220.4560		194.1464	202.2023	209.3937

	196.7773	211.3291	219.3498

Table- 5: Analysis of Errors

Analysis of Errors	Test Periods		
	α=0.1	α=0.5	α=0.9
Mean Error (ME)	29.25908	7.468683	4.644369
Mean Absolute Error (MAE)	47.87948	42.70894	50.88885
Mean Absolute Percentage Error (MAPE)	24.05667	24.7226	29.33219
Standard Deviation of Error (Unbiased)	47.51607	55.46098	62.57421
Mean Square Error (MSE)	2988.439	2960.8175	3719.572

The results are presented graphically in figure-3, it has been observed that, a large value of α (0.9) gives very little smoothing in the forecast whereas smaller value of α (0.1) gives considerable smoothing. The Mean Square Error (MSE) is computed over a test set, and then another α value is tried. The MSEs are compared to different α values that gives minimum MSE. Since MSE is less when α = 0.5, we can choose α = 0.5 as an appropriate value of smoothing coefficient for the present data collected. The analysis is represented graphically in figure 3.

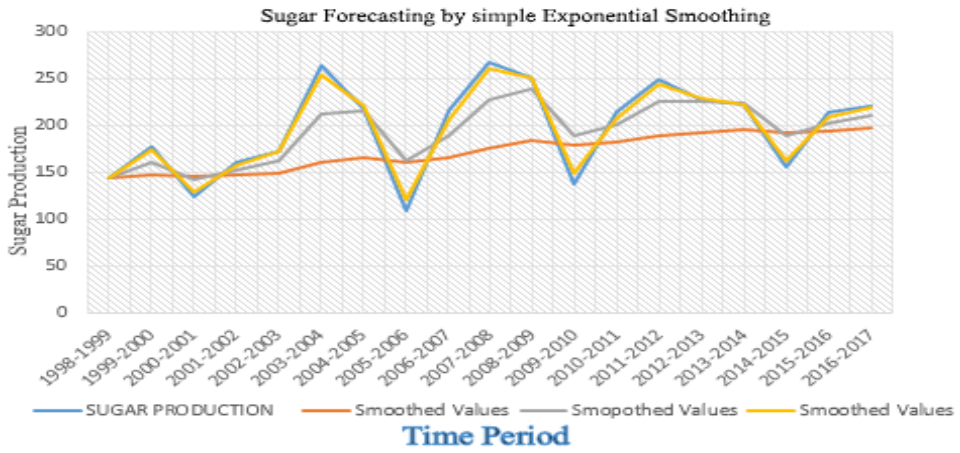

Figure -3: Forecasting of yearly output of sugar using linear (Holt's) exponential smoothing technique

Let S_t = single exponential smoothed value

α = Smoothing coefficient

T_t = indicates smoothed trend in the data series

Then the set of three equations useful in Holt's Exponential Smoothing technique are expressed as

$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + T_{t-1})$ (8)

$T_t = \alpha(S_t - S_{t-1}) + (1-\alpha) T_{t-1}$ (9)

And $F_{t+m} = S_t + T_t m$ (10)

It should be noted that, while applying this technique, the value of S is updated first and then the trend is updated. Since, the forecasting is done successively; we can take m equals to one. The best value of α is found by trying various combinations of values between zero and one and choosing that set of values that minimizes the Mean Square Error or Mean Absolute Deviation. The best possible values to the data are found to be α = 1.0 Using this value, from equations 8,9,10 when t = 2,

$S_2 = \alpha X_2 + (1-1)(S_1 + T_1) = 177.5020$

$T_2 = \alpha(S_2 - S_1) + (1-1) T_1 = 33.3910$

For period t = 3, the results are

$S_3 = \alpha X_3 + (1-1)(S_2 + T_2) = 123.9560$

$T_3 = \alpha(S_3 - S_2) + (1-1) T_2 = -53.5460$

Computing these calculations for intervening period, The results for period t = 19 are:

$S_{19} = \alpha X_{19} + (1-1)(S_{18} + T_{18}) = 220.4560$

$T_{19} = \alpha(S_{19} - S_{18}) + (1-1) T_{18} = 5.7973$

Therefore, the forecasted equation for period (t + m) is given by

$F_{t+m} = 220.4560 + 5.7973m$

Using this model the value of production can be forecasted for the given periods and the forecast values are given in table-6.

Table- 6: Forecasting of Yearly Output of Sugar Using Linear (Holt's) Exponential Smoothing

S. No.	Time period	Sugar production	Forecasted values
1	1998-1999	144.1110	116.1046
2	1999-2000	177.5020	121.9019
3	2000-2001	123.9560	127.6992
4	2001-2002	161.3180	133.4965
5	2002-2003	173.1080	139.2938
6	2003-2004	263.7700	145.0911
7	2004-2005	216.7500	150.8884
8	2005-2006	108.7750	156.6857
9	2006-2007	215.4050	162.483
10	2007-2008	267.3650	168.2803
11	2008-2009	250.1750	174.0776
12	2009-2010	137.8660	179.8749
13	2010-2011	214.1120	185.6722
14	2011-2012	248.9231	191.4695
15	2012-2013	227.3124	197.2668
16	2013-2014	222.3040	203.0641
17	2014-2015	155.2345	208.8614
18	2015-2016	214.6587	214.6587
19	2016-2017	220.4560	220.456

- Analysis of Errors

Mean Error (ME) = 28.7250

Mean Absolute Error (MAE) = 44.2292

Mean Absolute Percentage Error (MAPE) = 22.3832

Standard Deviation of Error (unbiased) = 46.7143

Mean Square Error (MSE) = 2892.506

The analysis is graphically represented in figure 4.

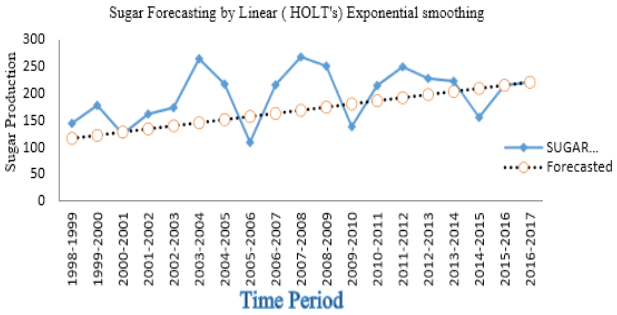

Figure-4: Sugar Forecasting by Linear (Holt's) Exponential Smoothing

4. Conclusion

A forecasting method is said to be suitable for which a measure of error is minimum. In other words we calculate the Mean Absolute Error (MAE) for each method and calculate that the method for which this Mean Absolute Error value is minimum as suitable method. Comparison of Mean Absolute Error is given in table-7.

Table -7: MAE for different Methods

S.no.	Method applied	Mean absolute error (MAE)
1.	Simple Linear Regression Method	34.0784
2.	Simple Non - Linear Regression Method	43.9504
3.	Simple Exponential Smoothing Method	42.7089
4.	Simple Linear (Holt's) Exponential Smoothing Method	44.2292

It is observed that measure of error MAE is minimum for "Simple Linear Regression Method" for forecasting yearly output of sugar. Thus, it is concluded that for forecasting yearly output of sugar, the Simple Linear Regression Method is appropriate. The future values predicted using the above methods are given below. The data were collected with regard to the production of sugar for a period of nineteen years, i.e., from 1998-1999 to 2016-2017. Forecasting was also made for the years up to 2022-2023 by the method of Simple Linear Regression. The Mean Absolute Error for each method was calculated and they show that this MAE is minimum for Simple Linear Regression Method. This gives a conclusion that this is the most optimum for forecasting.

5. Future Scope

It is observed that measure of Error MAE is minimum for Simple Linear regression method for forecasting yearly output of sugar. Thus, we conclude that for forecasting yearly output of sugar, the Simple Linear Regression method is appropriate. The future values are predicted by using the method of Simple Linear Regression is given in table-8.

Table -8: Predicted future values

Period	Yearly Output Qtls)
2017-2018	229.7404
2018-2019	233.0139
2019-2020	236.2874
2020-2021	239.5609
2021-2022	242.8344
2022-2023	246.1079

It can be concluded that the Simple Linear Regression method is appropriate for forecasting yearly output of sugar.

References

1. M.Adami,B.F.T.Rudorff,R.M.Freitas,D.A.Aguiar,L.M.Sugawara and M.P.Mello. Remote Sensing Time Series to Evaluate Direct Land Use Change of Recent Expanded Sugarcane Crop in Brazil. Sustainability, 4: 574-585, 2012.
2. M.A.Vieira,A.R.Formaggio,C.D.Renno,C.Atzberger,D.A.Aguiar and M.P.Mello. Object based image Analysis and Data Mining Applied to a

Remotely Sensed Landsat Time-Series to Map Sugarcane over Large Areas. Remote Sensing of Environment, 123:553-562, 2012.

3. R.R.V.Gonçalves,J.,Zullo,L.A.S.Romani,A.Nascimento and A.T.M.Traina. Analysis of NDVI time series using cross-correlation and forecasting methods for monitoring sugarcane fields in Brazil. International Journal of Remote Sensing, 33(15):4653-4672, 2012.
4. R.R.V.Gonçalves,J.Zullo,C.S.Ferraresso,E.P.Sousa,A.S.Romani and A.J.M.Traina Analysis of NOAA/AVHRR multi temporal images climate conditions and cultivated land of sugarcane fields applied to agricultural monitoring: 6° International Workshop on the Analysis of Multi-Temporal Remote Sensing Images 2011- Trento. Multi temp Proceedings, 2011.
5. R.G.Brown. Smoothing, forecasting and Prediction of Discrete time series.Prentice hall, 1962.
6. N.R.Draper and Smith. Applied regression analysis.Third Edition, A Wiley International Publications, New York, 1998.
7. W.Gilchrist. Statistical forecasting. Wiley, New York, 2000.
8. I.C.Johnston. Econometric methods.third edition, Mc Graw Hill, New York, 1995.
9. G.C.Judge. Theory and practice of econometrics. John Wiley, New York, 1985.