Journal Crop and Weed, 11(Special Issue):67-70(2015)

Forecasting groundnut production of India using nonlinear growth models

S. PAL AND [1]D. MAZUMDAR

Division of Computer Applications, ICAR-I.A.S.R.I., New Delhi-110012
[1]Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, Nadia-741252, West Bengal

Received: 15-10-2014; Revised:22-11-2014; Accepted:28-12-2014

ABSTRACT

Groundnut is one of the major sources of edible oil in India. Around one-fourth of country's total edible oil is produced from groundnut. This paper deals with a critical study of groundnut production of India with a non-linear approach. Different nonlinear growth models viz. Monomolecular, Logistic and Gompertz models have been employed for modeling of India's total groundnut production during the period 1950-51 to 2011-12. The parameters of these models were estimated using Gauss-Newton algorithm. It was observed that Monomolecular and Logistic models performed better followed by Gompertz for this dataset based on various goodness of fit criteria viz. Coefficient of determination (R^2), Mean absolute error (MAE), Root mean square Error (RMSE) and Mean absolute percentage error (MAPE). Finally, India's total groundnut production for 2014-15 to 2019-20 has been forecasted by using the Monomolecular and Logistic models.

Keywords: Forecasting, groundnut production, logistic, monomolecular, model, nonlinear growth model

India is the second largest producer of groundnut in the world after China with around 20 per cent of world's total production. Almost every part *viz*. oil, kernels, shell and straw of groundnut is of commercial value. It is called as the 'king' of oilseeds and accounts for nearly 25 percent of the total oilseed production of the nation. Although there has been a significant increase in oilseed production since 1960s, the demand for oilseed production is continuously going up due to increase in population growth rate and per capita edible oil consumption. However, due to the gap between domestic availability and actual consumption of edible oils, India has to resort to import of edible oils. The gap between demand and production is widening and this will continue in the foreseeable future. In case of groundnut, the production has been falling due to erratic monsoon and higher production cost compared to options of cotton and soya available to growers. Additionally, the demand of groundnut for direct human consumption has increased immensely. These led to less groundnut being available for the oil crushing industry, causing diminishing supply of groundnut oil. Despite various incentives offered by government there has not been adequate growth in oilseeds production of the country as well. In fact, according to figures available from Solvent Extractors Association of India (SEA), the production of oilseeds has grown marginally by close to 2% from 2003-04 to 2012-13 (Choksi, 2013). Therefore, a proper import policy for edible oil is of immense importance to meet the demand of growing population. It would be easier to formulate and initiate appropriate policy measures if the data with regard to the trend of production is obtained and analysed in advance (Dhekale *et al*., 2014). Being one of the key sources of edible oil, the future prediction of the groundnut production is a major concern to the policy makers of the country.

Nonlinear growth models have widely been used to measure agricultural growth in terms of growth rate (Joshi and Saxena, 2002; Sarma, 2005; Patil *et al.*, 2009; Rajarathinam *et al*., 2010). These are also employed in modeling (Iquebal and Sarika, 2013) and forecasting production, productivity (Panwar *et al*., 2014), area, etc. of various commodities. Thus, in framing of optimal agricultural policies, like import and export policies for various agricultural produce of a country, these nonlinear models play a crucial role. In the present investigation, three realistic nonlinear growth models *viz*. Monomolecular, Logistic and Gompertz have been used. Prajneshu and Chandran (2005) discussed in details about these models and their uses. A special feature of these models is that they are *Mechanistic models* in which parameters have specific biological interpretation. The paper is organized as follows: Section 2 deals with methodological aspect; results and discussions are mentioned in Section 3 followed by conclusions in section 4.

MATERIALS AND METHODS

Let $y(t)$ denote the response variable at time t, *e.g.* agricultural production, or productivity and r is the intrinsic growth rate. K signifies the resources yet to be achieved. Some important nonlinear growth models are (Seber and Wild, 2007):

(I) Monomolecular Model:

This model is given by

$$y(t) = K - (K - y_0)\ exp\ (-rt). \qquad (1)$$

Email: soumen.4345@gmail.com

where y_0 is the value of $y(t)$ at $t = 0$.

(ii) Logistic Model:

This model is given by

$$y(t) = K/[1 + (K/y_0 - 1) \exp(-rt)] \quad (2)$$

This model has sigmoid behaviour.

(iii) Gompertz Model:

This model also has sigmoid type of behaviour and is found quite useful in the biological work. It is given by

$$y(t) = K \exp[\ln(y_0/K) \exp(-rt)] \quad (3)$$

The above models are deterministic in nature. An error term is added on the right hand side of each one of these, thereby making these as 'Nonlinear statistical models' (Draper and Smith, 1998). The error term is assumed to be independently and identically distributed with equal variances. Nonlinear estimation procedure Gauss-Newton algorithm is employed for fitting the models. Goodness of fit (GOF) of fitted models is examined by computing Coefficient of determination (R^2), Mean absolute error (MAE), Root mean square Error (RMSE) and Mean absolute percentage error (MAPE) using the following formula:

$$R^2 = 1 - \sum_{t=1}^{n} (y_t - \hat{y}_t)^2 / \sum_{t=1}^{n} (y_t - \bar{y})^2 \quad (4)$$

$$\text{MAE} = \sum_{t=1}^{n} |y_t - \hat{y}_t| / n \quad (5)$$

$$\text{RMSE} = \left[\sum_{t=1}^{n} (y_t - \hat{y}_t)^2 / n \right]^{1/2} \quad (6)$$

$$\text{MAPE} = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (7)$$

where y_t is the actual observation for time period t, $\hat{y}_t$ the predicted value for the same period and $\bar{y}$ is the overall sample mean of observations.

In the present study, India's total groundnut production data (in Million tonnes) during the period 1950-51 to 2011-12 is considered. This data has been published by Department of Agriculture and Cooperation, Directorate of Economics and Statistics. Several sets of initial values for the parameters y_0, K and r have been tried and it is found that the final estimates remain the same.

RESULTS AND DISCUSSION

In the first instance, attempts were made to identify the model that best described this data set. Statistical software R (Ritz and Streibig, 2008) and SPSS have been employed for data analysis. A number of widely separated initial values were tried to ensure 'global convergence' and the results have been reported in table 1.

Table 1: Fitting of nonlinear growth models

Parameters/Statistics	**Growth model**		
	Monomolecular	**Logistic**	**Gompertz**
yo	3.2246 (0.4956)*	3.4043 (0.4075)	3.3231 (0.4440)
K	7.9193 (0.8411)	7.5505 (0.5083)	7.6872 (0.6251)
r	0.0361 (0.0163)	0.0623 (0.0195)	0.0491 (0.0178)
Goodness of fit statistics			
R^2	0.5547	**0.5571**	0.5560
MAE	**0.8103**	0.8155	0.8134
RMSE	1.0889	**1.0859**	1.0872
MAPE	**13.8848**	14.0384	13.9720

* *Figures in the parentheses indicate corresponding standard error*

It has been observed from table 1 that all the three models have been fitted well with the dataset. However, when we compared the models based on various GOF criteria, it was found that Monomolecular and Logistic models performed better than the Gompertz model. In case of MAE, RMSE and MAPE, the bold entries indicate the minimum values whereas for R^2, the bold figure specifies the maximum value. Graphs of fitted Monomolecular and Logistic models along with observed data are exhibited in Fig. 1 and 2. The x-axis and y-axis indicate time in year (t) and production in Million tonnes (y) respectively. In the figures, points represent the observed y-values and the curved lines represent estimated y-values.

However, before taking final decision on suitability of model, assumption of independence of errors needed to be examined by employing Run test on residuals. The results of the test obtained in respect of the three models, are reported in table 2. As calculated value of Z-statistic, in each case, is less than the tabulated value, it can be concluded that null hypothesis of independence of errors was not rejected at 5% level of significance for any model.

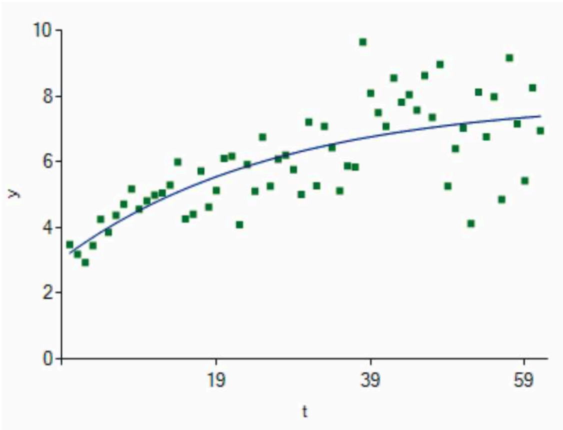

Fig. 1: Graph of fitted Monomolecular model along with observed data

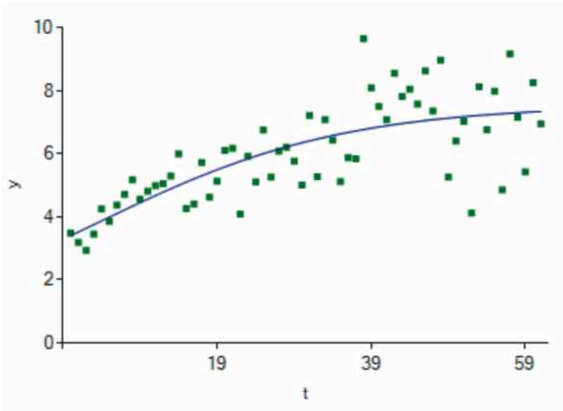

Fig. 2: Graph of fitted Logistic model along with observed data

Table 2: Run test

	Residuals from fitted model		
	Monomolecular	**Logistic**	**Gompertz**
Total Cases	62	62	62
Number of Runs	30	28	28
Value of Statistic	-0.442	-0.996	-0.996
Asymp. Sig. (2-tailed)	0.658	0.319	0.319

An important assumption of regression analysis, either linear or nonlinear, is that the residuals from fitted model should be normal. The assumption of normality of errors is checked by Kolmogorov-Smirnov and Shapiro-Wilk tests on residuals are depicted in table 3. It is observed from the table that for all the three models, the residuals passed the test of normality.

Table 3: Tests of Normality

Residuals from fitted model	**Kolmogorov-Smirnov**			**Shapiro-Wilk**		
	Statistic	**df**	**Sig.**	**Statistic**	**df**	**Sig.**
Mono-molecular	0.084	62	0.2	0.984	62	0.585
Logistic	0.075	62	0.2	0.984	62	0.596
Gompertz	0.074	62	0.2	0.984	62	0.590

As it has been observed that Monomolecular and Logistic models perform better than Gompertz model, point forecasting of India's groundnut production has been done by using these two models as represented in table 4.

Table 4: Forecasting groundnut production of India

	Production in million tonnes	
Year	**Monomolecular**	**Logistic**
2014-15	7.453	7.383
2015-16	7.469	7.393
2016-17	7.485	7.403
2017-18	7.501	7.411
2018-19	7.516	7.420
2019-20	7.530	7.427

Three nonlinear growth models *viz*. Monomolecular, Logistic and Gompertz have been employed for modeling time series data of India's total groundnut production over more than 60 years. All the three models were fitted well for the dataset under study, however, based on different measures of in-sample forecasting accuracy, Monomolecular and Logistic models outperform the Gompertz model. The residuals from fitted models also obey the assumptions of independence and normality which are validated by run test and test of normality respectively. Finally, point forecasting of India's total groundnut production for the year 2014-15 to 2019-20 has been carried out by using the suitable models. It has been observed that the country's groundnut production is likely to be increased with a slow and steady rate. Considering the increased population growth of the country, and the other oilseed production, this may be a matter of concern for the policy makers of our country. Finally, the approaches advocated here are very general and may be used, if fitted, for forecasting of time series data of other crops as well.

REFERENCES

Choksi, R. 2013. Consumption of edible oil in India - A paradigm shift. http://www.fnbnews.com/article/detnews.asp?articleid=34433&sectionid=32.

Dhekale, B. S., Sahu, P. K., Vishwajith, K. P., Mishra, P. and Noman, M. 2014. Modeling and forecasting of tea production in West Bengal. *J. Crop Weed*, **10**: 94-103.

Draper, N.R. and Smith, H. 1998. *Applied Regression Analysis*. 3rd Edn. John Wiley & Sons. U.S.A., pp. 543-49.

Iquebal, M. A. and Sarika. 2013. Nonlinear growth models for describing country's Lentil (*Lens culinaris* M.) production. *J. Fd. Legumes*, **26**: 79-82.

Joshi, P. K. and Saxena, R. 2002. A profile of pulses production in India. Facts, trends and opportunities. *Ind. J. Agric. Econ.*, **57**: 326-39.

Panwar, S., Singh, K. N., Kumar, A., Sarkar, S. K., Paul, R., Rathore, A. and Sivaramane, N. 2014. Forecasting of growth rates of wheat yield of Uttar Pradesh through non-linear growth models, *Ind. J. Agric. Sci.*, **84**: 856-59.

Patil, B. N., Bhonde S. R. and Khandikar, D. N. 2009. Trends in area, production and productivity of groundnut in Maharashtra. *Financing Agriculture*, March-April: 35-39.

Prajneshu and Chandran, K.P. 2005. Computation of compound growth rates in agriculture : Revisited. *Agril. Eco. Res. Rev.*, **18**: 317-24.

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. *J. App. Sci.*, **10**: 2419-25.

Ritz, C. and Streibig, J. C. 2008. *Nonlinear Regression with R*. Springer.

Sarma, M. 2005. Study on agricultural growth performance of Assam. *Econ. Affairs*, **50**: 38-41.

Seber, G. A. F. and Wild, C. J. 2007. *Nonlinear Regression*, 2nd Edn. John Wiley & Sons, New York, U.S.A.