Agriculture Update
Volume 7 | Issue 1 & 2 | February & May, 2012 | 63-67

Research Article

Coffee production modelling in India using nonlinear statistical growth models

T.L. MOHAN KUMAR, C.S. SATHISH GOWDA, M.B. DARSHAN AND S. SHEELA RANI

Article Chronicle :
Received :
12.10..2011;
Revised :
20.01.2012;
Ac1cepted :
17.02.2012

SUMMARY : Efforts have been made in this paper, to develop appropriate nonlinear statistical models with a view to provide analytical approach to describe the coffee production trends in India. To this end, attempts were made to apply six nonlinear statistical growth models. The parameters of each model were estimated using Levenberg -Marquardt (LM) iterative method. The main assumptions of 'independence' and 'normality' of error terms were examined by using respectively, the 'Run-test' and 'Shapiro-Wilk test'. The best model was selected based on the performance of several model goodness of fit criteria *viz.*, R^2, MAE, MSE, RMSE, MAPE, AIC and BIC. MMF and Logistic models were found to be quite successful for describing the pattern of coffee production. Forecast values were also computed using two best fitted models. A comparative study indicated that both selected models were performed similarly for forecasting coffee production for the years 2015 and 2020.

How to cite this article : Mohan Kumar, T.L., Gowda, Sathish C.S., Darshan, M.B. and Sheela Rani, S. (2012). Coffee production modelling in India using nonlinear statistical growth models. *Agric. Update*, **7**(1&2): 63-67.

Key Words :
Coffee production, Nonlinear growth models, Levenberg-Marquardt iterative method, Run-test, Shapiro-Wilk test

Author for correspondence :
T.L. MOHAN KUMAR
Division of Biometrics and Statistical Modelling, Indian Agricultural Statistical Research Institute (IASRI), NEW DELHI (INDIA)
Email: monis.iasri@gmail.com
See end of the article for authors' affiliations

BACKGROUND AND OBJECTIVES

Coffee is the most preferred beverage in the world by virtue of its special flavour and taste. Another feature of coffee is that the crop is more sensitive to climatic and soil requirements, hence, it is confined mostly to the tropical situations in South India. India ranks sixth in world coffee production after Brazil, Vietnam, Columbia, Indonesia and Ethiopia. Coffee production is mainly confined to the Southern States of India *viz.*, Karnataka, Kerala, and Tamil Nadu, which form the traditional coffee tracts. For the past five to six years, the productivity of coffee in India has been around 800 kg/ha. The India's coffee production was only 1.80 lakh tonnes in 1991-1992, thereby continued to add to the production by accounting an all time highest production of 3.02 lakh tonnes in 2010-2011. According to post blossom estimate, it will continue to add India's coffee production by 3.22 lakh tonnes for the year 2011-2012. The share of India's coffee production compared to world production was 3.18 per cent during 1992-1993, which was peaked to 4.46 per cent during 2001-2002 and again production fell down to around 3.78 per cent in 2010-2011. India is the largest coffee exporter in Asia. Coffee is predominantly an export oriented commodity in India with 70 to 75 per cent of the production being exported. Total quantity of 2.19 lakh tonnes of coffee was exported from India in 2011. The share of India's export to global trade was accounted around 4.14 per cent in 2009-2011 (Anonymous, 2011). Although there has been an increasing trend in the production of coffee from the country, there also are apparent wide fluctuations from year to year.

Growth rate analyses are widely employed to describing the long-term trends in variables over time in various agricultural crops. Growth models are generally 'mechanistic' and the parameters have meaningful biological interpretation. Some isolated attempts have been made in the past to investigate quantitatively the growth pattern of coffee production in the country. Chengappa

(1981) studied the growth rates of area, production and productivity of coffee in India using linear and exponential growth models. Prakash (1986) analyzed growth rates of production, consumption and export of Indian coffee using a modified exponential growth function. Bastine and Palanisamy (1994) analyzed the growth rates of area, production and productivity of major crops in Kerala including the cereal crops and the plantation crops using linear models. Pavitha (2005) applied cubic functional form model in order to analyse the performance of Indian coffee production and export. However, the assumption of linearity or exponential functional form may not hold for the real data in nature. There are few research studies in many branches of science which have demonstrated that more complex nonlinear functions are justified and required, since relationships among variables in agricultural sciences are normally "nonlinear" in nature. Prajneshu and Das (1998) carried out a detailed study dealing with modelling of wheat production data at State level in post-Green revolution era using several mechanistic nonlinear growth models. Prajneshu and Das (2000) applied some important nonlinear mechanistic growth models to examine critically state wise wheat productivity pattern in India. Khamis *et al.* (2005) studied the basic needs of parameters estimation for nonlinear growth model such as partial derivatives of each model and determination of initial values for each parameter. Prajneshu (2008) highlighted the deficiencies and emphasized that nonlinear estimation procedures should be employed for fitting the original model. Recently, Narinc *et al.* (2010) fitted several nonlinear growth models for modeling age-related changes of body weight growth in Japanese quail. This is partly attributed to the fact that the statistical methodology used for fitting nonlinear growth models to study coffee production trends in India has been not explored yet.

Keeping above points in view, the present study was aimed to develop appropriate nonlinear statistical models with a view to provide analytical approach to describe the coffee production trends in India. Based on performance of models fit, two best nonlinear models were chosen for future projection of coffee production trends in India.

RESOURCES AND METHODS

The secondary data pertaining to production of coffee (quantity in lakh tonnes) in India were collected for the period of 21 years (1991-92 to 2011-2012) from Database on coffee, Market Intelligence and Statistical Unit, Coffee Board, Bangalore. With a view to describing the pattern of coffee production, six nonlinear statistical growth models (Seber and Wild 1989) were employed and equations are presented in Table A.

In Table A, Y(t) represents the coffee production quantity at time t; α, β, k, m are parameters and ε denotes the error term.

Table A: Nonlinear statistical growth models considered in this study

Sr. No.	Models	Expression
1.	Monomolecular	$Y(t) = \alpha - (\alpha - \beta)\exp(-kt) + \varepsilon, \quad \beta = Y(0)$
2.	Logistic	$Y(t) = \dfrac{\alpha}{1 + \beta\exp(-kt)} + \varepsilon, \ \beta = \dfrac{\alpha}{Y(0)} - 1$
3.	Gompertz	$Y(t) = \alpha\exp(-\beta\exp(-kt)) + \varepsilon, \ \beta = l/Y(0))$
4.	Richard's	$Y(t) = \dfrac{\alpha}{[1 + \beta\exp(-kt)]^{1/m}} + \varepsilon, \ \beta = \dfrac{\alpha^m}{Y(0)^m} - 1$
5.	Weibull	$Y(t) = \alpha - \beta\exp(-kt^m) + \varepsilon$
6.	Morgan-Mercer-Flodin (MMF)	$Y(t) = \dfrac{\alpha\beta + kt^m}{\beta + t^m} + \varepsilon,$

The parameter 'k' is the 'intrinsic growth rate', while the parameter 'α' represents the 'carrying capacity'. For the third parameter, although the same symbol 'β' was used, yet this represented different functions of the 'initial value' Y(0) for different models and '*m*' is the added parameter (Prajneshu and Das, 2000).

It may be noted that all the above six models are 'nonlinear', as each one of these involves at least one parameter in a nonlinear manner. Parameter estimates in nonlinear case also can be obtained by minimizing the residual sum of squares. However, because of nonlinearity, the resulting normal equations are nonlinear in parameters and so cannot be solved exactly. Accordingly, a number of the iterative procedures have been developed to obtain approximate solutions. Four main methods are available in the literature to obtain estimates of the unknown parameters of a nonlinear regression models which are: (i) Linearization method, (ii) Gradient method, (iii) Does not use derivatives (DUD) method, and (iv) Levenberg-Marquardt (LM) method.

The linearization method uses the results of linear least squares theory in a succession of stages. However, the drawback of this procedure is that it may converge very slowly or oscillate widely or may not converge at all. In short, neither this method nor Gradient method can be recommended in practice. The LM procedure represents a compromise between these methods and combines successfully the best features of both as well as avoids their serious disadvantages. The DUD procedure, as the name suggests, is a derivative-free method. Nowadays most of the standard statistical software packages like SPSS, and SAS contain programmes for fitting nonlinear models by LM and DUD procedures. However, convergence to biologically meaningful values cannot always be guaranteed by any procedure and the success rate is quite low. The details of these methods are given in Seber and Wild (1989). The present statistical analysis was carried out by using the LM procedure available in PROC NLIN facility of SAS software package. The LM iterative method requires specification of the initial estimates of each parameter of the models to be estimated. Initial value specification is one of the most difficult problems encountered in estimating parameters of nonlinear models. In appropriate initial values will result in longer iteration, greater execution time, nonconvergence of the iteration and possibly convergence to unwanted local minimum sum of squares residual. To start the iterative procedure, many sets of initial values were tried to ensure global convergence. The iterative procedure was stopped when the reduction between successive residual sums of squares was found to be negligibly small. More details on methods of finding initial estimates of the parameters of models can be found in Seber and Wild (1989).

The main assumptions of 'independence' and 'normality' of error terms were examined by using, respectively the 'Run-test' and 'Shapiro-Wilk test' (D'Agostino and Stephens, 1986). It may be mentioned that none of these assumptions was violated for any data set and model combination considered in this study. Finally, the goodness of fit of a model is assessed by computing the Co-efficient of Determination (R^2), Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), Akaike's Information Criteria (AIC) and Schwarz Bayesian Information Criteria (SBC or BIC) (Narinc *et al.,* 2010).

OBSERVATIONS AND ANALYSIS

Six nonlinear statistical growth models, *viz.,* Monomolecular, Logistic, Gompertz, Richard's, Weibull and Morgan Mercer Flodin (MMF) were applied for coffee production data (in lakh tonnes) in India. Study revealed that the series had upward trends with fluctuation in quantity produced in India. The parameters estimate of the fitted models along with standard errors are presented in Table 1. The LM iteration procedure was converged for all growth models with various numbers of iterations. Statistical significance of the parameters of the nonlinear models was determined by evaluating 95 per cent asymptotic confidence intervals of the estimated parameters. The Null hypothesis was rejected when 95 per cent asymptotic confidence interval of corresponding parameter estimate does not include zero. Except for the estimate of parameter β in Gompertz model, all other parameters estimates of the above six nonlinear growth models were found to be statistically significant at the 5 per cent level. The main assumptions of 'independence' and 'normality' of error terms of each model were examined by using, respectively the 'Run-test' and 'Shapiro-Wilk test', and test statistic along with probability values which are presented in Table 1. Results revealed that the number of runs was found to be non-significant ($Z < 1.95$ or p-value > 0.05) and Shapiro-Wilk statistic was also non-significant (P-value > 0.05) at 5 per cent significance level for all the six models. Hence, the assumptions of randomness of residual and normal distribution of residual are satisfied.

The test statistics to check the goodness of fits were evaluated for each model that are presented in Table 1. All the six models are significantly fitted to production data. The best fitted model based on goodness of criteria could be ranked as follow; MMF, Richards, Logistic, Monomolecular, Gompertz and Weibull growth models. From the test for goodness of fit

Table 1: Parameters estimates of different models for coffee production (in lakh tonnes) data

Parameter	Models					
	Monomolecular	Logistic	Gompertz	Richard's	Weibull	MMF
α	3.032* (0.175)	2.946* (0.112)	2.381* (0.123)	2.974* (0.329)	3.000* (1.873)	2.598* (0.055)
β	1.615* (0.166)	0.826* (0.155)	-0.854^{NS} (0)	0.4322* (7.008)	2.000* (7.244)	2.0E-11* (3.6E-10)
k	0.147* (0.056)	0.235* (0.068)	-0.022* (0.004)	0.175* (0.291)	0.500* (2.080)	1.968* (0.092)
m	-	-	-	0.640* (8.888)	0.500* (2.009)	-12.410 * (9.723)
Test for randomness and normality of residuals						
Runs test Z: (p –value)	-0.817^{NS} [0.414]	0.011 ^{NS} [0.992]	-1.785 ^{NS} [0.074]	-0.887 ^{NS} [0.375]	-0.650 ^{NS} [0.516]	-0.689 ^{NS} [0.491]
Shapiro-Wilk (W): (p-value)	0.952 ^{NS} [0.364]	0.937 ^{NS} [0.190]	0.931 ^{NS} [0.146]	0.953 ^{NS} [0.390]	0.958 ^{NS} [0.511]	0.970 ^{NS} [0.750]
Goodness of fit criteria						
R^2	0.782	0.994	0.989	0.994	0.657	0.832
MAE	0.167	0.167	0.203	0.168	0.202	0.144
MSE	0.049	0.047	0.082	0.053	0.075	0.037
RMSE	0.206	0.201	0.265	0.210	0.245	0.171
MAPE	6.920	6.820	8.339	6.672	7.790	5.811
AIC	-60.312	-61.354	-49.846	-57.751	-52.177	-67.112
SBC(BIC)	-57.179	-58.221	-46.712	-53.573	-47.910	-62.934

* indicate significance value and NS=Non-significance value at 5% level of significance;
Values in bracket indicate Standard errors; Values in square bracket indicate probability values.

to coffee production data, lowest value of statistics such as MAE (0.144), MSE (0.037), RMSE (0.171), MAPE (5.811), AIC (-67.112) and BIC (-62.934) showed that MMF model performed better than other models, but on the basis of R^2 (0.994), both Richard's and Logistic model were performed similarly and fitted better than other models. On the contrary, statistics such as MAE (0.203), MSE (0.082), RMSE (0.265), MAPE (8.339), AIC (-49.846) and BIC (-46.712), showed that Gompertz model was least performed than other models, but based on statistics, R^2 (0.657), Weibull model was least performed than other models.

Based on the performance of model fit and goodness of fit criteria, two best nonlinear models were chosen for future projection of coffee production in India. As depicted in Table 1, MMF, and Logistic models fitted very well to production data sets in view of the low goodness of criteria values and highest R^2. Thus, both of these models may be selected as the best model for describing the coffee production trends. The identified MMF and Logistic models were utilized for forecasting the coffee production for the years 2015 and 2020, and forecasted values are exhibited in Table 2. Forecasting of coffee production in India on the basis of MMF model was found to be 2.859 and 2.810 lakh tonnes in 2015 and 2020 respectively. Forecasting of coffee production in India on the basis of Logistic model was found to be 2.941 and 2.947 lakh tonnes in 2015 and 2020, respectively. Forecasting value of both models reveals that, there will be no massive difference in coffee production in 2015 and 2020; its shows that there was constant growth pattern in coffee production after 2001 up to 2009, thereafter increase and again decreases in production in future. Forecasted values of MMF models were slightly smaller than the Logistic model.

Table 2: Forecasting of coffee production in India for the year 2015 and 2020

Years	MMF model	Logistic model
2015	2.859	2.941
2020	2.810	2.947

The objective of this research was to develop nonlinear statistical growth models with a view to provide analytical approach to describe the coffee production trends in India. To evaluate growth properly, it is required to select a suitable growth curve and its parameters should be able to be interpreted biologically (Karadavut *et al.,* 2010). The overall goodness of fit statistics has shown that the MMF and Logistic models have the best fitting to coffee production data set. These results are in good agreement with following research: Prajneshu and Das (2000) reported that Logistic model was found to be quite successful in describing the path of wheat productivity in India (Khamis *et al.,* 2005). Research study was found that Logistic model is the best model for modeling oil palm yield growth. Panwar *et al.* (2009) reported that the Logistic model was found to be suitable model for describing trends in onion production of India.

In short, it is evident from the analysis that production series had upward trends with fluctuation in quantity produced. Study also revealed that, among six nonlinear models, MMF and Logistic models were found to be most suitable models to describe trend in coffee production data. On the contrary Gompertz and Weibull models were found to be least suitable to describe trend in coffee production data. Both MMF and Logistic models were performed almost similarly for forecasting production quantity of coffee for the years 2015 and 2020. Forecasted values of MMF models were slightly smaller than the Logistic model.

Authors' affiliations:

C.S. SATHISH GOWDA, Department of Agricultural Economics, University of Agricultural Sciences, BENGALURU (KARNATAKA) INDIA (Email: cssg86@gmail.com)

M.B. DARSHAN, Division of Post Harvest Technology, Indian Agricultural Research Institute (IARI), NEW DELHI (INDIA) (Email: darshandachan@gmail.com)

S. SHEELA RANI, Department of SS and AC, University of Agricultural Sciences, BENGALURU (KARNATAKA) INDIA (Email: sheelusmasi @gmail.com)

REFERENCES

Anonymous (2011). Database on coffee, Market Intelligence & Statistical Unit, Coffee Board, BENGALURU, KARNATAKA (India).

Bastine, C.L. and Palanisami, K. (1994). An analysis of growth trends of principal crops in Kerala. *Agric. Situ. India*, **18**(12) : 885-891.

Chengappa, P.G. (1981). Growth rates of area, production and productivity of coffee in India. *J. Coffee Res.,* **11**(2) : 19-26.

D'Agostino, R.B. and Stephens, M.A. (1986). *Goodness of fit techniques*. Marcel Dekker, NEW YORK.

Karadavut, U., Palta, C., Kokten, K. and Bakoglu, A. (2010). Comparative study on some non-linear growth models for describing leaf growth of maize. *Internat. J. agric. Biology,* **12**(2) : 227-230.

Khamis, A., Ismail, Z., Haron, K. and Mohammed, A.T. (2005). Nonlinear growth models for modeling oil palm yield growth. *J. Mathematics & Statistics,* **1**(3): 225-233.

Narinc, D., Karaman, E., Firat, M.Z. and Aksoy, T. (2010). Comparison of nonlinear growth models to describe the growth in Japanese quail. *J. Animal & Vet. Advances*, **9**(14): 1961-1966.

Panwar, S., Sharma, S. C. and Kumar, A. (2009). A critical study of onion (*Allium cepa*) export of India: non-linear approach. *Indian J. agric. Sci.,* **73**(3): 178-180.

Pavitha, M. T. (2005). Post liberalization scenario of coffee industry -A case study of coorg district. M.Sc. (Ag.) Thesis, University of Agricultural Sciences, DHARWAD, KARNATAKA (India).

Prajneshu (2008). Fitting of Cobb-Douglas production functions: Revisited. *Agric. Eco. Res. Rev.,* **21**: 289-92.

Prajneshu and Das, P.K. (1998). Modelling wheat *(Triticum aestivum)* production at state level in post-green revolution era. *Indian J. agric. Sci.,* **68**: 23-26.

Prajneshu and Das, P.K. (2000). Growth models for describing state-wise wheat productivity. *Indian J. agric. Res.,* **34**(3):179-181.

Prakash, T.N. (1986). A study on supply response, market margins and growth rates of Indian coffee. M.Sc. (Ag.) Thesis, University of Agricultural Sciences, BENGALURU, KARNATAKA (India).

Seber, G.A.F. and Wild, C. (1989). *Nonlinear regression*. John Willy, NEW YORK.