Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 76(2) 2022 59–65

Modeling on Apple Production in Himachal Pradesh

Anju Sharma[1], Med Ram Verma[2] and Satish Kumar Sharma[1]

[1]Dr Y.S. Parmar University of Horticulture and Forestry, Nauni, Solan

[2]ICAR- Indian Veterinary Research Institute, Izatnagar

Received 05 February 2022; Revised 09 September 2022; Accepted 15 September 2022

SUMMARY

Apple is one of the important horticultural crops of Himachal Pradesh. The present study was conducted to see the apple production trends in Himachal Pradesh. The apple production data for the period 1973-74 to 2017-18 was used to fit various linear and non-linear models to study the apple production trends. The area under apple increased with compound growth rate 2.89% and production increased with compound growth rate 3.67%. The parameters of the various linear and non-linear models were estimated for apple production..Theil's inequality and Chow test indicated that it was not appropriate to predict the area under apple crop in the state however the quadratic model fitted well among the three linear models to the production data of apple with highest adjusted R^2 (0.526) and lowest RMSE (134.216) values. Among the various non-linear models the Rational model was the best fitted model based on various goodness of fit criteria viz., MSE (17105.973), RMSE (130.79), MAE (96.308), and AIC(446.623) values. The assumptions of independence and normality of error terms were examined by the 'Run-test' and Shapiro-Wilk's test respectively. Durbin Watson test was used to examine autocorrelation among residuals for the various fitted models. From the present analysis of data it was observed that all the models followed the assumptions of linear and non-linear models. On comparing different statistics of analysis of both linear and non-linear models, the non-linear Rational model performed better for describing apple production in Himachal Pradesh.

Keywords: Adjusted R^2, MSE, MAE, AIC, Durbin Watson test, Non-linear models.

1. INTRODUCTION

Apple (*Malusdomestica*) is one of the important fruit crops of Himachal Pradesh. The contribution of area under apple production is 48.79% of the total area under fruit crops. The share of apple crop in total fruit production is 78.99% (Anonymous, 2018). Himachal Pradesh is the second largest apple producing state of the country and is responsible for improving the socio-economic status of rural people in the apple growing belt of the state. The area under apple and its production has increased from 91,804 hectares and 2,68,402 Metric Tonnes in2006-07 to 1,11,896 hectares and 4,68,134 Metric Tonnes in 2016-17, respectively. However, the overall fruit production as well as the apple production slumped from 6.12 and 4.68 lakh tones in 2016-17 to 5.00 and 4.46 lakh tonnes in 2017-18, respectively (Anonymous, 2018).

Growth rate analyses are widely employed to describe the long-term trends in variables over time in various agricultural crops. The non-linear models play an important role for forecasting of agricultural production, animal population and import/export of the commodities in the country. Non-linear growth models are generally 'mechanistic' in nature rather than empirical and the parameters provide meaningful biological interpretation The purpose of analyzing trends in area and production is to examine the performance of apple cultivation in the state and see if there had been any noticeable changes during the last few years, which would help in finding out the factors responsible for such performance and thereby permit a broad judgment about the overall production possibilities in near future. Non-linear statistical models are more reliable in judgments and more frequently used by different authors [Venugopalan and Shamasundaran, 2003; Singh *et al.*, 2012; Panwar *et al.*, 2014; Singh

Corresponding author: Anju Sharma
E-mail address: anjusharma_uhf@yahoo.com

et al., 2014; Kumar *et al.*, 2018, Mishra and Thakur, 2021]. It is very important to have timely forecast in order to plan the strategies to increase the yield.

The objective of the present study was to compare linear and non-linear statistical models with a view to provide analytical approach to describe the apple production trends in Himachal Pradesh.

2. MATERIALS AND METHODS

For the present study time series data on area and production of apple in Himachal Pradesh were compiled for the period 1973-74 to 2017-18 from Directorate of Horticulture, Shimla H.P.

The compound growth rates (CGR) were computed for area, production and productivity of apple for the period of 1973-1974 to 2017-2018 using the function $Y = ab^t$ and compound growth rates were calculated by using the following formula:

CGR = (Antilog b-1)* 100

$$\text{Standard error (SE) of CGR} = \frac{100\text{ b}}{\text{Log}_e 10} \times \text{SE Log b}$$

Student's t test was used to test the significance of growth rates.

For analyzing the pattern of apple production the data on production various linear and non-linear models were fitted. Linear models included simple linear, quadratic and cubic models while non-linear models included Brody Modified (Brody *et al.* , 1924), Monomolecular, Gompertz, Logistic, Bertalanffy, Rational, and Hoerl models. Jhade and Singh (2020) used various nonlinear models for description of growth of the production and productivity of sugarcane crop in Uttar Pradesh. The descriptions of many nonlinear statistical models are given in Ratkowsky (1983), Seber and Wild (1989), Draper and Smith (1998) and France and Thornley (2006). Archontoulis and Miguez (2015) described many nonlinear models which are used in agriculture. The mathematical equations of the non-linear models which have been used in the present study are given in the Table 1.

In non-linear models, Y_t represents the observed apple area/production quantity at time t (year); a, b, c are parameters and 'e' denotes the error term. The parameter 'a' represents the asymptotic weight or the carrying capacity for each model; while 'b' denotes different functions of the initial value Y(0) or the scaling parameter; 'c' is the intrinsic growth rate and 'd'is the added parameter.

Table 1. Linear and nonlinear growth models

S. No.	Name of Model	Equation
	Linear models	
1	Linear model	$Y_t = a + bt + e$
2	Quadratic model	$Y_t = a + bt + ct^2 + e$
3	Cubic model	$Y_t = a + bt + ct^2 + dt^3 + e$
	Non- linear models	
1	Brody Modified model	$Y_t = a\exp(-bt) - a\exp(-ct) + e$
2	Monomolecular model	$Y_t = a(1 + b\exp(-ct)) + e$
3	Gompertz model	$Y_t = a\exp(-b\exp(-ct)) + e$
4	Logistic model	$Y_t = \frac{a}{1 + b\,\mathbf{exp}(-ct)} + e$
5	Bertalanffy model	$Y_t = a\left[1 - b\exp(-ct)\right]^3 + e$
6	Rational model	$Y_t = \frac{a + bt}{1 + ct + dt^2} + e$
7	Hoerl model	$Y_t = ab^t t^c + e$

These models are non-linear because each one of these involves at least one parameter in a non-linear manner. Parameter estimates can be obtained by minimizing the residual sum of squares. However, because of nonlinearity, the resulting normal equations are non-linear in parameters and hence cannot be solved exactly. Accordingly, a number of the iterative procedures have been developed to obtain approximate solutions. To obtain estimates of the unknown parameters of a non-linear regression model, Levenberg-Marquardt (LM method) technique was used.

The goodness of fit explains how well the given standard model fits a dataset. Measures of goodness of fit typically summarize the discrepancy between observed and predictive values under the model in question. In this research work the performance parameter of linear regression models was checked by adjusted R^2 and Root Mean Square Error (RMSE), while goodness of fit of the non-linear models was determined by

Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and The Akaike information criterion (AIC). The formulae used are as follows:

Adjusted Coefficient of determination ($\bar{R}^2$)

$$\bar{R}^2 = 1 - \frac{(1-R^2)(n-1)}{n-p-1}$$

where, n = number of observations, p = number of parameters R^2 is the

$$R^2 = 1 - \frac{\text{Residual Sum of Squares}}{\text{Total Sum of Squares}}$$

and coefficient of determination.

Theil's inequality coefficient (U)

$$U = \frac{\sum_{t=1}^{n-1}\left(\frac{\hat{Y}_{t+1} - Y_{t+1}}{Y_t}\right)^2}{\sum_{t=1}^{n-1}\left(\frac{Y_{t+1} - Y_t}{Y_t}\right)^2}$$

where Y_t is the actual value of a point for a given time period t, n is the number of data points, and $\hat{Y}_t$ is the forecasted value.

Interpreting Theil's U

Theil's U Statistic	Interpretation
Less than 1	The forecasting technique is better than guessing.
1	The forecasting technique is about as good as guessing.
More than 1	The forecasting technique is worse than guessing.

Chow test

When a regression model consists time series data, the values of the regression coefficients may vary with respect to time. If this variation in regression coefficients is significant, then the model is not valid for prediction. Chow test is one of the approaches to test the significance of change in the regression coefficients over the time by using the statistic.

$$F = \frac{\frac{\left[\sum e_p^2 - \left(\sum e_1^2 + \sum e_2^2\right)\right]}{p}}{\frac{\left(\sum e_1^2 + \sum e_2^2\right)}{n_1 + n_2 - 2p}}$$

where,

$\sum e_p^2$ = pooled error variance , $\sum e_1^2$ = First half sample error variance, $\sum e_2^2$ = Second half sample variance, n_1 = number of observations in first sample, n_2 = number of observations in second sample, p = number of regression coefficients in the model.

If $F \geq F_{tab}$, for ($k, n_1 + n_2 - 2p$) degrees of freedom then give data between F value is significant and the model is not valid for prediction.

Mean Squared Error (MSE)

$$MSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{(n-p)}$$

where, y_i = observed value, $\hat{y}_i$ = predicted value and p = number of parameters

Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{MSE}$$

Mean Absolute Error (MAE)

$$MAE = \frac{\sum_{i=1}^{n}|(y_i - \hat{y}_i)|}{n}$$

where, y_i = observed value and $\hat{y}_i$ = predicted value

The Akaike information criterion (AIC)

$$AIC = n \log_e MSE + 2P$$

where, n = number of observations and P = number of parameters

The Durbin-Watson (DW) statistic was used to check the presence of autocorrelation among the residuals of different fitted models. An important assumption of non-linear regression is that the residual 'e', or the dependent variable ' Y_t ' follows normal distribution. The assumption of 'normality' of error terms was examined by 'Shapiro-Wilk test' (Shapiro *et al.*, 1968). To check the randomness within residuals the Run test was used. The analysis of the data was done using SPSS software.

3. RESULTS AND DISCUSSION

3.1 Trends in area, production and productivity using compound growth rate

Himachal Pradesh has undergone a revolution in the apple production during last few decades. Apple production, which constituted 73.77 per cent of the total fruit production in the state during the year 2011-12 has gone up to 79% in 2017-18(Anonymous 2018). Its production has taken new strides and has increased from 2890 tonnes in 1966-67 to 4.46 lakh tonnes in 2017-18 in the state. The area under apple and its production has increased significantly at a compound growth rate (CGR) of 2.89 and 3.67 per cent per annum respectively between 1973-74 and 2017-18 as shown in Table 2. While the compound growth rate of productivity increased from 0.44 per cent in 2011-12 to 0.76 per cent in 2017-18. During 1966-67 the area under apples constituted nearly 58 per cent of the total fruit area in the state. In the later years, there has been relatively more emphasis of planting of other fruit trees in the state as a consequence of which the proportionate share of apple area has come down to nearly 49 per cent in the year 2017-18.

Table 2. Compound growth rate of area, production and productivity of apple in Himachal Pradesh during 1973-74 to 2017-18

Compound growth rate		
Area	Production	Productivity
2.89** (1.27)	3.67** (1.62)	0.76 (0.33)

Figures in parentheses are Standard errors;
** Significant at 1percent level of significance

The growth rate in productivity is an important determinant of agricultural transformation and is considered as the engine of growth to the farm economy. The crop productivity growth is an indicator of use of farming knowledge, technology, infrastructural development, farm investments, and development of suitable price policy. The scenario of continued deceleration in apple productivity is a cause of concern. This dismal growth in yield may be attributed to predominance of old and senile orchards, development of apple industry in rainfed conditions, global warming, low density of plantation, lack of efficient use of irrigation water, quality seeds and planting material, pollination problems, site selection, imbalanced use of resources etc.

3.2 Estimation of parameters and measures of goodness of fit

Fitting of linear models

Area and production as a function of time were investigated by fitting simple linear, quadratic and cubic models. The regression equations and corresponding values of adjusted R^2 and RMSE are presented in Table 3. The maximum value of adjusted R^2 and minimum value of RMSE values suggested that cubic and quadratic models were best fit for area and production respectively. The significance of best fitted function was tested through Theil's U statistic which was found to be 1.232 (>1) for area and 0.285 (<1) for production which indicated that model fitting for area under apple is worse than guessing. Further, the significant values of Chow F statistic in all the three models showed that none of the models were suitable for predicting the growth trends of area under apple production. Whereas, minimum non-significant Chow F statistic (0.87) value indicated the cubic model as best fit for apple production. Hence, for further analysis, non-linear models were tried on production data only.

Table 3. Fitting of basic regression models for area and apple production in Himachal Pradeshduring 1973-74 to 2017-18

S. No.	Model	F Statistics	Regression Equation	Adj R^2	RMSE	Theil's U	F(Chow test)
Area							
1	**Linear**	2476.121***	$X = 29.344 + 1.914\ t$	0.983	3.352	1.242	8.78*
2	**Quadratic**	1314.321***	$X = 27.147 + 2.194\ t - 0.006\ t^2$	0.984	3.255	1.182	12.70*
3	**Cubic**	904.150***	$X = 29.340 + 1.652\ t - 0.023\ t^2 - 0.0004\ t^3$	**0.984**	**3.206**	1.232	9.40*
Production							
1	**Linear**	49.494***	$Y=73.998 +10.856\ t$	0.524	134.438	0.285	1.40
2	**Quadratic**	25.400**	$Y=125.114 + 4.331\ t + 0.142\ t^2$	**0.526**	**134.216**	**0.285**	**0.87**
3	**Cubic**	16.636***	$Y=101.998+10.049\ t - 0.166\ t^2 + 0.004\ t^3$	0.516	135.607	0.291	1.34

***P<0.001

Fitting of non-linear models

The estimated parameters along with their standard errors and goodness of fit for seven different non-linear models are presented in Table 4 for apple production in Himachal Pradesh. A perusal into the goodness of fit values for apple production indicated that among the different non-linear models fitted, the lowest MAE (96.308) and RMSE (130.79) values were observed in Rational model as compared to all other non-linear models. Also, the AIC value was the lowest in case of Rational model. Fig. 1 shows the graphical representation of apple production using best fitted Rational model.

Table 4. Estimated parameters alongwith goodness of fit of different non-linear models

Model	Parameter	Estimate	Std. Error	MSE	RMSE	MAE	AIC
Brody Modified	**a**	138.498	28.101	17987.23	134.12	101.066	446.883
	b	-0.033	0.006				
	c	0.662	1.129				
Monomolecular	**a**	83348.151	1.501E7	18508.74	136.05	98.110	448.170
	b	-0.999	0.167				
	c	0.000	0.025				
Gompertz	**a**	34802.302	569576.987	17992.45	134.14	99.254	446.897
	b	5.646	16.064				
	c	0.007	0.026				
Logistic	**a**	1832.404	4875.301	17982.44	134.10	99.332	446.872
	b	13.790	35.784				
	c	0.043	0.029				
Bertalanffy	**a**	205055.805	8724664.344	18016.66	134.23	98.727	446.956
	b	0.918	1.205				
	c	0.002	0.026				
Rational	**a**	155.136	22.389	**17105.97**	**130.79**	**96.308**	**446.623**
	b	-3.371	0.558				
	c	-0.041	0.002				
	d	0.0001	0.000001				
Hoerl	**a**	106.638	69.861	17961.98	134.02	99.488	446.821
	b	1.029	0.016				
	c	0.119	0.320				

* Significant at 5 percent level of significance; ** Significant at 1percent level of significance; a – carrying capacity or intercept; b- function of initial value, c- intrinsic growth rate or slope; d- added parameter; MAE – Mean Absolute Error; MSE – Mean Square Error; RMSE – Root Mean Square Error

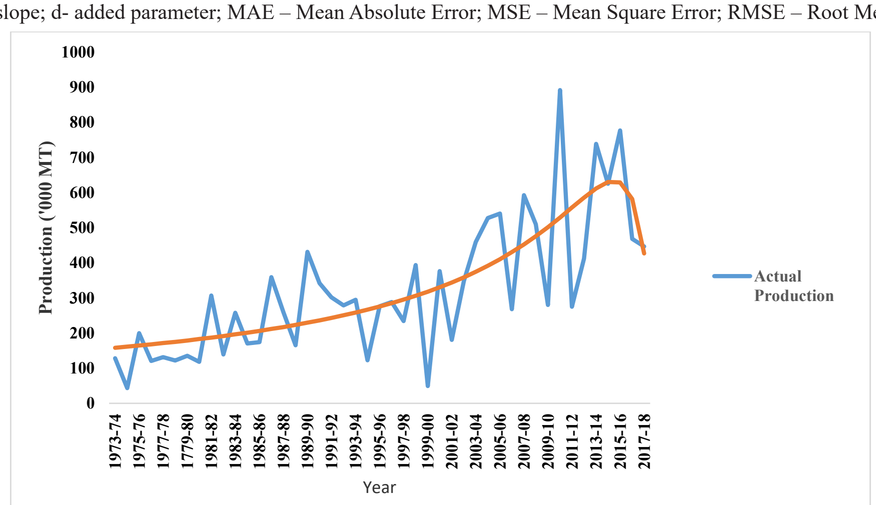

Fig. 1. Graph of actual values and best fitted 'rational' model for apple production in Himachal Pradesh

Testing of normality, independence and autocorrelation among the residuals for apple production

In order to test the autocorrelation among the residuals of different fitted models Durbin-Watson test was used while Shapiro-Wilk's test was used to test normality present in residuals. Error terms of each model were examined for independence by using Run test. The test statistic along with probability values are presented in Table 5. After the fitting of different models, residuals were estimated for all the models. After applying the Durbin-Watson test to the residuals of different models, it has been observed that in all the models residuals were not auto correlated. Shapiro-Wilk's test and Run test indicated that residuals of all the fitted models were random and normally distributed. Hence, none of these assumptions randomness of residual and normal distribution was violated for data set and model combination considered in this study. The P value of Shapiro-Wilk's test statistic (0.554) and the Run test statistic (1.210) to test for assumptions indicated that the residuals of the Rational model were normal and random respectively. The Durbin-Watson statistic recorded the value of 2.59 which indicated that there was no autocorrelation present in the dataset.

Table 5. Tests for the presence of normality, randomness and autocorrelation in residuals of apple production

Model	Shapiro Wilk Test		Run Test			DW statistic
	Statistic	P value	Runs	Z value	P value	
Brody Modified	0.980	0.605	26	0.607	0.544	2.52
Monomolecular	0.978	0.535	26	0.607	0.544	2.44
Gompertz	0.980	0.600	26	0.607	0.544	2.52
Logistic	0.980	0.617	26	0.607	0.544	2.52
Bertalanffy	0.981	0.646	26	0.607	0.544	2.51
Rational	0.978	0.554	28	1.210	0.226	2.59
Hoerl	0.980	0.604	26	0.607	0.544	2.52

Thus, based on the goodness of fit criteria the Rational model with a carrying capacity of 155.136 and the intrinsic growth rate of -0.41 can be regarded as the best fit for describing the trend in apple production during 1973-74 to 2016-17. Hence, Rational model can be used for projection of apple production in the state for future years. The non-linear models have been used by many researchers for the forecasting of agricultural production in India. Prajneshu and Chandran (2005) used the non-linear models for the computation of growth rates in agriculture. Iquebal and Sarika (2013) used non-linear models for describing the lentil production in India. Panwar *et al.* (2014) used non-linear models for forecasting of growth rates of wheat yield of Uttar Pradesh. Singh *et al.* (2014) used non-linear models for describing soybean production in Madhya Pradesh. Pal and Mazumdar (2015) used non-linear growth models for forecasting groundnut production of India. In a study carried out by Jhade and Singh (2020), Rational model was found to be most suitable for trend analysis of sugarcane in Uttar Pradesh. Dilliwar *et al.* (2016) used Brody Modified model along with other nonlinear models to study the growth trend of sheep and goat populations in India. Rajarathinam and Vetriselvi (2017) used Rational and Hoerl models along with other non-linear models to study the growth trends of cereals crop production in Gujarat. Rational and Hoerl models along with other non-linear models were used by Kumar *et al.* (2018) to study the growth trend of sugarcane production in Tamilnadu.

The apple production is influenced by weather conditions as well as other climatic factors besides the attack of insect-pests and diseases like wooly aphid, root borer, apple scab, early leaf fall etc. Thismay be the reason for ample fluctuation in apple production every year.

4. CONCLUSION

The estimated compound annual growth rates revealed that the area, production and productivity of apple has shown an increasing trend over the study period. Non-linear growth models proved to be an improvement over the conventional models and are more appropriate to visualize the temporal trend of production of apple in Himachal Pradesh. It is concluded from the study that the Rational model gave a better insight about estimates of the carrying capacity of the system and the intrinsic growth rate and is found to be the most suitable fitted model which clearly explained the trend of apple production in Himachal Pradesh.

ACKNOWLEDGEMENTS

The authors are highly thankful to the learned reviewer for his valuable comments on the original version of the paper.

REFERENCES

Anonymous. (2018). Directorate of Horticulture, Shimla, Himachal Pradesh, India.

Archontoulis, S.V. and Miguez, F.E. (2015) Nonlinear Regression Models and Applications in Agricultural Research. *Agronomy Journal*, **107(2)**, 786-798.

Brody, S.,Turner, C.W. and Ragsdale, A.C.(1924). The relation between the initial rise and subsequent decline of milk secretion following parturition. *Journal of General Physiology*, **6**, 541-45.

Dilliwar, L., Verma, M.R., Singh, Y.P., Sharma, V.B., Kumar, Sanjay and Shukla, A. (2016) Nonlinear modelling of sheep and goat populations in India. Journal of Applied and Natural Science, **8(4)**, 1766-1769.

Draper, N.R. and Smith, H. (1998): Applied Regression Analysis, Third Edition, John Wiley.

France, J. and Thornley, J.H.M. (2006): Mathematical Models in Agriculture, Second Edition, Butterworths.

Iquebal, M.A. and Sarika. (2013). Non-linear growth models for describing country's Lentil (*Lens culinaris* M.) production. *Journal of Food Legumes* 26(1&2): **79**, 82.

Jhade, S. and Singh, A. (2020). Comparison of non-linear models on area, production and productivity of sugarcane crop in Uttar Pradesh. *Pantnagar Journal of Research*, **18(3)**, 237-242.

Kumar, D.P.,Bakshi, B, and Manjunath,V.(2018). Non-linear Modeling of Area and Production of Sugarcane in Tamil Nadu, India. *International Journal of Current Microbiology and Applied Sciences*, 7(10), 3136-3146.

Kumar, M., Gowda, T.L., Sathish, C.S., Darshan, M.B. and Sheela Rani, S. (2012).Coffee production modelling in India using non-linear statistical growth models. *Agric. Update*. **7(1&2)**, 63-67.

Mishra, M. and Thakur, R.K. (2021) Comparative Study of Non-linear Growth Models for Area and Production of Sugarcane Crop in Uttar Pradesh. International Journal of Agricultural and Statistical Sciences, **17(2)**, 533-538.

Pal, S. and Mazumdar, D. (2015). Forecasting groundnut production of India using non-linear growth models. *Journal Crop and Weed.* 11 (Special Issue): 67-70.

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. *Indian Journal of Agricultural Sciences,* **84**, 856-859.

Prajneshu and Chandran, K.P.(2005).Computation of compound growth rates in agriculture :Revisited. Agricultural Economics Research Review, **18**, 317-324.

Rajarathinam, A. and Vetriselvi, R. (2017) Modeling cereals crop production based on parametric and non-parametric models. International Journal of Agricultural and Statistical Sciences, **13(1)**, 17-28.

Ratkowsky, D.A. (1983). Nonlinear Regression Modeling. A unified Practical Approach. Marcel Dekker New York.

Seber, G.A.F. and Wild, C.J. (1989). Non-linear Regression. John Wiley and Sons, New York.

Shapiro, S.S., Wilk, M.B. and Chen, H.J.(1968). A comparative study of various tests for normality. *Journal of the American Statistical Association*, **63(324)**, 1343-1372.

Singh, R.B., Meena, R., Tiwari, K.B. and Patidar, M. (2012). Non-linear model for prediction of area under wheat crop in Madhya Pradesh. *JNKVV Res J*. **46(2)**, 253-257.

Singh, U., Singh, R.B. and Gautam,S.S.(2014). Non-linear stochastic model for describing growth of soybean production in Madhya Pradesh and India. *International Journal of Agricultural and Statistical Sciences*, **10(2)**, 375-79.

Venugopalan, R. and Shamasundaran, K.S.(2003). Non-linear Regression: A realistic modeling approach in Horticultural crops research. *J.Ind. Soc.Ag.Statistics*, **56(1)**, 1-6.