Original Research Article

Non-Linear Growth Models for Acreage, Production and Productivity of Food-grains in Haryana

ABSTRACT

In order to understand the growth patterns of diverse commodities, agricultural research heavily relies on the computation of growth rates. For the computation of the growth rate, many researchers adopted the parametric approach rather than the non-linear model. In this paper, an attempt has been made to develop non-linear growth models for acreage, production and productivity of total (kharif + rabi) food-grains in Haryana from 1966 to 2021. We discussed different non-linear growth model viz. Logistic, Gompertz and Monomolecular and also determined the initial value for each parameter. The parameters were estimated using Levenberg - Marquardt's iterative method of non-linear regression. Best model was selected based on goodness of fit statistics such R^2, RMSE and MAE. Finally we concluded that Logistic model was found suitable to fit for acreage, production and productivity of food-grains grown in Haryana followed by Gompertz model. Forecasting for the period 2022–2026 was done using the selected best non linear growth model.

Keywords: Coefficient of determination, Gompertz, Logistic, Monomolecular, Non-linear growth model

Introduction

Food-grains production is by far the major activity in agriculture, covering about 80 percent of the cropped area in India and providing the main source of food. India is the second largest producer of food-grains in the world. India is also gifted with the diverse agro climatic conditions suitable for production of many important food-grain crops such as wheat, rice and bajra. In natural form, food-grains are a rich source of vitamins, minerals, carbohydrates, fats, oils, and proteins. Food-grains are deficient in the essential amino acid. That is why vegetarians combine their diet of grains with legumes. Thus, a combination of legumes and grains form a well-balanced diet for vegetarians. Food-grains provide almost all the calories and proteins consumed by the poor and also provide the employment to rural and poor people. Government policy in India has always given substantial importance to food-grains production. Such support, particularly since the beginning of the green revolution in the mid-

1960s, has contributed to noteworthy growth in this sector despite many constraints. With increasing the demand of poor for food-grains with a growing population, the country will require continuing high growth in production. Now a day's India has become self sufficient in food-grains, as the food-grains production tremendously increasing day by day. A number of studies have been undertaken to understand the growth pattern of food-grains in Haryana using the statistical tool and techniques ranging from graphs, charts, and diagrams to regression analysis and time series analysis. However, most of the studies are unable to explain the potential of food-grains production and productivity in Haryana and the period when this target is likely to be achieved. Hence, there is an urgent need to study the growth process of area, production and productivity by employing sound statistical methodology. Sharma (2013) analyzed the trend of area, production and productivity of foodgrains in Northeastern states by using the linear, quadratic and exponential functional forms. To fit the trend, linear functional form was used due to its higher R^2 value as compared to other two forms. Besides these, compound growth rate, coefficient of variation and instability index were also estimated. The effects of area, productivity and their interaction towards increasing production were also estimated in the study. They concluded that the increase in production is due to increase in area as well as interaction of area and productivity of food grain crops in the state. Borah and Mahanta (2013) determined a suitable method to estimate the parameters of some non-linear growth models using various methods of estimation. Six different estimation methods were used for estimating the parameters of the Monomolecular, Gompertz and Logistic growth models. Three different forestry data sets were used for testing the validity of the methods. The best fit method was selected on the basis of root mean square error. It was found that the most suitable growth model for the top height growth data is the Gompertz model which may be obtained by estimating the parameters using the Newton-Raphson method. Akshu and Sharma (2017) worked on district-wise trends in the growth of foodgrains production in Haryana and find growth in the production of foodgrains was argely affected by the increase in area under foodgrains. Singh and Supriya (2017) studied growth rate and trend analysis of wheat crop to evaluate growth in area, production and productivity of wheat crop in Uttra Pradesh during 1971 to 2001 by using different growth models they found a significantly increasing trend in area, production and productivity.

Material and methods

The present study was conducted to analyze the growth behavior of major food-grains production in Haryana for the time period 1966 to 2021. The secondary data of food-grains were obtained from Statistical Abstract of Haryana and other document published by Government of Haryana. The study compared different non linear growth models i.e. Logistic, Gompertz and Monomolecular models for estimating the growth in area, production and productivity of food-grains. To find the best fit model using the selection criteria such as R^2, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The data have been analysed using the statistical computing package Rstudio and SPSS 12.0.

Linear and Non-linear models

A linear model is one which output is directly proportional to input variables. In such a model, all the parameters appear a linear in the relationship. In contrast, non-linear model is one in which the linear relationship does not hold for at least one of the parameters. For example, the following equations:

$$y = a + b * t$$

$$y = a + b * t + c * t^2$$

represents a linear model, whereas equation:

$$y = y_0 * \exp(-a * t)$$

represents a non-linear model. Two types of non-linear models are found in practice. The following are the equations:

$$y = \exp\,(a + b * t^2)$$

and $$y = \exp(-b * t) - \exp\,(-a * t)$$

Although, both these models are non-linear in parameters a and b but they are of essentially different characters. Above equation can be transformed, by taking natural algorithm, into the form:

$$\ln\,(y) = a + b * t^2$$

which is linear in parameters.

Important Non-linear Growth Models:

Logistic model

Logistic model was developed by Pierre Verhulst in the mid of 1830s, who suggested that the rate of population increase may be limited, that is, it may depend on population size. The parameters of Logistic model have a simple physical interpretation. Logistic growth model provide an **S-shaped curve** and mainly used in the field where growth is symmetric about its point of inflection. Hence the Logistic growth model is given by

$$y = \frac{a}{(1 + b * \exp(-k * t))}$$

Monomolecular model

Monomolecular model describes the progress of a growth situation in which it is believed that the rate of growth at any time is proportional to the resources yet to be achieved. Monomolecular model is mainly used in the field where the growth rate decrease linearly as size increases and there is no inflection point. Monomolecular model may be written as:

$$y = a * [1 - b * \exp(-k * t)]$$

Gompertz model

Gompertz model was given by Gompertz in 1825 for the hazard in life tables. Gompertz model is very popular and mainly used in the field where growth is asymmetrical about the point of inflection. Gompertz model has sigmoid type of behaviour and is found quite useful in the biological phenomena. Gompertz model may be written as:

$$y = a * \exp -b * \exp(-k * t)$$

where y denote the growth at time t and a, b, k are the parameters. Parameter a represents carrying capacity; k is the intrinsic growth rate and b represents different functions of the initial value.

Initial Value Specification and Goodness of Fit

Starting value specification is one of the most difficult problems encountered in estimating parameters of non-linear models. However, the problem of specifying initial values of parameters can be solved with proper understanding of the definition of the parameters in the context of the phenomenon being modelled. Wrong starting values result in longer iteration, greater execution time, non-convergence of the iteration, and possibly convergence to an unwanted local minimum. The expressions that provide good starting values for some of the parameters have been developed (Borah and Mahanta, 2013). The most efficient order for determining starting values is a, b and k. Here we describe the method of three equidistant points for obtaining initial parameter values for rapid parameter used. Estimation of three parameter non-linear growth models used are, i.e Monomolecular, Gompertz and Logistic models. In this method, we select three equidistant points, t_1, t_2, t_3, from the given data of food-grains in Haryana. Let n be the number of observations, t_2 be the $\frac{t_1+n}{2}t$ observation and t_1 be the observation between the first observation and the $(n-2)^{th}$ observation so that the RMSE is least corresponding to that observation. And t_3 be the $(t_2+d)^{th}$ observation where $d = t_2 - t_1$. Then the required initial parameters for the Monomolecular, Gompertz and Logistic growth model are:

Parameter	Initial values		
	Monomolecular	Gompertz	Logistic
a	$\frac{y_2^2 - y_1 y_3}{2y_2 - y_1 - y_3}$	$exp\left(\frac{w_2^2 - w_1 w_3}{2w_2 - w_1 - w_3}\right)$	$\frac{2z_2 - z_1 - z_3}{z_2^2 - z_1 - z_3}$
b	$\frac{(y_2 - y_1)^2}{y_2^2 - y_1 y_3}\left(\frac{y_2 - y_1}{y_3 - y_2}\right)^{\frac{t_1}{d}}$	$-\left(\frac{(w_2 - w_1)^2}{w_3 - 2w_2 + w_1}\left(\frac{w_2 - w_1}{w_3 - w_2}\right)\right)^{\frac{t_1}{d}}$	$\frac{\left(\frac{(z_2-z_1)^2}{z_3-2z_2+z_1}\left(\frac{z_2-z_1}{z_3-z_2}\right)^{\frac{t_1}{d}}\right)}{\left(\frac{z_2^2-z_1z_3}{2z_2-z_1-z_3}\right)}$
k	$\frac{1}{d} ln\left(\frac{y_2 - y_1}{y_3 - y_2}\right)$	$\frac{1}{d} \ln\left(\frac{w_2 - w_1}{w_3 - w_2}\right)$	$\frac{1}{d} ln\left(\frac{z_3 - z_1}{z_3 - z_2}\right)$

Where $w_i = lny_i$ and $z_i = \frac{1}{y_i}$ for i = 1, 2 and 3.

Model fitting method and selection criteria

In non-linear model if it is not possible to solve the non-linear equations exactly this iterative procedure followed for estimation of parameters. Important iterative procedures to approximate analytical solution are: Linearization (Taylor series) method, Steepest descent method and Levenberg-Marquardt's method.Most often used method for computing non-linear least squares estimators is the Levenberg Marquardt's method. It is good in the sense that it almost always converges and does not 'slow down' even at the latter part of the iterative process. A brief description of this method is given below:

Let us consider the model:

$$y_i = f(x_i, \theta) + e_i \qquad i = 1,2,3, \ldots \ldots \ldots . . n.$$

Where y_i is the i^{th} observation of the dependent variable, x_i is the i^{th} value of the independent variable, $\theta = (\theta_1, \theta_2, \ldots..\theta_p)'$ are parameters, e_i is the error terms. e_i are assumed to be independent and follow N (0, σ^2). The residual sum of squares is:

$$S(\theta) = \sum [y_i - f(x_i, \theta)]^2$$

Let $\theta = (\theta_{10}, \theta_{20}, \ldots \ldots . \theta_{p0})$ be the vector of initial parameter values. Then the algorithm for obtaining successive estimate is essentially given by:

$$(H + \lambda I)(\theta_0 - \theta_1) = g,$$

Where

$$g = \frac{\delta S(\theta)}{\delta \theta}, \qquad H = \frac{\delta^2 S(\theta)}{\delta \theta^2}$$

I is the identity matrix and λ is a suitable multiplier.

There is no single best method to assess the goodness of fit of a model. There are many different methods that highlight different features of the data and the model. Graphical comparison provides a quick visual assessment of the goodness of fit. To examine model performance, coefficient of determination (R^2) is used. In addition to R^2, root mean square error (RMSE) and mean absolute error (MAE) are also used as the goodness of fit criterion to assess the suitability of fitted model. Coefficient of determination is calculated using the formula:

$$R^2 = 1 - \left(\frac{RSS}{TSS}\right)$$

Where RSS is the residual sum of square and TSS is the total sum of square. The R^2 value is an indicator measuring the proportion of total variation about the mean of the trait explained by the growth curve models. The coefficient of determination lies always between 0 to 1, and the fit of a model is satisfactory if R^2 is close to unity.

Root Mean Square Error is a kind of generalized standard deviation and is calculated as follows:

$$RMSE = \sqrt{\frac{\sum(y - \hat{y})^2}{n}}$$

Where y is the value of observed data and $\hat{y}$ is the predicted value, n is the number of observation. RMSE value is one of the most important criteria to compare the suitability of growth models. The best model is one with the lowest RMSE. Mean Absolute Error was calculated as using the equation:

$$MAE = \sum \frac{|y - \hat{y}|}{n}$$

Results and discussion

In the first instance, attempts were made to identify the non-linear growth model that best described in the food-grains data of acreage, production and productivity in Haryana. Firstly we divided the whole data of food-grains into two parts namely training data and testing data. Training data was taken from 1966 to 2016 and testing data was taken from 2017 to 2021. The best model was selected from the training data of food-grains. Testing data was used for validation purpose. The best model selected from the testing data was used for forecasting of acreage, production and productivity of foodgrains from 2022-2026. Various growth models like Logistic, Gompertz and Monomolecular were considered for the modelling of area, production and productivity under food-grains for the period 1966 to 2021 in state of Haryana. The most appropriate model for the area, production and productivity total (rabi+kharif) food-grains in Haryana was determined on the basis of training data. Different sets of initial parameter values was tried to ensure convergence of the parameter estimates for total food-grains. Set of estimated parameter values after convergence of total food-grains are presented in table 1. To examine model performance R^2, RMSE and MAE were used as the goodness of fit criterion and are presented in table 2. It was observed that Logistic was the best fitted model for production and productivity of total (kharif +rabi) food-grains with R^2 values 0.39, 0.97 and 0.98 respectively. In case of area all the models have poor performance as compared to production and productivity. Performance of various non-linear growth models for acreage, production and productivity of testing data are represented in table 3. It is observed that monomolecular model is the best fitted model with minimum percent relative deviation for area and productivity of total (kharif+rabi) food-grains in Haryana. However, gompertz model is the best fitted model on test data for production of total (kharif+rabi) food-grains in Haryana. Plot of fitted non-linear growth model along with observed data for area, production and productivity of food-grains are also represented, x-axis and y-axis indicate time in year and acreage/production/productivity of food-grains respectively. Plot show a good resemblance between observed and predicted values for production and productivity. Forecasted values of total food-grains for acreage, production and productivity in Haryana for the period 2022 to 2026 are presented, by using the best fitted model from testing data. It has been observed from forecasted value that Haryana total food-grains (kharif+rabi) area, production and productivity is likely to be increased with a slow and steady rate.

Table 1: Initial and estimated parameter values

Model		Initial values			Estimated values		
		a	**b**	**k**	**a**	**B**	**k**
Area	**Logistic**	5464.72	0.54	0.025	732618641.5	197325.59	0.004
	Gompertz	5677.10	0.47	0.23	82582.80	3.093	0.001

	Monomolecular	21656.84	22	0.096	35183.46	0.892	0.0001
Production	**Logistic**	19287.44	6.98	0.08	25118	7.38	0.05
	Gompertz	252.66	2.30	0.03	51240	2.88	0.02
	Monomolecular	21545.94	21	0.09	2391556	0.99	0.0001
Productivity	**Logistic**	4720.64	1.07	0.03	4287.75	4.85	0.06
	Gompertz	4867.53	1.01	0.02	5763.25	2.10	0.03
	Monomolecular	6450.06	0.90	0.01	193740.65	0.99	0.0001

Table 2: Goodness of fit statistics

Kharif+Rabi		**Logistic**	**Gompertz**	**Monomolecular**
Area	**R^2**	0.39	0.39	0.37
	RMSE	242.89	274.76	420.56
	MAE	189.22	220.19	369.19
Production	**R^2**	0.97	0.97	0.96
	RMSE	641.79	653.87	1602.71
	MAE	517.45	521.99	1205.65

Productivity	**R^2**	0.98	0.97	0.97
	RMSE	129.35	138.71	1278.01
	MAE	110.85	116.74	1128.21

Table 3: Performance of various non-linear growth models

Total (Kharif +Rabi)	**Year**	**Logistic**	**Gompertz**	**Monomolecular**
		RD (%)	**RD (%)**	**RD (%)**
Area	**2017**	3.40	8.70	17.20
	2018	2.21	3.01	12.56
	2019	2.90	2.48	11.93
	2020	0.81	4.73	13.95
	2021	1.89	3.27	12.0
	Average of RD (%)	**2.24**	**4.43**	**13.52**
Production	**2017**	12.35	11.25	26.31
	2018	3.56	3.82	18.03
	2019	0.46	2.86	20.49
	2020	12.5	15.68	8.89
	2021	7.95	11.01	12.4
	Average of RD (%)	**7.36**	**8.92**	**17.22**
Productivity	**2017**	11.49	8.76	62.90
	2018	3.56	1.85	59.73
	2019	6.52	3.97	62.21
	2020	8.22	12.29	54.43
	2021	0.98	4.87	56.73

	Average of RD (%)	**6.15**	**6.37**	**59.2**

Table 4: Forecasted values for the period 2022 to 2026

Year	Area	Production	Productivity
2022	4626.60	18348.94	4078.56
2023	4698.55	18556.56	4107.32
2024	4709.17	18816.23	4153.36
2025	4755.50	18998.23	4198.56
2026	4783.03	19178.26	4246.89

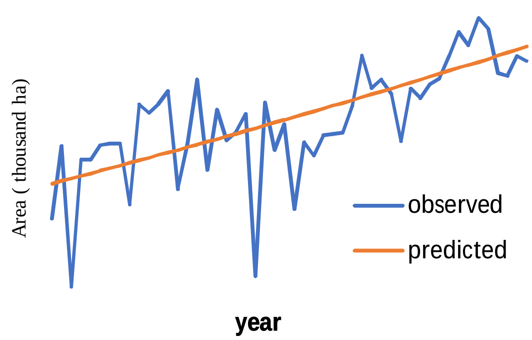

Fig. 1: Plot of observed vs predicted value for total (kharif +rabi) foodgrains area for the year 1966-2020

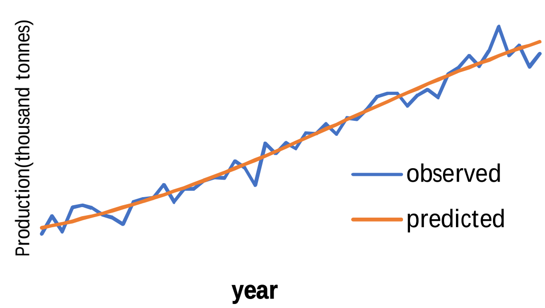

Fig. 2: Plot of observed vs predicted value for total (kharif +rabi) foodgrains production for the year 1966-2021

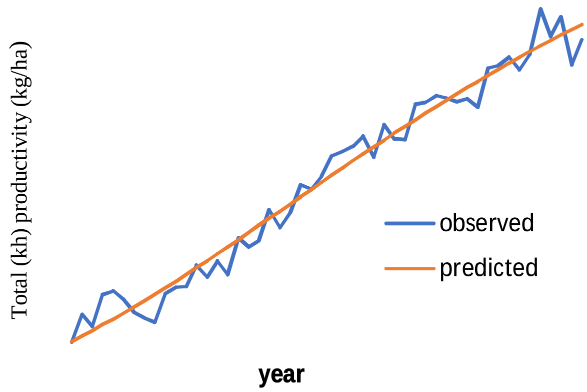

Fig. 3: Plot of observed vs predicted value for total (kharif +rabi) foodgrains productivity for the year 1966-2021

Conclusion

Several package like R, SPSS and SAS are readily available to fit the non linear growth models for computation of growth rates. In this paper R and excel software has been used for computation and forecasting of area, production and productivity of foodgrain in Haryana. Various growth models like Logistic, Gompertz and Monomolecular were considered for the modelling. To examine model performance R^2, RMSE and MAE were used as the goodness of fit criterion It was observed that Logistic was the best fitted model for production and productivity of total (kharif +rabi) food-grains with R^2 values 0.97 and 0.98 respectively.

REFERENCES

1. Akshu & Sharma, L. (2017). Production of foodgrain in Haryana: A District wise analysis. *International Journal of Academic Research and Development*, **2**(3), 182-187.
2. Archontoulis, S. V., & Miguez, F. E. (2015). Nonlinear regression models and applications in agricultural research. *Agronomy Journal*, *107*(2), 786-798.
3. Borah, M & Mahanta, D (2013). Rapid parameter eastimation of three parameter non linear growth models. *International Journal of Mathematical Archive*, ***4****(2), 274-282.*
4. Clark, A. J., Lake, L. W., & Patzek, T. W. (2011, October). Production forecasting with logistic growth models. In *SPE annual technical conference and exhibition*. OnePetro.
5. Dhekale, B. S., Sahu, P. K., Vishwajith, K. P., & Mishra, P. (2014). Modeling and forecasting of tea production in West Bengal. *Journal of crop and weed*, **10**(2), 94-103.
6. Draper,N.R., & Smith, H. (1998), Applied Regression Analysis. (vol.326). John Wiley & Sons.
7. Hojjati, F., & Ghavi Hossein-Zadeh, N. (2018). Comparison of non-linear growth models to describe the growth curve of Mehraban sheep. *Journal of Applied Animal Research*, *46*(1), 499-504.
8. Karadavut, U., Palta, Ç., Kökten, K., & Bakoğlu, A. (2010). Comparative study on some non-linear growth models for describing leaf growth of maize.
9. Panghal, P., Kumar, M. and Rani, S. (2019). Estimation of annual compound growth rates of guava (Psidium guajava L.) fruit in Haryana using Non linear model. Journal of Applied and Natural Science, 11(4), 778 – 784.
10. Rajan, S.M. &Palanivel, M. (2017). Comparison of non-linear models to describe growth of cotton. *International Journal of Statistics and Applied Mathematics*, **2**(4), 86-93.
11. Raji, A. O., Mbap, S. T., & Aliyu, J. (2014). Comparison of different models to describe growth of the Japanese quail (Coturnix Japonica). *Trakia Journal of Sciences*, **12**(7), 182-188.
12. Narinc, D., Karaman, E., Firat, M. Z., & Aksoy, T. (2010). Comparison of Non-linear Growth Models to Describe the Growth in Japanese Quail. *Journal of Animal and Veterinary Advances*, ***9***(14), 1961–1966.

13. Sharma, A. (2013). Trends of area, production and productivity of foodgrain in the north eastern states of India. *Indianjournal of agricultural research*, **47**(4), 341-346.
14. Singh, M., &Supriya, K. (2017). Growth Rate and Trend Analysis of Wheat Crop in Uttar Pradesh, India. *International Journal of Current Microbiology and Applied Sciences*, ***6***(7), 2295–2301.
15. Singh, R., Kumar, M., Dahiya, M. and Baloda, S. (2019). Development of growth model for Ber powdery mildew in relation to weather parameters. Indian Phytopathology. 72(1), 1-7.