International Journal of Chemical Studies

P-ISSN: 2349–8528
E-ISSN: 2321–4902
IJCS 2019; 7(3): 2544-2547

Received: 16-03-2019
Accepted: 18-04-2019

Gomathi T
Department of Physical Sciences & Information Technology, AEC&RI, TNAU, Coimbatore, Tamil Nadu, India

Vasanthi R
Department of Physical Sciences & Information Technology, AEC&RI, TNAU, Coimbatore, Tamil Nadu, India

Padmarani S
Department of Agricultural Economics, CARDS, TNAU, Coimbatore, Tamil Nadu, India

Patil SG
Department of Physical Sciences & Information Technology, AEC&RI, TNAU, Coimbatore, Tamil Nadu, India

Kalpana M
Department of Social Science, ADAC&RI, Trichy, Tamil Nadu, India

Correspondence
Vasanthi R
Department of Physical Sciences & Information Technology, AEC&RI, TNAU, Coimbatore, Tamil Nadu, India

Statistical analysis of area and production of banana in Trichy district: A nonlinear approach

Gomathi T, Vasanthi R, Padmarani S, Patil SG and Kalpana M

Abstract
This study aim to identify the best model for area & production of banana in Trichy. A secondary data was collected over a period of 30 years. There are four nonlinear models such as Sinusoidal, Rational, Gaussian, Hoerl model is used. The best model is selected based on R^2, RMSE and MAE values. According to area & production of banana, Rational model is found to be the best since it has highest R^2(71.8%), Lowest Root Mean Square Error (1515), lowest Mean Absolute Error (1189) for area and highest R^2 (57.4%), Lowest Root Mean Square Error (88589), lowest Mean Absolute Error(64846) for production. Hence, Rational model is the most suitable fitted model and it is very useful to study the trend analysis of area and production of banana in Trichy.

Keywords: Nonlinear models, Banana, R^2, RMSE, MAE

Introduction
In India, banana crop accounts for 2.8% of agricultural GDP. Tamil Nadu ranks first in area and production (5136.2MT) of banana**.** In 2015-2016, banana is cultivated an area of 841.2ha, with production of 29134.8MT and productivity of 34.6MT/ha for overall India. In 2016-2017, banana is cultivated an area of 858.1ha, with production of 29162.6 MT and productivity of 34 MT/ha for overall India (Source: Department of Agriculture and Cooperation). In Tamil Nadu, Trichy is the major district in banana production. This study is mainly focused on the computing the suitable Non-linear regression models. This study helps to know about the trend analysis in area &production of banana in Trichy district. A statistical model is a mathematical model that embodies a set of statistical assumptions concerning the generation of sample data (and similar data from a larger population. A mathematical model is an equation or a set of equations which represents the behavior of a system (France and Thornley, 1984). It can be either 'linear' or 'nonlinear'. A linear model is one in which all the parameters appear linearly. A 'nonlinear model' is one in which at least one of the parameters appears nonlinearly. More formally, in a 'nonlinear model', at least one derivative with respect to a parameter should involve that parameter. Examples of a nonlinear model are:

$$Y(t) = \exp(at+bt^2)$$

The term 'intrinsically linear' to indicate a nonlinear model which can be transformed to a linear model by means of some transformation. Parameters in a nonlinear model can also assessed by the method of least squares. Levenberg-Marquardt's method is the most widely used method for computing nonlinear least square estimators. Because it is always converges and does not slow down at the latter part of the iterative process. Prajneshu and P.K. Das (2000) [8] applied some important nonlinear growth models to inspect analytically state wise wheat productivity pattern in India.

Materials and Methods
Annual data of area and production of banana in Trichy for the period of 1985-2015 were collected from Season and crop Report Tamil Nadu. There are four nonlinear models used for this study.

Assumptions of error term
Nonlinear regression error term should follow normality and randomness. Shapiro-wilk test was used for normality.

The test statistic value of '*W*' ranges from 0 to 1. When *W* = 1 the given data are perfectly normal in distribution (Shapiro, *et al.,* 1968) [1]. When '*W*' is significantly lesser than 1, the assumption of normality is not met.

The test statistic is $W=\frac{(\sum_{i=1}^{n} a_i\ x_{(i)})^2}{\sum_{i=1}^{n}(x-\bar{x})^2}$ Run test: The run test can be used to decide if a dataset is from a random process. The test statistic is

$Z = \frac{r-\mu_r}{\sigma_r}$, Mean=$\mu_r = \frac{2n_1n_2}{n_1+n_2} + 1$

Standard deviation= $\sqrt{\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}$

n_1=Number of positive values in the series

n_2= Number of negative values in the series. The run test rejects the null hypothesis, if

$\frac{\propto}{2}$ |Z|>Z $_{1-\frac{\propto}{2}}$

Table 1: Nonlinear model equation

Name of the model	Model equation
Sinusoidal model	$Y = A + B\cos(CX - D) + e$
Gaussian model	$Y = \frac{A\exp\{-(B-X)^{2}\}}{2C^2} + e$
Rational model	$Y = \frac{A+BX}{1+CX+DX^2} + e$
Hoerl model	$Y = AB^X X^C + e$

Goodness of Fit of a Model

Goodness of fit of a model is assessed by computing the co-efficient of determination R^2, Root Mean square error (RMSE) and Mean Absolute error (MAE). Coefficient of determination is defined as the proportion of the variance in the dependent variable that is predictable from the independent variable(s). Kvalseth (1985) [9] has prioritize that

$R^2 = 1-\frac{\sum_{i=1}^{n}(Y_i-\widehat{Y_1})^2}{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}$ $0< R^2< 1$

Root Mean Square Error (RMSE)

RMSE is defined as the square root of the average of squared errors.

$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i-\widehat{Y_1})^2}{n}}$

Mean Absolute Error

The lower value of these statistics is considered as the best fitted model.

$MAE = \frac{\sum_{i=1}^{n}|Y_i-\widehat{Y_1}|}{n}$

Table 2: Estimation of parameters for fitted nonlinear models of area (ha) in Trichy district (1985-2015)

Parameters	Gaussian model	Hoerl model	Rational model	Sinusoidal model
a	14965**	12745**	10082**	10626**
b	-10	0.9**	2572**	3094**
c	32*	0.1	-0.07 d=0.01	0.1 ** d=-0.8*
R^2	0.643	0.687	0.718	0.708
RMSE	1702.79	1596	1515	1540
MAE	1352	1262	1189	1280
Shapiriowilks Test (w)	0.468^{NS} (0.968)	0.881^{NS} (0.983)	0.974^{NS} (0.988)	0.913^{NS} (0.984)
Runs test(z)	-2.920** (0.003)	-3.267** (0.001)	-2.529^{NS} (0.011)	-2.529^{NS} (0.011)

* indicate significance value at 5% level of significance and NS-Non Significant value

* * indicate significance value at 1% level of significance

Values in square bracket indicates probability values

Results and Discussion

There are four nonlinear models were used for studying the area and production of banana in Trichy district such as Gaussian, Rational, Hoerl, Sinusoidal models were used. The criteria for deciding the best model was based on R^2,RMSE, MAE. Table: 2 showed that Rational model was the best model among the four models. Because it has maximum R^2 (71.8%), minimum RMSE (1515) and minimum MAE (1189). According to Shapirio-wilks test and run test, the residuals of Rational model were normality and randomness attained. The next best fitted model is sinusoidal model which has R^2 value 70.8%. The scatter diagram (fig 1, 2) which confirmed error assumptions.

Rational model expression for area

$$Y = \frac{10082 + 2572 * X}{1 + (-0.07) * X + 0.01 * X^2}$$

Table 3: Estimation of parameters for fitted nonlinear models of Production of banana in Trichy district (1985-2015)

Parameters	Gaussian model	Hoerl model	Rational model	Sinusoidal model
a	513608**	380642**	527733**	45253*
b	9.416*	0.9**	-53917**	10865
c	20**	0.3	-0.08 d=-0.001	0.2** d=-2.054**
R^2	0.215	0.267	0.574	0.307
RMSE	120316	116267	88589	113069
MAE	88166	86239.93	64846	86396
Shapiriowilks Test(w)	0.008^{NS} (0.902)	0.039^{NS} (0.928)	0.364^{NS} (0.964)	0.093^{NS} (0.942)
Runs test(z)	-1.624^{NS} (0.104)	-1.350^{NS} (0.177)	-0.224^{NS} (0.823)	-1.457^{NS} (0.145)

* indicate significance value at 5% level of significance and NS-Non Significant value
* * indicate significance value at 1% level of significance
Values in square bracket indicates probability values

Table 3; which indicate that Rational model was the best model among the four models. Because it has maximum R^2 (57.4%), minimum RMSE (88589) and minimum MAE (64846).According to Shapirio-wilks test and run test, the residuals of Rational model were normality and randomness attained. The next best fitted model is sinusoidal model which had 30.7% R^2 value. The scatter diagram (fig 3, 4) which confirmed error assumptions.

Rational model expression for production

$$Y = \frac{527733 + (-53917) * X}{1 + (-0.08) * X + (-0.001)X^2}$$

Trend analysis which is helpful to study about the fluctuations. Due to climatic factors there is a lot of fluctuations in banana area and production in Trichy district. Here, Rational model was the best fitted model which is helpful to study about the fluctuations for banana area and production in Trichy district.

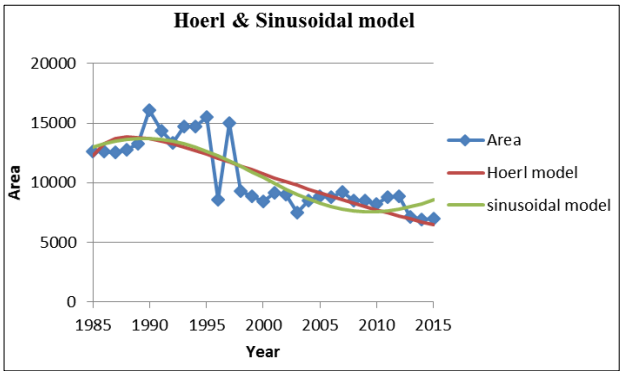

Fig 1: Graph for actual values and predicted values for banana area

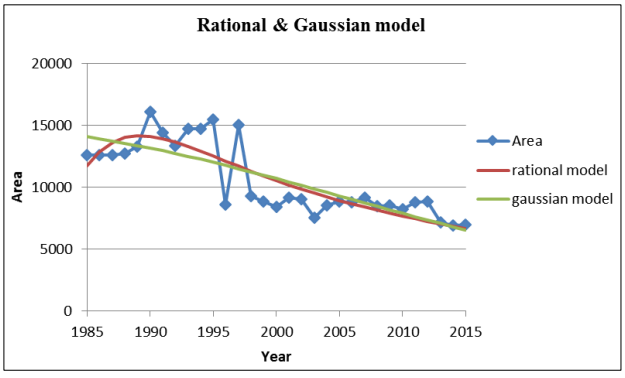

Fig 2: Graph for actual values and predicted values for banana area

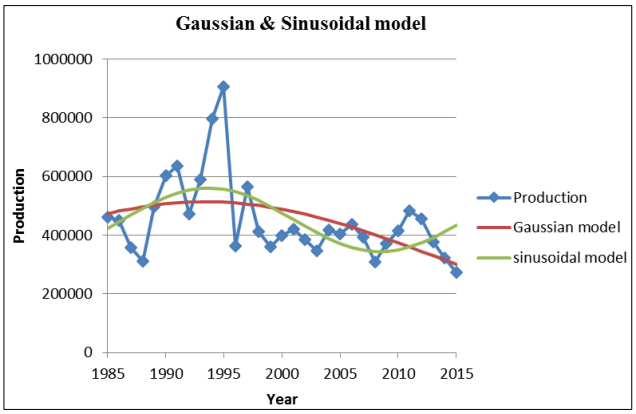

Fig 3: Graph for actual values and predicted values for banana production

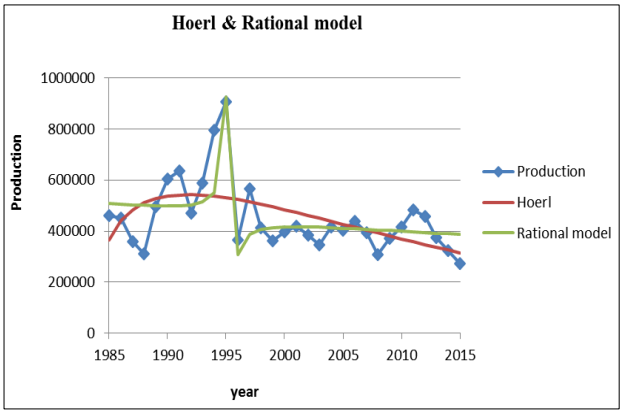

Fig 4: Graph for actual values and predicted values for banana production

References

1. Shapiro SS, Wilk MB, Chen HJA. Comparative study of various tests for normality Journal of the American Statistical Association. 1968; 63(324):1343-1372.
2. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika. 1965; 52(3-4):591-611.
3. Venugopalan R, Shamasundaran KS. Nonlinear Regression: A realistic modeling approach in Horticultural crops research. Journal of the Indian Society of Agricultural Statistics. 2003; 56(1):1-6.
4. Kumar TM, Gowda CS, Darshan MB, Rani SS. Coffee production modelling in India using nonlinear statistical growth models. Agriculture Update. 2012; 7(1/2):63-67.
5. Khan Shaista, Bhardwaj RK, Saxena RR, Bhagat RK, Jatav GK. Non-linear statistical modelling for area,

production and productivity of rice crop in chattisgarh. An Asian Journal of Soil Science. 2013; 8(1):98-102.
6. Ribeiro TD, Savian TV, Fernandes TJ, Muniz JA. The use of the nonlinear models in the growth of pears of 'Shinseiki' cultivar. Ciência Rural. 2018; 48(1).
7. Dinesh Kumar P, Bishvajit Bakshi, Manjunath V. Nonlinear Modeling of Area and Production of Sugarcane in Tamil Nadu, India. International Journal of Current Microbiology and Applied Sciences. 2018; 7(10):3136-3146.
8. Das PK. Growth models for describing state wise wheat productivity. Indian Journal of Agricultural Research. 2000; 34(3):179-181.
9. Kvålseth TO. Cautionary note about R^2. The American Statistician. 1985; 39(4):279-285.