Agric Res J **59** (5) : 985-989, *October* 2022
DOI No. 10.5958/2395-146X.2022.00139.9

FORECASTING COTTON PRODUCTION IN INDIA BY NONLINEAR GROWTH MODELS

S Govinda Rao[1]*, D Ramesh[2] and V Rajendra Prasad[3]

[1]*Department of Statistics and Computer Applications, , Agricultural College, Naira, Andhra Pradesh*
[2]*Department of Statistics and Computer Applications, Acharya N G Ranga Agricultural University, Agricultural College, Bapatla-522101, Andhra Pradesh*
[3]*Department of Agricultural Economics, Acharya N G Ranga Agricultural University, Agricultural College, Naira-532185, Andhra Pradesh*

Cotton is an important fibre and cash crop of India and plays a dominant role in the industrial and agricultural economy of the country. It provides the basic raw material (cotton fibre) to cotton textile industry. Cotton in India provides direct livelihood to 6 million farmers and about 40 -50 million people are employed in cotton trade and its processing (Anonymous-1). In India, there are ten major cotton growing states which are divided into three zones, viz. north zone, central zone and south zone. North zone consists of Punjab, Haryana, and Rajasthan. Central zone includes Madhya Pradesh, Maharashtra and Gujarat. South zone comprises Andhra Pradesh, Telangana, Karnataka and Tamil Nadu. Besides these ten states, cotton cultivation has gained momentum in the Eastern State of Orissa. Cotton is also cultivated in small areas of non-traditional states such as Uttar Pradesh, West Bengal & Tripura. Indian cotton is exported to countries. Cotton plays an important role in the Indian economy as the country's textiles industry is predominantly cotton based. India is one of the largest producers as well as exporters of cotton yarn. The Indian textiles industry contributes around 5 percent to country's gross domestic product (GDP), 14 percent to industrial production and 11 percent to total export earnings. The industry is also the second-largest employer in the country after agriculture, providing employment to over 51 million people directly and 68 million people indirectly, including unskilled women. Cotton yarn and fabrics export account for about 23 percent of India's total textiles and apparel export. In 2019–20, India's cotton production was 36.04 million bales of 170 kgs each. During 2019–20, export of cotton yarn, cotton fabrics, cotton made-ups and handloom products reached US$ 10.01 billion (Anonymous-1). Bangladesh and China are the largest export destinations for Indian cotton. Pakistan, Vietnam, Sri Lanka, Korea, Egypt, Portugal, Indonesia and Turkey are other main export destinations for Indian cotton. Nonlinear growth models have widely been used to measure agricultural growth in terms of growth rate (Patil *et al*., 2009; Rajarathinam*et al*., 2010). In view of its economic importance, present study is formulated to fit different nonlinear growth models for modeling and forecasting the area, production and productivity of cotton in India by using data from 1950-2019.

Data relating to area, production and productivity of Indian Cotton was collected for a period of seventy years from 1950 to 2019 from IndiaStat website (https://www.indiastat.com/), with a view to identify the growth as well as to predict for next five years. To study the growth trends of area, production and yield of Indian Cotton, the following nonlinear models were selected.

(1) Exponential Model: y=ae^(bx); (2) MMF Model: y=(a*b+c*x^d)/(b+x^d); (3) Richards Model: a/(1+exp(b-c*x))^(1/d); (4) Power Model: y=ax^b; (5) Hoerl Model: y=a*(b^x)*(x^c);

These models are deterministic in nature; hence 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. In the study, Goodness of fit for these fitted models was verified 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.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} \qquad RMSE = \left[\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 / n \right]^{1/2}$$

$$MAE = \sum_{i=1}^{n} \left| (Y_i - \hat{Y}_i \right| / n \qquad MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100\%$$

In this Y is the actual, $\hat{Y}$ is the predicted and 'n' indicates number of data points. Further, the assumptions of the model residuals were tested for randomness and normality with help of run test and Shapiro-Wilk test ,respectively (Dasyam *et al*., 2015).

*Corresponding author : govinda.seepana@gmail.com
Date of receipt: 26.04.2021, Date of acceptance: 31.10.2022

Table 1. Fitting of nonlinear growth models to cotton- area, production and yield in India

Parameter/ Statistics	Nonlinear model				
	Exponential	Power	Hoerl	Richards	MMF Model
Area under cotton Production during 1950 to 2019 in India					
a	6.34E+01	5.04E+01	7.93E+01	5.31E+02	6.43E+01
b	7.96E-03	1.57E-01	1.01E+00	2.04E+01	6.12E+02
c			-1.15E-01	7.63E-02	2.60E+02
d				9.59E+00	1.20E+00
Goodness of fit (GOF) Statistics					
R^2	0.60	0.37	0.65	0.60	0.54
RMSE	10.68	13.38	10.06	10.68	11.43
MAPE	8.60	10.46	8.43	8.60	9.04
MAE	10.13	11.96	10.08	10.18	10.42
Cotton Production during 1950 to 2019 in India					
a	7.06E+01	3.31E+00	9.38E+01	4.57E+02	3.75E+01
b	2.84E-02	1.16E+00	1.03E+00	2.93E+02	9.67E+02
c			-1.17E-01	4.62E+00	1.82E+03
d				1.47E+02	1.32E+00
Goodness of fit (GOF) Statistics					
R^2	0.91	0.84	0.92	0.93	0.85
RMSE	38.17	51.96	37.77	33.62	50.28
MAPE	26.49	42.52	25.28	24.81	40.45
MAE	11.31	25.34	12.15	11.25	20.58
Cotton Yield during 1950 to 2019 in India					
a	1.83E+01	2.61E-02	5.39E+01	4.21E+03	-1.12E+01
b	4.29E-02	2.24E+00	1.05E+00	1.61E+01	8.72E+02
c			-4.06E-01	6.63E-02	1.25E+03
d				8.22E+00	1.30E+00
Goodness of fit (GOF) Statistics					
R^2	0.90	0.84	0.90	0.81	0.73
RMSE	31.86	39.47	30.58	50.28	51.31
MAPE	22.65	32.78	20.69	40.45	42.00
MAE	23.91	43.15	20.05	48.58	45.28

Table 2. Run test and Shapiro Wilk test statistics for testing the residuals of models

Cotton	Residuals from the fitted model	Run test (Prob. Value)	Shapiro Wilk test (Prob. Value)
Area	Hoerl	0.15	0.12
Production	Richards	0.21	0.08
Yield	Hoerl	0.35	0.12

Test for the randomness of the residuals

Randomness of residuals can be tested by using non-parametric run test (Siegel and Castellan, 1988). Let m be the number of element sof one kind (+ve sign residuals) and n be the number of element of the other kind (-ve sign residuals) in sequence of N = m + n. Here r is the no. of runs. If m or n is larger than 20, determine the value of Z as under Mean = $\mu_r \frac{2mn}{N} + 1$ = and

Standard deviation = $\sigma_r = \left[\frac{2mn(2mn - N)}{N^2(N-1)}\right]^{1/2}$

H_0: Residuals sequence is random. Vs H_1: Residual sequence is not random

$$Z = \frac{r - \mu_r}{\sigma_r} = \frac{r + h - \frac{2mn}{N} - 1}{\sqrt{\frac{2mn(2mn - N)}{N^2(N-1)}}} \text{ where } h = +0.5 \text{ if } r < \frac{2mn}{N} + 1$$

and h = -0.5 if $r > \frac{2mn}{N} + 1$

The significance of any observed value of Z computed by using above statistic is determined from a normal distribution table.

Test for normality of the residuals

The Shapiro–Wilk (1965) statistic is used to test whether the residuals are normally distributed or not. The test is based on n residuals, which have to arrange in increasing order so that $e_1 \leq e_2 .... \leq e_n$.

H_0: The residuals are normally distributed Vs H_1: Residuals are not normally distributed.

Test statistic W is defined as $W=S^2 / b$

Where $S = \sum_{i=1}^{k} a_i \left[e_{n+1-i} - e_i\right]; \; b = \sum_{i=1}^{n} \left(e_i - \bar{e}\right)^2;$

The parameter k takes the values, K=n/2 (when data 'n' is even); k= (n-1)/2 when n is odd. The values of co-efficients "a_i" for different values of n and k are given in Shapiro – Wilk (SW) table.

An attempt was made to identify the appropriate model among selected nonlinear models so as to describe the behaviour of area, production and yield of Indian Cotton. For this, statistical software R was utilized for data analysis. In this, different initial values were tried for the selected models in order to ensure 'global convergence' and the results have been reported in Table 1.

It has been observed from Table 1 that among all the selected models, Hoerl was selected as appropriate based on various Goodness of fit criteria i.e, R^2, RMSE, MAPE and MAE to area under production and productivity of Cotton. Similarly, Richards model was found to be appropriate for cotton production, based on same diagnostic criterion. Graphs of actual and fitted series data was exhibited in Fig. 1, 2 and 3, respectively for area, production and productivity.

However, before taking final decision on suitability of model, assumption of normality and independence of errors needed to be examined, for this Shapiro-Wilk test and Run test were employed, respectively. It was revealed that in both these tests prob. values were not less than 0.05, hence it was confirmed that normality and randomness existed in residuals of all three selected models at 5 percent level. Finally, by using the selected models, forecastings were for next time periods as represented in Table 3.

To conclude, five nonlinear growth models viz. Exponential, Power, Hoerl, Richards and MMF Model have been employed for modeling time series data of India's area under cotton production, production and productivity over the period of 70 years. Among these models, hoerl model fitted well for the dataset under study of area and productivity where Richards was best for Production, based on different measures of in-sample forecasting accuracy i.e., R^2, RMSE, MAPE and MAE. The residuals from this fitted model 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 cotton production for the year 2024 has been carried out by using the Richards model and it resulted with increasing trend. It has been observed that the country's cotton production is likely to be increased with a slow and steady rate. Similarly, area under cotton production and productivity

Table 3. Forecasting cotton area, production and productivity of India

Year	Area under Cotton Production (in Lakh ha)	Cotton Production (in Lakh bales of 170 kg each)	Cotton Productivity (Kg. Ha)
2020	116.29	456.61	398.03
2021	117.54	456.85	417.13
2022	118.81	457.21	437.18
2023	120.09	457.51	458.23
2024	121.39	458.69	480.33

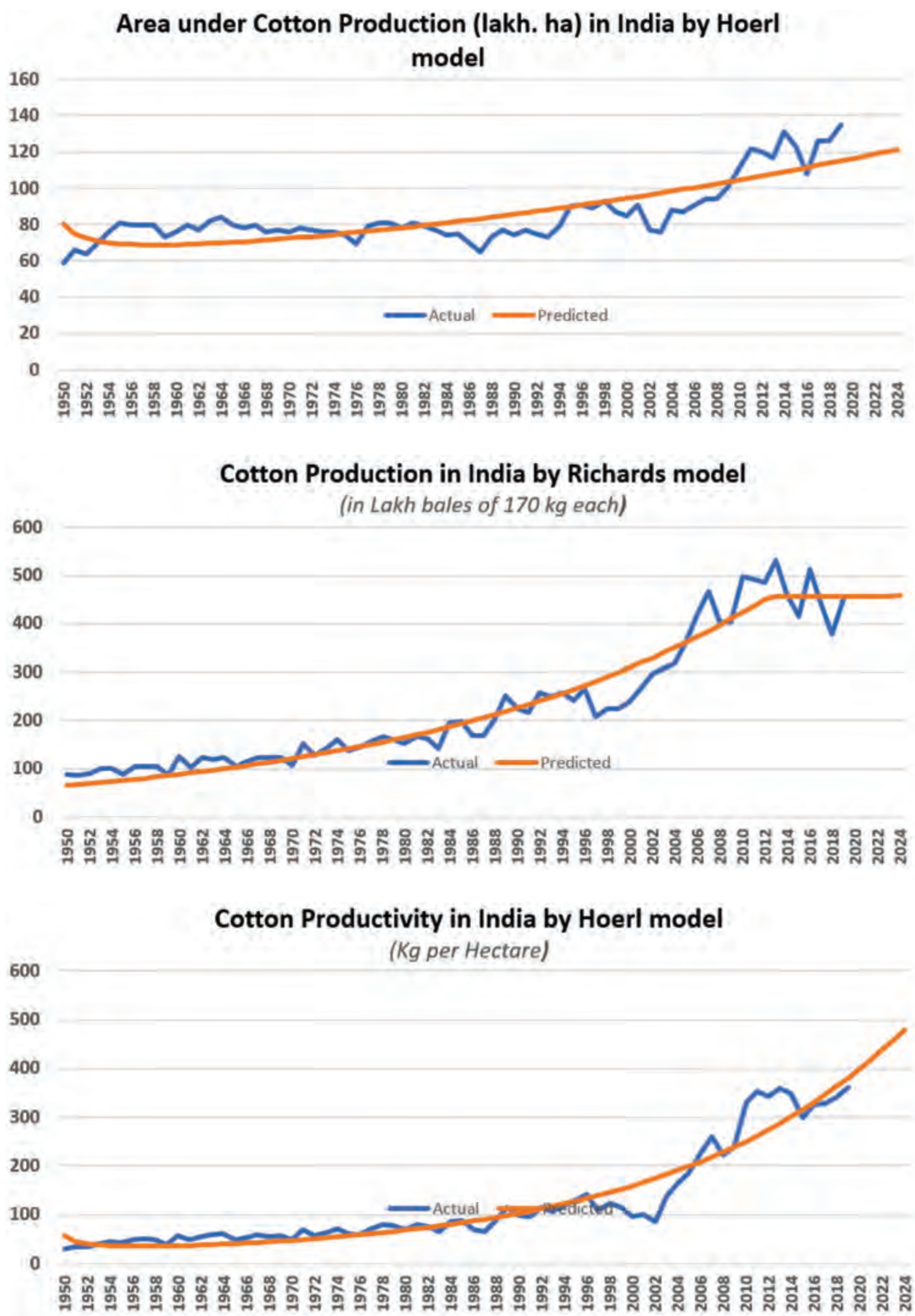

Fig. 1. Graph of fitted Hoerl model along with observed data

for the year 2024 were forecasted as 121.39(in Lakh ha), 480.33(Kg/ ha), respectively. These projections are important as these help in informing good policies with respect to relative production and price structure as well as consumption of cotton in our country. It can be concluded from the study, that the total cropped area can be increased in future, if land reclamation and conservation measures are mainly adopted. The projection shows that cotton will play dynamic role in expanding economic growth of India in future.

Authors' contribution

Conceptualization of research work and designing of experiments (SGR, DR, VRP); Execution of field/ lab work and data collection()SGR, DR, VRP; Data analysis and interpretation of results (SGR, DR, VRP); Preparation of manuscript(SGR, DR, VRP).

LITERATURE CITED

Dasyam R, Pal S, Rao V S and Banjul B 2015.Time Series modelling for trend analysis and forecasting Wheat production of India. *Int J Agri, Environ and Biotechnol* **8(2***):* 303-08.

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

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

Rajarathinam A, Parmer R S and Vaishnav P R 2010. Estimating models for productivity trends of Tobacco crops for Anand region of Gujarat State, India. *J Appl Sci* 10:2419-25.

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

Anonymous-1, https://plantix.net/en/blog/kharif-season-cultivating-cotton