Journal of Experimental Agriculture International

Volume 46, Issue 5, Page 202-219, 2024; Article no.JEAI.114659
ISSN: 2457-0591
(Past name: American Journal of Experimental Agriculture, Past ISSN: 2231-0606)

Modeling and Optimization of Groundnut Production in Vijayapura District of Karnataka, India

Sampatkumar S Biradar [a], Ragini H. R. [b++*] and Harshith K. V. [a]

[a] *Department of Agricultural Statistics, Applied Mathematics and Computer Science, University of Agricultural Sciences, GKVK, Bengaluru, India.*
[b] *Department of Agricultural Statistics, Uttar Banga Krishi Viswavidyalaya, West Bengal, India.*

Authors' contributions

This work was carried out in collaboration among all authors. Author SSB designed the study, conceptualized the primary idea, and formulated the research framework. Author SSB also played a pivotal role in developing the materials and methods section. Author RHR contributed significantly to data arrangement, meticulously organizing and structuring the research content. Author RHR expertise ensured the coherence and clarity of the study. Author HKV provided invaluable support throughout the research process. Author HKV actively participated in data analysis, offering critical insights, and consistently assisting in various aspects of the project. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JEAI/2024/v46i52371

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/114659

Original Research Article

Received: 11/01/2024
Accepted: 15/03/2024
Published: 18/03/2024

ABSTRACT

Groundnut is a significant oilseed crop in India, with Karnataka being one of the largest producers. The agricultural economy of Vijayapura district relies heavily on crop production, including groundnut. Understanding the production patterns and forecasting future yields is crucial for agricultural planning and economic sustainability. The study aimed to investigate the production

[++] *Ph.D. Scholar;*
**Corresponding author: E-mail: raginirgowda@gmail.com;*

patterns and forecast the groundnut yield in Vijayapura district, Karnataka. The research problem focused on understanding the trends in groundnut area, production, and productivity over time and developing a forecasting model for future yields. Secondary data from 1966-67 to 2020-21 was collected from various sources. Statistical models including linear, quadratic, cubic, exponential, log-logistic and GAM were used to analyze the trends. The ARIMA method was employed for forecasting. The models' adequacy was assessed using MAPE, R2, AIC, and BIC criteria. The log-logistic model was found to be the best fit for groundnut area trends, while the cubic model and GAM were best for productivity and production, respectively. Forecasting using ARIMA initially indicated a slight increase in groundnut yield, but the GAM model predicted a decrease in future production. The findings provide insights for policymakers, agricultural extension services and farmers to make informed decisions regarding crop planning, resource allocation and economic sustainability. Understanding the production patterns and forecasting future yields is crucial for agricultural planning and economic sustainability in Vijayapura district.

Keywords: Groundnut; production patterns; forecasting; Vijayapura district; Karnataka; agricultural economy; ARIMA model; log-logistic model; Generalized Additive Model (GAM).

1. INTRODUCTION

“Groundnut, also known as peanut, is a significant oilseed crop that plays a crucial role in the agricultural economy of India. It is grown in an area of about 85 lakh hectares with a total production of 84 lakh tones” [1]. “The crop is primarily cultivated in the states of Gujarat, Andhra Pradesh, Tamil Nadu, Karnataka, Maharashtra, Madhya Pradesh, Uttar Pradesh, Rajasthan, Punjab, and Orissa” [2,3]. “Groundnut is perfectly grown in well-drained sandy loam or sandy clay loam soil. Deep well-drained soils with a pH of 6.5-7.0 and high fertility are ideal for groundnut cultivation. The crop rotation of groundnut is very important as it helps in efficient nutrient utilization and reduces soil-borne diseases and nematodes” [4]. “In terms of groundnut cultivation in Karnataka, it is one of the largest producers of groundnut” [5]. “The North-Eastern dry zone of Karnataka, comprising Raichur, Gulbarga and Koppal, is a predominant groundnut-producing tract with an area of about 1.50 lakh ha, a production of 1.14 lakh tones, and a productivity of 742 kg/ha. Groundnut is grown throughout the year due to a two-crop cycle harvested in March and October” [6,7].

Crop patterns in the Vijayapura district the local crop patterns are influenced not only by agricultural climate conditions such as rainfall, soil and temperature, but also by government programs for crop production such as foundations and the speed of improvement of infrastructure. The crops of the district of Vijayapura are divided into the following categories.

- Food crops: Jowar, corn, red gram, and millets.
- Commercial crops: sugarcane, cotton, tobacco.
- Oil seeds: nuts, seeds, sunflowers, etc.
- Plantations crops: grapes, lemons and bananas.

Groundnuts offer a diverse array of nutrients and bioactive compounds that contribute to overall health. Incorporating them into the diet can be beneficial. With respect to nutritional composition groundnuts are a rich source of essential nutrients. They contain high-quality proteins, dietary fiber, vitamins, and minerals. Proteins in groundnut provide essential amino acids, making them valuable for overall health and the dietary fiber in groundnut is associated with reducing obesity risk and cardiovascular diseases [8].

“Agricultural production has seen significant changes over the past few decades. The amount of food we grow has increased rapidly due to two main drivers: the expansion of land used for agriculture and a rapid rise in crop yields. The past two decades have seen a steady upward trend in world agricultural production to meet expanding demand, with primary crops production growing by 54 percent between 2000 and 2021, meat production growing by 53 percent and milk production by 58 percent” [9]. “In India, agricultural production trends have been influenced by various factors. Prior to independence, agricultural production declined. However, post-independence, there was a steady rise in average yield per hectare, a steady rise in area under cultivation and due to these factors, total production of crops recorded a rising trend" [10,11]. "Crop yield forecasting is a critical predictive analytics technique in the agriculture industry. It involves accurately predicting the potential yield of a specific crop during a particular season in each region. Accurate yield predictability requires scientific expertise, local knowledge about the region and crops grown there" [12]. "Several factors, including weather vagaries, soil nutrition levels, fertilizer availability and cost, pest control and agrometeorological input parameters like temperature and rainfall, influence crop yield. Therefore, forecasting crop yield is a challenging task. Statistical models are the most used tools to forecast the crop yield. These models should be able to take advantage not only of historical data of crop yield but also the impact of various driving forces of the external environment" [13, 14].

2. MATERIALS AND METHODS

This section will provide a brief description of material and methods used in addition to the important statistical tools and techniques used in the analysis, to deal with the necessary database for the study. Method and analytical techniques incorporated in this study are listed as follows.

2.1 Description of The Study Area

Vijayapura, also known as Bijapur, is a district in the state of Karnataka (Fig.1), India. The city of Vijayapura is the district headquarters and is located 530 km northwest of Bangalore. The district is well known for its historical monuments built during the Adil Shahi dynasty [15]. In terms of agriculture, the Department of Agriculture in Vijayapura provides Agricultural Extension services to farmers and transfers the latest technical knowledge to the farming community. This includes the introduction of high-yielding varieties, laying demonstrations and imparting training to farmers to improve skills & knowledge to boost up agricultural production and productivity. The major crops grown in Vijayapura include cereals like Rice, Ragi, Jowar, Maize, minor millet and pulses like Red Gram, Horse Gram, Green Gram, Black gram and Bengal gram. The district is also known for the cultivation of vegetables as well as cash crops like sugarcane and cotton [16].

2.2 Nature and Data Sources

Secondary data on the area, production and productivity of selected crops have been collected from the Directorate of Economics and Statistics (GOK), the ICRISAT-Data Directorate and the District Administration statistical report, Vijayapura, Karnataka, for the 55-year period 1966-67 to 2020-21.

2.3 Analytical Tools and Techniques

In line with the objectives of the study, the following statistical tools and techniques have been used.

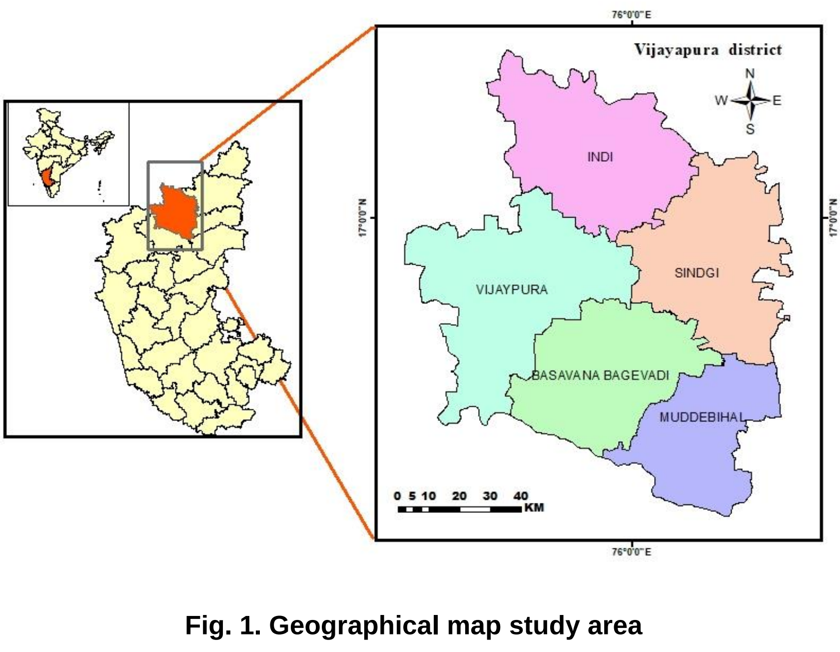

Fig. 1. Geographical map study area

2.3.1 Trend analysis

Trend refers to an increase or decline in data over a long period of time. It measures long-term changes in a time series without causing concern about short-term fluctuations between them. To estimate the long-term trend of agricultural area, production and productivity of large crops, the methodology of least square estimation was used [17]. In this method, the trend of area, production and productivity was measured by determining the mathematical relationship between time and response variables. The mathematical expression can be expressed as follows:

Linear model $$Y_t = \alpha + \beta t + \varepsilon$$

A linear model is one in which all the parameters are linear. The average trajectory for the data is a straight line corresponding to increasing or decreasing constant rate of change in time [14].

Quadratic model $$Y_t = \alpha + \beta t + kt^2 + + \varepsilon$$

A quadratic function is one having a peak or a trough in the data. (i.e., parabola). The average trajectory for the data contains a curve with variable degrees of steepness corresponding to rate of acceleration or deceleration.

Cubic model $$Y_t = \alpha + \beta t + kt^2 + \gamma t^3 + \varepsilon$$

The cubic is one having two troughs in the data. The average trajectory for the data behaves quadratically until a further curve occurs, which can correspond to an acceleration or deceleration with varied degree of steepness [18].

Where,

α: Intercept or Average effect; β, k, γ: Slope or Regression Coefficients (β: linear effect parameter, k: quadratic effect parameter and γ: cubic effect parameter); Yt: Area, production or productivity in time period t; ε: Error term or disturbance term.

2.3.1.1 Exponential model

If the values of t are arranged in an arithmetic series, the corresponding values of y form a geometric series, the relation is of the exponential type. The function of this type can be given as

$$Y_t = \alpha e^{ct} + \varepsilon$$

where, Y_t represents area, production or productivity in time period t , α and c are parameters, e is the exponential term and ε denotes the error term and α represents the value at t = 0, c represents the exponential rate

Log-logistic $$Y_t = \frac{\alpha}{1+\exp[-\beta\{|\log(t)-\log(\gamma)|\}]} + \varepsilon$$

Where, Y_t represents area, production, or productivity in time period t, α, β and γ are parameters and ε denotes the error term. The parameter ‘β’ is the ‘intrinsic growth rate’, while the parameter ‘α’ represents the ‘upper asymptote’ and ‘γ’ is the growth range.

2.3.1.2 Generalized additive model

The GAM is a non-linear combination of independent variables. The dependent variable can be represented by univariate smooth function of time. GAM is an additive modeling technique where the influence of the predictive variables is captured through smooth functions which depends on the underlying patterns in the data.

$$y_t = f(t) + \varepsilon$$

where, y_t represents area, production, or productivity in time t. $f(t)=\alpha_0+b_t\ (t)_\alpha$, represents the smooth function of the time t, where α is some parameter and b_t (t) is the basis function. The main assumptions of ‘independence of residuals’ and ‘normality of residuals’ was examined by using respectively the ‘Run-test’ and ‘Shapiro-Wilk test’.

2.3.2 Test for independence (randomness) of residuals by run test

“Non-parametric Run test was used to test the randomness of residuals. A Run is defined as ‘a succession of identical symbols which are followed and preceded by different symbols or no symbols at all’. If very few runs occur, a time trend or some bunching owing to lack of independence is suggested and if many runs occur, systematic short period cyclical fluctuations seem to be influencing the scores. The significance of any observed value of ‘Z’ computed using the equation may be determined from a normal distribution table”. [19].

2.3.3 Test for normality of residuals by Shapiro-Wilk's (W) test

"This is the standard test for normality. The test statistic W is the ratio of the best estimator of the variance (based on the square of a linear combination of the order statistics) to the usual corrected sum of squares estimation of the variance. W may be thought of as the correlation between given data and their corresponding normal scores. The values of W ranges from 0 to 1. When W=1 the given data are perfectly normal in distribution. When W is significantly smaller than 1, the assumption of normality is not met. A significant W statistic results to rejection of the assumption that the distribution is normal". [19] Shapiro-Wilk's W is more appropriate for small samples up to n=50

H_0: Samples x_1, . . ,x_n is from a normally distributed population.

H_1: Samples x_1, . . .,x_n is not from a normally distributed population.
Test statistic is given by:

$$W = \frac{\left[\sum_{i-1}^{n} a_i x_{(i)}\right]^2}{\sum_{i-1}^{n} (x - \bar{x})^2}$$

where, $x_{((i)\ )}$is the i^{th} order statistic, i.e.,, the i^{th} smallest number in the sample; $\bar{x}$ is sample mean and the constants a_i is given by

$$(a_1, a_2, \cdots, a_n) = \frac{m^T V^{-1}}{\sqrt{(m^T V^{-1} V^{-1} m)}}$$

Where $m^T = (m_1, m_2, \cdots, m_n)^T$ and $m_1, m_2, \cdots, m_n$ are the expected values of the order - statistics of independent and identically distributed random variables sampled from the standard normal distribution and V is the covariance matrix of those order statistics. Reject the null hypothesis if W is too small (near to zero) [17].

2.3.4 Model adequacy checking

A. Coefficient of determination (R^2)

The coefficient of determination (R^2) is a test statistic that will give information about the appropriateness of a model. R^2 value is the proportion of variability in a data set that is accounted for the statistical model. It provides a measure of how well the assumed model explains the variability in dependent variable.

$$R^2 = \frac{RSS}{TSS} = 1 - \frac{ESS}{TSS}$$

Where, ESS is error sum of squares, RSS is regression sum of squares; TSS is total sum of squares.

Computed R^2 value lies between zero and one. If R^2 value is closer to 1 indicates that the model fits the data. Adjusted R^2 and Root Mean Square Error (RMSE) are also used for the checking of the fit of model.

B. Adjusted R^2

The adjusted R-squared is a modified version of R-squared that has been adjusted for the number of predictors in the model. The adjusted R-squared increases only if the new term improves the model more than would be expected by chance. It decreases when a predictor improves the model by less than expected by chance. It is always lower than the R-squared.

$$Adjusted\ R^2 = \frac{RSS/_{df}}{TSS/_{df}}$$

where, RSS is regression sum of squares; TSS is total sum of squares; df is the respective degrees of freedom.

C. Root Mean Square Error (RMSE)

The Root Mean Square Error (RMSE) (also called the root mean square deviation, RMSD) is used to assess the amount of variation that the model is unable to capture in the data. The RMSE is obtained as the square root of the mean squared error hence considered as the model prediction capability and is obtained as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_t - \hat{Y_t})^2}{n}}$$

where, Y_t= observed value; $\hat{(Y_t\ )}$= predicted value; n= number of observations

D. Akaike Information criterion

The Akaike Information criterion (AIC) is a mathematical method for evaluating how well a model fits the data. AIC is calculated from the number of independent variables used to build the model and the maximum likelihood estimates of the model. The best fit model based on AIC is the one that explains the maximum amount of variation using the fewest possible independent variables. AIC is most often used for model selection.

The AIC is calculated using the formula,

$$AIC=2K-2\ln(L)$$

where, K = Number of independent variables; L = Log-likelihood estimate.

E. Bayesian Information Criterion

The Bayesian Information Criterion (BIC) is a method for scoring and selecting a model. BIC is a criterion for model selection among a finite set of models. It is closely related to AIC. It is named after the field of study from which it was derived i.e., Bayesian probability and inference. Like AIC, it is appropriate for models fit under the maximum likelihood estimation.

The formula for BIC is.

$$BIC=K\ln(n)-2\ln(L(\theta))$$

Where, n = sample size, K = Number of independent variables, θ = set of all the parameters, L(θ) – Loglikelihood estimate.

2.3.5 Forecasting using Auto Regressive Integrated Moving Average (ARIMA) methods

The Box Jenkins procedure involves the integration of the automatic regression Integrated Moving Average (ARIMA) model into a specific data set. The main objective of integrating this ARIMA model is to identify the stochastic process of time series and accurately predict future values. The ARIMA model with the parameter (p,d,q) was adapted with the Box-Jenkins technique (1976). This model includes the autoregressive order p, the differentiator to make the stationary series of degree d and the moving average of the order q [20].

A process $\{Y_t\}$ is said to follow an Integrated ARMA model, denoted by ARIMA (p, d, q). Then the ARIMA model can be written as

$$\varphi(B)\ 〖(1-B)d\ \ Y_t=\theta(B)\varepsilon_t$$

Where, $\varepsilon t \sim WN_{(0,\ \sigma\text{^}2)}$, WN indicating White Noise. The integration parameter d is a nonnegative integer. When d = 0,

$$ARIMA\ (p, d, q) \equiv ARMA\ (p, q)$$

3. RESULTS AND DISCUSSION

3.1 To Identify the Existing Patterns in Area and Production of Groundnut Crops

The data of each time series reflect the appropriate trend towards growth or decline. "Tends" refer to this tendency to increase or decrease. To estimate trends in area, production, and productivity of some major crops in the Vijayapura district of Karnataka. Data are collected for 24 years, from 1997 to 1998 and from 2020 to 2021. The trends were estimated using various models such as linear, quadratic, cubic, exponential, log-logistic and GAM models. The student t test was used to evaluate the statistical significance of parameters in linear models, square models, cubic models, exponential models, log logistics and GAM. For the other models, 95 per cent of the estimated parameter's asymptotic confidence range was calculated. If the estimated parameter of the adapted model is within a 95 percent confidence interval, this means that the parameter value is significant at a 5 percent level of significance. The basic assumptions of the randomness and normality of each model residual were examined using the Runs test and Shapiro test, respectively. Only models that meet the assumptions of residual independence and residual normality and in which all parameters are determined to be significant are considered suitable models.

Parameter estimates by six models for the annual area under Groundnut in Vijayapura district is given in Table 1, along with respective standard errors (in parenthesis). The test statistic, probability values, and model adequacy standards are presented in Table 2.

According to the results, the estimated parameters of the linear, log-logistic and GAM models were significant at 5 per cent level of significance. The results of Table 2. also revealed that all fitted models were found to be non-significant by the Run's test and Shapiro-Wilk's test statistic. This indicates, the data during the study period for the area under groundnut in the Vijayapura district was well fitted to the linear, log-logistic and GAM models.

The model with the lowest MAPE value was identified as the most appropriate model. Log-logistic model fitted for the data was found to be

Table 1. Parametric estimates by different models for the annual area under Groundnut in Vijayapura district

Parameters	Models					
	Linear	**Quadratic**	**Cubic**	**Log- Logistic**	**Exponential**	**GAM**
Intercept (α_0)	7,015.4*(5,019.3)	7,926.9* (7,242.4)	57,171.3(9,759.5)	4.58*(1.4)	9080.0 (5090.0)	16,524*(13,302)
α	-1,335*(366.1)	1,806.2 (1,390.3)	3,148.3(3,800.1)	63,953.8**(3,127.5)	14.3**(0.03)	20,664** (2,866)
β		-130.8 (56.2)	-365.7 (363.8)	22.4** (1.7)		
γ			10.0 (9.9)			

*** and * indicates significant at one and five per cent, respectively. Figures in parentheses indicate standard errors of respective coefficients*

Table 2. Test for randomness, normality of residuals and goodness of fit criteria of different models for the annual area under Groundnut in Vijayapura district

Criteria	Models					
	Linear	**Quadratic**	**Cubic**	**Log Logistic**	**Exponential**	**GAM**
Runs test(*p*-value)	-2.62 NS[0.08]	-3.05 NS[0.06]	-3.05*[0.03]	-3.05 NS[0.82]	0 NS[1.00]	-2.62 NS[0.08]
Shapiro-Wilk(*p*-value)	0.91 NS[0.48]	0.95 NS[0.29]	0.95 NS[0.29]	0.95 NS[0.37]	0.63 NS[2.32]	0.91NS[0.04]
RMSE	11,127.60	9,871.29	9,868.84	9,650.22	43,757.76	11,127.61
MAPE	17.36	16.05	25.01	15.58	34.87	17.36
AIC	499.86	496.35	498.33	495.30	562.84	499.86
BIC	503.26	500.89	504.01	499.85	566.25	503.26
R^2	0.38	0.90	0.51	0.94	0.56	0.38
Adj R^2	0.36	0.88	0.44	0.93	0.53	0.35

*NS: Non-Significant, *Significant at 5 per cent; Values in parenthesis indicate Probability value*

Table 3. Parametric estimates by different models for the annual production under Groundnut in Vijayapura district

Parameters	Models					
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log Logistic**	**GAM**
Intercept (α_0)	30,639.0*(2,428.7)	24,879.8*(3,526.4)	5,528.3*(5,301.4)	30600.0**[2460.0]	-1.0(1.0)	29,855.1*(847.5)
α	-62.7 (60.6)	1,266.3*(661.34)	983.5*(791.3)	-0.02 [0.01]	9,147.0*(6,430.9)	-21.1* (16.2)
β		-53.2 (25.2)	-25.4 (15.73)		4.7(4.33)	
γ			-4.7 (4.33)			

*** and * indicates significant at one and five per cent, respectively; Figures in parentheses indicate standard errors of respective coefficients.*

best, as shown by the results lowest MAPE (15.58) and the other criteria's, including highest R2 (0.94) and Adjusted R2 (0.93), lowest AIC (495.30), BIC (499.85), and RMSE (9,650.22). This indicates, Fig. 2. from 1997 to 2020, the data for the Vijayapura groundnut area has a log-logistic trend.

Parameter estimates and goodness of fit criteria of selected six models for production of Groundnut in Vijayapura district is given in Table 3 The test statistic, probability values and model adequacy standards are presented in Table 4. According to the results, the estimated parameters of linear, log-logistic and GAM model was significant at 5 per cent level of significance. The results of Table 4 also revealed that all fitted models were found to be non-significant by the Run's test and Shapiro- Wilk's test statistic. This indicates, the data during the study period for the production under groundnut in the Vijayapura district was well fitted to the linear and GAM model.

The best model was also chosen based on the lowest MAPE value. The GAM model fitted for the data was found to be best, as shown by the results lowest MAPE (9.20). And the other criteria's, including highest R2 (0.78) and Adjusted R2 (0.76), lowest AIC (475.38), BIC (484.12), and RMSE (3,553.32). GAM was found to be the best-fit model. This indicates, Fig. 3 from 1997 to 2020, the data for the Vijayapura groundnut production has a GAM trend.

Parameter estimates and goodness of fit criteria of different models for the productivity under Groundnut in Vijayapura district is given in Table 5. The test statistic, probability values, and model adequacy standards are presented in Table 6 According to the results, the estimated parameters of linear, cubic, exponential and GAM models were significant at 5 per cent level of significance. The results of Table 6 also revealed that all fitted models were found to be non-significant by the Run's test and Shapiro-Wilk's test statistic. This indicates, the data during the study period for the productivity under groundnut in the Vijayapura district was well fitted to linear, exponential and GAM models.

The most appropriate model was selected using the lowest MAPE value. The cubic model fitted for the data was found to be best, as shown by the results lowest MAPE (9.18). And the other criteria's, including highest R2 (0.68) and Adjusted R2 (0.63), lowest AIC (46.13), BIC (40.45) and RMSE (0.0714). This indicates, Fig. 4 from 1997 to 2020, the data for the Vijayapura groundnut productivity has a cubic trend.

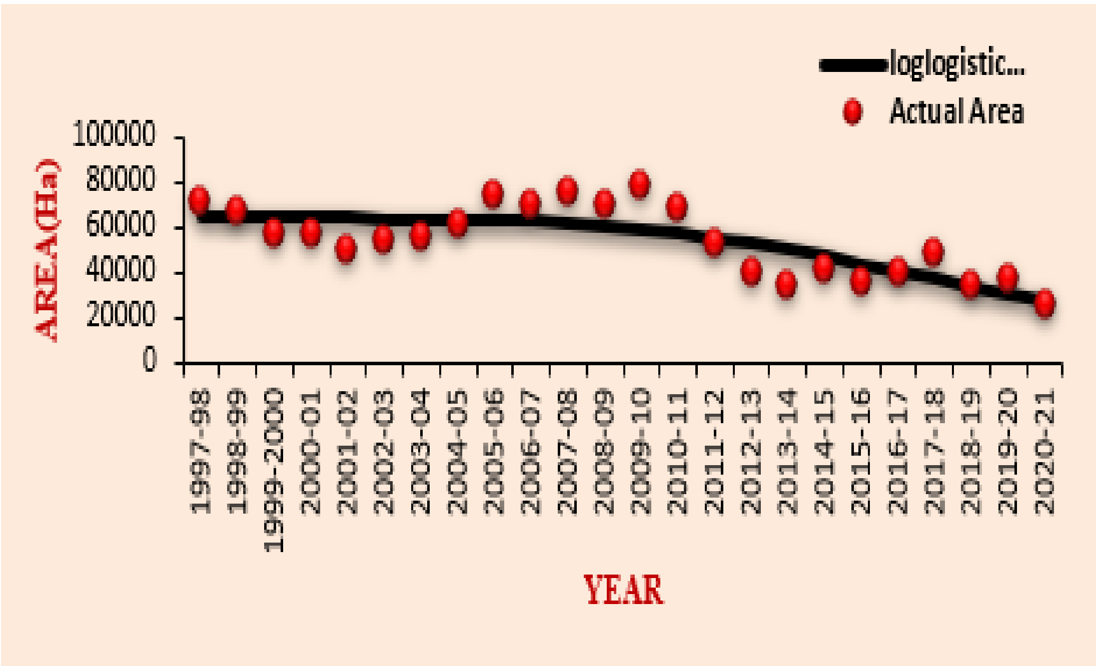

Fig. 2. Best fit model for area under Groundnut

Table 4. Test for randomness, normality of residuals and goodness of fit criteria of different models for the annual production under Groundnut in Vijayapura district

Criteria	Models					
	Linear	Quadratic	Cubic	Exponential	Log Logistic	GAM
Runs test(p-value)	-2.08^{NS} [0.08]	0 ^{NS} [1.00]	0 ^{NS}[1.00]	-2.82 ^{NS}[0.18]	-1.38*[0.01]	-0.41^{NS}[0.67]
Shapiro-Wilk(p-value)	0.89$^{N\ S}$[0.06]	0.93^{NS} [0.11]	0.93^{NS}[0.13]	0.89^{NS}[0.32]	0.96 ^{NS}[0.62]	0.98^{NS} [0.94]
RMSE	5,520.16	5,030.71	5,027.15	5,521.25	125,155.70	3,553.32
MAPE	15.42	13.98	14.00	15.43	88.12	9.20
AIC	487.68	485.22	487.19	487.69	639.50	475.38
BIC	491.22	489.94	493.08	491.22	644.21	484.12
R^2	0.09	0.67	0.70	0.09	0.13	0.78
Adj R^2	0.04	0.66	0.69	0.07	0.11	0.76

*NS: Non-Significant, *Significant at 5 per cent; Values in parenthesis indicate Probability value*

Table 5. Parametric estimates by different models for the annual productivity under Groundnut in Vijayapura district

Parameters	Models					
	Linear	Quadratic	Cubic	Exponential	Log-Logistic	GAM
Intercept (α_0)	0.38** (0.34)	0.43**(0.33)	0.49(0.08)	0.40**[0.27]	-0.73**(0.68)	0.38** (0.20)
α	0.15 **(0.42)	0.26(0.21)	-0.03*(0.02)	0.77**[0.40]	2.32(1.15)	0.54** (0.20)
β		0.05* (0.04)	0.03 (0.02)		413.61 (325.98)	
γ			-6.87e-05*(6.22e-05)			

*** and * indicates significant at one and five per cent, respectively; Figures in parentheses indicate standard errors of respective coefficients*

Table 6. Test for randomness, normality of residuals and goodness of fit criteria of different models for productivity under Groundnut in Vijayapura district

Criteria	Models					
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log Logistic**	**GAM**
Runs test(*p*-value)	-0.87^{NS} [0.38]	0 ^{NS}[1.00]	0.43^{NS}[0.66]	0 ^{NS}[1.00]	-1.74*[0.01]	-0.87^{NS}[0.38]
Shapiro-Wilk(*p*-value)	0.98^{NS} [0.96]	0.96^{NS} [0.57]	0.94^{NS}[0.24]	0.97*[0.01]	0.95^{NS}[0.46]	0.98^{NS}[0.96]
RMSE	0.07	0.06	0.07	0.07	0.08	0.09
MAPE	10.22	9.38	9.18	9.82	13.12	10.22
AIC	47.25	47.06	46.13	48.33	37.72	47.25
BIC	43.84	42.52	40.45	44.92	33.18	43.84
R^2	0.64	0.67	0.68	0.66	0.50	0.65
Adj R^2	0.62	0.63	0.63	0.64	0.48	0.62

*NS: Non-Significant, *Significant at 5 per cent; Values in parenthesis indicate Probability value*

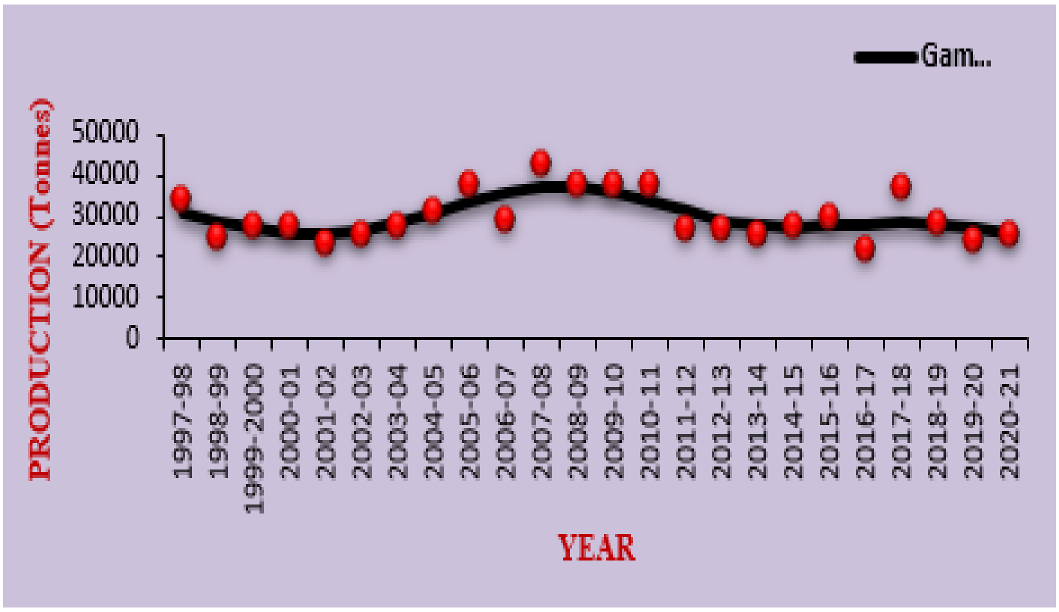

Fig. 3. Best fit model for production under Groundnut

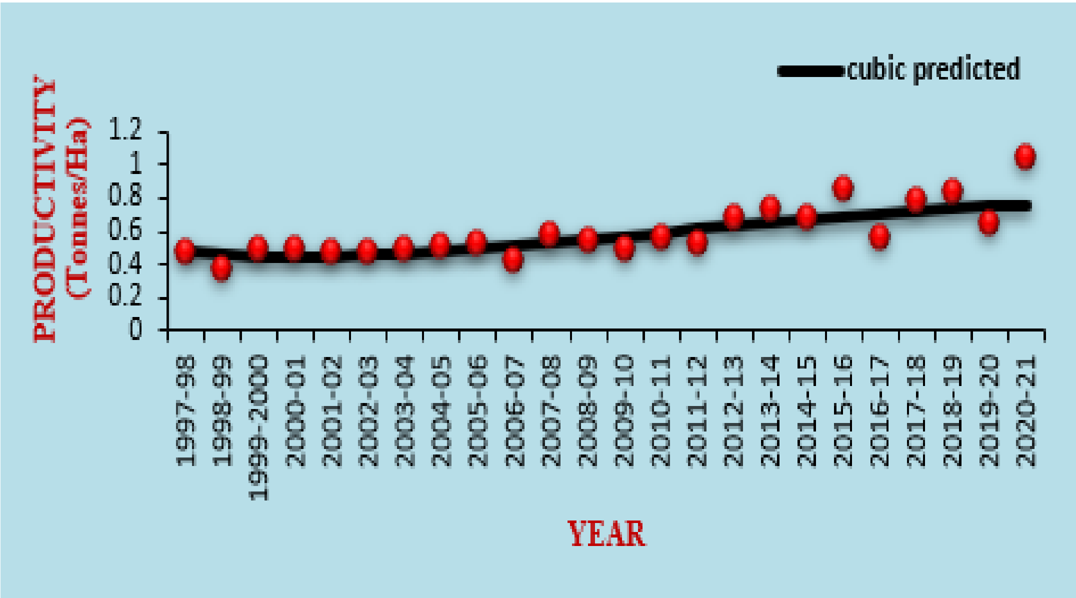

Fig. 4. Best fit model for productivity under Groundnut

3.2 To Develop ARIMA Model for Forecasting Production of Groundnut Crop

The selected model is finally used for forecasting yield of groundnut crop of Vijayapura district of Karnataka, Box-Jenkin's or ARIMA method was used based on the data considered for 55 years from 1966-67 to 2020-21. The tables and figures obtained from the analysis are presented at the end. The results of forecast yield of groundnut are discussed below. The data used of groundnut production for 55 years from 1966-67 to 2020-21. Throughout the study, 45 years of training data and 10 years of test data were considered. An examination of plot reveals positive trend over time indicating non-stationary nature of the series. From Table 7, it is observed that the stationarity in yield of groundnut can be obtained after first differencing of the original data as per the augmented Dickey-Fuller test (Fig. 6). The test statistic was found to be significant at 5 per cent level.

Table 7. Augmented Dickey-Fuller statistic for production of groundnut

District	cotton	Augmented Dickey-Fuller Test		
		Differencing	Statistic	p-value
Vijayapura	Production	1	-4.31	<0.001

3.2.1 Model identification

Based on autocorrelation function and partial autocorrelation function of the differenced series, models were fitted as shown in Fig.6. and Fig.7. two lagged value was outside the limit in PACF and one lagged value was outside the limit in ACF plot, Hence the suitable model was ARIMA (2,1,1). The MAPE and AIC values of the model were given in Table 8. and non-significance of L-Jung box test, which indicates the absence of auto correlation in the model for future predictions.

3.2.2 Model accuracy

Residual analysis was carried out to check the adequacy of the models. The residuals of ACF and PACF were obtained from the identified ARIMA model. All the lags were found to be non-significant, and the model fitted over test dataset gives a MAPE of 24.36 as indicated in Table 9. it also gives actual and forecasted values for test dataset gives satisfactory results to conclude ARIMA (2,1,1) as the suitable model (Depicted in Fig. 8).

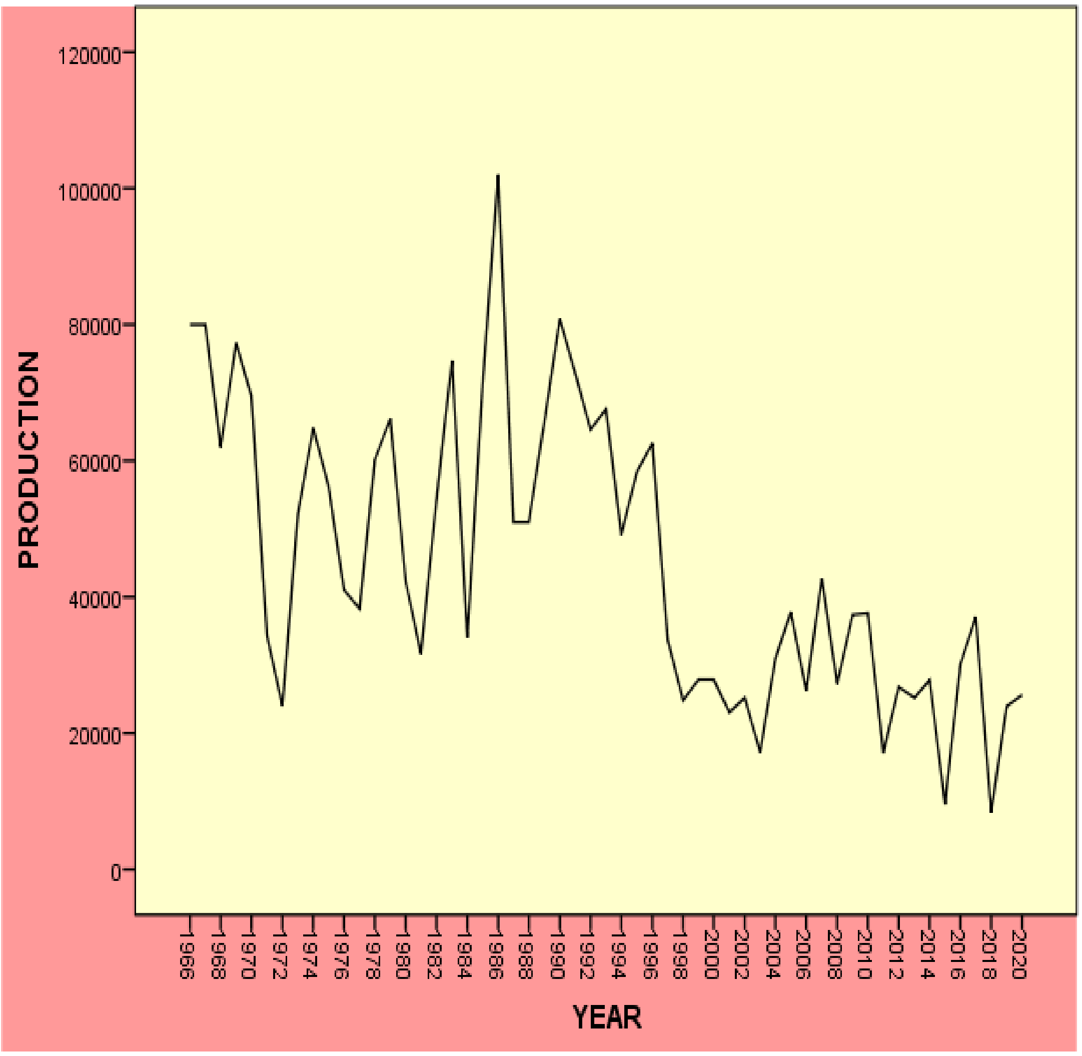

Fig. 5. Time plot for yield of Groundnut

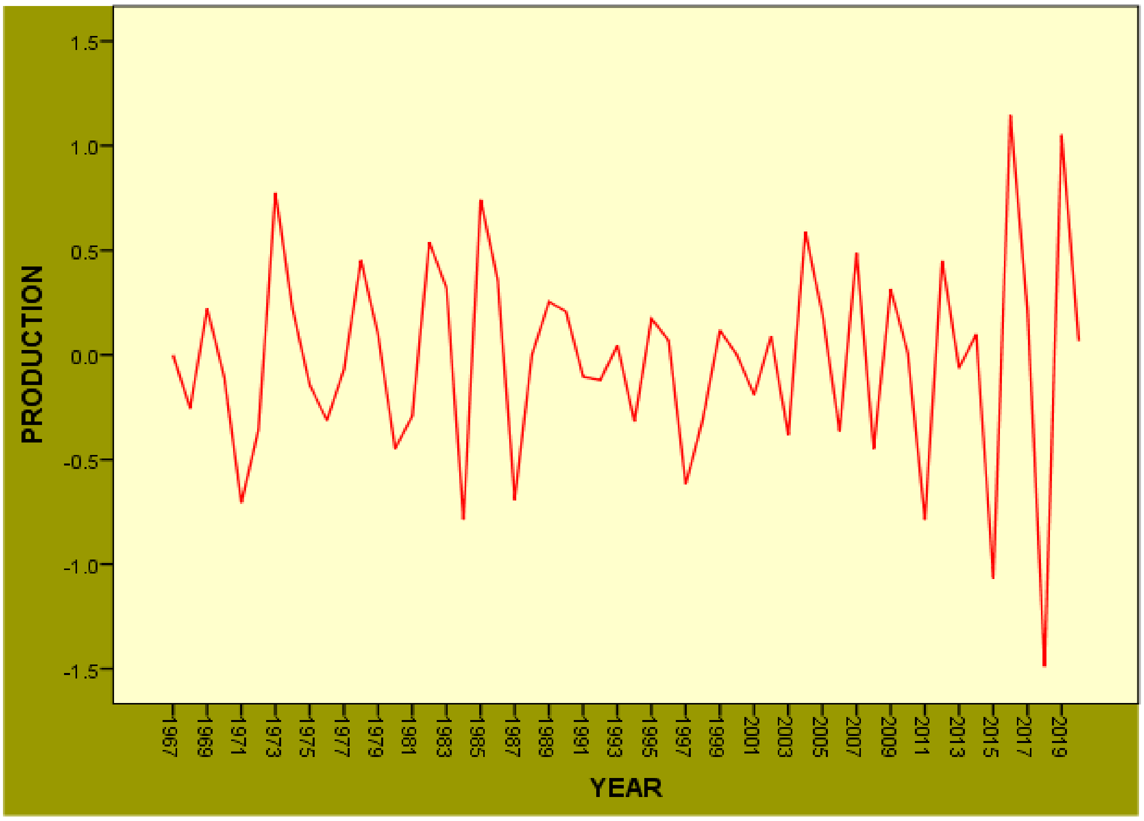

Fig. 6. Time plot for the differenced series

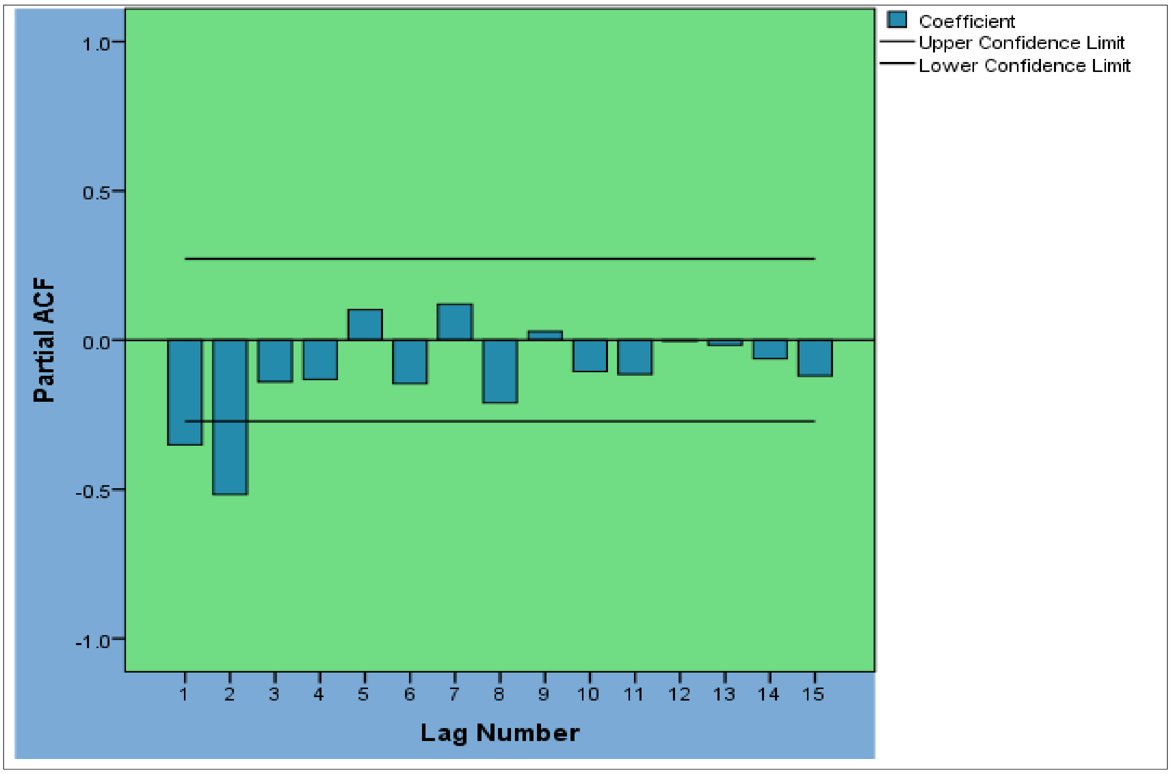

Fig. 7. Partial Autocorrelation at different lags of first differenced series for yield

Table 8. MAPE, AIC and L-Jung-Box Q for ARIMA model for groundnut production

Model	Model statistics		L-Jung-Box Q		
	MAPE	AIC	Statistic	DF	p-value
ARIMA (2,1,1)	25.32	974.74	14.28	15	0.50

Model	Ar1	Ar2	Ma1
ARIMA (2,1,1)	-0.19	-0.58**	-0.07

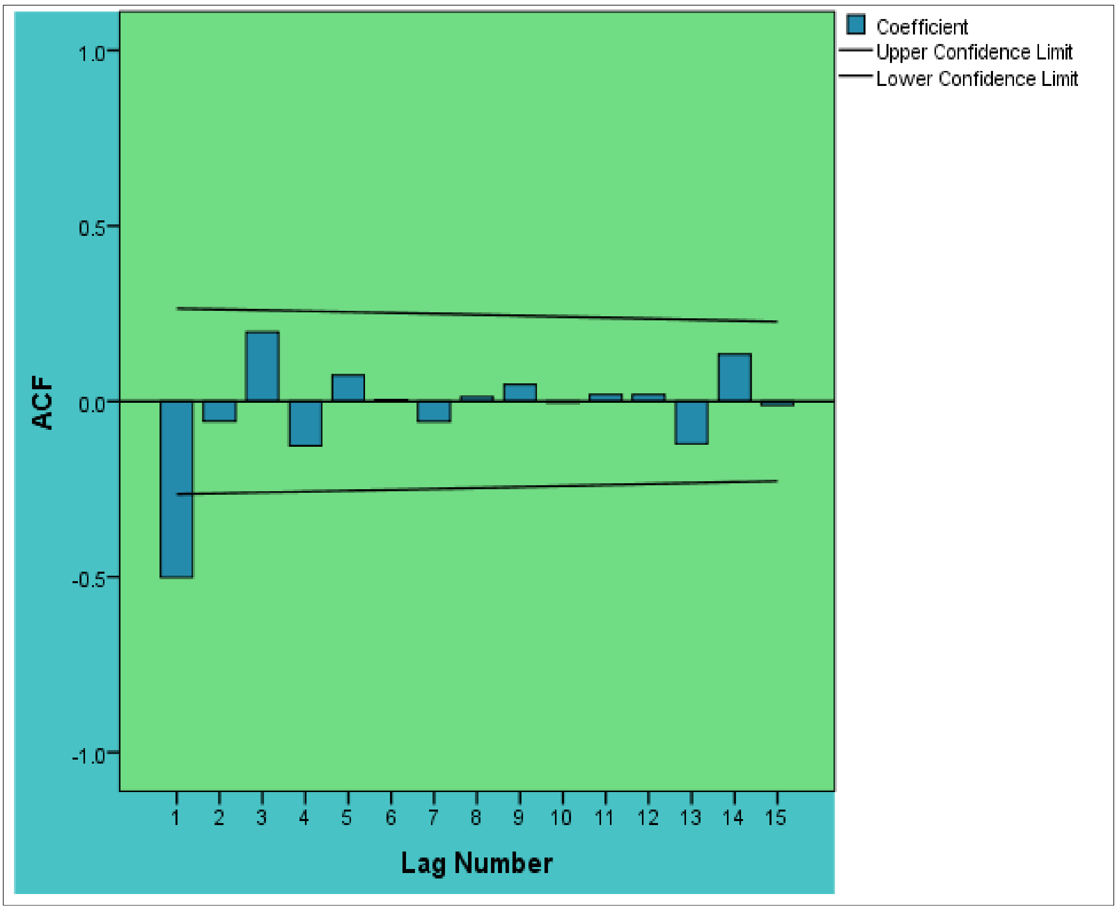

Fig. 8. Autocorrelation at different lags of first differenced series for yield of Groundnut

Table 9. Model fitted over test dataset with MAPE for production of groundnut

Year	Production (Tones) forecast using ARIMA (0,1,0)				MAPE
	Actual	Forecast	Lower Confidence limit	Upper Confidence limit	
2011 -12	17106	37684.6	8975.6	66393.6	24.36
2012 -13	26816	37540.8	1906.4	73175.2	
2013 -14	25194	37528.5	1409.3	73647.8	
2014 -15	27801	37614.4	850.2	76079.0	
2015 -16	9568	37605.0	6053.5	81263.5	
2016 -17	30113	37557.0	8598.2	83712.2	
2017 -18	37112	37571.6	10056.1	85199.3	
2018 -19	8372	37596.7	12541.2	87734.6	
2019 -20	23962	37583.4	15290.1	90456.9	
2020 -21	25629	37571.4	17201.6	92344.4	

3.3 Forecasting of the Production

The best fitted ARIMA (2,1,1) was used to forecast groundnut production for the period from 2021-2022 to 2025-26 for a five-year period presented in Table 10 and visualized in Fig. 10. The results (depicted in Fig. 9) indicated an increasing trend in yield of groundnut. The conclusion from the study is that groundnut production can be increased in future.

Using three linear models and two non-linear models and GAM were fitted to estimate the forecast value of production of groundnut crops in Vijayapura district of Karnataka. For forecasting, the best-fitted model was used based on the minimum value of MAPE among all the fitted models based on the data took of 24 years from 1997-2020 which are also used for trend analysis for predicting value in next 5 years from 2021-22 to 2025-26. Considering data from the previous 24 years (1997 to 2020), future yearly forecasts of groundnut production were generated. Using the MAPE criterion, it was concluded that the GAM provided the best fitting among the several models fitted. Using GAM, a forecast for the following five years (2021 to 2025) was generated. In the Table 11 and in Fig. 11, the forecasted production of groundnut was shown. The results in Table 11 indicated slight decreasing trend in groundnut output.

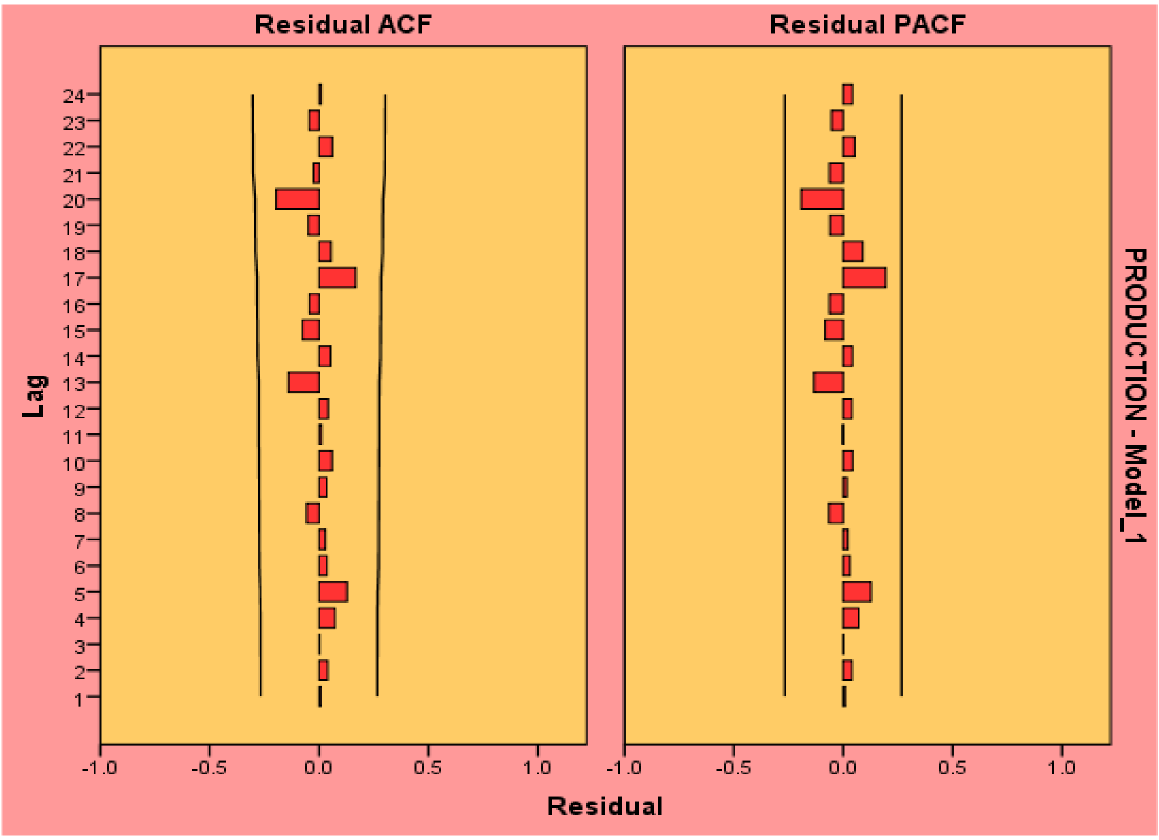

Fig. 9. Residual autocorrelation and partial autocorrelation for Groundnut yield

Table 10. Forecast of groundnut production for the period for a five-years

Year	Production (Tons) forecast by ARIMA (0,1,0)		
	Forecast	Lower Confidence limit	Upper Confidence limit
2021 -22	37581.4	18983.6	94146.5
2022 -23	37586.5	21091.0	96264.0
2023 -24	37579.7	23090.3	98249.7
2024 -25	37578.0	24823.4	99979.5
2025 -26	37582.3	26564.4	101729.1

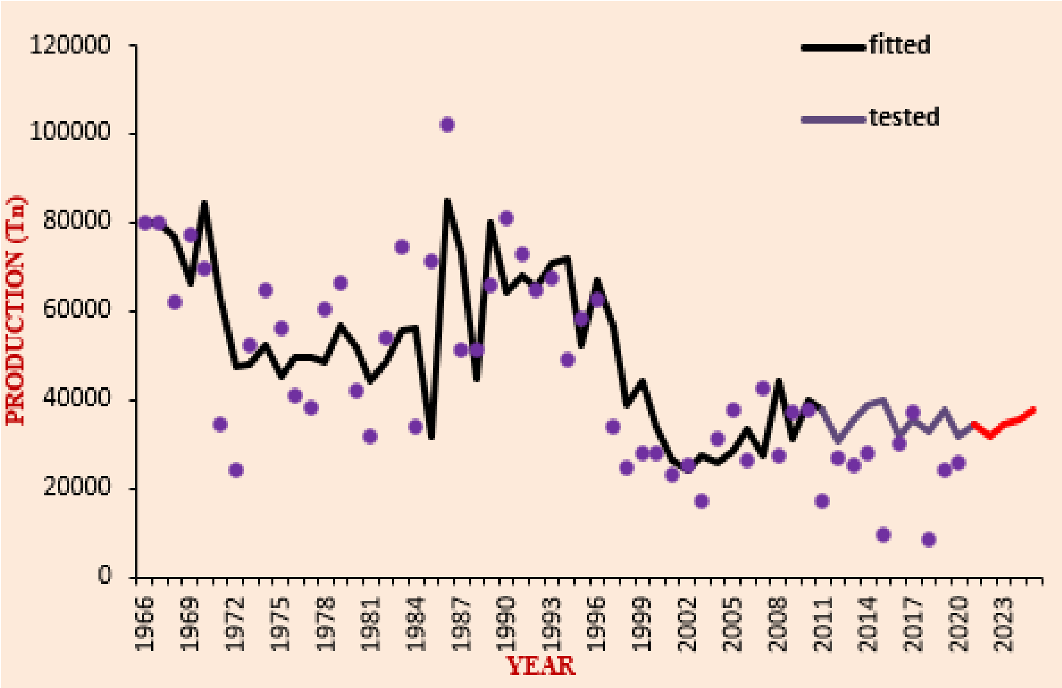

Fig. 10. Predicted values by ARIMA and actual Yield of Groundnut in Vijayapura district

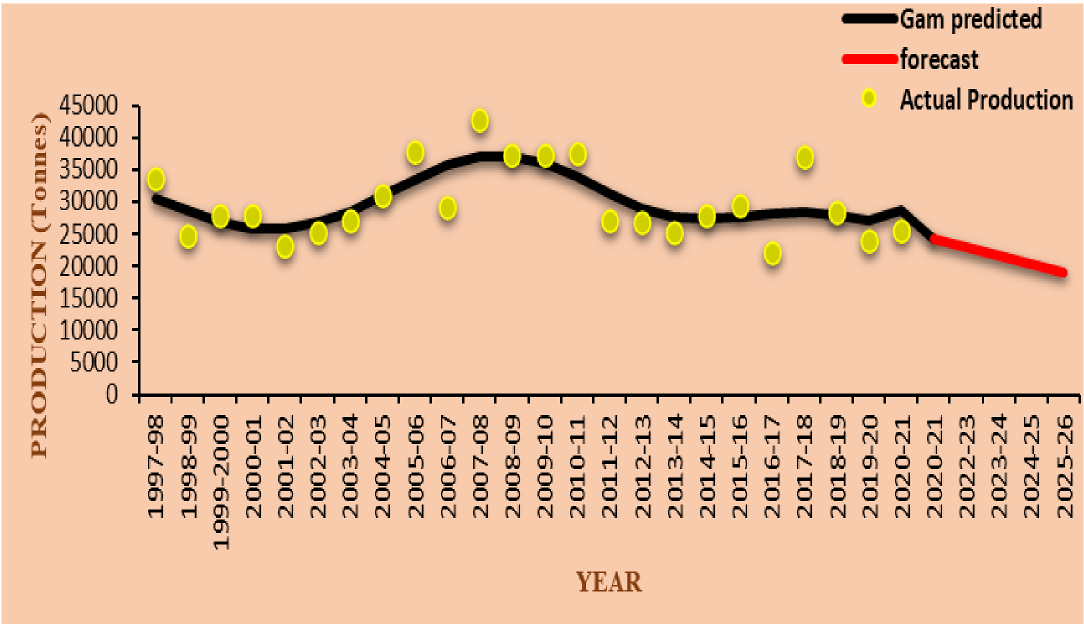

Fig. 11. GAM Forecasted production of Groundnut in Vijayapura district

Table 11. Forecasted production of groundnut in Vijayapura district

Year	Forecasted production (tons)
2021-22	24305.22
2022-23	22966.44
2023-24	21627.66
2024-25	20288.88
2025-26	18950.10

4. CONCLUSION

Agriculture has offered a living for a large percentage of people in the state for the past few decades. Agriculture plays a significant role in the economy of Karnataka and Vijayapura district. Approximately 76 per cent of the working population is employed in agriculture and farming. Agriculture accounts for approximately half of the district's revenue.

The log-logistic model that was found to be the best-fitted model for the area had the lowest MAPE value of 15.58. The cubic model with a minimum MAPE value of 9.18 is shown to be the best-fitted model for ground nut productivity, followed by the GAM model with a minimum MAPE value of 9.20. All the parameters in the selected models were significant and met the residuals' assumptions. According to the findings, the area experienced a minor growth pattern from 1997 to 2007, followed by a steady upward trend. Initially, the production pattern showed constant increase, but beginning in 2007, there was a major fall that lasted until 2019-20. Groundnut productivity first declined somewhat before increasing steadily from 2009 to 2020. Despite a fall in production as compared to prior years' trends, the ARIMA results indicated a little increase in anticipated values. As a result, ARIMA ignored recent trends, causing them to stand out as outliers. Because the ARIMA findings were inadequate, forecasting was performed using GAM, which was determined to be the best model based on the minimum MAPE value among all fitted models. The generalized additive model predicts that groundnut output will decrease somewhat, with production reaching 24,305.22 and 18,950.1 tons in 2021-22 and 2025-26, respectively.

To analyze the trend in area, production and productivity of Groundnut in Vijayapura district of Karnataka. Using three Linear, two non-linear and GAM models viz. linear, quadratic, cubic, exponential, Log-logistic and GAM models were fitted for 24 years. Groundnut crop area, production and productivity was fitted with Log-logistic, GAM and cubic models found to be the best fit.

The forecasted production of Groundnut crops for the period of 5 years by using ARIMA models and using time series data with the help of GAM. According to recent few years, sudden increase or decrease in production was observed. Due to this, ARIMA failed to take recent trends into account and therefore it was treated as outliers. Forecasting was performed since the results of ARIMA were inadequate, GAM revealed slight downward trend for a groundnut crop.

ACKNOWLEDGEMENTS

I extend my heartfelt gratitude to Harshith K V for his invaluable assistance in organizing and structuring the research paper. His meticulous efforts have significantly enhanced the clarity and coherence of our work. I would also like to express my sincere appreciation to Ragini H R for her diligent research analysis and unwavering support throughout this endeavor. Her insights and dedication have been instrumental in shaping the research outcomes. Together, their contributions have enriched this study, and I am truly thankful for their collaboration.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Patel P. Groundnut (Peanut/Monkey Nut). Journal of Agronomy. 2021;10(2):1-15.
2. Nautiyal PC. Groundnut: post-harvest operations. Research Centre for Groundnuts (ICAR) [www. icar. org. in] site visited. 2002;23(5):2013. Available:Post_Harvest_Compendium_-_Groundnut.pdf (fao.org).
3. Cultivation of Groundnut; A complete Information Guide. Agrovistafarming.com; 2021. Accessed 6 April 2022. Available: Cultivation of groundnut; A complete Information Guide - Agrovista-Farming (agrovistafarming.com).
4. Misra CM. Trends in area production and productivity of groundnut in India: Issues & Challenges. Journal of Research in Agriculture and Animal Science. 2017;4(7): 01-6.
5. Nayak A, Lokesha H, Gracy CP. Growth and instability analysis of Groundnut production in India and Karnataka. 2021: 61-69
6. Shivappa, Yogish SN. Reviews and reflections on area, production and productivity of groundnut crop in Tumkur District of Karnataka. Journal of Emerging

Technologies and Innovative Research .2019;6(4):479-483
7. Ashoka HG, Sreenatha A, Indrakumar N. Estimation of energy requirement for groundnut cultivation in Karnataka. International Journal of Agricultural Engineering. 2010;3(2):257-62.
8. Gonçalves B, Pinto T, Aires A, Morais MC, Bacelar E, Anjos R, Ferreira-Cardoso J, Oliveira I, Vilela A, Cosme F. Composition of nuts and their potential health benefits—An overview. Foods. 2023;12(5):942.
9. Agricultural production statistics. FAOSTAT Analytical Brief Series No. 60. Rome; 2000–2021..
10. Panigrahi R. Trends in Agricultural Production and Productivity Growth in India: Challenges to Sustainability. Business Governance and Society: Analyzing Shifts, Conflicts and Challenges. 2019;17-28.
11. Saxena R, Chauhan S, Ghosh DK, Kumar R, Kumar D, Das A. Trends and trajectory of Indian agriculture. In trajectory of 75 years of Indian agriculture after independence Singapore: Springer Nature Singapore. 2023;3-22.
12. Talaat FM. Crop yield prediction algorithm (CYPA) in precision agriculture based on IoT techniques and climate changes. Neural Computing and Applications. 2023, 30:1-2.
13. Pan H, Chen Z. Crop Growth Modeling and Yield Forecasting. Agro-geoinformatics: Theory and Practice. 2021: 205-20.
14. Dharmaraja S, Jain V, Anjoy P, Chandra H. Empirical analysis for crop yield forecasting in India. Agricultural Research. 2020;9:132-8.
15. Maradi DM. An explorative study on sustainable grape farming practices in Karnataka: A Case Study of Vijayapur District; 2021.
Available at SSRN 4090624.
16. Poddar RS, Lokesh S, Byahatti S. Impacts of don river flood on agriculture crops, dwelling units and economy of farmers: A Case study of Vijayapura District Karnataka, India. Research Journal of Agricultural Sciences. 2017;8(5):1225-9.
17. Ragini HR, Harshith KV. Identifying existing pattern in area and production of major food grains in Karnataka using linear and non-linear statistical models. International Journal of Statistics and Applied Mathematics. 2024;9(1):90-97.
18. Pázman A. Nonlinear statistical models. Springer Science & Business Media. 2013; 256.
19. Harish Nayak GH, Rajesh Reddy S, Revappa M, Rebasiddanavar G, Avinash, Veershetty, Tamilselvi. Forecasting area, production, and productivity of coffee in hassan district of Karnataka. The Pharma Innovation Journal. 2022;SP-11(11):1399-1406.
20. Shumway RH, Stoffer DS, Shumway RH, Stoffer DS. ARIMA models. Time series analysis and its applications: with R examples. 2017;75-163.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/114659