IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

STATISTICAL ANALYSIS ON AREA AND PRODUCTION OF MAIZE CROP IN KARNATAKA

1. Mr. Vijay M

Department of Agricultural Statistics, Applied Mathematics & Computer Science, CoA, UAS, GKVK, Bengaluru

2. Mr. H. S. Surendra

Department of Agricultural Statistics, Applied Mathematics & Computer Science, CoA, UAS, GKVK, Bengaluru

3. Dr. A. S. Kamble

Assistant professor (Ag. Statistics)

Department of Agricultural Statistics, Applied Mathematics & Computer Science, CoA, UAS, GKVK, Bengaluru

ABSTRACT

Maize crop is grown in most of the states in India as it can be cultivated in various geographical locations. Maize is cultivated for various purposes including grain, fodder, sweet corn, etc. The secondary data on area and production of maize crop along with the weather parameters of Karnataka was collected for the period of 30 years from 1991 to 2020. In this study an attempt has been made to elicit the trend in area and production of maize along with the forecasting of production of the above crop. Further an attempt has been made to know the impact of area and weather parameters on the production of maize crop. Trend analysis was performed by fitting different models such as linear, quadratic, cubic, exponential models and Generalized additive model (GAM). Based on the maximum R^2 value best fitted model was selected. Exponential and GAM was best fitted for production (49.3%) and area (95.1%) of maize crop respectively. Production of maize crop was forecasted for the period of 5 years from 2021 to 2025 using Exponential model.

INTRODUCTION

In India, after rice and wheat, maize is the third most widely grown and consumed staple crop. Maize is known as miracle or queen of cereals because it has the highest genetic yield potential among the cereals. In India the maize was introduced by Portuguese during seventeenth century. It is cultivated for various purposes including grain, fodder, green cobs, sweet corn, baby corn, and popcorn. It supports the livelihood of millions of people around the world. Because of its multiple uses, it is truly regarded as 4F crop *i.e.*, Food, Feed, Fuel and Fodder. Hence, it also contributes to the global economy.

Globally, the area under Maize (*Zea mays* L.) is dominated in China followed b United States and Brazil. Maize is the third most important cereal crop in the world after wheat and rice with an area of 1.86 lakh hectares, production of 121 lakh tonnes and average productivity of 6511 kg ha^{-1} (Anon., 2020a). In India, it is cultivated on an area of 98,90,000 hectares with a production of 2,59,000 tonnes and the productivity of 2690 kg ha^{-1} (Anon., 2021a). The top three states that account for highest area under maize cultivation are Karnataka, Maharashtra and Madhya Pradesh.

Various soil types, from loamy sand to clay loam, can support the growth of maize. A higher level of productivity is thought to be possible with soils that have higher organic matter, good water-holding capacity, and a neutral pH.

In Karnataka, it is cultivated on an area of 17,62,000 hectare with a production of 53,62,130 tonnes with an average productivity of 3043 kg ha^{-1} (Anon., 2021b). It comes up well under a wide range of soil and climatic conditions, there is a lot of scope to expand the present maize yields.

MATERIALS AND METHODS

Nature and Sources of data:

The secondary data on area and production of maize crop of Karnataka was collected from website ***www.data.gov.in*** and the statistical report of ‘Karnataka at a Glance’ published by the Directorate of Economics and Statistics, Government of Karnataka for the period of 30 years from 1991 to 2020.

An Objective formulated to study the trends in area and production of maize crops in Karnataka

Trend refers to the general tendency of the data to increase or decrease over a long period of time. It measures long-term changes occurring in a time-series without bothering about short-term fluctuations occurring in between.

For estimating the long-term trend of area and production of Maize crop, the method of least squares estimation has been employed. In this method, trend in area and production of Maize crop was measured by establishing mathematical relation between time and the response variable. The mathematical expression can be represented by:

A. Linear model

A linear model is one in which all the parameters appear linearly. The average trajectory for the data is a straight line corresponding to increasing or decreasing constant rate of change in time (Nini *et al*, 2017).

$$Y_t = \alpha_0 + \alpha t + \varepsilon \qquad \dots \dots \dots \dots \dots (1)$$

B. Quadratic model

A quadratic function is one which there is 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 and corresponding to an acceleration or deceleration (Nini *et al*, 2017).

$$Y_t = \alpha_0 + \alpha t + \beta t^2 + \varepsilon \qquad \dots \dots \dots \dots \dots (2)$$

C. Cubic model

The cubic is one which there are 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 variable degrees of steepness (Nini *et al*, 2017).

$$Y_t = \alpha_0 + \alpha t + \beta t^2 + \gamma t^3 + \varepsilon \qquad \dots \dots \dots \dots \dots (3)$$

Where,

α_0 : Intercept or Average effect

α, β, γ : Slope or Regression Coefficients (α : linear effect parameter, β : Quadratic effect parameter and γ : cubic effect parameter)

Y_t : Area or production in time period 't'

ε : Error term or disturbance term

The above linear models fitted by using 'lm' function in R programming software.

The coefficients α_0, α, β and γ are constant parameters are need to be estimated. Here, the relation is so derived that the sum of the squared deviations (errors) of the observed values from the theoretical values is least. The process of minimization of the sum of the squared errors results in some equations called normal equations. The normal equations are the equations, which are used for finding the coefficients of the relation, which is fitted by the method of least square.

Time-series are not likely to show either a constant amount of change or a constant ratio of change. The rate of growth is initially slow, and then it picks up and becomes faster and get accelerated, then becomes stable for some time after which it shows retardation.

The following are the two nonlinear growth curves, which were used to describe the growth of present time-series.

D. 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 by

$$Y_t = \alpha e^{\beta t} + \varepsilon \qquad \ldots\ldots\ldots\ldots\ldots (4)$$

Where,

Y_t : represents area or production in time period t

α and β are parameters, α represents the value at $t = 0$, β represents the exponential rate

e is the exponential term

ε denotes the error term

It may be noted that both the above growth models are 'nonlinear', which involves at least one parameter in a nonlinear manner. Exponential model was fitted by using 'SSexpf' function of the package named 'nlraa' in R software.

To evaluate the extent of nonlinearity between area, production and time, the generalized additive model was used.

E. Generalized Additive Model

$$y_t = f(t) + \varepsilon \qquad \ldots.\ \ldots..\ \ldots.(6)$$

Where,

y_t represents area or production in time period 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 Generalized additive model was fitted by using 'gam' function of package 'gam' with automatic smoothing in R software.

Once the parameters of the models were estimated, diagnostic check of residuals of the fitted models has to be analyzed to check any violations in the main assumptions of 'independence of residuals' and 'normality of residuals'. The main assumptions of 'independence of residuals' and 'normality of residuals' was examined by using respectively the 'Run-test' and 'Shapiro-Wilk test' (Prajneshu and Das, 2000).

Test for independence (Randomness) of residuals by Run Test

Non-parametric Run test can be used to test the randomness of residuals. A Run is defined as 'a succession of identical symbols in 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.

Null hypothesis (H_0) : Sequence is random

Alternative Hypothesis (H_1) : Sequence is not random

Let 'n_1', be the number of elements of one kind and 'n_2' be the number of elements of the other kind in a sequence of $N = n_1 + n_2$ binary events. For small samples *i.e.*, both $n_1$and n_2 are equal to or less than 20, if the number of runs r fall between the critical values, we accept the H_0 (null hypothesis) that the sequence of binary events is random otherwise, we reject the H_0.

For large samples *i.e.*, if either $n_1$or n_2 is larger than 15, a good approximation to the sampling distribution of r (runs) is the normal distribution, with

Mean $(\mu_r) = \frac{2n_1n_2}{n_1+n_2}+1$

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

Then H_0can be tested using test statistic:

$$Z = \frac{r-\mu_r}{\sigma_r{}^2} \sim N(0,1)$$

The significance of any observed value of 'Z' computed using the equation may be determined from a normal distribution table.

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 causes to reject the assumption that the distribution is normal. Shapiro-Wilk's W is more appropriate for small samples up to n=50.

H_0: Samples $x_1, \ldots, x_n$ is from a normality distributed population.

H_1: Samples $x_1, \ldots, x_n$ is not from a normality 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

$a_i's$ are constants, 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 (Shapiro *et al.*, 1968). Reject the null hypothesis if W is too small (near to zero).

Model Adequacy Checking

i.The 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.

ii. 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. The adjusted R-squared can be negative, but it’s usually not. 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

iii.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

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_t - \widehat{Y}_t\right)^2}{n}}$$

Where,

Y_t = Observed value

$\widehat{Y}_t$ = Predicted value

n= Number of observations

iv.Akaike Information criterion

The Akaike Information criterion (AIC) is a mathematical method for evaluating how well a model fits the data. AIC is used to compare different possible models and determine which one is the best fit or 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

AIC is calculated for each model and then the model with lowest value is selected and considered as the best fit for the data.

v. 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 (MLE) method. When fitting models, it is possible to increase the likelihood by adding parameters, but doing so may result in overfitting. The BIC resolves this problem by introducing a penalty term for the number of parameters in the model. The penalty term is larger in BIC than in AIC.

The formula for BIC is

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

Where,

n = Sample size

K = Number of independent variables

θ = Set of all the parameters

$L(\theta)$ = Loglikelihood estimate

The models are compared by calculating BIC for each model and then the model with lowest BIC is considered the best. Lower BIC value indicates lower penalty terms hence a better model.

Though these two measures are derived from a different perspective, they are closely related. Apparently, the only difference is BIC considers the number of observations in the formula, which AIC does not. In fact, BIC is always higher than AIC, lower the value of these two measures, better the model.

Forecasting the production of selected major crops in Karnataka

Forecasting is a process of estimating a future event by making use of past data. The past data are systematically combined in a predetermined way to obtain the estimate of the future value. Good forecast can be quite valuable and would be worth a great deal. In the present study, to estimate trend in area and production of selected crops, linear and non-linear models were fitted to find the best-fitted model. Statistical significance of the parameters of linear, quadratic, cubic and GAM model was determined by evaluating student t-test and for other remaining models were determined by computing the 95 per cent asymptotic confidence intervals of the estimated parameters.

Diagnostic check for residuals of the fitted models were checked to know if there any violations in the main assumptions of 'independence of residuals' and 'normality of residuals' using the 'Run-test' and 'Shapiro-Wilk test' respectively. Only for those models, all the parameters are found to be significant at given level of significance, and assumptions of 'independence of residuals' and 'normality of residuals' are satisfied were considered as good fitted models. Among all the good fitted models, the best-fitted model was selected based on minimum MAPE values. This selected best-fitted model was used to forecast the production of selected crops in Karnataka for the period of 5 years from 2020-21 to 2024-25.

For forecasting, the best model selected by testing goodness of fit of all the fitted models by computing Mean Absolute Percent Error (MAPE) which is given by:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| X\ 100$$

Where,

Y_t= Actual values

$\hat{Y}_t$= Predicted values

n = number of observations

Mean Absolute Percentage Error (MAPE) is the most widely used measure for checking forecast accuracy. It comes under percentage errors which are scale independent and can be used for comparing series on different scales. The goodness of fit of all the fitted models are assessed using MAPE. The model with the lowest MAPE value is considered to be the good fit. MAPE is often used as a loss function for regression problems and in evaluation of the model, because of its very intuitive interpretation in terms of relative error.

RESULTS AND DISCUSSION

Trend analysis in area and production of Maize crop in Karnataka

Descriptive statistics

To understand the nature of the data, the descriptive statistics such as mean, standard deviation, skewness, kurtosis, maximum and minimum were computed for the area and production of maize are tabulated in Table-1.

The average area under maize was 12,83,215 hectares and coefficient of variation calculated was 15.49 per cent. The value of skewness was found to be 0.10 which indicates that area was positively skewed, (large amount of data concentrated on the right tail). The kurtosis found to be 0.73 (data was platykurtic). For production mean value was 37,63,338 tonnes and coefficient of variation was 24.28 per cent for the same study period. The value of skewness was found to be 0.69 (Indicates that production was positively skewed). The kurtosis found to be -0.59 (Data was platykurtic). The minimum and maximum area under cultivation was 9,35,854 hectares and 17,26,000 hectares respectively, whereas the production lies between 26,41,529 to 55,52,090 thousand tonnes. The large deviation was found in production of maize over the years, when compared to area.

Table-1: Descriptive statistics for area and production of Maize crop in Karnataka

Measures	Variables	
	Area (ha)	Production (tonnes)
Mean	12,83,215	37,63,338
Median	13,15,563	35,34,421
Standard Deviation	1,98,775	9,13,588
CV (%)	15.49	24.28
Kurtosis	0.73	-0.59
Skewness	0.10	0.69
Range	790,146	29,10,561
Minimum	9,35,854	26,41,529
Maximum	17,26,000	55,52,090

Trend analysis for area of Maize crop

The estimated parameters of all the fitted models and along with their respective standard errors (given in parenthesis) for the area of maize were depicted in Table-2. Test statistic along with their probability values and the goodness of fit criteria are presented in Table-3. The results showed in Table-2 revealed that all the parameters of only linear, cubic GAM and exponential models were found to be significant at 5 per cent level, and the parameters quadratic model was found to be non-significant. Furthermore, results from Table-3 also implied that for the GAM, the Runs test and Shapiro-Wilk test statistic was found to be non-significant indicating that the assumptions of randomness and normality of residuals were satisfied.

The best-fitted model was also selected based on the minimum value of RMSE. As measures of accuracy, the RMSE was computed for all the models along with the other criteria presented in Table-3. The lowest value of RMSE (42,809) indicated that GAM performed better compared to the other models. The values of the other criteria such as R^2, Adjusted R^2, AIC, BIC and MAPE also indicated that GAM was the best fit. The estimated parameter values are $\alpha_0 = 9{,}72{,}391.32$, $\alpha = 36{,}569$. Therefore, GAM was found to be the best-fit model. Similarly, Mallikarjuna (2009) preferred spline model for forecasting cotton area and production in Dharwad. Here, the value α represents the coefficient of smooth function of time, which may be considered as the slope. From the above results it was observed that there was a increasing trend in the area of maize as represented in Fig-1. It is evident from the findings that the area under maize crop also established an increasing trend over the period specifying that there is a marginal replacement of other cropped area to the area concerned to maize crop, as it established a higher revenue in growing this crop (Ismail *et. al, 2019*).

Table-2: Parameter estimates by different models for the area under Maize crop in Karnataka

Parameters	Models				
	Linear	**Quadratic**	**Cubic**	**GAM**	**Exponential**
Intercept (α_0)	9,72,390** (52,070)	9,38,090.20** (87,232.40)	6,68,625** (89,252.60)	9,72,391.32** (52,071)	1.01e+06** (4.62e+04)
α	36,568** (5,385)	48,001 (23,617.50)	2,13,647.50** (44,091.50)	36,569** (5,383)	2.82e-02** (4.32e-03)
β		-672.60 (1,350.60)	-24,309.90* (5,933)		
γ			927* (229.90)		

**Significant at 1% level, * Significant at 5% level,
Values in parenthesis indicate standard error

Table-3: Test for randomness, normality of residuals and goodness of fit criteria of different models for the area under Maize crop in Karnataka

Criteria	Models				
	Linear	**Quadratic**	**Cubic**	**GAM**	**Exponential**
Runs test [*p*-value]	-1.55^{NS} [0.12]	-1.55^{NS} [0.12]	-1.04^{NS} [0.30]	2.07^{NS} [0.21]	-1.55^{NS} [0.12]
Shapiro-Wilk [*p*-value]	0.96^{NS} [0.73]	0.97^{NS} [0.90]	0.97^{NS} [0.86]	0.95^{NS} [0.47]	0.94^{NS} [0.39]
RMSE	92,881	92,007	59,955	42,809	94,431
MAPE	6.312	5.980	4.004	2.904	6.551
AIC	417.456	419.154	407.449	399.969	417.986
BIC	419.774	422.244	411.312	405.107	420.304
R^2	0.767	0.772	0.903	0.951	0.759
Adjusted R^2	0.751	0.736	0.879	0.928	0.747

NS: Non Significant; Values in parenthesis indicate Probability value

Trend analysis for production of Maize crop

The estimated parameters of all the fitted models and along with their respective standard errors for the production of maize crop were depicted in Table-4. Test statistic along with their probability values and the goodness of fit criteria are presented in Table-5. The results showed in Table- 4 revealed that all the parameters of only linear, GAM and exponential models were found to be significant at 5 per cent level of significance, and some of the parameters of other models namely quadratic and cubic models were found to be non-significant. Furthermore, results from Table-5 also implied that for the Exponential, the Runs test and Shapiro-Wilk test statistic was found to be non-significant indicating that the assumptions of randomness and normality of residuals were satisfied.

The best-fitted model was also selected based on the minimum value of RMSE. As measures of accuracy, the RMSE was computed for all the models along with the other criteria presented in Table-5. The lowest value of RMSE (6,26,280) indicated that exponential performed better compared to the other models. The values of the other criteria such as R^2, Adjusted R^2, AIC, BIC and MAPE also indicated that GAM was the best fit. The estimated parameter values are $\alpha = 2.727 \text{ x } 10^6$, $\beta = 3.626 \text{ x } 10^{-2}$. Therefore, Exponential model was found to be the best-fit model. Similarly, Sivasankari (2019) found that the exponential model was the best fit model to the data on green gram production in Tuticorin, Tamil Nadu. Here, the value β represents the coefficient of smooth function of time, which may be considered as the slope. From the above results and also from the Fig-2, it was evident that there is a increasing trend in the production of maize crop. Increase in area has established by statistical technique leading to slightly higher production in the maize crop. It is obvious that there is an interrelation between area and production of the crop (Kalita, 2008).

Table-4: Parameter estimates by different models for the production of Maize crop in Karnataka

Parameters	Models				
	Linear	Quadratic	Cubic	GAM	Exponential
Intercept (α_0)	26,17,740** (3,52,999)	27,80,231** (594,255)	20,65,517 (8,91,448)	37,63,338** (1,68,286)	2.73e+06** (2.94e+05)
α	1,34,776* (36,506)	80,613 (1,60,890)	5,19,964 (4,40,383)	1,34,774* (36,504)	3.63e-02* (9.84e-03)
β		3,186 (9,200)	-59,508 (59,258)		
γ			2,459 (2,296)		

**Significant at 1 per cent, * Significant at 5 per cent,
Values in parenthesis indicate standard error

Table-5: Test for randomness, normality of residuals and goodness of fit criteria of different models for the production of Maize crop in Karnataka

Criteria	Models				
	Linear	Quadratic	Cubic	GAM	Exponential
Runs test [*p*-value]	0.52^{NS} [0.61]	0.52^{NS} [[0.61]	1.04^{NS} [0.30]	0.52^{NS} [0.61]	0.52^{NS} [0.61]
Shapiro-Wilk [*p*-value)]	0.98^{NS} [0.99]	0.96^{NS} [0.99]	0.96^{NS} [0.81]	0.98^{NS} [0.99]	0.98^{NS} [0.99]
RMSE	6,29,667	6,26,982	5,98,825	6,29,667	6,26,280
MAPE	13.271	13.344	13.252	13.271	13.218
AIC	478.700	480.553	481.093	478.700	472.528
BIC	481.018	483.644	484.956	481.018	478.846
R^2	0.493	0.498	0.542	0.493	0.599
Adjusted R^2	0.457	0.421	0.427	0.457	0.585

NS: Non Significant; Values in parenthesis indicate Probability value

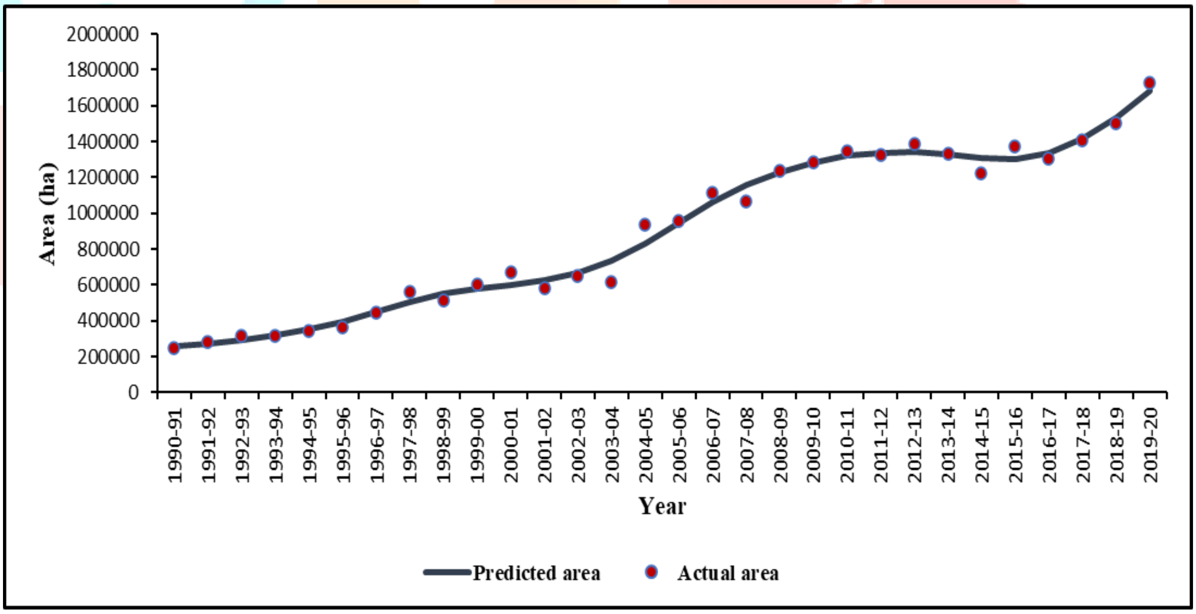

Fig-1: Best fit model for area under Maize crop

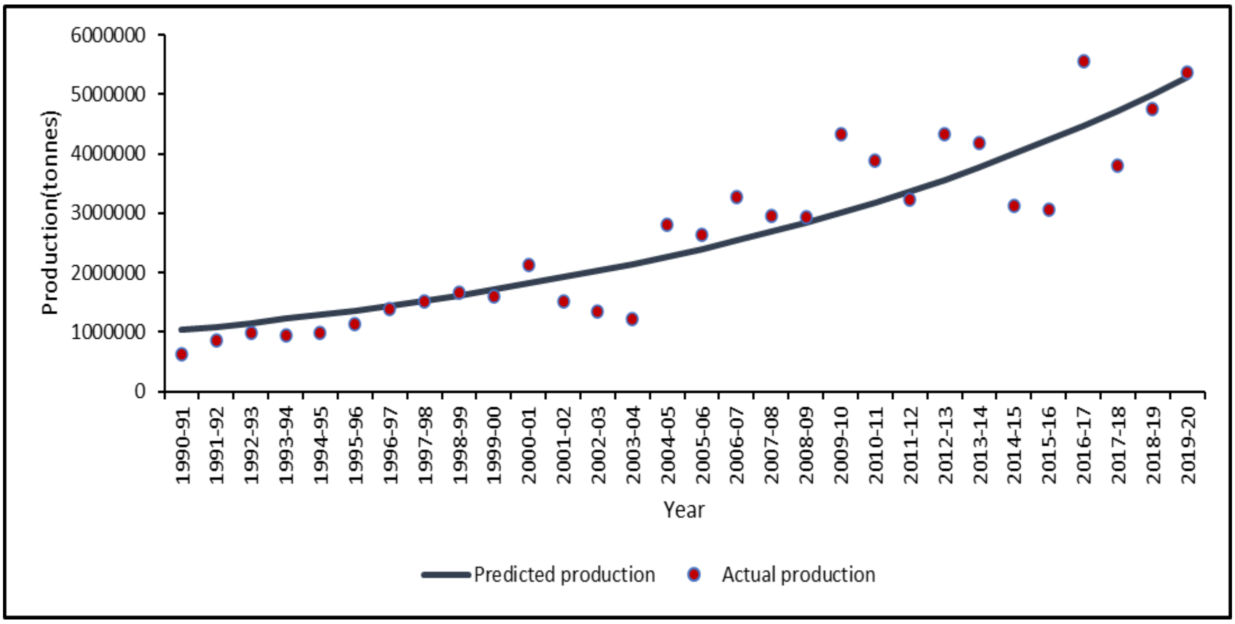

Fig-2: Best fit model for the production of Maize crop

To Forecast the Production of selected major crops in Karnataka

Linear and nonlinear models were fitted to estimate the forecast value of production of both sugarcane and maize in Karnataka. For forecasting, the best-fitted model was used based on the minimum value of MAPE among all the fitted models for predicting value in next 5 years from 2020-21 to 2024-25 and the same was dealt season wise hereunder.

Forecasting of Maize crop production in Karnataka

Different linear and nonlinear models were tried for fitting the data on production of maize crop by using 30 years data from 1991 to 2020. The exponential model was selected as the best-fitted model based on minimum MAPE. This result is similar to Sivasankari (2019). In this study they have used the exponential model for forecasting production of green gram Tuticorin district of Tamil Nadu district in India. The forecasting of maize production was made using this exponential model for next five years (2021 to 2025). The forecasted production of maize were presented in the Table-6 and Fig-3. The results in Table-6 indicated the upward trend. The obvious reason for increase in production of maize crop is based on the increase in maize area over the years. The same is reflected in forecasted production over 5 years. (Rajan *et. al*, 2015).

Table-6: Forecasted production of Maize crop in Karnataka

Year	Forecasted production (tonnes)
2021	50,50,254
2022	52,36,709
2023	54,30,047
2024	56,30,524
2025	58,38,402

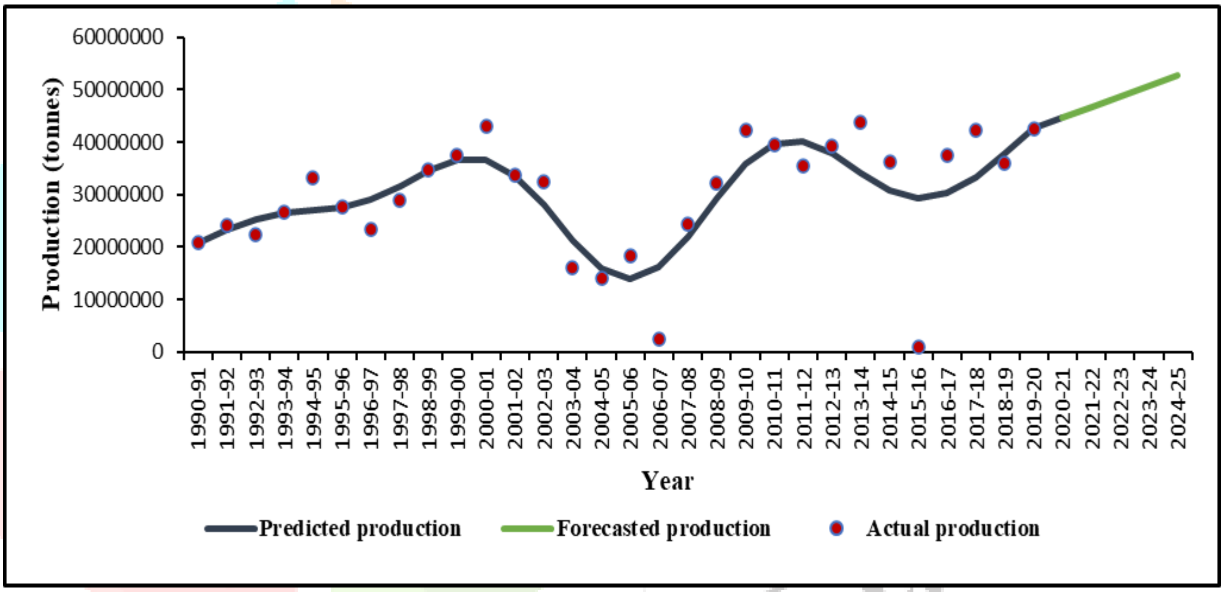

Fig-5: Forecasted production of Sugarcane crop

SUMMARY

Linear, quadratic, cubic, Generalized additive model and exponential model were adopted to analyse the trends in area and production of maize crop grown in Karnataka. The 'Run-test' and 'Shapiro-Wilk test' were used to examine the basic assumptions of independence and normality of residuals. The best-fitted model was chosen based on the lowest RMSE value and also other applied tools such as R^2, MAPE, etc., were used for checking the model adequacy. For forecasting the production of the selected crops efforts made to use time series data, the best fitted model was selected among the models fitted for trend analysis based on the minimum value of MAPE.

With respect to the area of Maize, Gereralized Additive Model was the best-fitted model with the minimum RMSE, whereas the Exponential model was found to be the best-fitted model with the minimum RMSE for the production of maize. Parameters of both GAM and Exponential model were significant and satisfied the assumption of residuals. It is evident from the findings that there exists the increasing trend in maize area over the study period which inturn also caused for the production to have the increasing trend.

The exponential model which showed the increasing trend was found to be the best fit model for maize, with future forecasts ranging from 50,50,254 to 58,38,402 tonnes for the years 2021 to 2025, respectively.

REFERENCES

ANONYMOUS, 2020b, Karnataka at a glance 2019, Directorate of Economics and Statistics, Bengaluru.

ANONYMOUS, 2021a, INDIASTAT- 2021-21.

DAS, P.K., 2000, Growth models for describing state wise wheat productivity. *Indian Journal of Agricultural Research*., **34**(3): 179-181.

ISMAIL, Y., MIR, S.A., NAZIR, N., WANI, M.H., WANI, S.A. AND PUKHTA, M.S., 2019. Trend analysis of area, production and productivity of Cherry in Jammu and Kashmir. *Int. J. Curr. Microbiol. App. Sci*., **8**(2): 2135-2144.

KALITA, S. A., 2008, Trends of area, production and productivity of major fruit crops in Jammu and Kashmir. *Agric. Situ. India*, **65**(7): 477-482.

MALLIKARJUNA, H., 2009, Analysis of spatial and temporal variations in area, production and productivity of cotton in northern Karnataka. M.Sc. Thesis (Unpub.), Acharya N.G. Ranga Agric. Univ., Hyderabad.

NINI, A., CORRADINI, C., GUO, D. AND GRIEVE, J., 2017, The application of growth curve modeling for the analysis of diachronic corpora. *Language Dynamics and Change*, **7**(1): 102-125.

RAJAN, M.S., PALANIVEL, M. AND MOHAN, S.K., 2015. Forecasting of cotton area, production and productivity using trend analysis. *Int. J. Res. in Appl. Sci. and Engg. Tech.*, **3**(12): 516-520.

SIVASANKARI, B., 2019, A statistical modeling approach for forecasting of pulses production in Tamil Nadu. *Int. J. Chem. Stud.*, **7**(3): 3112-3116.