International Journal of Statistics and Applied Mathematics 2023; SP-8(6): 155-164

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2023; SP-8(6): 155-164

https://www.mathsjournal.com
Received: 07-08-2023
Accepted: 10-09-2023

Ganesh V
Department of Agricultural Statistics, Applied Mathematics and Computer Sciences, University of Agricultural Sciences, Bangalore, Karnataka, India

Suman L
Ph.D. Scholar, Department of Agricultural Economics, University of Agricultural Sciences, Bangalore, Karnataka, India

Ajayakumar
Ph.D. Scholar, Department of Agricultural Economics, University of Agricultural Sciences, Bangalore, Karnataka, India

Madhu DM
Ph.D. Scholar, Department of Agricultural Economics, University of Agricultural Sciences, Bangalore, Karnataka, India

Corresponding Author:
Ganesh V
Department of Agricultural Statistics, Applied Mathematics and Computer Sciences, University of Agricultural Sciences, Bangalore, Karnataka, India

The changing landscape of paddy cultivation in Andhra Pradesh: A statistical assessment of area, production, and productivity

Ganesh V, Suman L, Ajayakumar and Madhu DM

DOI: https://dx.doi.org/10.22271/maths.2023.v8.i6Sc.1373

Abstract
The study an attempt was made using secondary data for the period of 23 years (1997 to 2019) to understand the trend in area, production and productivity of paddy crop in selected three districts of Andhra Pradesh *viz*., East Godavari, Krishna and West Godavari district. Trend analysis was performed by fitting different models such as linear, quadratic, cubic, exponential and log-logistic model. Based on the minimum MAPE value, it was evident that exponential model was best fit for area and cubic model were good fit for production and productivity of paddy in East Godavari district. Exponential model good fit model for area and productivity of paddy and linear model was found better model for production in Krishna district. For West Godavari district linear model was best for area and cubic model were deemed to be the good model for production and productivity.

Keywords: Paddy, trend analysis, model fitting, linear, quadratic, cubic, exponential and log-logistic model

1. Introduction
Agriculture is an important sector of Indian economy as it contributes about 20 per cent to the total GDP and provides employment to over 60 per cent of the population. Indian agriculture has registered impressive growth over last few decades. The foodgrain production has increased from 510 lakh tonnes in 1950-51 to 3,160 lakh tonnes during 2021-22. The rapid growth has helped Indian agriculture mark its presence at global level. India stands among top three in terms of production of various agricultural commodities like paddy, wheat, pulses, groundnut, rapeseeds, fruits, vegetables, sugarcane, tea, jute, cotton, tobacco *etc*.

Paddy (*Oryza sativa* L.) is one among the most important cereal crops in the world. Rice provides about 700 calories /day/ person for about 3,000 million people living mostly in developing countries (Sangeetha and Baskar, 2015) [10]. Rice is an important staple food crop for more than 60 per cent of the world people and contributes about 40 per cent of the total food grain production. More than 92 per cent of the world's paddy is produced and consumed in Asia. International year of rice -2004 had the slogan “Rice is life” as its broad meaning encompasses that rice is a staple food for more than half of the world's population and suggesting the role of paddy in poverty and malnutrition. The economic importance of paddy is primarily a high energy calorie food (Siju and Kombairaju, 2001; Madhu *et al.* 2023) [13, 4]. A major part of paddy consists of carbohydrate in the form of starch around 70 per cent. The flattened parboiled rice is known as flaked rice. Flaked rice is also used for preparing different kinds of food items. Bran is an important by-product of rice milling industry. It is used as a cattle feed. Paddy husk is used as fuel. It is also used in brick making. Straw is used as cattle feed, in the manufacture of strawboards and for making hats, ropes, mats, *etc*.

The global total paddy area is 1,620.60 lakh ha and production were about 5,031.70 lakh tonnes with productivity of 4,670 kg/ha in the year 2019 (Anon., 2020) [2]. According to the most recent official data, China was the world's leading paddy producer in 2019, with a production volume of over 2090 lakh tonnes accounting for more than 30 per cent of global paddy production.

India is in second place, contributing for roughly 22 per cent of global paddy output, followed by Indonesia with 8 per cent, Bangladesh with 7 per cent, and Vietnam with 6 per cent (Shah, 2019) [11].

India is the second largest producer and consumer of paddy in the world which occupies an area of 4366 lakh ha and production of 11,887 lakh tonnes with an average productivity of 2,722 kg/ha (Anon., 2020) [2]. In India, the major paddy producing states are West Bengal, Andhra Pradesh, Punjab, Uttar Pradesh and Tamil Nadu, Karnataka, Chhattisgarh, and Odisha, *etc*.

Andhra Pradesh State is called as "The bejewelled rice bowl of India". Agriculture plays an important role in the livelihoods of people as 63 per cent of the population in Andhra Pradesh live in rural areas and depend on agriculture and related livelihood opportunities. Andhra Pradesh is one of the leading producers of paddy which ranks 3rd position in Production and produces 128.95 lakh tonnes of paddy in India. It is a leading paddy producer with a production of 10 per cent of total paddy produced in the country. In Andhra Pradesh, paddy is the major crop cultivated more than 22 lakh ha during Kharif and Rabi seasons, 13 districts of Andhra Pradesh are producing paddy crop out of which West Godavari, East Godavari, Krishna, Guntur, Srikakulam, Vizianagaram and Chittoor are the major producers. West Godavari, East Godavari and Krishna are three most important paddy producing districts not only in Andhra Pradesh but in whole of India. West Godavari, and East Godavari districts considered a "Rice Bowl" of Andhra Pradesh (Pradeep and Krishna 2002; Sharma 2008) [8, 12].

Trend analysis is the phrase used to describe techniques for identifying hidden behavioural patterns in time series. Trend analysis can be used to assess complicated past occurrences even though it is normally employed to forecast the future (Mishra *et al.* 2012; Arunachalam and Balakrishnan, 2012; Sudha *et al*. 2013) [6, 14]. One of the main aims of time series analysis is forecasting, or the ability to predict "what will happen in the future" rather than "why". There are many forecasting models in use today. While several forecasting models may equally well fit a set of facts, they predict distinct futures (Choudhury *et al.* 2017) [3]. For proper, foresighted, and informed planning, especially in the uncertain world of agriculture, forecasting is a vital and essential input (Mohan Kumar *et al.* 2012; Rajan *et al.* 2015) [7, 9,]. The formulation of plans for stock distribution and the agricultural food supply to various regions of the nation benefits greatly from crop output estimates. It is also used to forecast the growth of crops. Considering this aspect study is confined only to the area, production and productivity of paddy crop.

2. Methodology

2.1 Nature and Sources of data

The secondary data on area, production and productivity of paddy crop in selected districts of Andhra Pradesh was collected from website *https://aps.dac.gov.in* and the statistical report from district administration, Andhra Pradesh for the period of 23 years from 1997 to 2019 was collected from the Department of Economics and Statistics, Government of Andhra Pradesh.

2.2 Analytical tools and techniques

For estimating the long-term trend in area, production, and productivity of paddy in selected districts of Andhra Pradesh, the method of least squares estimation has been employed. In this method, trend in area, production and productivity of paddy is measured by establishing mathematical relation between time and the response variable, which is depending on time. The mathematical expression can be represented by:

2.2.1 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).

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

2.2.2 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.

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

2.2.3 Cubic model

The cubic model is one in which there are two troughs in the data. The average trajectory for the bundle behaves quadratically until a further curve occurs, which can correspond to an acceleration or deceleration with variable degrees of steepness (Nini *et al*, 2017).

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

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 or productivity of paddy in time period t

ε: Error term or disturbance term

The above linear models fitted by using 'lm' function of R

Coefficients α_0, α, β and γ are constant parameters which 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.

2.2.4 Exponential model

If, when 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

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

Where,

Y_t represents area or production or productivity of paddy 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, and

ε denotes the error term.

2.2.5 Log-logistic

$$Y_t = \frac{\alpha}{1+exp[-\beta\{|\log(t)-\log(\gamma)|\}]} + \varepsilon \quad (5)$$

Where,
Y_t represents area or production or productivity of paddy in time period *t*
α, β and *γ* are parameters and
ε denotes the error term.

The parameter *'γ'* is the 'intrinsic growth rate', while the parameter '*α*' represents the 'upper asymptote' and *'β'* is the growth range.
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. The model loglogistic was fitted by using 'loglogistic' function of the package 'growthmodels' in R.
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'.

2.2.6 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. 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

$$\text{Mean } (\mu_r) = \frac{2n_1n_2}{n_1+n_2} + 1 \quad (6)$$

$$\text{Variance } (\sigma_r{}^2) = \sqrt{\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1n_2)^2(n_1+n_2-1)}} \quad (7)$$

Then H_0can be tested using test statistic:

$$Z = \frac{r-\mu_r}{\sigma_r} \sim N(0,1) \quad (8)$$

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

2.2.7 Test for normality of residuals by Shapiro-Wilk's (W) test
The Shapiro- Wilks test is the standard test checking normality in the data set. 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 ...x_n is from a normality distributed population.
H_1: Samples x_1 ,…x_n is not from a normality distributed population.
Test statistic is given by:

$$W = \frac{\left[\sum_{i-1}^{n} a_i x_{(i)}\right]^2}{\sum_{i-1}^{n}(x-\bar{x})^2} \quad (9)$$

$\bar{x}$is sample mean and the constants $a_i$is given by

$$(a_1, a_2, \cdots, a_n) = \frac{m^T V^{-1}}{\sqrt{(m^T V^{-1} V^{-1} m)}} \quad (10)$$

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).

2.3 Model Adequacy Checking
2.3.1 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 by the statistical model. It provides a measure of how well the assumed model explains the variability in dependent variable.

$$R^2 = \frac{RSS}{TSS} = 1 - \frac{ESS}{TSS} \quad (11)$$

Where,
ESS = Error sum of squares
RSS = Regression sum of squares
TSS = Total sum of squares

Computed R^2 value lies between zero and one. If R^2 value is closer to one 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.

2.3.2 Adjusted R^2
The adjusted R^2 is a modified version of R^2 that has been adjusted for the number of predictors in the model. The adjusted R^2 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^2 is always lower than the R^2.

$$Adjusted\ R^2 = \frac{RSS/_{df_R}}{TSS/_{df_T}} \quad (12)$$

Where,

RSS = Regression sum of squares
TSS = Total sum of squares
df_R = RSS degrees of freedom
df_T = TSS degrees of freedom

2.3.3 Mean Average Percentage Error (MAPE)

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.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i-\hat{Y}_i}{Y_i}\right| X\, 100 \qquad (13)$$

Where,
Y_i= Actual values
$\hat{Y}_i$= Predicted values
n= number of observations

2.3.4 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

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_t-\hat{Y}_t)^2}{n}} \qquad (14)$$

Where,
Y_t = Observed value
$\hat{Y}_t$ = Predicted value
n= Number of observations

2.3.5 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 for 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 formula for AIC is

$$AIC = 2K - 2\,ln(L(\theta)) \qquad (15)$$

Where,
K – Number of independent variables
$L(\theta)$ – Log-likelihood estimate
ln = log to the base e
AIC is calculated for each model and then the model with lowest value is selected and considered as the best fit for the data.

2.3.6 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 was 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

$$BIC = Kln(n) - 2ln(L(\theta)) \qquad (16)$$

Where,
n – sample size
K – Number of independent variables
θ – set of all the parameters
$L(\theta)$ – Loglikelihood estimate
ln = log to the base e

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.

3. Results and Discussion

In any time series data, there is a tendency to either increase or decrease over time. This tendency of data to increase or decrease over time is called trend. For estimating the trends in area, production and productivity of paddy in selected districts of Andhra Pradesh, the annual data for the period of 23 years from 1997 to 2019 were collected. Different polynomial models were fitted for the time series data on area, production and productivity of paddy in selected districts of Andhra Pradesh.
Statistical significance of the parameters of the linear, quadratic, cubic, exponential, and log-logistic model was determined by evaluating student *t*-test, and remaining models were determined by computing the 95 per cent asymptotic confidence intervals of the estimated parameters. The estimated parameter of the fitted model when lies within the 95 per cent confidence interval indicates then parameter values are significant at 5 per cent level of significance. The main assumption of normality of error terms of each model were examined by using “Shapiro-Wilk test”, and test statistic along with probability values are presented in tables given below. Among all the good fit models, the best-fitted model was selected based on the minimum MAPE value. As measures of accuracy, the MAPE was computed for all the models presented in tables given below.

3.1 Parameter estimates and goodness of fit criteria of different models for Paddy in East Godavari District

3.1.1 Parameter estimates and goodness of fit criteria of different models for the area under paddy in East Godavari district

Table 1 provides the parameter estimates for the five fitted models and their standard errors (in parenthesis) for the paddy area in the East Godavari district. Table 2 provides the test statistic, probability values, and model adequacy standards. According to the results, the calculated parameters of the linear, exponential models were significant at 5 per cent level of significance. Table 2 results also showed that the Run's test and Shapiro-test Wilk's statistic were non-significant for all fitted models. Therefore, the data of the area under paddy in East Godavari district for the study period fitted well to linear and exponential model.

Based on the lowest MAPE value amongst two well-fit models, the optimal model was chosen. It was observed that exponential MAPE (5.24) was the lowest. The exponential model had the best fit, as evidenced by the results of the other criteria, including R^2 (0.60), Adjusted R^2 (0.58), AIC (545.91), BIC (549.32) and RMSE (30283.41). So, it can be said based on the data from 1997 to 2019 for the paddy area in East Godavari the growth is exponential (Table 2).

Table 1: Parameter estimates by different models for the area under paddy in East Godavari district

Parameter	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Intercept (α_0)	387117.00** (13660.00)	394732.00 ** (22110.40)	397695.00** (32842.10)	3.87e+05** (1.36e+04)	1.24e+00** (5.81e+00)
α	-138.40** (996.30)	-1965.90 (4244.50)	-3307.00** (11599.50)	-3.61e-04** (2.58e-03)	3.86e+05** (1.91e+04)
β		76.10 (171.70)	212.90 (1110.50)		1.10e+03 (1.40e+04)
γ			-3.70 (30.40)		

**Significant at 1% level, *Significant at 5% level, Figures in parentheses indicate standard errors.

Table 2: Test for randomness, normality of residuals and goodness of fit criteria of different models for the area under paddy in East Godavari district

Criterion	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Runs test (*p*-value)	-0.43^{NS} (0.66)	-0.43^{NS} (0.66)	-0.43^{NS} (0.66)	-0.43^{NS} (0.66)	-0.43^{NS} (0.66)
Shapiro-Wilk (*p*-value)	0.76^{NS} (0.17)	0.78^{NS} (0.18)	0.78^{NS} (0.18)	0.76^{NS} (0.17)	0.76^{NS} (0.18)
RMSE	30283.51	30135.71	30123.38	30283.41	30289.64
MAPE	5.26	5.43	5.43	5.24	5.25
AIC	545.91	547.69	549.67	545.91	547.92
BIC	549.32	552.23	555.34	549.32	552.46
R^2	0.58	0.59	0.55	0.60	0.52
Adjusted R^2	0.54	0.56	0.52	0.58	0.48

**Significant at 1% level, *Significant at 5% level; NS: Non significant; Figures in parentheses indicate Probability value.

3.1.2Parameter estimates and goodness of fit criteria of different models for production of paddy in East Godavari district

Table 3 provides the parameter estimates for the five fitted models and their standard errors (in parenthesis) for the paddy production in the East Godavari district. Table 4 provides the test statistic, probability values, and model adequacy standards. According to the results, the calculated parameters of the linear, cubic, exponential, and log-logistic models were significant at 5 per cent level of significance. Table 4 results also showed that the Run's test and Shapiro-test Wilk's statistic were non-significant for all fitted models. Therefore, the data for production of paddy in East Godavari district for the study period fitted well to the linear, cubic, exponential, and log-logistic model.

Based on the lowest MAPE value amongst four well-fit models, the optimal model was chosen. The lowest MAPE value (9.66) was for the cubic model. The cubic model had the best fit, as evidenced by the results of the other criteria, including R^2 (0.53), Adjusted R^2 (0.46), AIC (625.59), BIC (631.27) and RMSE (156932.00). It can be concluded that growth is cubic for production of paddy for the year from 1997 to 2019 in East Godavari (Table 4).

Table 3: Parameter estimates by different models for production of paddy in East Godavari district

Parameter	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Intercept (α_0)	1120974.0 ** (85654.0)	1174177.0** (138477.0)	813708.0** (171095.0)	1.13e+06** (7.74e+04)	-2.85e-01** (9.13e-02)
α	19736.0** (6247.0)	6967.0 (26583.0)	170101.0* (60429.0)	1.47e-02** (4.61e-03)	2.33e+06** (1.00e+01)
β		532.0 (1075.0)	-16105.0* (5785.0)		2.97e+00* (1.31e+00)
γ			462.1** (158.7)		

**Significant at 1% level, *Significant at 5% level, Figures in parentheses indicate standard errors.

Table 4: Test for randomness, normality of residuals and goodness of fit criteria of different models for production of paddy in East Godavari district

Criterion	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Runs test (*p*-value)	-0.43 NS (0.66)	-0.43 NS (0.66)	-1.31 NS (0.18)	-0.43 NS (0.66)	-0.873 NS (0.38)
Shapiro-Wilk (*p*-value)	0.94 NS (0.18)	0.94 NS (0.28)	0.95 NS (0.44)	0.943NS (0.21)	0.928 NS (0.10)
RMSE	189890.20	188738.90	156932.00	189259.20	188706.90
MAPE	12.67	12.53	9.66	12.62	12.00
AIC	630.36	632.08	625.59	630.21	632.07
BIC	633.77	636.62	631.27	633.61	636.61
R^2	0.32	0.33	0.53	0.32	0.33
Adjusted R^2	0.28	0.26	0.46	0.29	0.29

**Significant at 1% level, *Significant at 5% level; NS: Non Significant; Figures in parentheses indicate Probability value.

3.1.3 Parameter estimates and goodness of fit criteria of different models for the productivity under paddy in East Godavari district

Table 5 provides the parameter estimates for the five fitted models and their standard errors (in parenthesis) for the paddy productivity in the East Godavari district. Table 6 provides the test statistic, probability values, and model adequacy standards. According to the results, the calculated parameters of the linear, cubic, and exponential and models were significant at 5 per cent level of significance. Table 6 results also showed that the Run's test and Shapiro-test Wilk's statistic were non-significant for all fitted models expect log-logistic model. Therefore, the data for the productivity under paddy in East Godavari district for the study period was well fitted to the linear, cubic, and exponential and models.

Based on the lowest MAPE value amongst three well-fit models, the optimal model was chosen. The cubic with the lowest MAPE (6.96) value was observed. The cubic model had the best fit, as evidenced by the results of the other criteria, including R^2 (0.65), Adjusted R^2 (0.60), AIC (22.76), BIC (28.44) and RMSE (0.31).

Table 5: Parameter estimates by different models for the productivity under paddy in East Godavari district

Parameter	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Intercept (α_0)	2.89** (0.19)	2.97** (0.30)	2.02** (0.34)	2.93** (0.172)	-0.28 (0.42)
α	0.05** (0.01)	0.03 (0.05)	0.46** (0.12)	0.01** (0.00)	6.33 (8.53)
β		0.00 (0.00)	-0.04** (0.01)		4.19 (40.80)
γ			0.00** (0.00)		

**Significant at 1% level, *Significant at 5% level, Figures in parentheses indicate standard errors.

Table 6: Test for randomness, normality of residuals and goodness of fit criteria of different models for the productivity under paddy in East Godavari district

Criterion	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Runs test (*p*-value)	-1.74NS (0.08)	-1.31 NS (0.18)	-0.43 NS (0.66)	-1.31 NS (0.18)	-2.62** (0.00)
Shapiro-Wilk (*p*-value)	0.95 NS (0.42)	0.95 NS (0.42)	0.93 NS (0.15)	0.95 NS (0.43)	0.95 NS (0.40)
RMSE	0.42	0.42	0.31	0.42	0.41
MAPE	9.83	9.76	6.96	9.78	8.98
AIC	31.71	38.15	22.76	31.60	32.78
BIC	35.12	38.15	28.44	35.01	37.32
R^2	0.40	0.40	0.65	0.40	0.42
Adjusted R^2	0.37	0.34	0.60	0.37	0.39

**Significant at 1% level, *significant at 5% level; NS: Non Significant; Figures in parentheses indicate Probability value.

3.2 Parameter estimates and goodness of fit criteria of different models for paddy in Krishna district

3.2.1 Parameter estimates and goodness of fit criteria of different models for the area of paddy in Krishna district

Table 7 provides the parameter estimates for the five fitted models and their standard errors (in parenthesis) for the paddy area in the Krishna district. Table 8 provides the test statistic, probability values, and model adequacy standards. According to the results, the calculated parameters of the linear and exponential models were significant at 5 per cent level of significance. The results also showed that the Run's test and Shapiro-test Wilk's statistic were non-significant for all the models. Linear and exponential models are good fit to the area under paddy in Krishna district for the study period (Table 8).

Based on the lowest MAPE value amongst two well-fit models, the optimal model was chosen. The lowest MAPE (13.39) value was observed for the exponential model. The exponential model had the best fit, as evidenced by the results of the other criteria, including R^2 (0.29), Adjusted R^2 (0.26), AIC (570.70), BIC (574.11) and RMSE (51909.44). It can be concluded that growth is exponential for the paddy area in Krishna for the years from 1997 to 2019 (Table 8).

Table 7: Parameter estimates by different models for the area of paddy in Krishna district

Parameter	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Intercept (α_0)	37502.00** (23468.00)	387077.10** (38014.90)	416695.76** (55702.66)	3.79e+05** (2.57e+04)	3.01e-01 (6.20e-01)
α	-4806.00* (1712.00)	-7698.12 (7297.69)	-21102.51 (19673.74)	-1.53e-02* (5.42e-03)	5.69e+05 (7.50e+05)
β		76.17 (171.71)	1487.54 (1883.58)		2.06e+01 (1.75e+02)
γ			-37.97 (51.66)		

**Significant at 1% level, *Significant at 5% level, Figures in parentheses indicate standard errors.

Table 8: Test for randomness, normality of residuals and goodness of fit criteria of different models for the area of Paddy in Krishna district

Criterion	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Runs test (*p*-value)	-1.74 NS (0.08)	-1.74 NS (0.08)	-1.74 NS (0.08)	-1.74 NS (0.08)	-1.74 NS (0.08)
Shapiro-Wilk (*p*-value)	0.94 NS (0.21)	0.94 NS (0.25)	0.94 NS (0.20)	0.94 NS (0.23)	0.94 NS (0.25)
RMSE	52028.24	51812.81	51091.45	51909.44	51408.03
MAPE	13.44	13.51	14.09	13.39	14.026
AIC	570.81	572.61	573.97	570.70	572.25
BIC	574.21	577.16	579.65	574.11	576.80
R^2	0.27	0.27	0.29	0.29	0.29
Adjusted R^2	0.23	0.20	0.18	0.26	0.25

**Significant at 1% level, *significant at 5% level; NS: Non Significant; Figures in parentheses indicate Probability value.

3.2.2 Parameter estimates and goodness of fit criteria of different models for the production under paddy in Krishna district

Table 9 provides the parameter estimates for the five fitted models and their standard errors (in parenthesis) for the paddy production in the Krishna district. Table 10 provides the test statistic, probability values, and model adequacy standards. According to the results, the calculated parameters of the linear and log-logistic models were significant at 5 per cent level of significance. Table 10 results also showed that the Run's test and Shapiro-test Wilk's statistic non-significant for all fitted models. Linear and Log-logistic models were better fit for the production data of paddy in Krishna district in the study period.

Based on the lowest MAPE value amongst two well-fit models, the optimal model was chosen. The lowest MAPE (15.16) value was for linear model. The linear model had the best fit, as evidenced by the results of the other criteria, including R^2 (0.36.), Adjusted R^2 (0.32), AIC (634.25), BIC (637.65) and RMSE (206633.40). Paddy production in Krishna district has shown a linear growth for the data from the year 1997 to 2019 (Table 10).

Table 9: Parameter estimates by different models for the production under paddy in Krishna district

Parameter	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Intercept (α_0)	1080135.00** (93206.00)	1175728.50** (149087.00)	1242786.95 ** (220516.04)	1.08e+06 ** (9.41e+04)	3.00e-02** (4.71e-03)
α	-1668.00** (6798.00)	-24609.90 (28619.90)	-54958.12 (77884.53)	-1.62e-03 (6.41e-03)	1.37e+06** (1.00e+01)
β		955.90 (1157.80)	4050.93 (7456.74)		3.02e+18** (1.00e+01)
γ			-85.97 (204.51)		

**Significant at 1% level, *Significant at 5% level, Figures in parentheses indicate standard errors.

Table 10: Test for randomness, normality of residuals and goodness of fit criteria of different models for the production under paddy in Krishna district

Criterion	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Runs test (*p*-value)	-1.31 NS (0.18)	-0.87 NS (0.38)	-0.87 NS (0.38)	-1.31 NS (0.18)	-1.31 NS (0.18)
Shapiro-Wilk (*p*-value)	0.96 NS (0.68)	0.98 NS (0.92)	0.98 NS (0.98)	0.96 NS (0.68)	0.97 NS (0.73)
RMSE	206633.40	203199.60	202261.20	206623.70	206275.80
MAPE	15.16	15.19	15.10	15.25	15.27
AIC	634.25	635.48	632.26	634.24	636.17

BIC	637.65	640.02	638.94	637.65	640.71
R^2	0.36	0.33	0.33	0.27	0.28
Adjusted R^2	0.32	0.26	0.27	0.24	0.24

**Significant at 1% level, *Significant at 5% level; NS: Non Significant; Figures in parentheses indicate Probability value.

3.2.3 Parameter estimates and goodness of fit criteria of different models for the productivity of paddy in Krishna district

Table 11 provides the parameter estimates for the five fitted models and their standard errors (in parenthesis) for the paddy productivity in the Krishna district. Table 12 provides the test statistic, probability values, and model adequacy standards. According to the results, the calculated parameters of the linear and cubic models were significant at 5 per cent level of significance. Table 12 results also showed that the Run's test and Shapiro-test Wilk's statistic were non-significant for all the models except linear and log-logistic model. Exponential model is a superior fit for the productivity of paddy in Krishna district for the study period.

The exponential has the lowest MAPE (7.21) value as observed from the table 12. The exponential model had the best fit, as evidenced by the results of the other criteria, including R^2 (0.64), Adjusted R^2 (0.61), AIC (15.78), BIC (19.19) and RMSE (0.29). As a result, from 1997 to 2019, the data for the paddy productivity in Krishna has shown an exponential growth (Table 12).

Table 11: Parameter estimates by different models for the productivity of paddy in Krishna district

Parameter	Model				
	Linear	Quadratic	Cubic	Exponential	Log-logistic
Intercept (α_0)	2.78** (0.13)	3.04** (0.21)	3.04e+00** (3.12e-01)	2.93** (0.17)	-0.16* (0.05)
α	0.04** (0.00)	-0.01 (0.04)	-1.41e02** (1.10e-01)	0.01** (0.00)	8.64* (3.99)
β		0.00 (0.00)	2.63e-03 (1.05e-02)		136.38 (687.47)
γ			-1.26e-06 (2.89e-04)		

**Significant at 1% level, *Significant at 5% level, Figures in parentheses indicate standard errors.

Table 12: Test for randomness, normality of residuals and goodness of fit criteria of different models for the productivity of paddy in Krishna district

Criterion	Model				
	Linear	Quadratic	Cubic	Exponential	Log-logistic
Runs test (p-value)	-0.87^{NS} (0.38)	-0.43^{NS} (0.66)	-0.43^{NS} (0.66)	-0.43^{NS} (0.66)	-2.18** (0.02)
Shapiro-Wilk (p-value)	0.84** (0.00)	0.94^{NS} (0.18)	0.94^{NS} (0.19)	0.94^{NS} (0.18)	0.89** (0.01)
RMSE	0.30	0.28	0.28	0.29	0.34
MAPE	7.30	6.88	6.89	6.21	8.11
AIC	16.45	15.72	17.72	15.78	24.95
BIC	19.86	20.27	23.40	19.19	29.50
R^2	0.52	0.58	0.58	0.64	0.37
Adjusted R^2	0.50	0.53	0.51	0.61	0.34

**Significant at 1% level, *Significant at 5% level; NS: Non Significant; Figures in parentheses indicate Probability value.

3.3 Parameter estimates and goodness of fit criteria of different models for paddy in West Godavari district

3.3.1 Parameter estimates and goodness of fit criteria of different models for the area under paddy in West Godavari district

Table 13 provides the parameter estimates for the five fitted models and their standard errors (in parenthesis) for the paddy area in the West Godavari district. Table 14 provides the test statistic, probability values, and model adequacy standards. According to the results, the calculated parameters of the linear and exponential models were significant at 5 per cent level of significance. Table 14 results also showed that the Run's test and Shapiro-test Wilk's statistic were non-significant for all fitted models. Therefore, the data of the area under paddy in West Godavari district for the study period was well fitted to the linear and exponential models.

Based on the lowest MAPE value amongst two well-fit models, the optimal model was chosen. The linear with the lowest MAPE (2.74) value was observed. The linear model had the best fit, as evidenced by the results of the other criteria, including R^2 (0.79), Adjusted R^2 (0.57) AIC (512.43), BIC (513.84) and RMSE (15205.61). As a result, from 1997 to 2019, the data for the paddy area in West Godavari has shown a linear growth (Table 14).

Table 13: Parameter estimates by different models for the area under paddy in west Godavari district

Parameter	Model				
	Linear	Quadratic	Cubic	Exponential	Log-logistic
Intercept (α_0)	453450.00** (6890.40)	452357.65** (11203.26)	456312.88** (16600.40)	4.54e+05** (7.15e+03)	7.83e-02** (2.45e-02)
α	-2450.50 ** (502.50)	-2188.38 (2150.66)	-3978.37 (5863.13)	-5.77e-03** (1.18e-03)	9.46e+05** (2.30e+05)
β		-10.92 (87.01)	171.62 (561.34)		6.54e-01 (4.04e+00)
γ			-5.07 (15.39)		

**Significant at 1% level, *Significant at 5% level, Figures in parentheses indicate standard errors.

Table 14: Test for randomness, normality of residuals and goodness of fit criteria of different models for the area under paddy in West Godavari district

Criterion	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Runs test (p-value)	-0.43 NS (0.66)	0.43 NS (0.66)	0.87 NS (0.38)	-0.43 NS (0.66)	-0.43 NS (0.66)
Shapiro-Wilk (p-value)	0.97 NS (0.81)	0.97 NS (0.82)	0.97 NS (0.81)	0.97 NS (0.82)	0.94 NS (0.25)
RMSE	15205.61	15269.59	15226.19	15285.68	16604.83
MAPE	2.74	2.92	2.92	2.96	3.48
AIC	512.43	516.41	518.28	514.46	520.27
BIC	513.84	520.95	523.96	517.87	524.81
R^2	0.79	0.62	0.63	0.53	0.64
Adjusted R^2	0.57	0.48	0.46	0.50	0.47

**Significant at 1% level, *Significant at 5% level; NS: Non Significant; Figures in parentheses indicate Probability value.

3.3.2 Parameter estimates and goodness of fit criteria of different models for production of paddy in West Godavari district

Table 15 provides the parameter estimates for the five fitted models and their standard errors (in parenthesis) for the paddy production in the West Godavari district. Table 16 provides the test statistic, probability values, and model adequacy standards. According to the results, the calculated parameters of the linear, cubic, and exponential models were significant at 5 per cent level of significance. Table 16 results also showed that the Run's test and Shapiro-test Wilk's statistic were determined to be non-significant for all the models. Therefore, the data for production of paddy in West Godavari district for the study period was well fitted to the linear and exponential model.

The cubic with the lowest MAPE (4.52) value was observed. The cubic model had the best fit, as evidenced by the results of the other criteria, including R^2 (0.76), Adjusted R^2 (0.71), AIC (599.73), BIC (605.41) and RMSE (89441.02). As a result, from 1997 to 2019, the data for the paddy production in West Godavari has shown a cubic growth (Table 16).

Table 15: Parameter estimates by different models for production of paddy in West Godavari district

Parameter	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Intercept (α_0)	1431974.00** (51872.00)	1508044.00** (81481.00)	1281732.51** (97513.30)	1.43e+06** (4.97e+04)	-5.11e-01** (5.21e-01)
α	9426.00* (3783.00)	-8830.90 (15641.00)	93589.52* (34440.23)	6.21e-03* (2.45e-03)	1.67e+06** (2.18e+05)
β		760.70 (632.00)	-9684.44** (3297.93)		6.39e-02 (1.16e-01)
γ			290.14** (90.40)		

**Significant at 1% level, *Significant at 5% level, Figures in parentheses indicate standard errors.

Table 16: Test for randomness, normality of residuals and goodness of fit criteria of different models for production of paddy in West Godavari district

Criterion	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Runs test (p-value)	-0.87 NS (0.38)	-1.74 NS (0.08)	-0.43 NS (0.66)	-0.87 NS (0.38)	0 NS (1.00)
Shapiro-Wilk (p-value)	0.94 NS (0.24)	0.96 NS (0.65)	0.93 NS (0.15)	0.94 NS (0.26)	0.94 NS (0.21)
RMSE	114998.20	111056.00	89441.02	114697.70	116948.60
MAPE	6.20	5.84	4.52	6.18	6.03
AIC	607.29	607.68	599.73	607.17	610.06
BIC	610.70	612.23	605.41	610.58	614.60
R^2	0.62	0.58	0.76	0.50	0.55
Adjusted R^2	0.57	0.53	0.71	0.46	0.50

**Significant at 1% level, *Significant at 5% level; NS: Non Significant; Figures in parentheses indicate Probability value.

3.3.3 Parameter estimates and goodness of fit criteria of different models for the productivity under paddy in West Godavari district

Table 17 provides the parameter estimates for the five fitted models and their standard errors (in parenthesis) for the paddy productivity in the West Godavari district. Table 18 provides the test statistic, probability values, and model adequacy standards. According to the results, the calculated parameters of the linear, cubic, and exponential models were significant at 5 per cent level of significance. Table 18 results also showed that the Run's test and Shapiro-test Wilk's statistic were non-significant for all fitted models. Therefore, the data of the productivity under paddy in West Godavari district for the study period was well fitted to the linear, cubic, and exponential models.

Based on the lowest MAPE value amongst three well-fit models, the optimal model was chosen. The cubic with the lowest MAPE (5.14) value was observed. The cubic model had the best fit, as evidenced by the results of the other criteria, including R^2 (0.59), Adjusted R^2 (0.54), AIC (10.10), BIC (15.78) and RMSE (0.24). As a result, from 1997 to 2019, the data for the paddy productivity in West Godavari has shown a cubic growth (Table 16).

Table 17: Parameter estimates by different models for the productivity under paddy in West Godavari district

Parameter	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Intercept (α_0)	3.15** (0.15)	3.46** (0.23)	2.72** (0.26)	3.15** (0.14)	-0.15** (0.08)
α	0.04** (0.01)	-0.03 (0.04)	0.29** (0.09)	0.01** (0.00)	8.24 (5.57)
β		0.00 (0.00)	-0.03** (0.00)		45.91 (403.56)
γ			0.00** (0.00)		

**Significant at 1% level, *Significant at 5% level, Figures in parentheses indicate standard errors.

Table 18: Test for randomness, normality of residuals and goodness of fit criteria of different models for the productivity under paddy in West Godavari district

Criterion	Model				
	Linear	**Quadratic**	**Cubic**	**Exponential**	**Log-logistic**
Runs test (p-value)	-1.74^{NS} (0.08)	-2.62 ^{NS} (0.08)	-0.43 ^{NS} (0.66)	-1.74 ^{NS} (0.08)	-2.62 ^{NS} (0.08)
Shapiro-Wilk (p-value)	0.97 ^{NS} (0.88)	0.94 ^{NS} (0.27)	0.98 ^{NS} (0.97)	0.97 ^{NS} (0.86)	0.95 ^{NS} (0.45)
RMSE	0.34	0.32	0.24	0.34	0.36
MAPE	7.55	7.24	5.14	7.49	7.59
AIC	22.41	21.34	10.10	21.98	26.68
BIC	25.82	25.88	15.78	25.39	31.22
R^2	0.37	0.45	0.59	0.38	0.34
Adjusted R^2	0.34	0.39	0.54	0.35	0.30

**Significant at 1% level, *Significant at 5% level; NS: Non Significant; Figures in parentheses indicate Probability value.

4. Conclusion

The study examined the trends in area, production, and productivity of paddy crop in three selected districts of Andhra Pradesh, namely East Godavari, Krishna, and West Godavari, spanning a 23-year period from 1997 to 2019. The analysis revealed distinct patterns in each district. In East Godavari district, the best-fitted models for area, production, and productivity of paddy were found to be exponential, cubic, and cubic models, respectively, with the lowest MAPE values. Notably, all parameters of these models were statistically significant and met the assumptions of residuals. The results indicated a declining trend in paddy cultivation area but an upward trend in both production and productivity over the study period. In Krishna district, the most suitable models for paddy area, production, and productivity were the exponential, linear, and exponential models, respectively, with low MAPE values. Similarly, all model parameters were statistically significant and satisfied residual assumptions. The findings indicated a decreasing trend in paddy cultivation area but an increasing trend in both production and productivity during the study period. In West Godavari district, the best-fitting models for paddy area, production, and productivity were identified as linear, cubic, and cubic models, with the lowest MAPE values. As in the other districts, all model parameters were statistically significant and adhered to residual assumptions. Results demonstrated a declining trend in paddy cultivation area and a simultaneous increase in both production and productivity over the 23-year period.

References

1. Anonumous, India Stat; c2019a.
2. Anonymous. Area, production and productivity of rice of major cereals in India; c2020. www.indiastat.com.
3. Choudhury N, Saurav S, Kumar RR, Budhlakoti N. Modelling and forecasting of total area, irrigated area, production and productivity of important cereal crops in India towards food security. Int. J Curr. Microbiol. App. Sci. 2017;6(10):2591-2600.
4. Madhu DM, Gangadhar K, Vineeth HT. Exploring the Untapped Potential of Millets in India's Agriculture Industry. Krishi Science: e-Magazine for Agricultural Sciences. 2023;04(05):32-35.
5. Mallikarjuna H. 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; c2009.
6. Mishra P, Sahu PK, Bajpai P, Nirnjan HK. Past trends and future prospects in production, and export scenario of tea in India. Int. Rev. Busi. Finance. 2012;4(1):25-33.
7. Mohan Kumar TL, Gowda CSS, Darshan MB, Rani SS. Coffee production modelling in India using nonlinear statistical growth models. Agriculture Update. 2012;7(1/2):63-67.
8. Pradeep KN, Krishna KG. Study on growth models: A critical Analysis with reference to Andhra Pradesh and India. Indian Agricultural Statistics Research institute; c2002.
9. Rajan MS, Palanivel M, Mohan SK. Forecasting of cotton area, production and productivity using trend analysis. Int. J Res. in Appl. Sci. and Engg. Tech. 2015;3(12):516-520.
10. Sangeetha C, Baskar P. Influence of different crop establishment methods on productivity of rice–a review. Agri. Review. 2015;36(2):113-124.
11. Shah IA. Trend Analysis of Area, Production and Productivity of Apple Fruit in Jammu and Kashmir. Production and Productivity of Apple Fruit in Jammu and Kashmir. J Gujarath res. Society. 2019;21(15):601-611.
12. Sharma Amod Kalita. Trends of area, production and productivity of major fruit crops in Jammu and Kashmir. Agric. Situ. India. 2008;65(7):477-482.
13. Siju T, Kombairaju S. Rice production in Tamil Nadu: A trend and decomposition analysis, Agricu. Situ. India. 2001;58(4):143-145.
14. Sudha CHK, RAO VS, Suresh CH. Growth trends of maize crop in Guntur district of Andhra Pradesh. Int. J Agric. Stat. Sci. 2013;10(2):115-121.