

Original Research Article

<https://doi.org/10.20546/ijcmas.2017.610.388>

## Analysis of Growth and Instability in Rice Production of Odisha by Fitting Appropriate Statistical Models

A. Dash<sup>1\*</sup>, D.S. Dhakre<sup>2</sup> and D. Bhattacharya<sup>2</sup>

<sup>1</sup>College of Agriculture, Chiplima (OUAT) 768025, Odisha, India

<sup>2</sup>Institute of Agriculture, PSB, Visva-bharati, Santiniketan, 731236, West Bengal, India

*\*Corresponding author*

### ABSTRACT

Agriculture in Odisha to a considerable extent means growing rice, the staple food of Odisha. The economic reforms in the 1991-92, affected the agriculture scenario of Odisha following which our study period (1970-71 to 2013-14) has been divided into pre-reform (1970-71 to 1991-92) and post reform (1992-93 to 2013-14) period. Attempts have been taken to make a comparative study of the growth rate and instability of area, production and yield of rice in the two periods. For studying the growth rate and instability, appropriate model which could best possibly describe the behavior of the phenomenon are fitted by applying spline regression technique with kink or knot placed at the year of transition from pre-reform period to post-reform period which is considered to be 1991-92. The possible spline regression models which could fit very well to the area, production and yield of rice are identified from the scatter plot of the data. Then by help of residual diagnostics and model fit statistics, the best fit models are obtained. Using the best fit model, average growth rates of area, production and yield are found. Coefficient of variation is used as a measure of instability. The difference in growth rates from the pre-reform to post-reform period is found to be negative but non-significant in case of area, negative and highly significant in case of production and significantly positive in case of yield. The difference in coefficient of variation between the two periods is highly significant and positive only for yield. For area and production, it is negative and significant only in case of area.

#### Keywords

Spline regression, kink, Economic reforms, Growth rate, Instability, Coefficient of variation.

#### Article Info

*Accepted:*  
26 September 2017  
*Available Online:*  
10 October 2017

### Introduction

Agriculture is the backbone of rural economy and livelihood of Odisha. Rice shares 66 per cent and 11 per cent of the total cropped area in the state, in kharif and rabi season respectively. In kharif season, rice shares nearly 90 per cent and 77 per cent of area under cereals and food grains respectively. In rabi season, rice also shares nearly 90 per cent of area under cereals but 17 per cent of area under food grains (Odisha Agricultural Statistics).

The economic reformation in 1991-92, is expected to have substantial effects in Indian agriculture which in turn affect the state agriculture. Although no direct attempt was made in the reforms process to give a boost to agriculture, it was argued that the new macroeconomic policy framework, more particularly, changes in exchange and trade policy, devaluation of the currency, gradual dismantling of the industrial licensing system and reduction in industrial protection

would benefit tradable agriculture by bringing an end to discrimination against it and making the agricultural terms of trade more favourable. This, in turn, was supposed to promote exports leading to rapid agricultural growth (Bhalla and Singh, 2009).

To describe the behavior of data and to take care of any significant jumps in the data over a long period of time, there is need for fitting of spline regression models. Keeping these perspectives in mind, the study has been made with objective of exploring appropriate model that best fits the area, production and yield of rice in Odisha and making a comparative study of growth rate and instability in pre-reform and post reform period, the growth rate and instability being estimated by using the best fit model.

## **Materials and Methods**

### **Period of study**

The study pertains to area, production and yield of rice in the state of Odisha from the year 1970-71 to 2013-14. The entire study period is divided into two periods – Pre-reform period (1970-71 to 1991-92) and Post-reform period (1992-93 to 2013-14). Pre-reform period is referred to as Period I and Post-reform period is referred to as Period II.

### **Sources of data**

The analysis is based on secondary source data relating to the area, production and yield of rice in Odisha for the period from 1970-71 to 2013-14. The data are obtained from various volumes of Odisha Agricultural Statistics published by the Directorate of Economics and Statistics, Government of Odisha. The area, production and yield are expressed in '000 ha, '000 MT and kg ha<sup>-1</sup>.

respectively.(1 ha = 10000 m<sup>2</sup>, 1 MT = 1000 kg)

## **Analytical techniques**

### **Fitting of appropriate model to the data**

A model is an equation or a set of equations which represents the behavior of the system. Models that are used to describe the behavior of the variables that vary with respect to time are called the growth models.

Fitting a single regression line for a long period to estimate the growth, may be misleading due to significant change (i.e. jumps or breaks) in direction of the growth path, which would make the estimate biased. To avoid this situation, the models are fitted by spline regression technique.

A spline regression model avoids the inappropriate “jump” (i.e., break) in the data for a long period of time by placing a kink in the line at the point of jump without allowing a break in the line.

Spline regression models are fitted with the help of dummy variables. Since the whole period of study (1970-71 to 2013-14) is divided into two periods – Pre-reform period (190-71 to 1991-92) and post-reform period (1992-93 to 2013-14) basing on the economic liberalization in 1991-92, the spline or kinked models have been fitted with one break point (i.e. kink) which is considered at the year 1991-92. The time variable for the year 1970-71 to 2013-14 is referred to as  $t = 1, 2, 3, \dots, 22, 23, 24, \dots, 43, 44$  i.e.,  $t = 1$  for the year 1970-71,  $t = 2$  for the year 1971-72,  $\dots$ ,  $t = 22$  for the year 1991-92,  $\dots$ ,  $t = 44$  for the year 2013-14.

The time variable for the year in which kink is placed (i.e. 1991-92) is referred as  $k$ . Thus  $k = 22$ .

The kinked models are obtained with the help of dummy variables (Paltasingh, R.K., 2013).

The kinked linear model, power model, compound model and quadratic model are obtained as:

**Linear model**

$$Y_t = \beta_0 + \beta_1 \cdot t \cdot I_{(1 \leq t \leq 22)} + \{ \beta_1 \cdot t + A_1 (t - k) \} \cdot I_{(23 \leq t \leq 44)} + \epsilon_t$$

**Power model**

$$Y_t = \beta_0 \cdot t^{\beta_1} \cdot I_{(1 \leq t \leq 22)} \{ t^{\beta_1} \cdot (t - k)^{A_1} \} \cdot I_{(23 \leq t \leq 44)} \cdot \exp(\epsilon_t)$$

The power model obtained is transformed to linear model by natural log transformation as,

$$\ln Y_t = \ln \beta_0 + \beta_1 \cdot \ln t \cdot I_{(1 \leq t \leq 22)} + \{ \beta_1 \cdot \ln t + A_1 \ln (t - k) \} \cdot I_{(23 \leq t \leq 44)} + \epsilon_t$$

**Compound model**

$$Y_t = \beta_0 \cdot \beta_1^t \cdot I_{(1 \leq t \leq 22)} \cdot \{ \beta_1^t \cdot A_1^{(t-k)} \} \cdot I_{(23 \leq t \leq 44)} \cdot \exp(\epsilon_t)$$

The compound model obtained is transformed to linear model by natural log transformation as,

$$\ln Y_t = \ln \beta_0 + t \cdot \ln \beta_1 \cdot I_{(1 \leq t \leq 22)} + \{ t \cdot \ln \beta_1 + (t - k) \ln A_1 \} I_{(23 \leq t \leq 44)} + \epsilon_t$$

**Quadratic model**

$$Y_t = \beta_0 + \{ \beta_1 \cdot t + \beta_2 \cdot t^2 \} I_{(1 \leq t \leq 22)} + \{ \beta_1 \cdot t + A_1 (t - k) + \beta_2 \cdot t^2 + A_2 (t - k)^2 \} I_{(23 \leq t \leq 44)} + \epsilon_t$$

Where  $I_{(A)}$  is the indicator function which is 1 if A holds and 0 else.

Using Ordinary Least Square technique, the estimated values of the coefficients  $\beta_0$ ,  $\beta_1$  and  $A_1$  are found out. The estimated values of  $\beta_0$ ,

$\beta_1$ ,  $\beta_2$ ,  $A_1$  and  $A_2$  are written as  $b_0$ ,  $b_1$ ,  $b_2$  and  $a_1$ ,  $a_2$  respectively.

The significance of the estimated coefficient is tested by applying t test statistic

$t = \frac{\beta_j}{S.E.(\beta_j)}$ , which follows ‘t’ distribution with  $n - p$  degrees of freedom,  $n$  is the number of observations.

The overall significance of the model is tested by applying F statistic

$F = \frac{MSM}{MSE}$ , which follows F distribution with  $(p - 1, n - p)$  degrees of freedom

MSM is the mean square of the model, MSE is the error mean square;

$MSM = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{p - 1}$ ,  $MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - p}$ ;  $n$  is the number of observations and  $p$  is the number of parameters involved in the model.

Assumptions in the model are: (i) Errors should be independent (ii) Error should have zero mean. (iii) Errors should have constant variance i.e. errors should be homoscedastic, (iv) Errors must follow normal distribution

The assumptions regarding the errors are tested by using (i) Durbin-Watson test for testing independence of residuals (ii) Park’s test for testing homoscedasticity of residuals. (iii) Shapiro-Wilk’s test for testing normality of residuals.

**Durbin-Watson test**

This test considers the first order autocorrelation among the residuals. (Montgomery, *et al.*, 2001)

Durbin-Watson test statistic i.e., D-W statistic,  $d = \frac{\sum_{t=2}^n (\epsilon_t - \epsilon_{t-1})^2}{\sum_{t=1}^n \epsilon_t^2}$ , where,  $\epsilon_t =$

$y_t - \hat{y}_t$ ,  $y_t$  and  $\hat{y}_t$  are respectively the actual and estimated values of the response variable in time t.

The value of 'd' ranges from 0 to 4. Upper and lower critical values,  $d_U$  and  $d_L$  have been tabulated for different values of k (no. of explanatory variables) and n (no. of observations) for corresponding level of significance ( $\alpha$ ) in the Durbin – Watson statistical table.

If  $d < d_L$ , it is significant. If  $d > d_U$ , then it is insignificant and the residuals are independent.

If  $d_L < d < d_U$ , test is inconclusive. For testing negative autocorrelation, the statistic  $4 - d$  is used to compare with  $d_U$  and  $d_L$ .

**Park’s test**

In this test, natural logarithm of the residual ( $\epsilon_t$ ) is regressed with natural logarithm of the independent variable (which is time, t) by fitting linear regression, i.e,  $\ln(\epsilon_t) = a + b \ln(t)$ . If the slope of the regression coefficient, b is found to be non-significant, then it is concluded that residuals are homoscedastic (i.e. constant error variance), otherwise, residuals are heteroschedastic (error variance not remaining constant).(Gujarati, D.N., 2003)

**Shapiro-Wilk’s test**

Shapiro-Wilk’s test statistic i.e., S-W test staitistic,  $w = s^2 / b$

Where,  $s^2 = \sum a(k) \{ x_{(n+1-k)} - x_{(k)} \}$  ;  
 $b = \sum_{t=1}^n (y_t - \bar{y})^2$  (Thode Jr., H.C., 2012)

The parameter k takes the values 1,2,...,n/2, when n is even and 1,2,...,(n-1)/2, when n is odd, n is the number of observations.  $x_{(k)}$  is the  $k^{th}$  order statistic of the set of residuals.

The values of coefficients a(k) for different values of n and k are obtained from the table of Shapiro-Wilk. If w is non-significant, then the residuals are normally distributed.

The model fit statistics, viz,  $R^2$ , adjusted  $R^2$  and RMSE (Root mean Square Error) are also computed. Among the models fitted for the dependent variable, which satisfy the error assumptions and show overall significance and significant parameter estimates, the one having highest adjusted  $R^2$  and lowest RMSE is considered to be the best fit model for that variable.

$R^2 = \frac{SSM}{SSE}$ , where, SSM is the sum of square due to model; SSE is the sum of square due to error.

$SSM = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$  ;  $SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$

Adjusted  $R^2 = 1 - (1 - R^2) \times \frac{n-1}{n-p}$  ; RMSE  
 (Root Mean Square Error) =  $\left( \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-p} \right)^{1/2}$

Where ‘n’ is the number of observations; p is the no. of parameters involved in the model.

**Estimation of growth rates**

Using this best fit model, the estimated/predicted values ( $\hat{Y}_t$ ) of the dependent variable (Area/Production/Yield) are found for the period I, period II and the whole period. By using the predicted values, the annual growth rates are found.

Annual Growth Rate for the year t,  $AGR_t = \left( \frac{\hat{Y}_t - \hat{Y}_{t-1}}{\hat{Y}_{t-1}} \right) \times 100$

Average Growth rate for the period I (1970-71 to 1991-92), period II (1992-93 to 2013-14) and the whole period is obtained by taking arithmetic mean of the annual growth

rates of the respective periods. (Prajneshu, 2005)

Also the difference in average growth rates of period I and period II is obtained as,  $\Delta GR = GR_2 - GR_1$

### **Study of instability in area/ production/ yield**

Coefficient of variation is used as a measure of instability. The simple c.v. often contains the trend component and thus overestimates the level of instability in time series data characterized by long term trends. So to eliminate the effect of trend, C.V. is estimated from the detrended values (i.e., trend eliminated values).

For linear and quadratic model, where the effects are assumed to be additive in nature, the detrended values are obtained by subtracting the predicted values (obtained from the model) from the actual values.

Thus, detrended value,  $y_D = y_t - \hat{y}_t$  (assuming additive model)

Where,  $y_t$  is the actual value of the variable in time t.

$\hat{y}_t$  is the predicted value obtained from the best fit model.

For power, compound and logarithmic model, where the effects are assumed to be multiplicative in nature, the detrended values are obtained by dividing the actual values with predicted values (obtained from the model).

Thus, detrended value,  $y_D = \frac{y_t}{\hat{y}_t}$  (assuming multiplicative model)

Where,  $y_t$  is the actual value of the variable in time t.

$\hat{y}_t$  is the predicted value obtained from the best fit model.

The detrended values are then centered by adding the mean of the actual values ( $\bar{y}_t$ )

The C.V. is found from these detrended and centered values.

$CV = \frac{\sigma_{yD}}{\bar{y}_D} \times 100$  C.V. is found for period I, period II and for the whole period. Also the difference in CV ( $\Delta CV$ ) between period I and period II is found.

### **Results and Discussion**

For selecting the appropriate models to be fitted for the variables, such as, area, production and yield of rice in Odisha, the scatter plot is made for the variables against the time.

Fig. 1 shows the different scatter plots for area, production and yield of rice. From the observation of the pattern of dots in the scatter plot, it is found that some linear and non-linear models are found suitable for being fitted to the data. The models like linear, power, compound and quadratic are fitted to each of the three variables, i.e., area, production and yield of rice of odisha.

Table 1 shows the parameter estimates, residual diagnostics and model fit statistics of the fitted models for area under rice. From the table it is found that for all the models fitted for area under rice, the estimate  $b_0$ , is significant at 0.01 level of significance. The estimate  $b_1$  is significant for compound and quadratic model at 0.01 level of significance and for power model at 0.05 level of significance. The estimate  $a_1$  is significant for all models except quadratic model. The estimate  $b_2$  and  $a_2$  for quadratic model is significant at 0.01 level of significance. From the result of the formal test of residual

diagnostics indicated in the table, it is found that for the models fitted for area under rice, D-W statistic is non-significant only for quadratic model which shows that the residuals are independent only for quadratic model. So, quadratic model is the only model which could be considered for selection as best fit model.

The quadratic model also satisfies the test of homoscedasticity and normality of errors as the coefficient of  $\ln(t)$  in the park's test for heteroscedasticity of residuals and the S-W statistic are both non-significant for the quadratic model. Thus, quadratic models fitted for area under rice is considered for selection. From the model fit statistics, F-value is significant for all the fitted models. Quadratic model also has the highest adjusted  $R^2$  and lowest RMSE than other fitted models. So, quadratic model is found to be the best fit model for area under rice in Odisha.

Table 2 shows the parameter estimates, residual diagnostics and model fit statistics of the fitted models for production of rice. From the table it is found that the estimate  $b_0$  is significant for all models at 0.01 level of significance and the estimate  $b_1$  is significant for all fitted models except linear model. The estimate  $a_1$  is significant only for compound and quadratic model.

The coefficient  $b_2$  is significant for quadratic model but the coefficient  $a_2$  for quadratic model is non-significant. From the result of the formal test of residual diagnostics indicated in the table, the D-W statistic and coefficient of  $\ln(t)$  in park's test are both nonsignificant for all fitted models.

This shows that all the models considered satisfy the assumptions of independency and homoscedasticity of errors. But the S-W statistic is found to be non-significant only for power model. This shows that the power

model only satisfies the assumption of normality of residuals, where as, in other fitted models, the residuals are having significant S-W statistic which shows that the residuals for these models do not follow normal distribution. So, power model is the best fit model for production of rice in Odisha despite of having comparatively low  $R^2$  value and high RMSE than other fitted models.

Table 3 shows the parameter estimates, residual diagnostics and model fit statistics of the fitted models for yield of rice. From the table it is found that all the fitted models, have significant estimates of coefficient  $b_0$  at 0.01 level of significance.

The coefficient  $b_1$  is significant for linear and compound models with 0.01 level of significance and significant at 0.05 level of significance for power and quadratic model. The coefficient  $a_1$  is significant at 0.01 level of significance for comound model and at 0.05 level of significance for quadratic model. For linear and power model,  $a_1$  is non-significant.

For quadratic the coefficient  $b_2$  is significant at 0.05 level of significance and the coefficient  $a_2$  is non-significant. From the result of the formal test of residual diagnostics indicated in the table, all the models fitted for yield of rice satisfy the assumptions of independence and homoscedasticity of errors as the D-W statistic and the coefficient of  $\ln(t)$  in Park's test are found non-significant for all the models.

But, except for quadratic model, all other fitted models have significant S-W statistic which shows that the assumption of normality of errors is satisfied only by quadratic model and for the remaining models the residuals do not follow normal distribution. Thus for yield of rice, the quadratic model is considered for selection.

**Table.1** Parameter estimates, residual diagnostics and model fit statistics of the fitted models for area under rice in Odisha

Fitted Models →		Linear	Power	Compound	Quadratic
Parameter estimates	b <sub>0</sub>	4506.01** (76.71)	4614.91** (103.3)	4501.29**(76.6)	<b>4756.57** (101.17)</b>
	b <sub>1</sub>	--11.36 (5.84)	-0.025*(0.01)	0.997**(0.013)	<b>-74.01** (20.26)</b>
	a <sub>1</sub>	23.36* (9.79)	0.037* (0.016)	1.005**(0.02)	<b>-24.73 (43.51)</b>
	b <sub>2</sub>	-	-	-	<b>2.72** (0.86)</b>
	a <sub>2</sub>	-	-	-	<b>-1.71** (0.57)</b>
Residual diagnostics	D-W Statistic	0.66**	0.936*	0.932*	<b>1.978</b>
	Coefficient of ln(t) in Park's test	0.677	0.85**	0.66	<b>0.93</b>
	S-W Statistic	0.933*	0.707	0.695	<b>-0.36</b>
Model fit statistics	R <sup>2</sup>	0.336	0.046	0.34	<b>0.563</b>
	Adjusted R <sup>2</sup>	0.305	0.022	0.304	<b>0.542</b>
	RMSE	203.62	180.15	203.52	<b>117.8</b>
	F Value	9.59**	9.9**	9.79**	<b>12.24**</b>

(Figures in the parentheses indicates the standard errors)

\* Significant at 0.05 level of significance \*\* Significant at 0.05 level of significance

**Table.2** Parameter estimates, residual diagnostics and model fit statistics of the fitted models for production of rice in Odisha

Fitted Models →		Linear	Power	Compound	Quadratic
Parameter estimates	b <sub>0</sub>	3296.2**	<b>3338.03**</b>	3406.36** (352.38)	4334.28** (495.48)
	b <sub>1</sub>	93.44	<b>0.15** (0.056)</b>	1.019** (0.007)	-166.08* (79.24)
	a <sub>1</sub>	-7.94	<b>0.013 (0.08)</b>	0.995** (0.009)	-483.15* (223.4)
	b <sub>2</sub>	-	-	-	11.28* (4.19)
	a <sub>2</sub>	-	-	-	0.02 (7.67)
Residual diagnostics	D-W Statistic	2.34	<b>1.96</b>	2.37	2.13
	Coefficient of ln(t) in Park's test	-0.251	<b>0.666</b>	0.175	0.733
	S-W Statistic	0.944*	<b>0.956</b>	0.947*	0.916*
Model fit statistics	R <sup>2</sup>	0.607	<b>0.532</b>	0.613	0.679
	Adjusted R <sup>2</sup>	0.598	<b>0.521</b>	0.600	0.663
	RMSE	972.90	<b>1061.38</b>	965.80	879.22
	F Value	63.33**	<b>47.79**</b>	66.45**	20.1**

(Figures in the parentheses indicates the standard errors)

\* Significant at 0.05 level of significance \*\* Significant at 0.05 level of significance

**Table.3** Parameter estimates, residual diagnostics and model fit statistics of the fitted models for yield of rice in Odisha

Fitted Models →		Linear	Power	Compound	Quadratic
Parameter estimates	b <sub>0</sub>	728.36** (69.43)	723.01** (113.87)	756.48*(68.01)	<b>911.79** (99.65)</b>
	b <sub>1</sub>	23.38** (5.29)	0.136* (0.05)	1.02** (0.01)	<b>-22.48*(9.96)</b>
	a <sub>1</sub>	-5.59 (10.72)	-0.023 (0.08)	0.989** (0.01)	<b>-107.88* (52.75)</b>
	b <sub>2</sub>	-	-	-	<b>1.99* (0.84)</b>
	a <sub>2</sub>	-	-	-	<b>1.22(1.65)</b>
Residual diagnostics	D-W Statistic	2.12	1.77	2.1	<b>2.02</b>
	Coefficient of ln(t) in Park's test	0.346	0.566	0.441	<b>0.628</b>
	S-W Statistic	0.906**	0.934**	0.901**	<b>0.966</b>
Model fit statistics	R <sup>2</sup>	0.627	0.553	0.625	<b>0.732</b>
	Adjusted R <sup>2</sup>	0.618	0.543	0.620	<b>0.719</b>
	RMSE	222.82	243.78	223.42	<b>188.7</b>
	F Value	68.86**	52.03**	69.94**	<b>25.99**</b>

(Figures in the parentheses indicates the standard errors)

\* Significant at 0.05 level of significance \*\* Significant at 0.05 level of significance

**Table.4** Average growth rates of coefficient of variation of area, production and yield of rice of Odisha in period I, period II and the whole period

(In per cent)

Variables	Average Growth Rate				Coefficient of Variation			
	Period I (GR <sub>1</sub> )	Period II (GR <sub>2</sub> )	Whole Period (GR)	$\Delta GR = GR_2 - GR_1$	Period I (CV <sub>1</sub> )	Period II (CV <sub>2</sub> )	Whole Period (CV)	$\Delta CV = CV_2 - CV_1$
Area	-0.25 (0.16)	-0.43** (0.11)	-0.32** (0.1)	-0.18 (0.27)	3.06** (0.46)	2.04** (0.31)	2.59** (0.28)	-1.02* (0.39)
Production	1.87** (0.43)	1.67** (0.39)	1.73** (0.29)	-0.20** (0.05)	19.15** (2.99)	18.21** (2.83)	18.69** (1.92)	-0.93 (2.98)
Yield	2.13** (0.48)	1.95** (0.53)	1.93** (0.37)	-0.18** (0.07)	13.22** (2.03)	22.19** (3.51)	14.91** (1.55)	8.97** (2.82)

(Figures in the parentheses represents the standard errors)

\* Significant at 5 % level; \*\* Significant at 1 % level

**Fig.1** Scatter diagrams of area, production and yield of rice in Odisha

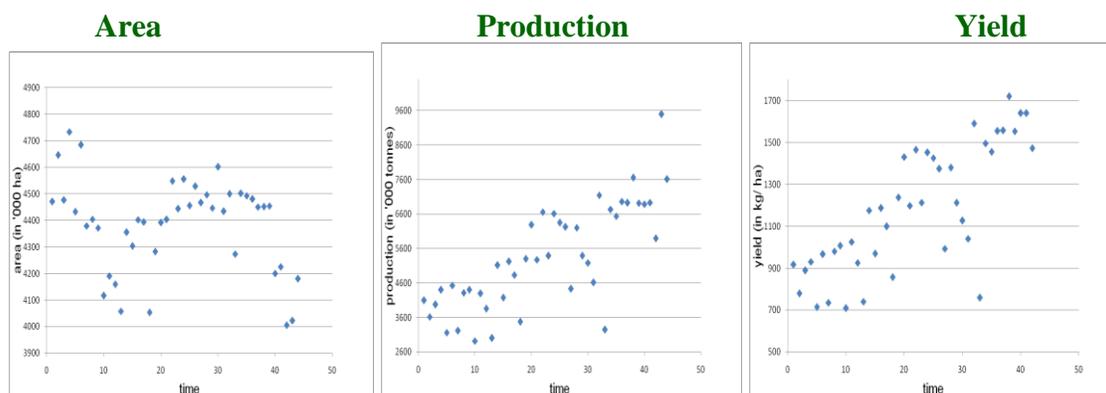


Table 4 shows the average growth rates and coefficient of variation of area, production and yield of rice in period I, period II and whole period. From deep perusal of table 4 regarding the average growth rates, it is found that the average growth rates of area under rice remained negative but non-significant in period I. In period II, the average growth rate becomes more negative and significant at 0.01 level of significance.

Thus, in the whole period of study the average growth rate remained negative and significant at 0.01 level of significance. The difference in average growth rates of area from period I to period II is negative but non-significant. The average growth rate of production and yield of rice remained positive and significant at 0.01 level of significance in both the periods

and in the whole period. The difference in the average growth rate from period I to period II is negative and significant at 0.01 level of significance for production and yield of rice.

From the scrutiny of table 4 regarding the coefficient of variation, it is found that the coefficient of variation is significant at 0.01 level of significance for all the variables under study i.e., area, production and yield in period I, period II and the whole period.

The difference in coefficient of variation from period I to period II is positive and significant at 0.01 level of significance only in case of yield of rice. For area, coefficient of variation is negative and significant at 0.05 level of significance, whereas, for production it is negative and non-significant.

The above discussion highlighted that though some models perform better in terms of model fit statistics, such as, high value of adjusted  $R^2$  and low value of RMSE, they are not worthy of being selected as the best model if they do not satisfy the error assumptions. This is the case with the statistical modelling in case of production of rice. Though the linear model, compound model and quadratic model fitted to the production of rice data, have higher values of adjusted  $R^2$  than the power model fitted to the same data, but these models show lack of normality of errors as indicated by their significant S-W statistic. So, for fitting a model, first of all the estimated parameters and the overall significance of the model should be checked. Then the error assumptions must be tested both formally and graphically by using appropriate techniques. After this the models that satisfies the error assumptions and having significant parameter estimates along with overall significance, are compared on basis of the model fit statistics and finally the best fit model that describes the data as best as possible, is selected. The spline regression technique also takes care of any abrupt jumps in the data.

From the study of average growth rate found with the help of best fit model, it can be summarised that though area under rice registers a negative average growth rate in both the periods which is significant only in post-reform period. It shows a declining trend and the difference in the average growth rate from pre-reform period to post reform period is significantly negative. This decline in area may be assigned to the rapid growth of industrialisation and shift of agriculture from rice to commercial crops in the post reform period. The average growth rate in case of production shows an increasing trend in both the periods, but the rate of increase is less in the post reform period. This might be due to increasing trend in yield of rice in both the

periods but the increase in average growth rate in yield of rice is at slower rate in post – reform period as compared to pre-reform period. This shows that attempts have been taken to improve the yield of rice in both the periods which was less successful to some extent in post-reform period.

From instability point of view, there is increase in instability of area, production and yield of rice in post-reform period as compared to pre-reform period which is significant only in case of yield. In case of area there is decrease in instability in post-reform period as compared to pre-reform period. In case of production of rice, there is decrease in instability in the post –reform period but the decrease is non-significant. Thus from the study it is revealed that the area under showing a low negative growth rate is almost stabilised. The high growth rate combined with high instability in case of yield of rice depicts that there is rigorous attempts to increase the yield by (i) applying developed technologies, viz., use of improved varieties, (ii) by ensuring adequate and timely operation of various cultural practices (iii) developing strategies to deal with the natural calamities, but some failures to maintain these developed activities uniformly throughout the period. This shows that high growth rate is more prone to high instability due to the fact that the activities that leads to increase in yield, may not be carried out with equal potency in all the years throughout the period. Due to the increase in yield of rice, the growth in production of rice stays at the positive end and also shows nearly same degree of instability in both the periods.

## References

- Bhalla, G.S., Singh, G., 2009. Economic liberalisation and Indian agriculture. A state-wise analysis. *Economic and Political Weekly* 54(52), 34-44.

- Gujarati, D.N., 2003. Basic Econometrics, Fourth Edition, Published by McGraw-Hill/Irwin
- Koopmans, L. H., Owen, D. B., Rosenblatt, J. I., 1964. Confidence intervals for the coefficient of variation for the normal and lognormal distributions. *Biometrika* 51, 25-32.
- Montgomery, D. C., Peck, E. A., Vining, G. G. 2001. Introduction to Linear Regression Analysis. 3rd Edition. New York. John Wiley & Sons.
- Paltasingh, R.K., Goyari, P. 2013. Agricultural Economics Research Review. Vol. 26 (Conference number) 67-78
- Prajneshu, Chandan, K.P., 2005. Computation of compound growth rates in agriculture, Revisited. *Agricultural Economics Research Review*. 18: July-December, 317-324
- Thode Jr., H.C., 2012. Testing for Normality. Marcel Dekker, Inc., New York.

**How to cite this article:**

Dash, A., D.S. Dhakre and Bhattacharya, D. 2017. Analysis of Growth and Instability in Rice Production of Odisha by Fitting Appropriate Statistical Models. *Int.J.Curr.Microbiol.App.Sci*. 6(10): 3313-3322. doi: <https://doi.org/10.20546/ijemas.2017.610.388>