International Journal of Plant & Soil Science

33(23): 178-188, 2021; Article no.IJPSS.76521
ISSN: 2320-7035

Growth Rate Estimation of Rabi Pulse Production of Odisha by Using Spline Regression Technique

Rakesh Kumar Rout [a#] and Abhiram Dash [a*†]

[a]Department of Agricultural Statistics, College of Agriculture, OUAT, Bhubaneswar, India.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/IJPSS/2021/v33i2330733
Editor(s):
(1) Prof. Hakan Sevik, Kastamonu University, Turkey.
Reviewers:
(1) Myo Than Tun, Seoul National University, Myanmar.
(2) Mgbasonwu Vincent Nwabinye, Michael Okpara University of Agriculture, Nigeria.
(3) Champa Lal Khatik, SKNAU, India.
Complete Peer review History: https://www.sdiarticle4.com/review-history/76521

Original Research Article

Received 06 September 2021
Accepted 12 November 2021
Published 17 November 2021

ABSTRACT

Pulses are considered to be important crop for ensuring nutritional security in Odisha. Proper estimation of growth rate in production of pulse crops allows for more effective cropping system planning and formulation of the agricultural policy of the state. To capture any abrupt changes and the variation in data in different phases of a long time period, spline regression technique is used as it can fit different models in different segments of the time period as necessary without losing the continuity of the model. The present study deals with the estimation of growth rate of area, yield and production of all rabi pulses in Odisha by using best fit spline regression model. To fit the spline regression model, the entire period of study is divided into different segments based on the scatter plot diagram which is further confirmed by testing the significance of change in coefficient of variation between the consecutive segments by chi square test. The regression model found to be suitable from the study of scatter plot of data are linear, compound, logarithmic, power, quadratic and cubic model. The best fit model is selected on the basis of error assumption test and model fit statistics such as R^2, adjusted R^2 and Mean Absolute Percentage error (MAPE). The respective selected best fit model is used for the estimation of growth rates of area, yield and production of rabi pulses in Odisha for each segment and the whole period of study. Among the spline regression models, the respective linear spline regression model is found to be best fit for area, yield and

#P.G. Scholar,
†Assistant Professor,
**Corresponding author: E-mail: abhidash2stat@gmail.com;*

production of rabi pulses and are used for growth rate estimation of these variables. It is found that though the growth rate in area and yield of rabi pulses are not significant, the growth rate of production is found to be significant for the whole period of study which shows that the interaction effect of area and yield on production seems to dominate.

Keywords: Rabi pulse; Growth rate; cropping system; spline regression; yield and production.

1. INTRODUCTION

Pulse crops comes under important crop groups in Odisha for ensuring nutritional security. Rabi pulses account for 69% of area and 64% of production, with a productivity of 510 kg/ha in the state. Estimation of growth rate in production of pulse crops is of utmost importance which is necessary for more effective planning and formulation of the agricultural policy of the state.

To capture the variation in data in different phases of a long time period, spline regression can be used as it can fit different models in different segments of the time period as necessary without losing the continuity of the model.

2. MATERIAL AND METHODS

The data on area, yield and production of rabi pulses are collected for the period 1970-71 to 2019-20 from Odisha Agricultural Statistics published by Directorate of Agriculture and Food Production [1], Odisha. The partitioning of data is initially based on scatter plot technique and then confirmed by chi-square test of significance of change in Coefficient of Variation (C.V.) in consecutive segments, in the following manner:

Let $\overline{x_i}$ and s_i be the mean and standard deviation of the i^{th} time period,
$m_i = n_i - 1$, where n_i is the no. of data points (i.e., no. of years) in the i^{th} period

Null hypothesis, H_o: No difference in C.V. of consecutive segments.

Alternative hypothesis, H_1: There is difference in C.V. of consecutive segments.

The chi-square test statistic to test the significance of difference between the C.V. of the consecutive segments is given by,

$$\chi^2 = C^{-2}(0.5+C^2)^{-1}\left[\sum_{i=1}^{2} m_i\left(\frac{s_i}{x_i}\right)^2 - M(C)^2\right] \quad [2]$$

Where, $M = \sum_{i=1}^{2} m_i$; $C = \sum_{i=1}^{2} m_i \frac{s_i}{x_i}$; If χ^2 is significant i.e., p-value of $\chi^2 \le 0.05$, then H_0 is rejected and it is concluded that there is significant difference in C.V. of the two time periods.

The regression models fitted with spline regression technique with two knots placed at time period, k_1 and k_2 in the following manner:

Linear spline model:

$$Y_t = \beta_0 + \beta_1 . t. I_{(1\le t\le k_1)} + \{\beta_1 . t + A_{11} (t - k_1)\}. I_{(k_1+1\le t\le k_2)} + \{\beta_1 . t + A_{11} t + A_{12} (t - k_2)\}. I_{(k_2+1\le t\le n)} + \varepsilon_t$$

Compound spline model:

$$Y_t = \beta_0 . \beta_1^{\,t} . I_{(1\le t\le k_1)} . \{\beta_1^{\,t} . A_{11}^{(t-k_1)}\}. I_{(k_1+1\le t\le k_2)} . \{\beta_1^{\,t} . A_{11}^{\,t} . A_{12}^{(t-k_2)}\}. I_{(k_2+1\le t\le n)} \exp(\varepsilon_t)$$

The compound spline model transformed to linear form by a natural log transformation is written as,

$$\ln(Y_t) = \ln \beta_0 + t . \ln(\beta_1) . I_{(1\le t\le k_1)} + \{t .\ln(\beta_1) + (t - k_1) . \ln(A_{11})\} I_{(k_1+1\le t\le k_2)} + \{t . \ln(\beta_1) + t . \ln(A_{11}) + (t - k_2) . \ln(A_{12})\} I_{(k_2+1\le t\le n)} + \varepsilon_t$$

Logarithmic spline model:

$$Y_t = \beta_0 + \beta_1.\ln(t). I_{(1\le t\le k_1)} + \{\beta_1.\ln(t)+A_{11} . \ln(t - k_1)\}. I_{(k_1+1\le t\le k_2)} + \{\beta_1 . \ln(t) + A_{11} . \ln(t) + A_{12} \ln(t - k_2)\} . I_{(k_2+1\le t\le n)} + \varepsilon_t$$

Power spline model:

$$Y_t = \beta_0 \cdot t^{\beta_1} \cdot I_{(1\le t\le k_1)} \{ t^{\beta_1} \cdot (t-k_1)^{A_{11}} \}. I_{(k_1+1\le t\le k_2)} \cdot \{ t^{\beta_1} \cdot t^{A_{11}} (t-k_2)^{A_{12}} \}. I_{(k_2+1\le t\le n)} \exp(\varepsilon_t)$$

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

$$\ln(Y_t) = \ln \beta_0 + \beta_1. \ln(t) . I_{(1\le t\le k_1)} + \{ \beta_1.\ln(t) + A_{11} \ln(t - k_1)\} . I_{(k_1+1\le t\le k_2)} + \{\beta_1.\ln(t) + A_{11}.\ln(t) + A_{12} \ln(t - k_2)\} . I_{(k_2+1\le t\le n)} + \varepsilon_t ,$$

Quadratic spline model:

$$Y_t = \beta_0 + (\beta_1 . t.+ \beta_2 . t^2) I_{(1\le t\le k_1)} +\{\beta_1 . t +A_{11} (t - k_1) + \beta_2 . t^2 + A_{21} (t - k_1)^2\}. I_{(k_1+1\le t\le k_2)} +\{\beta_1 . t + A_{11} t + \beta_2 . t^2 + A_{21} t^2 + A_{12} (t - k_2) + A_{22} (t - k_2)^2\}. I_{(k_2+1\le t\le n)} + \varepsilon_t$$

Cubic spline model:

$$Y_t = \beta_0 + (\beta_1 . t.+ \beta_2 . t^2 + \beta_3 . t^3) I_{(1\le t\le k_1)} +\{\beta_1 . t +A_{11} (t - k_1) + \beta_2 . t^2 + A_{21} (t - k_1)^2 + \beta_3 . t^3 + A_{31}(t - k_1)^3\}. I_{(k_1+1\le t\le k_2)} +\{\beta_1 . t + A_{11} t + \beta_2 . t^2 + A_{21} t^2 + \beta_3 . t^3 + A_{31} t^3 + A_{12} (t - k_2) + A_{22} (t - k_2)^2 +A_{32} (t - k_2)^3\}. I_{(k_2+1\le t\le n)} + \varepsilon_t$$

where, $I_{(P)}$ is the indicator function which is 1 if P holds and 0 otherwise.

In case of three knots the models are fitted in the similar manner with additional coefficients β_3, A_{31} and A_{32}.

In all the cases the parameters of the model are estimated by using Ordinary Least Square technique. The estimated values of β_0, β_1, A_{11}, A_{12}, β_2, A_{21}, A_{22}, β_3, A_{31} and A_{32} are written as b_0, b_1, a_{11}, a_{12}, b_2, a_{21}, a_{22}, b_3, a_{31} and a_{32} respectively.

The above models are fitted under the assumptions that errors are independently distributed, follow normal distribution and have constant variance i.e. heteroscedastic [3].

The following statistical tests are considered for testing the assumptions regarding errors in the model:

(i) Durbin-Watson test: This test considers the first order autocorrelation among the residuals [4].

Null hypothesis is taken as, H_0: the errors are independent.

And the alternative hypothesis as, H_1: the errors are not independent.

Durbin-Watson test statistic (DW statistic) is given by,

$$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2},$$

where, $e_t = y_t - \hat{y}_t$, y_t and $\hat{y}_t$ are respectively the actual and estimated values of the response variable at time t and *n* is the no. of observations.

If p-value of the test statistic 'd' is greater than 0.05, then the independency of errors can be assumed.

(ii) Shapiro-Wilk's test:

This test is used for testing normality of the residuals.

Null hypothesis here is H_0: the errors follow normal distribution.

Alternative hypothesis, H_1: The errors do not follow normal distribution.

To carry out the test, the data pertaining to errors are arranged in ascending order so that $e_{(1)} \le e_{(2)} \le \ldots \le e_{(n)}$

The Shapiro-Wilk's (SW) test statistic as given by,

$$SW = \frac{s^2}{b},$$

where, $s^2 = \sum_{k=1}^{m} a(k)\{e_{(n+1-k)} - e_{(k)}\}^2$; $b = \sum_{t=1}^{n} (e_t - \bar{e})^2$ [5]

If n is even, then m = $\frac{n}{2}$. If n is odd, then m = $\frac{n-1}{2}$;The parameter k takes the values 1,2,…, m.

n is the number of observations, $e_{(k)}$ is the k^{th} order statistic in the set of residuals,

e_t is the residual at time 't' and $\bar{e}$ is the mean of e_t .

If the SW statistic has a p-value below 0.05, then the null hypothesis is rejected and the residuals do not follow normal distribution and the fitted regression model used cannot be appropriate.

(iii) Breusch-Pagan test:

The homoscedasticity of errors obtained from the regression model can be tested by using Breusch-Pagan test [6]

Null hypothesis, H_0: Errors have constant variance i.e. homoscedastic.

Alternative hypothesis, H_0: Errors have non-constant variance i.e. heteroscedastic

Breusch-Pagan test statistic is given as,

BP = n x R^2; Where, n is the no. of observations, R^2 is the coefficient of determination of the regression of squared residuals (obtained from the original regression) on the independent variable (which is time, in the present study)

BP statistic follows chi-square distribution with 'k' degrees of freedom.

If the BP statistic has a p-value below 0.05, then the null hypothesis is rejected and heteroscedasticity is assumed to be present in the residuals and the regression model used can be considered to be inappropriate fit.

Among the fitted models having overall significance and satisfying the model diagnostics tests, the one having highest R^2, highest adjusted R^2 and lowest Mean Absolute Percent Error (MAPE) is considered to be the best fit model for that dependent variable.

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

SSR = $\sum_{t=1}^{n}\left(\hat{y}_t - \bar{y}\right)^2$, SSE = $\sum_{t=1}^{n}\left(y_t - \hat{y}_t\right)^2$;

Where, y_t and $\hat{y}_t$ are respectively the actual and estimated values of the response variable at time t, and $\bar{y}$ is the mean of y_t.

Adjusted R^2 is defined as Adjusted R^2 = 1 – (1-R^2) x $\frac{(n-1)}{(n-p)}$

where, p is the no. of parameters involved in the model.

Adjusted R^2 penalizes the model for adding some independent variables which are not necessary to fit the data and thus adjusted R^2 will not necessarily increase with the increase in the number of independent variables included in the model.

Significance of R^2 and Adjusted R^2 can be tested by calculating the F value given by:

$F = \frac{R^2/(p-1)}{(1-R^2)/(n-p)}$,

Where, p = number of coefficients involved in the model

n = total number of observations

If calculated F > critical $F_{\alpha,(p-1,n-p)}$ then R^2 is considered to be significant, otherwise it is insignificant.

RMSE = $\left\{\frac{\sum_{t=1}^{n}\left(y_t - \hat{y}_t\right)^2}{(n-p)}\right\}^{1/2}$;

$$\text{MAPE} = \left(\sum_{i=1}^{n}\frac{|P_i - O_i|}{O_i}\times 100\right)/n$$

, where P_i and O_i are respectively the predicted and observed values for the i^{th} year, i= 1, 2…, n.

2.1 Study of Average Growth Rates for Area, Production and Yield

Using the model which best fit the data, the estimated values, $\hat{y}_t$ of the dependent variable (Area/Production/Yield as the case may be) are found for the each segment of the time period By using the predicted values, the annual growth rates are found as

Annual Growth Rate for the year t, $AGR_t = \frac{\hat{y}_t - \hat{y}_{t-1}}{\hat{y}_{t-1}} \times 100$,

where $\hat{y}_t$ and $\hat{y}_{t-1}$ are the predicted values of the variable y at time t and (t-1) respectively.

Average Growth rates for the segment I, segment II and the whole period of study are obtained by taking arithmetic mean of the annual growth rates of the respective periods [7]

Thus, Average Growth Rate for the period k is obtained as, $GR_k = \frac{\sum AGR_t}{x}$,

Where x is the number of years in k^{th} period

Average Growth Rate for the whole period of study is obtained as,

$$GR = \frac{\sum_2^{5\,0} AGR_t}{49} \quad [8]$$

Also, the difference in average growth rates of two consecutive segments is obtained as,

$\Delta GR = GR_2 - GR_1$

The significance of the average growth rate for the different segments and for the whole period in the population is tested by using student's t-statistic.

The null hypothesis for the test is taken as H_0: Population GR = 0 which is tested against the alternative hypothesis H_1: Population GR ≠ 0.

The test statistic is $t = \frac{GR}{SE(GR)}$, which follows a t distribution with (n – 1) d.f., where n is the no. of observations.

The significance of the difference of average growth rate in the population is tested by using student's t-statistic. The null hypothesis for the test is taken as $H_0 : \Delta GR = 0$ which is tested against the alternative hypothesis $H_1 : \Delta GR \neq 0$.

The appropriate test statistic is $t = \frac{\Delta GR}{SE(\Delta GR)}$, which follows 't' distribution with $(n_1 + n_2 - 2)$ d.f., where n_1 and n_2 are the number of observations in segment I and segment II respectively.

$SE(GR) = \frac{\sigma_{AGR}}{\sqrt{n}}$, where AGR – Annual Growth Rates, n = no. of data points

$SE(\Delta GR) = \sqrt{\frac{\sigma^2_{AGR_i} + \sigma^2_{AGR_j}}{n_i + n_j - 2}}$, where AGR_i = Annual growth rates in i^{th} segment,
AGR_j = Annual Growth Rates in j^{th} segment,

n_i and n_j are the no. of data points in the i^{th} and j^{th} segment of the study period, respectively

3. RESULTS AND DISCUSSION

The study of scatter plots of data on area, yield, and production of pulses in Odisha for rabi season helps in the identification of regression models that could fit the data effectively. Figs. 1, 2 and 3 show the scatter plot of area, yield and production of rabi pulses respectively for the period 1970-71 to 2019-20. The regression model found to be suitable from the scatter plot are linear, compound, logarithmic, power, quadratic and cubic model. The scatter plot also gives an idea to partition the whole period of study into different segments in such a manner that there is less variation within the segment and sudden jump between the segments. This partitioning is further confirmed by chi square test of the difference in coefficient of variation between the consecutive segments.

In case of data on area under rabi pulses, from the scatter plot in Fig. 1 the knots can be placed at the year 1984-85 (first knot), 1995-96 (second knot) and 2004-05 (third knot) which correspond to the time period t = 15, 26 and 35 respectively. In case of data on yield of rabi pulses, from the scatter plot in Fig. 2 the knots can be placed at the year 1984-85 (first knot) and 1995-96 (second knot) which correspond to the time period t = 15 and 26 respectively. As evident from the scatter plot in Fig. 3, for data on production of rabi pulses the knots can be placed at the year 1984-85 (first knot) and 1995-96 (second knot) which correspond to the time period t = 15 and 26 respectively. Thus, the whole period of study is divided into three segments (segment I, II, III) for production and yield of rabi pulses in Odisha and into four segments (segment I, II, III and IV) for area under rabi pulses in Odisha which are represented in Figs. 1, 2 and 3 for area, yield and production of rabi pulses in Odisha, respectively.

The partitioning of data into segments can be further ascertained by calculating χ^2 values for CV of each segment and testing whether the difference in CV of consecutive segments is significant or not. Table 1 shows that partitioning of data on area, yield and production of rabi pulses based on the testing of Coefficient of Variation. The difference in CV of consecutive segments is found significant for area, yield and production of rabi pulses with p-value less than 0.05.

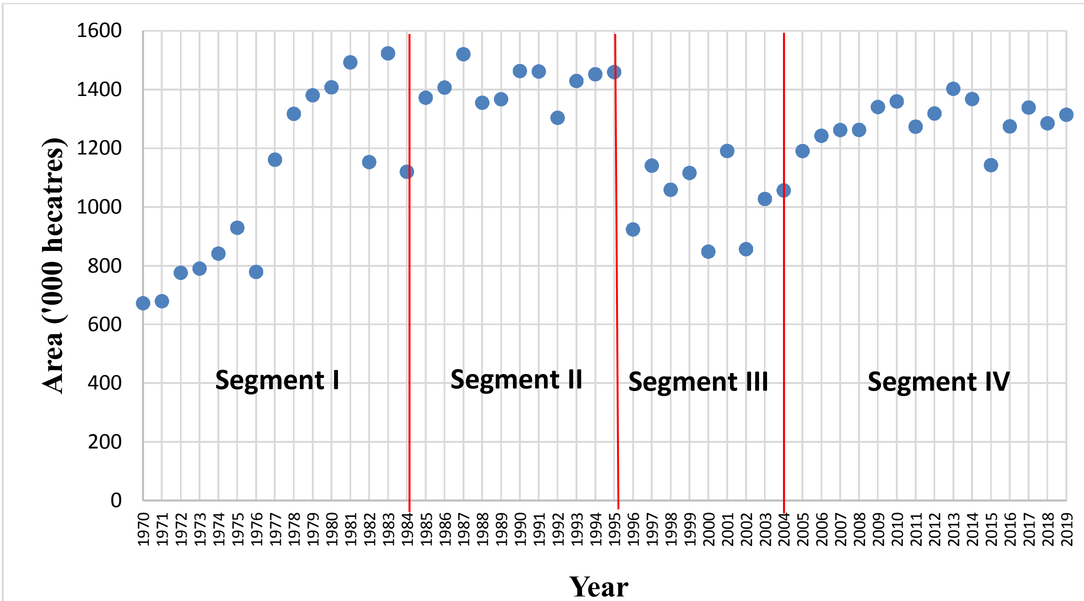

Fig. 1. Scatter plot of area under rabi pulses in Odisha

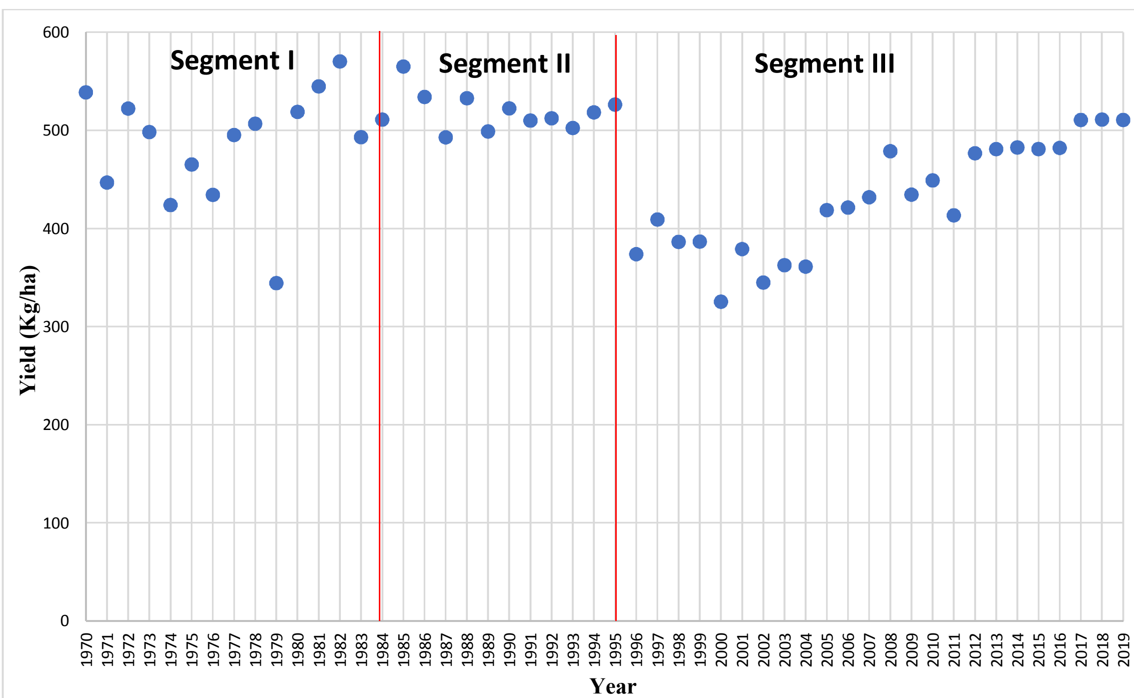

Fig. 2. Scatter plot of yield of rabi pulses in Odisha

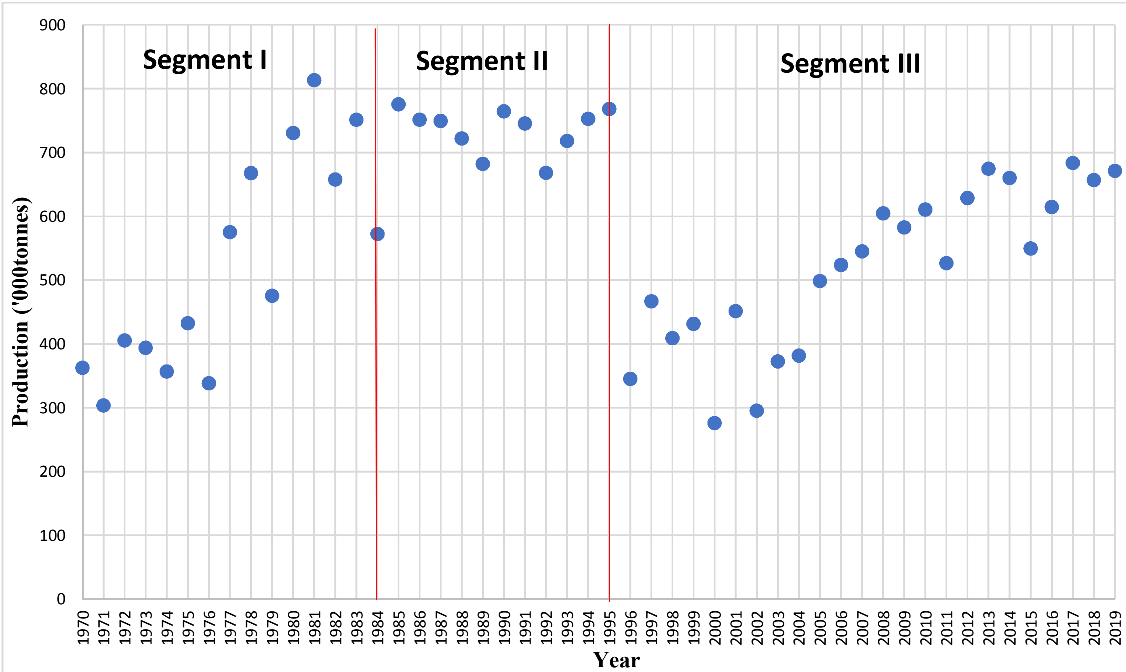

Fig. 3. Scatter plot of production of rabi pulses in Odisha

Table 1. Partitioning of data on area, yield and production of rabi pulses based on the testing of C.V.

Variable	Segments	Standard deviation	Mean	C.V.	D'AD		
					I - II	II - III	III – IV
Area	**I**	305.28	1068.7	28.56	18.603***	8.352**	5.555*
	II	62.58	1417.6	4.41	(<0.001)	(0.0038)	(0.018)
	III	123.10	1024.8	12.01			
	IV	73.70	1290.9	5.71			
Yield	**I**	57.07	487.82	11.70	9.849**	10.4546**	
	II	20.20	519.82	3.89	(0.002)	(0.002)	
	III	52.87	422.53	12.51			
Production	**I**	169.091	522.82	32.34	18.811***	14.773***	
	II	34.967	736.55	4.747	(<0.001)	(<0.001)	
	III	122.889	506.432	24.265			

The figures in the parentheses represents the p-value
**** p-value ≤ 0.001 ** 0.001 < p-value ≤ 0.01 * 0.01 < p-value ≤ 0.05*

3.1 Fitting of the Selected Regression Models

In case of area under rabi pulses, Table 2 shows that only linear spline model satisfies all error assumptions as all the test statistic in DW Statistic, SW Statistic and BP Statistic used for the purpose are insignificant and has moderately high and significant R^2 and adjusted R^2 with low value of MAPE than other models. Thus, linear spline model is the selected best fit model for data on area under rabi pulses in Odisha.

In case of yield of pulses in rabi season, Table 3 shows that only linear spline model satisfies all error assumptions in DW Statistic, SW Statistic and BP Statistic and has moderately high and significant R^2 and adjusted R^2 with low value of MAPE than the other models. Thus, linear spline model is the selected best fit model for data on yield of pulses in rabi in Odisha.

Table 2. Estimated parametric coefficients, model diagnostics measures and model fit statistics of spline regression models fitted to data on area under rabi pulses in Odisha

Model →	Linear Spline	Compound Spline	Logarithmic Spline	Power Spline	Quadratic Spline	Cubic Spline
Estimated Parametric Coefficients						
b_0	681.9***	60.85***	52.85	67.5***	62.18*	129.2***
b_1	46.92***	1.093***	32.33	1.3**	5.56	-26.85
a_{11}	-37.28***	1.005	110.23***	1.576***	17.57	66.38***
a_{12}	6.74*	0.952***	-92.50***	0.689***	-7.55	-39.95
a_{13}	-11.912	0.988	41.21**	1.132*	27.79	-1.09
b_2	-	-	-	-	0.25	3.74
a_{21}	-	-	-	-	-0.53	-12.28***
a_{22}	-	-	-	-	-0.18	-32.43
a_{23}	-	-	-	-	0.082	-79.35
b_3	-	-	-	-	-	-0.10
a_{31}	-	-	-	-	-	0.49**
a_{32}	-	-	-	-	-	1.19
a_{33}						-1.67
Model Diagnostics Criteria						
DW Statistic	1.882	0.876***	0.58***	0.769***	0.917***	1.618**
SW Statistic	0.969	0.976	0.975	0.976	0.957	0.930**
BP Statistic	6.38	9.431*	2.633	6.819	8.599	8.514
Model Fit Statistics						
R^2	0.751***	0.885***	0.822***	0.85***	0.917***	0.969***
Adjusted R^2	0.729***	0.875***	0.806***	0.836***	0.9***	0.959***
MAPE	7.902	15.256	17.947	15.256	12.812	6.904

**** p-value ≤ 0.001 ** 0.001 < p-value ≤ 0.01 * 0.01 < p-value ≤ 0.05*

Table 3. Estimated parametric coefficients, model diagnostics measures and model fit statistics of spline regression models fitted to data on yield of rabi pulses in Odisha

Model →	Linear Spline	Compound Spline	Logarithmic Spline	Power Spline	Quadratic Spline	Cubic Spline
Estimated Parametric Coefficients						
b_0	446.87***	442.94***	441.65***	438.91***	506.66***	545.41***
b_1	5.74***	1.01***	31.71	1.07	-11.57	-27.98
a_{11}	-9.54***	0.97***	-38.31**	0.91**	-32.99*	-32.69
a_{12}	10.46***	1.02***	10.48	1.026	-31.19	-25.55
b_2	-	-	-	-	0.93*	2.53
a_{21}	-	-	-	-	0.56	1.09
a_{22}	-	-	-	-	-1.31	2.75
b_3	-	-	-	-	-	-0.04
a_{31}	-	-	-	-	-	-0.01
a_{32}	-	-	-	-	-	0.01
Model Diagnostics Criteria						
DW Statistic	1.788	1.767	0.977***	0.966***	2.163	2.411
SW Statistic	0.891	0.888***	0.953*	0.888***	0.910**	0.919**
BP Statistic	4.600	3.444	0.934	0.515	5.888	11.306
Model Fit Statistics						
R^2	0.638***	0.628***	0.254**	0.237**	0.702***	0.735***
Adjusted R^2	0.614***	0.604***	0.205**	0.187**	0.660***	0.676***
MAPE	5.903	5.958	10.02	5.958	5.780	5.323

**** p-value ≤ 0.001 ** 0.001 < p-value ≤ 0.01 * 0.01 < p-value ≤ 0.05*

In case of production of rabi pulses, estimated parametric coefficients, model diagnostics measures and model fit statistics of spline regression models shown in Table 4 shows that in case of both the linear spline model and compound spline model satisfies all error assumptions as all the test statistic used for the purpose are insignificant. But between these two models, linear spline model has higher R^2 and adjusted R^2 and lower MAPE. Thus, linear spline model is the selected best fit model for data on production of rabi pulses in Odisha.

Table 4. Estimated parametric coefficients, model diagnostics measures and model fit statistics of spline regression models fitted to data on production of rabi pulses in Odisha

Model →	Linear Spline	Compound Spline	Logarithmic Spline	Power Spline	Quadratic Spline	Cubic Spline
Estimated Parametric Coefficients						
b_0	285.51***	316.48***	132.86	232.73***	276.01***	298.84***
b_1	29.60***	1.05***	224.69***	1.54***	30.14*	14.32
a_{11}	-28.03***	0.94***	-109.56***	0.80***	-38.03	-47.97
a_{12}	13.82***	1.02***	-2.13	1.01	3.44	150.47
b_2	-	-	-	-	0.08	2.50
a_{21}	-	-	-	-	0.29	7.42
a_{22}	-	-	-	-	-0.39	19.45
b_3	-	-	-	-	-	-0.09
a_{31}	-	-	-	-	-	-0.22
a_{32}	-	-	-	-	-	0.21
Model Diagnostics Criteria						
DW Statistic	1.994	1.994	0.996***	1.073***	2.028	2.289
SW Statistic	0.971	0.987	0.961	0.987	0.964	0.956
BP Statistic	2.614	5.226	5.394	4.731	11.553	17.1*
Model Fit Statistics						
R^2	0.790***	0.764***	0.456***	0.449***	0.792***	0.81***
Adjusted R^2	0.777***	0.749***	0.421**	0.414**	0.763***	0.768***
MAPE	11.359	11.372	19.717	11.372	11.283	10.938

**** p-value ≤ 0.001   ** 0.001 < p-value ≤ 0.01   * 0.01 < p-value ≤ 0.05*

Table 5. Average growth rates (in %) of area under rabi pulses in Odisha for different segments of growth from 1970-71 to 2019-20 obtained by using selected best fit linear spline model

	Segment I	Segment II	Segment III	Segment IV	WP
Average Growth rate (AGR)	4.70	0.67	-2.59	1.43	1.45
P value	<0.01	<0.01	0.604	0.241	0.103
HS/S/NS	HS	HS	NS	NS	NS
Difference in AGR between segment I and II	-4.03 (-85.74%)		-	-	-
P value	<0.01		-	-	-
HS/S/NS	HS		-	-	-
Difference in AGR between segment II and III	-	-3.26 (-486.56%)		-	-
P value	-	<0.01		-	-
HS/S/NS	-	HS		-	-
Difference in AGR between segment III and IV	-	-	4.02 (155%)		-
P value	-	-	<0.01		-
HS/S/NS	-	-	HS		-

HS – Highly Significant (p value ≤ 0.01); S - Significant (0.01 < p value ≤ 0.05
NS – Not Significant (P value ≥ 0.05)

Table 6. Average growth rates (in %) of yield of rabi pulses in Odisha for different segments of growth from 1970-71 to 2019-20 obtained by using selected best fit linear spline model

	Segment I	Segment II	Segment III	WP
Average Growth rate (AGR)	1.18	-0.74	0.35	0.34
P value	<0.001	<0.001	0.787	0.571
HS/S/NS	HS	HS	NS	NS
Difference in AGR between segment I and II	-1.92 (-162.71%)		-	-
P value	<0.01			-
HS/S/NS	HS			-
Difference in AGR between segment II and III	-	1.09 (147.29%)		-
P value		<0.01		-
HS/S/NS		HS		-

HS – Highly Significant (p value ≤ 0.01); S - Significant (0.01 < p value ≤ 0.05
NS – Not Significant (P value ≥ 0.05)

Table 7. Average growth rates (in %) of production of rabi pulses in Odisha for different segments of growth from 1970-71 to 2019-20 obtained by using selected best fit linear spline model

	Segment I	Segment II	Segment III	WP
Average Growth rate (AGR)	6.19	0.21	0.75	2.18
P value	<0.01	<0.01	0.766	0.027
HS/S/NS	HS	HS	NS	S
Difference in AGR between segment I and II	-5.98 (-96.61)		-	-
P value	<0.01		-	-
HS/S/NS	HS		-	-
Difference in AGR between segment II and III	-	0.54 (257.14%)		-
P value	-	0.14		-
HS/S/NS	-	NS		-

HS – Highly Significant (p value ≤ 0.01); S - Significant (0.01 < p value ≤ 0.05
NS – Not Significant (P value ≥ 0.05)

3.2 Estimating the Growth Rates Using the Best Fit Model

It can be seen from the Table 5 that average growth rates of area under rabi pulses in Odisha for different segments shows significant positive growth rate in first and second segment of the whole period. But there is insignificant growth rate in the third and fourth segment and also in the whole period. It is found that there is significant decrease of growth rate between first and second segment and second and third segment. It is found that there is significant increase in growth rate from third to fourth segment. Table 6 shows that average growth rates of yield of rabi pulses in Odisha for different segments shows significant positive growth rate in first segment of the whole period. There is also significant negative growth rate in the second segment. But there is insignificant growth rate in the third segment and also in the whole period. Significant decrease in growth rate is seen in second segment as compared to first. But from second to third segment there is significant increase in growth rate. It can be viewed from the Table 7 that average growth rates of production of rabi pulses in Odisha for different segments show significant positive growth rate in first and second segment of the whole period and also in the whole period. But there is insignificant growth rate in the third segment. There is significant decrease of growth rate from first to second segment, whereas from second to third segment there is no significant change in growth rate.

4. SUMMARY AND CONCLUSION

The respective linear spline model is found to be best fit for area, yield and production of rabi pulses. The study of growth rate in different segments gives an idea regarding the change in performance of pulse crop in Odisha during different time periods.

The growth rate in case of area under rabi pulses is positive and highly significant in the first two segments while in the last two segments it is significantly reduced and become insignificant. Due to this the growth rate in the whole period of study becomes insignificant. The change in growth rate from segment I to II and from II to III is negative and significant. The decrease in growth rate in area under rabi pulses might be due to shifting of pulses to cereals or other non-food grain crops.

The growth rate in case of yield of rabi pulses is positive and highly significant in the first segment while in the second segment it is significantly negative. In third segment the growth rate accelerated but could not be significant. Due to this the growth rate in the whole period of study becomes insignificant. The decrease in growth rate in yield of rabi pulses might be due to inadequate adoption of improved technology and farming practices in the second segment i.e., for the period 1985-86 to 1995-96

The growth rate in case of production of rabi pulses is positive and highly significant in the first segment while in second segment it is significantly reduced which also could not increase significantly in the third segment. But the growth rate in the whole period of study is significantly positive which is due to high growth rate during the first segment of the study period i.e., 1970-71 to 1984-85. The high growth rate during this period may be assigned to the immediate effect of green revolution.

Thus, it is seen that in rabi season, though the growth rate in area and yield of pulses are not significant, the growth rate of production is found to be significant for the whole period of study which shows that the interaction effect of area and yield on production seems to dominate.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Odisha Agricultural Statistics published by Directorate of Agriculture and Food Production, Odisha.
2. Feltz, Miller. An Asymptotic test for the equality of coefficients of variation from k populations, Statistics in Medicine. 1996;15(6): 647-658.
3. Dash A, Dhakre DS, Bhattacharya D. Fitting of appropriate model to study growth rate and instability of mango production in India. Agricultural Science Digest. 2017;37(3):191-196.
4. Montgomery DC, Peck EA, Vining GG. Introduction to linear Regression Analysis, 3rd Edition, New York, john Wiley & Sons, USA; 2001.
5. Lee R, Qian M, Shao Y. On rotational robustness of shapiro-wilk type tests for multivariate normality. Open Journal of Statistics. 2014;4(11):964-969.
6. Zaman A. The Inconsistency of the Breusch-Pagan Test, Journal of Economic and Social Research. 2000;2(1):1-11.
7. Prajneshu, Chandan KP. Computation of compound growth rates in agriculture, Revisited. Agricultural Economics Research Review. 2005;18(July-December):317-324.
8. Dash A, Dhakre DS, Bhattacharya D. Analysis of growth and instability in pulse production of Odisha during Rabi Season: A statistical modelling approach. International Journal of Current. Microbiology and Applied Science. 2017;6(8):107-115.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle4.com/review-history/76521