Environment & Ecology 35 (1A) : 219—223, January—March 2017
Website: environmentandecology.com ISSN 0970-0420

Pre-Harvest Forecast Models for Rice Yield Based on Biometrical Characters

Annu, B. V. S. Sisodia, Sunil Kumar

Received 15 March 2016; Accepted 18 April 2016; Published online 14 May 2016

Abstract Some statistical models for pre-harvest forecast of rice yield based on biometrical characters using data of varietal experiments have been developed in the present paper. On the basis of various statistical measures for comparison and validation of models, it has been found that the multiple linear regression model-I is most suitable followed by model-III for pre-harvest forecast of rice yield.

Keywords Biometrical characters, Experiment data, Regression models.

Introduction

Pre-harvest forecast of crop yield before the harvest constitutes a problem of topical interest. The government, agro based industries, traders agriculturists need statistics on pre-harvest crop yield for policy decisions in terms of in procurement, distribution, buffer-stocking, import-export, price fixation marketing of agricultural commodities, proper planning and measurement of operation. To meet such needs, crop yield forecast under the prevalent system in India are being issued by the Directorate of Economics and Statistics, Ministry of Agriculture, Govt of India, New Delhi. These forecast are, however, of a subjective since these are based on eye-estimate or personal judgment of agriculture officials. The final crop production estimates though based on objective crop-cutting experiments are of limited utility as these become available quite later after the harvest.

The main factors affecting crop yield are soil characteristics, crop-inputs, biomertical characters and weather variables among others during growth period. There are various approaches to develop statistical models for pre-harvest forecast of yield. One approach is based on weather variables including input factors, which requires time series data on yield of the crop and weekly data on weather variables. Recent works on this approach is due to Agrawal and Mehta [1]. Another approach is based on biometrical characters from experimental or survey data. Aneja et al. [2], Annu et al. [3], Draper and Smith [4] and Priya and Radhakrishnan [5] have made use of biometrical characters for development of forecast model for crop yield. An attempt has been made in the present paper to develop a suitable statistical model for pre-harvest forecast of rice yield based on biometrical characters

Annu*, B. V. S. Sisodia, S. Kumar
Department of Agricultural Statistics,
Narendra Deva University of
Agriculture & Technology,
Kumarganj, Faizabad 224229 (UP), India
e-mail : annustat@gmail.com
*Correspondence

Table 1. Detail of experiments. The fitted models along with value of R^2 and R^2_{adj} are given in the Table 2.

Sl. No.	Experiment	Design	Treatment	Replication	Plot size	Date of sowing
1.	I	Randomized Block Design	21 Varieties	03	2.0 m × 3.0 m	10th August, 2011
2.	II	Randomized Block Design	28 Varieties	03	4.0m × 2.0 m	16th August 2011

form experimental data in Faizabad district of Eastern Uttar Pradesh, India.

(The authors are thankful to Dr O. P. Verma, Assistant Professor & Rice Breeder, of N. D. U. A. & T. Kumarganj, Faizabad, for providing the data of rice experiment for the present study.)

Materials and Methods

The materials and the methodologies employed for development of forecast models based on biometrical characters are described below.

Study area

The present study is related to Faizabad district (Uttar Pradesh, India) which is situated between 26° 47' N latitude and 82° 12' E longitudes. It lies in the Eastern Plain Zone of Uttar Pradesh. It has an annual rainfall of about 1002 mm. Nearly 85% of total precipitation is received from south-west monsoon during the month of July to September. However, occasional mild shower occur during winter season. It is liberally sourced by the Saryu (Ghaghara) river and its tributaries. Soils are deep alluvial, medium to medium heavy textured but are easily ploughable. The favorable climate, soil and the availability of ample irrigation facility make growing of rice a natural choice for the area. Rice is generally cultivated during *kharif* season.

Sources and description of data

The data on yield of rice and related biometrical characters were obtained from two experiments conducted at Main Experimental Station of Narendra Deva University of Agriculture & Technology Kumarganj, Faizabad U.P. India. The details of the experiment are described below.

The names of 21 varieties of rice in the experiment-I are as follows : 1-NDRK-11-1, 2-NDRK-11-8, 3-NDRK-11-9, 4-NDRK-11-10, 5-NDRK-11-11, 6-NDRK-11-5, 7-NDRK-11-6, 8-NDRK-11-12, 9-NDRK-11-13, 10-NDRK-11-14, 11-NDRK-11-15, 12-NDRK-11-16, 13-NDRK-11-17, 14-NDRK-11-18, 15-NDRK-11-18, 16-NDRK-11-19, 17-NDRK-11-4, 18-NDRK-11-20, 19-CSR-36, 20-SARJOO-52, 21-IR-28.

Table 2. Forecast models for rice experiment-I. Figures in parentheses denote Standard Error of regression coefficient. *Significant at $p \leq 0.05$, **Significant at $p \leq 0.01$.

Models	Forecast regression equation	R^2 (%)	R^2_{adj} (%)
I	Yield= -34.869+.002X_1+2.693**X_2 + .022X_3 (15.125) (.002) (.853) (.0135) -.189X_4 -.148X_5+3.786X_6 -.049X_7 (.453) (1.867) (3.281) (.046)	84.80*	74.10
II	Yield= -142.448+9.917X_1+ 107.91*X_2 + 4.678X_3 (52.86) (8.93) (32.36) (23.15) -7.986X_4 -1.155X_5 +42.730X_6 -12.13X_7 (19.908) (41.70) (35.77) (10.84)	85.30*	75.10
III	Yield= -92.176 +.195X_1 +22.49**X_2 +.424X_3 (29.77) (.179) (6.932) (2.324) -1.62X4$_4$ -.638X_5 +16.71X_6 -1.02X (3.959) (11.64) (14.21) (.938)	85.10**	74.60
IV	Yield= 90.37 -8330.86X_1 -801.65**X_2 -180.61X_3 (14.64) (7291.64) (230.19) (758.58) +58.58X_4 -10.20X_5 -89.49X_6 +540.55X_7 (164.54) (175.15) (74.21) (468.52)	85.60*	75.50

Table 3. Actual and forecast yield of rice based on rice experiment-I. Figure in parentheses denoted % deviation of forecast, CV: Coefficient of variation.

Actual yield (q/ha)	Forecast Yield (q/ha) Model-I	Model-II	Model-III	Model-IV
28.00	26.61	27.36	26.95	28.34
	(4.92)	(2.28)	(3.71)	(1.21)
23.56	22.79	23.17	22.94	23.88
	(3.26)	(1.65)	(2.58)	(1.40)
21.85	21.33	21.29	21.30	21.37
	(2.36)	(2.51)	(2.51)	(2.51)
RMSE	0.95	0.53	0.76	0.38
PSE (CV)	8.15	9.24	8.63	10.97
PSE (CV)	9.02	10.60	9.70	13.20
PSE (CV)	9.86	10.68	10.23	11.90

The names of 28 varieties of rice in the experiment-II are as follows: 1-RAU-1428, 2-RAU-1-16, 3-CR-2218-64-1, 4-CR-2218-207, 5-CR-2461-1, 6-CR-2462-1, 7-CR-2219, 8-CARI, Dhan-2, 9-CARI Dhan-5, 10-NDRK-11-1, 11-NDRK-11-2, 12-NDRK-11-3, 13-NDRK-11-4, 14-NDRK-11-5, 15-NDRK-11-6, 16-RP-4353, 17-RP-4631, 18-PNL-9, 19-CSR-2K-219, 20-CSR-2K-242, 21-CSR-2K-255, 22-CSR-2K-262, 23-Check (CST-7-1), 24-Check (CST-27), 25-Check (CST-36), 26-Pusa Sugandha-1121, 27-Pusa Sugandha, 28-Narendra Usar Dhan-3.

The following biometrical characters for rice were measured from each plot of the experiment following standard methods of measurement. Their average values corresponding to the each variety were used for the study (Table 1). 1. X_1: Plant population/plot, 2. X_2: Plant height, 3. X_3: No. of tillers/plot, 4. X_4: Length of ear head/plant, 5. X_5: Green leaves/plant 6. X_6: Basal girth, 7. X_7: No. of grain/ear head.

Pre-harvest forecast model based on biometrical characters

The pre-harvest forecast models are proposed using the Multiple Linear Regression Techniques (MLR) by taking quantitative biometrical characters as regressor variables and variety yield as regressand. The following four models are postulated for the development of pre-harvest model.

Model-I $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots\ldots\ldots + \beta_7 X_7 + e$

Model-II $Y = \beta_0 + \beta_1 \log X_1 + \beta_2 \log X_2 + \ldots\ldots\ldots + \beta_7 \log X_7 + e$

Model-III $Y = \beta_0 + \beta_1 \sqrt{X_1} + \beta_2 \sqrt{X_2} + \ldots\ldots\ldots + \beta_7 \sqrt{X_7} + e$

Model-IV $Y = \beta_0 + \beta_1 \frac{1}{X_1} + \beta_2 \frac{1}{X_2} + \ldots\ldots\ldots\ldots + \beta_7 \frac{1}{X_7} + e$

where y is the yield i^{th} of variety (q/ha), $\beta_0, \beta_1, \beta_2 \ldots, \beta_7$ are model parameters and e is error term assumed to follow independently normal distribution with mean zero and constant variance δ^2.

The above models were fitted with the data by applying ordinary least square technique.

Measures for comparison and Validation of models

Coefficient of determination (R^2)

It is in general used for checking the adequacy of the model. R^2 is given by the following formula

$$R^2 = 1 - \frac{SS_{res}}{SS_t}$$

where SS_{res} and SS_t are the residual sum of square and the total sum of squares, respectively.

Adjusted R^2 is given by the following formula

$$R^2_{adj} = 1 - \frac{SS_{res} / (n - p)}{SS_t / (n - 1)}$$

where n is the number of observations and p is the number of regressor variables.

Percent deviation of forecast yield from actual yield

The formula for calculating the percent deviation of forecast yield is as follows.

Table 4. Forecast models for rice experiment-II. The fitted models along with values of R^2 and R^2_{adj}. Figures in parentheses denote standard error of regression coefficient. *Significant at $p \leq 0.05$, **Significant at $p \leq 0.01$.

Models	Forecast regression equation	R^2 (%)	R^2_{adj} (%)
I	Yield= 39.266 + .037**X_1 -2.046X_2 + 227**X_3 (23.82) (.007) (.961) (.201) -2.926*X_4 -.353X_5+1.556X_6 -.147X_7 (.917) (1.461) (3.268) (.074)	72.60**	61.20
II	Yield= -109.09+71.99**X_1 -61.87X_2 + 40.09X_3 (80.48) (14.20) (27.65) (30.76) -119.87*X_4 -3.97X_5 +14.60X_6 +32.14X_7 (36.86) (17.80) (30.64) (15.57)	72.70**	61.40
III	Yield= -50.84 +2.160**X_1 -14.91X_2 +4.01X_3 (47.21) (.492) (6.78) (3.27) -24.70*X_4 -1.52X_5 +6.16X_6 +2.87X_7 (7.65) (6.71) (13.12) (1.41)	72.70**	61.50
IV	Yield= -15.49 -25531.85**X_1 +341.20X_2 -1259.31X_3 (23.73) (1536.64) (151.44) (898.37) =920.26*X_4 -9.67X_5 -28.07X_6 -279.03X_7 (283.88) (41.28) (55.09) (616.38)	71.90*	60.30

$$\text{Percent deviational forecast yield} = \frac{\text{(actual yield - forecasted yield}}{\text{(actual yield)}} \times 100$$

Percent standard error of the forecast

Let $\hat{y}_f$ be forecast value of crop yield and γ_0 be the selected value of vector for X, say $X_0 = (1, \gamma_0)$, then variance of $\hat{y}_f$ as given in Draper and Smith [4] is obtained as

$$V(\hat{y}_f) = \hat{\sigma}^2 X'_0 (X'X)^{-1} X_0$$

where $X'X$ is the dispersion matrix of the sum of squares and cross products of regressors variable and σ^2 is the estimated residual variance. The percent standard error (CV) of the forecast is given by

Table 5. Actual and forecast yield of rice based on rice experiment-II. Figure in parentheses denoted % deviation of foreast, CV: Coefficient of variation.

Actual yield (q/ha)	Forecast yield (q/ha) Model-I	Model-II	Model-III	Model-IV
28.82	29.02 (0.69)	29.39 (2.01)	29.23 (1.45)	29.64 (2.84)
29.31	29.15 (0.54)	29.53 (0.75)	29.36 (0.17)	29.77 (1.56)
26.59	30.06 (13.05)	30.47 (14.63)	30.29 (13.91)	30.74 (15.64)
EMSE	2.01	2.27	2.15	2.46
PSE (CV)	10.01	10.68	10.75	10.61
PSE (CV)	10.21	10.31	10.25	10.20
PSE (CV)	8.10	9.27	8.18	8.35

$$\text{Percent standard error} = \frac{\sqrt{\hat{V}(\bar{y}_f)}}{\textit{forecast} \text{ yield}} \times 100$$

Root mean square error (RMSE)

It is also a measure for comparing two models. The formula of RMSE is given below

$$RMSE = \left[\left\{ \frac{1}{n} \sum_{i=1}^{n} (O_i - E_i)^2 \right\} \right]^{\frac{1}{2}}$$

O_i and the E_i are the observed and forecasted value of the crop yield, respectively, and n is the number of years for which forecasting has been done.

Results and Discussion

Using the data from both experiments, the models described in sub-section 2.3 were fitted by applying ordinary least square method. The data points used for fitted the models were 18 and 25 for experiments-I and II, respectively, as data points corresponding to three last varieties in both experiments were left for validation of the models. The results obtained are

presented below.

Results based on experiment-I

The forecast values of rice yield for last three varieties were computed using the forecast models I to IV and are presented in the Table 2. The percent deviation of forecast, RMSE and percent standard error (CV) of each forecast were also computed with respect to each model. These values along with forecast yield are given in Table 3.

Table 2 shows that only plant height (X_2) has exhibited significant effect on the rice yield in all the models. However, all the biometrical characters together have shown significant effect on the rice yield. The values of R^2 and R^2_{adj} for each model were obtained between 84.80 to 85.60% and 74.10 to 75.50%, respectively. That means all the models are at par in terms of the values of coefficient of determination (R^2). It shows that all the biometrical characters together are able to explain about 75% of the variability in yields of rice varieties and the remaining due to the other factors.

Table 3 reveals that the values of percent deviation of the forecast ranged between 1.21 and 4.92% for all the models, which are of small order. Similarly the values of RMSE were found to be below one for all the models. However, the percent standard error of forecast (PSE) based on the model-I was found little lower as compared to other models. Therefore, on the basis of R^2 and lower value of PSE it can be concluded that the model-I is most suitable followed by the model-III for pre-harvest forecasting of rice yield. However, the other models compete reasonably with midel-I.

Results based on experiment-II

The forecast yields for the last three varieties were computed for the models-I to IV and are presented in the Table 4. The percent deviation of forecast from actual yield, RMSE and percent standard error (CV) of the forecast yield were also computed with respect to each model. These values are also presented in the Table 5.

Table 4 reveals that the plant population per plot (X_1) and length of earhead (X_4) have contributed significantly to the yield of rice varieties while other biometrical characters have not shown significant effect on it. The values of R^2 and R^2_{adj} have been obtained almost at par for all the models, i.e. about 72 and 61%, respectively.

Table 5 indicates that the model-I is most suitable followed by the model III for pre-harvest forecast of the rice yield on the basis of the values of percent deviation of forecast, RMSE, and PSE (CV).

Conclusion

It can finally be concluded from the overall results that among the pre-harvest forecast models, the multiple linear repression model using the original data on biometrical characters as regressors variables (model-I) has been found to be most suitable followed by the model-III (the model in which square root of the data on biometrical characters were used as regressor variables). It is, therefore, recommended that simple multiple linear regression model (model I) can be easily be employed for pre-harvest forecast of rice yield using biometrical characters.

References

1. Agrawal R, Mehta SC (2007) Weather based forecasting of crop yields pest and diseases–IASRI model. J Ind Soc Agric Stat 61 : 255—263.
2. Aneja DR, Lajpat Rai, Deepak Grover, Batra SD (2008) Pre-harvest forecast models for cotton yield. J Cotton Res and Develop 22 : 129–134.
3. Annu, Sisodia BVS, Kumar Sunil (2015) Pre-harvest forecast models for wheat yield based on biometrical characters. Econ Affairs 60 : 89—93.
4. Draper NR, Smith H (1981) Applied regression analysis. 2nd edn. John Willy and Sons, New York.
5. Priya, Radhakrishnan SKK (2008) Multiple regression model for food crop yield forecast. Agric Situat in India 65 : 11—13.