Available online at www.ijpab.com

DOI: http://dx.doi.org/10.18782/2320-7051.5839

ISSN: 2320 – 7051

Int. J. Pure App. Biosci. **5 (5):** 1553-1558 (2017)

Research Article

International Journal of Pure & Applied Bioscience

OPEN ACCESS

Preharvest Forecasting of Rice Yield Using Biometrical Characters Along with Farmers Appraisal in Muzaffarpur District of Bihar

Mahesh Kumar*, M. M. Singh and Subhash Kumar

Department of Statistics, Mathematics and Computer Application

Dr. Rajendra Prasad Central Agricultural University, Pusa, Samastipur, Bihar -848125

*Corresponding Author E-mail: mahesh_smca@yahoo.co.in

Received: 18.09.2017 | Revised: 25.10.2017 | Accepted: 27.10.2017

ABSTRACT

This study is on development of pre-harvest yield forecast model of rice crop of Bihar. For this research observations on plant biometrical characters such as average plant population per m^2 *(*X_1*), average plant height in cm (*X_2*), average number of tillers per* m^2 *(*X_3*), average length of Panicle in cm (*X_4*), application of phosphorus (*P_2O_5*) in kg/ha (*X_6*), disease infestation in percentage (*X_9*) and average plant condition (*X_{10}*) on the basis of eye estimate corresponding to the rice yield were recorded from 50 farmer's field of Muzaffarpur district of Bihar using a Sample survey procedure. The villages were considered as first stage unit of selection while farmer's field as second stage unit of selection. All possible regression analysis was carried out to select the best combination of variables on the basis of some importance statistics such as C.V.,* R^2*, Adjusted* R^2*, RMSE, Residual & Cook's D statistic. However, 10 per cent of observations were kept for model validation purpose. The following models have been proposed. From the best five models one best model is selected for this district based on the above statistic value. Also we find that the actual yield and forecasted yield are very close with forecasting error is below 3.16. On the basis of above facts and below mentioned statistic values model of Muzaffarpur district is recommended for forecasting of rice productivity in Bihar. Thus preharvest forecasted yield of rice will be 16.45 q/h in Muzaffarpur district of Bihar for 2010-11.*

Model for Muzaffarpur district for 2010-11

$$\hat{Y} = 6.89084 + 0.25334\, X_1 + 0.04384\, X_2 + 0.09770\, X_3 - 0.18843\, X_4 - 0.01414\, X_6 - 0.22107\, X_9 + 1.62934\, X_{10}$$

CV = 12.281, $R^2 = 53.56\%$*, Adj* $R^2 = 44.77\%$*, RMSE = 1.85045*

Key Word: *Rice yield forecasting, Biometrical Character of rice, Farmers appraisal*

INTRODUCTION

Rice (*Oryza sativa* L.) is most important crop of Bihar. It is grown in approximately 3.16 million hectares' area under rice in India. Unfortunately, the average productivity even now is not more than 1.5 tonnes per hectare (Directorate of Statistics & Evaluation, Govt. of Bihar-2006-07).

Cite this article: Kumar, M., Singh, M.M. and Kumar, S., Preharvest Forecasting of Rice Yield Using Biometrical Characters Along with Farmers Appraisal in Muzaffarpur District of Bihar, *Int. J. Pure App. Biosci.* **5(5):** 1553-1558 (2017). doi: http://dx.doi.org/10.18782/2320-7051.5839

It is now a complex scientific activity aimed at producing maximum amount of agricultural produce with minimum expenditure in terms of time, space and energy to meet the needs of a growing population and economy. In Bihar rice is cultivated over an area of about 3.1 million ha with production of 3. million tonne and productivity 11.20 q/h (Directorate of Statistics & Evaluation, Govt. of Bihar-2006-07) which is much lower than most of the rice growing states of the country.

Several review were presented regarding Forecasting the yield of autumn rice in the 24-Pargans district of West Bengal has been studied through multiple linear regression analysis[4]. On these line the work done by Draper and Smith[2]. Sardana *et al*[5]., Singh *et al*[6,7]., Chandrahas *et al*[1]., can be cited. Sridharan[8] used multiple regression with principle component of biometrical characters as regressors for predicting the yield of Jowar. Yadav[9] used ridge regression technique to remove multicollinearity among explanatory variables for developing forecast models. A within year growth model, been developed for forecasting of rice and wheat yields components at maturity Jain *et al*[3].

MATERIALS AND METHODS

This study is to identify measurable and non-measurable characters of rice crop in Muzaffarpur districts of Bihar to develop the pre-harvest forecast models with good precision. In this regard entire technical programme of the work has been divided into following sub components.

1. Measurable and non-measurable characters for identification.
2. To obtaining data on rice crop after selection of villages and farmers field.
3. Using least square technique select the best combination of yield attributing characters.
4. Testing the validity of forecast model.

Measurable and non-measurable characters for Identification

Rice is kharif crop with average height 89.3 cm. The number of tillers varies from 5 to 8. It bears panicle on which the grain sets. The average length of panicle varies from 6 cm to 10 cm. For proper growth and development of the crop, chemical fertilizer is required in the form of nitrogen, phosphorus, and potassium as well as irrigation at various stages of crop development. Like other crops, the growth, development and yield of the rice crop are also affected by incidence of pest and diseases. In addition to this there are many unknown factors which affect the plant condition. The basis of observations, the following characters have been chosen for the development of forecast model.

Sl. No.	Variables	Codes of variables	Unit of Measurement	Type of characters
1.	Yield	Y	q/ha	Measurable
2.	Average plant population	X_1	per m^2	Measurable
3.	Average plant height	X_2	cm	Measurable
4.	Average number of tillers	X_3	per m^2	Measurable
5.	Average length of panicle	X_4	cm	Measurable
6.	Nitrogen (N)	X_5	kg/ha	Measurable
7.	Phosphorus (P_2O_5)	X_6	kg/ha	Measurable
8.	Potassium (K_2O)	X_7	kg/ha	Measurable
9.	Irrigation level	X_8	Numbers	Measurable
10.	Disease infestation	X_9	Percentage	Measurable
11.	Average plant condition	X_{10}	Eye estimate	Non-measurable

The last character is non-measurable since, it is based on human appraisal.

Obtaining data on rice after selection of villages and farmers field

Planning were developed to record observations from Muzaffarpur district of Bihar from where samples were obtained randomly at two stages considering village as first and the farmer's field as second stage of selection.

Using least Square technique fit the best combination of yield attributing characters

Out of 50 observations obtained from sampling, 10 per cent of recorded observations were kept for model validation purpose and 90 per cent of observations were used for developing forecast model.

Selection of best subset of regression analysis

Selection of variable through step up procedure and step down procedure does not give unique result whereas selection of variable through all possible regression requires heavy computation and it is possible only through high speed computing facilities. Since computational facilities are available for all possible regression equation, the best subset has to be chosen on the basis of following criteria.

i. R^2 criteria ii adj R^2 criteria
iiii Root Mean square criteria
iv Cooks' D v C.V.
vi Residual analysis

Testing the validity of forecast model

For the validity of regression model, the fulfillment of the following assumptions is required

i. Relationship between Y and X's is linear.
ii. The error has zero mean and constant variance
iii. The error is uncorrelated
iv. The error is normally distributed

Residual analysis and examination of predicted value are the two important aspects for testing the adequacy and the validity of model.

RESULTS AND DISCUSSION

The following stages has been carried out in present studies.

i. Sample selection
ii. Use all possible regression model for developing forecast model.
iii Fit model for the Muzaffarpur district for testing the validity of fitted model.

Forecasting through all possible regression

Sample collected from Muzaffarpur district of Bihar. On the basis of multistage random sampling the 50 farmers were selected randomly in different villages of different Blocks. For Developing forecast model in Muzaffarpur district, out of 50 observations, 5 observations were kept for model validation and 45 observations were put for developing forecast model. The all possible regression analysis was computed for 45 observations through SAS 9.3 The best five model were selected on the basis of RMSE, R^2, Adj R^2 value and C.V., has been presented in table 1.

Among five selected model if we compare between model three from all model the value of R^2, RMSE and AIC (Table- 1) are almost equal but if we compare regression analysis for model 3rd with all model we find that the 3rd model (Table-2) have almost equal of RMSE (1.85), almost equal coefficient of variation (12.28 %), high value of Adj R^2 (44.77%) and maximum R^2 (53.56%) among five regression analysis model. By the above statistic we find that the model third is best for forecasting of rice in Muzaffarpur district of Bihar. From the (Table- 4) maximum percentage forecasting error is 3.16. The value of standard error mean predicted, standard error residual, student residual and Cook's D are almost very low (Table 3) in comparison to other selected model in Muzaffarpur district. The actual yield and predicted value is also very close to each other in (Table 3). This indicated that, the regression subset consisting of X_1, X_2, X_3, X_4, X_6, X_9 and X_{10} could be considered the best subset for prediction purpose.

On the basis of above fact 3rd model is best model for forecasting rice yield in Muzaffarpur district and has been presented in Table 2. The analysis of variance also satisfied that F-value indicate that it is significant at 1 % and 5% level. The Table 2 consisting of regression subset X_1, X_2, X_3, X_4, X_6, X_9 and

X_{10} i.e. average plant population per m^2 (X_1), average plant height in cm (X_2), average number of tillers per m^2 (X_3), average length of Panicle in cm (X_4), application of phosphorus (P_2O_5) in kg/ha (X_6), disease infestation in percentage (X_9) and average plant condition (X_{10}) on the basis of eye estimate. This indicated that the characters such as average plant population, average plant height in cm, average number of tillers per m^2, average length of panicle, application of phosphorus in kg/ha, disease infestation in percentage, average plant condition explained 53.56 % of variation in rice yield of Muzaffarpur district.

The residual analysis for the subset has been presented in Table 3. The table indicated low value of residual for most of observations.

Proposed model in Muzaffarpur district of Bihar for validity test

$\hat{Y} = 6.89084 + 0.25334 X_1 + 0.04384 X_2 + 0.09770 X_3 - 0.18843 X_4 - 0.01414 X_6 - 0.22107 X_9 + 1.62934 X_{10}$

The 5 set of observations have been given in the Table 4 that corresponds to the variable include in the model. These observations have not been used in model building. The estimated deviation and per cent error of forecast for each observation set has been present in Table 4. It was observed that the per cent forecasts is below 3.16 per cent. This indicating the feasibility of the model with good precision. The forecasted yield of rice was predicted to be 16.45 q/hec. in Muzaffarpur district of Bihar in 2010-11.

Table 1: Select best 5 model among all possible regression statistic for forecasting rice in Muzaffarpur district of Bihar

Sl. No.	Independent variable sub set Model	R^2	RMSE	AIC
1.	X_4 X_9 X_{10}	48.75	1.8466	190.25
2.	X_1 X_8 X_9 X_{10}	49.49	1.8559	192.27
3.	X_1 X_2 X_4 X_9 X_{10}	51.13	1.8490	193.61
4.	X_1 X_2 X_3 X_4 X_9 X_{10}	52.29	1.8506	195.49
5.	X_1 X_2 X_3 X_4 X_6 X_9 X_{10}	53.56	1.8505	197.43

Table 2: Regression analysis for the 3rd model for forecasting rice yield in Muzaffarpur district

Regression equation

$\hat{Y} = 6.89084 + 0.25334 X_1 + 0.04384 X_2 + 0.09770 X_3 - 0.18843 X_4 - 0.01414 X_6 - 0.22107 X_9 + 1.62934 X_{10}$

Variable	Parameter Estimate	Standard Error	t-value	Pr >|t|
Constant	6.89084	4.220095	1.63	0.1111
X_1	0.25334	0.20778	1.22	0.2305
X_2	0.04384	0.02865	1.53	0.1344
X_3	0.09770	0.08300	1.18	0.2466
X_4	-0.18843	0.13050	-1.44	0.1572
X_6	-0.01414	0.01409	-1.00	0.3221
X_9	-0.22107	0.06885	-3.21	0.0027
X_{10}	1.62934	0.38384	4.24	0.0001

CV = 12.281, R^2 = 53.56%, Adj R^2 = 44.77%, RMSE = 1.85045

ANOVA

Source	d.f.	S.S.	M.S.	F-value	Pr > F
Model	7	146.10561	20.87223	6.10	< .0001
Error	37	126.69439	3.42417		
Total	44	272.80000			

Table 3: Residual analysis for forecasting Rice yield in Muzaffarpur district for 3rd model

Obs.	Dependent variable	Predicted value	Std. error mean predicted	Residual	Std. error residual	Student residual	Cook's D
1.	18.0000	15.4871	0.7778	2.5129	1.679	1.497	0.060
2.	14.0000	14.4733	0.5294	-0.4733	1.773	-0.267	0.001
3.	17.0000	15.7329	0.6694	1.2671	1.725	0.734	10.010
4.	11.0000	14.9007	0.8380	-3.9007	1.650	-2.364	0.180
5.	18.0000	17.7158	1.0405	0.2842	1.530	0.186	0.002
6.	15.0000	16.4181	0.7852	-1.4181	1.676	-0.846	0.020
7.	17.0000	14.6023	0.7379	2.3977	1.697	1.413	0.047
8.	16.0000	16.0376	0.5687	-0.0376	1.761	-0.0213	0.000
9.	18.0000	18.2745	0.9861	-0.2745	1.566	-0.175	0.002
10.	18.0000	19.0740	1.1158	-1.0740	1.476	-0.728	0.038
11.	12.0000	11.9769	1.0681	0.0231	1.511	0.0153	0.000
12.	19.0000	16.2635	0.7184	2.7365	1.705	1.605	0.057
13.	16.0000	15.1641	0.8196	0.8359	1.659	0.504	0.008
14.	14.0000	14.8001	0.5864	-0.8001	1.755	-0.456	0.003
15.	16.0000	13.4047	0.6520	2.5953	1.732	1.499	0.040
16.	12.0000	12.9307	1.0535	-0.9307	1.521	-0.612	0.022
17.	12.0000	14.4438	0.9010	-2.4438	1.616	-1.512	0.089
18.	19.0000	16.2916	0.5507	2.7084	1.767	1.533	0.029
19.	10.0000	11.7532	0.8486	-1.7532	1.644	-1.066	0.038
20.	12.0000	13.3178	0.9795	-1.3178	1.570	-0.839	0.034
21.	16.0000	15.6101	0.6815	0.3899	1.720	0.227	0.001
22.	11.0000	12.5322	0.8243	-1.5322	1.657	-0.925	0.026
23.	14.0000	16.4109	0.7606	-2.4109	1.687	-1.429	0.052
24.	16.0000	12.1791	0.8300	3.8209	1.654	2.310	0.168
25.	14.0000	14.7098	0.9321	-0.7098	1.599	-0.444	0.008
26.	16.0000	15.9548	0.4689	0.0452	1.790	0.0252	0.000
27.	19.0000	17.0138	0.7056	1.9862	1.711	1.161	0.029
28.	15.0000	18.0705	0.7900	-3.0705	1.673	-1.835	0.094
29.	16.0000	14.9022	0.4790	1.0978	1.787	0.614	0.003
30.	16.0000	14.6659	0.4220	1.3341	1.802	0.740	0.004
31.	15.0000	14.5517	0.9236	0.4483	1.603	0.280	0.003
32.	12.0000	12.1539	0.6873	-0.1539	1.718	-0.0896	0.000
33.	12.0000	12.9098	0.8194	-0.9098	1.659	-0.548	0.009
34.	18.0000	16.2622	0.6728	1.7378	1.724	1.008	0.019
35.	18.0000	17.6024	0.7803	0.3976	1.678	0.237	0.002
36.	12.0000	14.1348	0.8220	-2.1348	1.658	-1.288	0.051
37.	12.0000	13.9579	0.7328	-1.9579	1.699	-1.152	0.031
38.	13.0000	13.4413	0.7810	-0.4413	1.678	-0.263	0.002
39.	16.0000	15.2992	0.7427	0.7008	1.695	0.413	0.004
40.	14.0000	14.3103	0.9879	-0.3103	1.565	-0.198	0.002
41.	17.0000	16.4370	0.6752	0.5630	1.723	0.327	0.002
42.	13.0000	12.2213	0.7339	0.7787	1.699	0.458	0.005
43.	17.0000	16.0101	0.6553	0.9899	1.731	0.572	0.006
44.	17.0000	16.5021	0.5489	0.4979	1.767	0.282	0.001
45.	15.0000	17.0941	0.6389	-2.0941	1.737	-1.206	0.025

Table 4: Forecasting error for observations not included in model building for Muzaffarpur district of Bihar (3rd Model)

$\hat{Y} = 6.89084 + 0.25334\,X_1 + 0.04384\,X_2 + 0.09770\,X_3 - 0.18843\,X_4 - 0.01414\,X_6 - 0.22107\,X_9 + 1.62934\,X_{10}$

Sl. No.	X_1	X_2	X_3	X_4	X_6	X_9	X_{10}	Y	$\hat{Y}$	$Y - \hat{Y} = \hat{e}_i$	$\frac{\hat{e}_i}{\hat{Y}}$ x 100
1	14.5	110.6	14.2	21.4	50	14	4	15	15.48	-0.48	-3.10
2.	16	105.8	16.8	19.8	-	5	3	18	17.88	0.12	0.67
3.	16.5	112	13.8	21.8	-	12	2	14	13.82	01.18	1.30
4.	15	100	15.6	20.2	50	15	3	14	14.36	-0.36	-2.51
5.	15.5	106.8	16	18.2	50	6	3	17	16.48	0.52	3.16

REFERENCES

1. Chandrahas, Bapat, S.R. and Singh, B.H., Pre-harvest forecasting of yield of sugarcane. Agricultural Situation in India. **38(5):** 291-292 (1983).
2. Draper, N.R. and Smith, H., Application Regression Analysis. John Wiley and Sons, New York (1966).
3. Jain, R.C. and Agrawal, R., Probability model for crop yield forecasting. *Biom J*., **34(4):** 501-11 (1992).
4. Khan, S.A. and Chaterjee, B.N., Crop weather analysis for forecasting the yield of autumn rice in west Bengal. *Ind. J. of Agril. Sc.*, **57(11):** 791-794 (1987).
5. Sardana, M.G., Khosla, R.K., Ohri, N.K. and Mitra, P.C., Pre-harvest forecasting of yield rate of jute. *Jute Bull.* **35(7&8):** (1972).
6. Singh, D., Singh, H.P. and Singh, P., Pre-harvest forecasting of wheat yield. *Ind. J. Agril. Sci.*, **46(10):** 445-450 (1976).
7. Singh, D., Singh, P. and Jha, M.P., A study of pre-harvest forecasting yield of jute. *Ind. J. Agric. Res.*, 13(3): 167-68 (1979).
8. Sridharan, H., Pre-harvest forecasting of Jowar yield based on biometrical character. Thesis submitted to IARI for award of M.Sc. degree (1981).
9. Yadav, P.P., Application of ridge regression on collinear data in forecasting. Thesis submitted to IARI for award of M.Sc. Degree (1985).