Economic Affairs
DOI: 10.5958/0976-4666.2015.00012.1

Pre-harvest forecast models for wheat yield based on biometrical characters

Annu✉, B.V.S. Sisodia and Sunil Kumar

Department of Agricultural Statistics, Narendra Deva University of Agriculture & Technology, Kumarganj, Faizabad-224229, India.

✉Corresponding author: annustat@gmail.com

Paper No. 196 Received: 16 October, 2014 Accepted: 17 February 2015

ABSTRACT

Some statistical models for pre-harvest forecast of wheat yield based on biometrical characters in situation of normal and late sowing of wheat have been developed in the present paper. In both the situations, linear multiple regression model (model-I), where biometrical characters are used in the original form, has been found to be the best forecasting model as it has consistently smaller percent standard errors for the forecast yield of wheat alongwith maximum value of R2adj.(49 to 51%).

Keywords: Forecast model, biometrical characters, wheat crop.

A reliable forecast of crop yield before the harvest constitutes a problem of topical interest. Such forecast are needed by the Government, agro based industries, traders and agriculturists etc. Pre- harvest forecast of crop production forms a basis for its policy decision in regards to procurement, distribution, buffer- stocking, import- export and marketing of agricultural commodities etc. To meet such needs, crop forecast under the prevalent system in India are being issued by the Directorate of Economics and Statistics, Ministry of Agriculture, New Delhi. These forecasts are however, of a subjective nature 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 methodologies for pre-harvest forecast crop yield by modelling crop yield with weather variables have been developed by various research workers for different regions of the country in the past. Notably among them are Singh *et. al.* (1986), Singh and Bapat (1988), Singh *et. al.* (2007), Priya and Radhakrishnan (2008) etc. Jain *et. al.* (1984, 1985, 1992b) have developed statistical models for forecasting crop yield based on biometrical characters.

In the present papers, an attempt has been made to develop statistical model for forecasting wheat yield based on biometrical characters in the Eastern region of Uttar Pradesh, India.

Materials and Methods

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

Study area

The present study is related to Faizabad district (Eastern 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. The average minimum temperatures are 18.6°C and 31.3°C, respectively. It is liberally

Table 1. Details of the experiments

Sr. No.	Experiment	Design	Treatment	Replication	Plot Size	Date of Sowing
1	Experiment-I	Simple lattice Design	25 varieties	2	4.5 x3.0m	25th Nov. 2010(normal sowing period)
2	Experiment-II	Simple lattice Design	25 varieties	2	3.5 x2.5m	25th Dec. 2010(normal sowing period)

sourced by the Saryu (Ghaghara) river and its tributaries. Soils are deep alluvial, medium to medium heavy textured but are easily ploughable. The favourable climate, soil and the availability of ample irrigation facility make growing of rice and wheat a natural choice for the area. Wheat crop is generally cultivated during the Rabi season and rice is cultivated during Kharif session.

Source and description of data

The data on yield of wheat, 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 experiments are described below in Table -1.

The 25 varieties of wheat were same in the both the experiments. The name of varieties of are

1- AKDW4021, 2- NIAW1415, 3- DBW46, 4- USA316, 5- DBW51, 6- DBW52, 7- HD2864,

8- HD2932, 9-HD2997, 10-HD2985, 11- HI1563, 12-HI869, 13-HI977, 14-HUW234,

15- HW5207, 16- MP4010, 17- MP4106, 18- MACS3742, 19-NW4035, 20-PDW317

21-PBW590, 22-PBW621 23- PBW315, 24- RSP561 25-WHD943

The following biometrical characters were measured by standard techniques used by Plant breeders and agronomists.

1. Plant population /plot 2. Plant Height 3. No. of tillers/plot 4. Length of ear head/plant

5. Basal Girth 6. Green leaves/plant 7. No. of grain/ear head

Pre-harvest forecast model based on biometrical characters

The pre-harvest forecast models have been developed using Multiple Linear Regression Techniques (MLR) by taking original quantitative biometrical characters as regressor variables and plot yield as regressand. The following four models were postulated for the development of pre-harvest model.

Model- I

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots\ldots\ldots\ldots + \beta_k X_k + e$$

Model-II

$$Y = \beta_0 + \beta_1 \log X_1 + \beta_2 \log X_2 + \ldots\ldots\ldots\ldots\ldots + \beta_k \log X_k + e$$

Model-III

$$Y = \beta_0 + \beta_1 \sqrt{X}_1 + \beta_2 \sqrt{X_2} + \ldots\ldots\ldots\ldots\ldots + \beta_k \sqrt{X_k} + e$$

Model-IV

$$Y = \beta_0 + \beta_1 \frac{1}{X_1} + \beta_2 \frac{1}{X_2} + \ldots\ldots\ldots\ldots\ldots + \beta_k \frac{1}{X_k} + e$$

where Y is crop yield per plot (q/ha), Xi is the value of the ith biometrical Character (i= 1, 2, ...k), $\beta' = (\beta_0, \beta_1 \ldots \beta_k)$ is the vector of model parameter, and e is error term assumed to follow independently normal distribution with mean zero and variance 1. The above models will be fitted with the data by applying least square technique.

Percentage deviation =

$$\frac{\text{Actual yield - Forecast yield}}{\text{Actual yield}} \times 100$$

Root Mean Square Error (RMSE)

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

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

Oi and the Ei are the observed and forecasted value of the crop yield, respectively, and n is the number of

years for which forecasting has been done.

Percent Standard Error of the forecast

Let ŷf be forecast value of crop yield and X0 be the column vector of values of P independent variables at which y is forecasted then variance of ŷf is given by (Draper and Smith, 1998).

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

where $(X'X)$ is the matrix of the sum of square and cross products of regressors matrix X (independent variables) and $\hat{\sigma}^2$ is the estimated residual variance of the model. Therefore, the percent standard error (CV) of forecast is given by

$$\text{Percent S.E.} = \frac{\sqrt{V(\hat{y}_f}}{\text{forecastvalue}} \times 100$$

Results and Discussion

Pre-harvest forecast models based on experiment-I

The data of wheat yield (q/ha) and related biometrical characters corresponding to first 22 plots were used for the development of pre- harvest forecast model, and the corresponding data on last three plots were used for validation and comparison of the forecast models. The models were fitted with the data using least square technique. The fitted models along with value of R2 and R2adj are presented in Table 2.

Table 2. Forecast models for wheat experiment –I

Models	Forecast regression Model	R^2 (%)	R^2_{adj} (%)
I	Yield= 47.999 +.001X1 -.032X2 +.029X3 (11.26) (.001) (.133) (.106) +0.787*X4 -.028X5 +.105X6 -0.364**X7 (.383) (.360) (.612) (.089)	64.00*	49.90
II	Yield = 97.44 -4.484X1+3.288X2+4.524X3 (50.72) (6.82) (28.28) (9.99) +14.834X4+.329X5-.262X6-43.44**X7 (9.12) (9.75) (10.28) (10.84)	62.20*	43.20
III	Yield= 59.174 -.039X1-.167X2+.502X3 (22.65) (.080) (2.56) (1.36) + 4.56X4 -.067X5-.222X6-5.263**X7 (2.485) (2.479) (3.318) (1.294)	63.10*	44.60
IV	Yield= 26.772 +14681.57X1-510.24X2 -99.84X3 (11.59) (15133.21) (1117.65) (172.11) -46.67X4 -8.39X5+4.55X6-650.19**X7 (38.09) (48.69) (31.82) (246.78)	60.10*	40.10

Note: Figures in bracket denote Standard Error of regression coefficient

X1- Plant Population /plot, X2- Plant Height X3- No. tillers/plot X4- Length of ear head/plant , X5- Basal Girth (cm) X6- Green leaves/plant X7- No. of grain/ear head.

The yield of wheat was forecasted for remaining three plots using the models I to IV and are presented in Table 2. The percent deviation of forecast from actual yield, RMSE and percent standard error (CV) of forecast yield have been computed for each model and are presented in the Table 3 along with actual and forecast yield of wheat.

Table 3. Actual and forecast Yield of wheat based on wheat experiment –I

Actual Yield (q/ha)	Forecast Yield(q/ha)			
	I	II	III	IV
37.50	35.00	34.30	34.07	34.88
	(6.67)	(8.53)	(9.15)	(6.98)
35.00	35.33	35.94	35.63	36.52
	(0.94)	(2.68)	(1.8)	(4.34)
35.70	34.37	34.73	34.54	35.12
	(3.72)	(2.71)	(3.24)	(1.62)
RMSE	1.64	2.00	2.12	1.78
PSE(CV)	6.26	6.92	6.91	7.38
PSE(CV)	4.13	4.29	4.23	4.27
PSE(CV)	4.34	4.56	4.53	4.58

Note: Figure in brackets denoted percent deviation of forecast, CV: Coefficient of variation

It can be observed from the Table 2 that number of grain per earhead (X7) has exhibited significant (p<.01) effect on wheat yield in all the models I to IV. However, the length of earhead (X4) has also shown significant effect on wheat yield in the model –I. The values of R2 and R2adj have come out to be maximum (64.00 and 49%, respectively) for model-I followed by model III (63.10 and 44.10%, respectively). It can be observed from the Table (3) that the forecast yields were quite close to the actual yield. On the overall comparison of the values of percent deviation of forecast, percent RMSE, PSE (CV) and R2adj with respect to different models (Table 2 and 3). It can be concluded that the model –I is the best followed by model –III for pre-harvest of the wheat yield.

Pre-harvest forecast models based on Experiment-II

The data of wheat yield and related biometrical characters corresponding to first 22 plots were used for the development of pre- harvest forecast model and the reaming data on last three plots were used for validation and comparison of the forecast models.

Table 4. Forecast models for wheat experiment–II

Models	Forecast regression equation	R^2 (%)	R^2_{adj} (%)
I	Yield= 25.095 +.001X1 +.093X2 +.005X3 (9.290) (.001) (.106) (.116) +.929*X4 +040X5 -930X6 -.042X7 (.366) (.313) (1.126) (.075)	63.20*	50.80
II	Yield= 23.468 -6.409X1 +12.090X2 +3.368X3 (37.66) (8.58) (22.454) (11.52) +21.050*X4 -.266X5 -8.161X6 -.5.985X7 (8.35) (8.41) (16.94) (8.76)	55.10*	42.80
III	Yield= 18.242 -.072X1 +1.439X2 +.239X3 (18.88) (.093) (2.040) (1.544) + 5.876*X4 -.068X5 -3.783X6 -.660X7 (2.317) (2.997) (5.77) (1.072)	62.30*	43.50
IV	Yield= 38.53 +1688.89X1 -206.82X2 -120.13X3 (9.83) (22643.52) (872.38) (196.10) -84.10*X4 +3.787X5 +7.192X6 +154.98X7 (34.18) (21.40) (46.75) (185.37)	59.60*	39.50

Note: Figures in bracket denote Standard Error of regression coefficient

*Significant at $p \leq 0.05$, ** Significant at $p \leq 0.01$

The models-I to IV were fitted with the data using least square technique. The fitted models along with their values of R2 and R2adj are presented in Table 4.

The forecast yield of wheat was computed for the remaining plots using the models I to IV and are presented in Table 5. The percent deviation of forecast, percent RMSE and percent standard error (CV) of forecast yield have been computed for each model and are also presented in Table 5 along with actual and forecast yield of wheat.

Table 5. Actual and forecast yield of wheat based on wheat experiment –II

Actual Yield (q/ha)	Forecast Yield (q/ha)			
	Model –I	Model-II	Model-III	Model-IV
30.00	30.92	31.34	31.14	34.87
	(3.1)	(4.5)	(3.8)	(6.98)
30.50	31.51	31.76	31.63	36.52
	(3.33)	(4.16)	(3.73)	(4.34)
32.50	32.89	33.03	32.96	35.12
	(1.2)	(1.66)	(1.44)	(1.62)
RMSE (%)	.826	1.11	0.96	4.72
PSE(CV)	3.55	3.91	4.17	3.51
PSE(CV)	3.15	4.09	4.05	4.07
PSE(CV)	2.08	2.62	1.02	2.58

Note: Figure in brackets denoted% deviation of forecast, CV: Coefficient of variation

Perusal of Table 4 reveals that none of the biometrical characters except the length of earhead have shown significant effect on the wheat yield in all the models -I to IV. The values of R2 and R2adj have been found to be maximum (63.20 and 50.80%, respectively) in case of model-I followed by model-III (62.30 and

42.80%, respectively). On the perusal of value of percent deviation of the forecast, percent RMSE, R2adj and PSE (CV) of different models from the Table 5 alongwith their values of R2adj, the model-I has been found to be the most suitable one followed by the Model-III for pre- harvest forecast of the wheat yield.

Conclusion

Four types of forecast models based on linear multiple regression models, where biometrical characters were used in difference forms as the independent variables and wheat yield per plot as the dependent variables, have been developed for both normal and late sowing of wheat (Experiment-1 and 2). It can be concluded from the overall results of the study that linear multiple regression model (model-I) where biometrical characters were used in its original form as independent variables has performed best among all the forecast models for both the situations of normal and late sowing of wheat. Hence, the model-I can be recommended for the reliable forecast of wheat yield for both the situations.

Acknowledgement

The authors are thankful to Dr. B. N. Singh, Wheat Breeder, of N. D. U. A. & T. Kumarganj Faizabad for providing the data of wheat experiments for the present study.

References

Jain, RC., Sridharan, H. and Agrawal R .Principal component technique for forecasting of sorghum yield. Ind J Agril. Sci.1954; 54(6): p467-470.

Jain, RC., Sridharan, H. and Agrawal R. Use of growth indices in yield forecast. Biomet J. 1985; 27(4), 435-439.

Jain, RC., Agrawal, R and Singh, KN. A within year growth model for crop yield forecast. Biomet J. 1992; 34(7): p789-799.

Priya, and SKK. Radhakrishnan .Multiple regression model for food crop yield forecast. Agri Situ India. 2008; 65(1): p11-13.

Singh, BH. Bapat, SR. Pre-harvest forecast models for prediction of sugarcane yield. Ind J Agri Sci. 1988; 58(6): p465-469.

Singh, D., Singh, HP. and Singh, P. Pre- harvest forecasting of wheat yield. Ind. J. Agric. Sci.1986; 46(10): p445-450.

Singh, H., Singh, KN., Hassan, B. and Khan, AA. Agro-climatic models for prediction of growth and yield of rice (Oryza sativa) under temperature Kashmir condition. Ind. J. Agri. Sci. 2007; 80(3): p254-257.