Journal of Agrometeorology 9(2) : 158 - 166 (Dec. 2007)

Real time rice yield forecasting over Central Punjab region using crop weather regression model

KANISKA MALLICK[1], J. MUKHERJEE *, S. K. BAL, S. S. BHALLA and S.S. HUNDAL

Department of Agronomy and Agrometeorology, Punjab Agricultural University, Ludhiana 141004

[1]Space Application Centre, Ahmedabad 380015

ABSTRACT

An attempt was made to forecast rice yield over central Punjab by regression model. Correlation study was done for 29 years (1970 to 1998) between rice yield and corresponding weather data. Sensitive period for rice yield in respect to weather parameters were identified for different parameters. Basic regression models were developed by weather parameters in three forms (linear, exponential and power regression). In modified model the technology factor has been introduced to account for the technology inputs through suitable time scale dummy variable. All the three forms of equations for basic as well as modified models are highly significant at 1% level. The modified models viz. linear, exponential and power were obtained with multiple correlation coefficients of 0.86, 0.89 and 0.92, respectively. Power regression model (MCC, 0.92) predicted yield more accurately compared to linear and exponential models. The deviation of yield forecast from actual yield by modified models for 4 years (1999-2002) lies within -16% to +7% as compared to -27% to 13%.

Key words: Rice, regression model, technology trend, yield forecasting

Weather has long been recognized as one major control over the growth and yield of crops. Weather affects crop growth at different phenological phases. Therefore, large variation in yields from year to year and place-to-place is attributable to the weather. A number of statistical techniques such as multiple regression, principal component analysis (Jain *et al*., 1984), Markov chain analysis (Ramasubramanian & Jain, 1999) and agro-meteorological models (Walker, 1989) have been used to quantify the response of crops to weather.

Meteorological subdivision wise rice yield forecast models have been developed by Khan and Chatterjee (1987), Agrawal *et al*. (1986), Huda *et al*. (1976), and Appa Rao *et al*. (1978). By coupling technology trend with weather variables, models were developed for corn and soybean (Swanson & Nyankori, 1979 and Da Mota, 1983) and wheat (Supit, 1997 and Nain *et al*., 2002) crops.

* Corresponding author

Pre-harvest estimates of rice yield is of immense help to the policy makers and the planners for making advance planning for food and other relief measures in areas of impending crop failure, planning food imports and exports in case of an anticipated deficit and surplus, respectively and early recognition of areas of developing food crisis. Keeping this in mind an attempt has been made to develop agrometeorological rice yield-forecasting model for the Ludhiana district of Punjab. The study was conducted to examine the rice yields in Ludhiana district over the last 29 years (1970-98). The purpose of this examination is three fold. First, to search the evidence of a slackening in the rate of yield increase. Second, to ask if the rates of yield increase during this period have been positively or negatively affected by weather. Finally, what are the prospects in the near term for the adoption of available technology?

MATERIALS AND METHODS

The yearly production (q) and area (ha) under rice crop for the period 1970 to 1998 in respect of Ludhiana district have been collected from Statistical Abstract of Punjab. For each year the total production of the district was divided by the total acreage to calculate the rice productivity.

The weekly data of different weather elements for the years 1970 to 1998 were collected from the agrometeorological observatory, Department of Agronomy and Agrometeorology, PAU, Ludhiana for the months covering life cycle of the crop except the harvesting period, since the forecast is to be given before harvesting.

Technology trend

Since 1950 crop yield in many meteorological sub-divisions registered a general upward trend due to advances in the field of agricultural technology, like use of high yielding variety seeds, timely operations, better irrigation and drainage facilities, large scale use of fertilizers and pesticides etc. All these have contributed to the sharp rise in crop yield as compared to the yield of earlier years.

In this model, technology trend is used through suitable time-scale dummy variable to account for the effect of technology on rice yield (Prasad and Dudhane, 1989), wheat yield (Appa Rao, 1983 and Sarkar, 2000), and corn and soybean (Swanson and Nyankori, 1979) yield.

Method of analysis

The traditional multiple regression technique has been employed to develop yield-forecasting models. Firstly, a basic model has been developed by using weather parameters (Table 1) without including technological trend from a data series of 29 years, from 1970 to 1998. Secondly, a modified model has been obtained by introducing an assumed technological trend in the basic model keeping other independent variables constant. The development of modified model was intended to improve the accuracy of forecast of rice yield, by superimposing the impact of agricultural technology in the form of linear time scale dummy variable. The

Table 1: Basic data used in developing the model

Year	Actual yield (Ya) (kg ha^{-1})	Tmax (°C) (X_1)	Tmin (°C) (X_2)	Rainfall (mm) (X_3)	SSH (hrs) (X_4)	NORD (days) (X_5)	RHmax (%) (X_6)	Technology Trend (X_7)
1970	1800	33.9	24.9	124.3	5.5	12	66	1.00
1971	2125	32.9	24.5	209.2	5.8	10	80	3.00
1972	2342	33.8	23.8	226.1	5.0	9	65	4.00
1973	3123	32.4	23.5	12.0	6.7	6	86	8.00
1974	2979	32.4	24.4	1.5	9.3	3	85	8.00
1975	3383	33.3	24.5	89.0	6.9	6	83	11.00
1976	3614	31.8	22.9	58.8	6.2	4	88	12.50
1977	3720	32.3	24.7	270.8	8.5	8	88	13.00
1978	3776	32.5	25.1	34.3	6.9	6	86	13.25
1979	3443	32.9	25.5	17.1	9.0	10	86	13.25
1980	3790	31.5	23.9	58.1	9.8	4	87	13.50
1981	4130	32.8	25.9	38.7	9.8	4	85	13.50
1982	3941	32.5	26.0	5.0	7.3	1	87	13.50
1983	3677	31.2	25.6	84.7	6.8	4	82	13.50
1984	3568	32.4	26.4	159.9	5.7	3	85	13.50
1985	3812	29.2	25.8	44.6	7.3	4	93	14.00
1986	4274	30.5	25.8	154.2	9.4	5	92	14.00
1987	3927	32.2	26.1	4.3	9.1	8	76	14.00
1988	3242	32.1	26.3	157.7	7.8	2	87	14.50
1989	4146	33.1	25.7	28.8	7.0	3	86	14.50
1990	3673	31.2	25.8	28.8	8.0	8	86	15.00
1991	3579	32.4	25.9	0.0	8.1	2	84	15.50
1992	3879	32.0	25.7	28.7	6.0	4	85	15.50
1993	3900	33.0	27.3	66.5	10.8	3	87	16.00
1994	3557	31.7	25.7	111.6	5.7	4	90	16.50
1995	3231	32.8	25.9	6.0	5.9	3	85	17.00
1996	3604	30.9	26.3	50.0	6.2	5	88	17.50
1997	3841	29.3	28.0	28.6	5.8	4	89	18.00
1998	3273	30.8	26.8	308.1	6.3	9	91	18.00

X_1=Maximum temperature during 41st –42nd meteorological week, X_2= Minimum temperature during 32nd – 33rd meteorological week, X_3= Rainfall during 26th –27th meteorological week, X_4= Sunshine hours during 32nd –34th meteorological week, X_5= Number of rainy days during 25th –27th meteorological week, X_6= Morning RH during 29th –30th meteorological week, X_7= Technology trend

average reported crop-yield was taken as dependent variable with weather parameters and technology as independent variables. The general form of the model is:

$$Y_e = a_0 + \sum_{i=1} a_i x_i + \sum a_j x_j$$

Where,

Y_e = Estimated yield, kg ha^{-1}

a_0 = Regression constant

a_i = Regression coefficients for meteorological predictor variables

x_i = i th meteorological predictor variable i = 1, 2, ……n

a_j = Regression coefficients for technological trend variables.

x_j = j^{th} technological trend variable

In the correlation and regression technique significant correlation between yield and the meteorological parameters were identified. The critical periods when weather parameters exert significant influence on yield were located by analyzing the correlation coefficients for statistical and phenological significance. They were then used to calculate the multiple correlations with the yield. Multiple correlations of all the combinations were calculated by dropping one or more variables, which were found less significant. In the development of the models a total of nine parameters (maximum, minimum and mean temperature, rainfall, sunshine hour, number of rainy days, maximum, minimum and mean relative humidity) were tested and out of those parameters maximum temperature, minimum temperature, rainfall, sunshine hour, number of rainy days, and maximum relative humidity, which are statistically significant at the mandatory levels, were used in the final equation. Lastly, the model has been verified with independent data for the years from 1999 to 2002 outside their sampling series. The performance of the model has been examined critically by computing percentage deviations of estimates and forecast yield figures.

Sensitive periods and parameters

Out of all the periods, the sensitive periods of statistical and phenological significance were selected in terms of standard meteorological weeks (SMWs) for

Table 2: Sensitive period and corresponding weather parameters.

Meteorological Parameters	**Sensitive Period (SMWs)**	**Stage of rice crop**	**Effect on rice yield**
Maximum temperature	41^{st} -42^{nd}	late reproductive stage	- ve
Minimum temperature	32^{nd}-33^{rd}	early reproductive stage	- ve
Rainfall	26^{th}-27^{th}	late vegetative stage	- ve
Number of rainy days	25^{th} -27^{th}	late vegetative stage	- ve
Sunshine hours	32^{nd}-34^{th}	mid reproductive stage	+ ve
Maximum relative humidity	29^{th}-30^{th}	early reproductive stage	+ ve

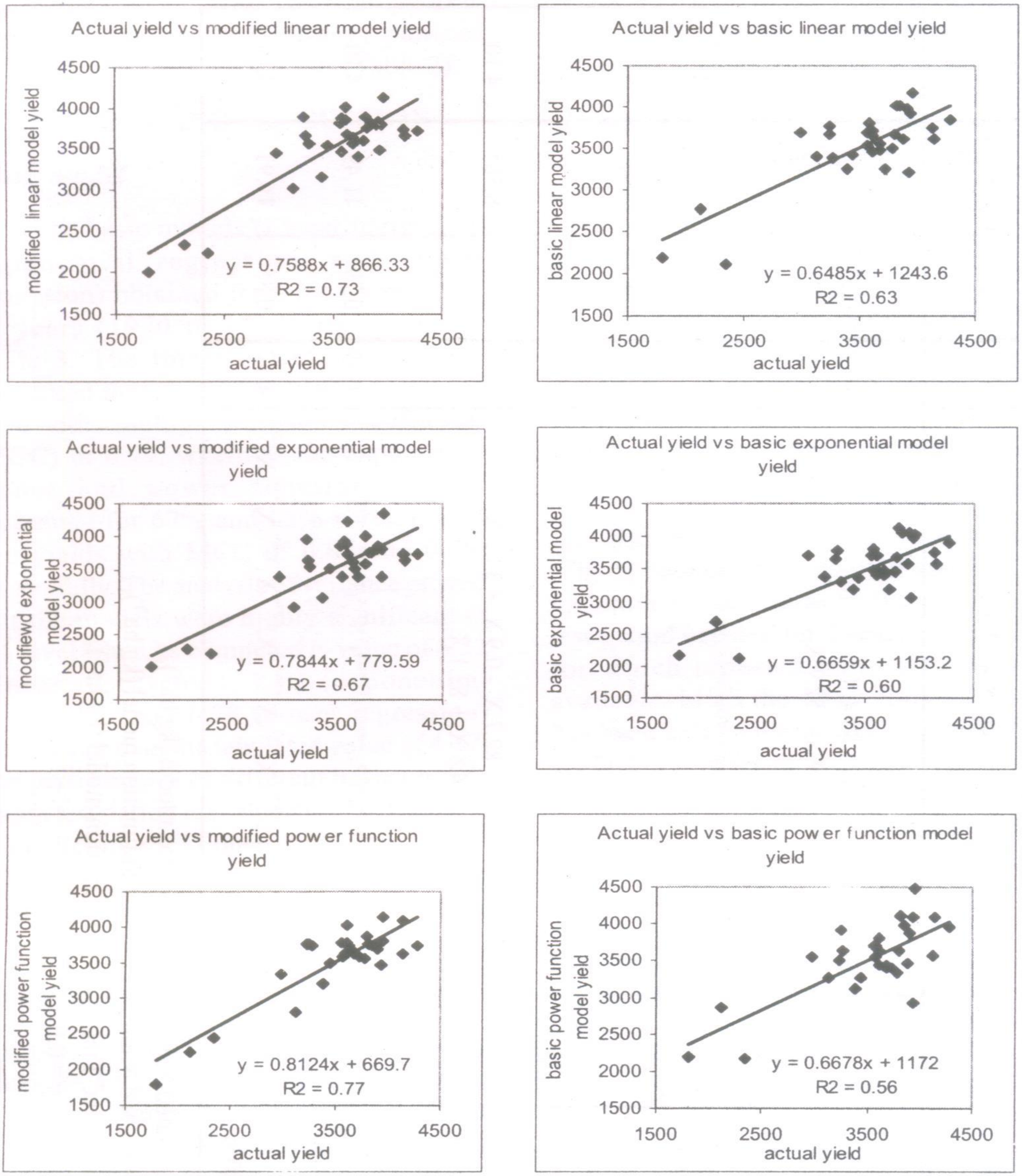

Fig. 1: Relationship between actual yield and predicted yield by different models

Table 3: Basic and modified models to forecast rice yields at Ludhiana

Model	Data series	Multiple Regression Equation	MCC	Percent variation	Computed F-value	Tabulated F-value
Basic	1970-98					
Linear Regression		$Yel = -955 - 24.4 X_1 + 81.4 X_2 - 1.2 X_3 + 36.9 X_4 - 52.9 X_5 + 38.7 X_6$	0.79	0.63	6.23*	4.17
Exponential Regression		$Yee = 728e^{-0.007X_1 + 0.022X_2 - 0.0004X_3 + 0.013X_4 - 0.018X_5 + 0.014X_6}$	0.82	0.67	7.73	4.17
Power Regression		$Yep = 29.2 . X_1^{-0.55} . X_2^{0.44} . X_3^{0.00044} . X_4^{0.08} . X_5^{-0.09} . X_6^{1.18}$	0.79	0.63	6.28	4.17
Modified	1970-98					
Linear Regression		$Yel = 3503 - 1.3X_1 - 82.7 X_2 - 0.9 X_3 + 33.7 X_4 - 34.5 X_5 + 14.8 X_6 + 82.1 X_7$	0.86	0.74	8.56*	4.10
Exponential Regression		$Yee = 3624e^{0.0014X_1 - 0.38X_2 - 0.0003X_3 + 0.012X_4 - 0.012X_5 + 0.006X_6 + 0.03X_7}$	0.89	0.79	11.49	4.10
Power Regression		$Yep = 30233 . X_1^{-0.30} . X_2^{-0.65} . X_3^{0.0029} . X_4^{0.036} . X_5^{-0.045} . X_6^{0.09} . X_7^{0.24}$	0.92	0.85	16.91	4.10

Yel = Estimated Yield by linear function model (kg ha^{-1}),
Yee = Estimated Yield by exponential function model (kg ha^{-1}),
Yep = Estimated Yield by power function model (kg ha^{-1}),
*Significant at $p = 0.01$

regression analysis. The sensitive periods represent maximum tillering, panicle initiation, heading and anthesis (Table 2).

RESULTS AND DISCUSSION

Basic model

The basic models (Linear regression, Exponential regression, and Power regression) obtained from a data series of 29 years (1970 to 1998) are shown in Table 3. The linear model successfully accounted for 63% of total variation in rice yield with a multiple correlation coefficient (MCC) of 0.79, where as the exponential model and power function models accounted for 67% and 63% variations in rice yields with MCC of 0.82 and 0.79, respectfully. The analysis of variance proved that the models were highly significant at 1% level when its computed F-value of 6.23 (linear regression), 7.73 (exponential regression), and 6.28 (power regression) was greater than the tabulated value of 4.17. The performance of different basic models within test period revealed that the deviation in predicting yield ranged between -13 to +30, -26 to +18 and -32 to +36 per cent for linear, exponential and power models, respectively.

Modified model

Improved agricultural technology necessitated the need to modify the basic model by introducing technological trend as an independent linear time scale dummy variable. The modified models (Table 3) gave multiple correlation coefficients of 0.86 to 0.92. The models accounted for 74% (Linear regression), 79% (Exponential regression), and 87% (Power regression) of total variation in yields. Analysis of variance for the modified models were found to bc highly significant at 1% level as seen from a comparison of the computed F-value of 8.56 (Linear), 11.49 (Exponential), and 16.91 (Power regression), respectively, with the tabulated value of 4.10. The performance of different modified models within test period reveales that the deviation in predicting yield ranged between -10 to +20, -14 to +23 and -18 to +17 per cent for linear, exponential and power models, respectively.

There was a fairly close agreement (Fig 1) between the reported and forecast yield during the test as well as outside the test period verified for 4 years (1999-2002) for which reported yield figures were available. In all the three years the model has been able to predict the rice yield over Ludhiana district more satisfactorily over the basic model (Table 4). The yield forecast by the models lies mostly within -16% to 7%. The analysis of variance confirmed that all regression models were highly significant at 1% level. Out of the three modified models the power regression model (MCC = 0.92) performed better compared to other two models. However, linear and exponential regression (modified) models can also be used to forecast of rice yield as both the models are having very high MCC and also the predictions (yield) are comparable with power regression model.

Table 4: Verification of the models over independent data set

Basic models

Year	Tech. Trend	Yebl (kg ha^{-1})	Yebe (kg ha^{-1})	Yebp (kg ha^{-1})	Yr (kg ha^{-1})	% deviation for Yebl	% deviation for Yebe	% deviation for Yebp
1999	-	3658	3416	3813	3611	+1	-5	+6
2000	-	4178	4036	4471	3947	+6	+2	+13
2001	-	3344	3055	3554	3898	-14	-22	-9
2002	-	3490	3153	3265	4322	-19	-27	-24

Modified models

Year	Tech. Trend	Yeml (kg ha^{-1})	Yeme (kg ha^{-1})	Yemp (kg ha^{-1})	Yr (kg ha^{-1})	% deviation for Yeml	% deviation for Yeme	% deviation for Yemp
1999	18.00	3805	3740	3880	3611	+5	+4	+7
2000	18.00	3932	3900	3995	3947	0	-1	+1
2001	18.00	3485	3781	3658	3898	-10	-3	-5
2002	18.00	3571	3772	3638	4322	-16	-13	-15

Yebl = yield estimated by linear basic model, Yebe= Yield estimated by exponential basic model, Yebp= Yield estimated by power basic model, Yeml= Yield estimated by modified linear model, Yeme= Yield estimated by modified power model, Yemp= Yield estimated by modified power model, Yr= Reported yield.

The satisfactory performance of the modified models suggests that it can be used to forecast rice yield in the Ludhiana district which represents the central region of Punjab.

REFERENCES

Agrawal R, Jain RC and Jha MP. 1986. Models for studying rice crop-weather relationship. *Mausam,* **37** (1): 67-70.

Appa Rao G, Sarwade GS, Jaipal J, Sarkar MB, Joseph, Lawrence and Jangle NK. 1978. Forecasting rice yieldsin India from weather parameters-Maharashtra, Rayalseema, Gujarat Region, and Himachal Pradesh. *Pre-published Scientific Report Number 15/78,* Indian Meteorological Department, Pune, 10pp.

Appa Rao GJ. 1983. Estimation of wheat yields over Punjab using district and state models. *Mausam,* **34** (3): 275-80.

Da Mota FS. 1983. Weather- technology models for corn and soybeans in south of Brazil. *Agri. Meteorol.,* **28**: 49-64.

Huda AKS, Ghildyal BP, Tomar VS and Jain RC. 1975. Contribution of climatic variables in predicting rice yield. *Agri.*

Meteorol., **15**: 71-86.

Jain RC, Sridharan H and Agarwal R.1984. Principal component technique eor forecasting of sorghum yield. *Indian J. Agric. Sci.,* **54** (6): 467-70.

Khan SA and Chatterjee BN. 1987. Crop weather analysis for forecasting the yield of autumn rice in West Bengal. *Indian J. Agric. Sci.,* **57** (11): 791-94.

Nain AS, Dadhwal VK and Singh TP. 2002. Real time wheat yield assessment using technology trend and crop simulation model with minimal data set. *Current Sci.,* **82** (10): 1255-58.

Prasad R and Dudhane SN. 1989. Forecasting rice yields in Gangetic West Bengal using rainfall and agricultural technology. *Mausam,* **40** (4): 441-46.

Ramasubramanian V and Jain RC. 1999. Use of growth indices in Markov Chain model for crop yield forecasting. *Biometrical J.,* **41** (1): 99-109.

Sarkar J. 2000. Wheat yield forecasting over Gujarat using agro meteorological model. *J. Maharashtra Agric. Univ.,* **25** (3): 294-97.

Swanson ER and Nyankori JC. 1979. Influence of weather and technology on corn and soybean yield trends. *Agri. Meteorol.*, **20**: 327-42.

Walker GK.1989. Model for operational forecasting of western Canada wheat yield. *Agri. Forest Meteorol.,* **44**: 339-51.