SHORT COMMUNICATION

Oryza Vol. 49. No. 4, 2012 (310-312)

Multiple linear regression models for prediction of rice production and productivity based on rainfall

A S Rama Prasad*, P Krishnamurthy, R Mahender Kumar and B C Viraktamath

Directorate of Rice Research, Rajendranagar, Hyderabad 500 030

ABSTRACT

A study on the influence of monthly total rainfall on rice production and productivity was carried out taking data for five decades from 1950 to 2003. Multiple linear regression technique was used to bring out a statistical model. The correlation studies revealed that there were no interrelationship between monthly rainfall and rice production parameters. In the absence of correlations, year variable was added in the study along with the monthly rainfall to take care of trend effects in rice production parameters. Step wise backward regression brought out that only year and July rainfall were contributing to rice productivity and production. The total rainfall of July for the years 2004 to 2007 were used to validate the models and proved for their accuracy.

Key words: *prediction model, monthly rainfall, rice production and productivity*

Climate and agriculture are very closely interrelated. The change in climate is causing global warming and it is projected to have significant impacts on factors like temperature and precipitation which affect crop production. Precipitation, especially rain, has a dramatic effect on agriculture. All plants need at least some water to survive; therefore, rain is important to agriculture as it is the main source of water either directly or indirectly. A regular rain pattern is usually vital for maximizing the farm output, while too much or too little rainfall can be harmful. Drought can kill crops in massive volumes, while overly wet weather can cause pests and diseases. Agriculture in all nations at least to some extent is dependent on rain; however the Indian agriculture, is heavily dependent on monsoon rains, especially crops like cotton, rice, oilseeds and coarse cereals.

Rice is grown in almost all ecosystems, most important being upland, irrigated, semi deep water, deep water and high altitude situations. Rainfall plays important role from sowing to harvesting of rice crop in influencing productivity, except in assured irrigation conditions.

In order to study effects of global warming on agriculture, different types of models, such as crop development models, yield prediction, quantities of water or fertilizer consumed, had been developed and are being used. Such models condense the knowledge accumulated related to the climate, soil and effects observed of the results of various agricultural experiments and practices. While using these models, it is possible to evaluate the strategies of adaptation to modifications of the environment; however, during last decade with scanty rainfall in the months of June and July, India is facing drought like situation; and in this context, it is very important to take up new studies, since earlier models either frequently inadequate or inconclusive to work out the rainfall effect on rice for all important decisions.

In view of the above problems and in order to develop simplified models, a study was initiated taking into consideration, the data of yearly rice production and productivity figures taken along with monthly rainfall from 1950 to 2003 (Directorate of Economics and Statistics, 1950 to 2003). Multiple linear regression techniques were used for this study (Snedecor and Cochran, 1967) with $Y= b0+b1X1+b2X2+b3X3+\ldots$

Where Y is dependent variable viz., yield (kg ha^{-1}) and production (m.t.) and X1, X2,…being independent variables viz., year and rainfall (mm) of months January to December.

Both stepwise forward as well as backward regression methods were utilized. The data considered for the study consists of yearly average rice productivity, production and monthly average rainfall of India from 1950 to 2003. The yearly rice yields ranged from 668 kg ha^{-1} in 1950 to 2077 kg ha^{-1} in 2003. The yearly rice production varied from 20.58 to 88.63 m. tonnes. The monthly rainfall ranges were 11.0-32.2, 9.2-42.5, 18.2-55.0, 24.2-58.3, 45.6-101.7, 20.3-212.2, 163.9-353.4, 209.8-313.9, 121.9-244.5, 39.0-149.7, 5.6-71.8 and 7.6-98.0 mm, respectively during the period from January to December. In general, January is the lowest rainfall month, while July is the highest rain fall month, in India. Considering the variability, August month had the lowest variability, While, December and November had the highest variability in rainfall.

The correlation studies revealed that there were no interrelations between the rainfall of months and to rice production parameters except that, September rainfall was associated to that of July (0.3169*), August (0.3384*) and October(0.3593**), while May rainfall is related to October rainfall(0.4742**). These all were positive and statistically significant. This showed that Even multiple regression analysis with monthly rainfalls as independent variables and rice productivity as dependent variable has not yielded any good result (R=0.4005 and adjusted R^2=0.08 not significant). It is quite common to use "year" as independent variable in forecasting models of time series data. Hence, year (in four digit format) variable was added along with the monthly rainfall to take care of trend effects in rice production parameters. The multiple R went up to 0.9837 and adjusted R^2 increased up to 0.9571 with "year" and "July rainfall" effects were only statistically significant. To reduce the independent variables from 13, step-down regression technique was used. Without losing much efficiency of the model, 13 independent variables could be reduced to 3 variables which included year, July rainfall and October rainfall; all these effects were statistically significant. The multiple R was 0.9787 and adjusted R^2 square was 0.9554. These three variables were almost as good as 13 variables in predicting the variability in rice yield. With the intention of finding more utility of prediction model by deleting October rainfall it was found that of R remained at 0.9731 and adjusted R square was 0.9449 and both variables had significant effects (Table 1).

Table 1. Multiple regression model for rice production and productivity

	Coefficients	Standard Error	t Stat	P-value	Multiple R	Adjusted R Square
Productivity						
Intercept	-50388.693	1723.420	-29.238	0.000	0.973	0.945
Year	25.894	0.859	30.158	0.000		
July rainfall	1.827	0.402	4.545	0.000		
Production						
Intercept	-2644.133	81.172	-32.575	0.000	0.978	0.955
Year	1.351	0.040	33.396	0.000		
July rainfall	0.094	0.019	4.964	0.000		

May rainfall could predict the extent of October rainfall at 99% confidence level.

In the absence of correlations of monthly rainfall with rice area, production and productivity, the prospectus of further study became bleak. It is obvious that rice production and productivity are dependent on many parameters like ecosystem, variety, duration, irrigation, fertilizer, soils, pest population etc. apart from the weather parameters like maximum temperature, minimum temperature, relative humidity, rainfall etc.

The equation for productivity developed = -50388.69 + 25.894*year + 1.827*July rainfall

The model can be used as early as, at the end of July to predict the productivity of that year. By using the model it was found that the rice productivity for the year 2009-10 could be 2.02 t ha^{-1} with July rainfall of 213.9 (75% of normal rainfall).

Similarly the following equation was arrived for forecasting the rice production

Table 2. Validation of the rainfall and rice productivity, and production model for the years 2004-07

Year	Productivity				Production		
	July Rainfall (mm)	Predicted Productivity (t ha^{-1})	Actual Productivity (t ha^{-1})	Difference (%)	Predicted production (t)	Actual production (t)	Difference (%)
2004	242.0	1.94	1.98	-2.0	85.05	85.13	-0.1
2005	334.0	2.13	2.10	1.8	95.05	91.78	3.6
2006	287.6	2.08	2.13	-2.4	92.03	91.35	0.8
2007	286.2	2.10	2.20	-4.5	93.26	96.43	-3.3

Production = 2644.13+1.351*Year+0.094*July rainfall

Both rice productivity and production models were validated with the rainfall data of 2004 to 2007, which were not included in the fitting of models and presented (Table 2). The deviation of predicted values in rice productivity is varied from 1.2 to 4.5 % only, while in rice production model it varied from 0.1 to 3.6 %. This showed the efficiency of these models in predicting the rice production and productivity accurately.

With multiple R being 0.9779 and adjusted R square being 0.9547, the rice production prediction for the year 2009-10 would be 89.16 m. t taking July rainfall as 213.9 (75% of normal rainfall). There may be fall of 10.3% from 99.37 m.t of 2008-09 year rice production.

Earlier Krishna Priya (2008) also carried out a multiple regression forecast model for paddy yield for Tamil Nadu and presented similar results. Models based on smaller manageable number of parameters in relation to entire weather distribution will not only enable reliable forecast of rice productivity but also reduce the computational errors.

Based on the results it is evident that July rainfall is the most important factor that influence the rice production and productivity in India. Further two multiple regression equations were developed for prediction of rice productivity and rice production and validated for their accuracy.

REFERENCES

S R Krishna Priya and Radhakrishnan, "Multiple regression Model for food Crop Yield Forecast", Agriculture situation in India Vol. LXVI, No.1 2008, pp11-13.

Directorate of Economics and Statistics, Government of India, http://www.indiastat.com

G.W. Snedecor and W.G. Cochran. Statistical methods, Oxford & IBH publishing company 1967. Pp 381-402