Statistics and Applications
Volume 7&8, Nos. 1&2, 2009-10 (New Series), pp. 1-8

A study on pre-harvest forecast of sugarcane yield using climatic variables

S.R. Krishna Priya and K.K. Suresh
Department of Statistics, Bharathiyar University, Coimbatore, India.

Abstract

A suitable statistical model has been developed for forecasting the yield of the sugarcane in Coimbatore district (1981-2004) using the yield data and fortnightly weather variable viz. average daily maximum and minimum temperature, relative humidity in the morning and evening and total fortnightly rainfall. The forecast model was developed using generated weather variables as regressors in model. The generated weather variables were developed using weighted accumulation of fortnightly data on weather variable, weights being the correlation coefficient of the weather variables, in respective fortnights with yield. The data for a period of (1981-2001) was used to develop the forecast model. The validation of the model was done using the data from (2002-2004).

The results revealed that the forecast model developed was able to explain 87% of variation in the sugarcane yield. And it is possible to forecast sugarcane yield successfully two months before harvest.

Key words: Generated weather variables; Weighted accumulation; Regressors; Fortnightly data; Correlation coefficient; Forecast model.

---

1 Introduction

Sugarcane is one of the important commercial crops in India. Sugarcane occupies about 3% of the total cultivated area and it is one of the most important non food grain crops which contributes to about 7.5% gross value of the agricultural production in the country. In India, sugar industry is the second largest agro-based industry, playing an important role. Sugar factories are considered as growth centers in rural areas. India produces 18 million tonnes of sugar annually through more than 453 sugar mills. The exploding population growth and increased per capita sugar consumption warrants increased sugar production with the available area. Proper forecast of such important commercial crops is necessary for future planning and policy making. In the year 2006, in its first advance estimate, the Ministry of Agriculture assessed 283.4 million tonnes of sugarcane crop that was successively revised upward to 355.5 million tonnes. These incorrect forecasts formed the basis of a ban on sugar exports and as a result there was a loss of export business when sugar price was high in the international market. Accurate early warning of crop failures can go a long way in mitigating the undesirable effects like price rise and agrarian distress through public policy. Since sugarcane productivity forecast could help in estimating the production and making decisions regarding export and import policies, distribution, price policies and for exercising measures for storage and marketing, an attempt has been made to develop suitable pre-harvest forecasting model for Coimbatore district of Tamilnadu state.

A number of statistical techniques such as multiple regression, principal component analysis, Markov chain analysis (Ram Subramanian and Jain, 1999), Discriminant function (Rai, 1999) and agro-meteorological models (Baweja, 2002, Bazgeer et al. 2008, Ravi Kiran and Bains, 2007, Muralidhara and Rajegowda, 2002) have been used to quantify the response of crops to weather.

Individual effects on weather factors on rice yield were studied by Jain et al. (1980) and Agrawal et al. (1986). Agrawal et al. (1983) studied the joint effect of weather variables on rice yield. In the above models generated weather variables were used. Weather indices and principal components of weather variables were used in the models developed by Agrawal et al. (1980). Composite models, combining biometrical characters and weather variables were developed by Mehta et al. (2000). Yield forecast models were developed for wheat and rice using weather variables and agricultural inputs on agro-climatic zone basis by Agrawal et al. (2001).Four different approaches, two on original weather variables and two on generated weather variables were used by Khistaria et al. (2004) and Varmola et al. (2004). By coupling technology trend with weather variables, models were developed by Mallick et al. (2007). The present study provides yield forecast models for sugarcane production of Coimbatore district using weather variables.

2 Materials and methods

The yield figures of sugarcane for a period of **(**1981-2004) collected from season and crop report, issued by the state government of Tamil Nadu has been used for the present study. The daily data on weather parameters such as Temperature (max. & min.), Relative humidity (morn. & even.), Amount of rainfall for 23 years period has been collected from weather station located at Sugarcane Breeding Institute, Coimbatore.

2.1 Data and variables used in the study

Fortnightly average data on weather variables have been used for the study namely, X_1 – Max. Temp (° C), X_2 – Min. temp (° C), X_3 – Rel. hum. Mor. (%), X_4 –Rel. hum. Eve. (%), X_5 – Rainfall (mm). The forecast models were developed using the partial crop season data. i.e. the data on weather variables during the active vegetative phase has been used for our study. The data from the period of (1981-2001) has been used in developing the forecast model and the remaining three years data from (2002-2004) has been used for the validation of the models.

2.2 Yield forecast model

The yield forecast model is given by

$$Y = A_0 + \sum_{i=1}^{p} \sum_{j=0}^{1} a_{ij} Z_{ij} + \sum_{i \neq i'=1}^{p} \sum_{j=0}^{1} a_{ii'j} Z_{ii'j} + cT + e, \quad (1)$$

where $Z_{ij} = \sum_{w=1}^{m} r_{iw}^{j} X_{iw}$ and $Z_{ii'j} = \sum_{w=1}^{m} r_{ii'w}^{j} X_{iw} X_{i'w}$.

Here Y is the sugarcane yield (tonnes/hec)

X_{iw} is the value of the i-th weather variable in the w-th fortnight.

$r_{iw}/r_{ii'w}$ is correlation coefficient of Y with i-th weather variable/product of i-th and i'-th weather variable in w-th fortnight.

'm' is the fortnight of forecast

'p' is the number of weather variables = 5

$i = i'$= 1,2,…..5 correspond respectively to maximum and minimum temperatures, relative humidity at 7hr , 14hr and rainfall.

a ,b and c are constants

'T' is year number included to correct for the long term upward or downward trend in yield

And 'e' is the error term.

For each weather variable, two variables were generated- one as simple accumulation of weather variable and the other one as weighted accumulation of fortnightly data on weather variable, weights being the correlation coefficients of the weather variables, in respective fortnights with yield. Similarly, for joint effect of weather variables, fortnightly interaction variables were generated using fortnightly products of weather variables taking 2 at a time.

Stepwise regression was used to select significant generated variables Z_{ij} and $Z_{ii'j}$. Further analysis was carried out including significant generated variables only.

In order to study the consistency of forecast, predicted yield values of subsequent years (not included in the forecast equation) were worked out. Yields of subsequent years were forecasted two months before harvest. For forecasting, observed weather was used up to the time of forecast and normal values of weather variables for the remaining period up to harvest.

3 Results and discussion

The results of ANOVA are presented in table 1. The results of F-test show that the regression equation is highly significant.

Table 1: The results of ANOVA for the regression equation

Model	Sum of Squares	df	Mean Square	F - value	Sig.
Regression	1638.3	10	163.83	6.393**	0.005
Residual	230.65	9	25.63		
Total	1868.95	19			

**- Significant at 1% level

The results of t-test show that the generated weather variables Z_{20} and Z_{40} are significant at 5% level. And further the generated variables Z_{25}, Z_{131} are significant at 1% level. The results of t-test along with the values of partial regression coefficients are presented in Table 2.

Table 2: The results of t-test and partial regression coefficients

Variables	Unstandardized Coefficients		t	Sig.
	B	Std. Error		
(Constant)	-627.984	230.282	-2.73	0.023
Z10	0.493	0.259	1.90	0.089
Z20	1.673	0.628	2.66*	0.026
Z40	0.125	0.049	2.55*	0.031
Z50	2.005	1.016	1.97	0.080
Z150	0.015	0.018	0.84	0.425
Z250	-0.100	0.024	-4.08**	0.003
Z350	-0.001	0.003	-0.50	0.629
Z450	-0.002	0.003	-0.62	0.553
Z131	0.062	0.013	4.73**	0.001
Z141	-0.009	0.008	-1.15	0.279

* - Significant at 5% level

**- Significant at 1% level

3.1 Yield Forecast Model

The yield forecast equation has been developed using the significant generated weather variables based on equation 1. The final yield forecast function using important weather variables along with its R^2 value has been presented below.

$$Y = -627.984 + 0.493\, Z_{10} + 1.673\, Z_{20} + 0.125\, Z_{40} + 2.005\, Z_{50} + 0.015\, Z_{150} - 0.100\, Z_{250} - 0.001\, Z_{340} - 0.002\, Z_{450} + 0.062\, Z_{131} - 0.009\, Z_{141}\,. \quad (2)$$

$R^2 = 0.877$

R^2 value which is measure of goodness of fit indicates that generated weather variables are able to explain 87% of variation in the sugarcane yield.

The performance of the sugarcane yield forecast equation has been tested by comparing the simulated values (which were not included in the forecast equation) with the observed values for a period of three years from (2002-2004) which are presented in table 3. The simulated values of sugarcane yield were 5.59 % lower than the actual yield values for the year 2002. For the year 2003, the simulated values were 3.47 % higher than the actual values. While comparing the results of 2004, it has been noticed that the deviations are little higher in comparison with the previous years. The results of table 3 indicate that the results of forecasted yield are satisfactory. These results reflect that the performance of the yield forecast model is acceptable.

Table 3: Performance of the sugarcane yield model

Year	Observed Yield (tonnes/hec)	Simulated Yield (tonnes/hec)	% of deviation (±)
2002	113	106.67	-5.59
2003	102	105.67	3.47
2004	116	95.62	-17.50

A comparison between the actual and predicted values of sugarcane yield, which were used in developing the forecast model, is presented in table 4. The results show that the percentage of deviations from the actual yield are within ±5 percent range. Figure 1 gives a graphical representation of actual and predicted values of sugarcane production during the period of 1981-2001.

Table 4: Comparison actual and predicted values of sugarcane yield of Coimbatore district, from 1981-2001

Year	Actual values	Predicted values	% of deviation (±)
	(tonnes/hectare)		
1981	113	114.7	1.49
1982	98	98.3	0.27
1983	90	85.9	-4.59
1984	99	101.9	2.91
1985	110	111.7	1.55
1986	106	107.6	1.47
1987	100	105.1	5.1
1988	102	101.3	-0.72
1989	103	107.1	3.95
1991	110	107.2	-2.53
1992	124	117.5	-5.25
1993	122	113.8	-6.74
1994	118	120.1	1.81
1995	117	120.2	2.72
1996	119	115.3	-3.13
1997	101	102	0.95
1998	89	90	1.08
1999	102	100.8	-1.16
2000	106	107	0.92
2001	112	113.8	1.61

4 Summary and Conclusion

Using the forecast model, pre-harvest estimates of sugarcane yield for Coimbatore district could be computed successfully very much in advance before the actual harvest. As the data used for developing this model is of high degree of accuracy, its reliability is also high. Further, this model will produce more accurate results depending on the accuracy of input data provided.

The district government authorities also can make use of the forecast model developed using weather indices, in this study, for obtaining accurate pre-harvest estimates of sugarcane crop.

Till the final production of crops becomes known, decisions have to be made on the basis of informed predictions or scientific forecasts. The main beneficiaries are farmers (decide their procurement prices), traders, exporters and importers (for planning their logistics, inventories and contracts). The processing companies can also plan in advance about the capacity, manpower and marketing strategy.

References

Agrawal, R., Jain, R.C., Jha, M.P. and Singh, D. (1980). Forecasting of rice yield using climatic variables. *Ind. J. Agri. Sci.* **50(9)**, 680-684.

Agrawal, R., Jain, R.C. and Jha, M.P. (1983). Joint effects of weather variables on rice yield. *Mausam* **34(2),** 189-194.

Agrawal, R., Jain, R.C., Jha, M.P. (1986). Models for studying rice crop-weather relationship, *Mausam* **37(1),** 67-70.

Agrawal, R., Jain, R.C. and Mehta, S.C. (2001). Yield forecast based on weather variables and agricultural inputs on agro climatic zone basis. *Ind. J. Agri. Sci.* **71(7),** 487-490.

Baweja, P.K. (2002). Predicting grain yield in maize: canopy temperature based regression indices. *J. Agromet.* **4 (2),** 177-179.

Bazgeer, S., Kamali, G.A., Eslamian, S.S., Sedaghakerdar, A., and Moradi, I. (2008). Pre-harvest wheat yield prediction using agrometeorological indices for different regions of Kordestan Province, *Iran. Res. J. Env. Sci.* **2(4),** 275-280.

Draper, N.R. and Smith, H .(1998). *Applied Regression Analysis*, 3rd ed. New York: John Wiley.

Khistaria, M.K., Vamora, S.L., Dixit, S.K., Kalola, A.D. and Rathod, D.N. (2004). Pre-harvest forecasting of wheat yield from weather variables in Rajkot district of Gujarat. *J. Agromet.* **6,** 197-203.

Mall, R.K. and Gupta, B.R.D. (2000).Wheat yield models based on meteorological parameters. *J. Agromet*. **2(1),** 83-87.

Mallick, K., Mukherjee, J., Bal, S.K., Bhalla, S.S., Hundal, S.S. (2007). Real time rice yield forecasting over Central Punjab region using crop weather regression model. *J. Agromet.* **9(2),** 158-166.

Mehta, S.C., Agrawal, R. and Singh V.P.N. (2000). Strategies for composite forecast. *J. Ind. Soc. Ag. Stat.,* **53(3),** 262-272.

Muralidhara, K.S. and Rajegowda, M.B. (2002). Stocheometric crop weather model to predict the fingermillet grain yield. *J. Agromet.* **4(1),** 53-57.

Ravi K. and Bains, G.S. (2007). Thermal time requirement and heat use efficiency in summer green gram. *J. Agromet.* **9(1),** 96-99.

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

Varmola, S.L., Dixit, S.K., Patel, J.S., and Bhatt, H.M. (2004). Forecasting of wheat yields on the basis of weather variables. *J. Agromet*. **6(2),** 223-228.

S.R. Krishna Priya
Research Scholar,
Department of Statistics,
Bharathiyar University, Coimbatore, India.
E-mail: *sr_priyamadhan@yahoo.co.in*

K.K. Suresh
Prof. & Head,
Department of Statistics, Bharathiyar University,
Coimbatore, India.