Agric. Res. J. Kerala 29 (1991) : 63 - 66

FORECASTING OF SUGARCANE YIELD

Reliable preharvest forecasts of crop production are of immense use as a decision making basis for planners, policy makers and agriculturists alike. Weather variables, agricultural inputs, biometric characters or their combinations serve as independent varibales in such prediction models. Some earlier workers like Jha *et al.* (1981) and Chandrahas *et al.* (1983) included biometrical variables in linear regression models. The present study is an attempt to develop linear prediction models for sugarcane with relatively lesser number of variables through the technique of principle component analysis.

The study was confined to two popular varieties of sugarcane namely Co 997 and Co 62175. The experiment was laid out at the Sugarcane Research Station, Thiruvalla, Kerala. Fifty plots of equal size (0.5 m x 0.5 m) were located in the field. In each plot three plants were randomly demarcated for recording biometrical observations. Observations were recorded at the fifth month after planting and thereafter at an interval of one month till harvest. Data on yield (Y kg), number of canes (X_5), number of tillers (X_6) and number of leaves (X_7) were recorded on a whole plot basis. In the first two months of study, number of canes and number of tillers could not be easily distinguishable and hence were considered as a single unit. The other characters included were height of cane (X_1 cm), girth of cane (X_2 cm), width of 3rd leaf (X_3 cm) and length of third leaf (X_4 cm).

The third leaf was considered to be the most stable leaf among the set of leaves of the sugarcane plant and so it was customary to choose it as the index leaf for recording biometrical observations. Observations on plant characters are the mean values of the three tagged plants.

Zero order correlation coefficients of the various characters among themselves and also with plot yield were calculated for each month separately. As there had been relatively higher degree of inter-relationship among the different variables, the method of principal component analysis was attempted on the intercorrelation matrix to avoid multicollinearity and to reduce dimensionality. The biometric variables so obtained were used as the independent variables in stepwise regression analysis with plot yield as the dependent variable.

The sets of correlation coefficients of the relevant biometric characters with the yield of the two varieties of sugarcane are presented in Table 1.

The cane yield is positively related to most of the above biometric caracters, the only exception being that with the length of the third leaf at specific periods. As the variables were strongly inter-correlated principal component analysis was applied on the inter-correlation matrix and the component vectors (Z_i) were identifed.

The per cent variability (Pj) explained by each component was also determined. The first four principal components that explained more than seventy per cent variation in plot yield were used as regressors in the respective prediction equations. The best subset of generated variables that contributed significantly towards the total variations in crop yield was identified for crop

Table 1. Correlation coefficients of plot yield with biometric characters

Characters	Co 997			Co 62175		
	5th month	6th month	7th month	5th month	6th month	7th month
Height of cane (X_1)	0.462**	0.521**	0.518	0.751"	0.659"	0.753"
Girth of cane (X2)	0.428"	0.567"	0.510"	0.406	0.512"	0.514"
Width of 3rd leaf (X_3)	0.434**	0286	0.077	0.641"	0.577"	0.322"
Length of 3rd leaf (X_4)	-0.199	0.031	0.101	0.570"	0.217	0.226
No. of canes (X_5)	0.685"	0.819"	0.826"	0.430"	0.619"	0.585**
No. of tillers (X6)	-	-	0.495"	-	-	0.539"
No. of leaves (X_7)	0.719"	0.893"	0.890"	0.776"	0.816"	0.644**

** significant at 1% level

Table 2. Principal components used for forecasting (Co 997)

Characters	5th month		6th month		7th month	
	Z_1	Z_2	Z_1	Z_3	Z_1	Z_3
X,	0.453	0.284	0.303	0.214	0.416	-0.189
X_2	0.398	0.431	0.481	0.199	0.387	0.332
X_3	0.438	0.433	0.439	0.260	0.047	-0.434
X_4	-0.185	0.190	-0.008	-0.823	0.095	0.564
X_5	0.439	-0.531	0.449	-0.324	0.488	-0.280
X_6	-	-	-	-	0.384	0.463
X_7	-0.471	-0.471	0.532	-0.255	0.530	-0.234
Pi	40.6	26.9	44.4	14.9	38.5	10.9

Table 3. Principal components used for forecasting (Co 62175)

Characters	5th month		6th month		7th month	
	Z_1	Z_4	Z_1	Z_2	Z,	Z_2
X,	0.478	-0.125	0.520	0.085	0.484	-0.175
X_2	0.327	-0.271	0.469	0.184	0.489	-0.050
X_3	0.452	0.708	0.449	0.309	0.476	0.231
X_4	0.438	-0.177	0.232	0.486	0.387	0.285
X_5	0.302	0.369	0.288	-0.622	0.048	-0.634
X_6	-	-	-	-	0.379	-0.118
X_7	0.419	-0.492	0.416	-0.491	0.055	-0.645
P_i	54.4	6.8	46.9	29.7	39.1	30.5

Table 4. Prediction equations at different months after planting

Variety	Months ft..	Equations	Coefficient of determination (R^2)
Co 997	5	Y = 9.664 + 1.654* Z_1 -0.455*	Z_2 65
	6	Y = 9.639 + 1.714* Z_1 -0.797*	Z_3 81
	7	Y = 9.681 + 1.807* Z_1 -0.475*	Z_3 84
Co 62175	5	Y = 7.081 + 0.972* Z_1 -0.773*	Z_4 61
	6	Y = 7.073 + 1.039* Z_1 -0.567*	Z_2 66
	7	Y = 7.115 + 0.871 Z_1 -0.794*	Z_2 62

(* Significant at 5% level)

forecasting through step-wise regression analysis. It was found that just two significant factors had to be retained in each of the prediction models. The eigen vectors (Z_i) used as regressors in prediction models and the per cent contribution of each of these vectors are given in Tables 2 and 3.

The principal components are linear functions of original variables and are given by

$$Z_j = \sum_{i=1}^{p} a_{ij} \quad X_i$$

Where a_{ij} denote the entries in the eigen vector of the jth component (which are given in Tables 2 and 3) and Xj the standarised values of the biometrical variables.

The percentage variability explained by the first principal component ranged from 38.5 to 54.4. Thus it could be inferred that the first principal component accounted for at least one third of the total variability and could be used as an index of growth during the entire period.

The prediction equations for the forecast of yield at 5, 6 or 7 months after planting are given in Table 4.

The results of analysis revealed that sugarcane yield could be predicted with sufficiently high degree of accuracy as early as in the sixth month after planting with the aid of principal components of biometrical characters observed during that period. In the case of variety Co 62175 there had not been any improvement in predictability after the 6th month whereas for the other variety the improvement was not substantially high. It was also found that variety Co 997 was more amenable to yield forecasting through biometric characters than variety Co 62175. This differential response of varieties may be due to the interaction effects of the genotypes to the environmental factors especially weather.

This research note forms a part of the M.Sc. (Ag. Stat) thesis of the first author submitted to the Kerala Agricultural University in 1984.

College of Agriculture
Vellayani 695 522, Thiruvananthapuram

Alphi Korath
P.V. Prabhakaran

REFERENCES

Chandrahas, S.R., Bapat, K. and Singh, B.H. 1983. On preharvest forecasting of yield of sugarcane. *Agric. Situation in India* 38: 291-292

Jha, M.P., Jain, R.C. and Singh, D. 1981. Preharvest forecasting of sugarcane yield. *Indian J. agric.* Sri. 51: 757-761

□□D