PRE-HARVEST FORECASTING OF SUGARCANE YIELD

BY

ALPHI KORATH

COLLEGE OF HORTICULTURE
VELLANIKKARA
LIBRARY

THESIS

Submitted in partial fulfilment of the requirement for the degree

Master of Science (Agricultural Statistics)

Faculty of Agriculture

Kerala Agricultural University

Department of Statistics

COLLEGE OF VETERINARY & ANIMAL SCIENCES

Mannuthy, Trichur

1984

DECLARATION

I hereby declare that this thesis entitled "PRE-HARVEST FORECASTING OF SUGARCANE YIELD" is a bonafide record of research work done by me during the course of research and that the thesis has not been previously formed the basis for the award to me of any degree, diploma, associateship, fellowship or other similar title of any other University or Society.

ALPHI KORATH

Mannuthy,
10-8-1984.

COLLEGE OF HORTICULTURE
VELLANIKKARA
LIBRARY

CERTIFICATE

Certified that this thesis, entitled "PRE-HARVEST FORECASTING OF SUGARCANE YIELD" is a record of research work done independently by Miss. ALPHI KORATH, under my guidance and supervision and that it has not previously formed the basis for the award of any degree, fellowship or associateship to her.

(P.V. PRABHAKARAN)
Chairman,
Advisory Committee,
Professor of Statistics,
College of Agriculture,
Vellayani.

Place: Vellayani

Date. 10.8.84

COLLEGE OF HORTICULTURE
VELLANIKKARA
LIBRARY

Dedicated To

Late Dr. P. U. Surendran

ACKNOWLEDGEMENT

I wish to express my deep sense of gratitude and indebtedness to Shri. P.V. Prabhakaran, Professor of Statistics, College of Agriculture, Vellayani, for suggesting the problem, valuable guidance, tireless help and constant encouragement throughout the course of the present investigation.

I like to express my gratitude to Dr.K.M. Narayanan Namboothiri, Professor of Agricultural Botany, College of Horticulture, Dr. R. Vikraman Nair, Professor (K.A.D.P.), Shri. K.L. Sunny, Assistant Professor, Department of Statistics, the members of the advisory committee who took keen interest in going through the material and making useful suggestions.

I remain grateful to Dr. M. Krishnan Nair, Dean, College of Veterinary and Animal Sciences, for providing all the necessary facilities for the conduct of the present investigation.

My thanks are also due to Sri. S. Sukumaran Nair, Associate Professor of Agricultural Botany, Sugarcane Research Station, Thiruvalla and other staff members of the Station for their sincere help rendered in the collection of data.

Author wishes to avail this opportunity to thank Shri.V.K.Gopinathan Unnithan, Associate Professor; Smt. K.P. Santha Bai, Technical Assistant; Smt. Narayanikutty, U., Junior Assistant Professor, Department of Statistics and Shri.P.C. Joseph, Kerala University for their help in the present investigation.

I am grateful to the Kerala Agricultural University for the award of junior fellowship during the course of the study.

Smt. L.Lalitha, Senior Office Superintendent also deserves my thanks for neatly typing this manuscript.

Thanks are also due to the staff and students of Department of Statistics, friends and all others whose timely help and encouragement enabled me to fulfill this venture.

ALPHI KORATH

C O N T E N T S

INTRODUCTION

INTRODUCTION

Reliable preharvest forecasts of crop production are of immense use as a decision making basis for planners, policy makers, traders and agriculturists alike. The Government need them for the formulation of various farm policies related to fixation of prices, procurement and distribution, buffer stocking, import, export and marketing of agricultural commodities. The agrobased industrialists and traders need them for formulating their strategies. Forecasts of yield of commercial crops like cotton, jute, tobacco and sugarcane are especially important for trade and industry, because the availability of raw materials is the basis of planning of manufacturing processes and trade operations.

Advance estimates regarding the average area under important crops and their expected yields are being published from time to time during the growth period of the crop in the form of bulletins. Such bulletins appear in the various issues of the Journal 'Agricultural Situation of India'. These bulletins are issued for most of the crops. The first bulletin is issued one month after sowing, second three months after sowing and the third one month before harvesting. As a matter of fact, the number of forecasts depends on the nature of the crop and its importance. For crops like groundnut and ginger, only a single forecast is made. But for rice and wheat five forecasts are made. For sugarcane six forecasts will be issued annually at different periods of plant growth. But all are based on an eye appraisal of the crop.

In India, forecast of crop yield is done by employing two methods. 1. Annawari system. 2. Random sampling method.

In the annawari system which was prevalent in India till recently, the total outturn of a crop was obtained as the product of the area under the crop and the average yield per hectare. Area under the crop is obtained from the village registers. Average yield is estimated by multiplying the normal yield of the crop and the condition factor, taking normal yield as 16 annas. The condition of a crop in a particular year can be described in relation to normal yield in terms of $\frac{x}{16}$ annas, where x is a variable assuming different values from 0 to 16 according to the condition of the crop. The anna condition of the crop is based on merely the eye estimate of the crop reporter. This method is of subjective nature and the final estimate though objective is of limited utility as it becomes available only after harvest.

In spite of its inherent drawbacks traditional method is still adopted in India in respect of minor food grains and some other less important crops. Yields of most of the crops are now based on crop cutting surveys. Although this method was first introduced in India by Hubback in 1925, it was Mahalanobis who introduced objectivity in the method through random sampling. Mahalanobis started crop cutting surveys on jute in Bengal in 1939. Now crop cutting surveys are being conducted for estimating the production of 37 crops and the results are published in the various publications of the Department of Economics and Statistics. In the case of plantation crops like tea, coffee and rubber, such estimates are being prepared by the respective Boards. During 1964-'65 about 95% of India's cereal production and 70% of pulse production were predicted based on crop cutting surveys. For rice and wheat, the percentages were 97 and 99 respectively.

But crop forecasting through crop cutting surveys, though objective and reliable, results in a considerable time lag between the date of sowing and harvesting of the crops and the availability of crop estimates. When the same crop is sown for more than one season, the crop estimate is scheduled to be issued only after the harvest of the second crop. There is also considerable time lag between the due date of the crop estimate and the date of its release. In fact, forecasts of most probable production of the crop should be available even while the crop is standing in the field. Such preharvest estimates of probable production are needed by traders, Government and industrial agencies for policy decisions and administration. The available system of forecasting in India is based on the eye estimate of the crop reporter and is totally unreliable. Thus, there is the need of developing an objective methodology for preharvest forecasts of yield of crops.

The proposed technique of crop forecasting should have some distinct advantages over the traditional method. These include objectivity of the estimates and a measure of precision to determine the reliability of crop forecasts which traditional method cannot provide. Another merit of a crop forecast through this technique is its ability to reflect the impact of the changes in the components of yield over time due to changes in the cultivation of crop varieties and cultural practices. Such changes do slightly affect the coefficients of the parameters in the forecast model, but the model's responsiveness to these changes is not affected as such.

Three objective approaches may be devised for the prediction of crop yields. They are: 1. Prediction based on climatological factors as explanatory variables. 2. Prediction based on agricultural inputs as explanatory variables. 3. Prediction based on biometric characters as explanatory variables.

It is generally believed that production of a crop in a season solely depend on the changes in technology and weather. A sound knowledge of climatic factors and their effect on growth and yield of crops is very much helpful in making reliable forecast of production. Influence of weather begins with preparatory tillage and continues throughout the crop growth period. But crop forecasts based on weather factors often become fruitless due to the variations in factors such as agricultural inputs, soil factors, technological factors and management factors. It will not be often possible to control the effect of all such variables in making a reliable crop forecast. Forecasting models based on macro climatological variables cannot be used in microlevel forecasting. Above all, several years' data are required for building up crop forecast models, based on climatological variables. Thus, preharvest forecasting based on climatological variables though useful has its own limitations, with regard to its range of validity and applicability.

The second approach which is based on agricultural inputs is also quite popular among agricultural researchers. An objective estimate of likely yield of a crop can be obtained by developing crop response models based on plant nutrients present in the soil. Additional variables such as systems of farming, soil management, frequency and level of irrigation etc, may be added to the response function to make it more efficient. A joint approach incorporating climatological as well as input variables will be more efficient in crop forecasting. But the chief defect of this approach is the non-availability of reliable farm data. The average Indian farmer is not at all an expert record keeper, and it is too much for him to keep a complete account of the details of crop cultivation.

Growth of plants is strongly influenced not only by genetical factors but also by environmental factors. The environmental factors include weather parameters, agricultural inputs, soil factors and managemental factors. The effect of all these factors is reflected through morphological characters. Thus a sizable part of variation in crop yield is explained through the variations in morphological characters during different stages of plant growth. If informations on such variables are available at different stages of plant growth, such data can be effectively utilised for building suitable forecasting models.

The present study deals with preharvest forecasting of yield of sugarcane (Saccharum officinarum Linn), which is one of the most commercially important crops of India. Sugarcane is the main source of sugar in India. A number of distillaries and paper factories are running in the country based on the by-products of sugar industry. Sugarcane industry and cultivation employs a considerable part of our population. Thus the crop occupies a prominent place in our national economy. Advance estimates of production of sugarcane will be useful for planners, policy makers, traders and agriculturists. Thus it was felt worthwhile to take up a study to develop suitable mathematical models for pre-harvest forecasting yield of sugarcane using biometric characters.

The main objectives of the study are:

1. To develop suitable models for the pre-harvest forecast of production of sugarcane using biometric characters.
2. To identify the important biometric characters contributing to yield of sugarcane and to assess the relative contribution made by each component.
3. To compare the adequacy of different models in describing sugarcane production.

REVIEW OF LITERATURE

2. REVIEW OF LITERATURE

Yield estimates are available for most of the field crops, but they need not always be unbiased. Several studies have been reported to be conducted to forecast crop yield of some major crops using biometrical characters, agricultural inputs or weather parameters. However in the case of sugarcane only a few studies have been conducted and the literature available is limited. A brief review of the work done so far for the identification of factors affecting yield of sugarcane, their relative importance and role in developing various forecasting models are presented below under three headings.

1. Regression analysis
2. Principal component analysis
3. Path analysis

2.1. Regression analysis

According to Sanderson (1954) the yields of short duration crops can be more accurately predicted by condition report. Wendell (1959) estimated the distribution of winter wheat yields based on climatological data. The fitted model has the property that the mean and variance of the estimated yields are the same as those of the observed yields. Partial regression coefficients of this equation were obtained by dividing the corresponding coefficients of a multiple regression equation by the coefficient of multiple correlation R.

Thomas and Osenburg (1959) conducted a study to determine the effects of manure, nitrogen, phosphorus and climatic factors on the production and quality of crested wheat grass. According to them low yields were associated with high seasonal temperatures. The study also indicated that the estimates of forage production should be based on observations on daily range of temperature.

Basu (1968) found that yields of sugarcane can be expressed as a joint function of crop characteristics namely, height, midgirth and number of canes per clump. The study showed that yield increased with weight and midgirth and assumed a constant value when these characteristics exceeded certain values. It also showed that there should be some optimum value of the number of canes per clump to get maximum yield. The study revealed that yield can be predicted two to three months before harvest using curvilinear models with an accuracy in the range 86 - 88 per cent.

Bohra _et al_. (1969) developed linear models for predicting forage yield based on certain biometric characters like basal diameter. Simple and multiple linear regression equations were worked out using a subset of variables selected according to their contribution towards forage production. The study indicated that the total yield was obtained as the product of estimated yield per plant and the true plant population in the pasture.

According to Singh and Sangha (1970) maximum relative contribution towards cane yield was due to juice percentage followed by girth, number of internodes and number of canes per clump.

Das and Ramchandra (1971) used 50 years meterological data for forecasting bajra yield of Ahmedbad district. They determined 'critical' periods in the growth stages of the crop and developed regression models with rainfall and temperature variables of the critical periods as explanatory variables.

Jitendra Mohan and Gyan Prakash (1971) predicted the yield of jute using a multiple linear regression equation with plant height, basal diameter and fibre content as explanatory variables. Partial correlation analysis revealed that basal diameter had the greatest influence on jute yield.

George and Vijayakumar (1979) fitted a multiple regression model of the form $y = b_0 + \Sigma b_i x_i$ for forecasting the yields of cashew trees (y) based on biometrical characters (x_i). Taking single spot observations on the characters at the first peanut stage, forecasts were made one to two months in advance of the first harvest. Another forecast model by taking the mean of the three observations starting from the first peanut stage at an interval of one month. Total numbers of nut alone was found to contribute substantially to yield.

Parameswaran (1979) conducted a study to identify different vegetative, flowering and fruiting characters influencing yield in cashew. He found that the most important vegetative character contributing towards yield was percentage of flowered shoots per unit area.

Jha et al. (1981) conducted a study in Meerut District, for the pre-harvest estimation of yield of sugarcane on the basis of biometrical characters. A stratified multi-stage random sampling design was adopted for the collection of data. Observations were recorded in the third month after planting in a monthly interval upto 8 to 9 months. High correlations were found to exist between yield and biometric characters. Linear models of the following form were used for the purpose of forecasting.

1. $Y = a_1 + \sum_{i=1}^{4} b_i x_i$

2. $Y = a_2 + \sum_{i=1}^{4} b_i \log x_i$

3. $Y = a_3 + \sum_{i=1}^{4} b_i \sqrt{x_i}$

4. $Y = a_4 + \sum_{i=1}^{4} b_i \frac{1}{x_i}$ where a_i and b_i are the parameters

of the regression models, Y the yield of sugarcane per plot, and x_i biometric characters. All equations were found to be almost equally efficient. Stepwise regression technique was applied to identify the contributing characters. They found that number of shoots per cane and girth of the cane were the most important characters in the first and third stages in the growth phase of sugarcane and using these two variables in a linear multiple regression model, the crop yield can be predicted about two to three months before harvest.

Chaube and Ratnalikar (1982) conducted a study to forecast production of cotton using pickingwise data before the completion of harvest. Yields of cotton from first picking to fifth picking were used as regressors in a forecasting model. It was found that data upto third picking was sufficient for forecasting the total yield.

Vaishnav and Patel (1983) fitted different statistical models for the corecasting of groundnut yield based on biometric characters in Gujarat State. They tried the linear models : 1. $Y = \beta_0 + \sum_i \beta_i x_i + E$ 2. $\log Y = \beta_0 + \sum_i \beta_i \log x_i + E$ 3. $\sqrt{Y} = \beta_0 + \sum_i \beta_i \sqrt{x_i} + E$

4. $\frac{1}{Y} = \beta_0 + \sum_i \frac{\beta_i}{x_i} + E$

Where Y represents the yield, and x_i's are the biometric characters. Multiple correlation coefficients showed that all models were almost equally efficient.

Multiple growth models of production of wheat, rice and sugarcane were tried to determine the factors controlling output by Singh et al. (1983) for 15 districts and whole of U.P., taking output of the crop as dependant variable, area under the crop, consumption of nitrogenous, phosphatic and potassic fertilizers, amount of rainfall and area under high yielding varieties as independent variables. The fitted models explained 31 to 78 percentage of variability in sugarcane production in different districts of U.P.

A study was conducted by Chandrahas et al. (1983) in Kolhopur district of Maharashtra to get objective pre-harvest estimates of production of sugarcane based on biometrical characters. Growth phase of the plant was divided into five stages like early growth, grand growth, flowering, maturity and harvest. Different forecast models were fitted using multiple regression technique. Stepwise regression analysis showed that height, girth and number of millable canes would be adequate to predict the yield at different stages of crop growth.

Pre-harvest forecasting of the yield of groundnut has been attempted by Kathri and Patel (1983) based on twenty two rainfall variables through stepwise regression analysis. It was possible to predict the groundnut yield using four explanatory variables with sufficient degree of precision.

Krishnakumar (1983) used the technique of multiple linear regression for predicting yield of coconut based on foliar nutrient contents and number of leaves retained in the palm. The study revealed that linear regression equation with 12 variables could be used for satisfactory prediction of yield of coconut, with a coefficient of determination 0.86.

2.2. Prediction using principal component analysis

According to Hotelling (1933) principal components are linear combinations of statistical variables which have special properties in terms of variances. The first principal component is the normalised linear combination with maximum variance.

Girshick (1936) has shown that principal components are linear functions of variates which have least variance ascribable to errors of measurement and factor loadings of the principal components are maximum likelihood statistics.

Grafius and Kiesling (1960) developed a technique of estimating the behaviour of oat varieties under different environments by vector analysis. Five vector sets comprised of twenty two varieties of oats grown under five different environments were picked to represent the effects of high night temperature, drought, lodging and composite effect of certain races of both stem and leaf rust. These vectors were used to predict the relative yields of the same twenty two varieties for nine other environments with an assurable degree of success.

An attempt has been made by Abraham and Koshla (1965) to form a single index of the level of incidence of pests and diseases in a field, using component analysis. The index of overall incidence of pests and diseases based on simple ranking method was found to agree closely with the one based on principal component analysis.

Centroid method of factor analysis based on 10 to 12 characters in two groups of populations was done by Murthy and Arunachalam (1967) to find diversity in genus sorghum. Three factors were found to be adequate to account for most of the intercorrelations in both the genotypic and environmental correlation matrices.

Tikka and Asawa (1978) used correlations of 28 genotypes for factor analysis in Lentil through principal component method. Only two factors were found important in explaining relationships in seven traits.

Agarwal et. al. (1980) used the method of principal components to develop weather indices. Further they used principal components as independant variables in a multiple regression equation. The first two principal components account about 80 percentage of variation in yield.

Centroid method of factor analysis was used by Sundaram *et al.* (1980) in cowpea to study its evolutionary pattern. The first three factors accounted for 98 per cent of the total variation in yield.

A forecast model has been obtained by Anonymous (1983) for hybrid Jowar using principal components of biometrical characters. According to them the forecast of hybrid jowar yield is possible one month before harvest for a crop of three and half months duration.

2.3. Path analysis

Wright (1921) proposed the method of path coefficient analysis to study the cause and effect relationship in correlated variables, which was first successfully used in animal breeding programmes.

Dewey and Lu (1959) carried out path analysis to study the components of crested wheat grass production, using six biometric characters. According to Li (1956) the method of path coefficients is essentially a device for analysis or decomposition of correlation coefficient under a structure of causal relationships among linearly related variables.

[illegible] Norman (1971) used path coefficient analysis to identify important component of sugarcane production. They found that number of millable stalks per unit area was the most important factor followed by stalk diameter and stalk length. Stalk density $V = \pi (D/2)^2 L$ contributed less to cane yield other than three variables.

Naphade (1972) revealed by path analysis that the number of leaves per plant was the most important component of fodder yield, followed by the plant height and leaf area.

Malhotra and Jain (1972) found positive correlations among yield, grains per ear and 1000 grain weight in barley. Multiple regression equation was fitted by taking yield as dependant, grain per ear and 1000 grain weight as independant variables. Maximum variation in yield was accounted by variation in grain per ear and 1000 grain weight.

Rao et al. (1973) conducted an experiment to identify the direct and indirect effects of plan5 height, curable leaf number, leaf length and leaf width in the yield of flue cured virginea tobacco.

The regression of yield on the morphological characters is obtained after logarithemic transformation of the data. According to them plants with broad and more logarithemic curable leaves were ideal for selection.

Hooda et al. (1979) applied path coefficient analysis on sugarcane genotypes for identifying a few morphological characters as reliable indices for selection at the settling stage. They observed that stalk weight and stalk height have positive direct effect on yield followed by brix content.

Sundaresan et al. (1979) observed significant positive genotypic correlations between biometric characters at seedling stage and settling stage. According to them cane thickness was the most reliable character for selecting genotype at both the stages.

An experiment was conducted by Singh et al. (1981 a) with 48 varieties of sugarcane in a R.B.D. with three replications. Ten clumps were selected atrandom in each plot at the time of germination, and data were collected from these clumps till harvest. Observations were recorded on 8 traits. According to this study selection on sugarcane should be based on stalk height, stalk girth, number of internodes per stalk, number of green leaves per stalk and brix. Since these characters are largely governed by additive genes.

Number of millable canes and number of internodes per clump were identified by Singh et al. (1981 b) as the important components having direct influence on brix quality. They got a high residual effect (0.956) indicating that some prominent characters were not included in the model.

In another study conducted by Singh and Sharma (1982) cane thickness and number of millable canes were found to be the major contributors towards sugarcane production.

MATERIALS AND METHODS

3. MATERIALS AND METHODS

Data used for the study were collected from the bulk crop of sugarcane available at the Sugarcane Research Station, Thiruvalla. The study was confined to two popular varieties of sugarcane namely CO-997 and CO-62175. Fifty plots of equal size were located in the experimental field under each variety. In each plot three plants were demarcated, the two end plants and the middle plant for recording biometric observations such as height of the cane (cm), girth of the cane(cm), width of the third leaf from the top (cm), length of the third leaf from the top (cm), and the number of green leaves. The identity of the three plants was retained till harvest. First observation was recorded in the fifth month after planting and thereafter at an interval of one month till harvest. The height of the cane was measured from the ground level to the last node and girth was taken at the middle of the cane. At harvest weights of the three selected canes (gms) were recorded separately in addition to the plot yield (kg). The total number of tillers, canes and leaves in each plot were also recorded in each month. In the first two months of study it was very difficult to distinguish between tillers and canes. So the observations on number of tillers and canes were considered together as a single observation.

In order to get a rapid method of determining leaf area using linear dimensions of the leaf the product of the maximum length and maximum breadth (say x) was correlated with actual leaf area (y) obtained by tracing the leaf on the graph paper. A sample of thirty leaves of different sizes was used for this purpose. Different regression equations were worked out for estimating leaf area with and without applying various transformations. An approximate prediction equation for estimating leaf area was selected on the basis of the twin criteria of simplicity and efficiency. In estimating leaf area it was assumed that shape of the leaf did not undergo significant changes during the entire growth period of the crop and hence the equation developed for estimating leaf area at the harvest period of the crop was also equally applicable during the early stages of crop growth. Further it was assumed that the leaf area of a leaf was not associated with its position. From these assumptions and the methods developed it was possible to estimate the leaf area of the third leaf of the tagged plants. The following models were tried for estimating the leaf area.

1. $y = bx$ 2. $y = a+bx$ 3. $\log y = a+bx$

4. $\log y = a+b \log x$ 5. $y = a+bx+Cx^2$ 6. $y = a+b\sqrt{x}+Cx$.

The total leaf area of the cane was estimated by multiplying the number of leaves per cane and mean leaf area of third leaf of the tagged plant from the top.

Pre-harvest forecasting of plot yield of sugarcane was attempted using

1. Cane-wise observations
2. Plot-wise observations

3.1. Prediction using cane-wise observations

Regression analysis was carried out using the observations recorded from the three selected plants for each of the fifty plots for the two selected varieties for each month separately to find the adequacy of different models for prediction. During the course of the plant growth some of the selected canes got damaged due to disease incidence, so observations on those plants were not accounted for the analysis. The biometric characters used for the prediction of cane yield were: 1. Height of the cane (x_1) 2. Girth of the cane (x_2) 3. Width of the third leaf from the top (x_3). 4. Length of the third leaf from the top (x_4). 5. Leaf area of third leaf from the top (x_5) 6. Total number of leaves in a cane (x_6) 7. Total leaf area of all the leaves of a cane (x_7). Correlation coefficients were worked out with the above mentioned seven morphological characters among themselves and with yield (y) for each month separately. The significance of correlation coefficient was tested using students 't' test. Selection of explanatory variables of the multiple linear regression equation was done on the basis of the relative influence of the various characters on cane yield. Only those characters which showed significant linear relationship with yield alone were retained in the model. A class of multiple linear regression equations was fitted for each month separately, coefficient of determination (R^2) calculated and tested for significance.

In multiple linear regression the values on P explanatory variables $x_1, x_2, --- x_p$ were used to predict the average value of the dependant variable y using the functional form

$y = a+b_1x_1+b_2x_2 + --- +b_px_p$. It was assumed that the independant variables were measured without error and errors in the dependant variable y followed a normal distribution with zero mean and constant variance. The parameters of the fitted equation were estimated by applying the ordinary principle of least squares. The set of normal equations for estimating the parameters $b_1, b_2, ---, b_p$ was

$$\underline{Siy}_{px1} = \underset{\sim}{Sij}_{pxp}\underline{B}_{px1}$$

where $\underline{Siy}$ is the vector of sum of products of the i^{th} explanatory variable with the dependant variable.

$\underset{\sim}{Sij}_{pxp}$ is the sum of product matrix of the explanatory variables, $\underline{B}_{px1}$ is the vector of parameters $b_1, b_2, ---, b_p$

$$\underline{B}_{px1} = \underline{C}ij_{pxp} \ \underline{S}iy_{px1}$$

where $\underline{C}ij_{pxp}$ is the inverse of Siy_{pxp}

$$a = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2 - - - b_p\bar{x}_p$$

The estimated value of the dependant variable can be obtained as

$$\hat{y} = \hat{a} + \hat{b}_1 x_1 + \hat{b}_2 x_2 = - - - + \hat{b}_p x_p$$

where $a, b_1 - - - b_p$ were the least square estimates of a, $b_1, b_2 - - - b_p$. The quantity $\sum_{i=1}^{p} \frac{b_i S_{iy}}{S^2 y}$ is called the multiple coefficient of determination, which is a measure of the percentage variation in the dependant variable, explained by the independant variables of the fitted equations. The significance of R^2 was tested using the variance ratio test given by

$$F = \frac{R^2}{1-R^2} \times \frac{n-p-1}{p}$$

The adequacy of a restricted model over the full model was tested using the F test

$$F = \frac{R_f^2 - R_r^2}{1-R_f^2} \times \frac{df}{dr-df}$$

where R_f^2 is the coefficient of determination for the full model, R_r^2 is the coefficient of determination for the restricted model, df is the degrees of freedom of R_f^2 and dr is the degrees of freedom of R_r^2.

A significant F test implies that the restricted model is inefficient to cope with the full model.

The following transformations were tried in search of a better fit to the data.

1. $y = a_1 + \sum_{i=1}^{7} b_{i1} \log x_i$. 2. $y = a_2 + \sum_{i=1}^{7} b_{i2} \frac{1}{x_i}$

3. $y = a_3 + \sum_{i=1}^{7} b_{i3} \sqrt{x_i}$ 4. $\log y = a_4 + \sum_{i=1}^{7} b_{i4} \log x_i$

where x_is are the biometric characters, a_j $j=1,2,3,4$ are the constants to be enumerated and b_{ij} are the respective regression coefficients.

The expected yield of a plot was determined as the product of the number of canes per plot and expected cane yield as obtained from the fitted model.

3.2. Prediction using per plot observations

An attempt was also made to predict the yield of sugarcane based on per plot observations. In the first two months of the study there were only eight biometrical characters and in the next three months there were nine characters. The characters included in the study were (1) Height of the cane (x_1), (2) Girth of the cane (x_2),

(3) Width of third leaf from the top (x_3), (4) Length of third leaf from the top (x_4), (5) Area of third leaf from the top (x_5), (6) Number of canes per plot (x_6), (7) Number of tillers per plot (x_7) (8) Total number of green leaves per plot (x_8) (9) Total leaf area of all leaves in a plot (x_9). The dependant variable was the plot yield in kgs. Observations on height, girth, width of third leaf, length of third leaf and area of third leaf are the mean values of measurements in the three tagged plants in each plot. Observations on other characters, viz., number of canes per plot, number of tillers per plot and number of leaves per plot were recorded on a whole plot basis. Correlation coefficients were calculated among the characters themselves and with yield and multiple linear regression equations of yield on the above mentioned biometrical characters were fitted as described in section 3.1. The four nonlinear models as explained in section 3.1. also were tried for all the months.

3.4. Prediction using principal Component analysis

In many of the experimental situations of multivariate data analysis the characters may be interrelated. In such situations in order to examine the relationships among the set of P correlated variables, it may be useful to transform the original set of P variables into a new set of P uncorrelated variables called principal components.

These new variables are linear combinations of original variables and are derived in decreasing order of their importance so that the first principal component accounts for maximum of the variation in the original data.

Plot wise observations were utilized for predicting plot yield using principal component analysis. The analysis was carried out for each month separately.

Since the variance covariance matrix of the standardised values is the same as the correlation matrix of the characters, the principal components were calculated from the matrix of correlation coefficients.

The eigen vector corresponding to the highest eigen value will be the first principal component. Similarly r^{th} principal component is the eigen vector corresponding to the r^{th} eigen value of the matrix.

Let $\underline{x}^1 = (x_1\ x_2\ ---\ x_p)$ be a P dimensional random variable with mean $\underline{\mu}$ and dispersion matrix Σ. The components $z_1, z_2\ ---\ z_p$ are defined as $\underline{z}_j = a_{ij}\underline{x}$ where $\underline{a}'_j = (a_{1j}\ a_{2j}\ ---\ a_{pj})$

The first principal component z_1 is to be so chosen as to have maximum variance and should satisfy the orthogonality condition $\underline{a}_1\underline{a}_1' = 1$

Variance of the first component $V(Z_1) = \underline{a}_1'\ \underline{R}\ \underline{a}_1$ where $\underline{R}$ represents the correlation matrix. Maximising

$\underset{\sim}{a}_1' \, \underset{\sim}{R} \, \underset{\sim}{a}_1$ using Legrange multiplier λ_1

$\underset{\sim}{a}_1' \, \underset{\sim}{R} \, \underset{\sim}{a}_1 - \lambda_1(\underset{\sim}{a}_1'\underset{\sim}{a}_1 - 1)$ (1)

$(\underset{\sim}{R} - \lambda_1 \underset{\sim}{I})\underset{\sim}{a}_1 = 0$

Equation (1) have a non-zero solution if $(\underset{\sim}{R} - \lambda_1 \underset{\sim}{I}) \, \underset{\sim}{a}_1$ is a singular matrix, λ_1 must be chosen such that $|\underset{\sim}{R} - \lambda_1 I| = 0$. A non-zero solution exists for equation (1) if λ_1 is an eigen value of $\underset{\sim}{R}$.

$$
\begin{aligned}
V(Z_1) &= \underset{\sim}{a}_1' \, \underset{\sim}{R} \, \underset{\sim}{a}_1 \\
&= \underset{\sim}{a}_1' \, \lambda_1 \underset{\sim}{a}_1 \\
&= \lambda_1 \quad \text{since } \underset{\sim}{a}_1' \, \underset{\sim}{a}_1 = 1
\end{aligned}
$$

If λ_i is the i^{th} eigen value then the variance of the i^{th} principal component is λ_i. To maximise the variances λ_1 must be chosen as the largest eigen value. From equation (1) the principal component $\underset{\sim}{a}_1$ must be the eigen vector of $\underset{\sim}{R}$, corresponding to the largest eigen value λ_i. In the same manner the second principal component will be the eigen vector corresponding to the second eigen value and so on.

The sums of variances of the original variables and their principal components are the same. The total variance in the system will be trace $(\underset{\sim}{R})$ which is the same as sum of the eigen values. The proportionate variation explained by the i^{th} component is $\frac{\lambda_i}{P}$, where P is the trace (R). The first m components account for

$$\sum_{j=1}^{m} \frac{\lambda j}{P} \times 100$$ percentage of the total variation.

Correlation between j^{th} component and i^{th} variable is given by $r(x_i, Z_i) = \lambda_j^{\frac{1}{2}} a_{ij}$. This is known as factor loadings or component loadings. If a variable has no significant correlation with a component then that variable is not contributing much to the variance of the component and hence unimportant in describing the causal structure. Thus variables can be ranked according to their relative importance and unimportant variables can be eliminated. But if the so called unimportant variables are correlated with any other component elimination will be dangerous. So elimination is done only after considering the major components.

The utility of a component depends upon the variability it accounts for. Thus the first step is to select the important components from the 'P' set of P components. A simple rule is to consider only these components which account for more than 75 percentage of the total variability. Another rule is to select only those components with latent roots numerically greater than unity. A third method used is that of testing the dissimilarity among latent roots by χ^2 test and retaining only those characteristic vectors corresponding to the roots which are distinct.

On several occasions principal component analysis is only the first step in multivariate data analysis. The derived observations can be subjected to further statistical analysis. In this study the values of the first three principal components are used as independent variables in multiple linear regression analysis with final yield as the dependant variable. Prediction equations during different months were worked out and corresponding coefficients of determination calculated.

3.3. Path coefficient analysis (Plant wise approach)

The simple correlation analysis does not take into account the cause and effect relationship between the related variables. The technique of path analysis developed by Wright (1921) is useful to study the functional relationship between causal factors and their effect. The method can be applied to assess the relative contribution of various biometric characters affecting yield of sugarcane so as to enable the researcher to identify the important variables to be retained in the prediction equation.

Path analysis is concerned with the decomposition of simple linear correlation coefficient between causal variable and the effect factor into numerous components due to direct effect of the causal factor and its indirect effect through other factors. The linear model used for the path analysis is of the form

$$Y = a+b_1x_1+b_2x_2+ - \quad +b_nx_n+b_ux_u$$

Where b_i's are partial regression coefficients, x_i's are the exogenous variables (the biometrical characters) and 'Y' the endegenous variable (cane yield).

Path coefficients are standardised regression coefficients and are given by $P_{iy} = b_i \frac{\sigma_i}{\sigma_y}$

Where σ_i and σ_y have their usual meanings. The indirect effect of x_i through x_j is $r_{ij}P_{iy}$.

Standardising the variables

$$\frac{y-\bar{y}}{y} = Y, \quad \frac{x_i-\bar{x}_i}{x_i} = x_i$$

$$Y = P_{1y}x_1+P_{2y}x_2+ - - - - +P_{ny}x_n+P_{uy}x_u$$

Let x_q be any variable

$$\text{Then } r_{yq} = \frac{1}{n}\sum_{i=1}^{U} Y\ x_q$$

$$= P_{1y}r_{1q}+P_{2y}r_{2q} + - - - -P_{ny}r_{nq}+ P_{uy}r_{uq}$$

$$r_{yq} = \sum_{i=1}^{u} P_{iy}r_{iq}$$

which is known as the first law of path analysis. If both the variables are purely exogenous, correlation cannot be broken down. If x_q is any variable among x_i's, $i=1,2....n$

Then $r_{yq} = \sum_{i=1}^{n} P_{iy} r_{iq}$

Since there are n exogenous variables the cause and effect relationship can be defined by n simultaneous equations in n unknowns. The coefficient matrix of normal equations will be the correlation matrix. The set of normal equations are

$$\underset{\sim}{R}_{nxn} \underline{P}_{iy_{nx1}} = \underline{r}_{iy_{nx1}}$$

Where $\underset{\sim}{R}_{nxn}$ represent the correlation matrix, $\underline{P}_{iy_{nx1}}$ the vector of path coefficients, $\underline{r}_{iy_{nx1}}$ the correlations of x_i with y. Solving these systems of equations the direct effect will be obtained.

The residual effect can be obtained as follows.

$$r_{yy} = 1 = \sum_{i=1}^{n} P_{iy} r_{iy} + P_{uy} r_{uy}$$

$$= \sum_{i=1}^{n} P_{iy} r_{iy} + r_{uy}^2 = 1$$

$$h^2 = r_{uy}^2 = 1 - \sum_{i=1}^{n} P_{iy} r_{iy}$$

$$h = \sqrt{1 - \sum_{i=1}^{n} P_{iy} r_{iy}}$$

where h is the residual effect and h^2 measure the degree of determination of Y by residual factors and $\sum_i r_{iy} P_{iy}$ measure the degree of determination of Y by the endogenous variables.

If the correlation coefficient is positive, direct effect is positive and indirect effects are negligible, then the direct selection of that trait will be effective.

If the correlation coefficient is positive but direct effect is negative or negligible the indirect effect seems to be the cause of correlation. In such situations the indirect causal factors are to be considered simultaneously.

In some cases correlation coefficients may be negative, but the direct effect is positive and high. In such situations a restricted simultaneous selection model is to be followed. Restrictions are to be imposed in order to nullify the undesirable indirect effect so as to make use of the direct effect.

Rules for translating equations into a path diagram are to draw an arrow from each cause to effect and between two purely exogenous variables, draw a curved line with arrow heads at each ends. The path diagram contains all the information of a system of equations, but for many models the diagram is easier to comprehend.

RESULTS

4. RESULTS

4.1. Estimation of leaf area

The different models fitted for estimating leaf area of sugarcane with corresponding R^2 values are given in table 4.1.1.

'x' is the product of maximum width and maximum length of the leaf. All the above mentioned six models were almost equally efficient in the estimation of leaf area. The coefficients of determinations were comparitively low for variety CO-62175. Considering the simplicity and convenience of calculations the equation y = bx was selected for the estimation of leaf area of sugarcane.

4.2. Pre-harvest forecasting of sugarcane yield - Method of multiple regression - Plant wise approach.

The inter-correlations among the different characters studied and those with yield were calculated for each month separately. The matrix of correlation coefficients, the regression equations fitted in each month for the two varieties and the value of coefficient of determination are given in tables 4.2(a) and 4.2(b) respectively. As mentioned in section 3.1. four types of transformations were tried on the experimental data and the regression equations fitted applying these transformations are denoted by letters a,b,c and d where a. $y=a_1+\sum_i \beta_i \log x_i$ b. $y= a_2+\sum_i \beta_i \frac{1}{x_i}$
c. $y = a_3+\sum_i \beta_i \sqrt{x_i}$ d. $\log y = a_4+\sum_i \beta_i \log x_i$

Table 4.1.1. Models for estimating leaf area of sugarcane

Sl. No.	Form of model	CO-997 Estimates of parameters a	b	c	R^2	CO-62175 Estimates of parameters a	b	c	R^2
1.	$Y = bx$		0.66		0.83		0.66		0.76
2.	$Y = a+bx$	63.59	0.61		0.83	51.21	0.57		0.75
3.	$\log Y = a+bx$	-2.29	-0.0005		0.81	-2.17	-0.001		0.75
4.	$\log Y = a+b \log x$	0.12	-0.98		0.81	-0.22	0.85		0.76
5.	$Y = a+bx+cx^2$	53.54	-0.82	0.0001	0.83	-42.72	-0.59	0.00003	0.75
6.	$Y = a+b\sqrt{x}+cx$	26.90	-20.29	-0.28	0.83	-587.44	2896	1.37	0.75

Table 4.2(a)1.1. Zero order correlation matrix of biometric characters and yield for CO-997 in fifth month of plant growth.

Characters	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No.of leaves/ cane x_6	Total leaf area/ cane x_7	Cane yield Y
x_1	1	0.445*	0.437*	-0.159	0.355*	0.358*	0.476*	0.472*
x_2		1	0.522*	-0.040	0.509*	0.210*	0.533*	0.526*
x_3			1	-0.081	0.917*	0.258*	0.849*	0.394*
x_4				1	0.286*	-0.136	0.179	-0.049
x_5					1	0.204*	0.894*	0.361*
x_6						1	0.546*	0.226*
x_7							1	0.408*
Y								1

Table 4.2(a)1.2. Zero order correlation matrix of biometric characters and yield for CO-62175 in fifth month of plant growth.

Characters	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No. of leaves/ cane x_6	Total leaf area/ cane x_7	Cane yield Y
x_1	1	0.489*	0.505*	0.469*	0.575*	0.712*	0.744*	0.545*
x_2		1	0.546*	0.454*	0.581*	0.404*	0.570*	0.590*
x_3			1	0.521*	0.937*	0.397*	0.807*	0.551*
x_4				1	0.777*	0.312*	0.654*	0.503*
x_5					1	0.426*	0.862*	0.612*
x_6						1	0.807*	0.413*
x_7							1	0.609*
Y								1

*Significant at 5% level P ($< .05$)

Table 4.2(a)2.1. Zero order correlation matrix of biometric characters and yield for CO-997 in sixth month of plant growth.

Characters	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No.of leaves/cane x_6	Total leaf area/cane x_7	Cane yield Y
x_1	1	0.023	0.215*	-0.095	0.153	0.079	0.151	0.375*
x_2		1	-0.016	0.104	0.034	0.037	0.046	0.055
x_3			1	0.154	0.842*	0.139	0.750*	0.420*
x_4				1	0.610*	-0.021	0.500*	0.115
x_5					1	0.098	0.862*	0.387*
x_6						1	0.577*	0.438*
x_7							1	0.543*
Y								1

Table 4.2(a)2.2. Zero order correlation matrix of biometric characters and yield for CO-62175 in sixth month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leef x_4	Area of 3rd leaf x_5	No.of leaves/cane x_6	Total leaf area/cane x_7	Cane yield Y
x_1	1	0.491*	0.464*	0.272*	0.458*	0.449*	0.554*	0.630*
x_2		1	0.585*	0.362*	0.593*	0.324*	0.581*	0.620*
x_3			1	0.426*	0.931*	0.262*	0.794*	0.542*
x_4				1	0.724*	0.039	0.811*	0.407*
x_5					1	0.210*	0.808*	0.574*
x_6						1	0.734*	0.415*
x_7							1	0.639*
Y								1

*Significant at 5% level $P (< .05)$

Table 4.2(a)3.1. Zero order correlation matrix of biometric characters and yield for CO-997 in seventh month of plant growth.

Character	Height	Girth	Width of 3rd leaf	Length of 3rd leaf	Area of 3rd leaf	No.of leaves/ cane	Total leaf area/ cane	Cane yield
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	Y
x_1	1	0.301*	0.153	-0.032	0.104	0.182	0.188	0.525*
x_2		1	0.277*	0.091	0.257*	0.265*	0.350*	0.591*
x_3			1	0.179	0.869*	0.300*	0.796*	0.478*
x_4				1	0.638*	-0.206*	0.317*	0.069
x_5					1	0.121	0.774*	0.394*
x_6						1	0.715*	0.522*
x_7							1	0.611*
Y								1

Table 4.2(a)3.2. Zero order correlation matrix of biometric characters and yield for CO-62175 in seventh month of plant growth.

Character	Height	Girth	Width of 3rd leaf	Length of 3rd leaf	Area of 3rd leaf	No.of leaves/ cane	Total leaf area/ cane	Cane yield
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	Y
x_1	1	0.598*	0.354*	0.212*	0.337*	0.061	0.240*	0.758*
x_2		1	0.495*	0.222*	0.488*	0.162	0.443*	0.696*
x_3			1	0.479*	0.865*	0.126	0.743*	0.581*
x_4				1	0.834*	0.233*	0.740*	0.449*
x_5					1	0.222*	0.885*	0.620*
x_6						1	0.629*	0.155
x_7							1	0.569*
Y								1

*Significant at 5% level P (< .05)

Table 4.2(a)4.1. Zero order correlation matrix of biometric characters and yield for CO-997 in eighth month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No.of leaves/ cane x_6	Total leaf area/ cane x_7	Cane yield Y
x_1	1	0.239*	0.227*	-0.039	0.155	0.050	0.139	0.551*
x_2		1	0.276*	0.004	0.196	0.220*	0.269*	0.560*
x_3			1	0.269*	0.798*	0.159	0.698*	0.431*
x_4				1	0.766*	0.109	0.614*	0.041
x_5					1	0.154	0.826*	0.325*
x_6						1	0.672*	0.408*
x_7							1	0.481*
Y								1

Table 4.2(a)4.2. Zero order correlation matrix of biometric characters and yield for CO-62175 in eigth month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No.of leaves/ cane x_6	Total leaf area/ cane x_7	Cane yield Y
x_1	1	0.529*	0.431*	0.285*	0.441*	0.078	0.387*	0.743*
x_2		1	0.453*	0.250*	0.424*	0.090	0.429*	0.666*
x_3			1	0.510*	0.897*	0.022	0.776*	0.608*
x_4				1	0.796*	-0.143	0.565*	0.363*
x_5					1	-0.035	0.808*	0.590*
x_6						1	0.512*	0.101
x_7							1	0.557*
Y								1

*Significant at 5% level P ($<$ .05)

Table 4.2(a)5.1. Zero order correlation matrix of biometric characters for CO-997 in ninth month of plant growth.

Character	Height	Girth	Width of 3rd leaf	Length of 3rd leaf	Area of 3rd leaf	No. of leaves/cane	Total leaf area/cane	Cane yield
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	Y
x_1	1	0.092	0.240*	-0.073	0.059	0.018	0.077	0.479*
x_2		1	0.109	-0.004	0.096	0.249*	0.216*	0.547*
x_3			1	0.348*	0.266*	0.326*	0.714*	0.336*
x_4				1	0.268*	0.208*	0.584*	0.143
x_5					1	0.123	0.288*	0.127
x_6						1	0.805*	0.512*
x_7							1	0.521*
Y								1

Table 4.2(a)5.2. Zero order correlation matrix of biometric characters and yield for CO-62175 in ninth month of plant growth.

Character	Height	Girth	Width of 3rd leaf	Length of 3rd leaf	Area of 3rd leaf	No. of leaves/cane	Total leaf area/cane	Cane yield
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	Y
x_1	1	0.561*	0.507*	0.261*	0.446*	0.107	0.389*	0.808*
x_2		1	0.456*	0.307*	0.406*	0.116	0.413*	0.769*
x_3			1	0.571*	0.849*	0.053	0.680*	0.585*
x_4				1	0.801*	-0.161	0.529*	0.384*
x_5					1	-0.031	0.748*	0.557*
x_6						1	0.606*	0.173
x_7							1	0.532*
Y								1

*Significant at 5% level $P (< .05)$

Table 4.2(b)1.1. Regression equations fitted for CO-997 in fifth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -478.382+4.834x_1+103.089x_2+68.01x_3+1.466x_4 -0.475x_5+5.43x_6+0.001x_7$	0.355**
2.	$Y = -241.445+4.84x_1+101.21x_2+22.426x_3+5.223x_6$ (S)	0.353**
3.	$Y = -246.674+4.855x_1+100.999x_2+19.536x_3+0.03x_4 +5.264x_6$ (S)	0.353**
4.	$Y = -244.691+4.994x_1+101.657x_2+0.188x_5 +5.701x_6$ (S)	0.353**
5.	$Y = -179.634+4.956x_1+101.425x_2+0.175x_7$ (S)	0.352**
6.	$Y = -206.619+5.03x_1+101.353x_2+23.54x_3$ (S)	0.352**
7.	$Y = 317.261+6.576x_1+0.047x_7$	0.266**
8.	$Y = 206.471+6.57x_1+61.09x_3+5.834x_6$	0.267**
a.	$Y = -3596.225+141.861x_1+56.991x_2+79.539x_3 -9.647x_4+172x_5+373.6x_6-90.8x_7$	0.346**
b.	$Y = 1455.171-1927.611x_1-2470.684x_2-507.803x_3 +12777.753x_4-9708.262x_5-2887.764x_6 +12996.929x_7$	0.373**
c.	$Y = -1377.347+71.394x_1+521.129x_2+96.357x_3 +6.423x_4-0.407x_5+25.879x_6+0.111x_7$	0.352**
d.	$Y = 7.821+0.319x_1+0.826x_2-0.215x_3-0.232x_4 +0.306x_5+0.029x_6+0.004x_7$	0.340**

**Indicates significance at 1% level P ($<$ .01)

Table 4.2(b)1.2. Regression equations fitted for CO-62175 in fifth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -446.404 + 4.052x_1 + 90.80x_2 - 18.31x_3 + 1.20x_4 + 0.885x_5 - 19.24x_6 + 0.069x_7$	0.500**
2.	$Y = -991.45 + 4.134x_1 + 88.726x_2 + 121.013x_3 + 3.961x_4$ (S)	0.498**
3.	$Y = -751.607 + 3.171x_1 + 96.823x_2 + 3.653x_4 + 0.063x_7$ (S)	0.488**
4.	$Y = -568.929 + 3.738x_1 + 95.04x_2 + 105.69x_3 + 0.045x_7$ (S)	0.484**
5.	$Y = -652.899 + 5.084x_1 + 97.482x_2 + 153.477x_3 - 3.493x_6$ (S)	0.480**
6.	$Y = -666.844 + 4.82x_1 + 97.091x_2 + 153.134x_3$ (S)	0.480**
7.	$Y = -388.771 + 3.014x_1 + 102.443x_2 + 0.095x_7$ (S)	0.475**
8.	$Y = -309.642 + 6.494x_1 + 236.445x_3 + 0.883x_6$	0.409**
9.	$Y = 589.858 + 3.952x_1 + 0.148x_7$	0.331**
a.	$Y = 19119.059 + 78.276x_1 + 268.04x_2 - 18478.245x_3 - 18502.271x_4 + 18379.113x_5 - 259.706x_6 + 279.204x_7$	0.454**
b.	$Y = 5762.09 - 3251.826x_1 - 4484.942x_2 - 8122.177x_3 - 383977.323x_4 + 467376.293x_5 - 4437.5x_6 + 900939.176x_7$	0.450**
c.	$Y = 5398.749 + 44.925x_1 + 470.053x_2 - 4251.957x_3 - 591.575x_4 + 485.938x_5 - 94.067x_6 + 6.942x_7$	0.487**
d.	$Y = 72.868 + 0.253x_1 + 0.635x_2 - 75.04x_3 - 74.977x_4 + 74.712x_5 - 0.560x_6 + 0.847x_7$	0.335**

**Significant at 1% level P ($<.01$)

Table 4.2(b)2.1. Regression equations for CO-997 in sixth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = 460.139+4.097x_1+1.244x_2+105.421x_3+1.294x_4-2.702x_5-47.251x_6+0.279x_7$	0.427**
2.	$Y = -614.966+3.73x_1+1.658x_2+109.873x_3+70.77x_6$ (S)	0.399**
3.	$Y = -604.76+3.744x_1+109.503x_3+70.959x_6$ (S)	0.397**
4.	$Y = -540.626+3.999x_1+0.807x_5+73.131x_6$ (S)	0.398**
5.	$Y = 54.752+4.019x_1+0.111x_7$ (S)	0.383**
6.	$Y = 174.957+4.013x_1+1.059x_2+0.111x_7$ (S)	0.384**
7.	$Y = -9.27+3.982x_1+2.235x_2+126.85x_3$	0.264**
a.	$Y = -3596.225+141.861x_1+56.991x_2+79.539x_3-9.647x_4+172x_5+373.6x_6-90.8x_7$	0.409**
b.	$Y = 3642.597-21729.061x_1-5229.521x_2-1657.684x_3-74871.427x_4-1422.685x_5-8131.048x_6+1120489.872x_7$	0.462**
c.	$Y = 2121.052+75.206x_1+29.693x_2+325.325x_3+11.751x_4-205.601x_5-977.602x_6+69.123x_7$	0.425**
d.	$Y = 2.549+0.586x_1+0.169x_2+0.238x_3+0.059x_4-0.259x_5+0.374x_6+0.440x_7$	0.399**

**Significant at 1% level P ($< .01$)

Table 4.2(b)2.2. Regression equations fitted for CO-62175 in sixth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -208.253+5.74x_1+91.801x_2+28.623x_3+4.032x_4 -3.532x_5-93.983x_6+0.430x_7$	0.586**
2.	$Y = -335.23+5.214x_1+86.275x_2+72.387x_3+4.611x_4 +27.542x_6$ (S)	0.571**
3.	$Y = -670.507+5.565x_1+100.123x_2+5.39x_4+28.462x_6$ (S)	0.563**
4.	$Y = -781.492+6.093x_1+117.943x_2+23.046x_6$	0.534**
5.	$Y = 313.919+6.310x_1+0.179x_7$	0.518**
6.	$Y = -1010.47+120.855x_2+153.589x_3+43.942x_6$	0.470**
7.	$Y = -893.067+5.486x_1+94.503x_2+103.12x_3+22.819x_6$	0.551**
8.	$Y = -422.823+3.175x_1+87.606x_2+0.123x_7$ (S)	0.571**
9.	$Y = -3260.9-5.599x_1+170.045x_2+972.177x_3$	0.403**
a.	$Y = -248872.77+153.851x_1+305.523x_2+255191.019x_3 +255303.979x_4+53.658x_5+255239.167x_6 -255159.688x_7$	0.532**
b.	$Y = 4456.763-22112.591x_1-5479.72x_2-4148.74x_3 -200321.249x_4+209363.675x_5-4816.034x_6 +837000.286x_7$	0.497**
c.	$Y = 2599.417+98.284x_1+530.478x_2-520.640x_3 -13.124x_4-209.654x_5-1437.791x_6+95.521x_7$	0.573**
d.	$Y = -1043.079+0.566x_1+0.969x_2+1083.538x_3 +1084.439x_4-2.059x_5+1081.621x_6-1081.394x_7$	0.511**

**Significant at 1% level P ($<$ .01)

Table 4.2(b)3.1. Regression equations fitted for CO-997 in seventh month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -4650.708+5.247x_1+88.047x_2+982.413x_3 +22.937x_4-9.157x_5-3.641x_6+0.111x_7$	0.677**
2.	$Y = 78.712+5.312x_1+90.389x_2+0.109x_7$ (S)	0.642**
3.	$Y = -1464.312x_5+5.111x_1+92.527x_2+115.312x_3 +54.052x_6$	0.615**
4.	$Y = -1240.993+5.544x_1+106.133x_2+149.104x_3$	0.572**
5.	$Y = -1180.223+6.408x_1+146.753x_3+64.305x_6$	0.549**
6.	$Y = -554.774+6.588x_1+0.135x_7$	0.548**
7.	$Y = -805.502+5.915x_1+127.422x_2$	0.482**

**Significant at 1% level P (< .01)

Table 4.2(b)3.2. Regression equations fitted for CO-62175 in seventh month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = 577.621+6.782x_1+70.969x_2-110.633x_3-2.298x_4+2.263x_5-1.768x_6+0.081x_7$	0.769**
2.	$Y = -822.361+7.117x_1+68.338x_2+0.160x_7$ (S)	0.760**
3.	$Y = -1298.886+6.411x_1+84.465x_2+112.931x_3+3.375x_4$	0.753**
4.	$Y = -1216.034+7.011x_1+70.268x_2+168.595x_3+28.059x_6$	0.735**
5.	$Y = -1032.488+8.48x_1+211.212x_3+36.326x_6$	0.708**
6.	$Y = -1044.620+6.621x_1+80.363x_2+174.646x_3$	0.722**
7.	$Y = -847.41+6.943x_1+118.239x_2$	0.667**
8.	$Y = -374.462+8.602x_1+118.239x_2$	0.645**
9.	$Y = -4896.192+6.419x_1+78.492x_2+1.698x_5$ (S)	0.761**

**Significant at 1 per cent level P ($<$ .01)

Table 4.2(b)4.1. Regression equations fitted for CO-997 in eighth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -381.292+5.535x_1+99.704x_2-15.140x_3$ $-2.503x_4-0.751x_5-33.015x_6+0.183x_7$	0.633**
2.	$Y = -1479.5+5.622x_1+100.881x_2+76.044x_3$ $+43.961x_6$ (S)	0.619**
3.	$Y = -1111.565+5.722x_1+109.021x_2+0.066x_7$ (S)	0.595**
4.	$Y = -1134.137+6.491x_1+103.314x_3+53.638x_6$	0.516**
5.	$Y = -1074.904+6.074x_1+132.561x_2$	0.498**
6.	$Y = 614.423+6.773x_1+0.084x_7$	0.471**
7.	$Y = -1249.947+5.538x_1+116.825x_2+87.913x_3$	0.447**

**Significant at 1% level P (< .01)

Table 4.2(b)4.2. Regression equations fitted for CO-62175 in eighth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -844.726+5.124x_1+93.113x_2+65.501x_3-0.562x_4$ $+0.034x_5-10.050x_6+0.081x_7$	0.716**
2.	$Y = -1064.146+5.113x_1+97.150x_2+135.967x_3$ (S)	0.711**
3.	$Y = -998.719+5.899x_1+125.079x_2$	0.656**
4.	$Y = -2502.262+5.329x_1+100.889x_2+0.905x_7$ (S)	0.699**
5.	$Y = -1069.825+5.094x_1+96.306x_2+136.559x_3$ $+4.227x_6$ (S)	0.712**
6.	$Y = -1523.632+6.381x_1+177.003x_3+6.208x_6$	0.656**
7.	$Y = -391.881+6.748x_1+0.121x_7$	0.637**
8.	$Y = -499.811+9.121x_1$	0.601**
9.	$Y = -943.218+5.076x_1+103.087x_2+0.974x_5$ (S)	0.704**

**Significant at 1% level P ($<$ .01)

Table 4.2(b)5.1. Regression equations fitted for CO-997 in ninth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -680.146+4.904x_1+108.879x_2-68.879x_3$ $-1.806x_4-0.015x_5-1.462x_6-0.092x_7$	0.663**
2.	$Y = -1195.981+4.669x_1+111.240x_2+20.628x_3$ $+38.557x_6$	0.641**
3.	$Y = -1015.126+4.563x_1+115.986x_2+0.057x_7$ (S)	0.635**
4.	$Y = -1132.072+4.330x_1+134.392x_2+58.709x_3$	0.518**
5.	$Y = -532.698+5.059x_1+21.668x_3+48.612x_6$	0.487**
6.	$Y = -270.015+4.923x_1+0.069x_7$	0.465**
7.	$Y = -880.117+4.999x_1+112.461x_2-63.583x_3$ $+0.778x_7$ (S)	0.654**

**Significant at 1% level P ($<$ .01)

Table 4.2(b)5.2. Regression equations fitted for CO-62175 in ninth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -1181.538+4.922x_1+115.727x_2-5.117x_3$ $+0.191x_4+1.135x_5+16.484x_6-0.043x_7$	0.826**
2.	$Y = -1074.184+4.971x_1+113.478x_2+0.704x_5$ $+7.610x_6$ (S)	0.825**
3.	$Y = -983.567+5.201x_1+113.362x_2+0.049x_7$ (S)	0.819**
4.	$Y = -1138.11+5.019x_1+114.943x_2+70.828x_3$ $+6.267x_6$ (S)	0.817**
5.	$Y = -1081.42+5.064x_1+116.474x_2+70.099x_3$ (S)	0.812**
6.	$Y =-0.1030.774+5.576x_1+125.855x_2$	0.798**
7.	$Y = -281.85+7.191x_1+0.079x_7$	0.708**
8.	$Y = -681.424+6.89x_1+120.452x_3+8.634x_6$	0.702**
9.	$Y = -1021.473+5.048x_1+115.769x_2+0.666x_5$ (S)	0.819**

**Significant at 1% level P ($\angle$.01)

A high positive significant correlation was found between height and cane yield in all periods of study for the two varieties and correlation coefficient lies in the range 0.375 to 0.808. Girth of the cane was highly correlated with yield in all periods except for CO-997 in the sixth month of plant growth. Product of number of leaves and area of third leaf, and width of third leaf are correlated with yield during all periods of plant growth. Length of third leaf was not correlated with yield of variety CO-997. But for CO-62175 length of third leaf was significantly and positively correlated with yield. Number of leaves per cane was positively correlated with cane yield except for the variety CO-62175 in the last three months. The biometric characters also showed high inter-correlations among themselves.

The value of the coefficient of determination of the fitted equations were in the range from 0.355 to 0.826. The predictability coefficient was relatively high in the later stages of plant growth than at early stages.

All the equations gave better fit to variety CO-62175 than the other variety. From the set of available regression equations listed for each month a single equation was identified and selected to be the best for prediction. The adequacy of the selected equations with fewer number of variables was tested against the full model by 'F' test and the differences were found to

Table 4.2(b)6. Forecasting models selected for CO-997 and CO-62175 in different months.

Month	Variety	Constant	Coefficient of x_1	Coefficient of x_2	Coefficient of x_3	Coefficient of x_5	Coefficient of x_6	Coefficient of x_7	R^2
		Estimates of parameters							
V	CO-997	-206.619	5.030	101.353	23.54				0.352**
	CO-62175	-668.844	4.820	97.091	153.134				0.480**
VI	CO-997	54.752	4.019					0.111	0.383**
	CO-62175	-422.823	3.175	87.606				0.123	0.571**
VII	CO-997	78.712	5.312	90.389				0.109	0.642**
	CO-62175	-4896.192	6.419	78.492		1.698			0.761**
VIII	CO-997	-1111.565	5.722	109.021				0.066	0.595**
		-1479.500	5.622	100.881	76.044		43.961		0.619**
	CO-62175	-943.218	5.076	103.087		0.974			0.595**
IX	CO-997	-1015.126	4.563	115.986				0.057	0.635**
	CO-62175	-1021.473	5.048	115.789		0.666			0.819**

**Significant at 1 per cent level P (< .01) be nonsignificant. The selected models with corresponding R^2 values for each of the two varieties in different months are given in table 4.2(b)6.

Yields of canes estimated by the above mentioned prediction equations when multiplied by the number of canes in different plots will give an estimate of plot yield of sugarcane in respective months for the two varieties with a sufficient degree of accuracy.

4.3. Pre-harvest forecasting of sugarcane yield - Method of multiple regression - Plot wise approach.

The intercorrelations among the characters themselves, and with plot yield were calculated for each month separately for the two varieties. The matrix of correlation coefficients the regression equations fitted and corresponding R^2 values for each month are given in tables 4.3(a) and 4.3(b) respectively. The regression equations fitted after transformations as mentioned in section 3.2 were represented by letters a, b, c and d. Where a. $y = a_0 + \Sigma b_i \log x_i$

b $y = a_1 + \Sigma \frac{b_i}{x_i}$ c. $y = a_2 + \Sigma b_i \sqrt{x_i}$

d. $\log y = a_3 + \Sigma b_i \log x_i$

It can be seen from tables 4.3(a) that cane yield was highly correlated with height of the cane, girth of the cane and the product of area of third leaf and the number of leaves in the plot in all periods of plant growth. During fifth and sixth months plot yield was positively correlated with width of third leaf. In later stages of plant growth leaf dimensions viz., leaf length and width were not found correlated with plot yield for CO-997, but in the case of CO-62175 leaf width and area of third leaf showed significant correlations with plot yield.

Number of leaves per plot was significantly correlated with plot yield except for CO-997 in the ninth month

Table 4.3(a)1.1. Zero order correlation matrix of biometric characters and yield for CO-997 in fifth month of plant growth.

Character	Height	Girth	Width of 3rd leaf	Length of 3rd leaf	Area of 3rd leaf	No.of canes/ tillers	No.of leaves/ plot	Total leaf area/ plot	Plot yield
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	Y
x_1	1	0.443*	0.527*	-0.141	0.479*	0.184	0.298	0.445*	0.462*
x_2		1	0.578*	0.047	0.587*	0.129	0.157	0.389*	0.428*
x_3			1	-0.133	0.933*	0.107	0.169	0.566*	0.434*
x_4				1	0.230	-0.218	-0.142	-0.025	-0.119
x_5					1	0.029	0.122	0.569*	0.378*
x_6						1	0.868*	0.724*	0.685*
x_7							1	0.877*	0.719*
x_8								1	0.771*
Y									1

Table 4.3(a)1.2. Zero order correlation matrix of biometric characters and yield for CO-62175 in fifth month of plant growth.

Character	Height	Girth	Width of 3rd leaf	Length of 3rd leaf	Area of 3rd leaf	No.of canes/ tillers/ plot	No. of leaves/ plot	Total leaf area/ plot	Plot yield
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	Y
x_1	1	0.501*	0.685*	0.640*	0.748*	0.308*	0.561*	0.716*	0.751*
x_2		1	0.506*	0.331*	0.524*	0.099	0.231	0.366*	0.406*
x_3			1	0.640*	0.959*	0.295	0.387*	0.714*	0.641*
x_4				1	0.761*	0.220	0.529*	0.713*	0.570*
x_5					1	0.228	0.432*	0.761*	0.661*
x_6						1	0.656*	0.557*	0.430*
x_7							1	0.905*	0.776*
x_8								1	0.849*
Y									1

*Significant at 5% level P (< .05)

Table 4.3(a)2.1. Zero order correlation matrix of biometric characters and yield for CO-997 in sixth month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No. of canes/tillers/plot x_6	No. of leaves/plot x_7	Total leaf area/plot x_8	Plot yield Y
x_1	1	0.513*	0.420*	-0.241	-0.203	0.339*	0.420*	0.409*	0.521*
x_2		1	0.744*	0.178	0.676*	0.213	0.425*	0.596*	0.567*
x_3			1	0.157	0.868*	0.146	0.345*	0.612*	0.386*
x_4				1	0.624*	-0.133	-0.065	0.206	0.031
x_5					1	0.054	0.244*	0.592*	0.325*
x_6						1	0.901*	0.762*	0.819*
x_7							1	0.922*	0.893*
x_8								1	0.855*
Y									1

Table 4.3(a)2.2. Zero order correlation matrix of biometric characters and yield for CO-62175 in sixth month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No. of canes/tillers/plot x_6	No. of leaves/plot x_7	Total leaf area/plot x_8	Plot yield Y
x_1	1	0.623*	0.613*	0.384*	0.614*	0.288*	0.498*	0.661*	0.657*
x_2		1	0.644*	0.282*	0.610*	0.133	0.330*	0.511*	0.512*
x_3			1	0.400	0.935*	0.005	0.244*	0.574*	0.577*
x_4				1	0.696*	-0.236	-0.054	0.214	0.217
x_5					1	-0.099	0.159	0.522*	0.526*
x_6						1	0.864*	0.702*	0.619*
x_7							1	0.922*	0.816*
x_8								1	0.904*
Y									1

*Significant at 5% level P ($< .05$)

Table 4.3(a)3.1. Zero order correlation matrix of biometric characters and yield for CO-997 in seventh month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No. of canes x_6	No. of tillers x_7	No. of leaves x_8	Total leaf area/plot x_9	Plot yield Y
x_1	1	0.424*	0.281*	0.131	0.303*	0.350*	0.319*	0.420*	0.505*	0.518*
x_2		1	0.220	0.058	0.194	0.209	0.448*	0.358*	0.396*	0.510*
x_3			1	0.084	0.855*	-0.133	-0.051	-0.087	0.200	0.077
x_4				1	0.585*	0.119	-0.074	0.095	0.289*	0.101
x_5					1	-0.042	-0.096	-0.019	0.316*	0.112
x_6						1	0.317*	0.895*	0.832*	0.826*
x_7							1	0.377*	0.326*	0.495*
x_8								1	0.941*	0.890*
x_9									1	0.878*
Y										1

*Significant at 5% level P (<.05)

Table 4.3(a)3.2. Zero order correlation matrix of biometric characters and yield for CO-62175 in seventh month of plant growth.

Character	Height	Girth	Width of 3rd leaf	Length of 3rd leaf	Area of 3rd leaf	No,of canes	No.of tillers	No.of leaves	Total leaf area	Plot yield
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	Y
x_1	1	0.630*	0.429*	0.317*	0.413*	0.241	0.474*	0.229	0.460*	0.753*
x_2		1	0.537*	0.327*	0.481*	0.087	0.398*	0.096	0.387*	0.514*
x_3			1	0.653*	0.921*	-0.138	0.318*	-0.198	0.406*	0.322*
x_4				1	0.890*	-0.214	0.180	-0.203	0.395*	0.226
x_5					1	-0.203	0.279*	-0.236	0.427*	0.297*
x_6						1	0.067	0.889*	0.673*	0.585*
x_7							1	0.175	0.324*	0.539*
x_8								1	0.767*	0.644*
x_9									1	0.795*
Y										1

*Significant at 5% level P (< .05)

Table 4.3(a)4.1. Zero order correlation matrix of biometric characters and yield for CO-997 in eighth month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No.of canes x_6	No.of tillers x_7	No.of leaves x_8	Total leaf area x_9	Plot yield Y
x_1	1	0.318*	0.326*	0.044	0.248	0.211	0.159	0.150	0.254	0.420*
x_2		1	0.435*	0.106	0.362*	0.248	0.255	0.351*	0.405*	0.530*
x_3			1	0.248	0.829*	-0.105	0.076	0.004	0.269*	0.152
x_4				1	0.747*	0.079	0.178	0.054	0.290*	0.100
x_5					1	-0.025	0.158	0.032	0.351*	0.163
x_6						1	0.413*	0.641*	0.636*	0.873*
x_7							1	0.384*	0.394*	0.492*
x_8								1	0.889*	0.672*
x_9									1	0.713*
Y										1

*Significant at 5% level P (< .05)

Table 4.3(a)4.2. Zero order correlation matrix of biometric characters and yield for CO-62175 in eighth month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No. of canes x_6	No. of tillers x_7	No. of leaves x_8	Total leaf area/plot x_9	Plot yield Y
x_1	1	0.645*	0.541*	0.518*	0.599*	0.201	0.247	0.206	0.522*	0.694*
x_2		1	0.445*	0.248	0.408*	0.175	0.326*	0.158	0.360*	0.471*
x_3			1	0.673*	0.915*	-0.145	0.306	-0.047	0.540*	0.376*
x_4				1	0.904*	-0.232	0.116	-0.192	0.419*	0.237
x_5					1	-0.177	0.227	-0.098	0.554*	0.374*
x_6						1	0.279	0.898	0.629*	0.593*
x_7							1	-0.164	0.017	0.164
x_8								1	0.760*	0.635*
x_9									1	0.769*
Y										1

*Significant at 5% level P ($< .05$)

Table 4.3(a)5.1. Zero order correlation matrix of biometric characters and yield for CO-997 in ninth month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No.of canes x_6	No.of tillers x_7	No.of leaves x_8	Total leaf area/plot x_9	Plot yield Y
x_1	1	0.052	0.213	-0.150	0.049	0.191	0.166	0.265	0.260	0.209
x_2		1	0.132	0.015	0.093	0.217	0.239	0.447*	0.426*	0.247
x_3			1	0.254	0.797*	-0.120	0.009	0.076	0.403*	-0.064
x_4				1	0.783*	-0.090	-0.090	0.062	0.370*	-0.007
x_5					1	-0.177	-0.055	0.086	0.492*	-0.053
x_6						1	0.649*	0.822*	0.635*	0.345*
x_7							1	0.620*	0.490*	0.197
x_8								1	0.899*	0.419*
x_9									1	0.353*
Y										1

*Significant at 5% level P ($< .05$)

Table 4.3(a)5.2. Zero order correlation matrix of biometric characters and yield for CO-62175 in ninth month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No.of canes x_6	No.of tillers x_7	No.of leaves x_8	Total leaf area x_9	Plot yield Y
x_1	1	0.650*	0.613*	0.510*	0.614*	0.212	0.187	0.387*	0.664*	0.673*
x_2		1	0.450*	0.375*	0.357*	0.066	0.149	0.168	0.351*	0.484*
x_3			1	0.590*	0.769*	-0.071	0.185	-0.048	0.455*	0.318*
x_4				1	0.789*	0.124	0.372*	0.208	0.645*	0.459*
x_5					1	0.053	0.033	0.066	0.681*	0.345*
x_6						1	-0.053	0.752*	0.544*	0.692*
x_7							1	0.167	0.115	0.118
x_8								1	0.759*	0.712*
x_9									1	0.770*
Y										1

*Significant at 5% level P ($< .05$)

Table 4.3(b)1.1. Regression equations fitted for CO-997 in fifth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -378.865 + 0.064x_1 + 0.913x_2 + 19.158x_3 + 0.491x_4 - 0.193x_5 + 0.269x_6 - 0.017x_7 + 0.0002x_8$	0.729**
2.	$Y = -11.556 + 0.046x_1 + 0.834x_2 + 0.761\ x_3 + 0.312x_6 + 0.041x_7$ (S)	0.686**
3.	$Y = -6.716 + 0.046x_1 + 0.81x_2 + 0.346x_6 + 0.0001x_8$ (S)	0.680**
4.	$Y = -9.402 + 0.070x_1 + 0.961x_2 + 0.144x_6 + 0.082x_7$ (S)	0.629**
5.	$Y = -2.617 + 0.033x_1 + 0.566x_2 + 0.0002x_8$(S)	0.623**
6.	$Y = -10.379 + 0.072x_1 + 1.174x_2 + 0.584x_6$ (S)	0.633**
7.	$Y = -6.258 + 0.064x_1 + 1.141x_3 + 0.593x_6$ (S)	0.633**
8.	$Y = 30.792 + 0.128x_1 - 5.729x_2 + 3.642x_3 + 0.643x_6$	0.509**
a.	$Y = -144.855 + 0.992x_1 + 2.355x_2 + 101.386x_3 + 101.145x_4 - 3646.429x_5 + 1.491x_6 - 3545.109x_7 + 3546.395x_8$	0.647**
b.	$Y = -5.855 - 237.035x_1 - 36.368x_2 + 184.532x_3 + 7454.087x_4 - 24752.439x_5 - 2.012x_6 - 38.176x_7 - 1015.287x_8$	0.655**
c.	$Y = -287.579 + 0.964x_1 + 4.906x_2 + 141.549x_3 + 22.646x_4 - 14.992x_5 + 1.768x_6 - 2.047x_7 + 0.147x_8$	0.718**
d.	$\text{Log}\ Y = -37.287 + 0.323x_1 + 0.880x_2 + 28.677x_3 + 28.452x_4 - 864.243x_5 + 0.254x_6 - 835.493x_7 + 835.912x_8$	0.739**

**Significance at 1% level P (< .01)

Table 4.3(b)1.2. Regression equations fitted for CO-62175 in fifth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -5.281+0.048x_1+0.099x_2+3.332x_3-0.124x_4$ $-0.018x_5-0.144x_6+0.07x_7+0.0001x_8$	0.799**
2.	$Y = -3.606+0.050x_1+0.022x_2+1.999x_3-0.027x_4$ $-0.125x_6+0.088x_7$(S)	0.794**
3.	$Y = 0.105+0.046x_1+0.105x_2-0.017x_8+0.0002x_6$(S)	0.778**
4.	$Y = -3.606+0.091x_1+0.035x_2+1.308x_3+0.158x_6$	0.630**
5.	$Y = -12.886+0.099x_1-0.04x_2+1.559x_6$	0.594**
6.	$Y = -5.343+0.04x_1+0.07x_2+1.717x_3-0.106x_6$ $+0.081x_7$(S)	0.781**
7.	$Y = -2.824+0.065x_1+0.271x_2-0.077x_6+0.078x_7$(S)	0.746**
8.	$Y = -2835.002+0.087x_1+1.482x_3+0.147x_8$	0.621**
9.	$Y = 0.396+0.206x_1+0.184x_2-0.0003x_8$	0.324**
a.	$Y = -8.113+0.704x_1-0.124x_2+3.809x_3-1.414x_4$ $+316.042x_5-0.757x_6+319.42x_7-316.631x_8$	0.767**
b.	$Y = 33.312-75.118x_1-2.52x_2-15.18x_3+708.432x_4$ $-2780.942x_5+12.021x_6-952.851x_7+148992.975x_8$	0.743**
c.	$Y = -5.134+0.634x_1+0.484x_2+10.677x_3-0.389x_4$ $-1.046x_5-0.956x_6+0.032x_7+0.088x_8$	0.784**
d.	$\log Y = -2.382+0.316x_1+0.111x_2+1.180x_3-0.319x_4$ $+180.463x_5-0.203x_6+181.59x_7-180.734x_8$	0.815**

**Significant at 1% level P ($<$ .01)

Table 4.3(b)2.1. Regression equations fitted for CO-997 in sixth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = 9.008+0.045x_1+1.401x_2-8.127x_3-0.148x_4$ $+0.084x_5+0.279x_6+0.134x_7-0.0002x_8$	0.883**
2.	$Y = -11.552+0.024x_1+1.382x_2+0.347x_6+0.058x_7$(S)	0.863**
3.	$Y = -9.058+0.028x_1+0.995x_2+0.$[illegible]$1x_7$(S)	0.847**
4.	$Y = -14.91$[illegible]$+0.027x_1+1.935x_2+0.743x_6$	0.837**
5.	$Y = -31639.711+0.025x_1+2.291x_2-0.226x_6$ $+0.808x_8$	0.833**
6.	$Y = -16.594+0.027x_1+2.194x_2-0.05x_3+0.74x_6$(S)	0.840**
7.	$Y = -14.849+0.026x_1+1.934x_2+0.742x_6$(S)	0.837**
8.	$Y = -9.731+0.055x_1+1.439x_3+0.758x_6$	0.770
a.	$Y = 349690.42+1.081x_1+5.507x_2-21.279x_3$ $-16.829x_4+1493.285x_5+0.958x_6+1477.113x_7$ $-1474.371x_8$	0.847**
b.	$Y =53.02-171.733x_1-83.389x_2-65.637x_3-4387.421x_4$ $+9114.868x_5-3.853x_6-622.006x_7+69$[illegible]$27.738x_8$	0.780**
c.	$Y = 13.586+0.704x_1+8.344x_2-39.832x_3-4.71x_4$ $+4.317x_5+1.581x_6+3.239x_7-0.092x_8$	0.868**
d.	$\text{Log } Y = -17.120+0.439x_1+1.353x_2+0.995x_3+1.566x_4$ $+28.011x_5+0.297x_6+29.88x_7-29.282x_8$	0.881**

**Significant at 1% level P (< .01)

Table 4.3(b)2.2. Regression equations fitted for CO-62175 in sixth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -11.269+0.012x_1-0.012x_2+1.744x_3+0.02x_4$ $+0.012x_5+0.063x_6+0.133x_7-0.0001x_8$	0.831**
2.	$Y = -2.6+0.014x_1+0.147x_2+0.028x_8+0.0004x_6$(S)	0.824**
3.	$Y = -1.05+0.133x_1+0.139x_2+0.0004x_8$(S)	0.823**
4.	$Y = -8.107+0.016x_1-0.004x_2+2.45x_3+0.007x_6$ $+0.109x_7$(S)	0.823**
5.	$Y = -4.769+0.034x_1+0.504x_2-0.171x_6+0.125x_7$	0.766**
6.	$Y = -10.901+0.036x_1+0.111x_2+2.901x_3+0.722x_6$	0.746**
7.	$Y = -7.629+0.061x_1+0.763x_2+0.629x_6$	0.666**
8.	$Y = -12.442+0.038x_1+3.002x_3+0.723x_6$	0.748**
a.	$Y = -46.883+0.337x_1+0.277x_2-6.416x_3-6.626x_4$ $-1335.279x_5+0.296x_6-1342.104x_7+1344.49x_8$	0.815**
b.	$Y = 8.122-55.132x_1-10.462x_2+50.691x_3+1964.663x_4$ $-5950.461x_5-4.384x_6-274.617x_7+3904.302x_8$	0.781**
c.	$Y = 70.133+0.224x_1+0.232x_2-4.863x_3-1.098x_4$ $+1.538x_5+0.435x_6+1.895x_7-0.021x_8$	0.824**
d.	Log $Y = -16.068+0.187x_1+0.075x_2-1.615x_3-1.711x_4$ $-1227.713x_5+0.163x_6-1229.654x_7+1230.324x_8$	0.851**

**Significant at 1% level P (< .01)

Table 4.3(b)3.1. Regression equations fitted for CO-997 in seventh month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -16.37+0.02x_1+0.702x_2+4.353x_3+0.083x_4 -0.026x_5+0.312x_6+0.146x_7+0.08x_8$	0.878**
2.	$Y = -9.668+0.028x_1+0.862x_2+0.297x_6+0.131x_7 +0.059x_8$ (S)	0.868**
3.	$Y = -8.044+0.014x_1+0.827x_2+0.395x_6+0.174x_7 +0.0001x_9$ (S)	0.873**
4.	$Y = -10.985+0.03x_1+1.034x_2+0.309x_6+0.061x_8$ (S)	0.859**
5.	$Y = -6.7-0.013x_1+0.837x_2+0.0002x_9$	0.803**
6.	$Y = -13.668+0.022x_1+0.961x_2+0.914x_3+0.323x_6 +0.063x_8$	0.866**
7.	$Y = -16.293+0.31x_1+1.349x_2+0.760x_3+0.781x_6$	0.822**
8.	$Y = -9.435+0.031x_1+0.856x_2+0.094x_8$ (S)	0.843**
9.	$Y = -15.346+0.045x_1+1.51x_2+0.867x_6-0.0004x_9$	0.803**
10.	$Y = -21.517+0.128x_1+0.837x_2+0.0002x_9$	0.803**
a.	$Y = -50.673+0.717x_1+2.123x_2+12.33x_3+10.383x_4 -6.415x_5+1.351x_6+0.295x_7+6.427x_8-4.269x_9$	0.878**
c.	$Y = -118.439+0.409x_1+3.125x_2+39.032x_3+5.513x_4 -2.538x_5+1.891x_6+0.604x_7+2.951x_8-0.084x_9$	0.862**
d.	$\text{Log } Y = -66.944+0.323x_1+0.805x_2-2.284x_3-2.421x_4 +2.45x_5+0.321x_6+0.077x_7+0.424x_8+0.127x_9$	0.897**

**Significant at 1% level P (< .01)

Table 4.3(b)3.2. Regression equations fitted for CO-62175 in seventh month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = 24.431+0.052x_1+0.07x_2-0.42x_3-0.014x_4 -0.024x_5+0.0111x_6+0.143x_7-0.086x_8+0.001x_9$	0.883**
2.	$Y = -6.068+0.055x_1+0.029x_2+0.149x_6+0.0002x_9$ (S)	0.836**
3.	$Y = -6.268+0.056x_1+0.261x_2+0.046x_1+0.139x_6 +0.059x_8$ (S)	0.835**
4.	$Y = -2.811+0.056x_1-0.095x_2+0.0003x_9$ (S)	0.823**
5.	$Y = -5.448+0.053x_1+0.249x_2+0.423x_6+0.165x_9$	0.796**
6.	$Y = -13.767+0.064x_1+0.392x_2+0.413x_6$	0.748**
7.	$Y = -8.734+0.06x_1+0.172x_2+1.045x_3+0.232x_6 +0.033x_8$	0.787**
8.	$Y = -8.168+0.061x_1+0.237x_2+0.804x_3+0.437x_6$	0.756**
9.	$Y = -6.086+0.067x_1+0.346x_2+0.051x_8$	0.753**
a.	$Y = -54.103+2.112x_1+0.604x_2-0.966x_3-0.884x_4 +5436.343x_5-0.005x_6+0.188x_7+5436.729x_8 -5434.471x_9$	0.843**
c.	$Y = 14.26+1.039x_1+0.924x_2+1.095x_3+0.022x_4 -2.118x_5+0.535x_6+0.483x_7-3.431x_8 +0.282x_9$	0.870**
d.	$\text{Log } Y = -67209.6+0.785x_1-0.031x_2-0.033x_3-0.087x_4 +1921.209x_5+0.042x_6+0.081x_7+1921.502x_8 -1920.811x_9$	0.883**

**Significant at 1% level P $(< .01)$

Table 4.3(b)4.1. Regression equations fitted for CO-997 in eighth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -22.338+0.036x_1+1.196x_2+2.615x_3+0.049x_4 -0.021x_5+0.654x_6+0.012x_7+0.001x_8+0.00\ \ x_9$	0.903**
2.	$Y = -15.661+0.041x_1+1.444x_2+0.704x_6$ (S)	0.892**
3.	$Y = -15.265+0.042x_1+1.329x_2+0.647x_6+0.010x_7 +0.010x_8$ (S)	0.897**
4.	$Y = -15.561+0.04x_1+1.429x_2+0.698x_6+0.021x_7$ (S)	0.892**
5.	$Y = -15.304+0.042x_1+1.334x_2+0.650x_6+0.01x_8$(S)	0.897**
6.	$Y = -16.406+0.037x_1+1.166x_2+0.655x_3+0.667x_6 +0.01x_8$ (S)	0.901**
7.	$Y = -16.742+0.036x_1+1.282x_2+0.639x_3+0.723x_6$ (S)	0.896**
8.	$Y = -14.911+0.062x_1+1.298x_2+0.052x_8$ (S)	0.609**
9.	$Y = -5.003+0.042x_1+1.624x_2+0.779x_6-0.003x_9$ (S)	0.882**
10.	$Y = -11.354+0.049x_1+1.195x_2+0.0001x_9$	0.613**
a.	$Y = -29.79+2.735x_1+3.296x_2-29.758x_3-30.778x_4 +30.291x_5+2.647x_6+0.093x_7+0.105x_8+0.376x_9$	0.839**
c.	$Y = -85.515+0.975x_1-5.78x_2+25.54x_3+3.69x_4 -2.291x_5+4.262x_6+0.007x_7+0.098x_8+0.005x_9$	0.878**
d.	$\text{Log } Y = 99.739+0.832x_1+1.069x_2-5.244x_3-5.308x_4 +5.134x_5+0.640x_6+0.033x_7+0.011x_8 +0.151x_9$	0.911**

**Significant at 1% level P (< .01)

Table 4.3(b)4.2. Regression equations fitted for CO-62175 in eighth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = 13.89+0.047x_1-0.480x_2-3.985x_3-0.109x_4$ $+0.042x_5+0.167x_6+0.148x_7-0.042x_8+0.0004x_9$	0.835**
2.	$Y = -13.996+0.054x_1-0.051x_2+0.878x_3+0.126x_6$ $+0.045x_8$	0.749**
3.	$Y = -4.241+0.047x_1+0.019x_2+0.23x_6+0.0001x_9$	0.753**
4.	$Y = -6.967+0.053x_1-0.126x_2+1.155x_6+0.439x_3$ (5)	0.767**
5.	$Y = -5.707+0.063x_1+0.017x_2+0.395x_6$	0.696**
6.	$Y = -14.969+0.025x_1+0.508x_2+0.002x_3+0.072x_8$	0.720**
7.	$Y = 1.297+0.041x_1+0.074x_2+0.0002x_9$	0.709**
8.	$Y = -6.781+0.063x_1+0.286x_2+0.039x_8$	0.629**
a.	$Y = -37.369+2.318x_1-2.043x_2-5.482x_3-6.213x_4$ $+5.70x_5+0.692x_6+0.251x_7-0.415x_8$ $+1.707x_9$	0.773**
c.	$Y = 25.953+1.022x_1-3.073x_2-16.03x_3-2.603x_4$ $+1.691x_5+0.734x_6+0.394x_7-0.443x_8+0.082x_9$	0.806**
d.	$\text{Log } Y = 1.231+0.928x_1-0.649x_2-1.878x_3-2.055x_4$ $+1.963x_5+0.246x_6+0.108x_7-0.115x_8$ $+0.509x_9$	0.822**

**Significant at 1 per cent level P (∠ .01)

Table 4.3(b)5.1. Regression equations fitted for CO-997 in ninth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -44.54+0.032x_1+0.541x_2+11.765x_3+0.348x_4 -0.141x_5+0.086x_6-0.128x_7-0.069x_8 +0.0002x_9$	0.253
2.	$Y = 2.304+0.03x_1+0.443x_2-1.035x_3-0.062x_6 +0.046x_8$	0.209
3.	$Y = -4.047+0.034x_1+0.93x_2-0.504x_3+0.237x_6$	0.197
4.	$Y = -0.773+0.023x_1+0.403x_2+0.039x_8$	0.191
a.	$Y = -825.691+9.858x_1+0.963x_2+274.451x_3 +272.406x_4+160866.358x_5-0.335x_6 -0.336x_7+161140.876x_8-161140.186x_9$	0.131
b.	$Y = -175.097+0.771x_1+2.175x_2+90.754x_3 +15.691x_4-10.619x_5+0.233x_6-0.37x_7 -2.385x_8+0.149x_9$	0.259
c.	$\text{Log } Y = -119.357+2.634x_1+0.5x_2+54.306x_3 +53.875x_4+40902.303x_5-0.007x_6 -0.091x_7+40956.63x_8-40956.499x_9$	0.065

Table 4.3(b)5.2. Regression equations fitted for CO-62175 in ninth month of plant growth.

Sl. No.	Regression equations	R^2
1.	$Y = -3.747+0.031x_1+0.412x_2+0.269x_3+0.032x_4 -0.032x_5+0.313x_6+0.015x_7-0.056x_8 +0.0004x_9$	0.834**
2.	$Y = -7.233+0.026x_1+0.575x_2+0.447x_6+0.0001x_9$ (S)	0.808**
3.	$Y = -4.142+0.038x_1+0.546x_2-0.239x_3+0.43x_6 +0.022x_8$ (S)	0.799**
4.	$Y = -7.319+0.046x_1+0.532x_2-0.639x_3+0.541x_6$ (S)	0.790**
5.	$Y = -3.156+0.015x_1+0.594x_2+0.0002x_9$	0.655**
6.	$Y = -6.884+0.032x_1+0.521x_2+0.068x_8$	0.713**
7.	$Y = -10.874+0.013x_1+0.783x_2+0.029x_5 +0.333x_6+0.056x_8$	0.602**
a.	$Y = -48.29+1.76x_1+1.41x_2+0.351x_3+1.24x_4 +0.663x_5+1.175x_6+0.01x_7+1.659x_8 -0.872x_9$	0.795**
c.	$Y = -2.334+0.71x_1+2.204x_2+0.876x_3+0.605x_4 -1.654x_5+1.955x_6+0.058x_7-2.239x_8 +0.19x_9$	0.829**
d.	$\text{Log } Y = -15.548+0.745x_1+0.384x_2+0.101x_3+0.552x_4 +0.215x_5+0.605x_6+0.036x_7+0.512x_8 -0.345x_9$	0.837**

**Significant at 1% level P ($\angle$ .01) of plant growth. Number of canes was found to be highly correlated with plot yield in all stages of plant growth for the two varieties. Number of tillers was also correlated with plot yield in the seventh month for both the varieties and in ninth month only for CO-997.

The biometric characters were highly interrelated among themselves. Height of the cane was correlated with girth which was also correlated with width of third leaf. Number of leaves was highly correlated w ith number of canes which was highly correlated with number of tillers and so on. This interrelationship among the characters introduced the problem of multicollinearity in the data.

The prediction equations fitted using all the biometric characters studied were found to give sufficiently high values of R^2 in different months. Maximum coefficient of determination was noticed for the full model in the eighth month after planting (0.903) for CO-997, and in the seventh month in the case of CO-62175 (0.883).

In ninth month none of the fitted model gave a significant R^2 in the case of variety CO-997. Predictability of the restricted models were compared with the full model using F test. Some of the simpler models were found to be equally efficient with the full linear model, these models have been indicated by the symbol (S).

Table 4.3(b)6. Regression equations selected in fifth and sixth months of plant growth.

Month	Variety	Constant	Estimates of parameters					R^2
			Coefficient of x_1	Coefficient of x_2	Coefficient of x_3	Coefficient of x_6	Coefficient of x_7	
V	CO-997	-11.556	0.046	0.834	0.761	0.312	0.041	0.686**
	CO-62175	-5.343	0.040	0.070	1.717	-0.106	0.081	0.781**
VI	CO-997	-11.552	0.024	1.382		-0.347	0.058	0.863**
	CO-62175	-8.107	0.016	-0.004	2.450	0.007	0.1 09	0.823**

**Significant at 1 per cent level P ($<$ .01)

Table 4.3(b)7. Regression equations selected in seventh, eighth and ninth month of plant growth.

Month	Variety	Estimates of parameters: Constant	Coefficient of x_1	Coefficient of x_2	Coefficient of x_3	Coefficient of x_6	Coefficient of x_7	Coefficient of x_8	R^2
VII	CO-997	-13.668	0.022	0.961	0.914	0.323		0.063	0.866**
	CO-62175	-6.268	0.056	0.261		0.139	0.046	0.059	0.835**
		-8.168	0.061	0.237	0.804	0.437			0.756**
VIII	CO-997	-16.406	0.037	1.166	0.655	0.667		0.010	0.901**
	CO-62175	-6.987	0.053	-0.126	0.439	1.155			0.767**
IX	CO-997+								
	CO-62175	-4.142	0.038	0.546	-0.239	0.430		0.022	0.799**
		-7.319	0.046	0.532	-0.639	0.541			0.790**

**Significant at 1% level P (< .01)

+ None of the R^2 values for the fitted models were significant (including the full model) and hence prediction equations were not given.

In most of the months R^2 values were found to be higher than 0.70, which indicated that more than 70 per cent variation in plot yield could be explained by the linear model of biometric characters.

For convenience of prediction a set of equations was selected in different months which is presented in tables 4.3(b)6 and 4.3(b)7. All the equations selected in different months were at par with the full model with respect to predictability. The coefficients of determination of the equations were fairly high indicating that yield of sugarcane can be successfully predicted with linear regression of yield on biometric characters.

4.4. Biometric characters influencing sugarcane yield - Method of path coefficients - Plant wise approach.

Path analysis was carried out for both the varieties in each of the months of observation. The direct and indirect effects towards yield for variety CO-997 and CO-62175 are given in tables 4.4(a) and 4.4(b) respectively. Path diagram was drawn only for the month having least residual variation. The residual effect is represented by h. Underlined figures denote the direct effects.

From the various tables under 4.4(a) and 4.4(b) it can be seen that height of the plant and stalk girth showed high positive direct effects uniformly during the entire period of observation. The direct effect of height on yield was slightly higher than that of girth in all the months except in the first month of observation for both the varieties. These two characters had low or negligible indirect effects through other characters. The direct effects of height on yield was found to be maximum (0.520) in the seventh month for variety CO-997 and for the other variety in the ninth month (0.440). The direct effects of girth on yield was maximum (0.424) for variety CO-62175 and for the other variety (0.398) in the ninth month of plant growth.

Width of third leaf showed relatively high positive direct effect on yield during seventh month for CO-997 but it had high negative effect through area of third leaf.

Table 4.4(a)1. Direct and Indirect Effects for Variety CO-997 in fifth month of Plant Growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No. of leaves/ cane x_6	Total leaf area/ cane x_7	Correlation with yield
x_1	0.264	0.162	0.111	-0.014	-0.067	0.012	0.003	0.472
x_2	0.117	0.364	0.133	-0.003	-0.096	0.008	0.004	0.526
x_3	0.115	0.190	0.254	-0.010	-0.173	0.009	0.006	0.394
x_4	-0.041	-0.014	-0.021	0.086	-0.054	-0.005	0.001	-0.049
x_5	0.099	0.185	0.233	0.024	-0.189	0.007	0.006	0.361
x_6	0.095	0.077	0.056	-0.012	-0.038	0.036	0.004	0.226
x_7	0.126	0.194	0.216	0.015	-0.169	0.019	0.007	0.408

h = 0.803

Table 4.4(b)1. Direct and Indirect Effects for Variety CO-62175 in fifth month of Plant Growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No. of leaves/ cane x_6	Total leaf area/ cane x_7	Correlation with yield
x_1	0.222	0.152	-0.015	0.024	0.091	-0.085	0.156	0.545
x_2	0.109	0.311	-0.016	0.023	0.092	-0.048	0.119	0.590
x_3	0.112	0.170	-0.029	0.026	0.150	-0.047	0.169	0.551
x_4	0.104	0.141	-0.015	0.051	0.122	-0.037	0.137	0.503
x_5	0.125	0.178	-0.027	0.038	0.161	-0.050	0.178	0.600
x_6	0.158	0.125	-0.012	0.016	0.067	-0.012	0.169	0.405
x_7	0.165	0.177	-0.024	0.033	0.137	-0.096	0.209	0.602

h = 0.715

Table 4.4(a)2. Direct and indirect effects for variety CO-997 in sixth month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No.of leaves/ cane x_6	Total leaf area/ cane x_7	Correlation with yield
x_1	<u>0.303</u>	0.001	0.056	-0.006	-0.142	-0.018	0.181	0.375
x_2	0.007	<u>0.030</u>	-0.004	0.007	-0.032	-0.008	0.005	0.055
x_3	0.065	-0.001	<u>0.262</u>	0.010	-0.784	-0.031	0.898	0.420
x_4	0.029	0.003	0.040	<u>0.065</u>	-0.568	0.005	0.599	0.115
x_5	0.046	0.001	0.221	0.040	-<u>0.931</u>	-0.022	1.032	0.387
x_6	0.024	0.001	0.036	-0.001	-0.091	-<u>0.222</u>	0.691	0.438
x_7	0.046	0.001	0.197	0.033	-0.803	-0.128	<u>1.197</u>	0.543

h = 0.758

Table 4.4(b)2. Direct and indirect effects for variety CO-62175 in sixth month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No. of leaves/ cane x_6	Total leaf area/ cane x_7	Correlation with yield
x_1	<u>0.362</u>	0.151	0.021	0.038	-0.289	-0.211	0.557	0.630
x_2	0.178	<u>0.307</u>	0.027	0.050	-0.373	-0.152	0.584	0.620
x_3	0.680	0.180	<u>0.046</u>	0.059	-0.586	-0.123	0.798	0.542
x_4	0.098	0.111	0.020	<u>0.139</u>	-0.456	-0.018	0.514	0.408
x_5	0.166	0.182	0.043	0.100	-<u>0.630</u>	-0.099	0.812	0.575
x_6	0.163	0.100	0.012	0.005	-0.132	-<u>0.469</u>	0.738	0.415
x_7	0.200	0.178	0.004	0.071	-0.509	-0.344	<u>1.006</u>	0.639

h = 0.644

Table 4.4(a)3. Direct and indirect effects for CO-997 in seventh month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No.of leaves/cane x_6	Total leaf area/cane x_7	Correlation with yield
x_1	0.339	0.099	0.316	-0.039	-0.274	0.004	0.082	0.525
x_2	0.102	0.329	0.572	0.113	-0.681	0.005	0.152	0.591
x_3	0.052	0.091	2.062	0.222	-2.302	0.006	0.346	0.478
x_4	-0.011	0.030	0.370	1.237	-1.690	-0.004	0.138	0.069
x_5	0.035	0.085	1.793	0.789	-2.648	0.002	0.337	0.394
x_6	0.061	0.087	0.618	-0.255	-0.321	0.020	0.311	0.522
x_7	0.064	0.115	1.641	0.392	-2.049	0.014	0.435	0.612

h = 0.568

Fig 4.4.1 Path Diagram for Variety Co 997 in Seventh Month of Plant Growth

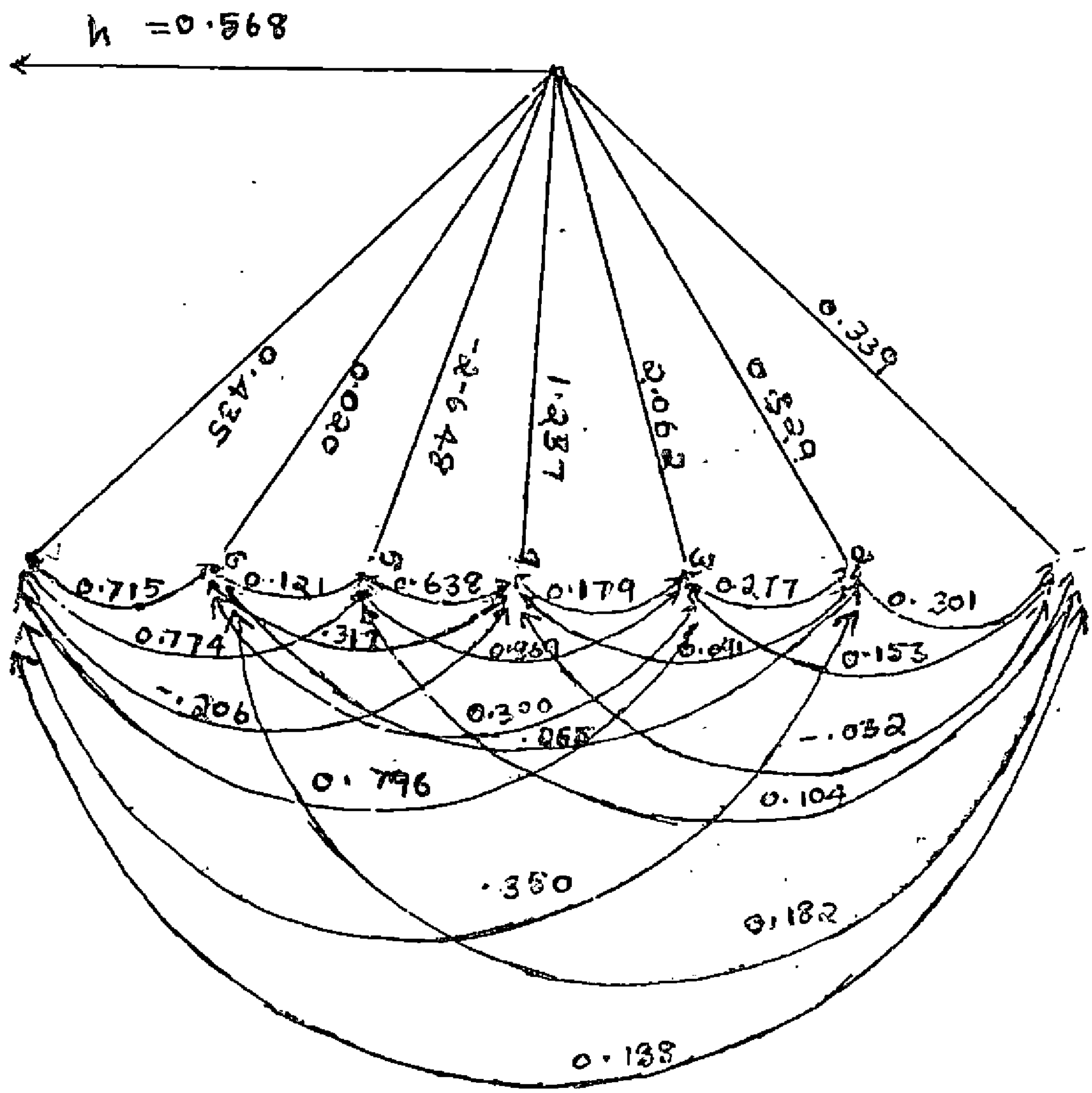


Table 4.4(b)3. Direct and indirect effects for CO-62175 in seventh month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No. of leaves/cane x_6	Total leaf area/cane x_7	Correlation with yield
x_1	<u>0.520</u>	0.135	-0.061	-0.029	0.155	-0.001	0.042	0.761
x_2	0.311	<u>0.227</u>	-0.086	-0.030	0.224	-0.001	0.077	0.721
x_3	0.184	0.112	-<u>0.174</u>	-0.065	0.397	-0.001	0.129	0.582
x_4	0.110	0.052	-0.083	-<u>0.137</u>	0.383	-0.002	0.128	0.450
x_5	0.175	0.111	-0.150	-0.114	<u>0.459</u>	-0.002	0.154	0.633
x_6	0.032	0.037	-0.022	-0.032	0.102	-<u>0.008</u>	0.109	0.218
x_7	0.125	0.100	-0.129	-0.101	0.406	-0.005	<u>0.174</u>	0.570

h = 0.465

Table 4.4(a)4. Direct and indirect effects for CO-997 in eighth month of plant growth

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No. of leaves/cane x_6	Total leaf area/cane x_7	Correlation with yield
x_1	<u>0.404</u>	0.082	-0.009	0.009	-0.048	-0.010	0.125	0.551
x_2	0.096	<u>0.341</u>	-0.011	-0.001	-0.060	-0.046	0.240	0.560
x_3	0.092	0.094	-<u>0.039</u>	-0.060	-0.246	-0.034	0.625	0.431
x_4	-0.016	0.001	-0.011	-<u>0.225</u>	-0.236	-0.023	0.550	0.040
x_5	0.063	0.067	-0.032	-0.172	-<u>0.309</u>	-0.032	0.740	0.324
x_6	0.020	0.075	-0.006	-0.025	-0.047	-<u>0.210</u>	0.610	0.408
x_7	0.056	0.092	-0.028	-0.138	-0.255	-0.141	<u>0.895</u>	0.481

h = 0.606

Table 4.4(b)4. Direct and indirect effects for CO-62175 in eight month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No. of leaves/ cane x_6	Total leaf area/ cane x_7	Correlation with yield
x_1	0.468	0.176	-0.028	-0.018	0.206	0.013	-0.073	0.743
x_2	0.248	0.332	0.030	-0.016	0.098	0.015	-0.081	0.666
x_3	0.202	0.150	-0.066	-0.028	0.117	0.071	-0.017	0.429
x_4	0.133	0.083	-0.030	-0.062	0.369	-0.024	-0.107	0.363
x_5	0.206	0.141	-0.016	-0.049	0.467	-0.006	-0.153	0.590
x_6	0.037	0.030	-0.028	-0.009	-0.016	0.167	-0.097	0.101
x_7	0.181	0.096	-0.006	-0.035	0.378	0.086	-0.188	0.557

h = 0.543

Table 4.4(a)5. Direct and indirect effects for CO-997 in ninth month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No. of leaves/ cane x_6	Total leaf area/ cane x_7	Correlation with yield
x_1	0.440	0.337	-0.053	0.008	-0.002	-0.000	0.050	0.479
x_2	0.041	0.398	-0.024	0.000	-0.003	-0.004	0.139	0.547
x_3	0.106	0.043	-0.222	-0.039	-0.009	-0.005	0.461	0.336
x_4	-0.032	-0.002	-0.077	-0.112	-0.009	-0.003	0.377	0.143
x_5	0.026	0.038	-0.059	-0.030	-0.033	-0.002	0.186	0.127
x_6	0.008	0.099	-0.072	-0.033	-0.004	-0.014	0.519	0.512
x_7	0.034	0.086	-0.158	-0.065	-0.009	-0.012	0.645	0.520

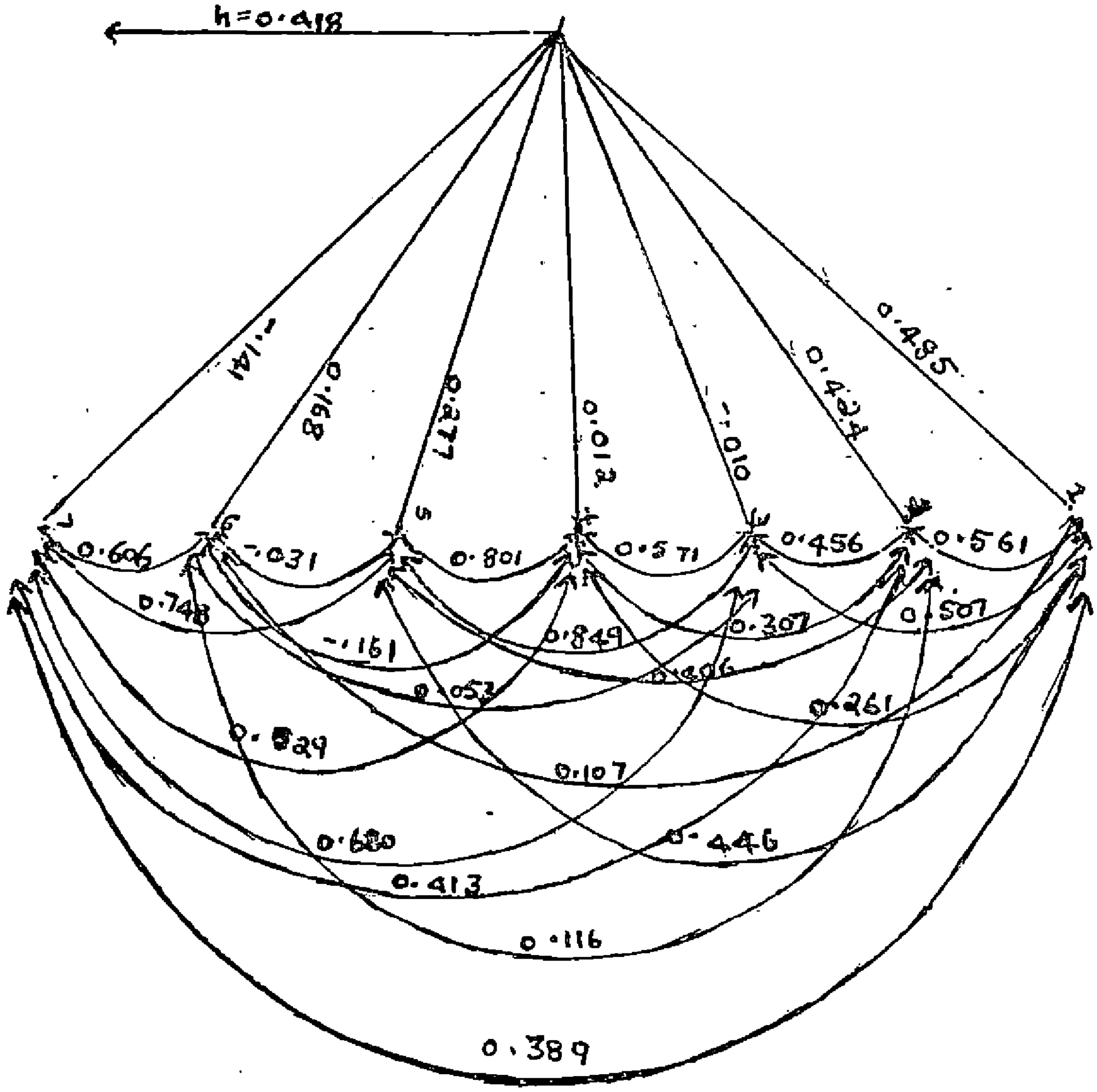
h = 0.581

Table 4.4(b)5. Direct and indirect effects for CO-62175 in ninth month of plant growth.

Character	Height x_1	Girth x_2	Width of 3rd leaf x_3	Length of 3rd leaf x_4	Area of 3rd leaf x_5	No.of leaves/cane x_6	Total leaf area/cane x_7	Correlation with yield
x_1	<u>0.485</u>	0.238	-0.005	0.003	0.124	0.018	-0.055	0.808
x_2	0.272	<u>0.424</u>	-0.005	0.004	0.113	0.020	-0.058	0.769
x_3	0.246	0.193	-<u>0.010</u>	0.007	0.236	0.009	-0.096	0.585
x_4	0.126	0.130	-0.006	<u>0.012</u>	0.222	-0.027	-0.075	0.384
x_5	0.216	0.172	-0.009	0.010	<u>0.277</u>	-0.005	0.105	0.557
x_6	0.052	0.049	-0.001	-0.002	-0.008	<u>0.168</u>	-0.085	0.173
x_7	0.188	0.175	-0.007	0.007	0.208	0.102	-<u>0.141</u>	0.532

h = 0.418

Fig 4.4.2. Path Diagram for Variety Co. 62175 in Ninth month of Plant Growth.



In the case of variety CO-997 area of third leaf showed a negative direct effect on all periods of study. But for the other variety in last three months of study area of third leaf exhibited a positive direct effect on yield. For the same variety, during the same period of time total leaf area showed negative direct effects on yield.

The direct effects of number of leaves on yield were either negative or negligibly small in all the months. Estimated total leaf area had a positive direct effect on yield in most of the months except for CO-997 in last three months.

Among the various plant characters the major contributors towards cane yield in all stages of plant growth were height of cane and girth of cane.

4.5. Pre-harvest forecasting of sugarcane yield - Principal component analysis - Plot wise approach.

The principal components and their loadings were calculated in each month separately for both the varieties. In most of the months the first three principal components explained as much as 75 per centage variation in the original data. But for CO-62175 in the seventh month the first four components explained only 62 per cent variation. Component loadings and proportionate percentage contribution of each component are given in table 4.5(a) for CO-997 and CO-62175. Regression equations fitted using the principal components as the explanatory variables are given in table 4.5(b).

Table 4.5(a)1. Component loadings and the percentage contribution of first three components in fifth month of plant growth.

Characters	CO-997 Component loadings I	II	III	CO-62175 Component loadings I	II	III
Height	0.6648	-0.1908	-0.3096	0.8420	-0.0983	0.0627
Girth	0.6419	-0.4034	-0.0346	0.5521	-0.4418	0.6646
Width of 3rd leaf	0.7908	-0.4765	-0.1909	0.8546	-0.3197	-0.0673
Length of 3rd leaf	-0.0720	-0.3203	0.9214	0.7995	-0.1339	-0.3835
Area of 3rd leaf	0.7556	-0.5807	0.1421	0.8966	-0.3501	-0.1436
No. of canes/ tillers per plot	0.5860	0.7428	0.0671	0.5092	0.7052	0.2586
No. of leaves/ plot	0.6838	0.6908	0.1408	0.7540	0.5658	-0.0275
Total leaf area/ plot	0.9082	0.2930	0.2229	0.9351	0.2516	0.0847
Percentage contribution	46.14	24.80	13.46	61.11	16.60	8.66
Total contribution		84.40			86.38	

Table 4.5(b)1. Regression equations in fifth month of plant growth

Variety	Regression equations	R^2
CO-997	$Y = 10.348 + 0.413x_1 + 0.264x_2 + 0.350x_3 - 0.303x_4 + 0.24x_5 + 0.783x_6 + 0.82x_7 + 0.778x_8$	0.568**
CO-62175	$Y = 7.937 - 0.04x_1 - 0.052x_2 + 0.31x_3 + 0.417x_4 + 0.342x_5 + 0.302x_6 + 0.008x_7 + 0.373x_8$	0.555**

**Significance at 1% level P ($<.01$)

Table 4.5(a)2. Component loadings and the percentage contribution of first three components in sixth month of plant growth.

Characters	CO-997 Component loadings			CO-62175 Component loadings		
	I	II	III	I	II	III
Height	0.5641	-0.2705	-0.7946	0.8342	0.1859	-0.0973
Girth	0.7940	-0.2758	-0.4033	0.7874	0.0227	-0.0204
Width of 3rd leaf	0.7880	-0.4077	-0.3743	0.8773	-0.1881	0.1886
Length of 3rd leaf	0.2102	-0.6805	0.7676	0.5731	-0.5121	-0.2281
Area of 3rd leaf	0.7403	-0.6496	0.7105	0.9022	-0.3523	0.0893
No. of canes/ plot	0.6396	0.6501	0.4420	0.2189	0.9251	-0.0756
No. of leaves/ plot	0.7973	0.5275	0.3240	0.4804	0.8268	-0.0750
Total leaf area	0.9373	0.1866	0.3284	0.0358	0.1040	0.9673
Percentage contribution	51.06	24.21	14.27	43.85	25.09	13.16
Total contribution		89.54			82.09	

Table 4.5(b)2. Regression equations in sixth month of plant growth

Variety	Regression equations	R^2
CO-997	$Y = 9.779 - 0.062x_1 + 0.163x_2 + 0.07x_3 - 0.124x_4 + 0.22x_5 + 0.148x_6 + 1.02x_7 + 0.85x_8$	0.814**
CO-62175	$Y = 7.386 + 0.34x_1 + 0.286x_2 + 0.307x_3 - 0.272x_4 + 0.22x_5 - 0.054x_6 + 0.466x_7 + 0.013x_8$	0.760**

**Significant at 1 per cent level P ($< .01$)

Table 4.5(a)3. Component loadings and the percentage contribution of first three components in seventh month of plant growth.

Characters	CO-997 Component loadings I	II	III	CO-62175 Component loadings I	II	III	IV
Height	0.6642	0.2751	0.1800	0.0372	0.1398	-0.5527	0.6620
Girth	0.5831	0.1401	0.6036	0.4698	-0.4453	-0.0435	0.4086
Width of 3rd leaf	0.2040	0.8642	0.0717	0.5183	-0.2831	0.0730	-0.2844
Length of 3rd leaf	-0.0333	0.6503	-0.2373	-0.1101	0.1190	-0.1418	0.6074
Area of 3rd leaf	0.2657	0.9203	-0.1073	-0.7238	0.3648	-0.1407	-0.0219
No.of canes/ plot	0.8178	-0.3363	-0.3680	-0.6332	0.5999	-0.1281	-0.0820
No.of tillers/ plot	0.5306	-0.2298	0.6116	-0.0892	-0.0197	-0.6369	0.2779
No.of leaves/ plot	0.9002	-0.3096	-0.2504	0.2595	0.7871	0.0176	-0.1444
Total leaf area	0.9411	0.0127	-0.2803	-0.3469	-0.1467	0.7571	-0.0994
Percentage contribution	39.34	26.39	0.125	18.03	16.07	14.98	13.00
Total contribution		78.18			62.09		

Table 4.5(b)3. Regression equations in seventh month of plant growth.

Variety	Regression equations	R^2
CO-997	$Y = 8.972+0.561x_1+0.504x_2-0.035x_3-0.2x_4+0.013x_5+0.792x_6+0.558x_7+0.005x_8+0.905x_9$	0.770**
CO-62175	$Y = 7.162+0.774x_1-0.758x_2-0.579x_3+0.282x_4+0.669x_5+0.319x_6+0.614x_7+0.66x_8-0.841x_9$	0.540**

**Significance at 1 per cent level P ($<$ .01)

Table 4.5(a)4. Component loadings and percentage contribution of the first three components in eighth month of plant growth.

Characters	CO-997 Component loadings I	II	III	CO-62175 Component loadings I	II	III
Height	0.4519	0.1482	0.6058	0.8057	0.0669	0.1465
Girth	0.6372	0.1302	0.4256	0.6472	0.0927	0.4207
Width of 3rd leaf	0.5271	0.8744	0.2640	0.8540	-0.3038	-0.0545
Length of 3rd leaf	0.4678	0.5964	-0.6453	0.7551	-0.4366	-0.2706
Area of 3rd leaf	0.6324	0.9428	-0.1908	0.8965	0.3724	-0.1761
No. of canes/plot	0.6002	-0.7582	-0.0748	0.2029	0.9382	-0.1932
No. of tillers/ plot	0.5459	0.3270	0.1480	0.3281	-0.0559	0.8463
No. of leaves/ plot	0.7059	-0.7303	-0.0536	0.2369	0.9460	-0.2160
Total leaf area	0.8668	-0.3874	-0.1143	0.7563	0.5461	-0.3271
Percentage contribution	37.924	23.626	12.380	43.870	27.90	13.47
Total contribution		73.93			85.24	

Table 4.5(b)4. Regression equations in eighth month of plant growth

Variety	Regression equations	R^2
CO-997	$Y = 10.205 + 0.605x_1 + 0.667x_2 - 0.014x_3 - 0.384x_4 - 0.228x_5 + 0.955x_6 + 0.745x_7 + 1.034x_8 + 0.9x_9$	0.736**
CO-62175	$Y = 7.234 + 0.220x_1 + 0.226x_2 + 0.054x_3 - 0.113x_4 + 0.142x_5 + 0.062x_6 + 0.06x_7 - 0.069x_8 - 0.105x_9$	0.709**

**Significant at 1 per cent level P ($<$ .01)

Table 4.5(a)5. Component loadings and percentage contribution of first three components in ninth month of plant growth.

Characters	CO-997 Component loadings I	II	III	CO-62175 Component loadings I	II	III
Height	0.3359	-0.0668	0.8317	0.7220	0.1957	0.4042
Girth	0.5058	-0.0387	-0.0807	0.7263	0.0244	0.3867
Width of 3rd leaf	0.3281	0.7330	0.4063	0.8293	-0.3621	-0.1256
Length of 3rd leaf	0.2485	0.7293	-0.4509	0.7281	-0.4111	-0.3862
Area of 3rd leaf	0.3664	0.9226	-0.0167	0.8552	-0.3461	-0.2632
No. of canes/plot	0.7445	-0.5194	-0.1467	0.1626	0.9206	0.1799
No. of tillers/plot	0.6766	-0.3785	-0.0355	0.5387	0.1290	0.5861
No. of leaves/plot	0.9272	-0.2545	-0.0729	0.1770	0.9544	-0.1330
Total leaf area	0.9545	0.1686	-0.0553	0.7094	0.5984	-0.3174
Percentage contribution	38.39	27.03	12.20	42.767	29.689	11.572
Total contribution		77.62			84.03	

Table 4.5(b)5. Regression equations in ninth month of plant growth

Variety	Regression equations	R^2
CO-997	$Y = 9.584 + 0.330x_1 + 0.501x_2 + 0.007x_3 - 0.048x_4 - 0.024x_5 + 0.926x_6 + 0.802x_7 + 0.992x_8 + 0.846x_9$	0.731**
CO-62175	$Y = 7.124 + 0.359x_1 + 0.279x_2 + 0.099x_3 + 0.026x_4 + 0.065x_5 + 0.497x_6 + 0.278x_7 + 0.499x_8 + 0.496x_9$	0.683**

**Significant at 1 per cent level P (<.01)

The coefficients of determination of the forecasting models using principal components ranges between 54 and 81.4 per cent. The highest predictability was noticed in sixth month of plant growth for the two varieties. This equation in table 4.5(b)2. can be used to forecast yield of sugarcane of the two varieties in sixth month with predictability 81.4 per cent for CO-997 and 76 per cent for CO-62175.

During the ninth month for variety CO-997 principal component method gave a high and significant R^2 value (0.731) when compared to a nonsignificant R^2 (0.253) by the regression analysis. Thus the regression equation through principal component analysis may be made use of for yield prediction in ninth month for variety CO-997.

DISCUSSION

5. DISCUSSION

Investigations on the pre-harvest forecasting of sugarcane yield on two popular varieties of sugarcane namely CO-997 and CO-62175 were made on the basis of periodical data on biometric characters gathered from the Sugarcane Research Station, Thiruvalla and the results obtained were discussed below:

5.1. Pre-harvest prediction of sugarcane yield - Method of multiple regression - plant-wise approach.

The simple correlation analysis revealed that height of the cane was positively and significantly correlated with yield in all stages of plant growth for the two varieties. The results are in agreement with the results of Hooda *et al*. (1979), Singh and Sangha (1970), Singh and Sharma (1982) and Norman (1971). Girth of the cane was also positively and significantly correlated with yield in all the months except in the sixth month for variety CO-997. Singh and Sharma (1982) and Norman (1971) reported positive significant correlation between girth and yield. Width of third leaf showed a positive significant correlation with yield in all periods of plant growth. Length of third leaf had a positive significant correlation with yield in the entire period of plant growth for CO-62175. In the case of the other variety the relationship was non-significant. The result is in contradiction with the findings of Hooda et al. (1979) who observed a positive significant correlation between length of third leaf and yield. Area of third leaf and estimated total leaf area were highly correlated with yield in different months. Number of leaves per cane had a positive significant correlation with yield in all stages of crop growth for CO-997 and only upto seventh month for CO-62175.

The coefficients of determination of the fitted regression equations were low but significant in the fifth and sixth months of plant growth. From the seventh month onwards the predictability increased considerably. A set of regression equations was selected for the pre-harvest prediction of yield and presented in table 4.2(b)6. The cane yield estimated from the equations multiplied by the number of canes in the plot will give an advance estimate of plot yield. Similar methodology was applied by Bohra et al. (1969) for the prediction of forage yield using biometric characters and there was close agreement between predicted yield and actual yield.

5.2. Pre-harvest prediction of sugarcane yield - Method of multiple regression - Plot wise approach.

From the correlation analysis of per plot observations it was found that height of the cane and girth of the cane had high positive significant correlation with yield except in the ninth month for variety CO-997. The result is in partial agreement with those of Jha *et al*. (1981) and Chandrahas *et al*. (1983). Width of third leaf had an insignificant correlation with plot yield from the seventh month onwards for variety CO-997. The other variety showed significant correlations between width of third leaf and yield during the entire period of observation. But Jha *et al*. (1981) noticed significant correlation only in five to six months after planting. Number of canes/tillers and number of leaves per plot showed high positive significant correlation with plot yield in all stages of plant growth. This result is in perfect agreement with those of Chandrahas *et al*. (1983) and Jha *et al*. (1981).

Coefficient of determination of the forecasting models of the present study ranged between 72.9 and 90 percentage when all the characters were taken into account. According to Jha *et al*. (1981) reliable forecasts can be made available only from the seventh month after planting onwards with an accuracy of about 68 percentage. All the statistical models developed in this study for yield prediction were more efficient than those proposed by Jha *et al*. (1981). Further these models could be used for yield prediction as early as in the fifth month after planting with sufficiently high degree of precision. The coefficients of determination of the proposed models were in the range from 68.6% to 90.1%. The predictability of the equation was considerably increased at the later stages of plant growth and finally attained a maximum value of 90 percentage for variety CO-997 in the seventh month and 88.6 percentage for variety CO-62175 in eighth month of plant growth. The results of the present study are also in confirmity with the findings of Chandrahas et al. (1983) who found that yield of sugarcane could be predicted from the fifth month onwards with satisfactory precision. But it is observed that the coefficient of determination of the models of the present study are higher than those reported by Chandrahas et al. (1983) and hence are more efficient for yield prediction. A possible reason for the high value of R^2 may be the inclusion of more characters in the model. Varietal differences might have also contributed to this discrepancy. The forecasting models for the two varieties were not identical in nature which indicated that the varietal aspect is also to be taken into account in building up suitable forecasting models for sugarcane crop. Chandrahas et al. (1983) used five explanatory variables viz., height of the cane, girth of the cane, width of third leaf, length of third leaf and number of canes/tillers per plot in the linear model and observed coefficient of determinations in the range 60 to 72 percentage. Whereas simpler models with the four characters, height of the cane, girth of the cane, width of third leaf and number of canes/tillers per plot developed in this study could explain about 63 to 89.6 percentage of variation.

It was found that the restricted models with five biometric characters, height of the cane, girth of the cane, width of third leaf, number of canes/tillers per plot and number of leaves per plot were sufficient to predict yield of variety CO-997 in all stages of the study. Yield of variety CO-62175 could effectively be predicted in fifth and sixth months of plant growth using the above mentioned five biometric characters. The informations on first four biometric characters are enough for the forecasting of yield of variety CO-62175 from seventh month of plant growth onwards.

The coefficient of determinations of none of the fitted models were found to be significant in the ninth month for variety CO-997. In this month plant-wise observations can be efficiently used for yield prediction.

The first three types of transformations mentioned in section 3.2 were tried by Jha <u>et al</u>. (1981) and Chandrahas <u>et al</u>. (1983), and they found that these three transformations were equally efficient as the original model. In this study also no significant gain was achieved by the square root, reciprocal and logarithemic transformations over the original data. The doubly logarithemic transformation resulted in a slight increase in the value of R^2 in some of the months. But this increase was negligibly small. So considering simplicity and convenience the linear models in the original non-transformed variables was used for yield prediction.

5.3. Biometric characters influencing sugarcane yield - Method of Path Coefficients - Plant-wise approach.

Path analysis was carried out in each month for the two varieties to identify the characters influencing the yield of sugarcane during different stages of plant growth.

Out of the seven characters studied only the height of the cane showed high positive direct effect on yield in all stages of plant growth for the two varieties and its influence on yield was higher in later stages of plant growth. The results are in agreement with the findings of Norman (1971), Hooda et al. (1979) and Singh and Sharma (1982), who observed high positive direct path due to stalk height on cane yield. Girth of the cane also had positive direct effect towards cane yield in all periods of study. According to Norman (1971) and Singh and Sharma (1982) girth of the cane had a positive direct effect, but Hooda et al. (1979) reported a negative direct effect. This result of the present study is in perfect agreement with that of

Norman (1971) and Singh and Sharma (1982) and contradictory to that of Hooda et al. (1979). Height and girth exerted very little indirect effects through other characters. The direct effects of width of third leaf were negative or negligible in all stages of plant growth for variety CO-62175. Whereas for other variety it showed high positive direct effects in early stages of study and negative effects in later stages. In the seventh month for variety CO-997 width of third leaf did show substantial positive direct effect, but its indirect effect through area of third leaf was negative and high. Length of third leaf and negative or negligibly small direct effects in entire period of observation except for CO-997 in the seventh month where it had a positive direct effect. But the indirect effect through area of third leaf was higher and negative. This result is in partial agreement with those of Hooda et al. (1979), who noticed a negative and low direct effect to leaf length on yield.

Area of third leaf had negative direct effect in all stages for variety CO-997, whereas in the case of other variety positive direct effect in almost all months except in sixth month. Number of leaves/cane had negative or negligible direct effects in all the months for the two varieties. Even though the number of leaves/cane and leaf dimensions showed negligible direct effects, the estimated total leaf area/cane exhibited high positive direct effect towards yield in the entire period of study for variety CO-997. But for other variety it had a positive direct effect in early stages of plant growth and negative in later stages.

The residual effect was comparitively high for both the varieties varying in the range from 0.418 to 0.803. The high residual effect indicated that there may be some important characters which could not be utilised in the present study. The residual effect was the least (0.568) for variety CO-997 in the seventh month of plant growth and for variety CO-62175 in the ninth month of plant growth (0.418).

The results of this study revealed that height of the cane and girth of the cane were the most important yield contributing characters in all stages of plant growth for the two varieties.

5.4. Pre-harvest forecasting of sugarcane yield - Principal component analysis - Plot wise approach.

On examining the values of the component loadings it was found that all of the explanatory variables exert their influence on the criterion variable through one or the other of the selected components. Hence elimination of variables is not advisable. Among the different variables the least contributing variable appears to be length of third leaf which has comparitively lower loadings than others in most of the months. For further screening of variables the axes are to be rotated through the technique of varimax rotation.

The first three principal components that explain more than 75 per cent variation in plot yield were used as regressors in the multiple linear regression equation for the two varieties in all the months except for variety CO-62175 in seventh month for which the first four components were used as regressors. The R^2 values are comparitively low for the prediction equations fitted using the method of principal components than the original equations by usual regression analysis. This may be due to the inclusion of only limited number of components in the regression model. It may be happen that some of the principal components with small variance may be related with the dependent variable. If more components were included, R^2 values would have been higher. However the accuracy of the forecasting model obtained through principal component analysis was higher than that reported by Anonymous. (1983) in hybrid Jowar.

SUMMARY

6. SUMMARY

A study was conducted to develop suitable statistical models for the pre-harvest prediction of sugarcane yield using biometric characters. Data were collected periodically with a monthly interval from the Sugarcane Research Station, Thiruvalla on two sugarcane varieties namely CO-997 and CO-62175. The first observation was recorded in the fifth month after planting. Prediction equations were evolved by the method of multiple linear regression using plant wise and plot wise observations. The characters influencing different stages of plant growth were identified by path analysis. Principal component analysis was carried out using the plot wise observations in each month separately for the two varieties. Regression equations were fitted using the principal components as explanatory variables. The salient findings in the study are summarised below:

6.1. The study revealed that yield of sugarcane per unit area could be successfully predicted with sufficient degree of accuracy in either case of utilising cane wise or plot wise observations.

6.2. The developed models based on plant wise observations were able to explain moderately high degree of variability during later stages of plant growth. The different linear and transformed models fitted on the basis of per plant observations failed to serve as efficient predictors of production during the early stages of plant growth. Some of the useful prediction models developed on the basis of per plant data are listed below:

Forecasting models selected for CO-997 and CO-62175 in different months

Month	Variety	Estimates of parameters							
		Constant	Coefficients of Height	Girth	Width of 3rd leaf	Area of 3rd leaf	No.of leaves/cane	Estimated total leaf area	R^2
VII	CO-997	78.712	5.312	90.389				0.109	0.642**
	CO-62175	-4896.192	6.419	78.492		1.698			0.761**
VIII	CO-997	-1479.500	5.622	100.881	76.044		43.961		0.619**
	CO-62175	-943.218	5.076	103.807		0.974			0.704**
IX	CO-997	-1015.126	4.563	115.986				0.057	0.635**
	CO-62175	-1021.473	5.048	115.789		0.666			0.819**

**Significance at 1 per cent level P ($<$.01)

The cane yield determined on the basis of the prediction equation when multiplied by the total number of canes/plot will give the expected plot yield.

6.3. Analysis of plotwise observations revealed that prediction of sugarcane yield could be effectively attempted as early as in the fifth month of plant growth with a sufficiently high degree of precision.

Maximum coefficient of variations (88.3%) was noticed in the seventh month for variety CO-62175 and 90.3 for variety CO-62175 in eighth month of plant growth. The selected models with corresponding coefficients of determination are tabulated below:

Forecasting models selected for CO-997 and CO-62175 in different months

Month	Variety	Constant	Estimates of parameters					R^2
			Coefficients of					
			Height	Girth	Width of 3rd leaf	No.of canes/ tillers/ plot	No. of leaves	
	CO-997	-11.556	0.046	0.834	0.761	0.312	0.041	0.686**
V	CO-62175	-5.345	0.040	0.070	1.717	-0.106	0.081	0.781**
	CO-997	-11.552	0.024	1.382		0.347	0.058	0.863**
VI	CO-62175	-8.107	0.016	-0.004	2.450	0.007	0.109	0.823**
	CO-997	-13.668	0.022	0.961	0.914	0.323	0.063	0.866**
VII	CO-62175	-8.168	0.061	0.237	0.804	0.437		0.756**
	CO-997	-16.406	0.037	1.166	0.655	0.667	0.010	0.901**
VIII	CO-62175	-6.987	0.053	-0.126	0.439	1.155		0.767**

**Significance at 1 per cent level P ($<$ .01)

6.4. Four types of transformations were applied in the experimental data, square root, reciprocal semi logarithemic and doubly logarithemic. It was found that none of the transformations resulted in a significant improvement over the original model. However the doubly logarithemic transformation was found to yield slightly better results in certain cases.

6.5. Path analysis revealed that among the various biometric characters the major contributors towards cane yield in all stages of plant growth were height of the cane and girth of the cane. The direct influence of number of leaves on cane yield were negligibly small in all stages of plant growth.

6.6. The principal component analysis failed to earmark a sub set of important characters at the expense of others. The only finding is that the contribution of length of third leaf to the divergence was negligibly small. The first three principal components explained more than 75 per cent of variation in the original data in most of the months. From the prediction equations fitted using the principal components as explanatory variables yield could be predicted from the fifth month onwards with an accuracy ranging from 54 to 81.4 percentage. Sixth month after planting was found to be the best for prediction using this method with a predictability coefficient of 81.4 per cent for variety CO-997 and 76 per cent for variety CO-62175. The equations in the original form are given in table.

Forecasting models using principal components in sixth month of plant growth

Variety	Regression equations	R^2
CO-997	Yield = 9.779-0.062 (height)+0.163 (girth) +0.07 (width of third leaf)-0.124 (length of third leaf)+0.22 (area of third leaf)+0.148 (number of canes/tillers)+1.02 (number of leaves)+0.852 (estimated total leaf area).	0.814**
CO-62175	Yield = 7.386+0.34 (height)+0.286 (girth) +0.307 (width of third leaf -0.272 (length of third leaf) +0.227 (area of third leaf)-0.054 (number of canes/tillers)+0.466 (number of leaves)+0.013 (estimated total leaf area).	0.760**

**Significant at 1 per cent level ($P < .01$)

REFERENCES

REFERENCES

Abraham, T.P. and Khosla, R.K. (1965). On the possible use of component analysis technique in pest and disease survey data. J. Indian Soc. agric. Stat., 17 : 208 - 219.

Agarwal, Ranjana; Jain, R.C.; Jha, M.P. and Singh, D. (1980). Forecasting of rice yield using climatic variables. Indian J. agric. Sci., 50 : 680 - 684.

Anonymous (1983). Statistical Newsletter of Indian Agricultural Statistics Research Institute, New Delhi. 9 (2) : 7.

Basu, A.N. (1968). A study on curvilinear relationship of yield as product of functions of crop characteristics in sugarcane. J. Indian Soc. agric. Stat., 20 : 29 -44.

Bohra, R.K.; Jain, M.B. and Chakravarthy, A.K. (1969). Estimation procedure for assessment of forage production in a pasture of Cenchrus ciliaric. Indian J. agric. Sci., 39 : 207 - 213.

Chandrahas, S.R., Bapat and Singh, B.H. (1983). On pre-harvest forecasting of yield of sugarcane. Agric. Situation in India, 38 : 291 - 292.

Chaube, N.A. and Ratnalikar, D.V. (1982). On forecasting cotton production using partial harvest data. J. Indian Soc. agric. Stat., 34 : 129.

Das, J.C. and Ramachandra, G. (1971). Forecasting of yield of bajra on the basis of weather parameters. J. Indian Soc. agric. Stat., 23 : 100.

Dewey, D.R. and Lu, K.H. (1959). A correlation and path coefficient analysis of components of crested wheat grass production. J. Agron., 51 : 515 - 518.

George, M.V. and Vijayakumar, K. (1979). Forecasting of cashew yield. J. Indian Soc. agric. Stat., 31 : 95.

Girshick,M.A. (1936). Principal components. J. Amer. Stat. Assoc., 31 : 519 - 528.

Grafius,J.E. and Kiesling,R.L. (1960). The prediction of the relative yields of different oat varieties based on known environmental variables. Agron. J. 52 : 396 - 398.

Hooda,R.S., Babu,C.N. and Khairwal,I.S. (1979). Association and path analysis of nine characters in progenies of four sugarcane crosses at settling stage. Indian J. agric. Sci., 49 : 931 - 933.

*Hotelling,H. (1933). Analysis of a complex of statistical variables into principal components. J. Educ. psych. 24 : 417 - 444.

Jha,M.P., Jain,R.C. and Singh,D. (1981). Pre-harvest forecasting of sugarcane yield. Indian J. agric. Sci., 51 : 757 - 761.

Jitendra Mohan,K.V. and Gyan Prakash (1971). Regression and correlation of fibre yield and its components in sannhemp (crotalaria juncea L.). Indian J. agric. Sci., 41 : 378 - 381.

Khatri,T.J. and Patel,R.M. (1983). Use of eye and rainfall variables for pre-harvest forecasting of grountnut yield in Gujarat. J. Indian Soc. agric. Stat., 35 : 69 - 78.

Krishnakumar,N. (1983). Yield prediction in coconut based on foliar N, P and K values. M.Sc. (Ag.) Thesis, KeralaAgricultural University.

Li,C.C. (1956). The concept of path coefficient and its impact on population genetics. Biometrics, 12 : 190 - 210.

Malhotra,R.S. and Jain,R.P. (1972). Path and regression analysis in barley. Indian J. agric. Sci. 42 : 404 -406.

Murthy,B.R. and Arunachalam,V. (1967). Factor analysis of diversity in the genus sorghum. J. Indian Soc. agric. Stat., 19 : 145.

Naphade,O.S. (1972). Correlation and path analysis for some characters contributing to fodder yield in sorghum. Indian J. agric. Sci., 42 (9) : 790 - 791.

Norman I. James (1971). Yield components in random and selected sugarcane populations. Crop. Sci. 2 : 906 - 908.

Parameswaran,N.K. (1979). Factors affecting yield in cashew. M.Sc. (Ag.) Thesis, Kerala Agricultural University.

Rao Venkata,C., Narasimhayya,G. and Rao Appa,K. (1973). Path analysis in flue - cured tobacco. Indian J. agric. Sci., 43 : 198 - 201.

* Sanderson,F.H. (1954). Method of crop forecasting. Harvard University Press, Cambridge.

Singh,M.N., Srivastava,S.K. and Ashok Kumar,K.A. (1983). Multiple growth models of production of wheat, rice and sugarcane in 15 Districts of Eastern U.P. J. Indian. Soc. agric. Stat., 35 : 189 - 190.

Singh,H.N., Singh,S.B. and Singh,T.K. (1981 a). Selection parameters in sugarcane. Indian J. agric. Sci., 51 : 562 - 566.

Singh,B.V. and Sangha,A.S. (1970). Correlation studies in sugarcane. Madras agric. J., 57 : (19) : 474 - 477.

Singh,H. and Sharma,H.C. (1982). Interrelationship between yield and yield components in sugarcane. Punjab agric. J. 19 : 185- 187.

Singh,H.N., Singh,S.B., Verma,P.S. and Singh,R.V. (1981 b). Correlation and path analysis in sugarcane. Indian J. agric. Sci. 15 : 126 - 130.

Sundaram,A., Ramakrishnan,P.V., Marappan,S. and Balakrishnan,G. (1980). Factor analysis in cowpea. Indian J. agric. Sci., 50 : 216 - 218.

Sundaresan, K., Balakrishnan, N. and Shanmugasundaram, S. (1979). Correlation between seedlings and their settlings in intervarietal crosses of sugarcane for major economic characters. Indian J. agric. Sci., 49 : 934 - 937.

Thomas, R. J. and Osenbrug, A. (1959). Effect of manure, nitrogen, phosphorus and climatic factors on the production and quality of Brone grass - crested wheat grass. Agron. J., 51 : 63 - 66.

Tikka, S. B. S. and Asawa, B. M. (1978). Factor analysis in Lentil. Indian J. agric. Sci., 48 : 643 - 646.

Vaishnav, P. R. and Patel, R. M. (1983). Evaluation of different statistical models for pre-harvest forecasting of groundnut yield. J. Indian Soc. agric. Stat., 35(2) : 101.

Wendell, C. J. (1959). A mathematical procedure for evaluating relationship between climate and wheat yield. Agron. J., 51 : 635 - 639.

* Wright, S. (1921). Correlation and causation. J. agric. Res., 20 : 557 - 585.

* Original not seen.

PRE-HARVEST FORECASTING OF SUGARCANE YIELD

BY

ALPHI KORATH

ABSTRACT OF A THESIS
Submitted in partial fulfilment of the
requirement for the degree
Master of Science (Agricultural Statistics)
Faculty of Agriculture
Kerala Agricultural University

Department of Statistics
COLLEGE OF VETERINARY & ANIMAL SCIENCES
Mannuthy, Trichur

1984

ABSTRACT

Several yield prediction models were tried to examine their suitability for the pre-harvest prediction of yield of two varieties of sugarcane namely CO-997 and CO-62175 in different months of plant growth using biometric characters based on the data collected from the Sugarcane Research Station, Thiruvalla. The methods of multiple regression analysis, path coefficient analysis and principal component analysis were used for the above purpose.

Multiple regression analysis using plant biometric characters revealed that cane yield could be predicted on the basis of observations on height of the cane, girth of the cane and estimated total leaf area per cane or area of third leaf from the seventh month after planting onwards with an accuracy in the range of 59.5 to 81.9 per cent. The estimated cane yield when multiplied by the number of canes in the plot will give an advance estimate of the plot yield.

Linear models with five biometric characters viz., height of the cane, girth of the cane, width of the third leaf determined from the selected plants of each plot and number of canes/tillers and number of leaves determined on a whole plot basis were sufficient to predict the plot yield of the crop as early as in the fifth month of plant growth with an accuracy in the range 68 to 90 per cent.

Path analysis revealed that height of the cane and girth of the cane were the two important characters contributing towards cane yield in all stages of plant growth.

Using the forecasting models fitted with principal components as explanatory variables, yield could effectively be predicted with 81.4 per cent accuracy for variety CO-997 and with 76 per cent accuracy for variety CO-62175 in the sixth month of plant growth.