Full Research Article

Multivariate Analysis of Yield Data of Apple Crop for Optimizing Productivity in Himachal Pradesh

Bharti[*], P. K. Mahajan and Ashu Chandel

Dept. of Basic Sciences Dr. Y. S. Parmar University of Horticulture and Forestry, Nauni-Solan, Himachal Pradesh (173 230), India

Abstract

The paper deals with the usefulness of Discriminant and Principal Component analyses for determining the relative contribution of morphological and reproductive characters responsible in increasing the yield of apple. The technique Discriminant analysis was applied to formulate categorization rule for allocating the apple tree to 'high' and 'low' yielder groups. This Discriminant equation revealed that the characters canopy spread (X_2), Fruit set (X_7), LD Ratio (X_9) and Fruit weight (X_{10}) are the most important characters that discriminated the two groups. The Principal Component Analysis was extracted for the assessment of relative contribution of morphological and reproductive characters responsible in increasing the yield of apple. In case of high yielders, three of the ten Principal Components (PCs) have Eigen values greater than unity (Gutman's lower bound) which played the main role in the analysis. These components were vegetative characteristics, Plant vigour and flowering characteristics and Fruiting characteristics which explained 33.29%, 22.24% and 13.55% respectively and collectively 69.08% of the total variation of the original variables. In case of low yielders, three principal components had been retained for the analysis. These components were Plant Vigour, Yield and Flowering Characteristics and Fruiting characteristics which explained 35.68%, 22.22% and 12.09% respectively and in aggregate, 69.99% of the total variation of original variables.

1. Introduction

Apple (*Malus domestica* Borkh) is a deciduous plant belonging to the family Rosaceae. In India the commercial cultivation of apple is largely confined to the states of Jammu & Kashmir, Himachal Pradesh and Uttarakhand. Together these states account for 99% of the total apple production in the country. Among various fruits of Himachal Pradesh, apple is of particular interest for a number of reasons. In terms of both area and production apple is the most important fruit crop. The production of this fruit provides a major source of income and employment. It occupies an area of 1,06,939 ha with production of 4,12,395 MT in Himachal Pradesh (Anonymous, 2014). Of the 12 districts of Himachal Pradesh apple is produced mainly in Shimla, Kullu and Kinnaur districts. Among these, Shimla is the main apple growing district of Himachal Pradesh.

Apple yield is a complex trait which is influenced by several factors namely; tree characteristics, blooming characteristics, fruit set and crop density. Identifying a single variable representative of the complex trait i.e. yield may not be reliable, so the researcher is faced with the possibility of separately examining many related variables. The attempt to conduct a series of univariate statistical analyses carried out for each of the variables does not hold promise as it ignores the correlation among the variables and sometimes the conclusions may be misleading. On the contrary, multivariate analysis takes into account the interdependence and relative importance of the various influencing characters and yields more meaningful information.

Bryant (1966) revealed that method of discriminatory analysis frequently yields more satisfactory results than a regression or correlation. Eaton and Lapins (1970) used discriminant function analysis for identification of standard and compact apple trees. Fisher (1993) also used discriminate analysis in apple root stock selection. Holland (1969) has discussed some applications of multivariate methods viz., principle component analysis and factor analysis in biological sciences. Chernudubov (1994) also used factor analysis for the morpho-

anatomical traits of *Pinus sylvestris* in Insular Pine Forest of the South Russian Plain. Schrevens et al. (1995) used principal component analysis to characterize the quality evaluations of tomatoes during shelf life in relation to specific treatments.

In this paper, discriminant analysis was carried out to formulate the categorization rule for allocating the apple tree to high yielder group and low yielder group. An attempt has also been made to bring out the basic components with linear combination of morphological and reproductive characters contributing significantly towards apple yield by using principal component Analysis. Lezzoni and Pritts (1991), Belie (2000); Qurrie et al. (2000) used Principal Component Analysis in agricultural research. Bernard and Gregory (1991) also used Principal Component regression to mitigate the effect of multicollinearity. The techniques are described in detail in standard statistical book such as (Anderson, 1958).

2. Materials and Methods

Field experiment was conducted during 2014-15 at farmer's apple orchard in Jubbal block of Shimla district as this area represent the main apple growing belt of the State. The farm is located at latitude of 31°06´ N, at a longitude of 77°39´ E and an altitude of 1,901 meters above mean sea level. An optimum sample size of 105 apple trees for 19 year old plantation was selected randomly by following a two-step approach as suggested by Stein (1945) and Cox (1958). Four branches from each of the tree in four directions as per the practice in vogue were selected and the following observations were recorded i.e. Yield (Y), Tree height (X_1), Canopy spread (X_2), Trunk girth (X_3), FD: Flower Density (X_4), FDI: Flower Density Index (X_5), FI: Flowering Intensity (X_6), FS: Fruit Set (X_7), CD: Crop Density (X_8), LD: Length Diameter Ratio (X_9) and Fruit weight (X_{10}).

The data collected were subjected to discriminant analysis to define a systematic and statistically valid procedure for categorizing the trees as high and low yielders. For these two populations, Principal Component Analysis was carried out to bring out the basic components associated with the above referred morphological characters of apple.

3. Results and Discussion

In the first stage, on the basis of yield, the observations were divided into two groups and discriminant function was fitted. The discriminant function was found to be D=-6.2+0.17 X_2+5.791 X_7-0.479 X_9+0.038 X_{10}. This equation reveals that the characters canopy spread (X_2), Fruit set (X_7), LD Ratio (X_9) and Fruit weight (X_{10}) are the most important characters that discriminant the two groups. To test the statistical hypothesis of no difference in mean vectors (μ_1 and μ_2) of ten characters for these two groups; the value of Wilk's lambda (Λ) was

obtained to be 0.383 and which, in turn, gave the computed value of chi-square (χ^2) as 92.139 and, hence, the hypothesis of equality of group mean vectors was rejected. Having found that the groups differ statistically, the trees were assigned to group I (High yielder) if D≥ m otherwise to group II (Low yielder), where m=0.05 is the average of groups centroids. The groups formed on the basis of allocation rule were subjected to Principal Components Analysis and population-wise the results are discussed below:

3.1. Population-I (high yielder)

The main results of Principal Component Analysis pertaining to this population have been presented in Table 1 and Figure 1. Perusal of Table 1 revealed that three of the ten Principal Components (PCs) have Eigen values greater than unity (Gutman's lower bound) which played the main role in the analysis. These components explain 33.29%, 22.24% and 13.55% respectively of the total variation. Together they account for 69.08% of total variation of the original variables.
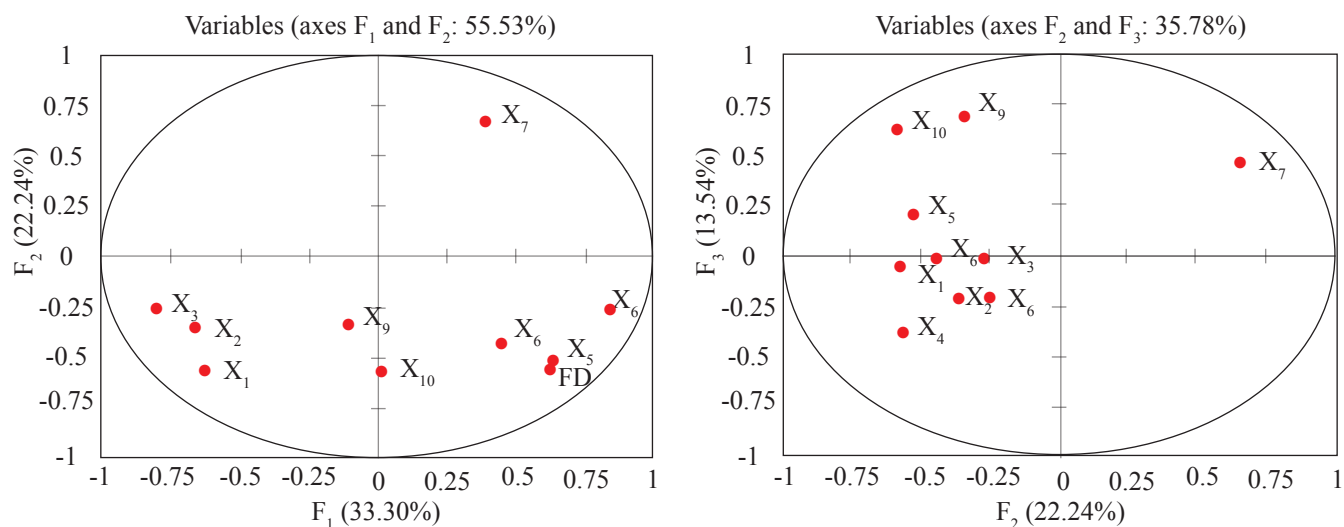
The variables loading for first principal component is highest for canopy spread, trunk girth and crop density. This component was interpreted as vegetative characteristics. The second principal component comprised of Tree Height, flower density and fruit set which were collectively termed as Plant vigour and flowering characteristics. The third principal component comprised of LD Ratio, fruit weight and Fruit Set which was interpreted as Fruiting Characteristics.

3.2. Population-II (Low yielder)

The summary of Principal Component Analysis of the data

Table 1: Eigenvectors of the Principal Component Analysis of high yielder group

Variables	Principal component 1	Principal component 2	Principal component 3
Tree height (X_1)	-0.342	-0.381	-0.046
Canopy spread (X_2)	-0.360	-0.236	-0.187
Trunk girth (X_3)	-0.436	-0.179	-0.019
Flower density (X_4)	0.343	-0.371	-0.335
Flower density index (X_5)	0.349	-0.351	0.168
Flowering intensity (X_6)	0.250	-0.291	-0.024
Fruit set (X_7)	0.215	0.441	0.394
Crop density (X_8)	0.460	-0.163	-0.182
LD ratio (X_9)	-0.059	-0.226	0.590
Weight of fruit (X_{10})	0.009	-0.388	0.534
Eigen value	3.330	2.224	1.354
% of variance	33.298	22.236	13.545
Cum. % of variance	33.298	55.534	69.078

Variables (axes F_1 and F_2: 55.53%)

Variables (axes F_2 and F_3: 35.78%)

pertaining to 'Low Yielder' group have been presented in Table 2 and Figure 2. Three principal components have been retained for the analysis. Together they account for 69.99% of the total variation of original variables. Table 2 explains that the variables loading for first principal component (PC_1) has Eigen value 3.57 and it is high for four characters Tree height, canopy spread, tree girth and crop density and explains 35.68% of total variation in the data set. This component may be interpreted as Plant Vigour. The second principal component (PC_2) explains 22.22% of total variation. In this component the variables viz., Flower density, flower density index, flowering intensity' and fruit weight had high weight and all together may be interpreted as yield and flowering characteristics. The third principal component (PC_3) explains 12.09% of total variation. The variables fruit weight and LD ratio gave maximum weight.

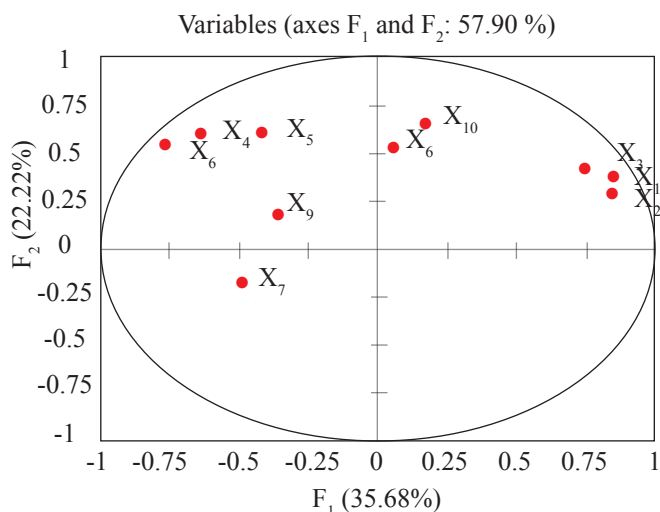This component may be interpreted as fruiting characteristics.

Variables (axes F_1 and F_2: 57.90 %)

Table 2: Eigenvectors PC analysis of low yielder group

Variables	PC_1	PC_2	PC_3
Tree height (X_1)	0.454	0.252	0.079
Canopy spread (X_2)	0.452	0.192	-0.076
Trunk girth (X_3)	0.398	0.286	0.123
Flower density (X_4)	-0.336	0.407	-0.089
Flower density index (X_5)	-0.219	0.409	-0.283
Flowering intensity (X_6)	0.032	0.361	-0.442
Fruit set (X_7)	-0.256	-0.113	0.273
Crop density (X_8)	-0.401	0.364	-0.004
Ld ratio (X_9)	-0.186	0.126	0.584
Weight of fruit (X_{10})	0.093	0.439	0.523
Eigen value	3.568	2.222	1.210
% of variance	35.679	22.219	12.099
Cum. % of Variance	35.679	57.898	69.997

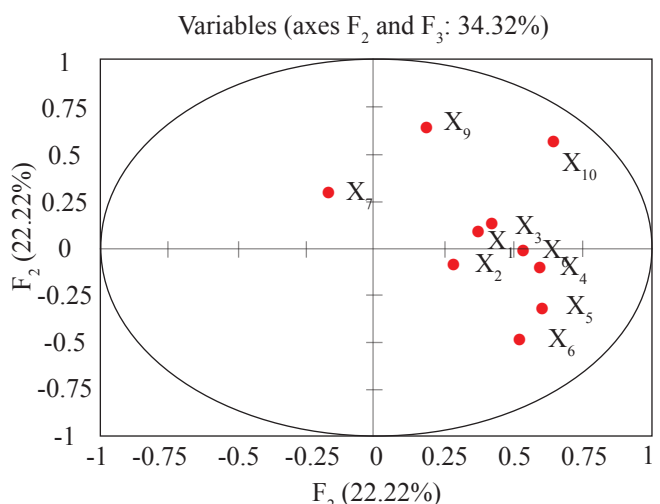Variables (axes F_2 and F_3: 34.32%)

Figure 2: Correlations between variables and factors (Low yielders)

4. Conclusion

The discriminant function revealed that canopy spread, fruit set, LD ratio and fruit weight are the most important characteristics that discriminate the groups. The groups were subjected to Principal Component Analysis. In both populations, three Principal Components were extracted, which played the main role in the analysis. Thus, the Principal Component Analysis has brought out some of the basic components associated with morphological characters of apple and could be considered as important tool in explanatory work for optimizing apple productivity.

5. References

Anderson, T.W., 1958. An Introduction to Multivariate Statistical Analysis. John Wiley & Co., London, 752.

Anonymous, 2014. National Horticultural Board Statistical Database. Shimla.

Belie, N.D., Smedt, V.D., Baerdemacker, J.D., 2000. Principal component analysis of chewing sound to detect differences in apple crispness. Post Harvest Biology and Technology 18(2), 109−119.

Bernard, J.M., Gregory, A.R., 1991. Principal components regression to mitigate the effects of multicollinearity. Forest Science 23(3), 291−296.

Bryant, E.C., 1966. Statistical Analysis. McGrew Hill, Tokyo, 321.

Chernudubov, A.T., 1994. The structure of the *Pinus sylvestris* L. in the Insular Pine forests of the South Russian plain. Silvae Genetica 43(4), 123−132.

Cox, D.R., 1958. Some problems connected with statistical inference. Annuals of Mathematical Statistics 29(2), 357−372.

Eaton, G.W., Lapins, K.O., 1970. Identification of standard and compact apple trees by discriminant function analysis. Journal of Applied ecology 7, 267−272.

Fisher, M., 1993. Use of discriminant analysis in apple root stock selection. 58(3), 137−143.

Holland, C., 1969. Forest typology and dendroecological studies of mountain pine forests in the French alps. Ecologie 26(4), 195−213.

Lezzoni, A., Pritts, M.P., 1991. Application of principal component analysis to horticultural research. Horticultural Science 26(4), 334−338.

Qurrie, A.J., Ganeshanandam, S., Noiton, D.A., Garrick, D., Shelbourne, C and Oraguzie, N. 2000. Quantative evalution of apple (*Malus domestica* Borkh.) fruit shape by principal component analysis of Fourier descriptors. Euphytica 111(3), 219−227.

Schrevens, E., 1995. Quality evolution of fruits by means of multivariate analysis on non-destructive parameters. Acta Horticulture 379, 569−578.

Stein, C., 1945., A two sample test for a linear hypothesis whose power is independent of the variance. Annuals of Mathematical Statistics 43, 243−258.