IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Study of various Multivariate techniques on Forestry: A case study of Wild Pomegranate (Punica granatum L)

Pawan Kumar, RK Gupta, Tara Gupta, PK Mahajan, Anmol Negi. and Sukhdeep Kaur
Department of Basic Sciences, College of Forestry, Dr YS Parmar UHF, Nauni, Solan (173230)

Abstract: The present study deals with the cluster, discriminant and principal component analysis. The data for the present study was collected. Data on different morphological and seedling characters, namely tree height (m), tree diameter (cm), crown spread E-W (m), crown spread N-S (m), fruit weight (g), leaf length(cm), internodal length(cm), collar diameter (mm), number of branches per plant and leaf petiole (cm) were considered from five different districts comprising of ten seed sources. Two seed sources were considered from each district viz; Narag and Neripul from Sirmour district, Waknaghat and Sadhupul from Solan district, Sundernagar and Rewalsar from Mandi district, Mohal and Banjar from Kullu district, Basantpur and Sunni from Shimla district. Different morphological and seedling characteristics of ten seed sources were evaluated. Cluster analysis was performed and the seed sources were grouped into three clusters. Discriminant analysis was carried out to categorize the seed sources into high and low yielders. Tree diameter, Tree height and crown spread E-W are the most important characters that discriminate the two groups. Six seed sources from Mandi, Kullu and Shimla district were high yielder whereas other four seed sources were low yielder. Three principal components (PC_S) were extracted out of ten which explained 34.675, 23.002, and 11.587 per cent of the total variation respectively amounting to 69.26 per cent of total variation.

Key words: Principle component analysis, cluster analysis, discriminant analysis, *Punica granatum* L, Wild Pomegranate

Introduction:

Wild pomegranate (*Punica granatum* L.) is commonly known as daru, dalim or dadima. It is one of the oldest known edible fruit crop. It is believed to be originated in South West Asia, probably in Iran and some adjoining countries (De Candolle, 1967). It is a deciduous agro-forestry as well as horticulture cash crop. It belongs to family Punicaceae. It grows as shrub or small tree and is often armed. It is found throughout the Mediterranean region. In India, it is found in vast tract of hill slopes of Himachal Pradesh, Jammu & Kashmir and Uttarakhand at an altitude of 900 to 1800 m amsl. In Himachal Pradesh, it is distributed in some pockets of Solan, Sirmour, Mandi, Shimla, Kullu and Chamba districts (Bhrot, 1998). The most important centre of wild pomegranate in Himachal Pradesh is Darlaghat, which is about 40 km from Shimla.

Wild pomegranate grows well on slightly hot climate characterized by dry summer and prolonged winter. This species provides a good source of income to villagers and farmers. Seed with aril are sun dried and commercially marketed as anardana. Anardana is a good source of Sugar, Vitamin C and iron. The powdered flower buds are used in bronchitis. Fruit rind, juice of leaves and young fruits are used for the treatment of gastrointestinal disorders. The fruit rind, root, stem bark and leaves are good sources of tannin, It is one of the esteemed dessert fruit and is very much liked by people for its cool refreshing juice, taste and being highly valued for its nutritional and medicinal properties. The pomegranate and its usage and deeply embedded in human history, and utilization is found in many ancient human cultures as food and as a medical remedy. Despite this fact, pomegranate culture has been restricted to generally considered as a minor crop. The pomegranate tree requires a long, hot and dry season in order to produce good yield of high quality fruit.

The yield of the tree may get affected by many growth characteristics, but the researchers are always interested to identify those set of characters (parameters) which significantly affect the yield of the tree. Keeping in view the economic importance of this species present study was conducted with the objective to discriminate high yielder and low yielder seed sources and to assess the relative contribution of different morphological and seedling characteristics. Five districts were selected randomly from the state of Himachal Pradesh and from each district two seed sources were selected randomly comprising of ten seed sources.

Material and methods

The data on various morphological characteristics were collected from each seed sources for the present study viz. Tree height (m), Tree diameter (cm), Crown spread E-W (m), Crown spread N-S (m), Fruit weight (g), Leaf length (cm), Internodal length (cm), also nursery was raised and data on characteristics i.e Collar diameter (mm), Number of branches per plant and Leaf petiole (cm) were recorded.

Cluster	No. of Seed sources	Seed source
I	3	Narag , Neripul, Sadhupul
II	4	Rewalsar, Sundernagar, Sunni, Mohal
III	3	Waknaghat, , Basantpur, Banjar

Different multivariate tools like Discriminant Analysis, Cluster analysis and Principle component Analysis were used to gather the information about the concerned objective.

Cluster Analysis: When certain multivariate populations are found to be heterogeneous and there is need to find out which subsets of populations are most alike and which are least alike. The cluster analysis was done by Mahalanobis. D^2 statistic as suggested by Rao (1952).

Discriminant analysis: Fisher (1936) introduced discriminant function analysis. With the help of this technique a set of multiple measurement were used to provide a discriminant function (linear) in the observation and having the property that, better than any other linear functions. Thus, the problem is reduced to that of single variable by choosing a linear component of the original variables and by constructing a statistic suitable for univariate case. The maximized value of this statistic obtained by a suitable choice of coefficients is taken as the appropriate test criteria. Under these circumstances the method of discriminant analysis frequently obtains more satisfactory result than a regression or correlation analysis. The real adequacy of the discriminant function, however, must be determined by how well it discriminates between classes I and II on a fresh sample of data. Ten seed sources of *Punica granatum* L. were divided into high and low yielder groups.

Principal Component Analysis: It is a multivariate statistical technique to reduce the data with large number of correlated variables in to a substantially smaller set of new variables. A principal component analysis is concerned with explaining the variance - covariance structure of a set of variables through a few linear combinations of these variables. Its general objectives are Data reduction and Interpretation The aim of principal component analysis is to ascertain new variables, called principal components, which carry most of the information present in original variable. Principal component analysis technique were used to identify the important characteristics contributing towards the yield of *Punica granatum* L.

Results:

CLUSTER ANALYSIS: On the basis of performance of various characteristics, the clustering pattern of different seed sources of *Punica granatum* has been presented in the Table 1. All the seed sources were grouped into 3 clusters. Maximum (4) number of seed sources i.e Rewalsar, Sundernagar, Sunni and Mohal were accommodated in cluster II followed by three seed sources in cluster I and cluster III respectively.

Table 1: Clustering membership**.**

Cluster means for different characters to assess the considerable amount of variation which shows the existence of diversity among seed sources. The cluster means of various characters are presented in Table 2. Maximum height was recorded in cluster II (6.95 m) and minimum height was recorded in cluster III (6.51 m). Maximum diameter was recorded in cluster I (13.12 cm) and minimum diameter was observed in cluster II (11.91 cm). Maximum crown spread E-W was recorded in cluster I (4.62 m) and minimum crown spread E-W was observed in cluster II (4.24 m). Maximum crown spread N-S was recorded in cluster I (6.47 m) and minimum crown spread N-S was observed in cluster II (6.06 m). Maximum fruit weight was observed in cluster II (26.43 g) and minimum was observed in cluster I (24.36 g). Maximum leaf length was recorded in cluster II (5.07 cm) and minimum leaf length was observed in cluster III (4.57 cm). Maximum internodal length was recorded in cluster II (3.57 cm) and minimum internodal length was observed in cluster I (3.35 cm). Maximum collar

diameter was recorded in cluster II (2.30 mm) and minimum collar diameter was observed in cluster I (1.81 mm). Maximum number of branches was recorded in cluster II (7.76) and minimum number of branches was observed in cluster I (6.35). Maximum leaf petiole was recorded in cluster I (0.47 cm) and minimum leaf petiole was observed in cluster III (0.43 cm).

Table 2: Cluster means.

Characters	Cluster		
	I	II	III
Tree height (m)	6.73	**6.95**	6.51
Tree diameter (cm)	**13.12**	11.91	12.06
Crown spread E-W (m)	**4.62**	4.24	4.37
Crown spread N-S (m)	**6.47**	6.06	6.20
Fruit weight (g)	24.36	**26.43**	25.65
Leaf length (cm)	4.58	**5.07**	4.57
Internodal length (cm)	3.35	**3.57**	3.43
Collar diameter (mm)	1.81	**2.30**	1.82
Number of branches	6.35	**7.76**	6.77
Leaf petiole (cm)	**0.47**	0.45	0.43

Average inter cluster distance values are presented in the Table 4.3. The inter cluster distance D^2 value was highest between cluster I and cluster II (18.73) and the lowest inter cluster distance D^2 value was observed in cluster II and cluster III (8.98). This indicates that seed sources included in cluster I and II have wide genetic diversity and could be used in hybridization programme aimed at direct selection for characters or improvement of seed sources. Characters like tree height, tree diameter and crown spread had more contribution towards genetic divergence, hence these characters are major determinants of genetic diversity in the present set of seed sources.

Table 3: Average inter cluster distance (D^2)

Clusters	I	II	III
I		18.73	10.18
II	18.73		8.98
III	10.18	8.98	

PRINCIPAL COMPONENT ANALYSIS

To interpret the data in a more meaningful form, it is necessary to reduce the number of variables to a few interpretable linear combinations of variables. Thus, principal component analysis was employed to reduce the observed variables into number of principal components that will account for most of the variation in observed variables.

The result pertaining to principal component analysis has been presented in Table 4.4. Table 4.14 revealed that three of ten principal components (PC_S) had eigen values greater than unity and therefore these three principal components are playing main role in the analysis. Thus first three principal components had been retained in the analysis, which explained 69.263 per cent of the total variation.

The first principal component had eigen value 3.468 and it explained 34.675 per cent of the total variation. The second principal component had eigen value 2.300 and explained 23.002 per cent of the total variation. The third principal component had eigen value 1.159 and explained 11.587 per- cent of the total variation.

Table 4: Eigen vectors of principal component analysis

Variables	PC_1	PC_2	PC_3
Tree height (X_1)	**0.674**	0.639	-0.079
Tree diameter (X_2)	0.520	**0.662**	-0.033
Crown spread E-W (X_3)	0.513	**0.629**	0.389
Crown spread N-S (X_4)	**0.418**	0.228	-0.327
Fruit weight (X_5)	-0.107	0.336	**0.750**
Leaf length (X_6)	-0.195	**0.374**	-0.353
Internodal length (X_7)	**0.847**	-0.441	0.089
Collar diameter (X_8)	**0.851**	-0.475	0.108
Number of branches (X_9)	**0.812**	-0.509	0.029
Leaf petiole (X_{10})	**0.394**	0.272	-0.330
Eigen values	3.468	2.300	1.159
% of variance	34.675	23.002	11.587
Cumulative % of variance	34.675	57.677	69.263

The first principal component was linear combination of tree height, crown spread N-S, inter-nodal length, collar diameter, number of branches and leaf petiole. This component may be interpreted as stem growth. The second principal component was linear combination of tree diameter, crown spread E-W and leaf length which represent tree characteristics. The third principal component comprised of only fruit weight which represent fruit characteristics.

DISCRIMINANT ANALYSIS: The approach of categorizing high and low yielder seed sources on the basis of randomly selected characteristics is statistically weak. The discriminant analysis is a systematic and statistically valid procedure for this purpose. In the present study, the seed sources were first divided into two groups namely 'high yielder' and 'low yielder' on the basis of average value of fruit weight and discriminant function was fitted. The discriminant function was found to be:

$$D = -41.41 - 1.68\ X_1 + 0.37\ X_2 - 1.67\ X_3$$

Where D stands for fruit weight, X_1 for tree height, X_2 for tree diameter and X_3 for crown spread E-W. Thus, this equation reveals that the characters tree height, tree diameter and crown spread E-W are the most important characters to develop discriminate rule which discriminate the two groups. The value of Wilk's lambda (λ) was obtained to be 0.140 and which in turn, gave the computed value of chi square (χ^2) as 7.862. The seed sources were assigned to group 1 (High yielder) if $D \geq m$ otherwise to group 2 (low yielder), where m = 0.45 is the average of group centroids. The groups formed on the basis of allocation rule are given in Table 4.5. Six seed sources were classified as high yielder, whereas four seed sources were low yielder. The seed source Rewalsar and Sundernagar from Mandi district, Mohal and Banjar from Kullu district and Basantpur and Sunni from Shimla district were categorized into high yielder seed sources.

Table 5: High and Low yielder seed sources in location wise

High yielder Seed Sources	Low Yielder Seed Sources
Rewalsar	Narag
Sundernagar	Neripul
Mohal	Waknaghat
Banjar	Sadhupul
Basantpur	
Sunni	

Thus Shimla, Mandi and Kullu districts was found to be high yielder, whereas Solan and Sirmour districts were found to be low yielder seed sources.

Summary: All Seed sources were grouped into 3 clusters and maximum inter cluster distance was between cluster I and II.

Tree diameter (cm), tree height (m) and crown spread (m) were the most important characters to discriminate different seed sources into high and low yielder groups. Seed sources Rewalsar and Sundernagar from Mandi district, Mohal and Banjar from Kullu district and Basantpur and Sunni from Shimla district were categorized into high yielder seed sources.

Three principal components (PC_S) were extracted out of ten which explained 34.675, 23.002, and 11.587 per cent of the total variation respectively.

References:

Bhrot NP. 1998. *Genetical analysis of wild pomegranate (Punica granatum L.) for some growth ecological and quality characters.* PhD. Thesis, Dr. Y S Parmar University of Horticulture and Forestry, Nauni, Solan, HP, India.

DeCandole A. 1967. *Origin of cultivated plants*. Hafner Publishing, New York. 237p.

.

Rao R.1952. *Advanced Statistical Methods in Biometrical Research*. John Wiley and Sons Inc., new York.

Fisher M. 1993. Use of discriminant analysis in apple rootstock selection. *Annals of Eugenics* 58:137-143