Journal of Agriculture and Natural Resources (2020) 3(1): 69-76
ISSN: 2661-6270 (Print), ISSN: 2661-6289 (Online)
DOI: https://doi.org/10.3126/janr.v3i1.27092

Research Article

Multivariate analysis of soybean genotypes

Pallavi Kumari Singh[1*], Jiban Shrestha[1] and Ujjawal Kumar Singh Kushwaha[1]

[1]Nepal Agricultural Research Council, Agriculture Botany Division, Khumaltar, Lalitpur, Nepal
*Correspondence: pallaviid.singh@gmail.com
ORCID: https://orcid.org/0000-0003-0544-4447
Received: July 21, 2019; Accepted: October 29, 2019; Published: January 7, 2020

ABSTRACT

The experiments were conducted using randomized complete block design with three replications at the research field of Agriculture Botany Division, Khumaltar, Lalitpur, Nepal in 2016 and 2017 to evaluate sixteen soybean genotypes using multivariate analysis. The results showed the significant ($p < 0.05$) differences among genotypes for plant height, days to maturity plant and hundred seeds weight and grain yield. Cluster analysis based on these traits, sixteen soybean genotypes were divided the genotypes into four clusters. The soybean genotypes grouped into cluster 1 showed the highest value for days to maturity. The genotypes belonged to cluster 2 had the highest values for grain yield and plant height. The principle components analysis showed that PC1 and PC2 having eigen values the highest than unity explained 76.6% of total variability among soybean genotypes attributable to plant height, days to maturity, number of pods/plant, 100 seed weight and grain yield. The genotypes showing wide diversity in cluster and principle component analysis can be used as parents in hybridization programs to maximize the use of genetic diversity and expression of heterosis and develop high yielding soybean varieties.

Keywords: Clustering, grain yield, principal component analysis and Soybean

Correct citation: Singh, P. K, Shrestha, J., & Kushwaha, UKS. (2020). Multivariate analysis of soybean genotypes. *Journal of Agriculture and Natural Resources*, *3*(1), 69-76.
DOI: https://doi.org/10.3126/janr.v3i1.27092

INTRODUCTION

Soybean [*Glycine max* (L.) Merrill] is a papillionoid legume that varies widely in its adaptability, although it grows mostly in rain-fed soils, covering a wide range of climatic, soil and growing conditions (Lackey, 1977; Fageria et al., 1997). It is currently cultivated for food, feed and medicine in East and Southeast Asia. It is an important legume of mid hill of Nepal, grown as intercrop with maize or as sole crop in paddy bund. About 80% of total soybean area and production is occupied by mid-hills. However, soybean is becoming popular as sole crop in terai and inner terai due to high yield potential and high demand of soymeal in poultry industry (Pokhrel et al., 2014).

Multivariate analytical techniques that analyze multiple measurements for each genotype at the same time are widely used in the analysis of genetic diversity. Cluster Analysis and Principal Component Analysis (PCA) is currently the most widely used methods. They seem particularly well suited for analyzing components contributing to yield (Melchinger, 1993; Johns et al., 1997; Brown-Guedira et al., 2000). The cluster analysis help to separate the varieties into homogeneous groups from heterogeneous group such that those varieties within a group have a similar to each other in response pattern across the locations (Kandel et al., 2018). Many researchers have used principal component analysis to assess genetic variability among maize genotypes because it retrieves small numbers of components that account for most of the variations in the data (Asare, 2016). Principal Component Analysis (PC) divides the total variance into several factors. Classification of genotypes according to agronomic characteristics made using multi-factor techniques can shorten the period and the cost of improving the yield. Information about cluster analysis on soybean genotypes is not sufficient in Nepal. Therefore, these studies were conducted to evaluate the performance of soybean genotypes for grain yield and yield attributing traits at Khumaltar, Lalitpur in the natural environment.

MATERIALS AND METHODS

1.1 Experimental site

This experiment was carried out at the experimental field of Agriculture Botany Division, Khumaltar, Lalitpur, Nepal in 2016-2017. Geographically, this place is located at about $27^0$4' N north latitude and $85^0$2'E east longitude and altitude of 1350 m.

2.2 Plant materials

Twenty soybean genotypes were used as the experimental materials. All the genotypes were received from National Grain Legumes Research Program, Khajura, Banke, Nepal.

Table 1. List of soybean genotypes used in the study

SN	Genotypes	SN	Genotypes
1	Ankur	9	G-1873
2	Chaingmow 60-63	10	GC-82234-22C
3	CM 9106	11	LS-77-16-16
4	CM-9125	12	Ransom
5	CM-9133	13	Sathiya
6	CN-60	14	SB-0095
7	G-1872	15	TH -227
8	G-4508	16	VLS-1

2.3 Experimental Design

The experiment was laid out in a randomized complete block design with three replication. Plot size was 4 x 1 m^2 with spacing of 50cm x 10cm.

2.4 Intercultural operations

The plots were fertilized with poultry manure (5 t/ha) and chemical fertilizers at the rate 20:40:20 NPK kg/ha. The plots were hand weeded twice at 25 and 45 days after seeding. Other crop management practices were adopted as per the standard package of practices.

2.5 Data collection

Data on plant height, early plant stands/m^2, final plant stands/m^2, day to flowering, day to maturity, number of pods/plant, no. of unfilled pods/plant, hundred seed weight and grain yield were recorded.

Table 2. List of different traits and their measurements

SN	Traits	Method of measurement
1	Plant height (cm)	The height from the base of the plant to the tip of last leaf
2	Early plant stands/m^2	Number of plants per square meter
3	Final plant stands/m^2	Number of plants per square meter
4	Days to flowering	More than 75% flowering in the plot
5	Days to maturity	More than 75% maturity in the plot
6	Pods per plant (number)	Total number of pods with seed in a plant
7	No. of unfilled pods/plant	Total number of unfilled pods with seed in a plant
8	100-Seed wt (g)	One hundred seeds counted and then weighted
9	Grain yield (kg/ha)	Weighing the seeds produced in a plot and then converted into kg per ha

(Source: Singh and Shrestha, 2019)

2.6 Statistical Analysis

The experimental data were processed by using Excel 2010 and analyzed by using Genstat 13.2. The treatment means were compared by the Least Significant Difference (LSD) test at 5% level (Gomez & Gomez, 1984; Shrestha, 2019). The collected data were subjected to multivariable analysis was done using statistical software packages of Minitab ver.14 (Mohammadi & Prasanna, 2003).

RESULTS AND DISCUSSION

3.1 Overall performance of the genotypes

Genotypes were found significantly different for 100 seed weight, days to maturity, and plant height. Maturity days ranged from 123 (Sathiya) to 136 (Chaingmow 60-63) days. Plant height ranged from 56.2 cm (VLS-1) to 90.1 cm (Ankur). Grain yield ranged from 1080.89 kg/ha (TH-227) to 2060.30 kg/ha (CN-60).

Table 3. Agro-morphological traits of sixteen soybean genotypes at Khumaltar, Lalitpur planted during 2016 and 2017

Genotypes	100 SW (g)	DM	Plant ht (cm)	Pods/plant	GY (kg/ha)
Ankur	17.12	130	90.1	51	1631.57
Chaingmow 60-63	18.78	136	84.1	43	1975.72
CM-9106	18.14	133	84.3	43	1394.12
CM-9125	16.24	134	80.8	51	1625.91
CM-9133	18.19	133	75.4	45	1850.61
CN-60	19.65	128	86.5	41	2060.30
G-1872	18.20	132	90.0	52	1816.52
G-1873	19.29	133	76.8	48	1464.34
G-4508	17.79	130	79.9	43	1955.67
GC-82234-22C	17.21	128	83.7	46	1805.66
LS-77-16-16	16.63	124	72	37	1971.56
Ransom	16.96	132	73.4	44	1711.26
Sathiya	20.43	123	65.9	42	1483.31
SB-0095	19.12	135	83.0	40	1216.89
TH-227	11.55	116	57.1	58	1080.89
VLS-1	17.45	123	56.2	37	1745.79
GM	17.67	129	77.5	45	1674.38
P value	0.001	0.001	0.001	0.091	0.001
CV%	7.9	7.4	21.6	29.6	27.6
LSD$_{0.05}$	1.298	8.934	15.61	12.5	430.528

DF= days to flowering, ES=early stand, FS= final stand, DM=days to maturity, SW= seed weight (g), GY=grain yield (kg/ha)

3.2 Cluster analysis

Cluster analysis showed that cluster 1 was comprised of 7 genotype, cluster 2 of 4, cluster 3 of 3, and cluster 4 of 2 (Figure 1). The soybean genotypes grouped into cluster 1 showed highest value for days to maturity. The genotypes belong to cluster 2 to showed the highest values for grainn yield and plant height. Estimating the relationship between genetic diversity and relationships among germplasm accessions facilitates the selection of parents with different genetic background necessary for the breeding program (Murphy et al., 1986; Souza & Sorrels, 1991; Shrestha, 2016). This clustering into natural groups is represented by a dendrogram, which gradually divides interventions into smaller and smaller groups (KC, 2001).

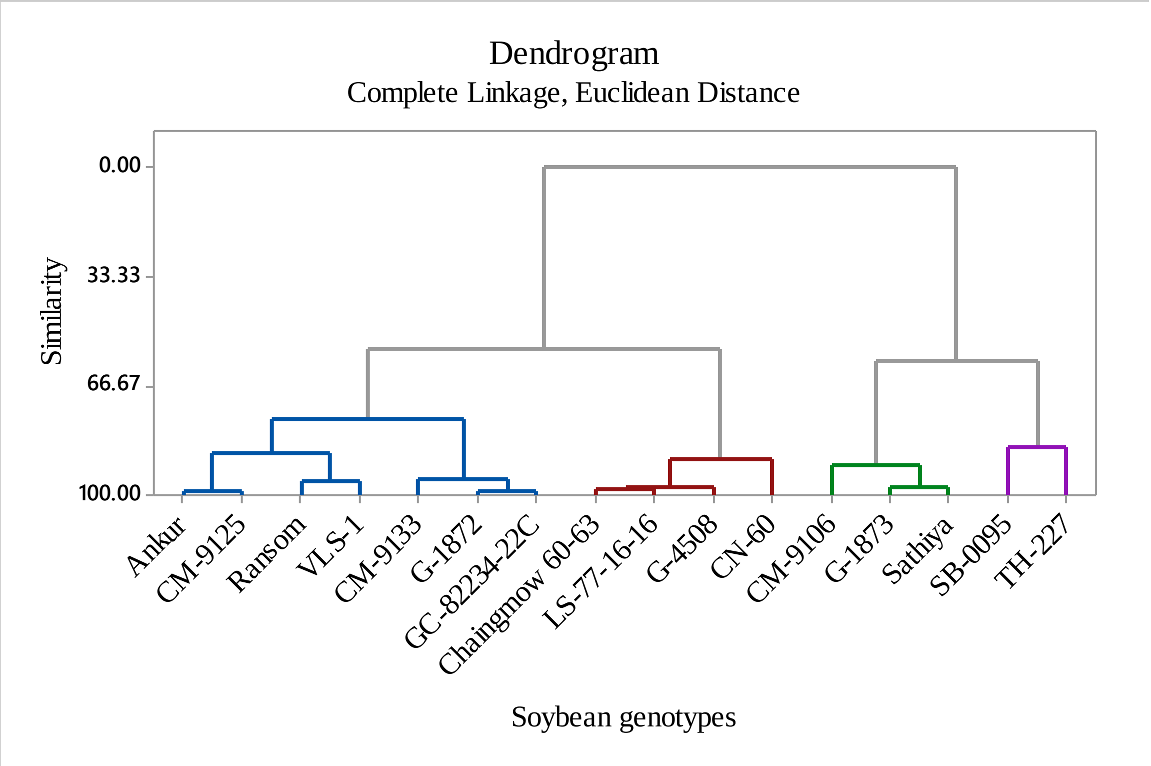

Figure 1. Cluster analysis of sixteen soybean genotypes evaluated for agro-morphological traits in 2016 at Khumaltar, Lalitpur, Nepal

Table 3. Grouping of Sixteen soybean genotypes into four clusters based on agromorphological traits

Cluster1	Cluster2	Cluster3	Cluster4
Ankur, CM-9125, Ransom, VLS-1, CM-9133, G-1872, GC-82234-22C	Chaingmow60-63, LS-77-16-16, G-45508, CN-60	CM-9106, G-1873, Sathiya	SB-0095, TH-227

Table 4. Agromorphological traits of sixteen soybean genotypes within and among four clusters

SN	Variable	Cluster1	Cluster2	Cluster3	Cluster4	Grand centroid
1	Plant height (cm)	78.51	80.63	75.67	70.05	77.45
2	Days to maturity	130.29	129.5	129.67	125.5	129.38
3	Number of pods/plant	46.57	41	44.33	49	45.06
4	100 seed weight (g)	17.34	18.21	19.29	15.34	17.67
5	Grain yield (kg/ha)	1741.05	1990.81	1447.26	1148.89	1674.38

3.3 Principle component analysis

Principal component analysis is often used to reduce the variables and genotypes of groups. Among these biometrical procedures, the main advantage of principal component analysis (PCA) is that each genotype can be attributed to only one group and also reflects the importance of the greatest contribution to the total variation in each differentiation axis (Sharma, 1998). Analysis of the principle components divides the total variance into several factors. By removing the relationship between the main component variables and the Eigen value, it reduces the dimensionality of multivariate data representing the variance of the main component (Matus, 1999). Principal component analysis is a powerful tool for obtaining parental lines for a successful breeding program (Akter et al., 2009).

Table 5. Eigen values, the percent of variance and cumulative variance of sixteen soybean genotypes for agro-morphological traits

SN	Variable	PC1	PC2	PC3	PC4	PC5
1	Eigen value	2.4848	1.3464	0.723	0.2681	0.1777
2	Proportion	0.497	0.269	0.145	0.054	0.036
3	Cumulative	0.497	0.766	0.911	0.964	1
4	Plant height (cm)	0.465	0.508	0.169	0.24	-0.662
5	Days to maturity	0.5	0.395	-0.215	-0.643	0.366
6	Number of pods/plant	-0.321	0.683	0.224	0.374	0.491
7	100 seed weight(g)	0.528	-0.198	-0.422	0.624	0.338
8	Grain yield (kg/ha)	0.388	-0.285	0.835	0.001	0.268

In this experiment, the principle component (PC) analysis partitioned the total variance into 4 PCs contributing maximum to the total diversity among the genotypes due to the study of various traits. Principal component analysis showed first 2 PCs having Eigen value >1 explaining 76.6% of the total variation (Table 5). The first principal component (PC1) which explained 49.7% variation was associated mainly with number of Plant height, days to maturity and hundred seed weight. The second principal component (PC2) was responsible for about 26.9% of the variation and was mainly related to plant height, days to maturity and number of pods per plant.

CONCLUSION

The soybean genotypes exhibited a wide range of variability for most of the traits. Genotypes CN-60 and Chaingmow 60-63 showed highest grain yield 2060 kg/ha and 1975 kg/ha respectively. The genotypes grouped in to cluster 2 were good because of their higher grain yield. Therefore, they can be grown for mild hills of Nepal for higher grain yield production

ACKNOWLEDGEMENT

The authors gratefully acknowledged the financial support from Nepal Agricultural Research Council, Agriculture Botany Division, Khumaltar, Lalitpur for this experiment. They were also thankful to National Grain Legumes Research Program, Khajura, Banke and Agronomy Division, Khumaltar for providing soybean genotypes for this experiment.

Authors' contribution

P.K. Singh	Conducted of experiment, data collection, soil analysis, data analysis, and manuscript writing
J. Shrestha	Data analysis, literature review and manuscript writing
UKS.Kushwaha	Data analysis and initial draft writing of manuscript

Conflict of interest

The authors declare that there is no conflict of interest regarding publication of this manuscript.

REFERENCES

Akter, A., Hasan, M.J., Paul, A.K., Mutlib, M.M., & Hossain, M.K. (2009). Selection of parent for improvement of restorer line in rice (*Oryza sativa* L.). *SAARC Journal of Agriculture, 7*, 43-50.

Asare, S., Tetteh, A.Y., Twumasi, P., Adade, K.B., & Akromah, R.A. (2016). Genetic diversity. *African Journal of Plant Science, 10*, 246-257.

Brown-Guedira, G.L., Thompson, J.A., Nelson, R.L., & Warburton, M. L. (2000). Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. *Crop Science, 40*, 815–823

Fageria, N.K., Baligar, V.C., & Jones, C.A. (1997). Growth and mineral nutrition of field crops. 2nd Ed. Marcel Dekker, Inc., New York.

Gomez, K. A., & Gomez, A.A. (1984). Statistical Procedure for Agricultural Research (2nd ed.). Int. Rice Res. Inst. And Willey, New York pp. 28-192.

Johns, M.A., Skrotch, P.W., Neinhuis, J., Hinrichsen, P., Bascur, G., & Munoz-Schick, C. (1997). Gene pool classification of common bean landraces from Chile based on RAPD and morphological data. *Crop Science, 37*, 605–613.

Kandel, M., Ghimire, S.K., Ojha, B.R., & Shrestha J. (2018). Cluster analysis among the Maize inbred lines (*Zea mays* L.) under heat stress condition. *International Journal of Global Science Research, 5*(1), 677-684.

KC, R.B., Sharma, M.D., Panthee D., & Gautam, D.M. (2001). Physio-morphological characterization of indigenous garlic germplasm of Nepal. Master Thesis. Institute of Agriculture and Animal Science Rampur, Nepal.

Lackey, J.A. (1977). A synopsis of Phaseoleae (Leguminosae, Papilionoideae). Ph.D. dissertation. Iowa State University. Ames, Iowa.

Matus, I., Gonzalez, M.I., & del. Pozo A. (1999). Evaluation of phenotypic variation in a Chilean collection of garlic clones using multivariate analysis. *Plant genetic resources Newsletter, 117*, 31-36.

Melchinger, A.E. (1993). Use of RFLP markers for analyses of genetic relationships among breeding materials and prediction of hybrid. International Crop Science Congress, 1st, Ames, IA. July 1992. CSSA, Madison.

Mohammadi, S.A., & Prasanna, B. M. (2003). Analysis of genetic diversity in crop plants-salient statistical tools and considerations. *Crop science, 43*(4), 1235-1248.

Murphy, J.P., Cox, T.S., & Rodgers D.M. (1986). Cluster analysis of red winter wheat cultivars based upon coefficients of parentage. *Crop Science, 26*, 672-676.

Pokhrel, K.R., Ghimire T.B., Pokhrel, A., & BC, K.B. (2014). Soybean Varietal Investigation in River Basin Area of Mid-Western Nepal. Proceedings of the 27th National Summer Crops Workshop (Vol: 1). Nepal Agricultural Research Council.

Sharma, J.R. (1998). Statistical and biometrical techniques in plant breeding. New Age International Limited Publishers, New Delhi, India.

Shrestha, J. (2016). Cluster Analysis of Maize Inbred Lines. *Journal of Nepal Agricultural Research Council, 2*, 33-36.

Shrestha, J. (2019). P-Value: A true test of significance in agricultural research. Retrieved from *https://www.linkedin.com/pulse/p-value-test-significance-agricultural-research-jiban-shrestha/*

Singh, P.K., & Shrestha, J. (2019). Evaluation of soybean [*Glycine max* (L.) Merrill] genotypes for agro-morphological traits using multivariate analysis. *Nepalese Journal of Agricultural Sciences, 18,* 100-107.

Souza, E., & Sorrells, M.E. (1991). Relationships among 70 North American oat germplasm: I. Cluster Analysis using quantitative characters. *Crop Science, 31*, 599-605.