

# Multivariate Analysis Among Soybean (*Glycine max* L.) Genotypes in Southwest Ethiopia

Masreshaw Yirga, Afework Legesse

Ethiopian Institutes of Agricultural Research, Jimma Agricultural Research Center, Pulse and Oil Crop Breeding and Genetics, Jimma, Ethiopia

**Email address:**

yirgamasresha@gmail.com (Masreshaw Yirga)

**To cite this article:**

Masreshaw Yirga, Afework Legesse. Multivariate Analysis Among Soybean (*Glycine max* L.) Genotypes in Southwest Ethiopia. *American Journal of Bioscience and Bioengineering*. Vol. 11, No. 2, 2023, pp. 20-26. doi: 10.11648/j.bio.20231102.12

**Received:** July 24, 2023; **Accepted:** August 9, 2023; **Published:** August 22, 2023

---

**Abstract:** Genetic variability information on any crop germplasm is imperative for conservation and effective utilization in the breeding program. The field experiment was conducted to estimate the extent of genetic variability of 64 soybean genotypes for grain yield and other agronomic traits at Jimma and Metu in 2017 and 2018 main cropping seasons. The trial was laid down using 8×8 simple lattice design. The data was subjected to statistical analysis using R-software. The combined analysis of variance revealed the presence of significant ( $P<0.01$ ) and wide range of variation among the tested genotypes for all of the traits. Based on the mean performance, latest maturing genotype was PI567104B (145 days), while the earliest was PI615437 (105 days). Genotype, PI567104B was the tallest in plant height (149.01 cm) while the shortest was PI507004 (44.24cm). Genotype PI567090 was found moderately susceptible in soybean rust (25.52%), while genotypes PI594538A was found the most tolerant (3.78%). Maximum hundred seed weight (24.01gm) was found from genotype PI506677, whereas the minimum seed weight (8.32gm) was recorded from genotype PI567068A. Coker240 scored maximum grain yield (3.09 t/ha) followed by genotype PI567104B (3.00 t/ha), while the minimum yield was scored from PI416826A (1.33tha<sup>-1</sup>). Cluster analysis categorized 64 soybean genotypes into five clusters. The maximum inter cluster distance was found between clusters-I and V, suggesting best recombinants can be found by crossing genotypes in these clusters. Principal component analysis (PCA) revealed that, the 1<sup>st</sup> four PCA with Eigen values exceeding one were responsible for about 88.74% of the total variation. Out of the entire variations, 1<sup>st</sup>PCA and the 2<sup>nd</sup>PCA accounted for more than two third of the total variations (68.47%). Generally, the present study indicated the existence of enormous genetic variability among soybean genotypes for various important morphological traits. Therefore, information and genetic variability obtained in this finding could be used to plan conservation, effective crossing and line selection in soybean variety improvement programs.

**Keywords:** Soybean, Cluster, D<sup>2</sup>, Eigen Value, PCA

---

## 1. Introduction

Soybean is one of the most important pulses and oil crops for food, feed and industries in the world. This is due to its high grain nutritional value with quality protein (35-40%) and vegetable oil (18-22%) content [1, 2]. The by-product or soy meal is also the main source of protein used for farm animal; livestock, poultry, and pig feeding [3, 4]. It is also important for soil fertility improvement and sustainable crop production due to its biological nitrogen fixing ability and the ability to break lifecycles of pests and diseases in cereal rotation system [5].

The vast majority of low land to mid land agro ecology

of Ethiopia is the potential arable land suitable for soybean production. However, according to CSA [6] report area coverage (83,797.2 hectare), annual soybean production (208,676.4 ton) and productivity (2.49t/ha) in Ethiopia is below its potential. Among the determinate factors for low productivity of the crop is lack of widely adapted improved variety, limited use of recommended technology, disease mainly soybean rust and limited genetic information.

Multivariate analysis is a useful tool in quantifying the degree of genotypic divergence among biological

populations and to assess the relative contribution of different components to the total divergence levels [7, 8]. Such a study also permits to select the genetically diverse parents to obtain the desirable recombinant in the segregating populations upon crossing. In the hybridization programs, inclusion of more diverse parents has been observed to increase the chance for obtaining strong heterosis and giving broad spectrum of variability in segregating generations [9].

Introduction and local hybridization are paramount important sources of germplasm to initiate soybean variety development program in Ethiopia. Since the inception of soybean research in Ethiopia, a number of germplasm were enhanced from these two methods. Some multivariate analysis studies were conducted on some germplasm by different researchers [10-13], each reported the presence of considerable genetic variations among the tested genotypes. Bearing in mind this, the study was conducted to estimate the extent of genetic variability among soybean genotypes for effective conservation and use in the breeding program with respect to quantitative traits.

## 2. Materials and Methods

### 2.1. Description of the Experimental Site

The experiment was conducted at Jimma and Metu, Southwestern Ethiopia during 2017-2018 main cropping seasons. Jimma has an altitude of 1754 m. a. s. l with the average annual temperature of from 26.3-26.3°C with its mean annual rainfall of 1,572mm, while Metu has an altitude of 1558m. a. s. l. and the mean annual temperature ranges from 12.7-28.9°C with annual rainfall of 1829 mm, The major soil type in southwest Ethiopia is Nitosols [14].

### 2.2. Experimental Materials, Design, and Field Management

Sixty four soybean genotypes including three standard check varieties (AFGAT, Clark-63K and Nyala) were evaluated in this study using simple lattice design. The genotypes were introduced from external source as rust tolerant genotypes (Table 1). Each genotype was planted in a plot of four rows and four meter length with regular spacing of five cm between plants and 60cm between rows. All the agronomic management practices were applied for the trial as per the recommendation.

Table 1. Details of the 64 Soybean genotypes.

Geno	designation	description	Source of materials	Geno.	designation	description	Source of materials
1	PI567099A	pure line	USA	33	PI567025A	USA	pure line
2	G 7955-C3RPP (C1)	Pipe line	USA	34	PI605838	USA	pure line
3	MKSOY-2N white (C2)	Variety	USA	35	PI567090	USA	pure line
4	PI605823	pure line	USA	36	PI605773	USA	pure line
5	PI567020A	pure line	USA	37	PI416810	USA	pure line
6	PI567102B	pure line	USA	38	PI605854B	USA	pure line
7	PI471904	pure line	USA	39	PI594767A	USA	pure line
8	PI567039	pure line	USA	40	PI566989A	USA	pure line
9	PI567058D	pure line	USA	41	PI200466	USA	pure line
10	PI605824A	pure line	USA	42	PI635999	USA	pure line
11	PI578457A	pure line	USA	43	PI423960A	USA	pure line
12	PI567046C	pure line	USA	44	PI417208	USA	pure line
13	PI615445	pure line	USA	45	PI567059	USA	pure line
14	PI567180	pure line	USA	46	PI567053	USA	pure line
15	PI606405	pure line	USA	47	PI567068A	USA	pure line
16	PI594760B	pure line	USA	48	PI567034	USA	pure line
17	PI605891A	pure line	USA	49	PI567189A	USA	pure line
18	PI567104B	pure line	USA	50	PI594538A	USA	pure line
19	PI567054C	pure line	USA	51	AFGAT (C5)	Ethiopia	Variety
20	PI605891B	pure line	USA	52	PI230970	USA	pure line
21	PI567069A	pure line	USA	53	PI615437	USA	pure line
22	Cocker 240	pure line	USA	54	PI416886	USA	pure line
23	PI606397B	pure line	USA	55	PI417085	USA	pure line
24	PI567056A	pure line	USA	56	PI203398	USA	pure line
25	PI628932	pure line	USA	57	PI423972	USA	pure line
26	PI587905	pure line	USA	58	PI423960B	USA	pure line
27	PI567061	pure line	USA	59	PI507004	USA	pure line
28	PI567024	pure line	USA	60	PI340898A	USA	pure line
29	PI605865B	pure line	USA	61	PI416873B	USA	pure line
30	PI416826A	pure line	USA	62	PI506677	USA	pure line
31	PI506939	pure line	USA	63	Clark	Ethiopia	Variety
32	PI587880A	pure line	USA	64	Nyal	Ethiopia	Variety

### 2.3. Data Collection

Data was recorded on Agronomic traits such as days to flowering, days to maturity, plant height (cm), number of pod per plant, number of seed per plant, number of branch per plant, crop lodging, shattering hundred seed weight (gm) and grain yield (t ha<sup>-1</sup>). Disease for soybean rust and frog eye leaf spot data was also recorded. The scoring system was 1-9 scale (1=immune, 9=susceptible, then 1-3=resistant, 4-6=moderately resistant and 7-9 = susceptible).

### 2.4. Data Analysis

#### 2.4.1. Analysis of Variance (ANOVA)

Data Analysis for simple lattice design was performed using the R-program software.

#### 2.4.2. Multivariate Analysis

Multivariate analysis techniques viz. cluster analysis and principal component analysis (PCA) was employed using SAS statistical package software. The numbers of clusters were determined by looking into three statistical approaches, namely, Pseudo-F, Pseudo-t<sup>2</sup> and cubic clustering criteria which is suggested by Copper and Miligan [15]. Accordingly, the number was decided where local peaks of Pseudo-F statistics and cubic clustering criteria combined with small values of Pseudo-t<sup>2</sup> statistics followed by a larger Pseudo-t<sup>2</sup> statistics for the next cluster fusion.

Divergence analysis (D<sup>2</sup>) was used to estimate the genetic distance/divergence of the soybean genotypes or to classify the divergent collections into different groups and it also measures the forces of differentiation at inter-cluster levels and determines the relative contribution of each component trait to the total divergent [16]. Genetic divergence between

clusters was determined using the generalized Mahalanobis's D<sup>2</sup> statistics [17] formula:  $D_{ij}^2 = (x_i - x_j)' s^{-1} (x_i - x_j)$ , where,  $D_{ij}^2$  = the distance between class i and j;  $X_i - X_j$  = the difference in the mean vectors of the two populations (class i and j);  $s^{-1}$  = the inverse of pooled variance covariance matrix. The D<sup>2</sup> values obtained for pairs of clusters were considered as the calculated values of Chi-square ( $\chi^2$ ) and tested for significance both at 1% and 5% probability levels against the tabulated value of ( $\chi^2$ ) for 'P' degree of freedom, where P is the number of traits considered [18].

The principal components analysis (PCA) was employed in order to minimize the traits into a new set of linearly combined measurements and to identify the traits contributing large part of the total variation among the collections. The analysis was performed using SAS software. In this analysis, only principal components with Eigen values greater than one were considered as important for the total variations.

## 3. Results and Discussion

### 3.1. Analysis of Variance

The combined analysis of variance across locations and year is presented in Table 2. The analysis of variance showed that mean squares due genotype was significant (P ≤ 0.01) for all the traits, indicating genotypes were responded differently for each trait. Means squares due to genotype x location x year interaction found significant effect (P ≤ 0.01) for hundred seed weight (HSW), soybean rust disease (SR) and grain yield (YLD). Generally, the observed variability of agronomic traits is dependent on genetic factors, environmental variables and the interaction factor.

**Table 2.** Mean squares of the combined analysis of variance for yield and related traits of 64 soybean genotypes at two location, evaluated during 2017-2018/

source of variation	Df	DTF	DTM	PH	NPP	NSP	HSW	Rust	YLD (t/ha)
Geno.	63	692.06**	927.72**	4344.80**	1452.52**	14867.35*	14867.35**	164.85**	14867.35**
Location	1	2516.06**	6202.20**	239.67ns	427.78ns	8294.72ns	8294.72**	22911.38**	8294.7**
Year	1	5466.66**	1582.03**	843.83*	6496.86**	370559.55**	370559.55**	117.24ns	370559.55**
Location: Rep.: Block	28	26.85ns	23.58ns	302.06*	507.80ns	5880.32ns	5880.32ns	100.18**	5880.32ns
Geno.: Location	63	17.48ns	60.42ns	72.05ns	432.77ns	1261.313ns	1261.31**	164.85**	1261.31**
Geno.: Year	63	58.55**	121.46**	383.96**	632.40ns	11487.84*	11487.84**	97.30**	11487.84**
Location: Rep.	2	14.063ns	107.28ns	1066.13ns	1485.79ns	13037.81ns	13037.81ns	19.16ns	13037.81ns
Location: Year	1	2501.60**	6159.85**	262.81**	2843.55*	20289.00*	20289.00**	116.09ns	20289.00**
Geno.: Location: Year	63	17.74ns	59.40ns	67.29ns	235.53ns	721.46Ns	721.46**	97.27**	721.46**
Residuals	226	21.87	49.99	169.77	525.72	4220.95	4220.95	53.75	4220.95

Where, \* = significant at (P ≤ 0.05) and \*\* = significant at (P ≤ 0.01), loc = location, Geno = genotype, Df = degree of freedom, DTF = days to 50% flowering, DTM = days to 95% pod maturity, PH = plant height, NPP = number of pod per plant, NSP = number of seed per plant, SR = soybean rust, HSW = hundred seed weight, YLD = grain yield per ha<sup>-1</sup>

### 3.2. Combined Mean Performance of the Genotypes

Based on the combined mean performance of the genotypes over environment (Table 3), latest physiological maturing genotype was genotype PI567104B (145.37 days) followed by PI567102B (144.75 days) and PI578457A (144.44 days), while the earliest genotype was PI615437

(104.58 days). Highest plant height was recorded from PI567104B (149.01 cm) followed by PI340898A (134.74 cm), while the lowest was recorded from genotype PI507004 (44.24 cm). The performance of the tested genotypes also showed resistance to moderately resistance to soybean rust. Maximum disease score was recorded from genotype PI567090 (25.52%), while the minimum was from

PI594538A (3.78%). Maximum hundred seed weight (24.01gm) was found from genotype PI506677, while the minimum seed weight (8.32gm) was recorded from genotype PI567068A. Coker240 scored maximum grain yield (3.09 t/ha) followed by genotype PI567104B (3.00 t/ha), PI567054C (2.85 t/ha) and G 7955-C3RPP (2.82 t/ha), while the minimum yield was scored from PI416826A (1.33t/ha).

**Table 3.** Range and Mean values of yield and other morphological traits of 64 soybean genotypes evaluated across two sites during 2017-2018.

No.	Designation	Trait							
		DTF	DTM	PH	NPP	NSP	SR (%)	HSW	yld
1	PI567099A	72.98	124.45	119.04	54.95	121.80	8.98	8.60	1.79
2	G 7955-C3RPP (C1)	63.54	131.56	106.23	44.70	95.15	12.67	18.83	2.82
3	MKSOY-2N white (C2)	68.76	138.47	94.80	71.06	121.28	9.33	15.39	1.76
4	PI605823	72.40	135.63	81.40	45.40	107.03	4.36	14.89	2.14
5	PI567020A	72.58	136.74	120.66	52.52	127.07	18.91	11.49	2.37
6	PI567102B	79.01	144.75	144.37	96.53	154.43	10.08	11.35	2.23
7	PI471904	68.01	130.76	109.40	53.57	116.38	11.19	13.19	2.83
8	PI567039	72.82	128.45	110.66	34.62	81.73	8.58	12.77	2.53
9	PI567058D	69.58	133.71	114.31	43.51	126.41	8.52	13.32	2.28
10	PI605824A	65.61	120.69	105.23	37.88	84.92	9.62	12.95	1.83
11	PI578457A	67.94	144.44	116.19	36.95	75.05	13.64	14.04	1.96
12	PI567046C	68.87	124.74	97.11	45.86	179.65	8.24	11.88	2.49
13	PI615445	64.48	119.25	85.22	34.35	73.85	9.55	14.35	2.05
14	PI567180	58.90	116.19	81.54	33.27	60.56	9.73	16.03	1.73
15	PI606405	61.40	118.89	75.04	29.89	56.12	8.01	15.42	1.75
16	PI594760B	63.23	122.00	98.49	46.43	112.50	8.12	15.29	2.27
17	PI605891A	61.82	119.77	84.54	29.67	64.39	12.11	14.91	2.38
18	PI567104B	79.18	145.37	149.01	74.61	143.37	11.48	12.30	3.00
19	PI567054C	68.82	127.19	98.76	52.83	118.26	8.95	12.65	2.85
20	PI605891B	66.43	134.13	98.05	44.49	95.72	12.54	13.58	2.54
21	PI567069A	79.08	140.30	104.69	58.82	130.23	16.76	8.42	2.36
22	Coker 240	59.51	124.58	86.00	33.99	65.47	21.17	18.50	3.09
23	PI606397B	65.41	126.71	76.72	38.17	72.27	21.21	14.74	2.45
24	PI567056A	87.39	142.03	120.84	62.20	123.61	21.91	9.67	2.00
25	PI628932	54.95	124.19	63.93	32.38	61.47	13.05	17.40	1.77
26	PI587905	59.06	112.85	56.90	63.07	246.25	6.46	13.86	2.02
27	PI567061	71.14	134.75	113.45	52.51	131.25	8.14	15.60	2.49
28	PI567024	65.23	121.15	99.92	38.73	179.61	18.58	16.07	2.45
29	PI605865B	63.12	118.29	89.47	40.41	90.36	14.10	15.76	2.63
30	PI416826A	47.19	108.68	51.68	18.67	44.07	9.87	14.11	1.33
31	PI506939	67.44	131.60	106.25	45.29	112.51	5.81	12.91	2.71
33	PI587880A	54.95	124.90	72.98	44.67	98.41	5.99	20.60	1.76
34	PI567025A	69.17	130.09	109.63	42.11	100.26	10.35	12.51	2.15
35	PI605838	72.50	129.30	110.85	48.84	109.62	9.79	12.86	2.41
36	PI567090	82.55	138.17	133.98	64.09	136.72	25.52	10.70	2.18
37	PI605773	59.29	125.79	88.44	42.79	103.29	10.63	16.31	2.08
38	PI416810	56.50	115.27	117.88	34.69	66.44	14.70	18.88	2.71
39	PI605854B	63.86	123.70	83.20	41.76	80.60	11.49	15.68	2.55
40	PI594767A	63.34	111.25	61.18	37.76	80.55	6.99	12.07	2.13
41	PI566989A	75.23	134.51	92.64	53.99	115.87	18.26	12.41	2.05
42	PI200466	49.49	113.89	71.89	43.45	100.89	12.56	19.25	2.35
43	PI635999	62.57	125.38	88.58	31.72	68.05	8.76	19.25	1.98
44	PI423960A	56.86	120.99	84.79	39.62	86.67	10.08	14.97	2.39
45	PI417208	55.61	108.42	66.35	32.06	134.55	14.06	21.90	2.56
46	PI567059	59.89	114.25	89.23	39.80	93.01	8.74	15.27	2.09
47	PI567068A	75.68	125.42	105.89	67.51	298.52	8.86	8.32	1.84
48	PI567034	54.54	108.52	77.76	33.65	73.09	11.17	11.58	2.07
49	PI567189A	61.37	120.46	83.81	40.94	95.20	9.99	14.12	2.16
50	PI594538A	55.45	108.45	52.25	30.62	72.54	3.78	18.86	2.51
51	Afgat (C5)	64.81	131.70	99.82	52.08	121.19	14.77	13.88	2.42
52	PI230970	59.44	131.85	101.06	60.85	152.68	5.89	12.35	2.62
53	PI615437	63.86	104.58	72.29	41.10	92.05	7.23	15.39	2.35
54	PI416886	50.43	106.72	44.87	28.43	71.17	10.01	18.39	1.78
55	PI417085	61.63	123.68	91.28	32.67	75.18	10.34	22.01	2.97
56	PI203398	58.14	122.29	87.52	40.38	82.52	8.44	14.63	2.47
57	PI423972	70.32	126.64	94.00	44.93	107.26	15.35	16.26	2.08
58	PI423960B	50.89	116.96	73.34	34.50	78.30	12.89	19.73	2.29
59	PI507004	49.07	109.11	44.24	20.48	31.49	6.92	19.40	2.04
60	PI340898A	88.65	145.51	134.74	56.51	116.67	15.08	11.26	1.99

No.	Designation	Trait							
		DTF	DTM	PH	NPP	NSP	SR (%)	HSW	yld
61	PI416873B	48.49	106.00	55.95	27.82	119.57	8.37	20.48	2.02
62	PI506677	53.31	108.65	57.39	29.60	74.16	9.34	24.01	2.19
63	Clarck63k	63.50	124.75	77.39	42.99	103.48	24.14	15.36	2.41
64	Nyal	58.15	115.11	65.88	34.21	70.02	21.82	15.43	2.68
	maximum	88.65	145.51	149.01	96.53	298.52	25.52	24.01	3.09
	mean	64.57	124.48	91.50	43.84	104.40	11.69	14.93	2.27
	minimum	47.19	104.58	44.24	18.67	31.49	3.78	8.32	1.33
	CV	7.24	5.68	14.24	52.30	62.23	62.72	11.15	17.51
	LSD	4.61	6.97	12.84	22.59	64.01	7.22	1.64	0.39
	P-value	**	**	**	*	**	**	**	**

### 3.3. Cluster Analysis

Cluster analysis categorized 64 soybean genotypes into five clusters (table 4). The grouping pattern showed; cluster-I contained highest number of genotypes (53%) followed by cluster-II (41%), cluster-III (3%), cluster-IV (2%) and

cluster-V (2%). In cluster analysis, if the categorization is successful, individuals within (homogenous) shall be closer and different clusters (heterogeneous) shall be farther apart [19]. The distribution pattern of genotypes in to different cluster might be difference in genetic background through their pedigree.

Table 4. The distribution of 64 soybean genotypes in to five clusters tested at Jimma and Mettu (2017-2018).

Cluster No.	No. geno.	(%)	genotypes
I	34	53	PI567039, PI605824A, PI578457A, PI615445, PI567180, PI606405, PI605891A, Cocker 240, PI606397B, PI628932, PI605865B, PI416826A, PI587880A, PI416810, PI605854B, PI594767A, PI200466, PI635999, PI423960A, PI417208, PI567059, PI567034, PI567189A, PI594538A, PI615437, PI416886, PI417085, PI203398, PI423960B, PI507004, PI416873B, PI506677, Clark, Nyal
II	26	41	PI567099A, G 7955-C3RPP (C1), MKSOY-2Nwhite (C2), PI605823, PI567020A, PI471904, PI567058D, PI567046C, PI594760B, PI567054C, PI605891B, PI567069A, PI567056A, PI567061, PI567024, PI506939, PI567025A, PI605838, PI567090, PI605773, PI566989A, PI567053, Afgat (C5), PI230970, PI423972, PI340898A
III	2	3	PI567102B, PI567104B
IV	1	2	PI587905
V	1	2	PI567068A

#### 3.3.1. Cluster Characterization Using Quantitative Traits

Mean performance of clusters (table 5) for the 8 traits reflected that the genotypes in cluster-III exhibited the highest in days to flowering, late maturing, longest plant height, maximum number of pod and high yielder. Genotypes in cluster I were characterized by early flowering, large seed size and lowest number of pod and medium value for the

remaining traits. Genotypes in Cluster-II were characterized by relatively highest severity in soybean rust and medium mean value for other traits. Shortest plant height, early maturity date and rust tolerant genotypes were found in cluster IV. Cluster-IV possessed genotypes maximum number of seed, low seed size and low yielder genotypes.

Table 5. Cluster mean values of 64 soybean genotypes tested at Jimma and Mettu (2017-2018).

Cluster	Traits							
	DTF	DTM	PH	NPP	NSP	SR	HSW	YLD
I	58.68*	117.75	77.83	34.80*	78.77*	11.43	16.71**	2.23
II	70.95	132.15	105.8	50.88	121.94	12.33**	13.12	2.33
III	79.10**	145.06**	146.7**	85.57**	148.90	10.78	11.83	2.61**
IV	59.06	112.85*	56.9*	63.07	246.25	6.46*	13.86	2.02
V	75.68	125.42	105.89	67.51	298.52**	8.86	8.32*	1.84*

\*\*,\* represents maximum and minimum values respectively, DTF = days to 50% flowering, DTM = days to 95% pod maturity, PH = plant height, NPP = number of pod per plant, NSP= number of seed per plant, SR=soybean rust, HSW=hundred seed weight, YLD= yield ton per ha<sup>-1</sup>

#### 3.3.2. Genetic Divergence ( $D^2$ )

Multivariate analysis by means of Mahalanobis'  $D^2$  statistics is a useful tool in quantifying the degree of genotypic divergence among biological populations and to assess the relative contribution of different components to the total divergence at inter-cluster levels [8]. The values of pair wise average inter-cluster divergence ( $D^2$ ) among 64 soybean

genotypes in five clusters based on 8 quantitative traits are presented in table 6.

The chi-square test revealed the existence of highly significant differences among the paired inter cluster distance except cluster I with II and IV with V. The maximum inter-cluster distance was found between cluster-I and V (144.65) followed by cluster-III and V (113.6), cluster-II and V

(91.59). The highest value of inter-cluster distance indicated that the genotypes belonging to these cluster were far diverged. The lowest inter-cluster distance was recorded between clusters-I and II (11.05) followed by cluster-IV and V (21.63), which means a close relationship between the genotypes.

Cluster-III was found divergent from other clusters chiefly due to days to maturity, days to maturity, plant height number of pods per plant and grain yield, indicating maximum contribution of these traits towards the divergence. Similarly, cluster-I was found diverged from cluster II and IV due to maximum hundred seed weight. On the other hand cluster II were diverged mainly due to soybean rust, while cluster IV was diverged mainly due to early maturing and rust tolerant genotypes.

Traits contributing maximum towards the divergence should be given greater emphasis for deciding the type of cluster for the purpose of further selection and choice of the parents for hybridization [20]. In this perspective, intra cluster mean performance for days to physiological maturity, plant height, number of pods and grain yield were maximum in cluster-III than other clusters, so these traits contribute a great role towards the divergence between cluster-III with other clusters. Generally, effective genetic recombination will be expected from cross that involve parents from the significant inter cluster distance. In the present investigation, superior recombinants can be exploited by crossing genotypes from between cluster-I and V followed by cluster-III and V, cluster-II and V. The current result is in support the previous findings [10, 11, 13].

**Table 6.** Pair wise average inters cluster divergence values ( $D^2$ ) among 64 soybean genotypes at Jimma and Metu (2017-2018).

Cluster	I	II	III	IV	V
I		11.05	61.89	86.91	144.65
II			30.47	55.21	91.59
III				79.31	113.16
IV					21.63
V					

\*, \*\*= significant, ( $p < 0.01$ )  $\chi^2=24.72$ , and ( $p < 0.05$ )  $\chi^2=19.67$ , respectively

### 3.4. Principal Component Analysis (PCA)

Principal component analysis was done using 8 quantitative traits (table 7). Accordingly, the first four principal components with Eigen values exceeding one were responsible for about 88.74% of the total variation among the genotypes. Maximum variation was accounted from the first principal component (52.49%) followed by the second (15.99%) principal components, which means, out of the entire variations, the first and the second principal components accounted for more than two third of the total variations (68.47%).

The first principal component that accounted maximum variation (52.49%) was due to the principal contribution of positive discriminatory traits like days to flowering, plant height, days to maturity, number of pod and number of seed. Considerable variation observed in the second principal

component (15.99%) was attributed to grain yield, soybean rust, hundred seed weight and number of seed. Traits which had substantial contribution to the third principal component (11.6%) were grain yield, number of seed, hundred seed weight, number of pod and soybean rust. On the other hand, number of seed and soybean rust predominantly influenced the variation in the fourth principal component. Consistent with this finding other researchers also found comparable result from different soybean genotypes [10-13].

Any traits with the largest absolute values closer to unit within the first principal component influence the clustering more than those with lower absolute values closer to zero [21]. Therefore, in the current investigation discrimination of the accessions in to different cluster was mainly due to days to flowering, plant height, days to maturity, number of pod and number of seed.

**Table 7.** Eigenvectors and Eigen values of the first four principal components for 8 traits of 64 soybean genotypes tested at Jimma and Metu (2017-2018).

Traits	Principal component			
	Prin1	Prin2	Prin3	Prin4
DTF	0.45	0.03	-0.16	-0.12
DTM	0.42	0.15	-0.12	-0.30
PH	0.43	0.15	0.01	-0.30
NPP	0.42	-0.18	0.22	0.13
NSP	0.29	-0.37	0.45	0.59
SR	0.14	0.58	-0.44	0.66
HSW	-0.37	0.25	0.27	0.00
YLD	0.09	0.62	0.67	-0.08
Eigen values	4.20	1.28	0.93	0.69
Total variance (%)	52.49	15.99	11.6	8.66
Cumulative variance (%)	52.49	68.47	80.07	88.74

## 4. Conclusions

The pre-request to perform selection in any breeding program is the presence of genetic variability. Form the current study the tested soybean genotypes were found genetically diverse in terms of different morphological traits. Wide range of variations and best performed genotypes were found for most of yield and yield related traits. A total 64 soybean genotypes were grouped into five clusters. The maximum inter cluster distance was found between clusters-I and V followed by cluster-III and V, suggesting superior hybrids or recombinants can be realized by crossing genotypes in these clusters. Principal component analysis (PCA) revealed that, the 1<sup>st</sup> four PCA with Eigen values exceeding one were responsible for about 88.74% of the total variation. Out of the entire variations, 1<sup>st</sup> PCA and the 2<sup>nd</sup> PCA accounted for more than two third of the total variations (68.47%). In conclusion, this research finding will enhance utilization of variation present with in soybean genotypes for variety development program.

## Acknowledgements

The author acknowledges Ethiopian Institute of

Agricultural Research for availing the resources to execute this research work. Appreciation also goes to Metu Agricultural Research sub-Center for cooperation in the field during the entire period of data collection.

---

## References

- [1] Singh, P., R., Kumar, S. N., Sabapathy, and Bawa, A. S., 2008. Functional and edible uses of soyprotein products. *Reviews in Food Science and Food Safety*, 7: 14–28.
- [2] Fekadu Gurmu, Hussein Mohammed and GetnetAlemaw, 2009. Genotype X environment interactions and stability of soybean for grain yield and nutrition quality. *African Crop Science Journal* 17: 87-99.
- [3] Clarke E. J. and Wiseman, J., 2000. Developments in plant breeding for improved nutritional quality of soybeans. Protein and amino acid content. *Journal of Agricultural Science*, 134: 111- 124.
- [4] Stevanović, P., V. Popović, V. Filipović, D. Terzić, V. Rajčić, D. Simić, M, Tatić, M. Tabaković, 2017. Proceedings of the Institute of PKB Agroecomics. Counseling of agronomists, veterinarians, technologists and agroeconomists 23: 119-127.
- [5] Graham, P. H. and Vance, C. P., 2003. Legumes: importance and constraints to greater use. *Plant Physiology*, 131: 872–877.
- [6] CSA (Central Statistical Authority of Ethiopia), 2021. Area and production of major crops, Central Statistics Agency (CSA), Addis Ababa, Ethiopia.
- [7] Murty, B. R. and Arunachalam, V. 1966. The nature of genetic divergence in relation to breeding system in crop plants. *Indian J. Genet*, 26A: 188-198.
- [8] Das, P. K. and Gupta, T. D., 1984. Multivariate analysis in Blackgram. *Indian Journal of Genetics*. 44: 243-247.
- [9] Joshi, A. B. and Dhawan, N. L. 1966. Genetic improvement of yield with special reference to self-fertilizing crops. *Ind. J. Genet. Pl. Br.*, 26A: 101-113.
- [10] Tadesse G. and Sentayehu A., 2015. Genetic Divergence Analysis on Some Soybean (*Glycine max L. Merrill*) Genotypes Grown in Pawe, Ethiopia. *Am-Euras. J. Agric. And Environ. Sci.*, 15: 1927-1933.
- [11] Abush Tesfaye, M. Githiri, J. Derera and TolessaDebele., 2017. Genetic Variability in Soybean (*Glycine max L.*) for Low Soil Phosphorus Tolerance. *Ethiopian Journal of Agricultural Sciences*, 27 (2), pp. 1-15.
- [12] Yechalew S., Andargachew G., Abush T., and Hailemariam, M., 2019. Contribution of Morphological Traits to the Total Variability in Soybean (*Glycine max (L.) Merr.*) Genotypes in Western Parts of Ethiopia. *Acad. Res. J. Agri. Sci. Res.* 7 (7): 408-413.
- [13] Masreshaw Yirga, Yechalew Sileshi, Abush Tesfaye, and Mesfin Hailemariam, 2022. Genetic Variability and Association of Traits in Soybean (*Glycine max (L.)*) Genotypes in Ethiopia. *Ethiop. J. Crop Sci.* Vol 9 No. 2.
- [14] Paulos Dubale, 2001. Soil and water resources and degradation factors affecting their productivity in the Ethiopian highland agro - ecosystems. Michigan State University Press, 8 (1): 1-18p.
- [15] Copper, M. C. and Milligan, G. W. 1988. The effect of error on determining the clusters. Proceedings on the International workshop on Data Analysis, Decision support and Expert Knowledge Representation in Marketing and Related Areas of Research, June 21–23, 1987, University of Karlsruhe, WestGermany. Pp. 319-328.
- [16] Sharma, J. R. 1998. *Statistical and biometrical techniques in plant breeding*. New Age International (P) limited publishers. New Delhi. 432p.
- [17] Mahalanobis, P. C. 1936. On the generalized distance in statistics. *National institute of science.*, 2: 49-55.
- [18] Singh, RK. and BD. Chaudhary, 1987. Biometrical methods in quantitative genetic analysis. *Kalyani publishers*, New Delhi-Ludhiana, India.
- [19] Hair, M. L., 1995. *Colloids and Surfaces A: Physicochemical and ...*, 1995 – Elsevier. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*. Elsevier. Volume 105, Issue 1, 1 December 1995, Pages 95-103.
- [20] Jagadev, P. N., Samal, K. M. and Lenka, D., 1991. Genetic divergence in rape mustard. *The Indian Journal of Genetics and Plant Breeding*, 51: 465-467.
- [21] Chahal, G. S. and Gosal, S. S., 2002. *Principles and Procedures of Plant Breeding: Biotechnology and Conventional Approaches*. Alpha Science International, United Kingdom, ISBN: 9781842650363: 604.