Electronic Journal of Plant Breeding, 7(3): 742-749 (September 2016)
ISSN 0975-928X
DOI: 10.5958/0975-928X.2016.00097.1

Research Note

Morphological characterization of newly introduced lettuce (*Lactuca sativa* L.) germplasm through principal component and regression analyses

Ramesh Kumar[1], Sangita Kaushal[1], Sandeep Kumar[1*], Dharminder Kumar[1], Y.R. Shukla[1] and B.S. Dogra[2]
[1]Department of Vegetable Science,
[2]College of Horticulture and Forestry, Neri, Hamirpur, HP
Dr. YS Parmar University of Horticulture and Forestry, Nauni, Solan – 173 230, HP, India
Email: sandeepkdhatwalia@gmail.com

(Received:28 Jul 2015; Accepted:21 Jun 2016)

Abstract
Twenty seven newly introduced genotypes of lettuce collected from different indigenous and exotic sources were characterized with respect to economically important traits by using principle component and regression analyses. Significant differences were observed among the genotypes for all the characters under study. CGN-10944 performed best among all the genotypes for most of the traits under study. Principal component analysis indicated that the first five factors had variances of 3.88, 2.78, 1.47, 1.42 and 1.24 with 25.87%, 18.57%, 9.81%, 9.49% and 8.26% of total variation, respectively and aggregating to 72.01% of total variation. The highest positive values of different traits in different components indicated that main emphasis should be given on the gross head weight (X_2), polar (X_6) & equatorial diameter (X_7) and seed vigour index-I (X_{12}) & II (X_{13}) to have higher yields in lettuce. To quantify the importance of each variable in predicting yield, multiple linear regression model was developed. The coefficient of determination (R^2) in linear regression model revealed that 94% of total variation in yield per plot (X_{15}) was influenced by days to marketable maturity (X_1), gross head weight (X_2), heading percentage (X_4), seed germination (X_{11}), seed vigour index-II (X_{13}) and severity of bacterial soft rot (X_{14}). Therefore, main emphasis should be given on days to marketable maturity, gross head weight, polar & equatorial diameter, heading percentage, seed germination, seed vigour index-I & II and severity of bacterial soft rot during selection and hybridization programmes for yield improvement in lettuce

Keywords
Principal component, regression analysis, morphological characterization lettuce, *Lactuca sativa* L.

Lettuce (*Lactuca sativa* L.) is an important salad crop and has attained elite place among the winter vegetables due to good economic returns, higher nutritional value and great production potential. It is grown as salad crop in some parts of the state- Himachal Pradesh. The crop is gaining importance because of its multifarious uses, liking by the tourists, local consumers and year round demand in cosmopolitan cities. It is an annual plant in temperate climates; leaf varies from yellow-green to dark red. It is an important component of the diet due to high vitamin A and minerals like Ca & Fe. The nutrient content is high in the darker green, outer leaves. The cos or leafy types have the highest levels followed by butter- and crisphead types. This increase in quality is due to the increased numbers of green leaves exposed to light (Ryder, 1979).

Variation in lettuce is mostly found in leaf length, shape, color, texture, size and heading type. All types need to be examined for identification of superior genotypes. Lettuce is consumed in fresh and processed forms. Lettuce cultivars grown in India are low in average yield due to undesirable cultivars/hybrids, biotic and abiotic stresses, genetic drift in cultivars, and development of new races of pathogens. To enhance productivity, genetic restructuring of lettuce germplasm is needed to develop/identify high yielding varieties/hybrids.

Utilization of genetic resources in developing sustainable solution to basic crops constraints have been suggested from time to time but these genetic resources could not be exploited fully due to their inherent problems of large size and lack of sufficient evaluation and classification (Dahberg, 1995). Germplasm maintenance, evaluation and characterization for economically important traits are pre-requisite for genetic improvement program of any crop. Genetic diversity in vegetable crops is important in selecting the best genotypes to bring improvement in yield. Qualitative and quantitative traits can be chosen in parents for hybridization to exploit heterosis or to select desirable segregants in subsequent generations. Knowledge of association of various characters provides the basis of selection for yield and its components for crop improvement. Since yield is a complex quantitative trait, simple correlation and regression of characters provides limited insight into the association of various traits to yield. Few reports are available on phenotypic variability, correlation and path analysis in lettuce

(Thakur *et al.,* 1997; Meglic and Vozlic 2000; Kumar *et al.,* 2010). Investigators, initially unaware of the relative importance of variables, try to include all possible variables which are likely to influence the outcome and the data matrices become unmanageable and complicated. Principle component analysis (PCA) helps in identifying the most relevant characters that can be used as descriptors by explaining as much of total variation in the original set of variables as possible with as few components as possible and reducing the dimension of the problem. Therefore, an attempt has been made in the present investigation to access and analyze the extent of genetic diversity through Principle component and regression analyses for yield improvement in lettuce.

The present investigations were carried out at Research Farm of the Department of Vegetable Science, Dr. YS Parmar University of Horticulture and Forestry, Nauni, Solan (HP) in *Rabi* 2008-2009. This location is at 30°50' N latitude, 77°11'30" E longitude and is 1260 m above mean sea level and represents the mid-hill zone of Himachal Pradesh. The climate of the Experimental Farm is generally characterized as sub-humid, sub-temperate with cool winters. The annual precipitation is about 1000-1300 mm, most of which is received during rainy season (June-September) in India.

The experimental material consisted of 27 diverse genotypes of lettuce, including check cultivars *i.e.,* Great Lakes and Alamo-1 (Table 1). The experiment was laid out in Randomized Complete Block Design (RCBD) with three replications of each genotype. All the recommended agronomic practices were followed to raise a healthy crop. Ten plants/heads in each replication were randomly selected to record observations on days to marketable maturity [X_1], gross head weight (g) [X_2], number of non-wrapper leaves [X_3], heading percentage (%) [X_4], head shape index [X_5], polar diameter (cm) [X_6], equatorial diameter (cm) [X_7], β-carotene content (μg/100g) [X_8], iron content (mg/100g) [X_9], 1000-seeds weight (g) [X_{10}], seed germination (%) [X_{11}], seed vigour index-I [X_{12}] and II [X_{13}], severity of bacterial soft rot (%) [X_{14}] and yield per plot (kg) [X_{15}]. Seed vigor indices I and II were determined by the formula of Abdul-Baki and Anderson (1973). The severity of bacterial soft rot was recorded periodically with the appearance of disease and severity determined by the formula of Mckinney (1923). Non-wrapper leaves were removed from heads that were weighed for yield and head shape index detemined (Odland and Noll, 1954). Contents of β-carotene and iron were determined according to methods of Ranganna (1986).

The data were subjected to analysis of variance as per the procedure described by Gomez and Gomez (1983). Contribution of different characters towards the divergence was estimated with the help of principle component analysis in accordance with Lawley and Maxwell (1963) and Ramchander *et al.* (1979). Multiple linear regression equation was used to predict average gross head weight and yield per plot.

Genetic variability is the basic fundamental need for any crop breeding programme. Analysis of variance (Table 2) indicated significant differences among the genotypes for all traits. These differences indicated the presence of variability and opportunity for improvement in yield and quality traits of lettuce.

Significant differences among the genotypes were observed for all the characters under study *viz*., days to marketable maturity, number of non-wrapper leaves, gross head weight, heading percentage, head shape index, equatorial and polar diameter of the head, β-carotene content, iron content, 1000-seeds weight, seed germination, seed vigour index-I and II, severity of bacterial soft rot and yield per plot (Table 3). Among horticultural traits, comparatively wide range was observed for gross head weight (196.00 - 450.00 g), number of non-wrapper leaves (6.46 - 14.33), heading percentage (60.54 - 89.15 %), polar diameter (8.10 - 11.80 cm), equatorial diameter (5.73 - 9.10 cm), β-carotene content (1.64 -7.06 μg/100g), seed vigour index-II (20.17 - 42.30) and severity of bacterial soft rot (6.68 - 35.23 %). Sufficient genetic variability for many traits had also been reported by Pearson (1931), Flory and Walker (1940), and Arumgam *et al.* (1978) for cabbage under their environmental conditions. CGN–10944 gave maximum yield per plot (8.00 kg) over both the check cultivars *viz.,* Great Lakes and Alamo-1. This genotype also performed well for other characters like days to marketable maturity, gross head weight, number of non-wrapper leaves and heading percentage but found moderately susceptible to bacterial soft rot. Besides CGN-10944, four genotypes namely CGN - 04508, CGN - 04987, CGN - 05167 and CGN - 11358 performed better for gross head weight and seven genotypes namely CGN - 04543, CGN - 04987, CGN – 05167, CGN – 05169, CGN – 09373, CGN – 14629 and CGN – 19009 performed better for heading percentage over both the checks. One genotype CGN-19088 performed better for β – carotene and iron content, CGN-05167 for 1000 - seeds weight, seed germination, seed vigour index I & II and genotypes CGN - 04778 for severity of bacterial soft rot performed better over the checks. Besides genotypes CGN – 04778 others *viz.* CGN - 04925, and CGN - 20721 are found resistant against bacterial soft rot (*Erwinia carotovora*) as compared to checks *viz.*, Great Lakes and Alamo -1 which were moderately resistant.

The characters contributing more to the divergence gave greater emphasis for deciding on the cluster for the purpose of further selection and the choice of parents for hybridization (Jagadev *et al.,* 1991). Principal component analysis indicated that the first five components account for the maximum estimated variation (72.01%). Factor analysis was applied to extract the basic factors underlying the observed traits of lettuce. The factors were extracted individually on the basis of eigen values (Table 3) and revealed the pattern and principal component analysis of the data. The first five components having eigen values greater than 1 were retained in the analysis because of the substantial amount of the variations. The factors corresponding to eigen values less than 1 were not considered (Fig 1). These factors were ignored due to " Guttmans lower bound principle" according to which eigen values less than unity ($\lambda<1$) should be ignored (Kaiser 1958). The orthogonal factors were extracted. The centroid method of analysis (Lawley and Maxwell, 1963) was used.

The first five factors had variances of 3.88, 2.78, 1.47, 1.42 and 1.24 with 25.87, 18.57, 9.81, 9.49 and 8.26 % of total variation, respectively and aggregating to 72.01% of total variation.

The first factor extracted had the combination of gross head weight (X_2), heading percentage (X_4), equatorial diameter (X_7), seed vigour index-I (X_{12}), severity of bacterial soft rot (X_{14}) and yield per plot (X_{15}). The first factor had high positive loadings for all the variables and differentiated between different accessions of lettuce germplasm that high gross head weight contained more equatorial diameter and yield per plot. The second factor was a combination of number of non-wrapper leaves (X_3), β-carotene content (X_8), 1000-seeds weight (X_{10}), seed germination (X_{11}) and seed vigour index-II (X_{13}) having high positive loadings. The third factor accounted for only one variable i.e., polar diameter (X_6). The fourth factor was for the combination of head shape index (X_5) and iron content (X_9). In the meanwhile, fifth factor had positive loadings only for days to marketable maturity (X_1). The highest positive values of different characters under study in different components indicated its importance in divergence among 27 genotypes of lettuce, whereas negative values showed the lowest contribution to the divergence (Table 4). Loading of different variables based on first two principle components indicated that iron content, 1000-seeds weight, seed germination per cent and seed vigour index-I and II were the main components of divergence between 27 genotypes of lettuce, whereas contribution of head shape index was found least in divergence (Fig 2). Hence, on the basis of this study main emphasis should be given on the gross head weight, polar & equatorial diameter and seed vigour index-I & II to increase the yield of lettuce crop. Similar findings had also been reported for cucumber, peppers and okra by Zhang and Cui (1993), Portis *et al.* (2006) and Koutsos *et al.* (2000), respectively.

To estimate the yield per plot, multiple linear regression model was extracted with yield per plot (x_{15}) as dependent variable and days to marketable maturity (X_1), gross head weight (X_2), number of non-wrapper leaves (X_3), heading percentage (X_4), head shape index (X_5), polar diameter (X_6), equatorial diameter (X_7), β-carotene content (X_8), iron content (X_9), 1000-seeds weight (X_{10}), seed germination (X_{11}), seed vigour index-I (X_{12}) and II (X_{13}) and severity of bacterial soft rot (X_{14}) as independent variables. After eliminating the independent variables with $p>0.10$, a simplified multiple linear regression model was obtained. The prediction of yield per plot with values in parentheses indicating standard errors of the regression coefficients had been given in the Table 5.

Multiple linear regression model indicated that yield per plot could be well predicted with the help of days to marketable maturity (X_1), gross head weight (X_2), heading percentage (X_4), seed germination (X_{11}), seed vigour index-II (X_{13}) and severity of bacterial soft rot (X_{14}). The characters namely gross head weight, heading percentage, seed germination and seed vigour index-II had positive effects on estimation of yield per plot, whereas days to marketable maturity and severity of bacterial soft rot had negative effects on yield per plot in lettuce (Fig. 3).

The coefficient of determination (R^2) in linear regression model was found high (0.942). It meant that 94 per cent of total variation in yield per plot was influenced by these characters. Hence, use of gross head weight, heading percentage, seed germination and seed vigour index-II was the best model for predicting yield per plot. Moreover, there is an opportunity for improvement through hybridization and selection due to genetic diversity in the newly introduced germplasm. Therefore, linear regression model can be used best for genetic improvement of lettuce by the development of new varieties/hybrids.

References

Abdul-Baki, A.A. and Anderson, J.D. 1973. Vigor germinated in soyabean seed by multiple criteria. *Crop Sci.,* **13**: 630–633.

Arumugam, R.S., Sambandamurthi and Natarajan, S. 1978. Studies on variability in cabbage (*Brassica oleraceae* var. *capitata* L.). *Prog. Horti.,* **10**: 11–15.

Dahberg, J.A. 1995. Operators hand book for sorghum germplasm maintenance. Agricultural Research Service, Beltsville, Md.

Flory, W.S. and Walker, J.C. 1940. Effects of different environment on head shape in Marion Market cabbage. *Proc. Amer. Soci. Horti. Sci.,* **37**: 778–782.

Gomez, K.A. and Gomez, A.A. 1983. Statistical procedures for Agricultural Research. John Wiley & Sons, New York. pp. 357-427.

Jagdev, P.N., Samal, K.M. and Lenka, L. 1991. Genetic divergence in rape mustard. *Ind. J. Genet.,* **51**: 465-466.

Kaiser, H.F. 1958. The varimax criteria for analytic rotation in factor analysis. *Psychomethka,* **23**: 187-200.

Koutsos, T.V., Koutsika-Sotiriou, M., Gouli-Vavdinoudi, E. and Tertivanidis, K. 2000. Study of genetic relationship of Greek okra cultivars (*Abelmoschus esculentus* (L.) Moench.) by using agronomic traits, heterosis and combining ability. *J. Veg. Crop Prod.*, **6**(1): 25-35.

Lawley, D.N. and Maxwell, A.E. 1963. Factor analysis as a statistical method. Butterworth, London. 117p.

Mckinney, H.H. 1923. Influence of soil temperature and moisture on infection of wheat seedlings by *Helminthosporium sativum*. *J. Agri Res.,* **27**: 195–197.

Meglic, V.S. and Vozlic, J. 2000. Genetic variability of lettuce. *Sodobno-Kanetijstvo (Slovenia)* **33**(5): 215–217.

Odland, M.L. and Noll, C.J. 1954. Vegetable variety trials. Pennsylvania agricultural experiment station progress report, Bulletien, 114.

Pearson, O.H. 1931. Methods of determining the solidarity of cabbage heads. *Hillgardia,* **5**: 383–393.

Portis, E., Nervo, G., Cavallanti, F., Barchi, L. and Lanteri, S. 2006. Multivariate analysis of genetic relationships between Italian pepper landraces. *Crop Sci.,* **46**: 2517-2525.

Ramchander, P.R., Biswas, S.R., Singh, D.P. and Pathan, C.S. 1979. Factor analysis in onion (*Allium cepa* L.). *Curr. Sci.,* **48**: 137- 141.

Kumar, Ramesh, Kaushal, Sangita and Shukla, Y.R. 2010. Variability, correlation, and path analysis studies in lettuce. *Internat. J. Veg. Sci.,* **16**(4): 299-315.

Ranganna, S. 1986. Handbook of analysis and quality control for fruit and vegetable products. 2nd ed. Tata McGraw Hill Publishing Co., New Delhi. 1112p.

Ryder, E.J. 1979. Leafy salad vegetables. AVI Publishing Co., Inc.,Westport, Connecticut. 266p.

Thakur, M.C., Joshi, A.K., Singh, N.P. and Shukla, Y.R. 1997. Variability studies in heading lettuce (*Lactuca sativa* L.). *Horti. J.,* **10**(2): 43–49.

Zhang, M. and Cui, H. W. 1993. Application of factor analysis to cucumber breeding. *Cucur. Genet. Cooper.,* **16**: 27-29.

Table 1. List of lettuce (*Lactuca sativa* L.) genotypes studied along with their sources

Genotypes	Sources
CGN-04508, CGN-04511, CGN-04543, CGN-04778, CGN-04925, CGN-04933, CGN-04934, CGN-04987, CGN-04988, CGN-04989, CGN-04990, CGN-05166, CGN-05167, CGN-05169, CGN-05198, CGN-09373, CGN-10944, CGN-11358, CGN-14629, CGN-14651, CGN-14688, CGN-17390, CGN-19009, CGN-19088 and CGN-20721	Crop Genetic Resources, the Netherlands
Great Lakes and Alamo-1 (Checks)	UHF, Nauni-Solan, India

Table 2. Analysis of variance for different characters in lettuce

Traits	Source of Variation		
	Genotypes (26)[a]	**Replication (2)**	**Error (52)**
Days to marketable maturity	152.86*	8.160	9.814
Number of non -wrapper leaves	13.02*	0.984	0.529
Gross head weight (g)	13.716*	3.603	1.786
Heading percentage (%)	135.872*	1.760	3.699
Head shape index	0.054*	0.00006	0.0008
Equatorial diameter (cm)	2.445*	0.118	0.069
Polar diameter (cm)	3.254*	0.323	0.248
β-carotene content (μg/100g)	5.357*	0.101	0.058
Iron content (mg/100g)	0.227*	0.025	0.020
1000- seeds weight (g)	0.0016*	0.0002	0.0005
Seed germination (%)	47.302*	0.405	5.215
Seed vigour index-I	0.000284*	0.000005	0.000006
Seed vigour index-II	72.955*	1.337	0.612
Severity of bacterial soft rot (%)	67.465*	2.942	2.840
Yield per plot (kg)	4.543*	0.547	0.396

* Significant at 5% level of significance
[a] values in parentheses are degree of freedom.

Table 3. Range, mean, and coefficient of variation of different characters in lettuce

Traits	Mean	Standard Error	Range		Coefficient of Variation
			Minimum	Maximum	
Days to marketable maturity	73.69	2.55	64.33	88.67	22.21
Gross head weight	339.53	34.51	196.00	450.00	75.44
Non wrapper leaves	9.46	0.59	6.46	14.60	45.06
Heading percentage	73.73	1.57	60.55	89.15	24.84
Head shape index	1.32	0.02	0.97	1.54	20.54
Polar diameter	9.84	0.40	8.10	11.80	21.93
Equatorial diameter	7.41	0.21	5.73	9.10	24.72
β- Carotene	3.59	0.19	1.63	7.06	74.76
Iron content	1.36	0.11	0.92	1.86	41.19
1000 seed weight	0.95	0.01	0.928	1.022	5.10
Seed germination	68.47	1.86	60.50	84.17	20.38
Seed vigour index-I	0.05	0.002	0.04	0.08	34.432
Seed vigour index-II	28.52	0.63	20.17	42.30	35.27
Severity of bacterial soft rot	18.00	1.37	6.68	35.23	72.28
Yield per plot	5.87	0.51	3.31	8.00	43.24

Table 4. Principle component analysis for different characters of lettuce germplasm

Characters	Principle Component[*]				
	PC_1[#]	**PC_2**	**PC_3**	**PC_4**	**PC_5**
Days to marketable maturity	-0.314	0.433	0.336	-0.250	**0.508**[**]
Gross head weight (g)	**0.824**	-0.182	-0.050	0.068	-0.087
Non wrapper leaves	-0.365	**0.541**	-0.297	0.456	0.252
Heading percentage(%)	**0.425**	-0.139	-0.305	-0.375	0.121
Head shape index	-0.578	-0.152	0.104	**0.578**	-0.216
Polar diameter (cm)	0.024	-0.311	**0.628**	0.506	0.322
Equatorial diameter (cm)	**0.712**	-0.062	0.330	0.042	0.465
β – carotene (µg/100g)	-0.063	**0.520**	-0.211	-0.275	0.387
Iron content (mg/100g)	0.154	0.326	-0.577	**0.461**	0.132
1000-seed weight (g)	0.311	**0.562**	0.267	-0.038	-0.498
Seed germination (%)	0.266	**0.618**	0.400	-0.073	-0.267
Seed vigour index-I	**0.739**	0.340	0.034	0.229	0.151
Seed vigour index-II	0.469	**0.755**	-0.059	0.198	-0.115
Severity of bacterial soft rot (%)	**0.411**	-0.552	-0.211	0.120	0.065
Yield per plot (Kg)	**0.905**	-0.254	-0.051	0.127	-0.051
Eigen values	**3.88**	**2.78**	**1.47**	**1.42**	**1.24**
Percentage of variance	**25.87**	**18.57**	**9.81**	**9.49**	**8.26**
Cumulative % of variance	**25.87**	**44.44**	**54.25**	**63.74**	**72.01**

[#]PC: Principal component

[*]Extracted through principle component analysis

[**]Bold value indicates the highest Eigen Vector for the corresponding trait amongst the five principal components

Electronic Journal of Plant Breeding, 7(3): 742-749 (September 2016)
ISSN 0975-928X

Table 5. Multiple linear regression coefficient model to predict yield per plot

Dependent variable	Independent variables influencing yield per plot							*R^2
	Intercept	(X_1)	(X_2)	(X_4)	(X_{11})	(X_{13})	(X_{14})	
Yield per plot (x_{15})	–11.70	-0.47 (0.015)	0.011 (0.002)	0.030 (0.015)	0.050 (0.024)	0.085 (0.044)	-0.038 (0.019)	0.942

*Coefficient of multiple determination

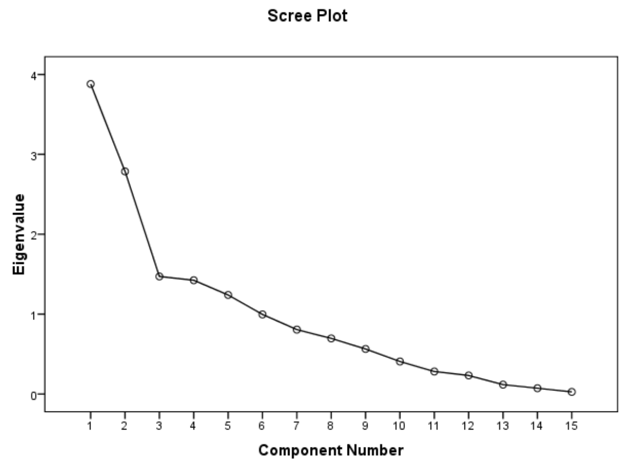

Fig. 1. Scree plot based on eigen values of different characters under study

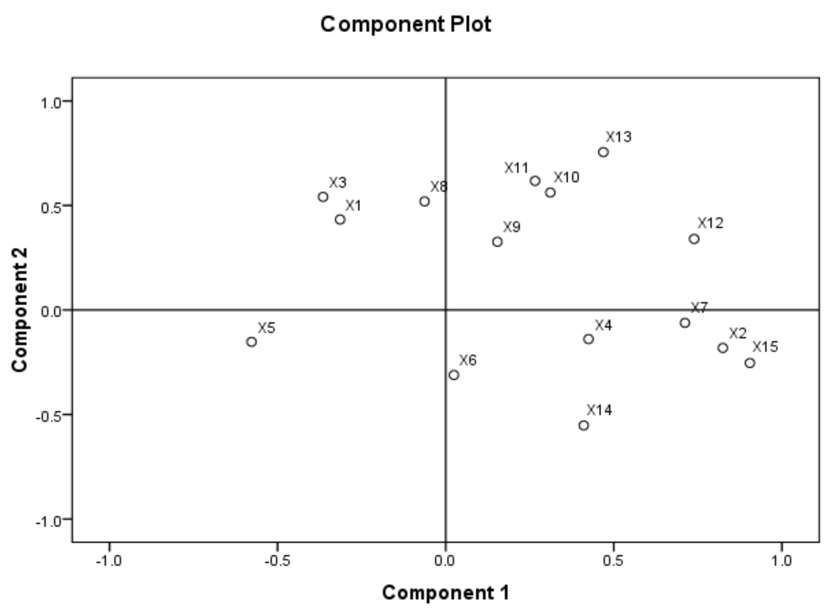

Fig. 2. Loading of different characters based on first two principle components

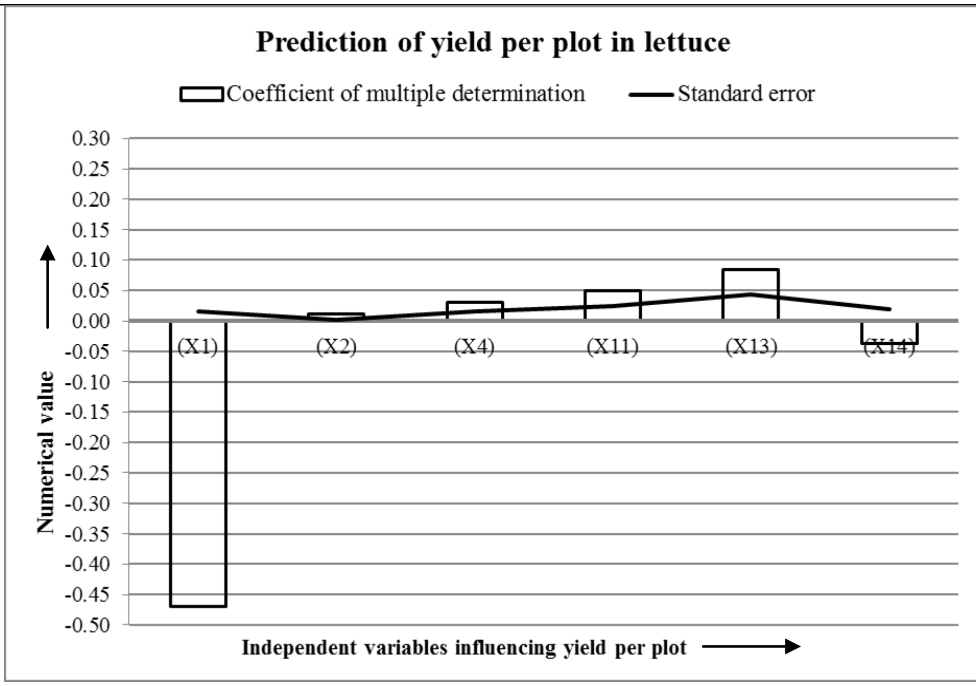

Fig. 3. Regression coefficient and standard error of multiple linear regression coefficient model based on independent variables influencing yield per plot in lettuce