International Journal of Chemical Studies

P-ISSN: 2349–8528
E-ISSN: 2321–4902
IJCS 2018; 6(4): 892-895

Received: 16-05-2018
Accepted: 24-06-2018

Narinder Panotra
OFRC, Sher-e-Kashmir University of Agricultural Sciences &Technology of Jammu, Jammu & Kashmir, India

Amitesh Sharma
OFRC, Sher-e-Kashmir University of Agricultural Sciences &Technology of Jammu, Jammu & Kashmir, India

M Iqbal Jeelani Bhat
Division of Statistics & Computer Sciences, Sher-e-Kashmir University of Agricultural Sciences &Technology of Jammu, Jammu & Kashmir, India

Rohit Sharma
RARS-Rajouri, Sher-e-Kashmir University of Agricultural Sciences &Technology of Jammu, Jammu & Kashmir, India

Correspondence
Narinder Panotra
OFRC, Sher-e-Kashmir University of Agricultural Sciences &Technology of Jammu, Jammu & Kashmir, India

Principal component analysis of yield characteristics of wheat (*Triticum aestivum* L. emend Fiori & Paol) in intermediate zone of Jammu and Kashmir

Narinder Panotra, Amitesh Sharma, M Iqbal Jeelani Bhat and Rohit Sharma

Abstract
The present study was carried out to study the variation among various yield characteristics of Wheat (*Triticum aestivum* L. emend Fiori & Paol) during Rabi 2013-14 and 2014-15. Nine yield characteristics *viz* Plant height, No. of effective tillers, No. of leaves/ plant, No. of grains/ ear, No. of grains/ plant, Grains yield quintal/ hectare, Straw yield quintal/ hectare, Biological Yield quintal/hectare and Harvest Index were studied from a wheat trail conducted at Regional Agricultural Research Station Rajouri, SKUAST-Jammu. The PCA analysis was used for depicting the proportion of variance among yield characteristics. PCA results revealed that first two components explain almost 92 % of the variation among yield characteristics, with highest proportion of contribution by Grains yield quintal/hectare attribute among nine attributes followed by No. of grains/ plant.

Keywords: wheat; PCA, scree plot, biplot

Introduction
Wheat (*Triticum aestivum* L.) is the world's leading cereal crop cultivated over an area of about 651 million tons making it the third most-produced cereal after maize and rice. India achieved remarkable progress in wheat production during the last four decades and is the second largest wheat producer in the world with the production touching a record level of 93.90 mt an area of around 28.40 m ha during 2011-12 (Anonymous, 2012), production has increased tremendously but is still far below the potential yield (11.2 tonnes/ha) (Singh *et al.*, 2010) [9]. Although, India is well placed in meeting its needs for food grains the major objective of food and nutritional secretary for its entire population has not been achieved. The demand for food grains is expected to rise not only as a function of population growth but also as more and more people cross the poverty line with economic and social development. Wheat production in India was about 6.46 million tonnes in 1950-51 and touched its peak of 76.37 million tonnes during 1999-2000. During the year 2006-07 its production was around 74.0 million tones. About 91% of the Indian wheat is produced in six states viz. U.P., Punjab, Haryana, M.P. Rajasthan and Bihar. Uttar Pradesh with 25.6 million tonnes in 2013 -14 continues to be the highest producer of wheat.
In the Jammu and Kashmir state about 4% area is under cultivation; 4% of this is irrigated and the rest is rainfed. And it has 13% area under forest and fruit trees. Maize and paddy are the principal crops, occupying 53% of the gross cropped area, followed by wheat (24%), oilseeds (7%) and pulses (4%); the remaining area is under barley, millets and other cereal crops.

Materials and Methods
A field experiment was conducted during winter season (*Rabi*) 2013-14 and 2014-15 at Regional Agricultural Research Station, Rajouri, SKUAST-Jammu. Nine yield characteristics *viz* Plant height, No. of effective tillers, No. of leaves/ plant, No. of grains/ ear, No. of grains/ plant, Grains yield quintal/ hectare, Straw yield quintal/ hectare, Biological Yield quintal/hectare and Harvest Index were studied under irrigated and rainfed conditions of Jammu and Kashmir. The soil of the experimental field was clay loam in texture with pH 7.2, EC 0.10 ds/m, organic carbon 0.68%, alkaline permanganate oxidizable available nitrogen 430 kg/ha, Olsen's available phosphorus 18.4 kg/ha and available potassium 246 kg/ha.

In order to study the proportion of varaince explained by each yield characteristics Principal components analysis was used. This technique was first described by Karl Pearson (1901) and the description of practical computing methods came much later from Hotelling (1933). The object of the analysis is to take p variables $X_1, X_2, \ldots X_p$ and find combinations of these to produce indices $Z_1, Z_2, \ldots Z_p$ that are uncorrelated. The absence of correlation means that the indices are measuring from different dimensions in the data. The indices are ordered so that Z_1 displays the large amount of variation, Z_2 displays the second largest amount of variation and so on. That is, $\text{var}(Z_1) \geq \text{var}(Z_2) \geq \cdots \text{var}(Z_p)$, where $\text{var}(Z_i)$ denotes the variance of Z_i in the data set. The Z_i are called principal components. Principal component analysis depends only on the covariance matrix Σ, or the correlation matrix of the variable under study. The best results are obtained when the original variables are very highly correlated, positively or negatively. In principal component analysis, the variance of most of the indices is low as to be negligible. In that case the variation in the data set can be described by the few Z variables with variances that are not negligible.

Let $X_1, X_2, \ldots X_p$ are variables under study, then the first principal component is the linear combination of the variables $X_1, X_2, \ldots X_p$,

$$Z_1 = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p$$

Such that variance of Z_1 is as large as possible subject to the condition that

$$a_{11}^2 + a_{12}^2 + \cdots + a_{1p}^2 = 1$$

This constraint is introduced because if this is not done, then $\text{var}(Z_i)$ can be increased simply by increasing any one of the a_{1j} values. The second principal component,

$$Z_2 = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p$$

is such that variance of Z_2 is as large as possible next to $\text{var}(Z_1)$ subject to the condition that

$$a_{21}^2 + a_{22}^2 + \cdots + a_{2p}^2 = 1 \text{ and } Cov(Z_1, Z_2) = 0$$

The third principal component,

$$Z_3 = a_{31}X_1 + a_{32}X_2 + \cdots + a_{3p}X_p$$

is such that variance of Z_3 is as large as possible next to $\text{var}(Z_1)$ and $\text{var}(Z_2)$

PCA analysis of yield characteristics of wheat was done in R-studio version 3.4.3 with library (FactoMineR) and library (devtools).R is very similar to S-Plus software and both are implementations of S language developed at Bell Laboratories, USA, which is the birth place of C language and unix operating system. Two fundamental books written by Becker *et al.* (1988) and Chambers & Hastie (1992) are of immense use for understanding this software. Khan and Mir (2005) and Jeelani *et.al* (2014) have discussed in detail the application of R- software in agricultural data analysis.

Results and Discussion

From Table 1 it is evident that first component accounts for 79% variability alone, remaining 21% variability is explained by the other 4 principal components, also from Figure 1 it is evident graphically that first principal component extracted maximum amount of variation among yield characteristics in the data. From Table 2 it can be seen that highest loading was observed in case of GYQPH followed by NOGPE with respect to PCA I, which suggests that these two characteristics are mainly responsible for highest amount of variation in the wheat data, with GYQPH contributing maximum to the first principal component followed by NOGPE, whereas NOGPP contributed minimum to the first principal component.

Conclusion

Multivariate data arise in most of the agricultural experiments. It is due to the fact that investigators often record data on more than one character of the same individual. These individuals may be plants, animals or other subjects related to agriculture. In most instances characters or variables are related to each other and hence principal component analysis is the obvious choice of data analysis in such cases. It has been found that these variables are isolated and univariate statistical tools are applied to analyze them, which is not fair. PCA has relevance in agricultural data analysis in the sense that most of the data is generated from the same individual which are often found to be highly correlated, and for a highly correlated data principal component analysis is one of the best multivariate techniques, especially in agronomical trails of wheat crop where various characteristics are related to each other.

Table 1: Summary results

	PC1	PC2	PC3	PC4	PC5
Eigen value	2.6811	1.0749	0.66999	0.45554	2.555e-15
Proportion of Variance	0.7987	0.1284	0.04988	0.02306	0.000e+00
Cumulative Proportion	0.7987	0.9271	0.97694	1.00000	1.000e+00

Table 2: (Loading; Plant height, No. of effective tillers, No. of leaves/ plant, No. of grains/ ear, No. of grains/ plant, Grains yield quintal/ hectare, Straw yield quintal/ hectare, Biological Yield quintal/hectare and Harvest Index.

	PC1	PC2	PC3	PC4	PC5
PLH	0.360	0.163	0.252	0.227	-0.642
NET	0.320	-0.383	-0.320	-0.476	-0.052
NOLP	0.325	0.421	0.241	0.203	0.314
NOGPE	0.356	-0.255	-0.084	0.205	-0.466
NOGPP	0.326	-0.098	0.627	-0.480	0.129
GYQPH	0.371	0.060	0.096	-0.091	0.252
SQPH	0.306	0.480	-0.355	0.130	0.124
BYQPH	0.347	0.114	-0.490	-0.223	0.069
HI	0.277	-0.573	0.008	0.577	0.411

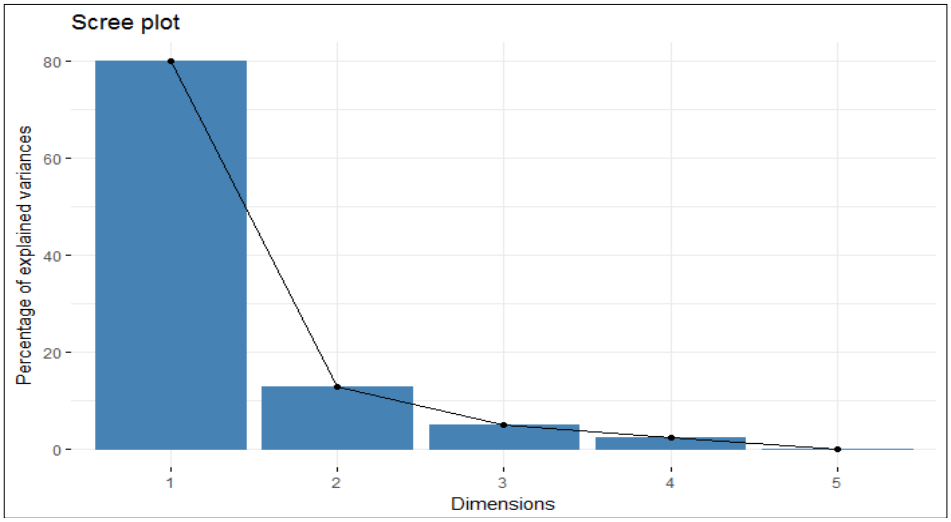

Fig 1: Scree plot of data showing that maximum variability is explained by first principal component

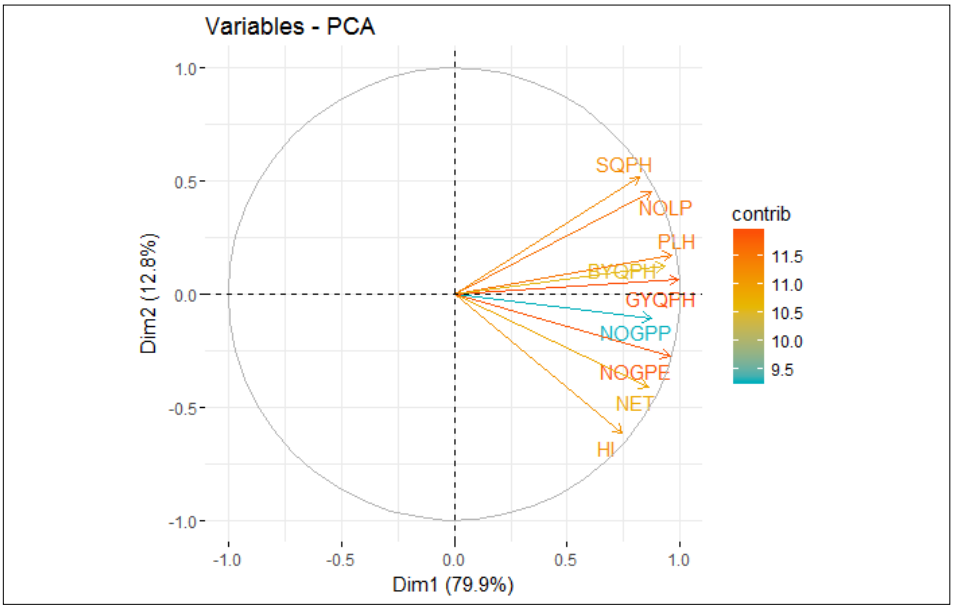

Fig 2: The length of arrows shows the proportion of contribution in the principal components and direction of arrows tells whether the proportion is positive or negative

References

1. Becker RA, Chambers JM, Wilks AR. The New S Language. Chapman & Hall, New York, 1988.
2. Chambers JM, Hastie TJ eds. Statistical Models in S. Chapman & Hall, New York, 1992.
3. Cheema MS, Akhtar M, Ali L. Effect of seed rates and NPK fertilizers on growth and yield of wheat variety punjnad-1. Pakistan journal of Agronomy. 2003; 2(4):185. 18
4. Jeelani MI, Nazir N, Mir SA, Jeealni F, Dar NA, Haq S *et al*. Application of Simple Random Sampling in Agriculture using R-software. Indian Journal of Science and Technology. 2014; 7(5):706-709.
5. Just, Richard E, Rulon D. Pope Production Function Estimation and Related Risk Consideration American Journal of Agricultural Economics. 1979; 276-284.
6. Khan AA, Mir AH. Applications of R- software in Agricultural Data Analysis. SKUAST Journal of Research. 2005; 7(1):36-64.

7. Mehra S. Instability in India Agriculture in the Context of the New Technology. International Food Policy Research Institute, Research Report No. 25, Washington DC, 1981.
8. Mishra B. Quality based procurement of wheat. The Hindu survey of Indian Agriculture. 2007, 32-35.
9. Singh AJ, Byerlee D. Relative Variability in wheat yields Across Countries and Over Time. Journal of Agricultural Economics. 1990, 21-32.
10. Xing MB, Tian YB, Song G, Cao S. Effects of fertilizers application on winter wheat root growth, grain yield and quality in purple soil, South West China Journal of Agricultural Science. 2005; 18(4):413-416.