

MULTIVARIATE ANALYSIS OF SPIKE YIELD TRAITS IN PRIMARY SELECTED NOVEL BREEDING HYBRID WINTER WHEAT GENOTYPES

Dekin Ghebriel Okbatinsae, postgraduate student of the Agrarian faculty of the Russian University; Peoples' Friendship University of Russia named after Patrice Lumumba, Russia, Moscow. ghebrielokba@gmail.com,

Burlutsky Valery Anatolievich, PhD in Agricultural Sciences, senior Lecturer of the Agrarian Faculty of the Peoples' Friendship University of Russia named after Patrice Lumumba, burlutskiy-va@rudn.ru

Abstract This study used cluster and principal component analysis to characterize spike yield traits in soft winter wheat hybrids. PCA identified four components explaining 94.67% of variance, with PC1 reflecting grain productivity and fertility as most dominant. Cluster analysis grouped genotypes into two: cluster one with superior spike productivity and fertility, and cluster two prioritizing spike length but with lower grain weight and fertility, indicating distinct yield strategies.

Key words: Cluster analysis, spike fertility, grain weight, principal component analysis, winter wheat

Introduction

Breeding of soft winter wheat aims to select optimal spike traits that strongly affect yield potential. Understanding trait interrelationships and their contributions to yield is essential for efficient breeding [5]. While univariate analyses provide insights, multivariate methods like Principal Component Analysis (PCA) and cluster analysis better capture the complex interplay of spike traits [3]. These approaches classify genotypes based on yield-related traits, revealing genetic diversity patterns that aid selection and help identify trait profiles maximizing grain yield [1]. This study utilized PCA and clustering on soft winter wheat spike trait data to identify optimal values and phenotypic groups, providing benchmarks to enhance breeding strategies focused on yield improvement.

Methods and materials

Plant material, Sowing, sample collection and parameters measured

A total of 47 genotypes/spikes, selected from initially evaluated material, were established in field plots measuring 25 meters in length and 2 meters in width. Each genotype was cultivated in an individual one-square-meter block. Sowing was performed at a density of 500 seeds per square meter. From each block, ten representative spikes were sampled to assess desirable trait profiles aligned with the objectives of optimal breeding. The sowing and subsequent harvesting of the selected top genotypes identified in the initial evaluation were conducted during the 2024/25 growing season. The parameters measured included spike length (sl), spikelet number (sn), grain number (gn), grain weight (gw), thousand grain weight (gw1000), derived

density distribution traits (ddsn, ddgn, ddgw), and efficiency-related traits such as fertility (%f), attributive value (atr%), and grain yield coefficient (gyc).

Data analysis: Principal component analysis (PCA) and cluster analysis were performed using Python version 3.11.7 within the Jupyter Notebook environment.

Results and discussion

Principal component analysis of spike yield-related traits

The principal component analysis (PCA) of the spike yield trait dataset revealed four principal components (PCs) accounting for a total of 94.67% of the variance (Table 1). The first principal component (PC1) explained the largest proportion of variance at 49.77%, indicating that nearly half of the total variability in spike traits is captured by this component. The second and third components (PC2 and PC3) explained 18.42% and 17.26% of variance respectively, with the top three components together accounting for 85.44% of the variability. The fourth component contributed an additional 9.23%, bringing the cumulative explained variance to 94.67%. The eigenvalue of PC1 was the highest at 5.594, confirming its dominant role in representing variation in the dataset. These results suggest that the majority of meaningful variability in spike yield traits can be effectively summarized by the first two to three principal components, simplifying the complexity of the data without significant loss of relevant biological information [3].

Table 1

Summary of principal components extracted from spike yield trait data

PC	Explained Variance (%)	Cumulative Variance (%)	Eigenvalue
PC1	49.77	49.77	5.594
PC2	18.42	68.19	2.070
PC3	17.26	85.44	1.939
PC4	9.23	94.67	1.037

Trait Loadings on Principal Components with eigenvalues greater than one

Trait loadings (Table 2) reflect the strength and direction of each trait's contribution to the principal components, with positive or negative values indicating the correlation sign. PC1 shows strong positive loadings from traits such as grain number (gn, 0.361), density distribution of grain number (ddgn, 0.404), density distribution of grain weight (ddgw, 0.363), percent fertility (%f, 0.339), and attributive value (%atr, 0.401), suggesting that PC1 captures variation primarily associated with grain productivity and fertility components. PC2 is characterized by notable positive loading for spike length (SL, 0.561) and strong negative loading for density distribution of spikelet number (ddsn, -0.550), indicating this component represents contrasting influences of spike size versus spikelet development. PC3 relates strongly to thousand kernel weight (gw1000, 0.624) and negatively to spikelet

number (sn, -0.444), suggesting this PC differentiates grain size from spikelet count effects. PC4 is less dominant but shows positive association with spikelet number (sn, 0.676) and negative with fertility percentage (-0.416), highlighting more nuanced aspects of spike fertility dynamics [4]. Understanding these trait loadings aids in targeted selection and breeding strategies by focusing on PCs representing important yield determinants in hybrid winter wheat.

Table 2

Loadings of spike yield-related traits on principal components extracted from phenotypic data, indicating trait contributions and associations for pcs with eigenvalues >1

Trait	PC1	PC2	PC3	PC4
sl	-0.174	0.561	-0.273	0.218
sn	0.157	-0.069	-0.444	0.676
gw	0.306	0.408	0.194	0.236
gn	0.361	0.202	-0.324	-0.025
gw1000	-0.102	0.215	0.624	0.311
ddsn	0.242	-0.550	0.019	0.231
ddgw	0.363	-0.046	0.363	0.094
ddgn	0.404	-0.173	-0.099	-0.135
%f	0.339	0.278	-0.137	-0.416
%atr	0.401	0.118	0.154	-0.163
gyc	0.285	0.039	0.111	0.252

Scree plot of principal component eigenvalues for winter wheat trait data

The scree plot(Figure 1) indicates the eigenvalues of principal components, highlighting that the first four components capture most of the variance in the dataset. Components with eigenvalues above 1 are considered meaningful, while those below 1 contribute less variance than an original variable and are usually discarded, helping to determine the optimal number of components to retain for effective data summarization [1].

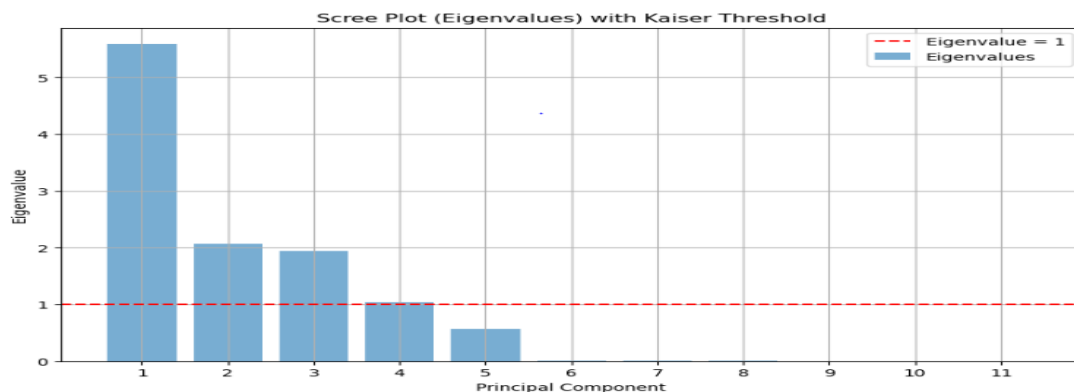


Figure 1. Scree plot showing the eigenvalues with Kaiser thresholds of principal components derived from winter wheat trait data

PCA Biplot of Spike Yield Traits in Winter Wheat Genotypes

The PCA biplot (Figure 2) shows winter wheat genotype distribution along with directions of principal spike yield-related traits on principal components PC1 and PC2. PC1, which accounts for 49.8% of variance, discriminates genotypes along lines of grain weight (gw), grain number (gn), and fertility (%f) and, attributive value (%atr) such that these significantly impact genotype discrimination on this axis. PC2 represents 18.4% of variance and discriminates genotypes along lines of spike length (sl) and density distribution of spikelet number (ddsn) in directions opposite to PC1. Angles between character vector directions indicate correlation; short angled arrows such as gw and gn imply positive correlation whereas nearly orthogonal directions imply poor correlation. Genotypes that are bunched close to character vectors exhibit those attributes strongly, which enhances targeted selection of lines with favorable yield attributes [1]. This biplot aids in interpreting multivariate association of traits with genotypes as well as identification of top genotypes of wheat for breeding applications.

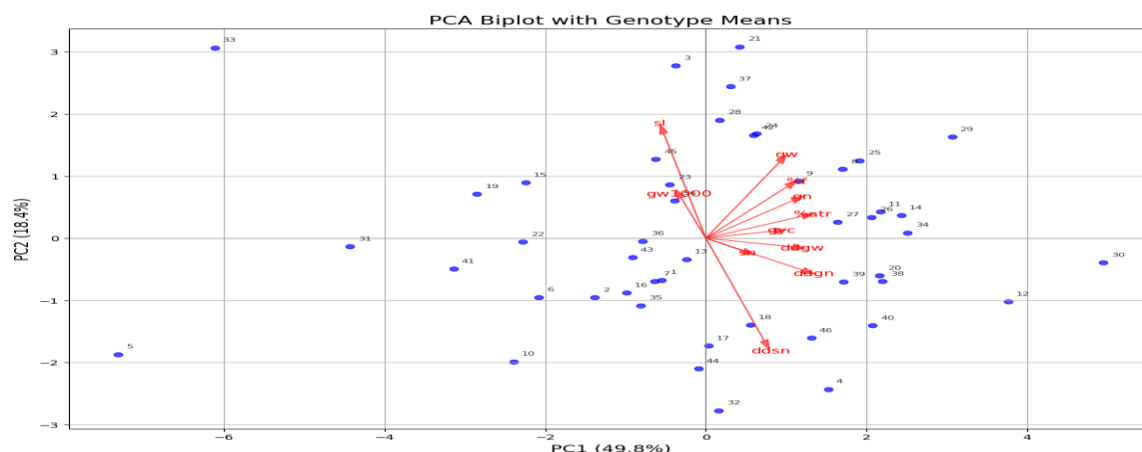


Figure 2: Principal Component Analysis (PCA) biplot displaying genotype (blue points) and trait loadings (red vectors) for winter wheat spike yield-related traits

Determination of optimal cluster number for grouping genotypes using spike traits by elbow and silhouette methods

The Elbow method (Figure 3) plot demonstrates a rapid drop in inertia from 1 to 2 clusters, indicating significant reduction of variance within clusters; beyond this, the decline flattens, suggesting 2 clusters may sufficiently capture natural groupings. Conversely, Silhouette scores, which measure cluster cohesion and separation, show a fluctuating trend with the highest peak at 2 clusters, reflecting better-defined cluster separation at that point. Using both methods, the cluster number

choice involves balancing inertia minimization and silhouette maximization based on analysis objective, with k value 2 being reasonable candidate for the winter wheat genotype dataset.

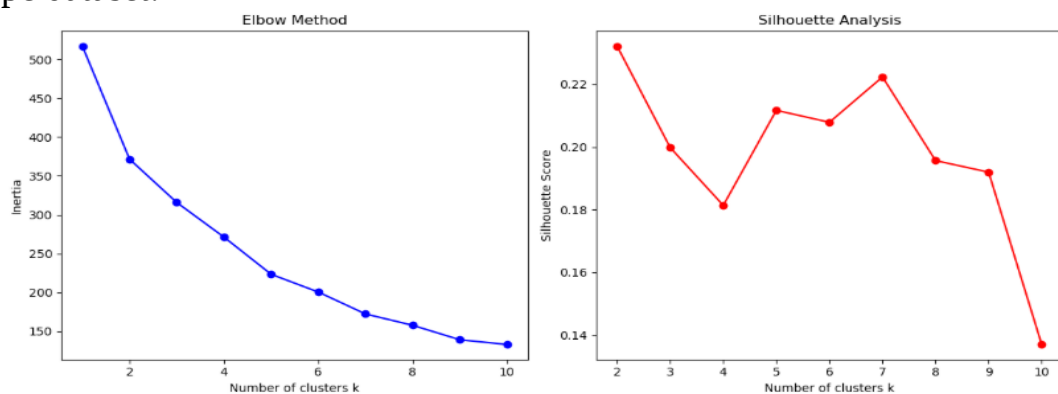


Figure 3 : Elbow Method and Silhouette scores indicating clustering performance for various k values, with optimal cluster count K=2

Assessment of phenotypic diversity and cluster structure using hierarchical clustering

The dendrogram (Figure 4) illustrates two primary clusters: one depicted in orange on the left and the other in green on the right. Within each primary cluster, smaller sub clusters categorize genotypes with heightened similarity as the vertical connecting height (distance) diminishes. The elevated point at where the orange and green branches converge signifies a stronger dissimilarity between these two primary groups. This hierarchical structure facilitates the visualization of relations among genotypes by displaying their grouping based on phenotypic or genetic similarities through a sequential agglomeration process [2]. It is useful for identifying genetically different groups and investigating genetic diversity in a dataset.

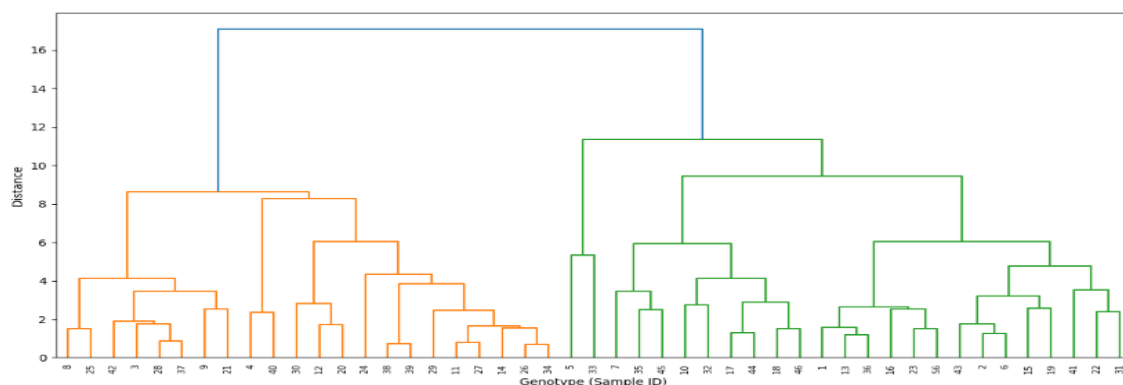


Figure 4: Dendrogram of agglomerative hierarchical clustering(ward linkage) demonstrating genotypic relationships among soft winter wheat hybrids

Hierarchical cluster analysis of genotypes based on spike yield related traits

The results (Table 3) presents the mean values of the traits for two clusters derived from hierarchical clustering for hybrid wheat genotypes using the characters for spike yield. There are 22 entries for Cluster 1 and 25 entries for Cluster 2. As a whole, Cluster 1 genotypes show their supremacy in terms of individual spike productivity and fertility-related factors possibly influencing positive outcomes in grain yield potential. Genotypes in Cluster 2 could possibly display different allocation strategies in terms of prioritizing higher length of spikes with reduced individual grain weight and fertility. Based on these observations, there could be scope for using Cluster 1 in prioritizing breeder selection for higher spike fertility values in terms of grain filling potential for improving yield enhancement. This multi-variant clustering technique helps to effectively identify the grouping of genotypes according to complex traits, thereby also helping in the desired selection of superior hybrid wheat genotypes for breeding [5].

Table 3

Mean trait values across hierarchical clusters of hybrid winter wheat genotypes

Cluster	Cluster size	sl	sn	gw	gn	gw1000	ddsn	ddgw	ddgn	%f	%atr	gyc
1	22	9.2	17.3	2.6	48.8	54.3	19.1	2.9	53.4	70.4	20.3	81.8
2	25	9.3	17.1	2.2	41.4	53.3	18.6	2.4	45.1	60.3	14.6	80.0

Conclusion

Principal component analysis identified four key components explaining 94.67% of phenotypic variance in soft winter wheat spike traits. PC1 was mainly related to fertility and grain weight, PC2 to spike morphology, and PC3 and PC4 to kernel size versus spikelet architecture. These components reveal important trait combinations underlying phenotypic diversity and yield potential. Cluster analysis distinguished two groups: Cluster 1 genotypes excelled in spike productivity and fertility traits likely enhancing grain yield, whereas Cluster 2 genotypes appeared to prioritize producing more spikes but with lower individual grain weight and fertility, indicating different resource allocation strategies.

References

1. Adilova S. S., Qulmamatova D. E., Baboev S. K., Bozorov T. A., Morgunov A. I. Multivariate cluster and principal component analyses of selected yield traits in Uzbek bread wheat cultivars // American Journal of Plant Sciences. 2020. Vol. 11, no. 6. P. 903. DOI: 10.4236/ajps.2020.116066.
2. Balcha U., Mekbib F., Lule D. Cluster and principal component analysis among bread wheat (*Triticum aestivum* L) genotypes in mid rift valley of Oromia, Ethiopia // Advances in Crop Science and Technology. 2022. Vol. 10, no. 8. P. 525.

3. Bibi S., Qayyum A., Bibi A., Begum K. Multivariate analysis of selected agro-morphological traits of common wheat genotypes for estimation of genetic diversity // Emirates Journal of Food and Agriculture. 2022. Vol. 34, no. 12. P. 1020–1027.

4. Pradhan S., Babar M. A., Robbins K., Bai G., Mason R. E., Khan J., Shahi D., Avci M., Guo J., Maksud Hossain M., Bhatta M., Mergoum M., Asseng S., Amand P. S., Gezan S., Baik B. K., Blount A., Bernardo A. Understanding the genetic basis of spike fertility to improve grain number, harvest index, and grain yield in wheat under high temperature stress environments // Frontiers in Plant Science. 2019. Vol. 10. DOI: 10.3389/fpls.2019.01481.

5. Zewdu D., Mekonnen F., Geleta N. Cluster and principal component analysis for yield and yield related traits of bread wheat (*Triticum aestivum* L.) genotypes // AGBIR. 2024. Vol. 40, no. 2. P. 962–967.

NANO-FERTILIZER AND GROWTH REGULATORS TECHNOLOGY ON: IMPROVING CROP PRODUCTIVITY, AND NUTRITIONAL CONTENT OF SPRING WHEAT (*TRITICUM AESTIVUM* L)

Shimendi G Okbagabir, Department of Agrobiotechnology, Agrarian Institute Technological, People's Friendship University (RUDN), Moscow, Russia, named after Patrice Lumumba, shimendigde20@gmail.com, (Tel. No +79777450344)

Elena N. Pakina, Department of Agrobiotechnology, Agrarian Institute Technological, People's Friendship University (RUDN), Moscow, Russia, named after Patrice Lumumba, pakina-en@rudn.ru, +79853472754

Abstract A 2023–2024 field study at Agropolygon Research Institute assessed foliar/nano-fertilizers (Terrasay) and growth regulators (Centrino) on wheat using a randomized complete block design. Treatments (F0–F3 at 0–2 l/ha; GR1–GR3 at 0–1.5 l/ha) significantly affected tiller count, flowering, grain yield (peaking at 4142.9 kg/ha with GR2), and nutrient content (e.g., 28.25% wet gluten with F3 × GR2). Notable interactions ($p \leq 0.048$) included F2 × GR2 and F3 × GR3, enhancing fertile tillers, spike elongation, and grain weight. Tailored nano-fertilizer and growth regulator combinations show promise for sustainable wheat production.

Key Words: Foliar Nutrients; Growth Regulators; Spring Wheat; Productivity; and Qualitative traits.

INTRODUCTION

Chemical fertilizers particularly micronutrients like manganese, copper, boron, iron, molybdenum, zinc etc. are important for the growth and development. A substantial increase of crop yields with the green revolution and new farming practices has progressively decreased the micronutrients of soil like zinc, iron, and molybdenum [1]. Foliar application of iron compounds with the technology of nanoparticles may be a solution to the problem. The concentrations of iron nano-oxide solution has been used in five levels (0, 0.01%, 0.02%, 0.03% and 0.04%) to check