https://doi.org/10.61308/QQFA6344

Use of cluster analysis for grouping and evaluation of landrace populations and foreign synthetic maize lines

Yuliana Yordanova[1]*, Nataliya Petrovska[1*], Petyo Dauldzhiev[1], Katya Uzundzhalieva[2]

[1]Maize Research Institute – Knezha, 5835, Agricultural Academy, 1373, Sofia, Bulgaria

[2]Institute of Plant and Genetic Resources "K. Malkov", 2 Druzhba str. 4122 Sadovo, Agricultural Academy, 1373, Sofia, Bulgaria

*E-mail: *yuliannayordanowa@abv.bg*

Citation: Yordanova, Yu., Petrovska, N., Dauldzhiev, P., & Uzundzhalieva, K. (2026). Use of cluster analysis for grouping and evaluation of landrace populations and foreign synthetic maize lines. *Bulgarian Journal of Crop Science, 63*(4) 23-29

Abstract: This study applies statistical methods to analyze the morphological and genetic variability of 20 landrace and 20 foreign synthetic maize lines (*Zea mays L.*). Using hierarchical cluster analysis (Ward's method, Euclidean distance), four main clusters were identified, reflecting differences in productivity, adaptability, and morphological characteristics. The evaluation of the tested landrace populations and foreign synthetics was based on the following biometric indicators: ear length (cm), number of rows per ear, number of kernels per row, kernel length (mm), 1000-kernel weight (g), total plant height (cm), height to tassel (cm), ear leaf length (cm), ear leaf width (cm), height to main ear (cm), number of leaves above the upper ear, number of ears per plant, and hectoliter mass (kg/l). The mathematical approach to grouping the lines allows for the identification of differences between them, highlighting valuable traits, and forming clusters based on their degree of similarity. The landraces are grouped into two clusters: compact and stress-resistant varieties and high-yielding adaptive lines, while the foreign lines are also divided into two subgroups: high-protein hybrids and intensive productive varieties.

Keywords: hierarchical cluster analysis; morphological variability; productivity; adaptability; biometric indicators; maize

INTRODUCTION

The value of genetic diversity lies in its ability to provide genetic barriers against various biotic and abiotic stress factors. Increasing genetic variability is essential for managing pest and disease resistance and offers opportunities for further crop improvement (Hughes et al., 2004; Hajjar et al., 2008). According to Miku (1980, 1981), in addition to the geographical distance between source materials, "the greatest genetic differences exist among forms from different maturity groups, due to limited gene flow between them." Crosses between such materials may result in novel breeding lines combining high productivity typically inherited from late-maturing lines with low grain moisture typically from early-maturing genotypes (Kang and Zuber, 1987; Medvedev, 2007).

Among the secondary traits in maize, grain number per unit area and the interval between tassel and silk flowering are of particular importance, as grain yield is directly influenced by the number of kernels at harvest (Vulchinkov and Vulchinkov, 2023). Cluster analysis enables the grouping of breeding materials based on similarities or differences across studied parameters, revealing relationships across the full dataset (Dimova and Bozhinov, 2002; Roychev et al., 2008; Zhuzhukin et al., 2010).

The application of cluster analysis in the evaluation of breeding material provides plant breeders with a powerful tool for planning and optimizing selection strategies (Ahmad et al., 2008).

In this context, various authors recommend the use of multivariate statistical methods (Brown, 1991; Jorge, 2003). Genotypic differentiation is most pronounced during the tasseling stage. Through clustering, genotypes can be grouped based on their phenotypic expression of specific traits or trait combinations (Ivanov, 2014).

The primary aim of this study is to apply cluster analysis for grouping based on key biometric traits, yield components, and seed productivity. This analytical approach has previously been employed in characterizing maize and sunflower germplasm (Binohd et al., 2007; Chambo, 2014; Ganusheva et al., 2010; Ilchovska et al., 2014; Lima et al., 2005).

MATERIALS AND METHODS

The study was conducted during the 2023–2024 period at the experimental field of the Maize Research Institute in Knezha. A total of 180 landrace populations and 120 foreign synthetic lines from the gene bank collection at the MRI-Knezha were evaluated. From these, 20 landraces and 20 foreign genotypes were selected for cluster analysis.

To assess the productivity of the maize populations, a randomized block design was used with three replications. Each experimental plot covered 0.5 hectares. Sowing was carried out manually, with a row spacing of 70 cm and in-row spacing of 25 cm.

The evaluation of the tested landraces and foreign genotypes was based on the following biometric traits: cob length (cm), number of rows per cob, number of kernels per row, kernel length (mm), 1000-kernel weight (g), plant height (cm), height to tassel (cm), length of the husk leaf (cm), width of the husk leaf (cm), height to the primary cob (cm), number of leaves above the uppermost cob, number of cobs per plant, and hectoliter weight (kg/l).

A mathematical approach was used to group the lines and assess genetic distances between them. Grouping was performed using the Ward's method of hierarchical clustering (Ward, 1963; Duran and Odell, 1977). Euclidean distance was applied as a similarity measure.

A dendrogram was constructed to visually present the resulting clusters. To avoid distortion due to differences in scale among variables, the data were standardized prior to / before analysis. Statistical processing was performed using SPSS software (Kuneva and Tahsin, 2015).

RESULTS AND DISCUSSION

Biometric traits of the plant and ear are critical components of yield and serve as key indicators of productivity in both landrace populations and foreign synthetics. Measurements, such as total plant height (cm), height to tassel (cm), length and width of the husk leaf (cm), height to the primary ear (cm), number of leaves above the uppermost ear, and number of ears per plant, were recorded during the tasseling stage when maize plants are most susceptible to environmental stress during the growing season.

Cob biometric parameters, 1000-kernel weight, and hectoliter weight were recorded after harvest. Collectively, these traits provide essential information for evaluating the adaptive potential and genetic diversity among the tested maize accessions.

The biometric characteristics of the plant and the ear are critical components of yield and serve as key indicators of productivity in both local landrace populations and foreign synthetic varieties.

Measurements such as total plant height (cm), height to tassel (cm), length and width of the husk leaf (cm), height to the primary ear (cm), number of leaves above the uppermost ear, and number of ears per plant were recorded during the tasseling stage, when maize plants are most susceptible to environmental stress during the growing season. These traits were considered in the characterization of the genotypes (samples) for drought tolerance (water deficit).

Biometric parameters of the ear, 1000-kernel weight, and hectoliter weight were recorded after harvest. Taken together, these character-

istics provide important information for evaluating the adaptive potential and genetic diversity among the tested maize accessions. All of these indicators were taken into account in the characterization of the genotypes according to their adaptive potential or based on preliminary evaluation data.

In the hierarchical cluster analysis, at a higher level of separation (at an Euclidean distance of approximately 200), two main clusters were formed, each containing two larger subgroups. The first main cluster (located on the right side of the figure) combines the local landrace populations, characterized by high adaptability and stable productivity under variable climatic conditions.

The second cluster includes the foreign synthetic lines, distinguished by their high breeding value and responsiveness to intensive maize cultivation.

Thus, the main division of the dendrogram clearly coincides with the origin and biological characteristics of the materials—local landraces versus foreign synthetics.

Cluster I – local maize landraces

Cluster I includes the local landrace populations that demonstrate strong ecological plasticity and stable yields under different soil and climatic conditions. Within this cluster, two clearly distinguishable subgroups are formed.

The first subgroup combines stress-tolerant genotypes such as Svilengrad, Karamantsi, Gyulevo, and Markovo, which are characterized by shorter plant height (120–135 cm) and good kernel filling even under drought or low-temperature conditions. These accessions are typical representatives of adaptive local material developed through natural selection under environmental stress. They are valuable as sources of genes for drought and cold tolerance, as well as for yield stability under climate change conditions.

The second subgroup includes high-yielding adaptive populations such as Bdintsi, Florosh, Pishurka, Bozhentsi, and Zemen, which exhibit medium plant height (140–160 cm), high 1000-kernel weight (>285 g), and hectoliter weight exceeding 770 kg/l. These genotypes show a favorable balance between productivity and stress tolerance, making them suitable for areas with better soil conditions and intensive cultivation systems. The high hectoliter weight indicates dense and high-quality grain, serving as an indicator of good nutrient accumulation and high energy value.

The overall characterization of the first cluster can be defined as an "adaptive type with resilience and stable productivity."

The differences between the subgroups are mainly due to morphological traits such as plant height, ear size, and grain mass, while their internal proximity indicates a common genetic origin and selection direction toward adaptability.

Cluster II – foreign synthetics

The second main cluster includes foreign synthetic genotypes, characterized by high breeding value, large kernels, and stable yield under intensive production conditions. Within this cluster, two distinct subgroups are also identified.

The first subgroup comprises populations with high technological and breeding value, such as Sinya L2 (Co), Sint L2 (C2), Sint 16 (C1), and Sint 2/96 "L" (Co). These genotypes are distinguished by high hectoliter weight (>780 kg/l), large kernel length (>8 mm), and yield exceeding 6 t/ha. They are suitable for inclusion in breeding programs aimed at increasing yield potential and grain quality. Their stability and high technological performance make them valuable sources of genetic material for improving hybrid maize varieties.

The second subgroup includes intensively productive synthetic populations, such as Po Synth, Synth min spons, No. 13 Synthetic (Argentina), and Mexican deute pon 831268-4-2-1, which exhibit exceptional responsiveness to improved agronomic practices and high production efficiency. These genotypes have been selected for conditions of intensive agriculture, where maximum yield and economic efficiency are the primary objectives. They are characterized by high kernel mass, elevated starch content, and excellent technological suitability for both industrial and feed use.

The overall characterization of this cluster can be defined as an "intensive type with high breeding and technological value."

Here, the main grouping factor is the breeding direction and the degree of responsiveness to intensive cultivation, which distinguishes these genotypes from the more adaptive local lines of the first cluster.

At a lower level of separation (i.e., at a smaller Euclidean distance - around 100–150), the dendrogram reveals the formation of three smaller clusters, each containing two subgroups.

This indicates that the internal variability within the main groups is substantial, particularly among the local populations, which include both high-yielding and moderately productive but highly stress-tolerant forms.

In this way, the dendrogram reflects the hierarchical nature of genetic relatedness - at higher levels, two main groups (local vs. foreign) are clearly distinguished, while at lower levels, additional substructures emerge, associated with specific morphological and yield-related traits.

The overall grouping pattern demonstrates that genetic origin and breeding history are the primary factors determining the dendrogram structure.

The local populations (Cluster I) are grouped due to their adaptability, stress tolerance, and stable productivity, whereas the foreign synthetic lines (Cluster II) are united based on high yield potential, large kernel size, and superior technological quality. This clearly expressed differentiation confirms that the adaptive strategies of the local material differ significantly from those of an intensively selected foreign population, in a manner that is both statistically supported and biologically meaningful.

The separation observed at an Euclidean distance of around 200 reflects the natural boundary of genetic similarity between the two major groups and demonstrates that the cluster model provides the most accurate description of the analyzed dataset. At the same time, the internal subdivision within each cluster emphasizes the richness of genetic resources and the potential for future breeding programs aimed at sustainable and efficient maize production under diverse environmental conditions.

The results are showin in Figure 1, **Cluster** I – the local maize populations, and Cluster II the foreign synthetics**.**

The graphical representation of the mean values of the analyzed traits for the two main clusters reveals distinct quantitative differences between the local and foreign genetic materials. As shown in Figure 2, the foreign synthetics (Cluster II) outperform the local landrace populations (Cluster I) in most morphological and productive traits related to yield potential and plant structure.

In the local populations, lower mean values were observed for plant height, cob length, and 1000-kernel weight. However, they maintain a high hectoliter weight, which indicates good kernel density and high physiological grain quality. This suggests efficient resource utilization and stable productivity under environmental stress—characteristic features of adaptive local forms that have evolved under variable climatic conditions.

Conversely, the foreign synthetics exhibit higher mean values for plant growth, kernel mass, and cob length. This confirms their breeding orientation toward intensive production and high yield potential, while also indicating a certain dependence on optimal agronomic conditions. The increased values of cob- and grain-related traits demonstrate a significant genetic potential for yield improvement within breeding and hybridization programs.

Overall, the graph highlights the complementary nature of the two clusters: the local populations are distinguished by resilience and adaptability, while the foreign genotypes demonstrate high productivity and technological value.

These differences in mean biometric parameters emphasize both the physiological specificity of each group and their potential importance for breeding and sustainable maize production**.**

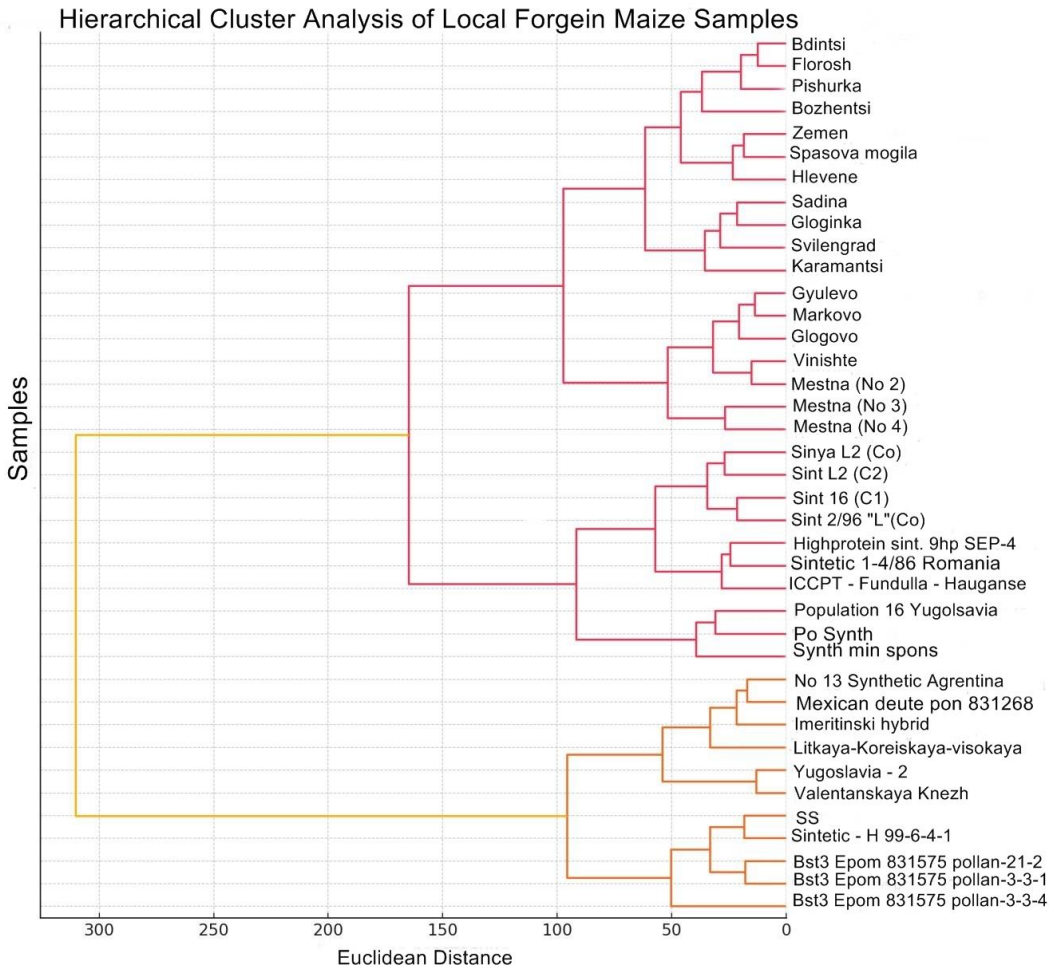


Figure 1. Dendrogram based on average intergroup Euclidean distances

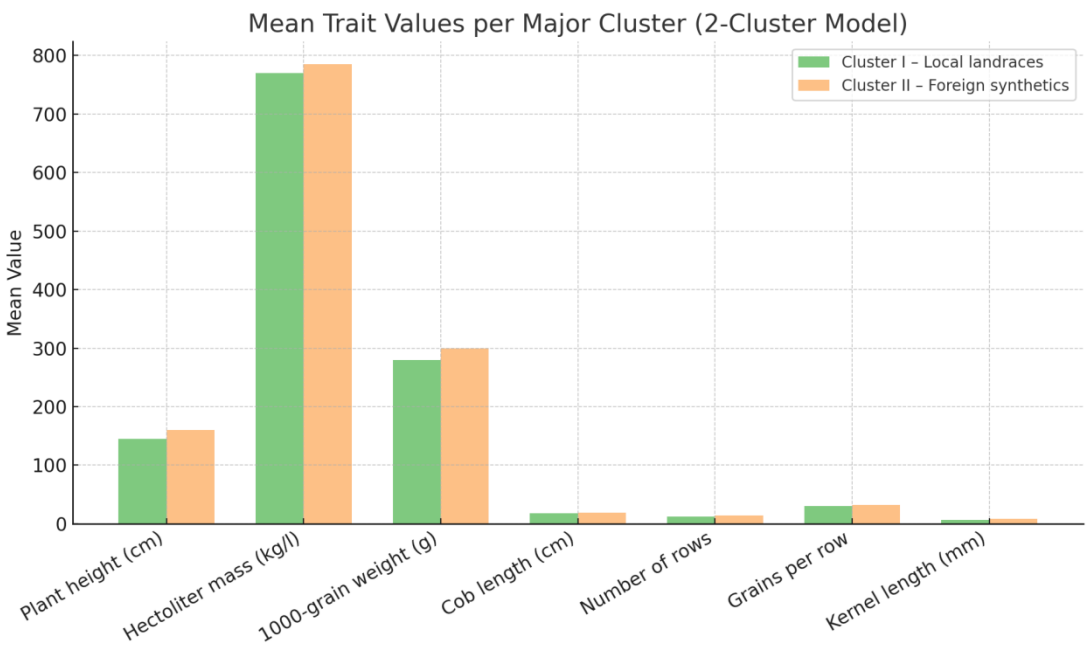


Figure 2. Mean values of key agromorphological traits across clusters of landraces and foreign synthetics

CONCLUSIONS

• The hierarchical cluster analysis (Ward's method, at an Euclidean distance of approximately 200) divided the 40 analyzed genotypes into two clearly distinct main clusters – local landrace populations (Cluster I) and foreign synthetic lines (Cluster II).

The main division of the dendrogram reflects the genetic origin and breeding history of the studied material – the local gene pool has been shaped by adaptive mechanisms, while the foreign synthetics are the result of targeted breeding for high productivity and technological value.

• The local populations are characterized by low to medium plant height, moderate yield parameters, and high hectoliter weight, which indicates good kernel density and physiological grain quality. These lines demonstrate high adaptability and yield stability under unfavorable climatic conditions. The foreign synthetics exhibit higher mean values for most biometric traits – plant height, ear length, 1000-kernel weight, and kernel size. They represent an intensive type of breeding material, optimized for high productivity and economic efficiency under favorable agronomic conditions.

• The obtained results confirm the complementary nature of the two clusters – the local populations can serve as sources of resilience and adaptability, while the foreign synthetics can act as donors of high yield potential and technological quality. The observed intra-cluster variability reveals valuable genetic potential that can be utilized in future breeding programs aimed at developing hybrids combining both stress tolerance and high productivity.

• The cluster model of grouping is confirmed both statistically and biologically, accurately describing the structure and relationships among the analyzed genotypes and emphasizing the potential of an integrated approach to evaluating maize genetic resources.

REFERENCES

Ahmad, I., Anjum, F. M., Butt, M. S., Hussain, S., & Khan, M. I. (2008). Predictive modeling of spring wheat varieties by cluster analysis. *International Journal of Food Properties, 11*(2), 310–320.

Binodh, A. K., Manivannan, N., & Varman, P. V. (2007). Cluster analysis of yield traits in sunflower (*Helianthus annuus* L.). *Madras Agricultural Journal, 94*(1–6), 27–31.

Chambo, E. D., Correia, A. F., Cunha, F., Garcia, R. C., Oliveira, N. T. E., Vasconcelos, E. S., & Silva, N. L. S. (2014). Cluster analysis of sunflower genotypes grown in conditions of free and restricted pollination. *Scientia Agraria Paranaensis, 13*(Suppl.), 323–328.

Dimova, D., & Bozhinov, B. (2002). Prilozhenie na klasternia analiz i analiza na osnovnite komponenti za otsenka na selektsionni materiali. In *50 godini Dobrudzhanski zemedelski institut – Yubileyna nauchna sesia* (pp. 308–312).

Ganusheva, N., Mokreva, T., Popova, Z., & Andonov, R. (2010). Genetic distance on quantitative traits in promising lines and varieties of malting barley. *Plant Science, 6*(1), 31–36.

Ilchovska, M., & Ivanova, I. (2014). Usage of cluster analysis for grouping hybrids and evaluation of experimental mutant maize hybrids. *Agricultural Science and Technology, 6*(1), 14–16.

Ivanov, L. (2014). Harakteristika na samooprasheni linii tsarevitsa chrez metodite klasteren analiz i analiz na komponentite. *Izsledvaniya varhu polskite kulturi, 9*(1), 49–56.

Jorge, F. J., Crossa, J., Taba, S., & Shands, H. (2003). A multivariate method for classifying cultivars and studying group × environment × trait interaction. *Crop Science, 43*, 1249–1258.

Kang, M. S., & Zuber, M. S. (1987). Grain dry-down studies in corn. *Proceedings of the 42nd Annual Corn and Sorghum Research Conference*, 220–226.

Kuneva, V., & Tahsin, N. (2015). Use of cluster analysis for grouping and evaluation of different oil sunflower hybrids. In *Jubilee Scientific Conference with International Participation „Traditions and Challenges of Agricultural Education, Science and Business". Scientific Works of Agricultural University – Plovdiv, 59*(2).

Lima, I. M., & Guraya, H. S. (2005). Optimization analysis of sunflower butter. *Journal of Food Science, 70*(6).

Medvedev, A. E. (2007). *Sozdanie rannespelykh samoopylennykh liniy kukuruzy metodom rekurrentnogo otbora v sinteticheskikh populyatsiyakh s razlichnoy geneticheskoy osnovoy* (Doctoral dissertation). Krasnodar.

Roychev, V., Bozhinov, B., & Dimova, D. (2008). Palino-biometrichni izsledvania na bezsemenni, semenni i *in vitro* razmnozheni sortove lozi. *Rastenievadni Nauki, 45*, 269–276.

Russell, J. R., Fuller, J. D., Macaulay, M., Hatz, B. G., Jahoor, A., Powell, W., & Waugh, R. (1997). Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theoretical and Applied Genetics, 95*(4), 714–722.

Vulchinkov, Z., & Vulchinkov, S. (2023). Variation and correlations between characteristics in maize hybrids from different vegetation groups. *Journal of Mountain Agriculture on the Balkans, 26*(4), 180–200.

Zhuzhukin, V. I., Gudova, L. A., & Solovov, D. P. (2010). Faktornyiy analiz vzaimosvyazey priznakov kukuruzyi. *Kukuruza i Sorgo, 1*, 18–20.

Received: May, 29, 2026; *Approved:* July, 08, 2026; *Published:* August, 2026