ONLINE JOURNAL OF BIOLOGICAL SCIENCES

RESEARCH ARTICLE

Cluster and Correlation Analysis of Valuable Traits of Bean Cultivars Under Irrigation Conditions in Kazakhstan

Kudaibergenov Mukhtar Sarsenbekovich [1], **Bastaubaeva Sholpan Orazovna** [1], **Kanatkyzy Makpal**[1], **Saken Gaukhar Sakenkyzy**[1,2]*, **Urazaliev Kairat Rahimovich** [2]

[1]Kazakh Research Institute of Agriculture and Plant Growing, Almaty region, Kazakhstan
[2]Kazakh National Agrarian Research University, Kazakhstan

*Corresponding Author: gaukharsaken@mail.ru

Abstract: This study aimed to investigate a diverse collection of beans accessions under irrigation conditions in the Almaty region of Kazakhstan to identify promising genotypes for breeding. Over a ten-year period (2014–2024), 300 accessions were analyzed, of which 100 were selected for detailed assessment based on key agronomic traits including earliness, productivity, seed weight, disease resistance, protein content, and adaptability to mechanized harvesting. Cluster and correlation analyses were employed to classify accessions and determine interrelations among traits. Cluster analysis revealed three major groups based on productivity and seed quality traits. Correlation analysis indicated strong positive associations between plant productivity components such as number of productive nodes, number of pods per plant, and seed weight per plant (r = 0.91). A moderate positive correlation (r = 0.52) was found between 1000-seed weight and overall plant weight, while negative correlations were observed between productivity traits and seed protein content. The findings provide valuable insights for developing high-yielding, protein-rich, and machine-harvest-adapted beans varieties suitable for the agroclimatic conditions of Kazakhstan.

Keywords: Beans, Genetic Diversity, Trait Correlations, Protein Yield, Morpho-Agronomic Traits, Germplasm

Received: 18-04-2025 | **Revised:** 13-06-2025 | **Accepted:** 15-07-2025 | **DOI:** 10.3844/ojbsci.2026.26.01.011

Introduction

Beans are one of the most important agricultural crops, with a centuries-old history of cultivation and valued for their rich nutritional profile and role in food security. They provide a significant source of protein, carbohydrates, dietary fiber, vitamins, and minerals, making them particularly valuable in regions with limited access to animal products [1].

According to the Food and Agriculture Organization of the United Nations (FAO), global common bean production in 2019 reached approximately 28.9 million metric tons, cultivated across 33.1 million hectares. The leading producers were India (5.31 million metric tons), followed by Brazil (2.9 million metric tons) and China (1.2 million metric tons). This data underscores the continued importance of common beans as a staple food and a source of income for millions, particularly in developing countries. The FAO's 2024 update on agricultural production statistics highlights these figures and provides a comprehensive overview of global crop production trends [2]. However, in recent decades, interest in this crop has increased significantly in other parts of the world, including Kazakhstan, where it is becoming a strategically important element of agricultural production.

Growing common bean is especially relevant in the southern regions of Kazakhstan, such as the Almaty region, due to the unique nutritional properties of the crop and high export potential. However, successful cultivation of beans in the region requires taking into account specific climatic and agrotechnical features. The Almaty region is characterized by a sharply continental climate with significant temperature fluctuations, hot summers, cold winters, and uneven distribution of precipitation, as stated in Abildaeva's research [3]. These factors create serious challenges for agriculture, necessitating the development of varieties that are resistant to extreme temperatures, droughts and other abiotic stresses.

In modern conditions, one of the priority tasks of breeding is the creation of common bean varieties with a set of economically valuable traits, including high yield, resistance to diseases and pests, suitability for mechanized harvesting and excellent grain quality [4]. In the conditions of intensive agricultural production in the Almaty region, mechanization of cultivation and harvesting processes is of particular importance, which directly affects economic efficiency. In this context, international experience is especially valuable: in Latin American countries such as Mexico and Brazil, beans varieties with a compact bush form and optimal pod attachment height (18-20 cm) have been successfully introduced, which significantly reduced losses during mechanized harvesting [5]. These developments are of particular interest for adaptation to the conditions of the Almaty region, where the efficiency of agricultural production largely depends on the technology of cultivation and harvesting.

Another key focus of breeding work is improving the quality of seeds, determined by such parameters as size, shape, color, protein content, and resistance to damage during harvesting and storage. High grain quality not only increases the market value of beans crops or bean seed, but also enhances their competitiveness in the international market, which makes the assessment of these indicators the most important stage of breeding work [6].

To successfully assess economically valuable traits in common bean, a genetically diverse collection including varieties, lines, and wild species is essential. This diversity ensures adaptability and resilience when breeding for yield, stress tolerance, and seed quality.

During large-scale research in the Almaty region from 2014 to 2024, more than 300 samples of beans of various origins were studied, of which 100 most promising were included in the collection. A comprehensive assessment was carried out for all key economically valuable traits: early maturity, productivity, disease resistance, suitability for mechanized harvesting, seed size and protein content. The work used genetic resources, including varieties, lines and wild species obtained through cooperation with international breeding centers and gene banks. The main areas of breeding work include increasing yields, resistance to pathogens and pests, as well as optimizing characteristics for mechanized harvesting.

The agronomic and physiological traits selected for this study such as plant height, pod attachment height, number of branches, productive nodes, seed weight, 1000-seed weight, seeds per pod, vegetation period, and protein content—play pivotal roles in beans crop performance. Plant height and pod attachment directly influence the efficiency of mechanized harvesting and resistance to lodging. Yield components like pod number and seed weight are primary determinants of productivity, while vegetation period governs the adaptability of varieties to the length of the growing season in diverse climatic zones. Protein content is crucial for seed nutritional quality, affecting both market value and breeding priorities aimed at improving food security and nutritional outcomes. Understanding these traits in combination allows for the development of cultivars that are both productive and regionally adapted.

Overall, beans are rightfully considered a strategically important crop for Kazakhstan, especially for its southern regions. Its cultivation makes a significant contribution to ensuring food security and increasing the economic efficiency of agriculture. At the same time, the collection of beans genetic resources plays a vital role in creating highly productive and resistant to abiotic stress which contributes to the development of competitive agricultural production and strengthening Kazakhstan's position in the international market.

Methodology and Conditions of the Research

Field Research

DOI: 10.3844/ojbsci.2026.26.01.011

The field experiments were established and conducted according to the methodology developed by [7]. This methodology was adapted to the specific agro-environmental conditions of the Almaty region in southeastern Kazakhstan, which is characterized by a sharply continental climate with hot, dry summers, cold winters, and irregular precipitation patterns. The experimental site is located at an altitude of 700–850 meters above sea level on light to medium loamy sierozem soils, which are low in organic matter and nitrogen but moderately supplied with phosphorus and potassium. The soil pH ranged from 7.6 to 8.1, indicating slightly alkaline conditions. These characteristics influenced the choice of fertilization and irrigation regimes.

A randomized complete block design with three replications was used to control for spatial variability in the field. The field layout and trial management were specifically adapted to suit the region's soil and climate conditions. Sowing dates were selected based on long-term meteorological data and adjusted annually to match spring temperature and moisture conditions. Irrigation was applied using furrow methods according to local water availability and key crop development stages. Weed and pest control measures followed regional agricultural recommendations and observed outbreak data.

Phenological observations of beans plant growth and development were carried out in accordance with the methodological recommendations of the All-Russian Institute of Plant Growing (N. I. Vavilov All-Russian Institute of Plant Genetic Resources-VIR), as edited by V.I. Bulanov [8]. These adaptations ensured that the experimental conditions closely reflected real-world agricultural practices in southern Kazakhstan, allowing for the reliable identification of genotypes suited for local production systems.

Evaluation of Resistance to Biotic Factors

The registration of plant damage by diseases and pests was performed using standard phytopathological methods; in accordance with the recommendations of the State Commission for Variety Testing of Agricultural Crops of the Republic of Kazakhstan; using the methodological developments of the International Center for Agricultural Research in the Dry Areas (ICARDA).

Determination of Drought Resistance

Drought resistance scoring was performed during flowering and pod-setting stages, correlating osmotic potential measurements with visible wilting symptoms under field drought conditions.

The assessment of drought resistance was carried out using the method of Oleynikova T.V., Kozhushko N.N. and Osipova Yu.F. [9], based on the correlation between the osmotic potential of cell sap and the resistance of plants to water deficiency.

Hybridization Technique

Manual hybridization was carried out in the flowering phase, using emasculation and tagging procedures adapted from [10], modified for Phaseolus floral morphology.

Hybridization methods were carried out according to the methodology developed by V.F. Dorofeev, Yu.P. Laptev and N.M. Chekalin [10].

Structural Analysis of the Crop

This included the assessment of the following: morphometric indicators (stem length, number of internodes); productivity elements (number of beans and seeds); quantitative characteristics of yield (weight of seeds per plant, weight of 1000 seeds).

Laboratory Studies of Grain Quality

The following standard methods were used: determination of crude protein content: GOST 13496.4; analysis of starch content: GOST 10845; assessment of amylose content - standard biochemical methods; determination of physical and technological properties: GOST 6201, GOST 51074, GOST 8756.

Bean`s Data Analysis

Statistical analyses were conducted using the R software environment (version X.X). Pearson's correlation coefficients were calculated to assess relationships among agronomic and biochemical traits, with significance levels set at $p < 0.05$ and $p < 0.01$. Cluster analysis was performed using hierarchical clustering based on Euclidean distances and k-means clustering, with the optimal number of clusters determined by the Elbow Method. Principal Component Analysis (PCA) was applied to visualize trait variation and cluster differentiation. Regression analyses were used to validate strong correlations between key

DOI: 10.3844/ojbsci.2026.26.01.011

traits. All assumptions for parametric tests, including normality and homoscedasticity, were checked prior to analysis. Relevant R packages such as factoextra, cluster, and corrplot were utilized for data visualization and statistical computations to ensure reproducibility.

To perform the beans data analysis, we followed a structured, multi-step procedure using the R programming language:

Step 1: Loading Data and Preparation

```
# Loading required libraries
library(dplyr)
library(ggplot2)
library(cluster)
library(factoextra)
library(corrplot)
# Loading data
beans <- read.csv("beans.csv")
# View data structure
str(beans)
summary(beans)
```

Step 2: Cluster Analysis

For cluster analysis we will use the k-means clustering method. The following parameters will be used as clustering variables: Height, Lower_beans, Branches, Nodes, beans, Weight, Mass_1000, Seeds_in_beans, Vegetation, Protein.

```
# Selecting variables for clustering
beans_cluster <- beans %>% select(Height, Lower_beans, Branches, Nodes, beans, Weight, Mass_1000,
Seeds_in_beans, Vegetation, Protein)
# Data normalization
beans_scaled <- scale(beans_cluster)
# Determining the optimal number of clusters using the "elbow" method
fviz_nbclust(beans_scaled, kmeans, method = "wss")
# Performing cluster analysis with 3 clusters
set.seed(123)
kmeans_result <- kmeans(beans_scaled, centers = 3, nstart = 25)
# Append clustering results to the original dataframe
beans $Cluster <- as.factor(kmeans_result$cluster)
# Visualization of clusters
fviz_cluster(kmeans_result, data = beans_scaled)
```

Step 3: Separation of Varieties by Vegetation Period

To divide varieties into early-ripening, mid-ripening and late-ripening, we can use quantiles of the vegetation period (Vegetation).

```
# Division into early, mid-season and late ripening
beans <- beans %>% mutate(Vegetation_Group = case_when(
Vegetation < quantile(Vegetation, 0.33) ~ "Early ripening",
Vegetation >= quantile(Vegetation, 0.33) & Vegetation < quantile(Vegetation, 0.66) ~ "Mid-season", Vegetation >=
quantile(Vegetation, 0.66) ~ "Late-ripening"))
# View distribution
table(beans$Vegetation_Group)
```

Step 4: Correlation Analysis

For correlation analysis, we examine the relationships between crop structure, growing season and protein content.

```
# Selecting variables for correlation analysis
cor_vars <- beans %>% select(Height, Lower_beans, Branches, Nodes, beans, Weight, Mass_1000, Seeds_in_beans,
Vegetation, Protein)
# Calculation of the correlation matrix
cor_matrix <- cor(cor_vars)
# Visualization of the correlation matrix
corrplot(cor_matrix, method = "circle")
```

Step 5: Effect of Productivity on Protein Content

Linear regression can be used to analyze the effect of productivity on protein content.

DOI: 10.3844/ojbsci.2026.26.01.011

```
# Linear regression to analyze the effect of productivity on protein content
model <- lm(Protein ~ Weight + beans + Mass_1000, data = beans)
summary(model)
# Visualization of results
ggplot(beans, aes(x = Weight, y = Protein)) +
 geom_point() + geom_smooth(method = "lm", col = "red") +
 labs(title = "Effect of Productivity on Protein Content", x = "Weight", y = "Protein Content")
```

Materials

This research was conducted on a collection of 300 samples of beans of various ecological and geographical origins. Based on a preliminary assessment, 100 of the most promising samples were selected for in depth study.

Weather conditions. In the conditions of the Almaty region, emphasis is placed on to the development of adaptive technologies for growing beans, which includes selection for resistance to drought and high temperatures, optimization of sowing dates, the introduction of resource-saving technologies and the use of modern methods for monitoring the condition of crops.

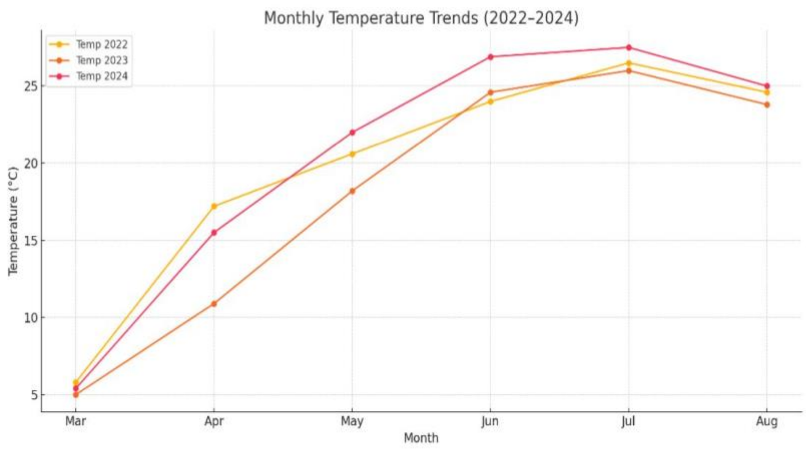


Fig. 1: Monthly temperature trends from March to August (2022–2024) at KazRIAPG experimental site, showing interannual variation in spring and summer heat stress conditions

In 2022, the common bean growing season at KazRIAPG LLP took place in favourable meteorological conditions. Spring came early, with abundant precipitation in the post-emergence period. In March, the average monthly temperature was 5.8 °C—5.1 °C above the long-term average while precipitation totaled 168.6 mm, more than three times the usual amount. In April, the temperature in the first ten-day period was 17.2 °C, there was little precipitation (6.5 mm), which made it possible to sow beans.The second and third ten-day periods of April were also warm, but there was little precipitation (10.9–12.2 mm). May was characterized by higher temperatures (2.6 °C above normal) and abundant precipitation (145.4 mm, which is 2.3 times higher than normal). In June, the temperature was moderate, there was little precipitation, but this did not affect vegetation. In July, the temperature rose to 26.5 °C, and there was only 15.1 mm of precipitation (56.7% of the norm). Favourable conditions allowed for harvesting common bean in the first and second ten-day periods of July, and harvesting continued until August with an average temperature of 24.6 °C and minimal precipitation (7.3 mm) (Figure 1,2).

In 2023, the meteorological conditions at KazRIAPG LLP were characterized by a cold spring and a dry summer. In the first and second ten-day periods of April, the air temperature was 9.9 °C and 10.9 °C, respectively, the amount of precipitation was 64.3 mm, which is 3.9 mm less than the norm. Cold conditions delayed the seeding emergence and increased the duration of the post-emergence phases. In the third ten-day period of April and the first ten-day period of May, the temperature was lower (14.9 °C and 13.0 °C, respectively), and there was very little precipitation (3.9 mm and 4.2 mm), which slowed down the growth of common bean. In the second and third ten-day periods of May, the temperature increased to 20.4 °C and 18.2 °C, but only 39.2 mm of precipitation fell (49.5% of the norm). In June, there was an abnormally high temperature (24.6 °C) and minimal precipitation (4.3 mm), and in July, the temperature remained high, precipitation was 33.6 mm, which is 61.1 mm less than the norm (Figures 1, 2).

In 2024, meteorological conditions at KazRIAPG LLP were characterized by elevated temperatures and uneven precipitation distribution. In March, the average monthly temperature was 5.4 °C (9.8 °C above normal), precipitation was twice as high as normal, which contributed to the early resumption of vegetation. In June, the temperature increased by 3.3 °C compared to normal, reaching 26.9 °C, with low precipitation. However, soil moisture reserves were preserved due to rains in May (Figures 1-2).

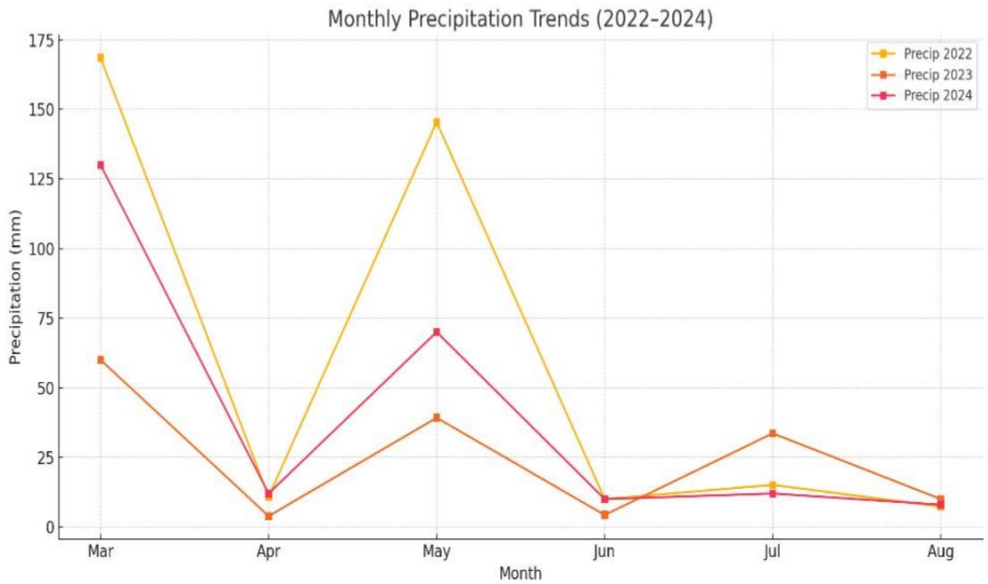


Fig. 2: Monthly precipitation patterns during the growing season (March–August) from 2022 to 2024 (Illustration of the seasonal rainfall variability impacting bean development and drought susceptibility)

Thus, meteorological conditions in 2022–2024 significantly affected the growth and development of common bean. In 2022, favorable conditions contributed to successful vegetation and harvesting, while the cold spring and dry summer of 2023 created challenges for crop growth. In 2024, rising temperatures and uneven precipitation patterns required adjustments in agricultural practices, such as shifting sowing dates and increasing the use of drought-resistant crop varieties to cope with changing conditions (Figure 2). The conducted research and analysis of meteorological data made it possible to identify key factors affecting beans productivity and develop recommendations for increasing its resilience to changing climatic conditions. The integration of modern technologies and breeding methods will contribute to the further development of sustainable agriculture in the context of global climate change.

Although meteorological conditions from 2022 to 2024 varied significantly—especially in terms of temperature extremes and precipitation anomalies—these environmental fluctuations were not statistically incorporated into the trait correlation or clustering analyses in the current study. The aim was to observe phenotypic responses across variable years under natural field conditions. However, we acknowledge that a more detailed multivariate approach (e.g., two-way ANOVA, mixed models, or genotype × environment interaction analysis) could strengthen conclusions about trait stability and environmental influence. Future work will involve such models to quantify the contributions of climate variation to trait performance and clustering patterns.

Results and Discussion

To achieve the research objective of identifying promising common bean genotypes adapted to irrigation conditions in Kazakhstan, we conducted a series of field, laboratory, and statistical analyses. These included trait evaluation, clustering, and correlation analysis to classify accessions and identify relationships between yield components and seed quality. The following sections present key findings grouped by trait categories and analytical methods. Developing varieties for each agroclimatic zone is an important task, since it is difficult to combine high productivity potential and ecological flexibility in one genotype in breeding. Using varieties created in other zones does not always lead to success, since varieties that have demonstrated high productivity in one zone often prove ineffective in others. Therefore, each agro-climatic zone requires the selection of crop varieties adapted to local soil and climate conditions and resistant to common pests and diseases, such as the beans weevil (Acanthoscelides obtectus), fusarium, rust, and bacterial diseases. From the 100 varieties of the collection common bean nursery that we studied, we selected samples that were at the standard level or exceeded the main indicators based on the main economically valuable characteristics (earliness, productivity, suitability for mechanical harvesting, disease resistance, seed size and high protein content).

Research on the accumulation, study and conservation of common bean genetic resources is actively carried out in the countries of Europe, as well as North and South America [11]. The main breeding centers engaged in the creation of new beans varieties are located in these countries. Thanks to cooperation with these centers and the exchange of genetic resources with various gene banks, from 2014 to 2024 we were able to study more than 300 common bean samples of various origins in the conditions of field experiments in Kazakhstan. More than 100 of them were included in our collection. The main criterion for including samples in the collection is their high seed productivity, which allows us to maintain samples in the collection and create seed reserves for further exchange or transfer to interested researchers.

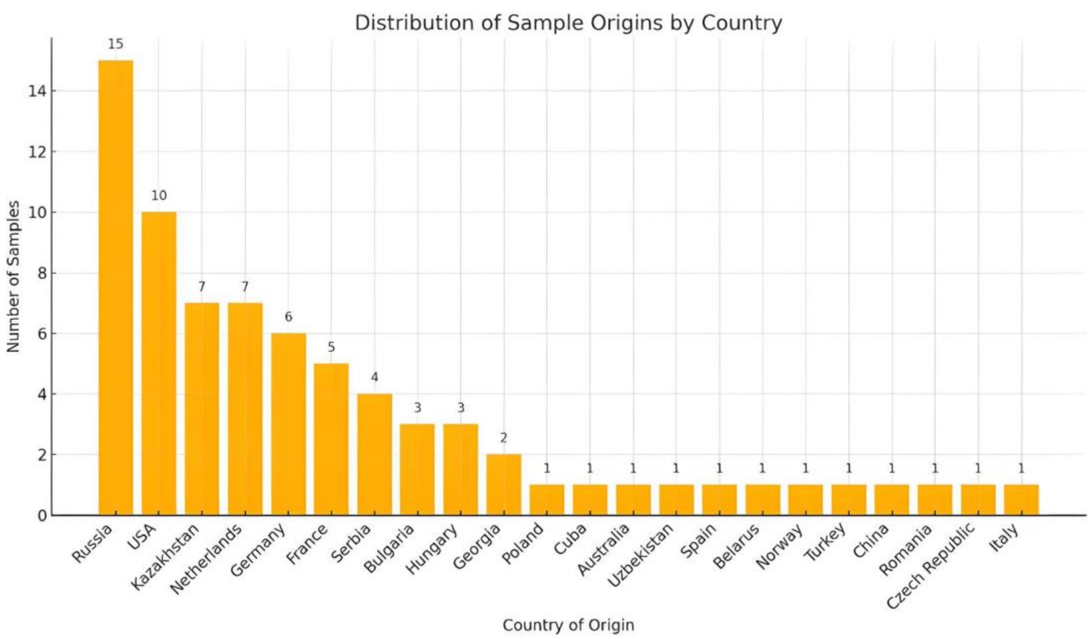


Fig. 3: Distribution of varieties by country of origin

From Figure 3 it is evident that the samples come from many countries, which are presented in the list, we can highlight several popular sources, such as the USA, Germany, the Netherlands and Russia. There are also samples originating from less common countries, such as Georgia, Turkey, Norway, Italy, and Serbia (Figure 3). This indicates that our institute maintains close ties with the All-Russian Institute of Plant Genetic Resources named after N.I. Vavilov (VIR), the All-Russian Research Institute of Legumes and Cereal Crops (Orel), the All-Russian Research Institute of Lupine Selection (Bryansk), the Institute of Field Crops and Vegetable Growing (Serbia), the Research Institute of Genetic Resources and Plants (Uzbekistan), and the Central Research Institute for Field Crops (Turkey).

The evaluation of collection samples was carried out in field conditions using the method [11], which includes the study of a complex of morphological, biological and economic characteristics. More than 100 samples, which are of the greatest value for the creation of new varieties due to the presence of a whole set of useful features, were characterized by us in terms of adaptive capacity and environmental stability, using the above-mentioned method.

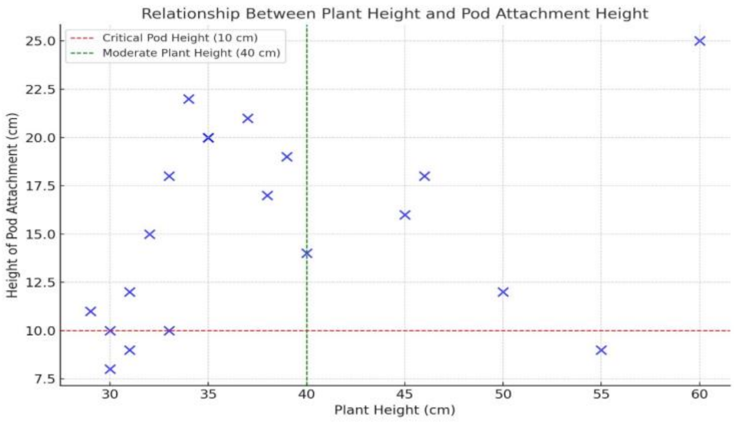
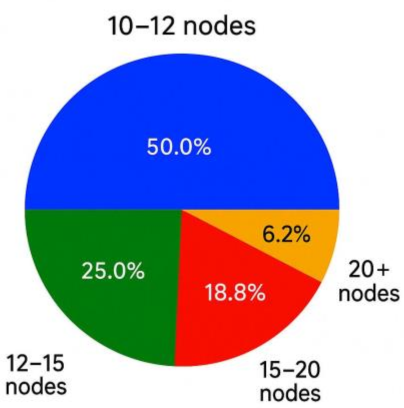


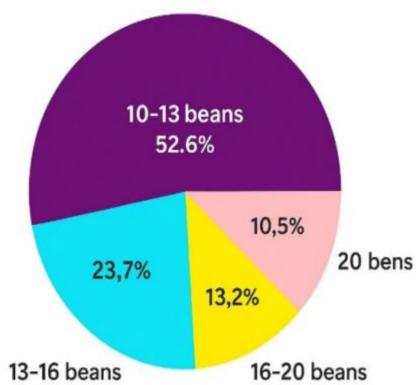
Fig. 4: Dependence of the height of the lower pod (Phaseolus vulgaris L.) attachment on the height of the plant

As shown in Figure 4, the studied cultivar samples differed in plant height, which ranged from 22.7 to 108.3 cm depending on the year of research (Figure 4). Based on the data obtained, 29 samples were included in the short-stemmed group (less than 30 cm), 33 samples in the medium-stemmed group (from 35 to 60 cm), and 10 samples in the long-stemmed group (from 61 to 108.3 cm). Medium-stemmed samples are the most suitable for developing an optimal model of a new generation variety.

"Lower pod attachment height" is a key technological feature that helps reduce common bean seed losses during harvesting using direct combining. This indicator depends on the total number of internodes on the plant and their distance, especially between the 4th and 5th internodes, since the first productive nodes are formed at these levels. According to the data obtained in Figure 4, the value of this feature varied from 9.6 to 29.8 cm. The group with low lower pod (Phaseolus vulgaris L.) attachment (up to 15 cm) included 35 samples, and with high attachment (more than 15 cm) - 65 samples. Pairs with the highest possible values of this feature in both parents are selected for hybridization. Sources highlighted: 185 (k-14990, Bulgaria), FSO-342 Cabretta (k-15352, USA), Zagadka (k-15256, Russia), 226 (k-12962, France), 12700 (Norway), 216 (77-10, Spain), 196 (Hungary), 194 (k-15281, Poland), 214 (Holberg, USA), Empress (k-15208, USA), Golden Sword (Netherlands), 224 (k-12962, France), 15189 (Hungary), 264 (Belarus), 233 (k-13748, Germany), 1071 (France), 153 (USA), etc.



A



B

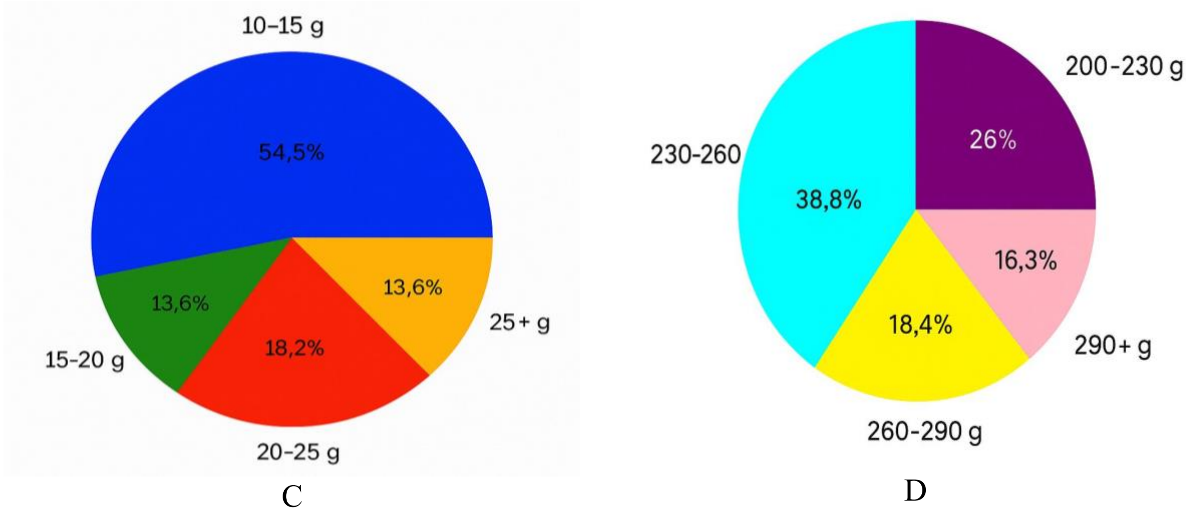


Fig. 5: Distribution of productivity parameters among beans varieties in the collection (A- Number of nodes per plant; B- Number of beans per pod; C- Seed weight per plant; D- Mass of 1000 seeds)

The productivity of grain common bean is a complex feature that depends on the constituent elements of the crop structure and, as shown in Figure 5, plays a key role in creating a new morphotype and assessing the prospects for mass cultivation of future varieties. On average, significant fluctuations in all of its parameters were recorded over the years of studying the collection material (Figure 5).

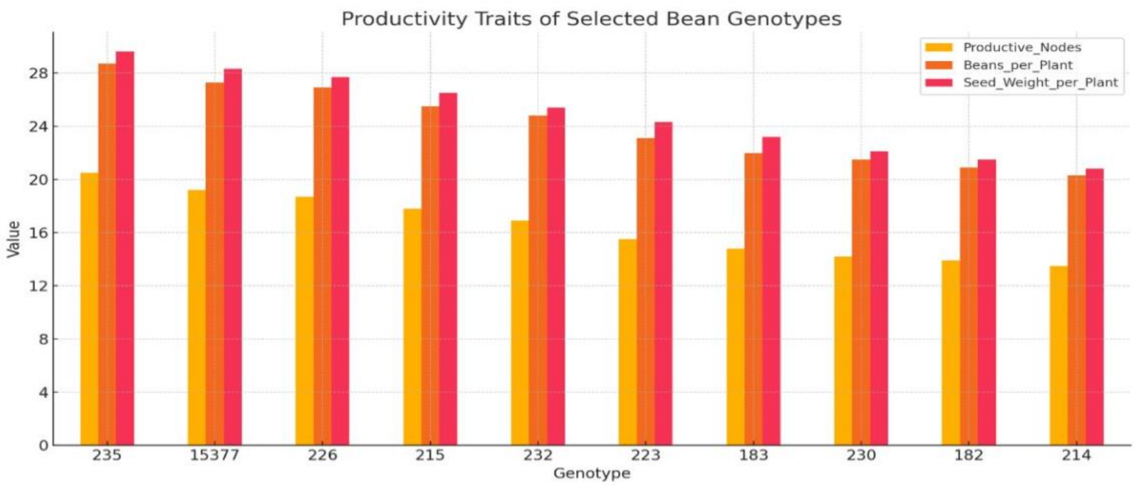


Fig. 6: Comparison of key productivity traits (productive nodes, beans per plant, and seed weight per plant) in selected high-performing bean genotypes grown under irrigation conditions in southeastern Kazakhstan.

As illustrated in Figure 6, accessions such as 235, 15377, and 226 showed consistently high values across all three productivity indicators, making them promising candidates for breeding programs targeting yield optimization (Figure 6).

Moreover, the following samples were selected: with a large number of productive nodes (10.1-21.1 pcs), with a high number of common bean per plant (10.2-28.7 pcs), with a high seed weight per plant (10-29.6 g): 235 (Uzbekistan), 15377 (USA), 226 (k-12962, France), 215 (Russia), 232 (k-13684, USA), 223 (Russia), 183 (k-14689, Romania), 230 (k-13531, USA), 182 (Gallatin-50, USA), Lathyrus-1 (Serbia), 214 (USA), 220 (Cuba), 15189 (Hungary), 228 (Hungary), Brilliant (k-15359, Germany), Fabiolo nano balong (Italy), 15185 (Georgia), 154 (USA), 222 (Germany), 217 (France), 15245 (Netherlands), FSO-348 Gom (Netherlands), FSO-338 Judia (Netherlands), 15134 (Turkey), 231 (k-13612, USA), Empress (k-15208, USA).

DOI: 10.3844/ojbsci.2026.26.01.011

Fig. 7: Samples of beans with different seed shapes: A-15134 HOROZ (Turkey), kidney-shaped mottled seeds; B-15185 (Georgia), elongated white seeds; C-2145 (Georgia), oval red seeds

The 1000-seed weight is a crucial economic trait, exhibiting relatively low variability. The 1000-seed weight varied from 128 to 524 g depending on the variety. The variety samples with a very low value of the trait "weight of 1000 seeds" (up to 200 g) were identified - 24 samples, low (200-250 g) - 29 samples, average (251-350 g) - 36 samples and high (> 350 g) - 14 samples.

For breeding work on creating varieties with a low "1000–seed weight" trait, that is, for small-seeded varieties, the following varieties can serve as a source: 220, FSO-348 Gom, 183, 164, FSO-338 Judia, 210, 170, 213, 172, 194, 237, 197, 233, 230, 1106, 211, 217, Strela, 15378, 176, and for large-seeded varieties – 15483, 15185, 2-vad, 165, 15134, 153, 264, 08-543, 167, 236, 15522, 160, Nina Megazin, 1986.

As you can see from Figure 7, the collection shows diversity not only in seed weight, but also in seed shape. Of great value for creating grain-type varieties are the round (Figure 7, A) and kidney-shaped (Figure 7, B, C) seed shapes, which is dictated by both consumer preferences and requirements for seed harvesting technology (large, long, narrow seeds are subject to a high risk of damage, such as crushing during threshing). Round and kidney-shaped seeds are preferred not only for their visual and market appeal, but also for their superior processing qualities. These shapes are less prone to cracking and mechanical damage during threshing and post-harvest handling, which is especially important in mechanized harvesting systems. Varieties with elongated or irregular seed forms tend to suffer higher breakage rates during mechanical processing, reducing seed quality and market value. Therefore, selection for favorable seed shape directly supports the development of cultivars that are compatible with large-scale production and modern processing infrastructure.

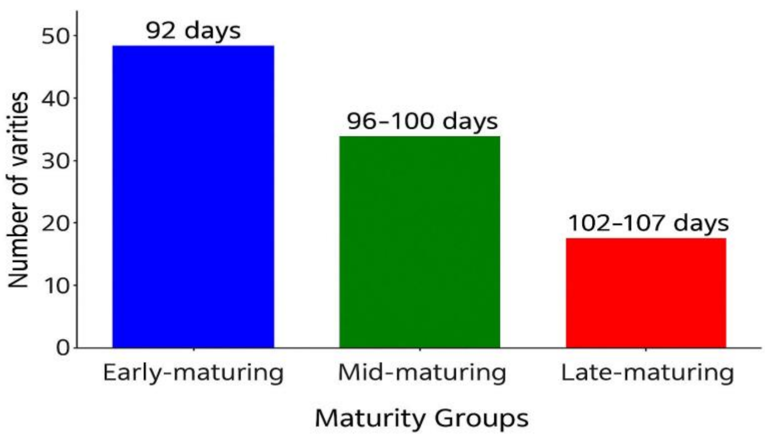


Fig. 8: Distribution of bean varieties by growing season

As can be seen in Figure 8, as a result of studying the duration of the full vegetation period (according to average long-term data), The samples were classified into three groups based on maturation timing: early, medium, and late maturation (such a qualitative division of samples into groups is not provided for by the universal descriptors of common bean used in the work) (Figure 8). More than half of the studied material (about 53%) was classified as early-maturation samples, the group of late-maturation ones amounted to about 12%.

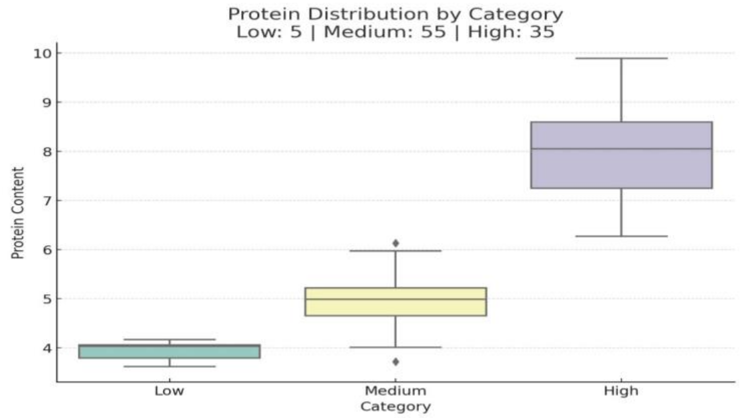


Fig. 9: Protein level indicator

Common bean is widely distributed due to the high protein content in their seeds (20–35% depending on the variety), which is an important indicator of their nutritional value. The protein level in common bean seeds is mainly determined by the genotype of the plants, to a lesser extent by agronomic, soil, climate and other factors, and the influence of the place of reproduction on this indicator is insignificant. Data analysis shows in Figure 9, that the largest number of samples (about 55) belong to the category with medium protein content, with an average protein content of 24.3%, indicating the predominance of this group among the studied legumes. The high protein category includes about 35 samples, while the low protein category includes the fewest samples (Figure 9).

To better classify the genetic diversity among common bean accessions and identify promising groups for breeding, we conducted hierarchical cluster analysis based on key morpho-agronomic and protein-related traits. This analytical approach supports our research objective of grouping accessions according to their performance under irrigation conditions in southeastern Kazakhstan. Cluster analysis enables the organization of accessions with similar or contrasting characteristics, which is particularly useful for identifying genetically diverse and complementary parent pairs for targeted crossing in breeding programs.

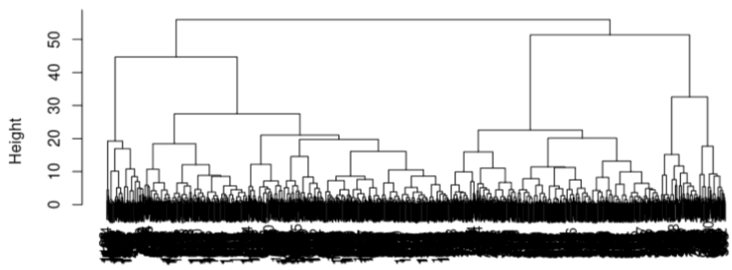


Fig. 10: Dendrogram of hierarchical clustering of beans varieties

The graph in Figure 10 shows a dendrogram of hierarchical clustering, as it illustrates the grouping of samples based on their similarity across key morphological traits, which illustrates the process of grouping common bean varieties. The Height axis displays the distance or dissimilarity between objects (common bean varieties) or clusters in the process of hierarchical clustering (Figure 10).

The dendrogram also illustrates how common bean varieties group into distinct clusters based on trait similarities. This visual hierarchy confirms significant genetic and phenotypic differences across accessions. Such structure is essential for identifying breeding pools with contrasting and complementary traits (Figure 10).

To determine the optimal number of clusters, the Elbow Method was applied by plotting the Within-cluster Sum of Squares (WSS) for k values ranging from 1 to 10. As shown in Figure 11, the curve displays a noticeable decline in WSS up to k = 3, after which the rate of improvement diminishes. This “elbow” point suggests that three clusters offer the best balance between compactness and interpretability.

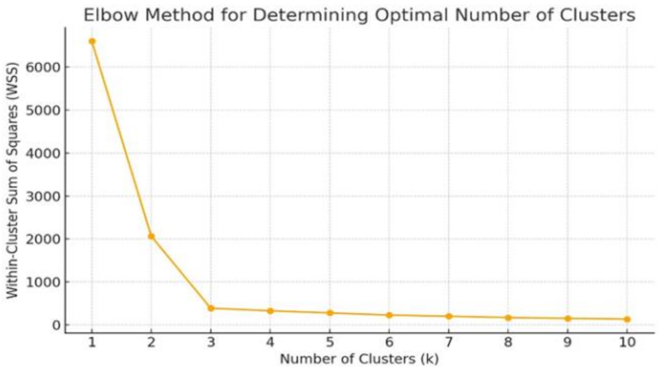


Fig. 11: Elbow plot showing the within-cluster sum of squares (WSS) for k = 1 to 10 clusters. The inflection point at k = 3 indicates the optimal number of clusters for this dataset

The primary objective of the cluster analysis was to group common bean accessions based on the similarity of key morpho-agronomic and physiological traits. The variables used for clustering included plant height, number of lower pods, number of branches, productive nodes, number of pods per plant, seed weight per plant, 1000-seed weight, number of seeds per pod, vegetation period, and protein content. These traits were selected due to their agronomic importance and relevance to yield performance and seed quality. As shown in Figure 11, the Elbow plot indicates an inflection point at k = 3, suggesting that three clusters represent the optimal grouping structure for this dataset (Figure 11).

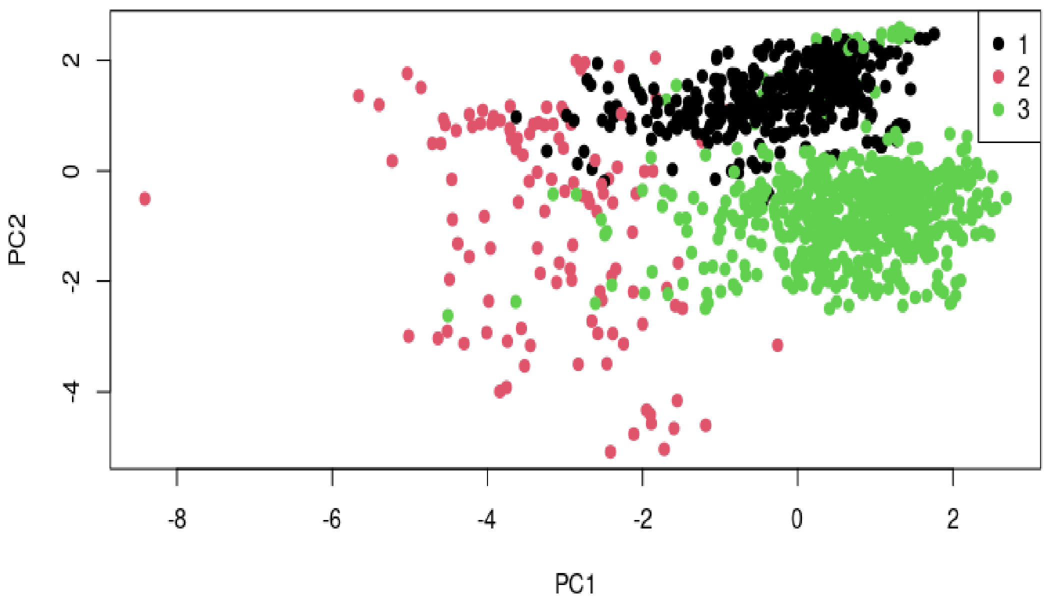


Fig. 12: Hierarchical clustering (PCA)

Based on this result, k-means clustering was performed using three clusters. To further assess clustering quality and visualize group separation in multivariate trait space, Principal Component Analysis (PCA) was conducted. In the resulting plot in Figure 12, the clusters are represented by distinct colors, clearly illustrating their separation and confirming the validity of the classification (Figure 12).

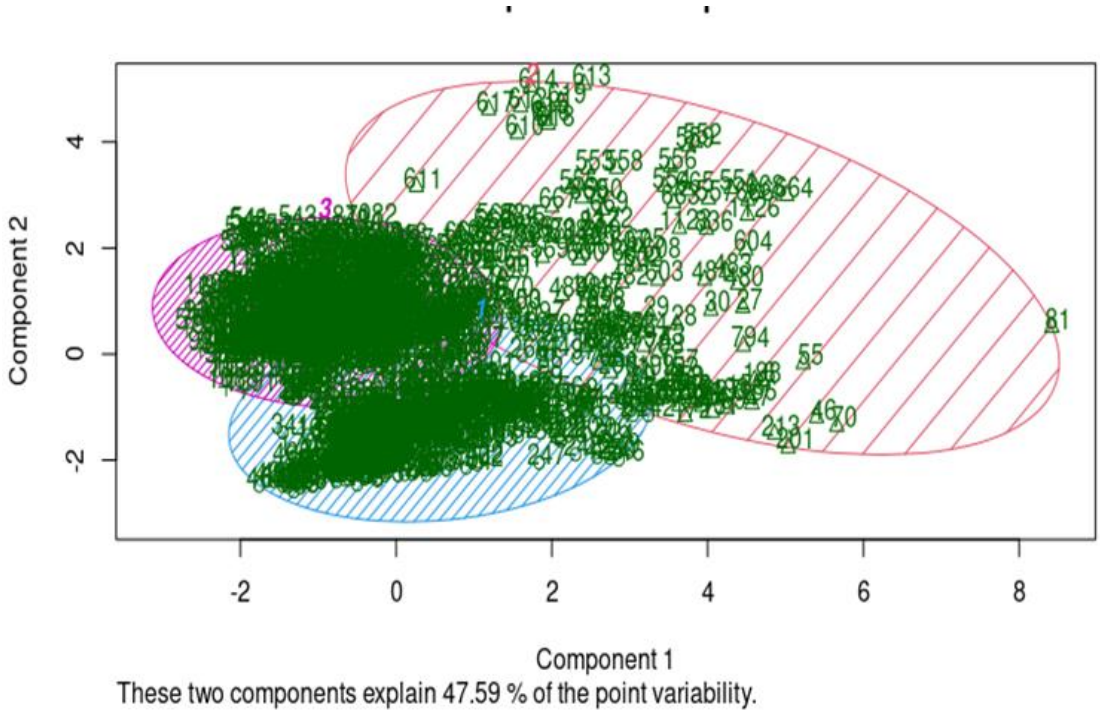


Fig. 13: Cluster analysis of beans varieties

To visualize these clusters in trait space and better understand their internal variation, we conducted a Principal Component Analysis (PCA). This approach revealed the relationships between genotypes and the traits driving differentiation. Figure 13 shows the result of cluster analysis of common bean accessions, where two principal components are used to visualize the data. The X-axis (Component 1) and Y-axis (Component 2) reflect the principal components, which explain 47.59% of the data variability, indicating a significant proportion of the explained variance (Figure 13).

The graph shows several individual point clouds, each representing a cluster of common bean varieties with similar characteristics. The points, indicated by green symbols, are grouped into several dense areas, indicating homogeneity of varieties within each cluster. These clusters likely differ in characteristics such as size, texture, or other physicochemical properties (Figure 13).

The areas in the graph marked with different colours and lines (red, blue and purple) represent confidence areas for each cluster, allowing us to assess the concentration of accessions in each group. These differences reflect the underlying genetic and phenotypic variation and support their use in developing diversified breeding lines. As a result of the analysis, it was decided to use 3 clusters. The clear separation of clusters in PCA confirms the reliability of the classification and highlights potential candidate groups for targeted breeding, especially high-yielding, protein-rich types suited for Kazakhstan's irrigated environments (Figure 13).

Cluster #1 – High productivity group includes 27 samples: high protein level from 25-30.0% (179, 158, 160, 235, 217, 222, 167), a large number of beans from 15.4-28.7 pieces (226, 225, 164, 230, 223, 182, Strela, Lathyrus-1, 171, 214, 220) and seeds in beans from 6-8 pcs (Fabiolo nano balong, 213, 237, 182, 15190, Empress, Nina Megazin, 230, 15377, 223, 228, 220, 172), high seed weight per plant from 15-41.5 g (15190, Clendick, Empress, 15522, Belko, 1986, 160, Zagadka, 08-543, SZ (Slavonski zhutozeleni), Nina Megazin, Lathyrus-1) and 1000 seed weight from 402-524 g (264, 08-543, 167, 236, 15522, 160, Nina Megazin, 1986).

Cluster #2 – Low productivity group: low protein level from 15-19.0% (214, 183, 1106, 15378, 171, 164, 209, 15134, 15185), fewer beans from 4-7 pieces (15571, Zolotaya Shpaga, Shokoladnitsa, 1280, 1279, 15529, 189, 160, 2145, 178, 2-vad, 11758, 11633, 1986, 1202, 15522, 10709, 1214, 210, 196, 208, Clendick, 1106, 185) and seeds in beans 3 pcs (165, Lathyrus-1, 208, 12170, 167, 197, 15366, 168), low seed weight per plant from 3-10 g (208, 197, 210, 1106, 233, 250, 194, 222, FSO-348 Gom, 164, 170, 224, 228, 220, 211, 1071, 190, 232, 214, 191, 12700, FSO-338 Judia, 231, 189, 216, 10709, 172, 165, 12170, 15366, 183, 15245, 209, 217, 235, 225, 179, 15260, 8400, 154, 176, 15480, 15378, 168, 15185, 152, 15529,

DOI: 10.3844/ojbsci.2026.26.01.011

158, Brilliant, FSO-342 Cabretta, 178, 196, 185, 11789, 173, 188, 153, 15189, 15452) and low 1000 seed weight from 128-200 g: (220, FSO-348 Gom, 164, 183, FSO-338 Judia, 210, 170, 213, 172, 194, 237, 197, 233, 230, 1106, Strela, 211, 217, 15378, 176, 228, 225, 15245, 15529).

Cluster #3 – Average productivity group: average protein level from 19.1-24.8% (225, 1986, 176, Brilliant, 182, 188, 15480, 194, 196, FSO-348 Gom, 213, 11633, Strela, 224, FSO-342 Cabretta, 185, 154, 208, 1279, Lathyrus-1, 12170, 210, 211, 189, 226, St.Inzhu, 216, Empress, 11789, 15260, 1280, Zolotaya Shpaga, 15522, 8400, 228, 10709, 15571, SZ (Slavonian zhutozeleni), Nina Megazine, 220, 15529, 2145, 173, Clendick, 191, 231, 15377, 232, 168, 15483, 12700, 215, 08-543, 1214, 15190, 152, Fabiolo nano balong, 15189, 15452, 165, FSO-338 Judia, 1202, 11758, 15245, 1071, 223, Zagadka, Shokoladnitsa, 15366, 153, 250, 1088, 197, 233, 230, Belko, 264, 178, 172, 170, 237, 2-vad, 236, 190), average number of common bean from 7.1-14.5 pieces (11789, 191, 233, 15378, Zagadka, 250, 08-543, 179, 158, St. Inzhu, 1071, 15260, 12170, 176, 236, 194, 216, 1088, 8400, Nina Megazin, 173, 15483, 153, 170, 12700, 167, 197, 15480, 15452, 15366, 237, FSO-342 Cabretta, 168, Belko, 190, SZ (Slavonski zhutozeleni), 224) and seeds in common bean from 4-5 pcs (15185, 152, 183, 15245, 215, 211, 171, 15571, Zolotaya shpaga, 1280, 1279, 15529, 2145, 15378, 158, 1071, 236, 1088, 153, 190, 209, Brilliant, 232, 217, 15134, 264, 235, 225, 214, Shokoladnitsa, 160, 11758, 1986, 1202, 15522, 210, 1106, 191, 233, 250, 08-543, 179, 15260, 194, 8400, 15483, 12700, FSO-342 Cabretta), the average weight of seeds per plant is 10-13.9 g (15452, 264, 1214, 11633, 15483, 15571, Zolotaya Shpaga, Strela, St.Inzhu, 1279, 182, 11758, 2145) and the average weight of 1000 seeds is 202-396 g: (190, 12170, 182, Empress, Clendick, 154, Brilliant, 1214, 188, 11633, 231, 189, 226, 222, St.Inzhu, 15190, 224, 171, 152, 235, 232, Shokoladnitsa, 168, Fabiolo nano balong, 1202, 15260, 15366, 15189, FSO-342 Cabretta, 191, 179, 8400, Lathyrus-1, 214, 209, 158, 1280, 2145, 250, 10709, 12700, 15377, 1071, 185, 15571, Belko, 1088, Zagadka, 11789, 15480, 15452, 208, 223, 11758, 215, SZ (Slavonian zhutozeleni), Golden sword, 173, 178, 216, 1279, 196, 15483, 15185, 2-vad, 165, 15134).

Thus, the result of the cluster analysis confirms the presence of several clearly distinguishable groups among the beans varieties, which allows for their effective classification based on the key characteristics identified using the principal components.

While hierarchical and k-means clustering were effective for initial classification, we acknowledge the potential advantages of more advanced methods such as principal component-based clustering, Random Forest classifiers, or unsupervised neural network models. This study was conducted to evaluate the performance of different bean cultivars under varying irrigation regimes in the Almaty region. The findings of this research provide valuable insights for improving bean production in the region.

Correlation analysis between the common bean traits revealed significant relationships presented in the correlation matrix. The exceptionally high correlation ($r = 0.91$) between productive nodes and seed weight per plant was validated through linear regression analysis ($R^2 = 0.83$, $p < 0.01$), consistent with findings by Dipp et al. for Brazilian common bean [5]. The strongest positive correlations are observed between plant height (Height) and position of lower pod (Lower_common bean) ($r = 0.47$), as well as between the number of nodes (Nodes) and the number of common bean) ($r = 0.91$), indicating the interdependence of these morphological characteristics. The mass of 1000 seeds (Mass_1000) demonstrates a moderate positive correlation with other parameters ($r = 0.10$-0.19), with the exception of a more pronounced relationship with weight ($r = 0.52$). Protein content (Protein) shows weak correlations with other traits ($r = 0.05$-0.24), indicating its relative independence from morphological characteristics. Of particular interest is the weak negative correlation between the length of the vegetation period (Vegetation) and other parameters ($r = 0.01$-0.21), which may indicate a compromise between early maturity and productivity. These data are important for breeding, allowing one to predict the joint change of traits and optimize the selection of plants with the desired set of characteristics. The negative correlation observed between seed protein content and yield components, such as seed weight and pod number ($r = -0.24$), aligns with findings from previous international studies. Saenz et al. reported similar trade-offs in soybean breeding programs, where selection for increased protein content often led to reduced yields [12].

Likewise, Jarecki et al. demonstrated that higher seed protein content frequently corresponded to lower productivity in multiple legume species [13]. These studies reinforce that this trade-off is a common challenge in legume breeding and must be addressed when developing new high-performing cultivars.

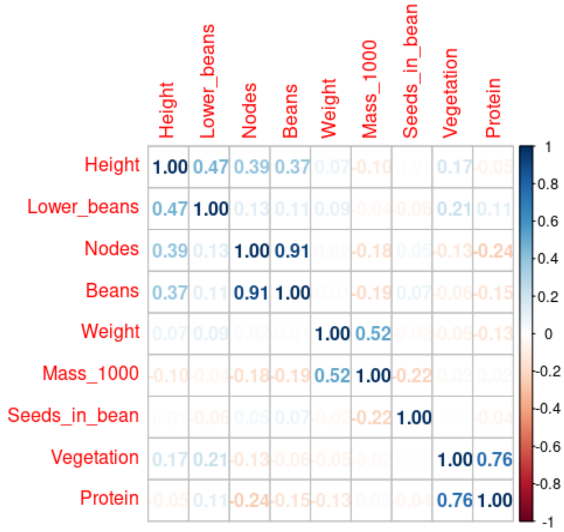


Fig. 14: Correlation matrix of economically valuable parameters of bean varieties

Figure 14 shows the correlation analysis as a correlation matrix. The color scale on the right shows the range of correlation values:

- Blue color – positive correlation (up to 1.0)
- Red color – negative correlation (up to -1.0)
- White or light gray color – no significant correlation (around 0) (Figure 14)

In our studies, the yield depended on the number of productive nodes, common bean per plant and bean weight per plant. All the presented traits closely correlated with this indicator. A direct strong correlation was established between seed yield and the number of productive nodes, the number of common bean per plant and seed weight per plant. According to these traits, the correlation coefficient varied: r = 0.91.

The strong correlation observed between the number of productive nodes and seed weight per plant (r = 0.91) was further validated by linear regression, which yielded an R^2 value of 0.83 ($p < 0.001$), confirming the robustness of this association. This finding is consistent with pod-setting behavior in determinate and semi-determinate genotypes of Phaseolus spp., where a greater number of productive nodes directly contributes to higher seed biomass.

According to these traits, the following were distinguished with a high number of productive nodes (10.1-21.1 pcs): 235, 15377, 226, 215, 232, 223, 183, 230, 182, Strela, 225, 164, 171, Lathyrus-1, 214, 220; with a large number of common bean per plant (10.2-28.7 pcs) – 15189, 228, 209, 188, Brilliant, Fabiolo nano balong, 15185, 154, 152, 232, 183, 165, 222, 217, 15245, FSO-348 Gom, FSO-338 Judia, 172, 15134, 231, Empress, 215, 15377, 213, 264, 211, 235, 15190, 226, 225, 164, 230, 223, 182, Strela, Lathyrus-1, 171, 214, 220.

A moderate positive correlation was found between the 1000-seed weight and the plant weight (r = 0.52), indicating that larger seeds contribute to an increase in the total plant weight. The following collection common bean samples with large seeds weighing more than 350 g were identified during the analysis of the results: 15483, 15185, 2-vad, 165, 15134, 153, 264, 08-543, 167, 236, 15522, 160, Nina Megazin, 1986.

In common bean production technology, harvesting mechanization is still the main problem, since it is largely hampered by the lack of specialized machines. The varieties with high attachment of the lower pod are of greatest interest.

In the common bean varieties, the height of attachment of the lower pod varied from 10.2 to 29.8 cm. In our experiment, the attachment of the lower pod depended on the height of the plant: r = 0.47.

Based on the results of the study of the suitability of common bean varieties for mechanized harvesting, it is recommended to use the following varieties as sources in the selection process: FSO-342 Cabretta, 185, 226, Zagadka, 1279, 1106, 12700, 214, 196, 216, 209, 194, 215, Empress, Zolotaya Shpaga, 224, 167, 158, 15189, 264, 233, 1071, 1280, 153, 15529, 211, 189, 165, 183, 191, 15190, 171, 15522, 160, 15245, 217, 1986, 15260, 236, 168, 15377, 15134, 2145, 250, 11789, 15571, FSO-348 Gom, Strela, 188, 220, 235, 225, FSO-338 Judia, 15378, 190, 12170, 08-543, Shokoladnitsa, 164, 182, 172.

The most significant negative correlations are observed between plant productivity (productive nodes, number of pods per plant, seed mass per plant, number of seeds per pod) and protein content: Nodes and Protein (r = (-0.24) this means that plants with more nodes tend to have lower protein content in seeds. Mass_1000 and Seeds in_common bean (r = (-0.22) indicates that the higher the 1000 seed mass, the fewer seeds per pod. Protein and Mass_1000 (r = (-0.15) the higher the 1000 seed mass, the lower the protein content, larger seeds often have lower protein concentrations as they may contain more starch or fat. This negative correlation may reflect a physiological trade-off, where high-yielding genotypes tend to accumulate more starch or carbohydrate reserves in seeds, leading to a relative dilution of protein concentration. This relationship has also been observed in international studies of legume crops. In our research, a negative correlation between seed protein content and yield-related traits (e.g., productive nodes and seed weight per plant, r = –0.24) was identified. Saenz et al. found similar results in soybean, where high protein levels were associated with yield penalties [12]. Jarecki et al. further confirmed this trade-off across multiple legume species and environmental conditions [13]. These findings highlight the widespread nature of this phenomenon and support the need for balanced breeding strategies that address both yield and nutritional quality.

This correlation analysis clearly shows which parameters have a strong connection with each other. Our goal is to create a model of a new common bean variety that will be characterized by a short growing season, high yield, bush form with high attachment of lower pod from 15 cm, branched with a large number of productive nodes and common bean on the plant, with an average weight of 1000 seeds and a protein content of 25-30%. The obtained analysis data will be used to achieve the set goal and optimize the selection process and agrotechnical methods in growing common bean (Figure 14).

The results of this study have important practical implications not only for beans, but also for the breeding and cultivar selection of legumes in Kazakhstan. By identifying genotypes with favorable combinations of productivity, protein content, and morphological traits, breeders can prioritize accessions for crossing to develop improved cultivars tailored to regional needs [14]. The cluster and correlation analyses facilitate the selection of varieties that perform well under specific agro-climatic conditions, such as drought-prone areas in southern Kazakhstan or cooler northern regions [15]. Moreover, traits associated with higher pod attachment and moderate plant height indicate potential for improved mechanization compatibility, which is essential for scaling up common bean production efficiently. These findings support breeding programs focused on enhancing yield stability, nutritional quality, and regional adaptability.

Conclusion

Research has shown that for efficient beans cultivation in various agroclimatic zones, it is necessary to develop varieties adapted to local conditions. The exchange of genetic resources with international research centers has enriched the collection with valuable varieties with high productivity rates, resistance to diseases and suitability for mechanical harvesting. Varieties such as Empress (USA) and Zagadka (Russia), with high pod attachment and moderate seed weight, are recommended for breeding programs targeting mechanized harvesting. High-protein donors like 235 (Uzbekistan) and 217 (France) should be used in breeding for nutritional enhancement. These specific lines offer practical value for both regional adaptation and commercial production. In addition to identifying promising genotypes, we recommend concrete actions to stakeholders. For breeders, we suggest crossing high-yielding lines with moderate-protein donor genotypes to optimize both productivity and nutritional value. For farmers in southeastern Kazakhstan, varieties such as Empress (USA) and Zagadka (Russia), with moderate plant height and high pod attachment, are particularly well suited to mechanized harvesting and local conditions. For policymakers, we recommend supporting breeding programs that prioritize drought-resilient, protein-rich, and mechanization-compatible varieties to strengthen domestic legume production and export potential.

Cluster analysis of common bean accessions revealed three distinct groups with different productivity based on characteristics such as protein level, common bean quantity, seed weight and other parameters. These groups allow for efficient classification of accessions and identification of high, medium and low productivity accessions. Visualization using the Principal Component Analysis (PCA) confirmed significant differences between clusters, which helps in further analysis and breeding work.

Correlation analysis showed that common bean yield closely depends on the number of productive nodes, beans per plant and seed weight. A strong positive correlation was found between these characteristics. A moderate positive correlation was also found between the weight of 1000 seeds and the total plant weight, which confirms the importance of a large seed for increasing plant weight. Noteworthy is the moderate relationship between the height of attachment of the lower pod and the height of the plant, which is important for mechanized harvesting. Negative correlations were found between productivity and protein content in seeds, a pattern also observed in studies like Dipp et al. [5]. This inverse relationship likely reflects a physiological trade-off where increased carbohydrate allocation to seeds reduces relative protein concentration. The data obtained will help in the selection process and optimization of agricultural technology for creating highly productive common bean varieties.

Acknowledgment

The work was carried out within the framework of Grant financing of the Ministry of Agriculture of the Republic of Kazakhstan under the budget for 2024-2026 (AP23489482).

Funding Information

This work was funded by the Ministry of Agriculture of the Republic of Kazakhstan under the Grant Financing program, budgeted for 2024–2026 (Grant No. AP23489482).

Authors Contributions

Kudaibergenov Mukhtar Sarsenbekovich: Conceptualization, Supervision, Funding acquisition, Project administration.

Bastaubaeva Sholpan Orazovna: Supervision, Investigation, Resources.

Kanatkyzy Makpal: Investigation, Data curation, Writing the original draft, Visualization.

Saken Gaukhar Sakenkyzy: Data curation, Translating into English, review and editing.

Urazaliev Kairat Rahimovich: Data analysis, Cluster analysis.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

1. Gobatto, D. R., Carneiro, E., Woyann, L. G., & others. (2025). Cold stress during beans reproductive stage: Effects in physiological and agronomic parameters. *Tropical Plant Biology, 18*, 38. https://doi.org/10.1007/s12042-025-09407-2
2. Food and Agriculture Organization of the United Nations. (2024). *World food and agriculture – Statistical yearbook 2024*. FAO Statistical Yearbook.
3. Abildaeva, J. B., Kudaibergenov, M. S., Kenenbayev, S. B., Kanatkyzy, M., Saken, G. S., Bekkuly, Y. E., & Batyrbekuly, N. (2025). Development of varietal technology for new varieties of peas under irrigation conditions of South-East Kazakhstan. *SABRAO Journal of Breeding and Genetics, 57*(1), 217–229. https://doi.org/10.54910/sabrao2025.57.1.21
4. Kudaibergenov, M. S., Baitarakova, K. Zh., Saikenova, A. Zh., Kanatkyzy, M., Abdrakhmanov, K. A., & Saken, G. S. (2024). Chickpea genotype selection based on economically valuable traits to develop high-yielding types. *SABRAO Journal of Breeding and Genetics, 56*(1), 119–128. https://doi.org/10.54910/sabrao2024.56.1.11
5. Dipp, C. C., Marchese, J. A., Woyann, L. G., & others. (2017). Drought stress tolerance in beans: What about highly cultivated Brazilian genotypes? *Euphytica, 213*, 102. https://doi.org/10.1007/s10681-017-1893-5
6. Saikenova, A. Z., Kudaibergenov, M. S., Nurgassenov, T. N., Saikenov, B. R., & Didorenko, S. V. (2021). Crop yield and quality of lentil varieties in the conditions of the southeast of Kazakhstan. *OnLine Journal of Biological Sciences, 21*(1), 33–40. https://doi.org/10.3844/ojbsci.2021.33.40
7. Dospekhov, B. A. (1985). *Field experimentation methodology*. Agropromizdat. (in Russian)
8. Vavilov Institute (VIR). (1988). *Diagnostics of plant resistance to stress influences*. (in Russian)
9. Osipov, Y. F., Oleynikova, T. V., & Kozhushko, N. N. (1970). Determination of drought resistance of wheat and barley varieties by seed germination in sucrose solutions with high osmotic pressure. In *Methodical instructions* (pp. 22–39). (in Russian)
10. Dorofeev, V. F., Laptev, Y. P., & Chekalin, N. M. (1990). *Flowering, pollination and hybridization of plants*. Agropromizdat. (in Russian)

11. Vishnyakova, M. A., Buravtseva, T. V., Bulyntsev, S. V., & others. (2010). *Collection of world genetic resources of grain legumes: Replenishment, conservation and study: Methodological guide*. VIR. (in Russian)
12. Saenz, D., Assefa, Y., Woli, P., & Naeve, S. (2023). An evaluation of the soybean (*Phaseolus vulgaris* L.) protein–yield trade-off and its implications for breeding. *Plant Breeding, 142*(1), 88–99. https://doi.org/10.1111/pbr.13062
13. Jarecki, W., Bobrecka-Jamro, D., & Grzebisz, W. (2022). Seed yield and protein content in selected legumes under Polish climate conditions. *Agronomy, 12*(11), 2667. https://doi.org/10.3390/agronomy12112667
14. Sarsenbekovich, K. M., Makpal, K., Sakenkyzy, S. G., Karkyn, N., Bakaevna, A. J., Zhumabaevna, S. A., & Zhumanovna, B. K. (2025). Optimizing pea cultivation in Kazakhstan: Breeding drought-resistant varieties for sustainable agriculture. *OnLine Journal of Biological Sciences, 25*(3), 633–644. https://doi.org/10.3844/ojbsci.2025.633.644
15. Zhumabaevna, S. A., Sarsenbekovich, K. M., Makpal, K., Sakenkyzy, S. G., Zhumanovna, B. K., Bakaevna, A. J. & Karkyn, N. (2025). Optimizing Lentil Breeding Models for South-East Kazakhstan: Yield, Adaptability, and Protein Enhancement. *OnLine Journal of Biological Sciences*, *25*(3), 674-683. https://doi.org/10.3844/ojbsci.2025.674.683