International Journal of Design & Nature and Ecodynamics
Vol. 19, No. 1, February, 2024, pp. 249-257
Journal homepage: http://iieta.org/journals/ijdne

Comparative Analysis of Soybean (*Glycine max* (L.) Merrill) Varieties: Implications for Yield and Quality

Zarina Tleulina*, Gulden Kipshakbayeva, Ingkar Ashirbekova, Bekzak Amantayev, Assemgul Kipshakbayeva

Department of Agriculture and Plant Growing, S. Seifullin Kazakh Agro Technical Research University, Astana 010011, Republic of Kazakhstan

Corresponding Author Email: zarinatleulina42@gmail.com

https://doi.org/10.18280/ijdne.190127

Received: 15 September 2023
Revised: 20 December 2023
Accepted: 22 January 2024
Available online: 29 February 2024

Keywords:
protein, thousand-kernel-weight, cluster analysis, productivity, correlation, soybean varieties, product quality

ABSTRACT

Soybean, an essential crop with diverse industrial uses, benefits from selecting climate-adapted varieties to enhance yields and product quality. The purpose of study was to compare the yield of soybean varieties of different origins to identify the most productive and quality varieties. To fulfil purpose, a field experiment was conducted in 2022, where 130 soybean varieties and lines were investigated. As a result of the study, it was found that the number of beans per plant was 4-38 pieces, the maximum value was found in the following varieties: Suiyang 1, SC Farta, Jinyaan 55, Kenfeng 20, Dongnong 63, Chera 1-1, Progress, lines No. 33, No. 92, No. 77. High yields were found in varieties Ivushka, Suiyang 1, SC Farta, Beidou 43, Beidou 19, Kenfeng 20, Huajiong 2, Kenfeng 6, Chera 1-1, Beidou 47, Progress, lines: No. 78, No. 33, No. 83, No. 16. 78, No. 33, No. 83, No. 16. The protein content of soya seeds ranged from 38.66% to 47.25%. Varieties with high seed protein content: Beidou 26, Kendou 60, Alaska, Osmon, K-0126, K-0128. Quality indicators weakly correlate negatively with soybean yield, influenced by factors like some productive nodes, seed weight per bean, thousand-kernel weight, and other characteristics.

1. INTRODUCTION

It is important to note that due to climate change and the increase in extreme weather events, drylands are becoming increasingly important for agricultural production. The study of crops such as soybeans in such conditions reflects the urgency of their adaptation to change. Soybeans are an important crop for protein and oil production, and they are essential to the global food chain. Research on its cultivation under drought conditions can be of great importance for food security in a changing climate. In some regions of the world, drought has become a widespread problem, and this is particularly true in the southern states of the United States and other arid regions. Research aimed at growing soybeans under similar conditions may have practical applications for local agricultural industries.

Animal products are a traditional source of protein, but they can be limited in availability, economic cost, and environmental sustainability. In the near future, it will be difficult to fully meet the demand for dietary protein from animal products alone. Therefore, plant sources of protein, including soya protein, are becoming increasingly important to address this issue as they can offer sustainable and environmentally friendly alternatives [1]. Research by Achina et al. [2] shows that increased production of high-protein grain legumes, especially soybean, can considerably help to address dietary protein deficiency. Soybeans are of particular importance as their grain contains up to 45% protein and 20-23% fat. Soy protein contains all essential amino acids and is close to proteins of animal origin in terms of biological value. In fact, soya protein is the highest plant protein source [3].

Soybean (*Glycine max* (L.) Merrill) is one of the key crops in the world due to its versatility and wide range of uses. According to Baraskar et al. [4], increasing soybean production is important as it can be a key component of a strategy to increase production of high-protein crops. Furthermore, increased soybean cultivation will not only reduce dietary protein deficiency but improve food safety in many countries around the world. However, soybean cultivation is limited by its high heat requirement and photoperiodic response. Most soybean varieties are short-day plants. In regions with higher latitudes, where day length is longer, there are problems with plant development and maturation before autumn frosts. Global warming, characterized by rising temperatures and changing climate patterns, may favor the expansion of the growing area of heat-loving species, including soybean. However, it's essential to consider that even with global warming, weather conditions related to day length, temperature extremes, and moisture deficit or surplus will continue to have the greatest impact on soybean yields [5].

Cai et al. [6] observed that the biochemical composition of soybean seeds can vary depending on various factors such as variety, climatic conditions, growing phase, and agronomic practices. All these factors can affect the protein, fat, carbohydrate, mineral, and other nutrient content of soybean seeds. The development and introduction of new soybean varieties is an important aspect for the development of soybean cultivation technologies. Breeders are working on developing soybean cultivars that have optimal adaptation to different agro-climatic conditions and demonstrate high yields, as well as increased protein and fat content in seeds. The process of breeding new soybean cultivars involves selecting and crossing parental lines and then conducting multi-year trials to evaluate and select the best progeny [7].

Thus, the study of yield and quality of soybean varieties of different origins is important for eliminating the deficiency of dietary protein, diversity of food and feed products, as well as improving public health, which determines the relevance of the study. Furthermore, soybean yield studies can identify varieties that may exhibit higher productivity under different cultivation conditions, which can help farms improve their efficiency and sustainability. Consequently, the development and introduction of new soybean varieties that have high yields, high protein and fat content in seeds is of important practical interest. The introduction of such soybean varieties can considerably increase the total protein and fat yield per unit area, which is important for meeting the requirements for these components. Given the above, the purpose of this study was to compare the yield of soybean varieties of different origins to identify the most productive and quality varieties under very arid climatic conditions. To fulfil this purpose, the following tasks were set and implemented: to study the indicators of yield structure and yield level of different soybean varieties, as well as to investigate the qualitative characteristics of soybean grain (fat and protein content).

2. MATERIALS AND METHODS

Studies on yield and quality of soybean varieties of different origin were conducted in 2022 on the fields of S. Seifullin Kazakh Agro Technical Research University located in Northern Kazakhstan. The region is characterised by a sharply continental climate with arid conditions and average heat levels. The average annual rainfall is 240-330 mm, the average annual sum of positive temperatures is about 1900-2000℃, and the length of soybean growing season is about 89 to 101 days. Temperature conditions during the period under study had insignificant differences, while moisture content differed significantly from the long-term average, as conditions were very dry. The hydrothermal coefficient differed insignificantly by year (Table 1).

Table 1. Hydrothermal coefficient during the growing season of soybean

Year	Month					Average
	V	VI	VII	VIII	IX	
2022	0.35	0.37	0.81	0.47	0.2	0.44

Note: Selyaninov index (K): wet zone, 1.6-1.3; slightly arid zone, 1.3-1; arid zone, 1-0.7; very arid zone, 0.7-0.4; dry zone, <0.4.
Source: compiled by the authors of this study.

Soils of the experimental plot were represented by southern carbonate chernozem with neutral or slightly alkaline reaction, containing 56.5% of physical clay and 43.5% of physical sand in the arable layer. The humus content was about 4.5-5%, nitrogen content – 28-30%, phosphorus content – 0.13-0.14%, and potassium content – 2.1-2.2%. Soil quality index was 65. The relief was flat, not characterised by forest vegetation [8]. More than 130 soybean varieties and lines of different origins were studied. The standard soybean variety approved for use in the area, namely Ivushka, was used for comparison and evaluation of these varieties. The other varieties were arranged in the list at 10-number intervals relative to this standard. Sowing of the collection nursery was done in optimal terms for the zone according to the methodology of implementation of innovative solutions (IIS).

Soybean grain yield was estimated by harvesting at full ripeness from each variant separately. After threshing, seeds were weighed, and yields were recalculated in kg/ha. The Kjedahl method was applied to carry out biochemical analysis of soybean seeds including protein and fat content.

The Kjeldahl method is a widely used chemical technique for determining the nitrogen content in organic and inorganic substances, particularly in samples like soil, plant material, and food products. Named after its inventor, Johan Kjeldahl, this method involves several key steps. The sample is digested by heating it in the presence of concentrated sulfuric acid. During this process, organic nitrogen compounds are converted into ammonium sulfate $(NH_4)_2SO_4$. After digestion, the acidic solution is neutralized with sodium hydroxide (NaOH) to convert the ammonium sulfate to ammonia (NH_3). The liberated ammonia is distilled from the neutralized solution using steam distillation. It is then collected in a receiving flask containing a known amount of boric acid (H_3BO_3) solution. The ammonia forms ammonium borate. The collected ammonium borate solution is titrated with a standard solution of sulfuric acid (H_2SO_4). The acid reacts with the ammonium ions (NH_4^+) to form ammonium sulfate, while the boric acid remains unchanged. By measuring the volume and concentration of the titrant (sulfuric acid) used in the titration, the nitrogen content in the original sample can be determined.

The study carefully selected over 130 soybean varieties of diverse origins to ensure a comprehensive evaluation. A standard regional variety, Ivushka, was chosen as a reference point, with other varieties systematically listed at 10-number intervals relative to Ivushka. This approach aimed to capture a wide range of genetic backgrounds and traits, providing a robust dataset for analysis.

Cultivation and experimentation took place at the fields of S. Seifullin Kazakh Agro Technical Research University in Northern Kazakhstan, a region characterized by a sharply continental climate with arid conditions. This choice aligned with the study's goals of assessing soybean performance in challenging environments. While specific latitude and longitude coordinates were not provided, the description of Northern Kazakhstan's climate and soil characteristics offered sufficient context.

Regarding replication, the methodology does not specify the number of replications for each variety. Soybean grain yield was estimated individually for each variant, implying a potential use of single replications. However, the study's primary focus appears to be the assessment of a wide variety of soybean types rather than statistical replication.

Using the mean standardised values of these biochemical parameters, cluster analysis was performed using Ward's method. In addition, during the study, correlations between the main elements of yield structure and yield of soybean varieties were established. The results were processed for reliability using multivariate analysis of variance MANOVA using Microsoft Excel software and Statistical 10 software package.

Differences in the obtained results are possible at the significance level of $P \leq 0.05$ by Student's t-test.

3. RESULTS

The collection of soybean varieties under study corresponds to the natural-climatic conditions of the north of Kazakhstan with regard to the duration of the growing season, which varies in a narrow range from 89 to 101 days. Notably, the duration of individual development phases was optimal, which favourably affected the formation of high yields (Table 2).

Soybean plant height varies not only because of varietal characteristics but also because of weather conditions. In the study conducted, drought at the beginning of the growing season resulted in the formation of stunted plants, which affected the yield of soybean varieties. Plant height ranged from 17.4 to 42.8 cm. The height of the lower bean attachment on the plants is an important technological indicator. It was found that only 28% of the variability of this trait is due to hereditary factors, while 72% depends on soybean growing conditions. The lower bean attachment height ranged from 8 cm to 10.92 cm, with line No. 83 having the highest value (10.92 cm). The number of beans per plant varied significantly – from 4 to 38 beans per plant. Some varieties such as Suiyang 1, SC Farta, Jinyaan 55, Kenfeng 20, Dongnong 63, Chera 1-1, Progress, and lines No. 33, No. 92 and No. 77 had maximum number of beans. The coefficient of variation for this trait was 30.5%.

The number of productive nodes on a plant is highly dependent on growing conditions. It was found that this trait has a positive correlation with yield. High values of productive nodes were characteristic of varieties Suiyang 1, SC Farta, Beidou 52, Jinyaan 55, Kenfeng 20, Nuralem-1, Omskaya 4, as well as lines No.78, No.16, No.113, No.114, No.5, K-0125, K-0126, K-0127, Victoria, K-0134. The coefficient of variation for this trait was 23%. The soybean varieties under study differed in seed size as evidenced by thousand-kernel-weight, which ranged from 90 g to 223.1 g. The study identifed varieties with maximum values of this trait (192.3-223.1 g): Suiyang 1, Heihe 43, Heihe 58, Heihe 49, Beidou 36, Beidou 26, Juisan 14-99, Huajiong 2, Heihe 33, Kendou 69, Heike 59, Chera-1-3, Mezenka, Line No. 83, Omskaya 4 and Beidou 43. The share of these varieties was 12.2%. Varieties with seed weight between 90 g and 192.3 g accounted for 63.4%, while small grain varieties accounted for 24.4%. The coefficient of variation for this trait was 18.3%.

The yields of the soybean varieties studied ranged within 7.8-12.2 kg/ha. High yielding varieties accounted for 11.5% of the total and included Ivushka, Suiyang 1, SC Farta, Beidou 43, Beidou 19, Kenfeng 20, Huajiong 2, Kenfeng 6, Chera 1-1, Beidou 47 and Progress, as well as promising lines No. 78, No. 33, No. 83, and No. 16.

Table 2. Soybean yield structure and yield level indicators, average for 2022

Variety	H1, cm	H2, cm	Q1, pcs.	N, pcs.	Q2, pcs.	Q3, pcs.	W1, g	W2, g	W3, g	Y, c/ha
Ivushka	28.22	8.4	1.3	8.6	16	2.8	5.39	0.49	156.8	7.9
Suiyang 1	27.4	8.4	2.8	12.4	25.6	3	9.94	0.52	198.7	12.2
SC Farta	42.8	9.8	2.2	15.6	30.8	3	13.87	0.5	115.5	8.5
Beidou 43	30.1	9.6	2.25	11.2	18.6	3	10.63	0.71	223.1	10.1
Beidou 19	33.4	8	2	11	18.2	3.4	9.78	0.63	156	7.8
Kenfeng 20	37.6	9.1	4	13.8	25	3.2	4.125	0.51	130.2	8.3
Huajiong 2	28	10.1	1.6	11.6	19.2	3	9.99	0.66	197.2	9.1
Kenfeng 6	31.6	10.6	0.1	8	16.8	3.2	6.5	0.54	184	7.9
Chera 1-1	35.8	9.7	1.7	12.2	21.8	2.8	4.48	2.92	160	7.8
Beidou 47	23	9.5	3	13	38	3	10.7	0.55	190	10.1
Progress	26.5	9.8	2.25	15	22.6	2.6	11.79	0.31	175.6	8.3
Line 78	32.9	10.4	2	19.4	37	3	10.03	0.53	117.3	10.4
Line 33	29.7	10.24	2.8	13	22.4	2.6	9.02	0.52	170.4	7.9
Line 83	27.4	10.92	5.8	10	17.8	2.8	8.09	0.51	192.3	7.8
Line 16	39.4	10.5	2.5	15.25	34.25	2.75	5.75	0.45	161.1	12.1
CV, %	17.1	39.8	52.8	23	30.5	7.6	32.9	90.4	18.3	16.9

Note: H1 – Height of plant; H2 – Height of lower bean attachment; N – Number of productive nodes; Q1 – Quantity of lateral branches; Q2 – Quantity of beans on 1 plant; Q3 – Quantity of grains in 1 bean; W1 – Seed weight from 1 plant; W2 – Seed weight from 1 bean; W3 – thousand-kernel weight; Y – Yield.
Source: compiled by the authors of this study.

Table 3. Correlations of main elements of yield structure and yield of soybean varieties, average for 2022

	Y	H1	H2	Q1	N	Q2	Q3	W1	W2
Y									
H1	-0.01								
H2	-0.56	0.59							
Q1	0.04	-0.11	0						
N	0.42	0.36	-0.15	0.14					
Q2	0.6	0.22	-0.31	0.14	0.8				
Q3	-0.03	0.2	0.21	-0.19	-0.18	-0.06			
W1	0.16	-0.18	-0.04	0	0.35	0.25	0.01		
W2	-0.22	0.2	0.11	-0.17	-0.09	-0.13	-0.09	-0.39	
W3	0.19	-0.68	-0.37	0.05	-0.57	-0.41	-0.12	0.09	-0.02

Note: H1 – Height of plant; H2 – Height of lower bean attachment; N – Number of productive nodes; Q1 – Quantity of lateral branches; Q2 – Quantity of beans on 1 plant; Q3 – Quantity of grains in 1 bean; W1 – Seed weight from 1 plant; W2 – Seed weight from 1 bean; W3 – thousand-kernel weight; Y – Yield.
Source: compiled by the authors of this study.

The coefficient of variation for this trait was 16.9%. The coefficients of variation for plant height, thousand-kernel-weight and yield were about 16.9-17.7%, indicating average variability in these traits. The other traits under study such as height of lower bean attachment, number of productive nodes, number of lateral branches, number of beans per plant, seed weight per plant and seed weight per bean had coefficient of variation more than 20% indicating significant variability. The coefficient of variation for the number of beans per plant was 7.6%.

The study of soybean varieties revealed that the number of beans was strongly related to the number of productive nodes, having a stable high correlation (r=0.8). Moderate relationship was found between productivity and number of productive nodes (r=0.42), and between productivity and seed weight per bean (r=0.16), and thousand-kernel-weight (r=0.19). A negative weak relationship was observed between productivity and seed weight per bean (r=-0.22). A moderately negative relationship was found between yield and lower bean attachment height (r=-0.56). There was little or no relationship between productivity and plant height and number of grains in the bean (Table 3).

The analysis of the obtained dendrograms helped group the soybean varieties of different origins under study based on minimum and maximum interactions both within the group and between clusters. The study found that with minimal interaction within the group, multiple clusters were formed. This is explained by the large number of soybean samples under study, their origin, and the diversity of yield structure elements. A comparable situation is observed on the dendrogram on yield, where also more than 7 main "large" clusters were formed, which are of practical interest for breeding in the conditions of northern Kazakhstan. Cluster analysis by Ward's method was performed to investigate the structure and level of yield formation of soybean varieties. The results of the analysis are presented in the form of dendrograms that show the sequence of association and separation of the soybean varieties of different origins under study into clusters (Figures 1 and 2).

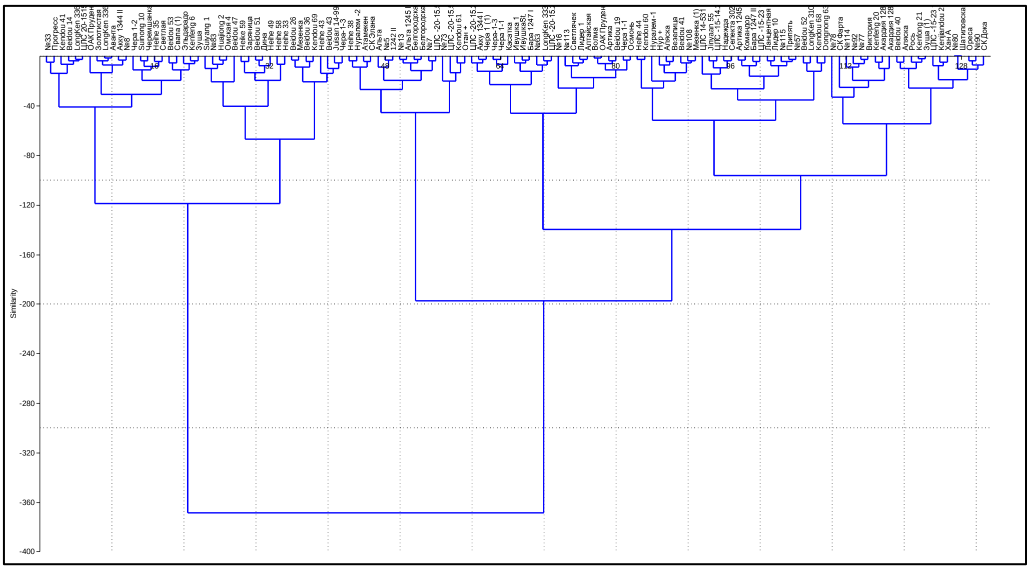

Figure 1. Dendrogram of similarity-difference of soybean varieties of different origin by elements of yield structure
Source: compiled by the authors of this study.

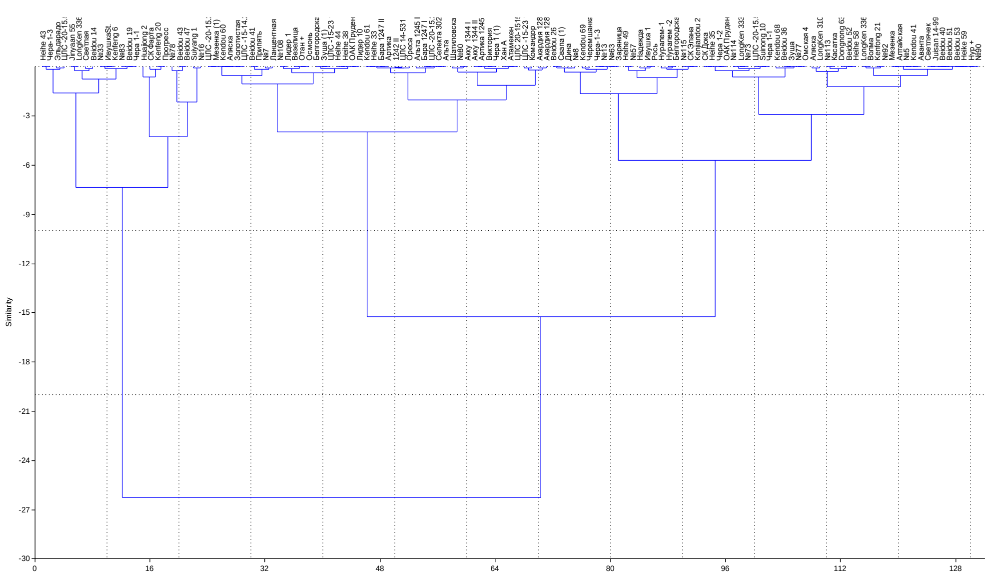

Figure 2. Dendrogram of clustering of soybean varieties of different origins by yields
Source: compiled by the authors of this study.

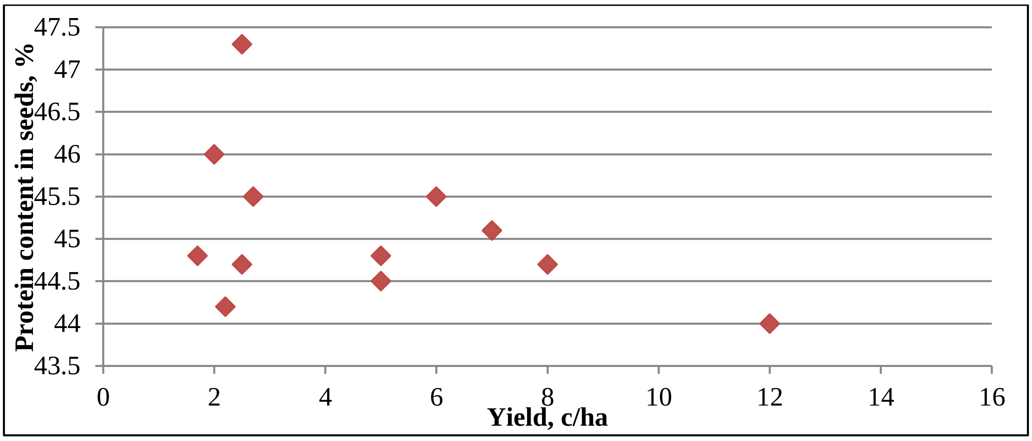

Figure 3. Cluster analysis of similarities and differences between soybean varieties in terms of yield and protein content in seeds
Source: compiled by the authors of this study.

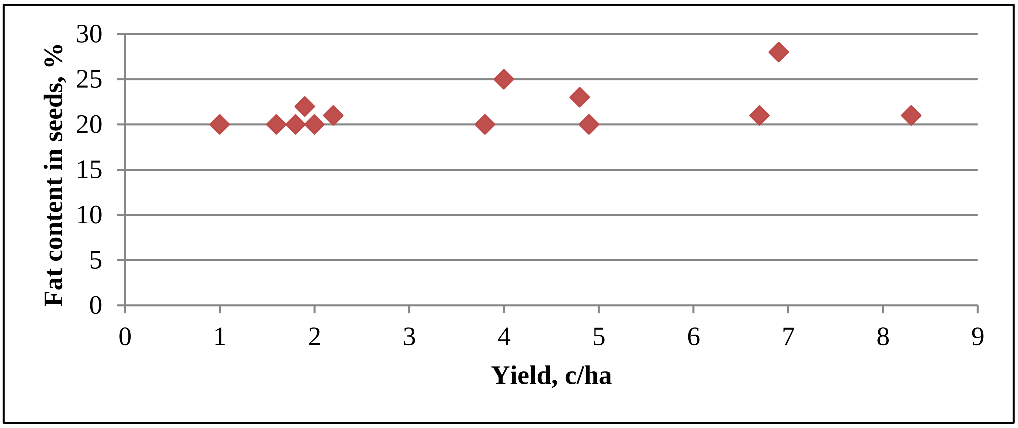

Figure 4. Cluster analysis of similarities and differences between soybean varieties in terms of yield and fat content in seeds
Source: compiled by the authors of this study.

Protein content is a quantitative trait that is subject to polygenic inheritance and high sensitivity to changes in environmental conditions. Under favourable moisture conditions, soybean varieties can use biologically fixed air nitrogen and reach a protein content of up to 48% in seeds. Therefore, the search for and isolation of high-protein accessions and their use in the breeding process to create new varieties are still relevant. However, protein accumulation in soybean seeds depends on external conditions, soil and climatic factors, and varietal characteristics (Table 4).

Table 4. Variation in protein and fat content of soybean varieties and their coefficient of variation, average for 2022

Quality Indicators	Range of Variation	Average	Coefficient of Variation V, %
Protein	38.66-47.25	42.1±1.73	4.1
Fat	15.82-27.8	19.1±1.72	9

Source: compiled by the authors of this study.

In this study, the protein content of soya seeds was found to vary from 38.66% to 47.25% irrespective of variety origin. Samples with high seed protein content such as Beidou 26, Kendou 60, Alaska, Osmon, K-0126, K-0128 were identified. The rest of the varieties had average values of seed protein content, which was 95.5% of the total number of samples under study. The variation of biochemical parameters, including protein content, had acceptable values of the coefficient of variation, indicating the need to use more adaptive soybean varieties in practical breeding. In terms of the fat content of soybean seeds, it was observed that some varieties such as Heihe 49 (China), Kasatka (Russia) and Eldarado (Russia) had high fat content ranging from 25.4% to 27.8%. At the same time, some varieties including Suiyang 1, SC Farta, Svetlyachek, Beidou 43, Heihe 58, Heihe 33, Kendou 60, Heike 59, Chera-1-3, Nuralem-2, Ros, Zolotistaya, Nur+, Otan+, Line No. 8, Line No. 113, Oresa and Dina had low fat content of 13.6% of the total samples under study. Theoretical regression lines with straight-line correlation between yield and protein content, and yield and fat content, show that grain yield level with high protein content has an insignificant relationship (r=0.004, significance P=99), and with fat content has a weak relationship (r=0.1, significance P=99). This indicates a weak relationship between the traits under study. A cluster analysis of the best soybean varieties was conducted to examine the similarity of soybean source material for yield and protein content traits (Figure 3).

Thus, the studies showed differences in protein and fat content of soybean seeds among the varieties under study. The possibility of obtaining high-protein soybean varieties with high yield seems to be a promising task for breeding. However, the weak relationship between yield and protein content and the negative correlation between the two must be considered when developing new soybean varieties with desirable traits. In the analysed experimental sample, the division of soybean varieties into 6 clusters depending on protein content was observed. Cluster No. 1 consists of one sample (Osmonj variety) which has high seed protein content, but yield is not characterised by high value. Clusters No. 2 and No. 3 have yields at 1 kg/ha, with cluster No. 2 varieties having high protein content and cluster No. 3 varieties having medium protein content. The varieties of cluster No. 4 have predominantly medium protein content and yields range from 2 to 6 kg/ha. Cluster No. 6 is low in protein but has high yields. However, no soybean varieties were found in the sample that combined high yield and high protein content. This complicates the breeding process and requires the use of complex crosses and prolonged selection of hybrid progeny to obtain varieties that fulfil the required characteristics. A cluster analysis of the best soybean varieties was also conducted to examine the similarity of soybean source material for yield and fat content traits (Figure 4).

In analysing the experimental sample, it was found that the first group consisted of one sample of medium fat content and very low grain productivity. Group No. 2 includes 6 samples with low yield and medium fat content. In cluster No. 3, the yields of varieties ranged within 4–6 kg/ha, with two samples standing out for their high fat content and the remaining three having average values. Group No. 4 was characterised by higher productivity as compared to other groups. Three varieties were identified in this group: one variety with high fat content and two varieties with medium fat content in soybean seeds. The analysis shows a weak negative correlation ($r=-0.133$) between grain fat content and yield. The trait “fat content” also had the least variability ($CV=9.08\%$). Analysis of variance revealed a relationship between soybean yield and grain quality. The correlation coefficient between yield and grain protein content was negative ($r=-0.067$), indicating a weak relationship between these parameters. Similarly, the correlation coefficient between yield and fat content was also negative ($r=-0.13$). These results indicate that an increase in yield does not necessarily lead to an increase in grain protein and fat content. The regression coefficient for yield and protein content was 0.005 and for yield and fat content was 0.1. The t value (statistical significance) at the significance level of 0.05 was -2.23 for both cases. Thus, the analysis conducted showed that soybean yield depends on a range of factors such as productive nodules, thousand-kernel-weight, and other plant characteristics. However, the relationship between yield and grain protein or fat content is weak and ambiguous. These results may be useful in selecting high yielding soybean varieties and determining their quality characteristics.

4. DISCUSSION

Current research confirms the importance of increasing soybean production and is being conducted to investigate various aspects of soybean yield and quality. This is important for several reasons [9-11]:

1. Food safety. Increasing soybean production plays a vital role in ensuring food security as soybean is one of the major sources of vegetable protein and vegetable fat in the world. Increased soybean yields provide access to food resources for the population, especially in countries where protein deficiency is an issue.
2. Economy support. Soybean is a major crop that creates economic prosperity in many parts of the world. Increased soybean yields contribute to agricultural development, create jobs, increase the incomes of agricultural producers, and support economic growth.
3. The importance of high-protein crops. Soybean is one of the major high-protein crops that play an essential role in balancing the diet of the population. The protein found in soybean has high nutritional value and is an important source of amino acids for humans. Increased soybean production reduces dietary protein deficiencies and improves nutrition.
4. Improving environmental sustainability. Soybean yield and quality research is also aimed at developing varieties with increased resistance to diseases, pests, and agro-climatic conditions. This helps reduce the use of pesticides and fertilisers and reduce the adverse environmental impact of agriculture.

The weak and negative relationships between yield and protein content, as well as yield and fat content, have important implications for soybean cultivation and its impact on food security and quality. While this study does not delve deeply into the discussion of these relationships, we can explore potential implications and explanations. A negative relationship between yield and protein content suggests that as soybean yields increase, the protein content may decrease [12-14]. This has implications for the nutritional quality of soybean-based products. Lower protein content could impact the suitability of soybeans for use in protein-rich foods and animal feed. This may not align with the goal of improving food security, especially in regions where soybeans are a significant source of protein [15].

Similarly, a negative correlation between yield and fat content implies that as yields increase, the fat content may decrease. The fatty acid composition of soybeans is essential for various applications, including cooking oils and biodiesel production. A reduction in fat content could affect the suitability of soybeans for these purposes. The negative relationships may be attributed to crop management practices. High-yielding varieties may require different agronomic practices that prioritize yield over nutritional quality. For example, increased plant density and nitrogen fertilization may boost yields but potentially dilute protein and fat concentrations. Understanding these trade-offs is crucial for farmers and policymakers when selecting soybean varieties and implementing farming practices.

In summary, the weak and negative relationships between yield and protein/fat content in soybeans have complex implications for food security, nutritional quality, and agricultural practices. Further research and a nuanced approach to crop management and breeding are necessary to address these challenges effectively.

Limitations related to heat, photoperiodic response, and weather conditions can be significant factors affecting soybean production. Kazakhstan has a sharply continental climate characterised by dry and hot summers, cold winters and significant temperature fluctuations throughout the year. Late spring and early autumn frosts can be dangerous for many crops, including soybeans. Under conditions of long daylight hours and insufficient sum of active temperatures during the growing season, it is important to choose soybean varieties with weak photoperiodic sensitivity. Such varieties can adapt to conditions with long daylight hours and start flowering relatively early, allowing seed formation in a shorter growing season [16].

According to Ghani et al. [17], yield and quality studies on soybean varieties of different origins help expand the genetic resource of the crop and ensure adaptation to changing conditions. This is particularly important considering climate change, which requires crops that can withstand extreme temperatures, drought, and other stressful conditions. The development and introduction of new soybean varieties has already led to several results. For instance, soybean varieties have been developed that have increased resistance to drought, pests, and diseases, thereby increasing yields and reducing losses. Some new varieties also show improved quality characteristics such as higher protein and fat content in the seeds. This allows farmers to grow soybean in a wide range of regions, thereby promoting diversity of production and increasing its sustainability [18]. Years of trials and progeny evaluation help select the best soybean varieties with high yields and desirable traits. This is a crucial step in the breeding process that allows discarding less successful variants and focus on developing the most promising varieties.

Adetiloye et al. [19] believe that cluster analysis method is an important tool in the study of yield and quality of soybean varieties of different origins, allowing systematisation of data, identification of patterns and drawing conclusions regarding different groups of varieties. Furthermore, the method of cluster analysis for grouping soybean varieties based on their yield and quality characteristics, as well as determining the correlation relationship between the main elements of yield structure and yield of soybean varieties has been widely applied by different researchers. Thus, in a study by Li et al. [20], in which 40 soybean varieties were clustered based on their yield and quality characteristics such as protein, fat content, and other components, several clusters were identified to classify soybean varieties according to their performance and quality.

A study by Mahmoud et al. [21] also applied cluster analysis method to investigate 25 soybean varieties based on their yield and phytochemical parameters such as isoflavones content and antioxidant activity. Cluster analysis allowed the identification of groups of varieties with comparable characteristics, which helped assess differences in yield and quality among them. According to the study by Matsuo et al. [22], in which data on yield and quality characteristics of 20 soybean varieties were analysed, positive correlations between yield and such elements of yield structure as number of beans per plant and seed weight per plant were found. Furthermore, a negative correlation relationship was found between yield and seed fat content, which also supports the findings of the current study.

Further support for the study can be found in a study by Mofokeng and Mashingaidze [23], where yield data of 30 soybean varieties were analysed. Positive correlations between yield and such elements of yield structure as number of lateral branches, plant height, and number of beans per plant were found. A positive correlation between yield and thousand-kernel-weight was also found. A plethora of studies conducted on the development of high-yielding soybean varieties with increased seed protein content have revealed some inverse correlation between yield and protein content [24-26]. Thus, in the study of Agegn et al. [5], a negative correlation relationship was observed between yield and protein content of soybean seeds (r=-0.068) and the coefficient of variation for this trait was 4.1%, which also resonates with the study conducted.

Staniak et al. [27] suggest that one plausible reason for this inverse correlation may be that when genetic characteristics of a variety are improved to increase protein content, plant growth intensity and yield may decrease slightly. This is because plants may devote more energy and resources to protein accumulation in seeds, to the detriment of other aspects of growth and development. However, Xia et al. [28] note that this inverse correlation is not always observed in all studies and varieties of soybean. In some cases, it has been possible to develop varieties that combine both high yields and increased protein content. This can be achieved through breeding and genetic improvement of varieties and optimisation of growing conditions and agronomic practices. According to Adamič and Leskovšek [1], one of the challenges for researchers and breeders is to strike a balance between yield and protein content in soybean seeds. The aim is to obtain varieties that provide high yields and at the same time have a high enough protein content to meet the needs of the food and feed industry.

Considering the number of scientific studies [29-31], as well as the results of the performed research on the yield and quality of soybean varieties of different origins, it can be stated that the yield and quality of soybean varieties can vary significantly depending on their origin and growing conditions. Studies conducted in Northern Kazakhstan allow assessing the adaptation and productivity of soybean varieties under specific agro-climatic conditions, and help identify the most suitable varieties, considering factors such as climate, soil, moisture, and other conditions

A study, like any other, has its limitations, which may affect its results and the adequacy of its conclusions. One possible limitation is that soybean varieties were sampled from only a limited number of sources. This may lead to the fact that the results do not fully reflect the diversity of soybean varieties and their characteristics. For more representative conclusions, it would be useful to expand the sample to include more varieties from different sources and regions.

Since the study was conducted in a specific region, changes in climatic conditions may affect the results. Changes in temperature, humidity, and other factors can cause changes in plant growth and development, which in turn can affect the protein and fat content of soybean seeds. The results of the study may also be limited by the agronomic practices used during soybean cultivation. Different approaches to tillage, fertilization, and irrigation can affect soybean yield and quality. Optimizing these practices can improve results.

Data quality and measurement accuracy can also affect results. If data sources were not reliable or measurements were made with errors, this can lead to inaccurate results. Differences in the genetic diversity of soybean varieties can also be a factor affecting results. Different varieties may have different genetic potential for growing proteins and fats, and this can be an important factor. Limitations related to environmental conditions and economic circumstances can also affect the results of a study. For example, economic instability or changes in market conditions could affect soybean production.

In summary, the weak and negative relationship between yield and protein/fat content of soybeans has complex implications for food security, nutrition, and agriculture. To effectively address these issues, more research and a nuanced approach to plant cultivation and breeding are needed.

5. CONCLUSIONS

As a result of the study, more than 7 main large clusters were formed, in which soybean varieties are of practical interest for breeding in the conditions of northern Kazakhstan. The number of beans per plant was found to range from 4 to 38 per plant. The protein content of soybean seeds ranged from 38.66% to 47.25% depending on the origin of the variety.

Varieties with high seed protein content are Beidou 26, Kendou 60, Alaska, Osmon, K-0126, K-0128. The following varieties were noted for high fat content (25.4-27.8%) in soybean seeds: Heihe 49 (China), Kasatka (Russia) and Eldarado (Russia). The analysis indicates that there is a weak negative correlation relationship between quality parameters (protein and fat) and yield in the studies: r=-0.068 and r=-0.133 respectively. The results suggest a need for further research into the relative influence of genetic factors and cultivation conditions.

The study also found that soybean yield depends on various factors such as number of productive nodes, seed weight per bean, thousand-kernel-weight, and other plant characteristics. However, the relationship between yield and grain protein or fat content is not strong and unambiguous. This means that high protein or fat content in soybean seed does not always guarantee high yield, and conversely, high yield does not always mean high protein or fat content. The practical significance of the findings lies in the fact that they can help breeders and farmers to select the most suitable varieties for optimal cultivation, increasing yields and improving the nutritional value of products. The prospect for further research is to assess the impact of various agronomic practices on the yield and quality of soybean varieties. This may include determining the optimal doses of fertilisers, sowing methods, irrigation system, and other factors. Potential methodologies and expected results of such research include optimization of fertilizer doses. Conducting controlled experiments with the application of different doses of fertilizers (nitrogen, phosphorus, potassium, etc.) to soybean crops. The result is to determine the ideal fertilizer dose that maximizes yield while maintaining acceptable levels of protein and fat in soybean seeds. The seeding methodology is to compare different seeding methods, such as conventional tillage, no-till, and precision seeding, in field trials. The result is the determination of the sowing method that provides the highest yield and optimal nutritional quality of the soybean crop.

The findings of this study have several potential impacts on yield, nutritional value, and farming efficiency in soybean production. Understanding that the relationship between yield and grain protein or fat content is not always strong and unambiguous allows farmers to focus on maximizing yield without sacrificing nutritional quality. This can lead to increased soybean production, providing higher yields and more food resources. Farmers and breeders can use this knowledge to select soybean varieties that strike a balance between yield and nutritional quality. They can target varieties that offer both high yields and acceptable levels of protein and fat content, ensuring better nutritional value in soybean-based products. The study paves the way for research into agronomic practices that can further optimize yield and quality. This includes determining the optimal use of fertilizers, sowing methods, irrigation systems, and other factors. By fine-tuning these practices, farmers can enhance efficiency in soybean cultivation.

ACKNOWLEDGEMENTS

This research was carried out under grant funding for scientific and (or) scientific-technical projects for 2022-2024 years of the Ministry of Science and Higher Education of the Republic of Kazakhstan IRNAP14870651 "Production early maturing soybean material with high productive and adaptive potential using molecular breeding methods for conditions of Northern Kazakhstan". Programme-targeted financing of scientific research (IRN BR10764991), on the theme "Creation of high-yielding varieties and hybrids of oilseeds, grain crops on the basis of modern scientific achievements for sustainable production in different zones of Kazakhstan".

REFERENCES

[1] Adamič, S., Leskovšek, R. (2021). Soybean (*Glycine max* (L.) Merr.) growth, yield, and nodulation in the early transition period from conventional tillage to conservation and no-tillage systems. Agronomy, 11(12): 2477. https://doi.org/10.3390/agronomy11122477

[2] Achina, T., Twumasi, P., Quain, M.D. (2019). Use of morphological and molecular profiling for the establishment of relatedness between commercially viable and elite soybean genotypes. European Journal of Biotechnology and Genetic Engineering, 6(1): 1-17.

[3] Ignatova, L., Brazhnikova, Y., Berzhanova, R., Mukasheva, T. (2015). The effect of application of micromycetes on plant growth, as well as soybean and barley yields. Acta Biochimica Polonica, 62(4): 669-675. https://doi.org/10.18388/abp.2015_1100

[4] Baraskar, V.V., Kachhadia, V.H., Vachhan, J.H., Barad, H.R., Patel, M.B., Darwankar M.S. (2014). Genetic variability, heritability and genetic advance in soybean [*Glycine max* (L.) Merrill]. Electronic Journal of Plant Breeding, 5(4): 802-806.

[5] Agegn, A., Bitew, Y., Ayalew, D. (2022). Response of yield and quality of soybean [*Glycine max* (L.) Merrill] varieties to blended NPSZnB fertilizer rates in Northwestern Ethiopia. Heliyon, 8(5): 09499. https://doi.org/10.1016/j.heliyon.2022.e09499

[6] Cai, Y., Chen, L., Zhang, Y., Yuan, S., Su, Q., Sun, S., Wu, C., Yao, W., Han, T., Hou, W. (2020). Target base editing in soybean using a modified CRISPR/Cas9 system. Plant Biotechnology Journal, 18(10): 1996-1998. https://doi.org/10.1111/pbi.13386

[7] Alsajri, F.A., Wijewardana, C., Irby, J.T., Bellaloui, N., Krutz, L.J., Golden, B., Gao, W., Reddy, K.R. (2020). Developing functional relationships between temperature and soybean yield and seed quality. Agronomy Journal, 112(1): 194-204. https://doi.org/10.1002/agj2.20034

[8] Tokarieva, O., Puzrina, N., Vorotynskyi, O. (2022). Pyrological characteristics of forest edges under intensive recreational loads. Ukrainian Journal of Forest and Wood Science, 13(2): 51-57. https://doi.org/10.31548/forest.13(2).2022.51-57

[9] Dayoub, E., Piva, G., Shirtliffe, S.J., Fustec, J., Corre-Hellou, G., Naudin, C. (2022). Species choice influences weed suppression, N sharing and crop productivity in oilseed rape – Legume intercrops. Agronomy, 12(9): 2187. https://doi.org/10.3390/agronomy12092187

[10] Gardarin, A., Celette, F., Naudin, C., Piva, G., Valantin-Morison, M., Vrignon-Brenas, S., Verret, V., Médiène, S. (2022). Intercropping with service crops provides multiple services in temperate arable systems: A review. Agronomy for Sustainable Development, 42(3): 39. https://doi.org/10.1007/s13593-022-00771-x

[11] Pospielova, G., Kovalenko, N., Nechyporenko, N., Kocherga, V. (2020). Influence of agroclimatic factors

on the development of common soybean diseases. Ukrainian Black Sea Region Agrarian Science, 24(3): 45-52. https://doi.org/10.31521/2313-092X/2020-3(107)-6
[12] Milenko, O., Shevnikov, M., Solomon, Yu., Rybalchenko, A., Shokalo, N. (2022). Influence of foliar top-dressing on the yield of soybean varieties. Scientific Horizons, 25(4): 61-66. https://doi.org/10.48077/scihor.25(4).2022.61-66
[13] Ivaniv, M., Vozniak, V., Marchenko, T., Baklanova, T., Sydiakina, O. (2023). Varietal features of elements of soybean cultivation technology during irrigation. Scientific Horizons, 26(6): 85-96. https://doi.org/10.48077/scihor6.2023.85
[14] Hnatiuk, T., Kravchenko, O., Abarbarchuk, L., Churilov, A., Chobotar, V. (2023). Influence of drugs produced by electropulse ablation methods on the development of soybean phytopathogenic bacteria. Plant and Soil Science, 14(3): 22-34. https://doi.org/10.31548/plant3.2023.22
[15] Turmagambetova, A.S., Sokolova, N.S., Bogoyavlenskiy, A.P., Berezin, V.E., Lila, M.A., Cheng, D.M., Dushenkov, V. (2015). New functionally-enhanced soy proteins as food ingredients with anti-viral activity. VirusDisease, 26(3): 123-132. https://doi.org/10.1007/s13337-015-0268-6
[16] Gawęda, D., Haliniarz, M., Bronowicka-Mielniczuk, U., Łukasz, J. (2020). Weed infestation and health of the soybean crop depending on cropping system and tillage system. Agriculture, 10(6): 208. https://doi.org/10.3390/agriculture10060208
[17] Ghani, M., Kulkarni, K.P., Song, J.T., Shannon, J.G., Lee, J.D. (2016). Soybean sprouts: A review of nutrient composition, health benefits and genetic variation. Plant Breeding and Biotechnology, 4: 398-412. https://doi.org/10.9787/PBB.2016.4.4.398
[18] Hou, Z., Fang, C., Liu, B., Yang, H., Kong, F. (2023). Origin, variation, and selection of natural alleles controlling flowering and adaptation in wild and cultivated soybean. Molecular Breeding, 43(5): 36. https://doi.org/10.1007/s11032-023-01382-4
[19] Adetiloye, I.S., Ariyo, O.J., Alamu, O., Osewa, S.O. (2020). Agronomic potential and genetic diversity of 43 accession of tropical soybean (*Glycine max* (L) Merrill). International Journal of Plant Research, 10(2): 33-39. https://doi.org/10.5923/j.plant.20201002.02
[20] Li, M., Liu, Y., Wang, C., Yang, X., Li, D., Zhang, X., Xu, C., Zhang, Y., Li, W., Zhao, L. (2020). Identification of traits contributing to high and stable yields in different soybean varieties across three different Chinese latitudes. Frontiers in Plant Science, 10: 1642. https://doi.org/10.3389/fpls.2019.01642
[21] Mahmoud, R., Casadebaig, P., Hilgert, N., Alletto, L., Freschet, G.T., de Mazancourt, C., Gaudio, N. (2022). Species choice and N fertilization influence yield gains through complementarity and selection effects in cereal-legume intercrops. Agronomy for Sustainable Development, 42: 12. https://doi.org/10.1007/s13593-022-00754-y
[22] Matsuo, N., Tsuchiya, S., Nakano, K., Fukami, K. (2019). Design and evaluation of a one-operation shallow up-cut tillage sowing method for soybean production. Plant Production Science, 22(4): 465-478. https://doi.org/10.1080/1343943X.2019.1625274
[23] Mofokeng, M.A., Mashingaidze, K. (2018). Genetic differentiation of ARC soybean [*Glycine Max* (L.) Merrill] accessions based on agronomic and nutritional quality traits. Agriculture and Food Sciences Research, 5(1): 6-22. https://doi.org/10.20448/journal.512.2018.51.6.22.
[24] Rivière, C., Béthinger, A., Bergez, J.E. (2022). The effects of cover crops on multiple environmental sustainability indicators – A review. Agronomy, 12(9): 2011. https://doi.org/10.3390/agronomy12092011
[25] Nakagawa, A.C.S., Ario, N., Tomita, Y., Tanaka, S., Murayama, N., Mizuta, C., Iwaya-Inoue, M., Ishibashi, Y. (2020). High temperature during soybean seed development differentially alters lipid and protein metabolism. Plant Production Science, 23(4): 504-512. https://doi.org/10.1080/1343943X.2020.1742581
[26] Shuvar, I., Korpita, H. (2023). Herbicide influence on the agrocenose of soy and its photosynthetic activity in the western Forest Steppe of Ukraine. Ukrainian Black Sea Region Agrarian Science, 27(2): 21-27. https://doi.org/10.56407/bs.agrarian/2.2023.21
[27] Staniak, M., Stępień-Warda, A., Czopek, K., Kocira, A., Baca, E. (2021). Seeds quality and quantity of soybean [*Glycine max* (L.) Merr.] cultivars in response to cold stress. Agronomy, 11(3): 520. https://doi.org/10.3390/agronomy11030520
[28] Xia, Z., Zhai, H., Wu, H., Xu, K., Watanabe, S., Harada, K. (2021). The synchronized efforts to decipher the molecular basis for soybean Maturity Loci E1, E2, and E3 that regulate flowering and maturity. Frontiers in Plant Science, 12: 632754. https://doi.org/10.3389/fpls.2021.632754
[29] Tsyhanskyi, V., Tsyhanska, O., Didur, I. (2023). Influence of biologisation of the nutrition system on the transformation of biological nitrogen and formation of soybean productivity. Plant and Soil Science, 14(4): 86-97. https://doi.org/10.31548/plant4.2023.86
[30] Sydiakina, O., Ivaniv, M. (2023). Productivity of soybean varieties of different maturity groups depending on plant density under drip irrigation in the South of Ukraine. Scientific Horizons, 26(11): 100-110. https://doi.org/10.48077/scihor11.2023.100
[31] Litvinov, D., Olefirenko, O. (2023). Assessment of the tillage impact on soybean productivity. Plant and Soil Science, 14(3): 75-83. https://doi.org/10.31548/plant3.2023.75