

## Discrimination of radish seed cultivars using fluorescence spectroscopy and advanced machine learning algorithms

T. Yari Dramane<sup>1</sup>, V. Slavova<sup>2\*</sup>, A. Yasar<sup>3</sup>, T. L. Dimitrova<sup>4</sup>

<sup>1</sup>Graduate School of Natural and Applied Sciences, Department of Computer Engineering, Selcuk University, Konya, Türkiye

<sup>2</sup>Department of Breeding, Maritsa Vegetable Crops Research Institute, Brezovsko shosse 32, 4000 Plovdiv, Agricultural Academy, Bulgaria

<sup>3</sup>Department of Mechatronics Engineering, Selcuk University, Konya, Türkiye

<sup>4</sup>University of Plovdiv "Paisii Hilendarski", 24 Tsar Assen Str., 4000, Plovdiv, Bulgaria

Received: July 30, 2025; Revised: October 19, 2025

The differentiation of seed varieties plays a vital role in modern agriculture, since seed type and quality directly influence germination rate, plant vigor, crop yield, and ultimately, farmers' profitability. In recent years, the integration of decision support systems with advanced analytical tools has become increasingly important to ensure accurate and practical solutions for farmers. Within this scope, the present study addresses the classification of radish seed varieties, namely French breakfast, Nacional 2, Espresso F1, and Red large, by combining fluorescence spectroscopic techniques with state-of-the-art machine learning algorithms. The rationale behind the proposed approach lies in the complementary strengths of both methods: fluorescence spectroscopy provides a non-destructive, rapid, and sensitive characterization of seeds, while machine learning algorithms enhance the ability to recognize subtle spectral differences and achieve reliable classification outcomes.

In this study, four classifiers: Fine Tree, Quadratic Support Vector Machine (SVM), Fine k-Nearest Neighbor (KNN), and Neural Networks were employed to evaluate their performance on fluorescence spectral data. The findings revealed that all applied algorithms produced satisfactory classification accuracies exceeding 90%, thereby confirming the robustness of the proposed framework. Notably, the Quadratic SVM model achieved the highest performance with an accuracy of 100%, demonstrating its superior capability in distinguishing radish seed varieties with complete precision.

These results highlight that the synergy between spectroscopic data and machine learning models can be effectively utilized in practice for agricultural decision support. The developed methodology offers farmers a practical, non-invasive, and highly accurate tool for radish seed variety identification, which may significantly contribute to improved seed management, resource optimization, and sustainable agricultural practices.

**Keywords:** Radish seeds, fluorescence spectroscopic techniques, machine learning techniques

### INTRODUCTION

Radish (*Raphanus sativus*) is an annual root vegetable of the *Brassicaceae* family. Its origin traces back to Central Asia [1], and it has been cultivated as a food crop since approximately 1000 BC in regions including China, Japan, Egypt, Rome, and Greece. Varieties of radish are typically divided into two main categories: European-type and Chinese-type [2]. Radishes are primarily eaten raw, often in salads, as they are not well-suited for cooking. They are easy to grow, adaptable to a variety of soils, and have a rapid growth cycle [3]. Cultivation usually involves direct seeding in the soil, where the plants reach full maturity [4].

The development of rapid and precise non-destructive techniques, such as optical analysis, can facilitate the study and monitoring of seed germination, ensuring high-quality planting material for both farmers and the agricultural industry [5].

Optoelectronic approaches offer advantages including non-contact operation, speed, selectivity, and preservation of sample integrity. These benefits make it feasible to implement non-invasive methods for assessing radish seeds. To date, however, no studies have applied these techniques to radish.

Belyakov (2019) [6] presented findings from research on cereal seeds. Characteristic excitation wavelengths reported were 362 nm [7], 424 nm [8], and 485 nm [9]. These studies demonstrated that during the maturation of cereal seeds such as wheat, oats, and corn, immature seeds tend to show excitation at shorter wavelengths, while mature seeds predominantly emit longer-wavelength radiation. The ratio of long- to short-wavelength emissions changes over time and can be statistically described using linear models, which are essential for constructing reference databases.

Furthermore, Belyakov *et al.* (2021) [10]

\* To whom all correspondence should be sent:  
E-mail: [vanya\\_plachkova@abv.bg](mailto:vanya_plachkova@abv.bg)

developed a sensor capable of determining seed physiological maturity by irradiating seeds with two specific wavelengths and measuring the resulting photoluminescence. Maximum luminescence was observed to be less pronounced compared to the excitation spectrum [11]. Spectral luminescence of forage seeds was also measured after scarification, revealing that repeated scarification induced qualitative changes, including the appearance of a new peak at 423 nm. These observations can similarly be applied to radish seeds, both standard varieties and first-generation hybrids, enabling the design of fiber-optic systems to evaluate planting material.

Optical properties of seeds from other crops, such as legumes [12] and tomatoes [13], have been studied using comparable methods. Typically, the excitation spectrum falls between 355–500 nm, showing a primary peak at 424 nm and a secondary peak at 485 nm. The emission spectrum ranges within 420–650 nm, with maxima around 500–520 nm. The water content of seeds including any impurities, is expressed as a percentage of the total seed mass [14].

A recent study demonstrated that fluorescence spectroscopic data on garlic cultivars can reliably represent different species, achieving a classification accuracy of 99.93% using Neural Network algorithms after hyperparameter optimization [15].

Therefore, this study aims to develop a non-invasive methodology for differentiating between radish seed varieties through integration of fluorescence spectroscopy with advanced machine learning algorithms. This research contributes to the literature by being the first to propose an optoelectronic framework specifically for radish seeds, thereby supporting modern decision support systems in agriculture.

The main contributions of this study can be summarized as follows:

- **Novel methodology:** We propose, for the first time, an optoelectronic framework that integrates fluorescence spectroscopy with advanced machine learning algorithms for the classification of radish seed varieties.
- **Comprehensive dataset:** Experimental data were obtained from four distinct radish cultivars (French breakfast, Nacional 2, Espresso F1, and Red large) with sufficient replicates, ensuring robust evaluation.
- **Detailed performance evaluation:** Multiple machine learning models (Fine Tree, Quadratic SVM, Fine KNN, and Neural Network) were comparatively assessed using 10-fold cross-

validation, with reported classification accuracies consistently above 90%.

- **Insightful analysis:** The study highlights the discriminative spectral region (475–555 nm) and provides a discussion of why Quadratic SVM achieved superior performance.
- **Practical implications:** The proposed non-destructive approach offers a fast, accurate, and reproducible method for seed quality assessment, which can support farmers, seed producers, and breeding programs in agricultural decision-making.
- **Future perspectives:** The methodology sets the foundation for future integration of deep learning algorithms and portable sensor-based systems, enabling real-time applications in precision agriculture.

## MATERIALS AND METHODS

### *Materials*

For this investigation, four radish genotypes were selected, including three conventional cultivars and one first-generation hybrid. The characteristics of each genotype are summarized as follows:

*French breakfast:* This cultivar demonstrates adaptability to both spring and autumn field production. The roots are generally solitary and elongated, with a distinctive white tip. The internal tissue is crisp and pale, providing a favorable texture. The crop reaches physiological maturity within approximately 30 days, and the recommended seeding density is between 1.5 and 2.5 kg per hectare.

*Nacional 2:* An early-maturing cultivar, Nacional 2 produces large, globular roots exhibiting a vibrant red exterior with contrasting white tips. The flesh is tender and crisp, with a mild and pleasing flavor profile.

*Espresso F1:* This early-harvest hybrid is characterized by round, red roots and a robust root structure with firm attachment to the foliage. The variety shows excellent post-harvest durability and transportability, making it suitable for winter cultivation and very early-season field production.

*Red large:* A medium-early cultivar designed for both spring and autumn planting, Red large develops deep-red, single roots with white, flavorful flesh. This variety is distinguished by its resistance to cracking and high sensory quality, which enhances marketability.

These genotypes were chosen to encompass a range of morphological traits, maturation rates, and post-harvest qualities, providing a representative sample for the study of seed and root characteristics.

### Fluorescence spectroscopy

The mobile fiber-optical spectral installation for the study of fluorescence signals is specifically designed for the rapid analysis of plant biological samples. 120 seeds from each of the 10 variants for the varieties Espresso F1, Red large, French breakfast and Nacional 2 were tested. The graphs presented for AI processing are averaged. It was chosen to work with a laser diode with emission radiation of 245 nm because the system is designed for wide application in plant breeding for the analysis of various types of plant samples, and some of them have an emission wavelength with low intensity levels. Figure 1 presents the averaged spectral distributions of seeds of the varieties Espresso F1, French breakfast, Nacional 2 and Red large. This source was chosen after preliminary studies and experimental tests showed that this is the optimal wavelength for the analysis of optically active biological media. The mobile experimental setup used for fluorescence spectroscopy is presented in Figure 2 and includes the following components:

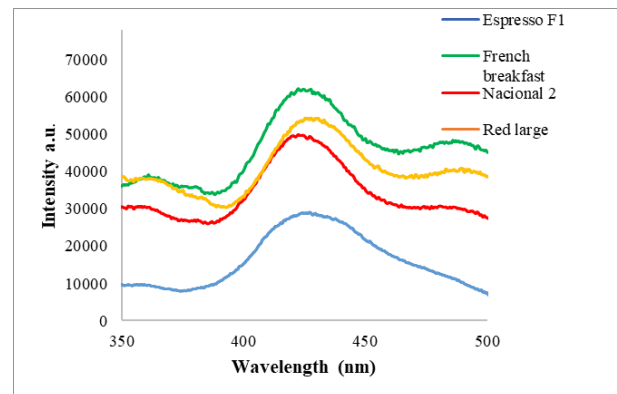
- Laser diode (LED): Emission at 245 nm, supplied at 3 V, enclosed in a hermetically sealed TO39 metal housing. It operates with a voltage drop of 1.9–2.4 V and a current consumption of 0.02 A.
- Rod lens (achromatic doublet type): Consists of two bonded lenses with different dispersion coefficients (Schott and Corning) and an anti-reflective coating. The optical design minimizes chromatic aberration by compensating for lens differences.
- Quartz glass (4 cm<sup>2</sup>): Transparent to visible, ultraviolet, and infrared light, with weak light absorption. Its high purity ensures superior optical and thermal properties compared to conventional glass.
- CMOS detector: Photosensitive area of 1.9968 × 1.9968 mm, sensitivity range of 200–1100 nm, and spectral resolution of  $\delta\lambda = 5$ . It is optimized for precise signal detection with minimal data loss.
- Multimode optical fiber (FG200LEA): Core diameter of 200  $\mu\text{m}$ , with a step-index refractive structure.

During operation, the seed sample is irradiated by the LED, inducing fluorescence emission. The emitted signal is collected at a 45° angle using the rod lens, transmitted through the optical fiber, and detected by the CMOS sensor.

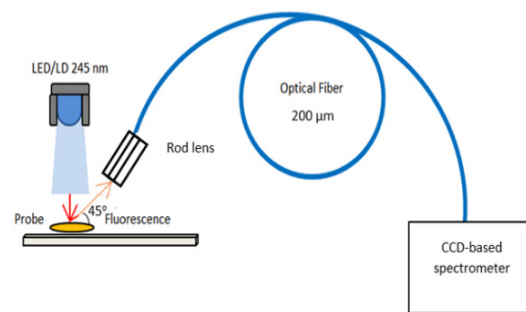
This system offers several advantages:

1. Improved transmission efficiency through the inclusion of a rod lens, reducing air gaps between optical elements.

2. Optimized fiber coupling design, achieved with a durable housing that minimizes signal loss.
3. Efficient signal collection at 45°, enhancing sensitivity and accuracy of fluorescence detection.



**Figure 1.** Averaged spectral distributions of seeds of the varieties Espresso F1, French breakfast, Nacional 2 and Red large



**Figure 2.** Mobile experimental installation used for fluorescence spectroscopy

No signal normalization is required for spectral analysis processing. The software used to generate the spectra from the spectrometer used in the scheme of Figure 2 allows for editable data acquisition parameters per channel, such as detector integration time, automatic dark current correction, signal averaging, spline interpolation, and spectral smoothing. The software was chosen because of its ability to generate time series in which the output of user-defined functions, integrals, peaks (intensity, wavelength) can be tracked simultaneously over time.

### RADISH SEEDS CLASSIFICATIONS

In the classification of the radish seeds (French breakfast, Nacional 2, Espresso F1 and Red large) obtained by using fluorescence spectroscopy techniques, machine learning (ML) algorithms such as Fine Tree, Quadratic Support Vector Machine (SVM), Fine K Nearest Neighbor (KNN) and Neural Network (NN) were applied. During the application process of these algorithms cross-validation

approach was used. Then, 10-fold cross-validation approach and ML algorithms were applied to make classification and analysis process more objective and free from randomness. All algorithms were implemented in MATLAB R2022b using the built-in Classification Learner Toolbox. The default hyperparameter settings were applied as follows: for classification, several machine learning models were employed using preset default configurations to maintain reproducibility and avoid manual tuning of hyperparameters. The Fine Decision Tree was limited to a maximum of 100 splits, using Gini's diversity index to determine node splits. The Quadratic SVM utilized a quadratic kernel, with a box constraint (C) value of 1, while the kernel scale was automatically calculated. The Fine k-Nearest Neighbor (KNN) model was configured with a single neighbor (k=1), Euclidean distance measurement, and uniform weighting for all neighbors. The Neural Network (NN) consisted of a single hidden layer containing 10 neurons, applying the tan-sigmoid activation function and trained using the scaled conjugate gradient method.

Model evaluation was carried out using a 10-fold cross-validation approach. The dataset was divided into ten equal segments; in each iteration, nine segments were used for model training and one for testing. This cycle was repeated ten times so that each segment served as a test set once. The final performance metric was obtained by averaging accuracy results across all folds, ensuring comprehensive utilization of the dataset for both training and testing. The cross-validation procedure applied in this study is illustrated in Figure 3.



Figure 3. k folds Cross-Validation

In the training and testing processes, the selected algorithms provided the most satisfactory results. Generally, the performance of the classification algorithms is measured using classification metrics contained in the confusion matrix (Figure 4) such as accuracy, sensitivity, specificity and F1-score given in equations (1) to (4), respectively. From these metrics, only the accuracy is used in this work to evaluate the performance of the used algorithms.

The accuracy value is the measure of prediction made by the algorithm and its value is obtained using equation (1) given below. In this equation, TP means True Positive, FN means False Negative, FP means False Positive, and TN means True Negative.

Multi Class Confusion Matrix		Predicted Class			
		Class <sub>1</sub>	Class <sub>2</sub>	.....	Class <sub>N</sub>
Actual Class	Class <sub>1</sub>	C <sub>1,1</sub>	FP <sub>2,1</sub>	.....	C <sub>1,N</sub>
	Class <sub>2</sub>	FN <sub>2,1</sub>	TP <sub>2,2</sub>	.....	FN <sub>2,N</sub>
	.....	.....	.....	.....	.....
	Class <sub>N</sub>	C <sub>N,1</sub>	FP <sub>2,N</sub>	.....	C <sub>N,N</sub>

Figure 4. Example of multiclass confusion matrix [16].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} * 100 \quad (1)$$

$$Sensitivity = \frac{TP}{(TP+FN)} * 100 \quad (2)$$

$$Precision = \frac{TP}{(TP+FP)} * 100 \quad (3)$$

$$F1 - Score = \frac{(2TP)}{(2TP+FP+FN)} * 100 \quad (4)$$

From the given metrics, only the classification accuracy was used to evaluate the performances of the each algorithm. The applied procedure was used to classify radish seed variety based on fluorescence spectroscopic data using machine learning algorithms (Figure 5).

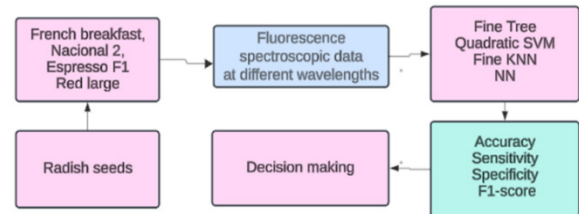


Figure 5. Procedure applied to classify radish seeds.

## RESULTS AND DISCUSSION

This study presents the analysis of radish seed varieties using fluorescence spectroscopic data combined with machine learning algorithms. Four classification models: Fine Tree, Quadratic SVM, Fine KNN, and Neural Network were applied to distinguish between different radish varieties, with classification accuracy serving as the primary performance metric.

Prior to classification, the raw spectral data were pre-processed to enhance signal quality and comparability across samples. Baseline correction and spectral normalization were implemented to remove intensity variations caused by measurement

conditions, while a smoothing filter was applied to reduce random noise. From the pre-processed spectra, it was observed that the wavelength region between approximately 475 and 555 nm contained the most discriminative information for distinguishing seed varieties.

True Class	1	1333	4	2	14
	2	1	1293	53	6
	3	4	38	1303	8
	4	9	9	16	1310
		1	2	3	4
		Predicted Class			

Figure 6a. Confusion matrix results - Fine Tree.

True Class	1	1353	0	0	0
	2	0	1353	0	0
	3	0	0	1353	0
	4	0	0	0	1353
		1	2	3	4
		Predicted Class			

Figure 6b. Confusion matrix results - SVM.

True Class	1	1351	0	0	2
	2	0	1329	19	5
	3	0	39	1309	5
	4	1	7	0	1345
		1	2	3	4
		Predicted Class			

Figure 6c. Confusion matrix results - KNN.

True Class	1	1351	0	0	2
	2	0	1329	19	5
	3	0	39	1309	5
	4	1	7	0	1345
		1	2	3	4
		Predicted Class			

Figure 6d. Confusion matrix results - Neural Network.

This interval corresponded to the region where emission spectra exhibited the greatest variance among cultivars, whereas wavelengths outside this range contained largely redundant or noisy information. For this reason, the 475–555 nm interval was selected to maximize classification accuracy while minimizing computational complexity.

In the classification process, machine learning techniques such as Fine Tree, Quadratic SVM, Fine KNN, and Neural Network models were used to classify radish varieties. The implementation of these models was carried out using Matlab R2022b software. Classification of fluorescence spectroscopy data measured from radish seeds was performed with 10-fold cross-validation using standard machine learning parameters. The confusion matrices summarizing the classification results for each model are presented in Figures 6a–d.

The classification performance of four machine learning models Fine Tree, Quadratic SVM, K-Nearest Neighbors (KNN), and Neural Network was systematically evaluated using the metrics accuracy, precision, sensitivity (recall), and F1-score, derived from the confusion matrices (Tables 1–4). The averaged performance metrics for each classifier are summarized in Table 5.

Among the evaluated models, the Quadratic SVM classifier achieved flawless classification, attaining 100% accuracy, precision, sensitivity, and F1-score across all classes. This result indicates that SVM effectively captures the discriminative features present in the fluorescence spectra of radish seed varieties, providing highly reliable and consistent predictions. Similarly, the Neural Network model exhibited near-optimal performance, with an overall accuracy of 99.28%, precision of 98.57%, sensitivity of 98.56%, and F1-score of 98.56%. The slightly lower but still excellent performance of the Neural Network suggests its strong capacity to model complex nonlinear relationships inherent in the spectral data.

Table 1. Classification performance results of the Fine Tree model

	Accuracy	Precision	Sensitivity	F1-Score
Class 1	89.91	76.92	80	78.43
Class 2	86.24	69.23	72	70.59
Class 3	85.32	73.33	73.33	73.33
Class 4	94.5	92.59	86.21	89.29

Table 2. Classification performance results of the SVM model

	Accuracy	Precision	Sensitivity	F1-Score
Class 1	100	100	100	100
Class 2	100	100	100	100
Class 3	100	100	100	100
Class 4	100	100	100	100

**Table 3.** Classification performance results of the KNN model

	Accuracy	Precision	Sensitivity	F1-Score
Class 1	\	98.90	99.70	99.30
Class 2	93.46	87.25	86.47	86.86
Class 3	93.59	86.77	87.73	87.25
Class 4	99.37	99.25	98.23	98.74

**Table 4.** Classification performance results of the Neural Network model

	Accuracy	Precision	Sensitivity	F1-Score
Class 1	99.94	99.93	99.85	99.89
Class 2	98.71	96.65	98.23	97.43
Class 3	98.84	98.57	96.75	97.65
Class 4	99.63	99.12	99.41	99.26

**Table 5.** Classification performance metrics using standard parameter values

Classifier Model	Accuracy	Precision	Sensitivity	F1-Score
Fine Tree	88.99	78.02	77.89	77.91
Quadratic SVM	100.00	100.00	100.00	100.00
Fine KNN	96.52	93.04	93.03	93.04
Neural Network	99.28	98.57	98.56	98.56

The KNN classifier also demonstrated robust performance, with 96.52% accuracy, 93.04% precision, 93.03% sensitivity, and 93.04% F1-score. While marginally lower than SVM and NN, these results confirm that KNN is capable of accurate classification, albeit with slightly reduced discrimination in cases of overlapping spectral features.

In contrast, the Fine Tree classifier presented comparatively lower performance, achieving 88.99% accuracy, 78.02% precision, 77.88% sensitivity, and 77.91% F1-score. Although the Fine Tree successfully distinguished between the four radish seed classes, its lower precision and sensitivity values highlight potential limitations in handling subtle spectral variations, indicating that tree-based methods may be less suited for highly complex spectral datasets.

Overall, these findings underscore the superiority of advanced classifiers, particularly Quadratic SVM and Neural Networks, in providing both high accuracy and robustness in the classification of radish seed varieties. The results also suggest that model selection is critical for achieving reliable performance, with nonlinear and kernel-based

approaches outperforming simpler, tree-based algorithms when dealing with high-dimensional spectral data.

## CONCLUSION

This study demonstrates the successful application of fluorescence spectroscopy, combined with machine learning algorithms, for the non-invasive classification of radish seeds. Among the evaluated models, the SVM and Neural Network classifiers exhibited superior performance, achieving nearly perfect accuracy, precision, sensitivity, and F1-score, while KNN also provided robust results. The Fine Tree classifier, although less accurate, still achieved a reasonable classification, highlighting the importance of model selection for handling complex spectral data.

The novelty of this work lies in applying optoelectronic analysis specifically to radish seeds, a crop for which such an approach has not been previously reported. The proposed methodology enables rapid, accurate, and non-destructive evaluation of seed quality, reducing reliance on traditional germination tests and supporting informed decisions in seed selection and crop management.

Future research should focus on expanding the dataset to include more diverse radish varieties and growth conditions to enhance model generalizability. Additionally, integrating advanced deep learning techniques and developing portable, sensor-based systems could facilitate real-time, field-level applications. These improvements would further strengthen the reliability of seed quality assessment and support the adoption of smart agricultural technologies in precision farming.

## REFERENCES

1. A. Kjær, J. Øgaard Madsen, Y. Maeda, Y. Ozawa, Y. Uda, *Agricultural and Biological Chemistry*, **42**, 1721 (1978).
2. Y. Kaneko, C. Kimizuka-Takagi, S. Bang, Y. Matsuzawa, *Vegetables*, 160 (2007).
3. R. M. P. Gutiérrez, R. L. Perez, *The Scientific World Journal*, **4**, 837 (2004).
4. K. M. Bae, S.C. Sim, J.H. Hong, K.J. Choi, D.H. Kim, Y.S. Kwon, *Horticulture, Environment, and Biotechnology*, **56**, 224 (2015).
5. Z. Huyan, S. Ding, X. Liu, Y. Yu, *Analytical Methods*, **10**, 3207 (2018).
6. V.M. Belyakov, *Inzhenernye Tekhnologii i Sistemy*, **2**, 319 (2019).
7. V.K. Rewatkar, *Romanian Journal of Biophysics*, **30**(3) (2020).
8. W. Zhang, R. Lv, Y. Sun, H. Gu, *Agricultural Sciences*, **12**, 574 (2021).

9. S.P. Singh, K. Vogel-Mikuš, P. Vavpetič, L. Jeromel, P. Pelicon, J. Kumar, R. Tuli, *Planta*, **240**, 289 (2024).
10. M. Belyakov, E. Sokolova, V. Listratenkova, N. Ruzanova, L. Kashko, in: *E3S Web of Conferences*, **273**, 1003 (2021).
11. B.V. Olegovna, D.V. Vitalyevna, D.A. Grigoryevna, *International Scientific Review*, **4**, 31 (2017).
12. W.H. Su, S.A. Fennimore, D.C. Slaughter, *Biosystems Engineering*, **186**, 167 (2019).
13. C. Li, X. Wang, Z. Meng, *Optical Design and Testing VII*, **10021**, 516 (2016).
14. J. Boothe, C. Nykiforuk, Y. Shen, S. Zaplachinski, S. Szarka, P. Kuhlman, M. Moloney, *Plant Biotechnology Journal*, **8**, 606 (2010).
15. A. Yasar, V. Slavova, S. Genova, *Tehnički Glasnik*, **18**, 531(2024).
16. A. Golcuk, A. Yasar, *Journal of Food Composition and Analysis*, **119**, 105253 (2023).