

Different Chemometric Approaches to Detect Adulteration of Cold-Pressed Flaxseed oil with Refined Rapeseed Oil Using Differential Scanning Calorimetry

[Mahbuba Islam](#) , [Anna Kaczmarek](#) , [Magdalena Montowska](#) , [Jolanta Tomaszewska-Gras](#) *

Posted Date: 20 July 2023

doi: 10.20944/preprints202307.1441.v1

Keywords: DSC melting profile; Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA); Artificial neural networks (ANN); Multiple Linear Regression (MLR); MARS; SVM; Food fraud; Oils adulteration



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Different Chemometric Approaches to Detect Adulteration of Cold-Pressed Flaxseed oil with Refined Rapeseed Oil Using Differential Scanning Calorimetry

Mahbuba Islam ¹, Anna Kaczmarek ¹, Magdalena Montowska ² and Jolanta Tomaszewska-Gras ^{1*}

¹ Department of Food Quality and Safety Management, Poznań University of Life Sciences, ul. Wojska Polskiego 31/33, 60-624 Poznań, Poland; mahbuba.islam@up.poznan.pl (M.I.); anna.kaczmarek@up.poznan.pl (A.K.); jolanta.tomaszewska-gras@up.poznan.pl (J.T.-G.)

² Department of Meat Technology, Poznań University of Life Sciences, ul. Wojska Polskiego 31/33, 60-624 Poznań, Poland; magdalena.montowska@up.poznan.pl (M.M.)

* Correspondence: jolanta.tomaszewska-gras@up.poznan.pl

Abstract: Flaxseed oil is one of the best sources of *n*-3 fatty acids, thus its adulteration with refined oils can lead to a reduction in its nutritional value and overall quality. The purpose of this study was to use the differential scanning calorimetry (DSC) technique to detect adulterations of cold-pressed flaxseed oil with refined rapeseed oil (RP). Based on the melting phase transition curve, parameters such as peak temperature (T), peak height (h), and percentage of area (P) were determined for pure and adulterated flaxseed oils with a RP concentration of 5, 10, 20, 30, 50% (w/w). Significant linear correlations ($p \leq 0.05$) between the RP concentration and all DSC parameters were observed, except for h1. In order to assess the usefulness of the DSC technique for detecting adulterations, three chemometric approaches were compared: 1) classification models (Linear Discriminant Analysis, LDA Adaptive Regression Splines, MARS, Support Vector Machine, SVM, Artificial Neural Networks, ANNs); 2) regression models (Multiple Linear Regression, MLR, MARS, SVM, ANNs, PLS) and 3) a combined model of Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA). With the LDA model, the highest accuracy of 99.5% in classifying the samples, followed by ANN > SVM > MARS was achieved. Among the regression models, the ANN model showed the highest correlation between observed and predicted values ($R = 0.996$), while other models showed goodness of fit as following MARS > SVM > MLR. Comparing OPLS-DA and PLS methods, higher values of $R^2X(\text{cum}) = 0.986$ and $Q^2 = 0.973$ were observed with the PLS model than OPLS-DA. These results demonstrate the usefulness of the DSC technique combined with chemometrics for predicting the adulteration of cold-pressed flaxseed oil with refined rapeseed oil.

Keywords: DSC melting profile; chemometric analysis; oil adulteration; plant oils; regression analysis; classification model; artificial neural networks

1. Introduction

Recent research has found persuasive evidence that combining data fusion techniques with chemometric methodologies can produce extraordinary results in the assessment and classification of food quality. Researchers have proved the considerable potential of data fusion-multivariate statistical analysis in the field of food authenticity research by merging multiple sources of data and employing advanced statistical analysis [1–3]. This novel approach provides for a more thorough and accurate evaluation of foodstuffs, allowing for the detection of any adulteration [4] or mislabeling. These findings highlight the need to combine data and use multivariate statistical analysis tools to verify the authenticity and quality in the food supply chain. Adulteration of high-priced edible oils continues to be a major concern for both the edible oil industry and consumer health, despite the fact that experts recognized the problem centuries ago [5]. This deceptive technique is primarily

motivated by individuals seeking to increase their revenues by boosting the volume of the product [6] by taking advantage of the lack of effective quality assessment tools for unreliable food products, as suggested by the Food and Agricultural Organization (FAO) [7]. Food adulteration occurs frequently in many wealthy countries around the world, despite the researchers' attention to the fraudulent phenomenon [8]. Addressing food fraud poses a significant financial burden, with estimates suggesting that the global food industry incurs an annual cost of approximately EUR 30 billion. This substantial expense highlights the economic impact of combating fraudulent practices in the food sector [9].

Several approaches to investigating and detecting adulterations in food products have been proposed by various food scientists. Researchers have shown the successful application of combining analytical experimental results with other linear and non-linear chemometric tools [10] to build classification and regression models for oil samples i.e., linear discriminant analysis (LDA)[11], multiple linear regression (MLR)[12], multivariate adaptive regression splines (MARS) [13], support vector machine (SVM) [14], artificial neural networks (ANNs) [15], principle component analysis (PCA) [16], orthogonal partial least squares discriminant analysis (OPLS-DA) [17] and partial least squares regression (PLS) [18]. For example, to detect adulteration in extra virgin olive oils, UV-IMS (ultraviolet ion mobility spectrometry) combined with chemometric analysis like PCA and LDA [19], near-infrared spectroscopy with chemometric techniques [20], and DSC combined with SVM [21]. As in this study, the detection of adulteration in flaxseed oil has also been reported by other authors using different analytical methods coupled with statistical approach, e.g., mid-infrared spectroscopy (MIR) associated with chemometric technique of PLS [22], low-field nuclear magnetic resonance relaxation fingerprints [23], gas chromatography-mass spectrometry (GC-MS) coupled with PCA and recursive support vector machine (R-SVM) [24], HPLC-ELSD profiling of triacylglycerols and chemometrics [25], dielectric spectroscopy with PCA and LDA analysis [11], Fourier transform infrared spectroscopy (FTIR) and MLR [26]. From all the references mentioned above, the importance of coupling chemometric study for describing the multiple variables obtained from analysis to classify the adulteration detection efficiently was found in most of the aforementioned studies, along with other studies [2,27,28]. Different studies were conducted to show the applicability of using DSC for the adulteration assessment for different fats and oils which are comparatively expensive and acclaimed as being nutritious by scientists, e.g., olive oils and other vegetable oils [29–34], animal fats [35–38].

Derived from the flax plant (*Linum usitatissimum* L.), flaxseed is a seed that is widely grown in countries like Canada, America, China, and India [39]. The oil obtained from cold-pressing flaxseeds, known as flaxseed oil, is highly regarded for its exceptional content of α -linolenic acid (ALA) [40,41], an essential fatty acid that can be converted into beneficial compounds like eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) in the human body [42]. Additionally, flaxseed contains abundant phenolic compounds such as lignans, ferulic acid, and p-coumaric acid, as well as mucilage. These bioactive components have demonstrated positive effects on intestinal function [43]. Flaxseed oil has been found to offer numerous health benefits, particularly for the cardiovascular and skeletal systems. It has also shown positive effects in inflammatory conditions like rheumatoid arthritis, psoriasis, ulcerative colitis and colon tumor [44,45]. Authenticity analysis is crucial for cold-pressed oils like flaxseed oil, as they are commonly targeted by fraudulent practices, such as adulteration with lower-quality oils or blending with cheaper oils. Refined rapeseed oil can be used as an adulterant to flaxseed oil, since it is cheaper and widely used as cooking oil known for its neutral flavor, high smoke point, and longer shelf life. Adulteration of flaxseed oil with refined rapeseed oil compromises its authenticity and can lead to a reduction in its nutritional value and overall quality [46]. Therefore, the aim of the study was to study the feasibility of the DSC technique combined with chemometric methods for the detection of adulteration of cold-pressed flaxseed with refined rapeseed oils in different concentrations. The novelty of the study lies in evaluating the potential of DSC coupled with various chemometric methods, which were employed to create predictive classification and regression models for the detection of oil adulteration. The classification approach categorized the level of oil adulteration, while the regression approach treated the concentration of

refined rapeseed oils as a continuous variable. A number of chemometric methods, i.e., LDA, MLR, MARS, SVM, ANNs, PCA, OPLS-DA, PLS were used to classify, describe and generate prediction models of the adulteration phenomena. This approach evaluates the significance of combining data fusion techniques with chemometric methodologies for assessing the authenticity of cold-pressed oil.

2. Materials and Methods

2.1. Materials

Sample oil seeds for cold-pressed flaxseed oils were obtained from different Polish cultivars and then pressed mechanically. Seeds from four different cultivars were purchased i.e. *Bukoz* (Polish Institute of Natural Fibers and Medicinal Plants in Poznan), *Dolguniec*, *Szafir* (SEMCO manufactory in Śmiłowo, the Hodowla Roślin Strzelce Sp. z o.o. in Strzelce), and one sample of an unidentified variety from the VitaCorn company in Poznan. The oils were obtained through screw-pressing the seeds while keeping temperature below 50 °C. After the pressing process, the oils were left for 24 hours for decantation, and subsequently stored in brown glass bottles. Refined rapeseed oil was purchased from the market. Cold-pressed flaxseed oil samples were adulterated by adding refined rapeseed oils in varying concentrations (0, 5, 10, 20, 30, 50% w/w). Prepared samples were analyzed in three replications.

2.2. Methods

2.2.1. Melting phase transition analysis by differential scanning calorimetry (DSC)

Melting analysis of oil samples was carried out with modifications according to the method used for butterfat [47]. A Perkin Elmer differential scanning calorimeter DSC 8500 PerkinElmer (Waltham, Massachusetts, USA), equipped with an Intracooler II and running with Pyris software (Perkin Elmer, Waltham, Massachusetts, USA), was used. Nitrogen (99.999% purity) was the purge gas. Samples of ca. 6-7 mg were weighed into aluminum pans of 20 µL (Perkin Elmer, No. 0219-0062, Waltham, Massachusetts, USA) and hermetically sealed. An empty, hermetically sealed aluminum pan used as reference. Analysis started with cooling the oil sample at a scanning rate of 2 °C/min from a temperature of 30 °C to -65 °C, after which it was heated at scanning rates 5 °C/min from -65 °C to 30 °C. For each measurement at a given scanning rate, the calibration procedure was completed with the correct scanning rate. After the analysis, the DSC files were converted to the ASCII format and then were analyzed using Origin Pro software, version 2023 (OriginLab Corporation, Northampton, MA, USA). The Origin PeakFit module was used to project the DSC curves of all investigated samples. Different DSC parameters i.e., peak temperature (T , °C), peak height (h , W/g), and percentage of area (P , %) were measured from the melting curves. Peak temperature (T) was determined on the temperature axis (X) by locating the maximum point of heat flow for each peak. Peak height (h) was determined at the heat flow maximum on the axis Y for each peak. The percentage of each peak area (P) was calculated as the ratio of the area of each peak to the total area of the melting phase transition curve.

2.2.2. Data analysis

The recorded data was subjected to statistical analysis using Statistica 13.3 software, developed by TIBCO Software Inc. in the USA. A significance level of $\alpha=0.05$ was chosen for the analysis. The outcomes were reported as the mean and standard deviation. The initial step in the statistical analysis involved assessing the assumptions of ANOVA, which included testing for variance homogeneity and checking data normality. If these assumptions were met, one-way analysis of variance (ANOVA) was performed, followed by the application of Tukey's test to establish statistically homogeneous groups. To assess the impact of adding adulterants at varying percentages to cold-pressed flaxseed oils, linear regression analysis was employed, utilizing the least squares estimation method. This analysis considered the DSC parameters extracted from the melting curves. Classification and regression approaches were used to build predictive models of oil adulteration. For all models, the

predictors were DSC parameters, while the dependent variable was the level of oil adulteration. In the classification approach, the dependent variable (level of oil adulteration) was categorical (6 classes - one for each level of oil adulteration), while in the regression approach the dependent variable was a continuous variable (concentration of refined rapeseed oils). Artificial Neural Networks (ANNs), Support Vector Machine (SVM) and Multivariate Adaptive Regression Splines (MARS) were used to build classification as well as regression models. In addition, Linear Discriminant Analysis (LDA) was used to build a classification model and Multiple Linear Regression analysis (MLR) to build a regression model. The goodness of fit of the regression models was estimated based on R (correlation coefficient), R^2 (determination coefficient), adjusted R^2 (modified version of R^2 was adjusted for the number of predictors in the model), Akaike information criterion (AIC), Bayesian information criterion (BIC) and Root Mean Square Error (RMSE). The confusion matrix was used for selecting the best classification model. The confusion matrix represents the counts of predicted and true values. The score "TN" stands for True Negative, which shows the number of correctly classified negative examples. Similarly, "TP" stands for True Positive, which indicates the number of correctly classified positive examples. "FP" stands for False Positive, which is the number of actual negative examples classified as positive; and "FN" stands for False Negative, which is the number of actual positive examples classified as negative. Based on the confusion matrix, the performance parameters (Accuracy, Misclassification Rate, Precision, Sensitivity, Specificity and F1-score) of the classification models were calculated. To perform PCA and OPLS-DA the SIMCA software version 16.1 (Sartorius Stedim Data Analytics AB, Umea, Sweden) was utilized. Cross-validation was performed on both the PCA and OPLS-DA models, and the OPLS-DA models were validated using permutation testing. For these models, DSC parameters were considered as X variables, and Y variables consisted of different levels of concentrations added as adulterants to the flaxseed oils. The metrics (R^2_X , R^2 and Q^2) were analyzed from the models to collectively provide information about how well the OPLS-DA model fits the X data (DSC parameters).

3. Results

3.1. DSC melting profiles of cold-pressed flaxseed oil adulterated with refined rapeseed oil

In Figure 1, the DSC melting curves obtained for all cultivars of flaxseed oils (pure and adulterated) are presented. Each Figure (1a, 1b, 1c, and 1d) demonstrates the melting curves for different cultivars of flaxseed oils (*Bukoz*, *Unknown*, *Dolguniec* and *Szafir* variety, respectively) with different concentrations of refined rapeseed oil. In this study, the thermal profile of flaxseed oils was examined using melting DSC curves to track the alterations resulting from the addition of adulterants (rapeseed oil). Differential scanning calorimetry analysis provided unique and substantial information regarding the thermal profiles of pure cold-pressed flaxseed oils and adulterated with refined rapeseed oils in different concentrations (0, 5, 10, 20, 30, 50% w/w). All the samples were first crystallized to -65°C at a $2^\circ\text{C}/\text{min}$ cooling rate prior to the heating program. On the melting curve of pure flaxseed oil, three endothermic peaks were identified as a result of the melting of crystals and nuclei.

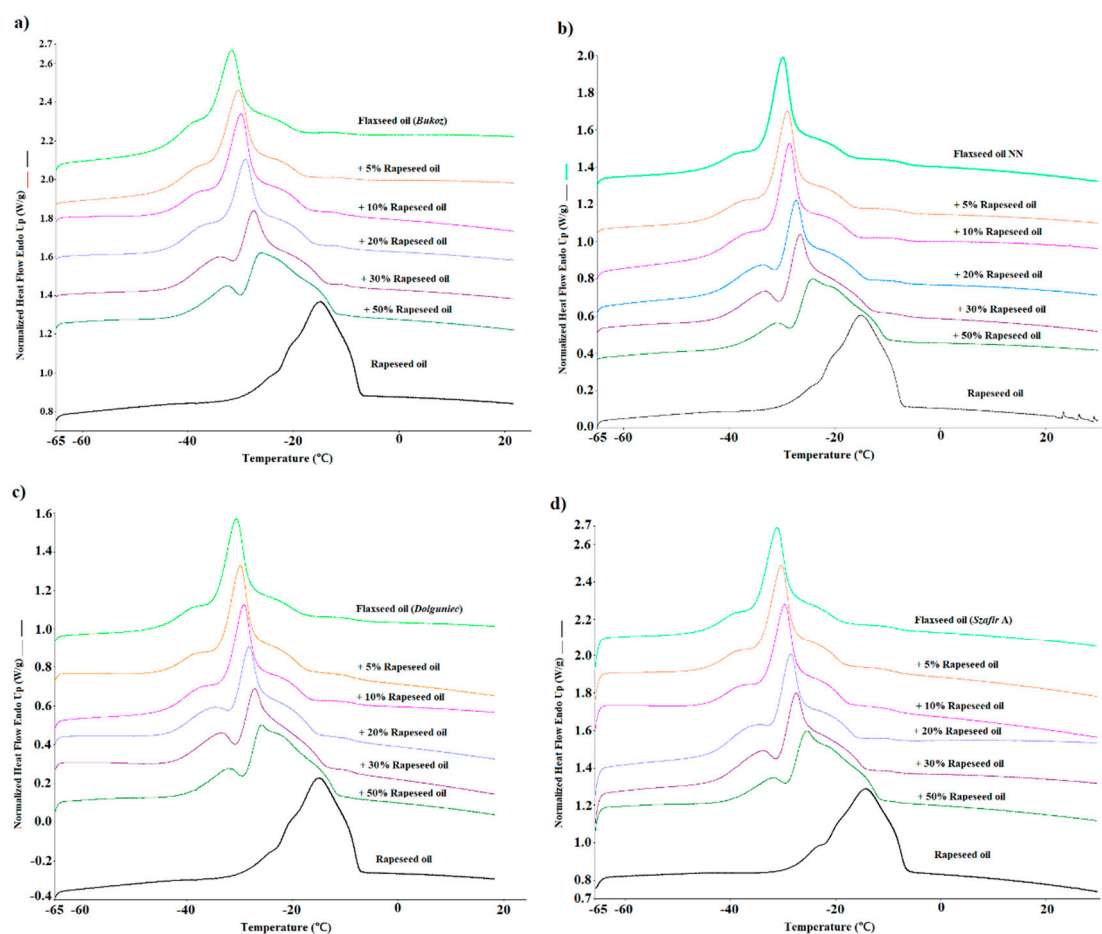


Figure 1. DSC melting curves obtained at a 5 °C/min heating rate for different cultivars of cold-pressed flaxseed oils adulterated with refined rapeseed oils in different concentrations 0, 5, 10, 20, 30, 50% w/w. (a) Bukoz cultivar, (b) Unknown cultivar, (c) Dolguniec cultivar, (d) Szafir cultivar.

The first shoulder peak was detected at around -36°C , the second, as a major peak, occurred at around -30°C and the third peak appeared at a temperature of approx. -25°C (Figure 1). Apart from the flaxseed oil, the melting curves of refined rapeseed oil, which gives two endothermic peaks, are also shown. Between the DSC curve for flaxseed oil and refined rapeseed oil curves, the curves of adulterated flaxseed oil with rapeseed in different concentrations are shown. Comparing all figures (1 a, 1b, 1c, 1d), it can be seen that all the varieties have shown similar changes of thermal transition upon the addition of refined rapeseed oil. With an increasing refined oil concentration, gradual changes in the formation of all peaks can be observed in Figure 1. The addition of rapeseed oil caused all peaks to shift to higher temperatures. Every flaxseed oil variety showed a similar thermodynamic alteration in the melting curve. Apparently, all three peaks' positions shifted to a higher temperature with the gradual addition of rapeseed oil. Consequently, the peak height also changed, as the major peak height dropped, while the height of the shoulder peaks (first and third peak) rose with an increasing concentration of rapeseed oils in the mixture.

Alterations in the thermodynamic characteristics of the target samples with the addition of adulterants have been reported in other studies as well. A similar approach to adulterating cold-pressed with refined and cheaper oils was taken by other authors studying adulteration. For instance, using mid-infrared spectroscopy (MIR), authors performed a quantitative analysis of soybean oil and sunflower oil as adulterants with concentrations from 3.5 to 30% (w/w) in extra virgin flaxseed oil [22]. Another study showed the possibility of detecting the adulteration of flaxseed oil samples by various concentrations of sunflower oil (10, 20, 30, 50, 70, 90% v/v) using magnetic resonance

fingerprinting (MRF) [23]. Successful adulteration analysis has also been performed using the DSC technique for butterfat adulterated with different concentrations of refined palm oil (0, 2, 5, 10, 15, 20, 25, 30 and 35% w/w) [47]. The addition of refined palm oil was manifested in changed melting parameters of peak temperature, area and height of peaks [47]. Another study considered both the cooling and heating profiles obtained from DSC to analyze the changes occurring for beef and chicken fat contaminated with lard (0.5 - 5%), where the heating profiles turned out to be more efficient for detecting adulteration than the crystallization profiles [48].

In Table 1, changes in DSC parameters are presented within an increasing concentration of refined rapeseed oil. Anova analysis was performed to show significant differences between six concentrations of refined rapeseed oil for all the parameters measured, i.e. peak temperature (T1, T2 and T3); peak height (h1, h2 and h3) and the percentage of the peak area (P1, P2 and P3). Considering the changes in peak temperatures resulting from a 0 to 50% addition, it can be seen that the first peak (T1) shifted from -36 to -32°C, the second from -31 to -25°C and the third peak from -25 to -19°C. In contrast, peak height (h) did not change in the same way for all peaks. The first (h1) and third (h3) peaks increased with the addition of an adulterant, while the main peak, the second peak (h2), reduced from a value 0.6 W/g for pure flaxseed oil to 0.37 for a sample with 50% refined rapeseed oil and (Table 1). Similarly, different characteristics were exhibited by the percentage of area parameter (P), where P3 was increased while P1 and P2 decreased with the addition of refined oil.

Table 1. DSC thermodynamic parameters of melting phases for cold-pressed flaxseed oils adulterated with different concentrations of refined rapeseed oils.

DSC parameters	Concentrations					
	0%	5%	10%	20%	30%	50%
T1	-36.15 ± 0.35 ^a	-35.44 ± 0.23 ^{bc}	-34.95 ± 0.35 ^{bc}	-35.32 ± 0.36 ^b	-34.57 ± 0.59 ^c	-32.39 ± 0.6 ^d
T2	-30.67 ± 0.69 ^a	-30.26 ± 0.67 ^{ab}	-29.55 ± 0.68 ^b	-28.31 ± 0.54 ^c	-27.08 ± 0.61 ^d	-24.78 ± 0.77 ^e
T3	-24.73 ± 1.04 ^a	-24.08 ± 0.84 ^{ab}	-23.44 ± 0.72 ^{bc}	-22.92 ± 0.81 ^c	-22.53 ± 0.48 ^c	-19.12 ± 1.07 ^d
h1	0.12 ± 0.02 ^a	0.14 ± 0.02 ^{abc}	0.14 ± 0.02 ^{abc}	0.14 ± 0.01 ^{bc}	0.15 ± 0.01 ^c	0.13 ± 0.02 ^{ab}
h2	0.60 ± 0.03 ^e	0.57 ± 0.03 ^{de}	0.55 ± 0.02 ^d	0.50 ± 0.02 ^c	0.44 ± 0.03 ^b	0.37 ± 0.02 ^a
h3	0.16 ± 0.01 ^a	0.16 ± 0.01 ^a	0.17 ± 0.01 ^a	0.21 ± 0.02 ^b	0.23 ± 0.02 ^c	0.27 ± 0.02 ^d
P1	24.12 ± 3.58 ^{bc}	27.50 ± 4.00 ^c	27.55 ± 3.17 ^c	24.41 ± 3.04 ^{bc}	23.09 ± 1.74 ^b	17.08 ± 1.84 ^a
P2	39.77 ± 2.16 ^c	38.32 ± 1.96 ^c	37.49 ± 2.51 ^c	34.52 ± 2.35 ^b	32.43 ± 1.69 ^{ab}	30.27 ± 2.00 ^a
P3	30.49 ± 2.57 ^a	28.78 ± 3.07 ^a	30.75 ± 2.31 ^a	38.3 ± 2.52 ^b	41.5 ± 2.26 ^b	48.71 ± 2.96 ^c

All values are mean ± standard deviation (n=10), (a-e) - means with the same letters within the column are not different (*p* > 0.05). T1, T2, T3 represent the first, second and third peak temperatures, respectively; h1, h2, and h3 means peak height for the first, second and third peak, respectively; P1, P2, P3 represent percentage of peak area for first, second and third peak, respectively.

3.2. Changes in DSC parameters of flaxseed oil melting phase transition depending on adulterant concentrations

Parameters determined from DSC curve for three peaks i.e. peak temperature (T), peak height (h), and percentage of peak area (P) versus rapeseed oil concertation were analyzed by linear regression (Figure 2). Linear regression analysis was also used by other authors to explain adulteration phenomena [34,47,49]. Thus, in this study, all variables were analyzed with linear regression model to find out the trend of changes in DSC parameters resulting from the addition of the adulterant. The data in Figure 2 (a) indicates that peak temperatures always rose to higher temperatures linearly with an increasing quantity of refined rapeseed oil added to the target oil, and it can be seen that all changes were significant (*p* ≤ 0.05). Among the three peaks, the strongest correlation was observed for the second peak, T2 (*r*= 0.95). In Figure 2 (b), for the first peak, the parameter of peak height (h1) is comparatively stable (slightly increased, which is not statistically significant at that level (*p* > 0.05) upon addition of refined oils. In contrast, for the third peak, h3 increased significantly (*p* ≤ 0.05), and correlated strongly with the concentration of adulteration (*r*= 0.92). On the other hand, for the second peak, h2 was significantly lowered with the increment of added adulterants (50%), with strong negative correlation, *r*= -0.96. The parameters for percentage of area calculated for each peak P1, P2 and P3 were also plotted against concentration data. Clearly, the

first and second peaks' area proportion to the total melting curve decreased significantly, while the percentage area of the third peak increased significantly ($p \leq 0.05$), with a strong correlation to the adulterant concentration ($r=0.92$).

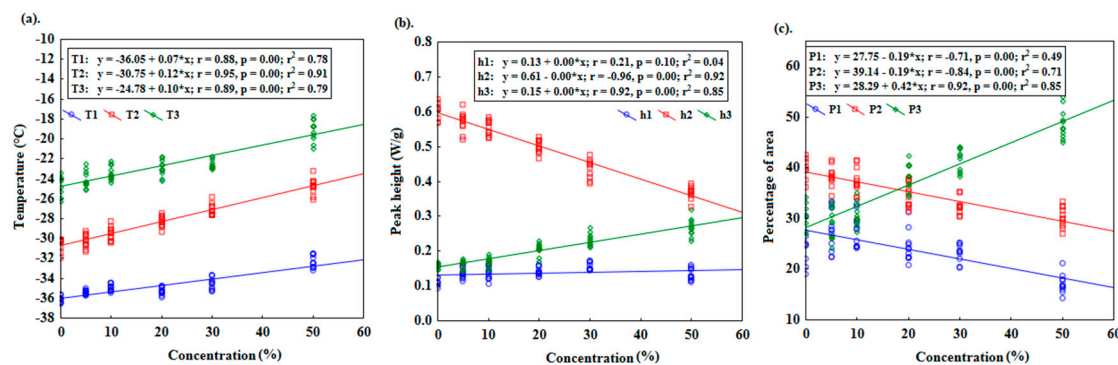


Figure 2. Regression analysis of adulterated flaxseed oil samples for DSC parameters **(a).** peak temperature (T, °C), **(b).** peak height (h, W/g), and **(c).** percentage of area (P).

4. Discussion

4.1. Classification models for predicting cold-pressed flaxseed oil adulteration levels

In order to assess the ability to build models that classify adulterated oil samples into appropriate classes, the LDA, MARS, SVM and ANNs methods were used. Linear discriminant analysis (LDA) was used to build the first classification model. LDA and the related Fisher's linear discriminant (FLD) are used in machine learning to find the linear combination of features that best distinguish between two or more classes of objects. The resulting combinations are used as a linear classifier. Discriminant analysis resulted in a statistically significant model with Wilks' Lambda = 0.00119 and $p \leq 0.05$. All variables (DSC parameters) except h3 have significant statistical discriminant power. From our study, five discrimination function were obtained based on Wilk's Lambda statistics, with $p \leq 0.05$ for the first three functions. For purposes of classifying the cases, six classification functions were calculated. Each classification function represents a linear equation that combines the input variables (DSC parameters) to discriminate between six groups (G_1 to G_6) and thus provides different classes (C1 to C6). The case is classified by evaluating the values of the classification functions for that case and assigning it to the class associated with the highest C value. The six classification functions are as following:

$$C1 = -71.3 \cdot T1 - 48.9 \cdot T2 - 127.8 \cdot T3 + 2483 \cdot h1 + 2487.6 \cdot h2 + 2172.7 \cdot h3 + 175.4 \cdot P1 + 206.1 \cdot P2 + 170 \cdot P3 - 13488.8 \quad (1)$$

$$C2 = -68.7 \cdot T1 - 47.6 \cdot T2 - 125.2 \cdot T3 + 2672.8 \cdot h1 + 2442.4 \cdot h2 + 2240.3 \cdot h3 + 175.8 \cdot P1 + 206.1 \cdot P2 + 170.4 \cdot P3 - 13326.2 \quad (2)$$

$$C3 = -65.5 \cdot T1 - 46.6 \cdot T2 - 125.1 \cdot T3 + 2740.0 \cdot h1 + 2413.7 \cdot h2 + 2284.9 \cdot h3 + 177.4 \cdot P1 + 207.5 \cdot P2 + 172.2 \cdot P3 - 13332.2 \quad (3)$$

$$C4 = -69.5 \cdot T1 - 44.8 \cdot T2 - 121.4 \cdot T3 + 3034.4 \cdot h1 + 2242.8 \cdot h2 + 2478.0 \cdot h3 + 176.9 \cdot P1 + 207.6 \cdot P2 + 173.2 \cdot P3 - 13341.7 \quad (4)$$

$$C5 = -67.6 \cdot T1 - 40.9 \cdot T2 - 119.8 \cdot T3 + 3346.2 \cdot h1 + 2059.7 \cdot h2 + 2606.5 \cdot h3 + 175.4 \cdot P1 + 207.1 \cdot P2 + 173 \cdot P3 - 13058.4 \quad (5)$$

$$C6 = -59.4 \cdot T1 - 52.2 \cdot T2 - 103.2 \cdot T3 + 3328.9 \cdot h1 + 1783.4 \cdot h2 + 2843.7 \cdot h3 + 168.8 \cdot P1 + 198.3 \cdot P2 + 168.8 \cdot P3 - 12082.0 \quad (6)$$

Where, variables (i.e., T1, T2, T3, h1, h2, h3, P1, P2 and P3) are DSC parameters related to the case being classified, the constants (e.g., -71.3, -48.9, -127.8, etc.) are regression coefficients (slope), the constant term (e.g., -13488.8, -13326.2, etc.) represents the intercept in the linear equation.

Figure 3. presents the results of the discriminant analysis. From each classification function, a (C) value can be calculated based on the linear combination of the DSC variables and their corresponding coefficients. The higher the C value, the more likely the case belongs to the corresponding class. It's important to note that the coefficients in the classification functions are obtained through the LDA algorithm, which allowed the separation between classes to be maximized, based on the available data from DSC melting curves. The confusion matrix indicated (Table 2) that only one oil sample with 5% adulteration was classified as a 10% adulterated sample. Thus, the accuracy of the LDA model was 99.5%. A similar approach to detecting adulterations in peanut oil was adopted by other authors, where the identification accuracy was 97% [14].

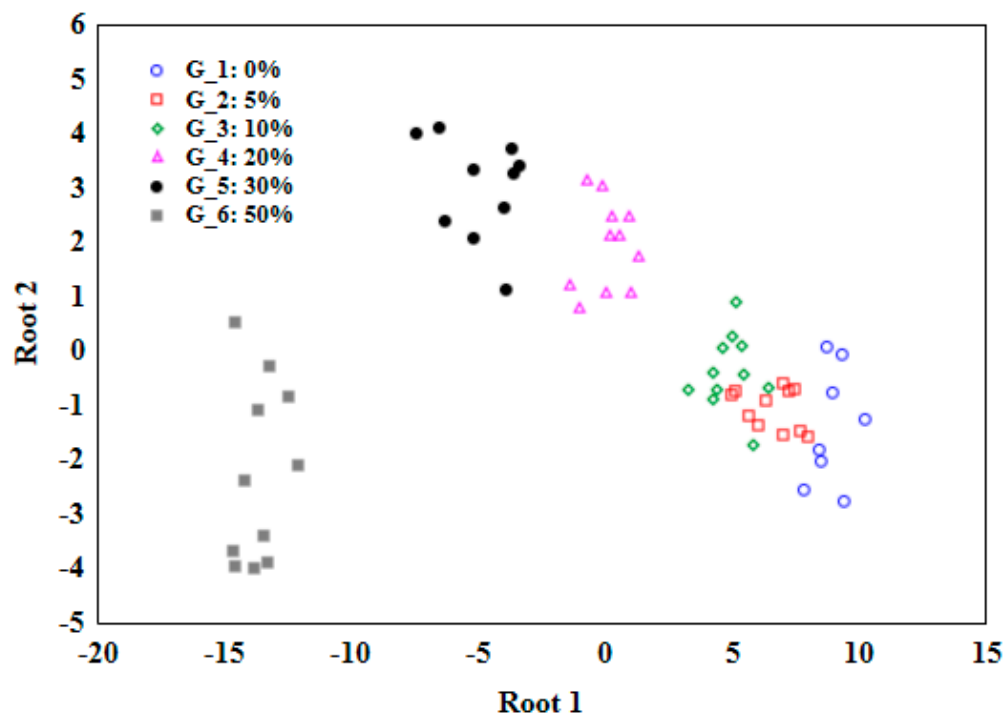


Figure 3. Linear discrimination analysis plot (LDA) for cold-pressed flaxseed oil adulterated with various concentration of refined rapeseed oils (0, 5, 10, 20, 30, 50% w/w).

Table 2. Confusion matrix of cold-pressed flaxseed oils adulterated with different concentrations of refined rapeseed oils.

Observed							
Predicted		0%	5%	10%	20%	30%	50%
0%	LDA	8	0	0	0	0	0
	ANN	8	0	0	0	0	0
	SVM	6	2	0	0	0	0
	MARS	7	1	0	0	0	0
5%	LDA	0	10	1	0	0	0
	ANN	0	10	1	0	0	0
	SVM	0	10	1	0	0	0
	MARS	0	9	2	0	0	0
10%	LDA	0	0	11	0	0	0
	ANN	0	3	8	0	0	0

	SVM	0	2	9	0	0	0
	MARS	0	1	10	0	0	0
20%	LDA	0	0	0	11	0	0
	ANN	0	0	0	11	0	0
	SVM	0	0	0	11	0	0
	MARS	0	0	0	9	2	0
30%	LDA	0	0	0	0	10	0
	ANN	0	0	0	0	10	0
	SVM	0	0	0	0	10	0
	MARS	0	0	0	2	8	0
50%	LDA	0	0	0	0	0	11
	ANN	0	0	0	0	0	11
	SVM	0	0	0	0	0	11
	MARS	0	0	0	0	0	11

MARS regression was used to build the second classification model. Multivariate Adaptive Regression Splines (MAR Splines) is the implementation of a generalization of a technique introduced into wide use by Friedman [50] and used to solve both regression and classification problems. MARS is a non-parametric procedure requiring no assumptions about the functional relationship between the dependent and independent variables. MAR Splines models this relationship with a set of coefficients and so-called basis functions that are entirely determined from the data. In this study, MARS models created for the data matrix included a maximum of 21 basis functions. The penalty was set to 2, and the threshold to 0.0005. The MARS model of the first order was created for classification purposes and the maximum number of terms was limited by pruning. The model has 6 basis functions and 7 terms with GCV=0.516. Increased numbers of basis functions did not decrease the GCV error. MARS model coefficients and knots are presented in Table 4. The model developed here allows 90.3% correct classifications to be obtained. The confusion matrix indicated that six samples were incorrectly classified. Therefore, the accuracy of the model based on MARS analysis was about 95.7%, as presented in Table 3. The MARS regression model was also used by other researchers to define the discriminant surface for studying the authentication of cod liver oil [15].

Table 3. Performance parameters of models for classification of cold-pressed flaxseed oils adulterated with different concentrations of refined rapeseed oils.

Performance parameter Model	Accuracy	Misclassification Rate	Precision	Sensitivity	Specificity	F1-score
LDA	99.46%	0.54%	98.39%	98.39%	99.68%	98.39%
ANN	97.85%	2.15%	93.55%	93.55%	98.71%	93.55%
SVM	97.31%	2.69%	91.94%	91.94%	98.39%	91.94%
MARS	95.70%	4.30%	87.10%	87.10%	97.42%	87.10%

Table 4. MARS model coefficients and knots calculated for classification of cold-pressed flaxseed oils adulterated with different concentrations of refined rapeseed oils.

	Intercept	Term 1	Term 2	Term 3	Term 4	Term 5	Term 6
0%	-1.01	-3.39	-9.46×10^{-1}	8.85×10^{-1}	2.39×10^{-2}	9.37×10^{-1}	8.68
5%	1.27	-1.39	1.35	-9.96×10^{-1}	3.12×10^{-1}	-1.28	4.00
10%	-1.02×10^{-1}	1.81×10^1	-3.68×10^{-1}	-2.01×10^{-1}	-7.60×10^{-2}	2.72×10^{-1}	-2.12×10^1
20%	3.38×10^{-1}	2.92	1.83×10^{-1}	3.18×10^{-1}	-3.07×10^{-1}	-2.26×10^{-1}	-9.92
30%	5.08×10^{-1}	-1.58×10^1	-6.08×10^{-1}	-1.67×10^{-2}	4.51×10^{-2}	2.55×10^{-1}	1.79×10^1
50%	1.39	-2.29	5.25×10^{-1}	-1.03×10^{-1}	3.13×10^{-2}	-8.92×10^{-2}	4.43
Knots T1			-3.44×10^1	-3.44×10^1		-3.55×10^1	

Knots T2		-2.88 x 10 ¹	
Knots h2	4.83 x 10 ⁻¹		5.28 x 10 ⁻¹

Another model which was examined for its usefulness in classifying oil samples into different adulteration classes was the Support Vector Machines (SVM) model. SVM is a method for classifying samples on the basis of the variables (predictors) that describe them. It is a supervised technique, that is, with a supervisor, i.e., there are both variables describing the samples and their membership in defined classes in the learning sample. The support vector method performs classification tasks by constructing hyperplanes in a multidimensional space that separates samples belonging to different classes. For SVM model calculations, the datasets were divided into three subsets in a ratio of 2:1:1 (training, validation, and test set). Samples were classified by the C-SVM method with a linear Kernel type. As a result of learning, a model was obtained that allowed an almost 92% (Table 2) correct classification of oil samples with 97.3% accuracy (Table 3). Another study showed the classification accuracy of SVM as 96.25% while comparing chemometrics and AOCS official methods for predicting the shelf life of edible oil [51].

The last classification model built was an Artificial Neural Network model (ANN). For calculating ANN model, the datasets were divided into three subsets in a ratio of 2:1:1 (training, validation, and test set). The ANN model was trained using selected parameters from the data set and was subsequently validated using an independent data set. Multilayer feed-forward connected ANN was trained with the Broyden-Fletcher-Goldfarb-Shanno (BFGS) learning algorithm (200 epoch). The search for an appropriate ANN model was done using multilayer perceptron (MLP) and radial basis function (RBF) networks. In total, 20 networks were evaluated and the best five were retained. The neural network consists of an input layer, one hidden layer and one output layer. The network architecture, mainly the size of the hidden layer, was selected empirically, taking into consideration the accuracy of predicting the results. The best five ANN-MLP networks are presented in Table 5.

Table 5. ANN models calculated for the classification of cold-pressed flaxseed oils adulterated with different concentrations of refined rapeseed oils.

Net architecture	Training accuracy	Test accuracy	Validation accuracy	Training algorithm	Error function	Hidden activation	Output activation
MLP 9-9-6	88.636	100.000	77.778	BFGS 10	Entropy	Linear	Softmax
MLP 9-11-6	88.636	88.889	100.000	BFGS 11	Entropy	Linear	Softmax
MLP 9-8-6	81.818	88.889	100.000	BFGS 9	Entropy	Linear	Softmax
MLP 9-9-6	84.091	77.778	100.000	BFGS 32	SOS	Exponential	Exponential
MLP 9-4-6	84.091	88.889	88.889	BFGS 15	Entropy	Tanh	Softmax

In the neural network obtained for oil sample classification, the Linear, Exponential and Tanh functions were used in the hidden layer, while Softmax and Exponential functions were used in the output layer. In the input layer, there are 9 neurons, which are DSC parameters. The number of neurons in hidden layer varies from 4 to 11, while the output layer contains 6 neurons representing each class of oil adulteration. A model consisting of the best five networks was used for oil sample classification. The accuracy of the resulting ANN model is almost 98% (Table 3) with only 3 samples misclassified (Table 2). This finding can be compared with the study conducted by Firouz et al. [52], who employed the classification and quantification of sesame oil adulteration, and acquired 100% accuracy.

On the basis of the models' performance parameters, it was determined that the best model is the LDA model, with the highest values for accuracy (99.5%) precision (98.4%), sensitivity (98.4%), specificity (99.7%) and F1-score (98.4%), and the lowest value of misclassification rate equal 0.54%. In contrast, the worst one was the MARS model, which had the lowest values for accuracy (95.7%), precision (87%), sensitivity (87%), specificity (97.4%) and F1-score (87%) and the highest value of misclassification rate equal 4.3%. The second-best model was the ANN model and the third was the

SVM model. The accuracy of all these models was very high, which suggested its ability to predict adulterated oil samples into appropriate classes.

4.2. Regression models for predicting the concentration of refined rapeseed oil in cold-pressed flaxseed oil

Multiple regression analysis (MLR) was performed to formulate a general linear equation which will fit the variables from DSC melting curves against different concentrations of adulterants. This will provide the possibilities to detect the percentage of adulterants in any sample. The MLR model that was obtained was statistically significant with $F(9.52) = 364.57$ ($p \leq 0.05$), $R^2=0.9844$ and adjusted $R^2=0.9817$. The standard error of estimation was 2.3028.

Table 6 demonstrates the summary of DSC parameters, where (b^*) values refer to the standardized regression coefficient, and (b) values refer to the regular regression coefficient. Determining (b^*) allowed for a direct comparison of the magnitude and importance of the independent variables, where we can see the highest values are presented for h_2 , T_3 and h_3 as -0.32 , 0.27 and 0.16 , respectively. On the other hand, (b) values signify the slope coefficient associated with an independent variable. It represents the change in the dependent variable for a one-unit change in the corresponding independent variable, while holding all other independent variables constant. Table 6 shows that the h_2 (-65.52) variable has the strongest and negative relationship with the concentration variables, indicating that with a decreased value of h_2 , the concentration of adulterants increased. On the other hand, h_3 and T_2 variables consequently increase or decrease linearly with the concentration values of adulterants. Accordingly, a model with statistically significant predictors was built.

$$\% \text{ adulterant} = 2.185 \cdot T_3 - 109.995 \cdot h_2 + 96.576 \cdot h_3 + 105.319 \pm 2.567 \quad (7)$$

Where T_3 represents the third strongest significant independent variable ($p = 0.000$), h_2 represents the highest strongest significant independent variable ($p = 0.000$), and h_3 represents the second strongest significant independent variable ($p = 0.000$).

Table 6. Summary of independent variables in multiple regression analysis.

DSC parameters	b^* (standardized co-efficient)	b (raw co-efficient)	p -value
		145.3464*	0.000059*
T_1	0.090680	1.2225	0.126748
T_2	-0.052531	-0.4175	0.627294
T_3	0.273274*	2.3374*	0.000332*
h_1	0.080773	73.6583	0.056670
h_2	-0.324189^*	-65.5228^*	0.000014*
h_3	0.168666*	65.1633*	0.004121*
P_1	-0.128623	-0.4718	0.151392
P_2	-0.145236	-0.6314	0.083058
P_3	0.019714	0.0435	0.896428

* Coefficients are significant statistically ($p \leq 0.05$).

The goodness of fit of the model to the experimental data and the coefficient of determination R^2 and the coded coefficient of determination were 0.978 and 0.977, respectively. Equation no. 7 can be used to estimate the percentage of adulterants (for this study, refined rapeseed oil) in a cold-pressed flaxseed oil sample based on three dependent variables: T_3 , h_2 , and h_3 . The equation implies that the variables T_3 , h_2 , and h_3 are assumed to have a linear relationship with the percentage of the adulterant. The correlation between observed and predicted values was 0.992 with a low RMSE value of 2.12 (Table 7). A similar study by Sim et al. [12] showed that it was possible to predict adulteration of lard in palm oil olein using the MLR model, where the prediction performance was measured based on the percentage root mean square error (%RMSE).

Table 7. Goodness of fit parameters between observed and predicted regression models for the prediction of concentrations of refined rapeseed oils in cold-pressed flaxseed oil.

Model	R	R ²	Adjusted R ²	AIC	BIC	RMSE
ANN	0.996	0.992	0.992	233	240	1.51
MARS	0.995	0.990	0.990	244	251	1.65
SVM	0.992	0.985	0.984	274	280	2.10
MLR	0.992	0.984	0.984	275	281	2.12

MARS regression was used to build the second regression model. In this study, MARS models created for the data matrix included a maximum of 21 basis functions. The penalty was set to 2, and the threshold to 0.0005. The MARS model of the first order was created for classification purposes and the maximum number of terms was limited by pruning. The model has 10 basis functions and 11 terms with GCV=6.252. Equation 8 represents the MARS model for predicting the concentration of refined oil in the samples.

$$\begin{aligned} \% \text{ adulterant} = & 1.955e - 1.9203 \cdot \max(0; T2+2.763e) - 3.192 \cdot \max(0; -2.763e-T2) + 3.6434 \cdot \max(0; \\ & T1+3.437e) + 4.572 \cdot \max(0; T3+2.254e) + 3.465e^2 \cdot \max(0; h1-1.241e^{-1}) - 4.490e^2 \cdot \max(0; h1-1.4729e^{-1}) \\ & + 2.214 \cdot \max(0; 3.240e-P2) - 7.047e^{-1} \cdot \max(0; P1-2.769e) + 4.543e^{-1} \cdot \max(0; 2.769e-P1) - \\ & 7.618e \cdot \max(0; h2-4.510e^{-1}) \end{aligned} \quad (8)$$

The correlation between the observed and predicted value was 0.995 with a low RMSE value of 1.65 (Table 7).

Another model which was examined for its usefulness in predicting oil sample adulteration was the Support Vector Machines (SVM) model. SVM can be used for both classification and regression problems. In SVM regression, the search is for a functional dependence of the dependent variable y (% of adulteration) on a set of independent variables x (DSC parameters). For calculating the SVM model, the datasets were divided into three subsets in a ratio of 2:1:1 (training, validation, and test set) for model regression type 1 (C=10.000000, epsilon=0.100000) with radial basis function (gamma=0.111111) kernel type. Samples were classified by C-SVM method with linear kernel type. The correlation between the observed and predicted values was 0.992 with a low RMSE value of 2.1 (Table 7).

The last regression model built was an Artificial Neural Network model (ANN). For calculating the ANN model, the datasets were divided into three subsets in a ratio of 2:1:1 (training, validation, and test set). The ANN was trained using selected parameters from the data set and was subsequently validated using an independent data set. Multilayer feed-forward connected ANN was trained with the Broyden-Fletcher-Goldfarb-Shanno (BFGS) learning algorithm (200 epoch). The search for an appropriate ANN model was performed using multilayer perceptron (MLP) and radial basis function (RBF) networks. In total, 20 networks were evaluated and the five best were retained. The neural network consists of an input layer, one hidden layer and one output layer. The network architecture, mainly the size of the hidden layer, was selected empirically, taking into consideration the accuracy of the results prediction. The five best ANN-MLP networks are presented in Table 8.

Table 8. ANN models calculated for the prediction of refined rapeseed oil concentrations in cold-pressed flaxseed oils.

Net architecture	Training accuracy	Test accuracy	Validation accuracy	Training algorithm	Error function	Hidden activation	Output activation
MLP 9-9-1	0.9957	0.9919	0.9901	BFGS 43	SOS	Logistic	Logistic
MLP 9-13-1	0.9964	0.9961	0.9941	BFGS 46	SOS	Logistic	Tanh
MLP 9-11-1	0.9965	0.9913	0.9885	BFGS 61	SOS	Logistic	Exponential
MLP 9-13-1	0.9961	0.9899	0.9866	BFGS 54	SOS	Logistic	Exponential
MLP 9-11-1	0.9963	0.9937	0.9972	BFGS 37	SOS	Tanh	Tanh

In the neural network obtained for predicting oil adulteration, the Logistic and Tanh functions were used in the hidden layer, while Logistic, Tanh and Exponential functions were used in the output layer. In the input layer, there are 9 neurons, which are DSC parameters. The number of neurons in the hidden layer varies from 9 to 13, while the output layer contains 14 neurons representing the refined oil concentration. A model consisting of the best five networks was used for prediction purposes. The correlation between the observed and predicted values was 0.996 with a low RMSE value of 1.51 (Table 7). The study found that ANN regression analysis demonstrated robust models for adulteration phenomena in sesame oil generated by sunflower oil, canola oil and sunflower + canola oils quantitatively [52]. Another study stated that using ANN as a pattern recognition technique for the data obtained from electronic nose could not detect the proportion of adulteration in camellia seed oil, but successfully quantified adulteration in sesame oil [53].

Table 7 presents the goodness of fit parameters of the regression models obtained. The best model is based on ANN algorithms and exhibits the highest R (0.996), R^2 (0.992), adjusted R^2 (0.922) values with the lowest values of AIC (233), BIC (240) and RMSE (1.51). The MARS model was next, followed by the SVM, and the least fitting was the MLR model.

4.2.1. Principle component analysis (PCA) and Orthogonal partial least squares discriminant analysis (OPLS-DA)

This adulteration detection study involves analyzing multiple variables of DSC parameters simultaneously. Chemometric techniques like PCA and OPLS-DA are designed to handle multivariate data, allowing for a comprehensive analysis of the oil samples. For instance, PCA presented in Figure 4 (a) reduced the dimensionality of the dataset by transforming the variables into a smaller set of principal components, capturing the most important variations in the data. Hence, as an unsupervised method, PCA represents the combinations of the original variables and can be difficult to interpret in the context of class separation. To solve this issue, OPLS-DA analysis was adopted for assessing the discrimination and classification of the adulterated flaxseed oil samples. As a fast and efficient screening tool for large datasets, OPLS-DA allowed us to evaluate the effectiveness of DSC melting profiles of adulterated flaxseed oils in classifying and detecting the percentage of adulterants concentration by differentiating them (Figure 4 b.).

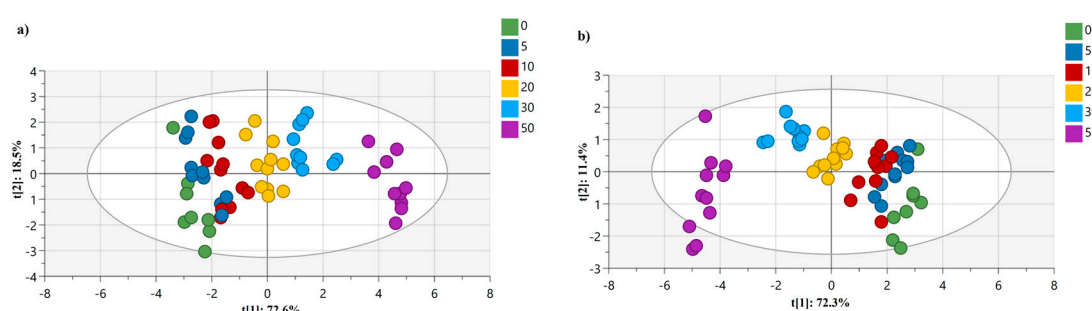


Figure 4. Score plots obtained by (a) PCA and (b) OPLS-DA for cold-pressed flaxseed oil adulterated with various concentration of refined rapeseed oils (0, 5, 10, 20, 30, 50% w/w).

In Figure 4 (a), DSC data matrix serves as the basis for conducting PCA analyses, which provided a visual representation of the data pattern for six concentrations of adulterants mixed with pure flaxseed oils. In the score plot, each point represents a sample (a specific concentration of adulterant mixed with flaxseed oil) in the space of two principal components, $t[1]$ and $t[2]$, which were able to explain 91.1% of the variation of the normalized heat flow results. Additionally, $R^2X(\text{cum})$ and $Q^2(\text{cum})$ values (Table 9) are the quantities useful for PC model diagnostics as the fraction of the explained variation R^2X and the fraction of predicted variation Q^2 . The more significant a principal component, the closer its R^2X and $R^2X(\text{cum})$ will be to value 1 for a PC model with a sufficient number of components. For the PCA analysis presented in Figure 4 (a), the $R^2X(\text{cum})$ value was 0.973, which indicates that the retained principal components capture a larger proportion of the overall variation

in the dataset. This finding can help in determining the appropriate number of components to retain for further analysis. Besides this, the Q^2 (cum) value was 0.897 for the PCA model, which shows that the cumulative sum of the cross-validated predictive ability is high for the variables of the normalized heat flow of phase transition curves. This approach of employing PCA analysis can be compared to other studies, where researchers detected (with 100% accuracy) adulteration of flaxseed oil with rapeseed, corn, peanut, sunflower seed, soybean, and sesame oils [25], or adulteration of virgin coconut oil with refined coconut oil [16].

Table 9. Summary of the fit of different models for discrimination (differentiation) of samples with different concentrations of adulterants (refined rapeseed oil).

Model	R ² X(cum)	R ² (cum)	Q ² (cum)
PCA	0.973		0.897
OPLS-DA	0.986	0.465	0.330

The next chemometric approach was analyzing the dataset of multiple variables using OPLS-DA, which can effectively enhance the separation of classes while maintaining the predictive power of the model by utilizing orthogonal projection in the score plot. The analysis aims to classify and distinguish different concentrations of adulterants (ranging from 0% to 50%) added to pure flaxseed oil. The model consists of 15 variables, where a total of 9 DSC parameters are considered as X variables and 6 different concentrations of adulterants added are considered as Y variables representing the 6 classes. Five predictive components (P1 to P5) capture the between-class variation, meaning they account for the differences between the different concentrations of adulterants. The orthogonal components capture the within-class variation, representing similarities within each concentration group. Within the framework of the OPLS-DA model, the systematic variation in data was described by two distinct components. The first component, known as the predictive component, exhibits a linear relationship with the classes (Y) and possesses the ability to make accurate predictions. In Figure 4 (b), the X-axis represents the first component (t1=72.3%) and the Y-axis the second principal component [t2=11.4%]. The observations in the scatter plot are colored, based on their class, which corresponds to the different concentrations of adulterants added to pure flaxseed oil. The scatter plot serves as a visualization of how the modeled observations in the X space are positioned relative to each other. Observations that are close to each other in the plot indicate a higher degree of similarity compared to those that are farther apart.

Also, in Table 9, the R^2X (cum) value was presented as 0.986, indicating that the OPLS-DA model fits the X data well, capturing a large portion of the variation present in the DSC parameters. On the other hand, a Q^2 (cum) value of 0.33 indicates that the OPLS-DA model can predict approximately 33% of the variation in the Y data, according to cross-validation. The range of Q^2 values suggests that the model has reasonable predictive ability for the concentration of adulterants based on DSC variables. OPLS-DA was also adopted by other authors to determine important variables when detecting flaxseed oil multiple adulteration by near-infrared spectroscopy. These authors also adopted the one-class partial least squares (OCPLS) method to build a detection model which provided a high accuracy of 95.8% [17].

Although the model demonstrates a good fit to the X data (DSC parameters), a low Q^2 (cum) value (0.33) indicates that the model's ability to explain and predict the variation for the Y data (adulterant concentrations) was poor. Thus, the authors decided to explore an alternative modeling technique i.e., the Partial Least Squares (PLS) technique. In Figure 5 (a), a loading plot of PLS analysis is presented for DSC parameters obtained from the melting curves. To obtain a comprehensive understanding of the model's performance and predictive ability, the R^2X (cum) value was calculated at a level of 0.953 and was lower than for OPLS-DA, while the predictive Q^2 (cum) value was higher than for OPLS-DA (0.973). The results obtained for R^2X (cum) and Q^2 (cum) indicate that the PLS model had higher predictive power than OPLS-DA. Additionally, Figure 5 (b) presents the variables' influence on the projection (VIP) plot, which provides information about the importance of variables

(DSC parameters) which are above value 1. As was the case with the MLR model (Table 6), the parameter for the first peak height (h1) and percentage area (P1) were not significant.

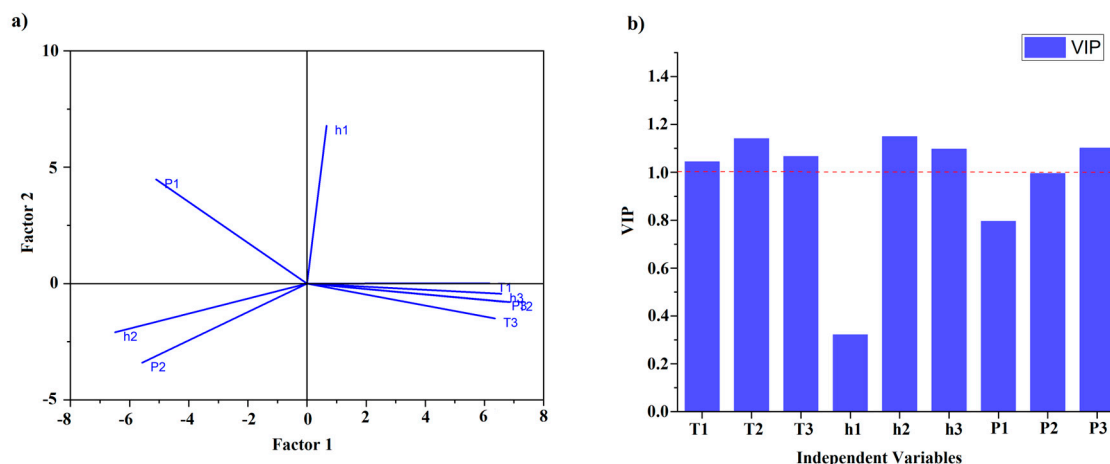


Figure 5. Loading plot of (a) PLS analysis for all DSC parameters determined for cold-pressed flaxseed oils adulterated with various concentration of refined rapeseed oil (0, 5, 10, 20, 30, 50% w/w). (b) the variables' influence on the projection (VIP) graph.

In addition to the approaches presented in the study, the observed and predicted value graph from the PLS model is presented in Figure 6. By plotting the observed concentrations of adulterants (actual values) against the predicted concentrations (values predicted by the PLS model) on a graph, it was possible to assess visually how well the model predicts the adulterant levels in the flaxseed oil samples, based on the DSC parameters from melting curves. We can see that the observed and predicted values align closely along a diagonal line, which indicates that the PLS model accurately predicts the adulterant concentrations based on the DSC parameters. A Pearson's correlation coefficient (r) of 0.995 between the observed and predicted values indicates an extremely strong positive linear relationship between the two sets of values. This graph also shows that this model can effectively differentiate between pure flaxseed oil and adulterated samples, providing a reliable means of detecting and estimating the adulterant concentrations. By assessing this graph, it is also evident that the PLS model has successfully learned the relationship between the DSC parameters and the adulterant concentrations, which validates the model for this purpose. This finding can be compared with the study conducted by Rocha et al., who adopted the PLS method for the classification and quantification of different types of blended biodiesel synthesized from peanut, corn, and canola oils, and observed a Pearson's correlation coefficient of 0.969 between the real and predicted concentrations [13].

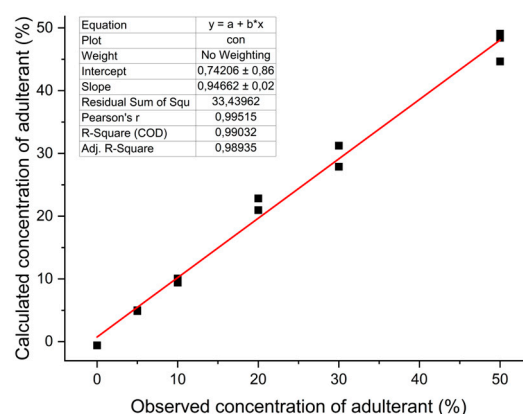


Figure 6. Observed and predicted values in a Partial Least Squares (PLS) model based on the DSC parameters.

5. Conclusions

By detecting adulteration in flaxseed oil, the study contributes to addressing the ongoing concern of food fraud and its economic impact on the global food industry. The DSC melting curves provided unique and substantial information about the thermal profiles of the oils and showed distinct changes when adulterants were added. The second peak in the DSC curves was identified as the major peak, and its characteristics, such as peak temperature, peak height, and percentage of peak area, were found to be significantly affected by the concentration of adulterants. Nevertheless, the findings of this study demonstrate the efficacy of DSC coupled with chemometric analysis in detecting and classifying adulterations in cold-pressed flaxseed oil. Of the classification models built, the LDA model exhibited the best performance, underlining its potential for accurate identification of adulterated oil samples. On the other hand, the regression model based on the ANN algorithm showed best goodness of fit for DSC parameters regarding the prediction of adulterant concentrations. The equation and PMML codes derived from the MLR, MARS, SVM and ANN regression analysis can be used to estimate the percentage of adulterants in flaxseed oil samples based on the values of peak temperature, peak height, and percentage of peak area. The study also employed other chemometric techniques, such as PCA, OPLS-DA and PLS to effectively classify and describe the adulteration phenomena. The resulting plots demonstrated that the PLS model showed the greatest accuracy in predicting adulterant levels in flaxseed oil samples based on DSC parameters, as indicated by a strong positive linear relationship (Pearson's correlation coefficient of 0.995). The PLS model effectively differentiated between pure flaxseed oil and adulterated samples, providing a reliable means of detecting and estimating adulterant concentrations. This study highlights the significance of combining DSC with chemometric analysis for detecting adulterations in flaxseed oil and emphasizes the importance of quality assessment and authenticity verification in the food industry.

Author Contributions: Conceptualization, M.I., and J.T.-G.; methodology, M.I., A.K., and J.T.-G.; formal analysis, M.I., A.K., M.M., and J.T.-G.; investigation, M.I. and J.T.-G.; resources, J.T.-G.; data curation, M.I., A.K., and J.T.-G.; writing—original draft preparation, M.I., A.K., and J.T.-G.; writing—review and editing, M.I., A.K., and J.T.-G.; visualization, M.I., A.K. and J.T.-G.; supervision, A.K., J.T.-G., project administration, J.T.-G.; funding acquisition, J.T.-G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the NATIONAL SCIENCE CENTRE, POLAND, grant number: 2018/31/B/NZ9/02762.

Data Availability Statement: The data presented in this study are available upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Drivelos SA, Higgins K, Kalivas JH, Haroutounian SA, Georgiou CA. Data fusion for food authentication. Combining rare earth elements and trace metals to discriminate “fava Santorinis” from other yellow split peas using chemometric tools. *Food Chem* [Internet]. Elsevier Ltd; 2014;165:316–22. Available from: <http://dx.doi.org/10.1016/j.foodchem.2014.03.083>
2. Hong Y, Birse N, Quinn B, Li Y, Jia W, Mccarron P, et al. Data fusion and multivariate analysis for food authenticity analysis. *Nat Commun*. 2023;14:3309.
3. Márquez C, López MI, Ruisánchez I, Callao MP. FT-Raman and NIR spectroscopy data fusion strategy for multivariate qualitative analysis of food fraud. *Talanta* [Internet]. Elsevier; 2016;161:80–6. Available from: <http://dx.doi.org/10.1016/j.talanta.2016.08.003>
4. Jović O, Jović A. FTIR-ATR adulteration study of hempseed oil of different geographic origins. *J Chemom*. 2017;31:1–9.
5. Carter OC. S. On the Detection of Adulterations in Oils. 1885;22:296–9.
6. FAO 2009. Economically motivated adulteration; public meeting; request for comment. 2009. p. 69–71.

7. FAO. Food fraud-Intention, detection and management. Food safety technical toolkit for Asia and the Pacific No. 5. Bangkok [Internet]. 2021. Available from: <https://www.fao.org/3/cb2863en/cb2863en.pdf>
8. Gelpi E, Posada de la Paz M, Terracini B, Abaitua I, Gómez de la Cámara A, Kilbourne EM, et al. The spanish toxic oil syndrome 20 years after its onset: A multidisciplinary review of scientific knowledge. *Environ Health Perspect.* 2002;110:457–64.
9. European Commission. European Commission The EU Food Fraud Network and the System for Administrative Assistance-Food Fraud [Internet]. 2018. Available from: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:772:FIN>
10. Rocha WFC, Prado CB, Niksa B. Comparison of Chemometric Problems in Food Analysis using Non-Linear Methods. *Molecule.* 2020;25.
11. Zhang L, Chen J, Jing B, Dong Y, Yu X. New Method for the Discrimination of Adulterated Flaxseed Oil Using Dielectric Spectroscopy. *Food Anal Methods.* 2019;12:2623–9.
12. Sim SF, Chai MXL, Jeffrey Kimura AL. Prediction of lard in palm olein oil using simple linear regression (SLR), multiple linear regression (MLR), and partial least squares regression (PLSR) based on fourier-transform infrared (FTIR). *J Chem. Hindawi;* 2018;2018:1–8.
13. Rocha WFC, Vaz BG, Sarmanho GF, Leal LHC, Nogueira R, Silva VF, et al. Chemometric Techniques Applied for Classification and Quantification of Binary Biodiesel/Diesel Blends. *Anal Lett.* 2012;45:2398–411.
14. Peng D, Shi C, Nie Q, Xie S, Bi Y, Li J. Qualitative and quantitative detection of peanut oils adulteration based on fatty acid information fusion coupled with chemometrics. *Lwt [Internet]. Elsevier Ltd;* 2023;181:114785. Available from: <https://doi.org/10.1016/j.lwt.2023.114785>
15. Giese E, Rohn S, Fritsche J. Chemometric tools for the authentication of cod liver oil based on nuclear magnetic resonance and infrared spectroscopy data. *Anal Bioanal Chem. Analytical and Bioanalytical Chemistry;* 2019;411:6931–42.
16. Bao R, Tang F, Rich C, Hatzakis E. A comparative evaluation of low-field and high-field NMR untargeted analysis: Authentication of virgin coconut oil adulterated with refined coconut oil as a case study. *Anal Chim Acta [Internet]. Elsevier B.V.;* 2023;1273:341537. Available from: <https://doi.org/10.1016/j.aca.2023.341537>
17. Yuan Z, Zhang L, Wang D, Jiang J, Harrington P de B, Mao J, et al. Detection of flaxseed oil multiple adulteration by near-infrared spectroscopy and nonlinear one class partial least squares discriminant analysis. *Lwt [Internet]. Elsevier;* 2020;125:109247. Available from: <https://doi.org/10.1016/j.lwt.2020.109247>
18. Hao S, Zhu L, Sui R, Zuo M, Luo N, Shi J, et al. Identification and quantification of vegetable oil adulteration with waste frying oil by laser-induced fluorescence spectroscopy. *OSA Contin.* 2019;2:1148.
19. Garrido-Delgado R, Eugenia Muñoz-Pérez M, Arce L. Detection of adulteration in extra virgin olive oils by using UV-IMS and chemometric analysis. *Food Control.* 2018;85:292–9.
20. Vanstone N, Moore A, Martos P, Neethirajan S. Detection of the adulteration of extra virgin olive oil by near-infrared spectroscopy and chemometric techniques. *Food Qual Saf.* 2018;2:189–98.
21. Huang K, Zhong P, Xu B. Discrimination on potential adulteration of extra virgin olive oils consumed in China by differential scanning calorimeter combined with dimensionality reduction classification techniques. *Food Chem [Internet]. Elsevier Ltd;* 2023;405:134996. Available from: <https://doi.org/10.1016/j.foodchem.2022.134996>
22. De Souza LM, De Santana FB, Gontijo LC, Mazivila SJ, Borges Neto W. Quantification of adulterations in extra virgin flaxseed oil using MIR and PLS. *Food Chem.* 2015;182:35–40.
23. Huang ZM, Xin JX, Sun SS, Li Y, Wei DX, Zhu J, et al. Rapid identification of adulteration in edible vegetable oils based on low-field nuclear magnetic resonance relaxation fingerprints. *Foods.* 2021;10.
24. Sun X, Zhang L, Li P, Xu B, Ma F, Zhang Q, et al. Fatty acid profiles based adulteration detection for flaxseed oil by gas chromatography mass spectrometry. *Lwt [Internet]. Elsevier Ltd;* 2015;63:430–6. Available from: <http://dx.doi.org/10.1016/j.lwt.2015.02.023>
25. Wang J, Han Y, Wang X, Li Y, Wang S, Gan S, et al. Adulteration detection of Qinghai-Tibet Plateau flaxseed oil using HPLC-ELSD profiling of triacylglycerols and chemometrics. *Lwt [Internet]. Elsevier Ltd;* 2022;160:113300. Available from: <https://doi.org/10.1016/j.lwt.2022.113300>
26. Elzey B, Pollard D, Fakayode SO. Determination of adulterated neem and flaxseed oil compositions by FTIR spectroscopy and multivariate regression analysis. *Food Control [Internet]. Elsevier Ltd;* 2016;68:303–9. Available from: <http://dx.doi.org/10.1016/j.foodcont.2016.04.008>

27. Maléchaux A, Le Dréau Y, Artaud J, Dupuy N. Control chart and data fusion for varietal origin discrimination: Application to olive oil. *Talanta* [Internet]. Elsevier B.V.; 2020;217:121115. Available from: <https://doi.org/10.1016/j.talanta.2020.121115>
28. Wang T, Wu HL, Long WJ, Hu Y, Cheng L, Chen AQ, et al. Rapid identification and quantification of cheaper vegetable oil adulteration in camellia oil by using excitation-emission matrix fluorescence spectroscopy combined with chemometrics. *Food Chem* [Internet]. Elsevier; 2019;293:348–57. Available from: <https://doi.org/10.1016/j.foodchem.2019.04.109>
29. Chiavaro E, Vittadini E, Rodriguez-Estrada MT, Cerretani L, Bendini A. Differential scanning calorimeter application to the detection of refined hazelnut oil in extra virgin olive oil. *Food Chem*. 2008;110:248–56.
30. Jafari M, Kadivar M, Keramat J. Detection of adulteration in Iranian olive oils using instrumental (GC, NMR, DSC) methods. *JAOCS, J Am Oil Chem Soc*. 2009;86:103–10.
31. Van Wetten IA, Van Herwaarden AW, Splinter R, Boerrigter-Eenling R, Van Ruth SM. Detection of sunflower oil in extra virgin olive oil by fast differential scanning calorimetry. *Thermochim Acta*. Elsevier B.V.; 2015;603:237–43.
32. Karbasian M, Givianrad MH, Ramezan Y. A rapid method for detection of refined olive oil as adulterant in extra virgin olive oil by differential scanning calorimetry. *Orient J Chem*. 2015;31:1735–9.
33. J.M.N. Marikkar. Differential Scanning Calorimetric Analysis of Virgin Coconut Oil, Palm Olein, and their Adulterated Blends. *Cord*. 2019;35:9.
34. Marina AM, Che Man YB, Nazimah SAH, Amin I. Monitoring the adulteration of virgin coconut oil by selected vegetable oils using differential scanning calorimetry. *J Food Lipids*. 2009;16:50–61.
35. Coni E, Di Pasquale M, Coppolelli P, Bocca A. Detection of animal fats in butter by differential scanning calorimetry: A pilot study. *J Am Oil Chem Soc*. 1994;71:807–10.
36. Tomaszewska-Gras J. Multivariate analysis of seasonal variation in the composition and thermal properties of butterfat with an emphasis on authenticity assessment. *Grasas y Aceites*. 2016;67.
37. Upadhyay N, Goyal A, Kumar A, Lal D. Detection of adulteration by caprine body fat and mixtures of caprine body fat and groundnut oil in bovine and buffalo ghee using differential scanning calorimetry. *Int J Dairy Technol*. 2017;70:297–303.
38. Marikkar JMN, Ghazali HM, Che Man YB, Lai OM. The use of cooling and heating thermograms for monitoring of tallow, lard and chicken fat adulterations in canola oil. *Food Res Int*. 2002;35:1007–14.
39. FAO. Production/Yield quantities of Oil of linseed in World + (Total) [Internet]. 2020. Available from: <https://www.fao.org/faostat/en/#data/QCL/visualize>
40. Tomaszewska-Gras J, Islam M, Grzeca L, Kaczmarek A, Fornal E. Comprehensive Thermal Characteristics of Different Cultivars of Flaxseed Oil (*Linum usitatissimum* L.). *Molecules*. 2021;26:1–20.
41. Khattab RY, Zeitoun MA. Quality evaluation of flaxseed oil obtained by different extraction techniques. *LWT - Food Sci Technol* [Internet]. Elsevier Ltd; 2013;53:338–45. Available from: <http://dx.doi.org/10.1016/j.lwt.2013.01.004>
42. Visentainer J V., De Souza NE, Makoto M, Hayashi C, Franco MRB. Influence of diets enriched with flaxseed oil on the α -linolenic, eicosapentaenoic and docosahexaenoic fatty acid in Nile tilapia (*Oreochromis niloticus*). *Food Chem*. 2005;90:557–60.
43. Tuncel NB, Uygur A, Karagül Yüceer Y. The Effects of Infrared Roasting on HCN Content, Chemical Composition and Storage Stability of Flaxseed and Flaxseed Oil. *JAOCS, J Am Oil Chem Soc*. 2017;94:877–84.
44. Xu J, Yang W, Deng Q, Huang Q, Yang J, Huang F. Flaxseed oil and α -lipoic acid combination reduces atherosclerosis risk factors in rats fed a high-fat diet. *Lipids Health Dis*. 2012;11:1–7.
45. Dwivedi C, Natarajan K, Matthees DP. Chemopreventive effects of dietary flaxseed oil on colon tumor development. *Nutr Cancer*. 2005;51:52–8.
46. Popa S, Milea MS, Boran S, Nițu SV, Moșoarcă GE, Vancea C, et al. Rapid adulteration detection of cold pressed oils with their refined versions by UV–Vis spectroscopy. *Sci Rep*. 2020;10:1–9.
47. Tomaszewska-Gras J. Rapid quantitative determination of butter adulteration with palm oil using the DSC technique. *Food Control*. Elsevier Ltd; 2016;60:629–35.
48. Dahimi O, Rahim AA, Abdulkarim SM, Hassan MS, Hashari SBTZ, Siti Mashitoh A, et al. Multivariate statistical analysis treatment of DSC thermal properties for animal fat adulteration. *Food Chem* [Internet]. Elsevier Ltd; 2014;158:132–8. Available from: <http://dx.doi.org/10.1016/j.foodchem.2014.02.087>
49. Indriyani L, Rohman A, Riyanto S. Authentication of avocado oil (*Persea americana* Mill.) Using differential

- scanning calorimetry and multivariate regression. *Asian J Agric Res* [Internet]. Science Alert; 2016;10:78–86. Available from: <http://dx.doi.org/10.3923/ajar.2016.78.86>
50. Friedman JH. Multivariate Adaptive Regression Splines. *Ann Stat* [Internet]. 1991;19:1–67. Available from: <http://www.jstor.org/stable/2241837>
51. Karami H, Rasekh M, Mirzaee – Ghaleh E. Comparison of chemometrics and AOCS official methods for predicting the shelf life of edible oil. *Chemom Intell Lab Syst* [Internet]. Elsevier Ltd; 2020;206:104165. Available from: <https://doi.org/10.1016/j.chemolab.2020.104165>
52. Firouz MS, Omid M, Babaei M, Rashvand M. Dielectric spectroscopy coupled with artificial neural network for classification and quantification of sesame oil adulteration. *Inf Process Agric* [Internet]. China Agricultural University; 2022;9:233–42. Available from: <https://doi.org/10.1016/j.inpa.2021.05.001>
53. Hai Z, Wang J. Detection of adulteration in camellia seed oil and sesame oil using an electronic nose. *Eur J Lipid Sci Technol*. 2006;108:116–24.