Available online at www.sciencedirect.com
SCIENCE DIRECT®
Talanta 66 (2005) 1108–1116
Talanta
www.elsevier.com/locate/talanta

Gas-chromatographic fatty-acid fingerprints and partial least squares modeling as a basis for the simultaneous determination of edible oil mixtures

M. Hajimahmoodi [a,*], Y. Vander Heyden [b], N. Sadeghi [a], B. Jannat [a], M.R. Oveisi [a], S. Shahbazian [a]

[a] *Department of Food Science and Nutrition, Faculty of Pharmacy, Tehran Medical Sciences University, Tehran 14174, Iran*
[b] *Vrije Universiteit Brussel (VUB), Farmaceutisch Instituut, Department of Pharmaceutical and Biomedical Analysis, Laarbeeklaan 103, 1090 Brussels, Belgium*

Received 28 June 2004; received in revised form 15 December 2004; accepted 11 January 2005
Available online 5 February 2005

Abstract

Partial least squares modeling and gas-chromatographic fatty-acid fingerprints are reported as a method for the simultaneous determination of cottonseed, olive, soybean and sunflower edible oil mixtures. In this work, two sets of three- and four-component combinations of oils were prepared, hydrolyzed and the obtained free fatty acids analyzed by gas chromatography (GC) without any further derivatization. The normalized percentages of the myristic (14:0), palmitic (16:0), palmitoleic (16:1), stearic (18:0), oleic (18:1), linoleic (18:2) and linolenic (18:3) acids were chromatographically measured in samples and used for constructing calibration matrix. The cross-validation method was used to select the number of factors and the proposed methods were validated by using two sets of synthetic oil mixture samples. The relative standard error for each oil in mixture samples was less than 10%. This approach allows determining possible adulteration in each of the four edible oils.

Keywords: Fatty-acid fingerprint; Partial least squares; Edible oil; Gas chromatography

1. Introduction

Inclusion of high-quality vegetable oils is of great importance in a fully balanced diet. Hence, there is a need for analytical techniques to control the purity and authenticity of such oils. Quality assurance methods on fatty acids (FAs) are the determination of the peroxide value, iodine value, anisidine value, rancimat induction time and *trans*-value [1–3]. While these methods are still useful with pure samples, the current adulteration methods require more sophisticated techniques. Recently, various techniques have been proposed to establish the authenticity of vegetable oils and to detect the occasional adulteration level. UV spectrophotometry [4], Raman spectroscopy [5,6], ^{13}C NMR analyses [7] and IR spectrometry [8,9] has been used for many years in the oil and fat industry. However, these techniques have limitations as detection methods because the spectral differences of most vegetable oils, which contain the same FAs, mainly C16 or C18, and their triglycerides (C50, C52, C54), are quite small. However, each vegetable oil has its own characteristic fatty acid composition [10]. Therefore, determining the fatty acid "fingerprint" by chromatographic methods provides useful information concerning authenticity and possible adulteration of oils. Usually, fatty acid determination is carried out with liquid chromatography (LC) [11,12] and gas chromatography (GC) [13,14]. Since most LC methods require derivatization in order to enhance detectability of fatty acids [15], GC is now extensively used for the compositional determination of fatty acids [16]. Nowadays, attempts

* Corresponding author. Tel.: +98 21 6959090; fax: +98 21 6461178.
E-mail address: hajimah@sina.tums.ac.ir (M. Hajimahmoodi).

0039-9140/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2005.01.011

are made to separate free underivated fatty acids by HPLC and GC and to detect them without any further derivatization [11,17].

On the other hand, the interest in chemometric methods has now spread to numerous fields of analytical chemistry [18,19]. Data analysis is greatly improved by exploring chromatographic datasets with chemometric techniques, such as principal component analyses (PCA) [12,20], neural networks [21] and discriminant analyses [22,23]. The total triacylglycerol (TAG) profile, measured by LC, was evaluated with principal component analysis to detect adulteration levels between 10 and 20% of four different vegetable oils in olive oil [24]. Kaufmann [25] applied PCA and partial least square (PLS) regression analysis in binary mixtures of four different oils in olive oil and obtained 1.5% detection level. The efficiency of different multivariate calibration models was compared and discussed in many papers and it can be resumed that PLS seems to predict better than principal component regression (PCR) when there are random linear baselines and for low precision data [26,27].

The aim of this paper is to establish the level of different oils present in a mixture, based on the fatty acid profiles determined for these mixtures. The results then can be used to control the purity and authenticity of oils. In the present work, two sets of three- and four-oil combinations were prepared. The oils considered are the commonly used edible cottonseed, olive, soybean and sunflower oils. A GC method without derivatization process in combination with PLS was evaluated to determine the levels of the different oils present in the mixtures.

2. Experimental

2.1. Reagents and chemicals

Standard grades of cottonseed, olive, soybean, sunflower oils and free fatty acid standards were purchased from Sigma (St. Louis, MO, USA). Potassium hydroxide, sulphuric acid and the organic solvents hexane, butanol and isopropanol were obtained from Merck (Darmstadt, Germany).

2.2. Sample preparation

Two sets of 42 and 36 samples with potentially three- and four-component combinations of oils were prepared. In ternary oil mixtures, 23 samples, and in quaternary oil mixtures, 21 samples were selected as calibration sets (Table 1), 19 samples were selected as validation set for ternary mixtures and 15 samples for quaternary (Tables 2 and 3). The fraction of each oil in the samples varied between 0 and 100%, with 20% steps for the quaternary mixtures and 10% ones for the ternary mixtures. The samples were treated as in the AOAC method for preparation of fatty acid solution [1]. A mass of 0.5 g well-mixed sample was transferred to a tube with lip and 5 ml of 50 mg/ml of KOH in isopropanol was added. The tube was placed 5 cm deep in a boiling water bath for 30 min to saponify the oil and to evaporate the isopropanol. A volume of 0.35 ml of a mixture of concentrated H_2SO_4/distilled H_2O (2:1) was added dropwise to the soap while the tube was cooled in cold water. Lumps at the bottom of the tube were broken using a glass-stirring rod. A yellow mixture containing fatty acids clinging to a viscous aqueous K_2SO_4 layer is obtained. A volume of 10 ml hexane–butanol (100:1) solution is added to the tube and mixed thoroughly with a glass rod. The organic solution was filtered using a 0.45 μm filter. A volume of 1 μl of the filtrate was analyzed by GC.

Table 1
Composition of the two calibration sets for either the mixtures with potentially three or four components (a = soybean oil; b = sunflower oil; c = olive oil; d = cottonseed oil)

Sample number	Fraction (%)						
	Ternary oil mixtures			Quaternary oil mixtures			
	a	b	c	a	b	c	d
1	60	20	20	40	0	40	20
2	20	20	60	60	0	40	0
3	40	20	40	0	0	0	100
4	30	50	20	0	40	40	20
5	20	30	50	20	40	20	20
6	30	30	40	0	20	40	40
7	10	0	90	20	20	40	20
8	10	90	0	100	0	0	0
9	100	0	0	40	40	0	20
10	0	20	80	0	40	60	0
11	50	30	20	20	0	40	40
12	30	40	30	20	40	40	0
13	90	0	10	20	20	60	0
14	0	0	100	20	60	0	20
15	0	30	70	0	100	0	0
16	30	70	0	40	40	20	0
17	50	0	50	40	20	0	40
18	40	10	50	0	60	20	20
19	0	70	30	20	0	60	20
20	10	40	50	0	0	100	0
21	40	40	20	0	80	20	0
22	20	40	40				
23	0	100	0				

2.3. GC analyses

GC analyses were performed on a model 1000 DPC gas chromatograph (DANI, Milano, Italy) equipped with a split/splitless capillary (1:40) injector and a flame ionization detector. Analytical separation was achieved on an AT-1000 capillary column (25 m × 0.32 mm i.d.) with 0.2 μm film thickness (Alltech, Illinois, USA). Nitrogen was used as carrier gas (with a constant flow rate of 1 ml/min). The air, hydrogen and auxiliary gas (N_2) pressures for detector were kept 1.1, 0.9 and 0.8 bar, respectively. Temperature setting was as follows: injector, 250 °C; detector, 280 °C. The oven temperature was held at 180 °C for

Table 2
Composition of the validation set for the ternary mixtures, their prediction by the PLS models and the percentage recovery found for the PLS predictions (a = soybean oil; b = sunflower oil; c = olive oil)

Sample number	Fraction added (%)			Fraction recovered (%)			Recovery (%)		
	a	b	c	a	b	c	a	b	c
1	20.0	60.0	20.0	20.2	58.0	20.5	101.1	95.8	102.3
2	0.0	60.0	40.0	−0.8	58.7	41.3		99.1	103.2
3	0.0	80.0	20.0	−3.2	82.4	19.5		101.8	97.6
4	20.0	50.0	30.0	15.9	51.5	31.8	79.3	102.8	105.9
5	0.0	50.0	50.0	−0.2	52.0	48.2		102.8	96.4
6	20.0	0.0	80.0	17.4	3.2	80.7	87.1		100.8
7	80.0	0.0	20.0	77.4	1.6	19.9	96.8		99.6
8	20.0	80.0	0.0	21.5	76.6	−0.3	107.6	95.1	
9	60.0	40.0	0.0	54.6	43.8	−0.5	91.0	109.6	
10	60.0	0.0	40.0	57.2	0.7	41.9	95.3		104.7
11	40.0	30.0	30.0	33.1	36.0	30.3	82.7	112.9	100.9
12	30.0	20.0	50.0	28.6	21.9	49.8	95.3	102.3	99.6
13	30.0	0.0	70.0	27.5	2.2	71.5	91.7		102.1
14	50.0	20.0	30.0	47.4	19.7	32.5	94.9	92.8	108.5
15	40.0	60.0	0.0	39.8	60.6	−2.4	99.6	96.9	
16	0.0	40.0	60.0	1.7	39.9	58.8		99.1	98.0
17	50.0	50.0	0.0	48.9	51.3	−2.2	97.8	97.2	
18	50.0	40.0	10.0	46.0	42.4	10.3	92.0	99.8	103.2
19	70.0	0.0	30.0	69.7	0.7	29.6	99.5		98.7
Mean recovery							94.1	100.6	101.4

2 min, then programmed to 240 °C at 22 °C/min, and held for 25 min at 240 °C. The peak areas were used to calculate the FA content (expressed as percentage of total fatty acids). The major fatty acids chosen to characterize the different oil mixtures were myristic (14:0), palmitic (16:0), palmitoleic (16:1), stearic (18:0), oleic (18:1), linoleic (18:2) and linolenic (18:3) acids. For computation purposes an IBM Pentium 100 computer and the Matlab program version of 4.2 (MathWorks, Natick, MA) were used.

3. Results

3.1. Fatty acid composition of vegetable oils

The chromatograms of the fatty acid profiles obtained from the four types of commercial vegetable oils are shown in Fig. 1. The fatty-acid fingerprints are different for the four oils. It allows to identify the type of oil and to detect adulteration by comparing peak areas and heights [28]. The obtained experimental values of fatty acid contents from

Table 3
Composition of the validation set for the quaternary mixtures, their prediction by the PLS models and the percentage recovery found for the PLS predictions (a = soybean oil; b = sunflower oil; c = olive oil; d = cottonseed oil)

Sample number	Fraction added (%)				Fraction recovered (%)				Recovery (%)			
	a	b	c	d	a	b	c	d	a	b	c	d
1	0.0	60.0	40.0	0.0	−1.1	60.5	39.3	−0.7		100.8	98.2	
2	40.0	20.0	20.0	20.0	46.2	16.1	19.7	20.0	115.4	80.7	98.6	100.2
3	20.0	40.0	40.0	0.0	20.9	37.5	39.4	0.1	104.7	93.7	98.6	
4	0.0	0.0	60.0	40.0	3.4	4.7	61.2	35.3			102.0	88.2
5	40.0	20.0	40.0	0.0	37.6	21.6	43.3	0.9	94.1	108.1	108.2	
6	0.0	20.0	40.0	40.0	−0.3	22.7	39.1	38.8		113.7	97.7	97.0
7	20.0	40.0	0.0	40.0	22.8	37.3	−2.5	35.2	114.0	93.4		88.0
8	0.0	20.0	60.0	20.0	3.7	22.2	58.3	19.1		111.2	97.2	95.6
9	0.0	20.0	80.0	0.0	4.3	21.6	77.9	0.2		107.9	97.4	
10	20.0	0.0	80.0	0.0	24.0	2.0	79.5	−0.2	119.8		99.4	
11	20.0	80.0	0.0	0.0	18.3	80.8	−2.4	1.4	91.4	101.0		
12	60.0	40.0	0.0	0.0	58.8	41.8	0.7	0.8	98.0	104.5		
13	0.0	40.0	20.0	40.0	−0.8	40.4	20.0	37.8		101.0	100.0	94.6
14	20.0	20.0	0.0	60.0	22.2	20.4	−1.7	57.8	111.0	102.1		96.3
15	60.0	20.0	20.0	0.0	59.4	20.8	21.7	2.1	99.0	103.8	108.7	
Mean recovery									105.3	101.7	100.5	94.2

these oils were in good agreement with the values proposed by the Codex alimentary standard [29]. The data are listed in Table 4. As can be seen, cottonseed oil compared to the other oils has a higher percentage of myristic and palmitic acid. Oleic acid is rich in olive oil and linolenic acid is more important in soybean oil. Sunflower oil and soybean oil showed similar FA composition but in soybean oil palmitic acid is about two times higher than in sunflower oil.

3.2. PLS analyses

Modeling and prediction of mixture composition was done by partial least squares regression, a technique widely used in the chemometric literature for multivariate calibration. The theory and practice of PLS has been described in detail [30]. The first step in the PLS determination of the different oils in their mixtures involves constructing the calibration matrix. In this work, the quantitative analyses of two sets of three-

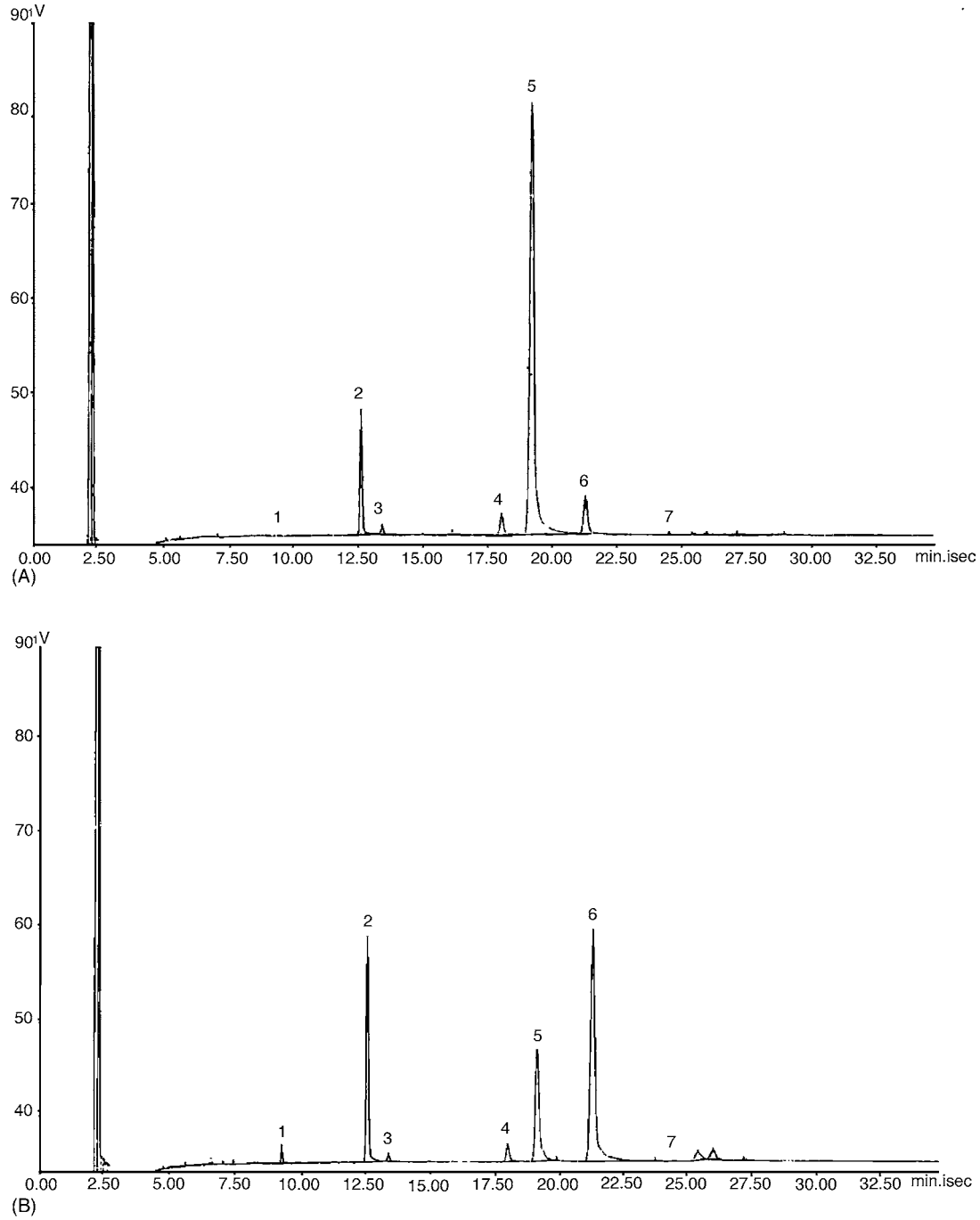

Fig. 1. Chromatographic fatty acid profiles of the edible oils: (A) olive oil; (B) cottonseed oil; (C) sunflower oil; (D) soybean oil. (1) Myristic acid; (2) palmitic acid; (3) palmitoleic acid; (4) stearic acid; (5) oleic acid; (6) linoleic acid; (7) linolenic acid.

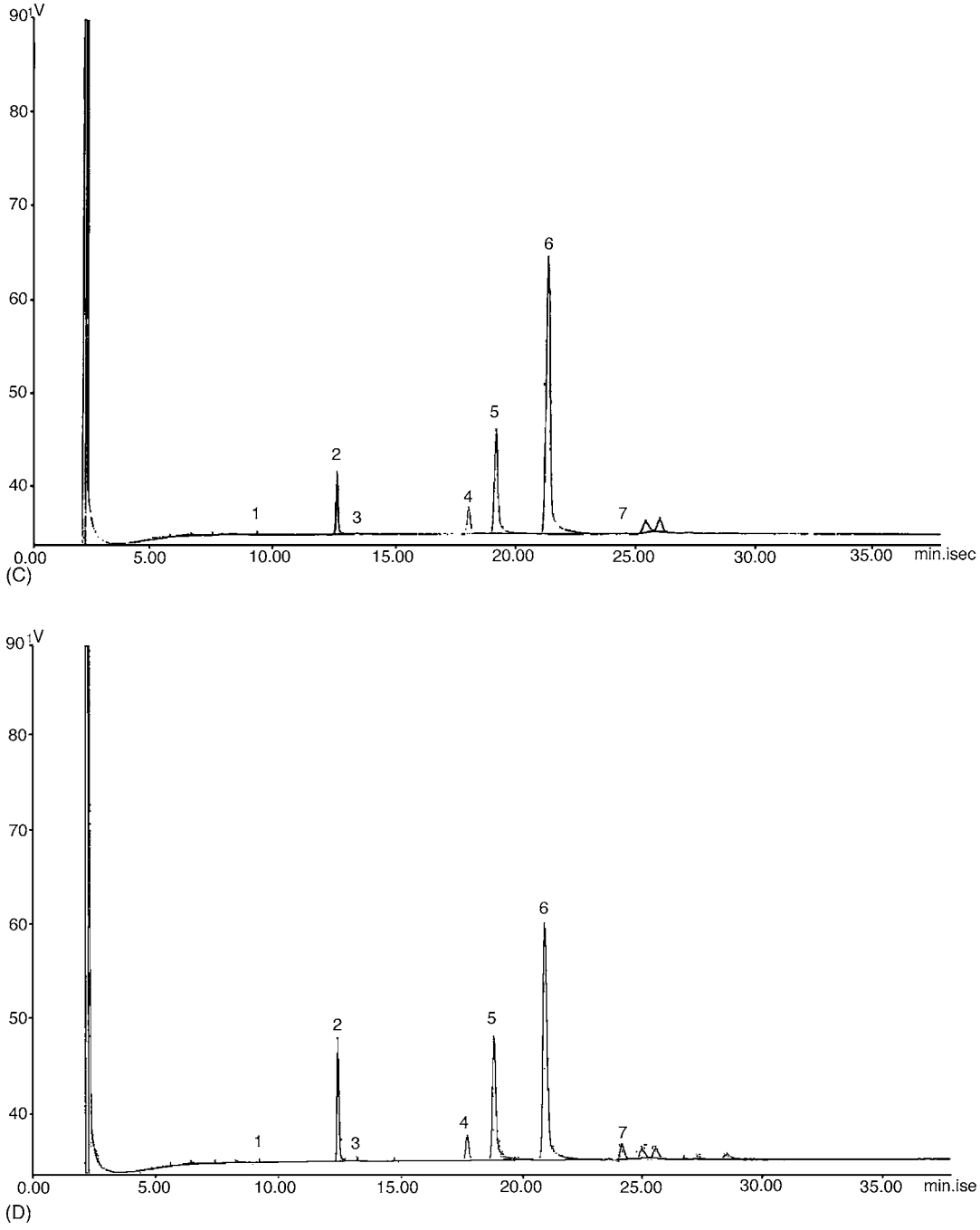

Fig. 1. (*Continued*).

Table 4
Fatty acid contents expressed as percentage of the total fatty acid content in edible oils

		Myristic acid	Palmitic acid	Palmitoleic acid	Stearic acid	Oleic acid	Linoleic acid	Linolenic acid
Cottonseed oil	Experimental value	1.1	22.0	0.7	2.6	21.0	49.3	0.0
	Codex	0.6–1	21.4–26.4	ND–1.2	2.1–3.3	14.7–21.7	46.7–58.2	ND–0.4
Sunflower oil	Experimental value	0.1	6.5	0.0	4.3	19.8	63.1	0.0
	Codex	ND–0.2	5–7.6	ND–0.3	2.7–6.5	14–39	48.3–74	ND–0.3
Olive oil	Experimental value	0.0	11.0	0.8	3.0	78.1	6.0	0.3
	Codex	ND–0.1	7.5–20	0.3–3.5	0.5–5	55–83	3.5–21	ND–1.5
Soybean oil	Experimental value	0.1	11.8	0.1	3.9	22.9	51.5	3.1
	Codex	ND–0.2	8–13.5	ND–0.2	2–5.4	17–30	48–59	4.5–11

ND = not detected.

Table 5
PRESS values for both mixtures with potentially three or four components (a = soybean oil; b = sunflower oil; c = olive oil; d = cottonseed oil)

Number of factors	Ternary oil mixtures			Quaternary oil mixtures			
	a	b	c	a	b	c	d
1	0.2926	0.2976	0.208	0.1567	0.1343	0.2078	0.1588
2	0.2751	0.1808	0.0046	0.1319	0.0874	0.0020	0.15
3	0.0041[a]	0.0028	0.0011	0.1341	0.0413	0.0010[a]	0.0447
4	0.0036[b]	0.0033	0.0011	0.0365	0.0019[a]	0.0009[b]	0.0181
5	0.004	0.0021[a,b]	0.0008[a,b]	0.0065[a]	0.0016[b]	0.0009	0.0046
6	0.0039	0.0023	0.0016	0.0063	0.0017	0.0009	0.0016[a,b]
7	0.0042	0.0021	0.0016	0.0048[b]	0.0019	0.0009	0.0023

[a] The first value of PRESS not significantly greater than the minimum one according to F-statistic comparison ($\alpha = 0.25$).
[b] The minimum value of PRESS for selected methods.

and four-component combinations of oils were performed. A first calibration set of 21 samples was selected by mixing maximally four oils and a second set of 23 samples was selected by mixing maximally three of them (Table 1). The normalized percentages of the C14:0, C16:0, C16:1, C18:0, C18:1, C18:2 and C18:3 fatty acids were chromatographically measured in these samples and used for constructing calibration matrixes. For each calibration set, PLS calculations performed modeling the fraction of each oil in a given mixture. To select the optimal number of factors for the PLS models, a leave-one-out cross-validation method, was employed [31]. The predicted and actual fractions of the oils in each sample were compared and the prediction error sum of squares (PRESS) was calculated as follows:

$$\text{PRESS} = \sum_{i=1}^{N}\sum_{j=1}^{M}(C_{ij} - \hat{C}_{ij})^2 \quad (1)$$

where C_{ij} is the true fraction of component i in sample j, and $\hat{C}_{ij}$ is its predicted fraction, N is total number of samples and M is total number of components used in the prediction set. The PRESS was calculated in the same manner each time a new factor was added to the PLS model. The PRESS values for the different models are shown in Table 5. A good criterion for selecting the optimum number of factors, involves the comparison of PRESS from model with the model which involves the number of factors yielding the minimum PRESS. Haaland and Thomas [32] used a F-statistic to make the significant determination and empirically determined that the number of factors for the first PRESS value whose F-ratio probability drops below 0.75 was selected as optimum.

3.3. Validation set

In order to test the performance of the proposed models, they were applied to the two validation sets of artificial mixtures with maximally three and four oils (Tables 2 and 3). The above-mentioned models were used to predict the fraction of each oil in the two validation sets of ternary and quaternary synthetic mixtures. The recovery of the system was calculated by comparing the composition of the synthetic samples and their prediction by the PLS models. The results are shown in Tables 2 and 3. The prediction error of each oil in these two validation sets was calculated as the relative standard error (RSE) of the prediction concentration [33]:

$$\text{RSE (single)} = 100 \times \left[\frac{\sum_{j=1}^{N}(C_j - \hat{C}_j)^2}{\sum_{j=1}^{N}C_j^2}\right]^{0.5} \quad (2)$$

where C_j is the fraction of the oil in the jth sample and $\hat{C}_j$ is the estimated fraction and N is total number of samples. The total prediction error for the mixtures is calculated as:

$$\text{RSE (total)} = 100 \times \left[\frac{\sum_{i=1}^{N}\sum_{j=1}^{M}(C_{ij} - \hat{C}_{ij})^2}{\sum_{i=1}^{N}\sum_{j=1}^{M}C_{ij}^2}\right]^{0.5} \quad (3)$$

where C_{ij} is the true fraction of the ith oil in the jth sample and $\hat{C}_{ij}$ is its estimation. N is total number of samples and M is total number of components used in the prediction set. Table 6 shows the single and total relative standard errors for both validation sets. It can be seen that the relative standard error for each oil in the mixtures was less than 10%. To compare precision in calibration and prediction steps, stan-

Table 6
Single and total relative standard error for the mixtures of the validation sets

	Ternary mixtures			Quaternary mixtures		
	RSE	SEC	SEP	RSE	SEC	SEP
Olive oil	3.48	0.008	0.006	3.80	0.013	0.008
Sunflower oil	4.70	0.013	0.010	6.36	0.014	0.012
Soybean oil	7.08	0.015	0.015	9.97	0.027	0.022
Cottonseed oil	–	–	–	7.80	0.011	
Total mixture	5.60			5.86		

dard error of calibration (SEC—data for training samples; Table 1) and standard error of prediction (SEP—data for validation sets; Tables 2 and 3) were calculated according to the formula:

$$\mathrm{SEC\,(SEP)} = \sqrt{\frac{1}{m-1}\sum_{j=1}^{n}(\hat{C}_j - C_j)^2} \quad (4)$$

The obtained results are presented in Table 6.

3.4. Outlier detection

It is important to choosing the optimum number of factors for the model to check that the calibration samples are representative of the unknown samples to predict or not, i.e., that the prediction samples are comparable to the calibration ones. If a prediction samples is different from the calibration samples, it can be considered to be an outlier. During prediction, outliers in the unknown samples can be detected from the chromatogram *F*-ratio. For unknown samples, the chromatogram *F*-ratios are calculated by using the final calibration of all *m* calibration samples [32]. Therefore, the chromatogram *F*-ratio for an unknown sample is given by:

$$F_{x,y}(a_s) = m\left(\frac{\sum_{k=1}^{n} e_{a_k k}^2}{\sum_{i=1}^{m}\sum_{k=1}^{n} e_{a_i k}^2}\right) \quad (5)$$

where a_s represents the chromatogram of the unknown sample and e_{a_s} the residue of it. The subscripts on $F(a_s)$, indicate the degrees of freedom and *n* is the number of peak included in the analysis. Unfortunately, there is some debate over the actual number of degrees of freedom to use for chromatogram residuals. Lindberg et al. [34] believe that $x = (n-r)/2$ and $y = (n-r)(m-r-1)/2$ are good estimates of *x* and *y*. In Table 7, the chromatogram *F*-ratios for the two ternary and quaternary validation sets are shown. On the other hand, to emphasize on the outlier detection method, the quaternary validation set was predicted by the supposed ternary training set. The chromatogram *F*-ratios for each sample are calculated and are shown in Table 7. The chromatogram *F*-ratios can be used as guide in the calibration and prediction steps to flag possible outlier samples. This is especially important when these methods are used for quality control application.

3.5. Application

The above method can be applied to the determination of the olive, soybean, cottonseed and sunflower oil fractions in mixtures. The technique is interesting since olive oils are frequently adulterated with other vegetable oils. Therefore, five natural olive oils were spiked with a combination of cottonseed, sunflower and soybean oils. These mixtures have a composition as described in Table 8. They were treated as in the experimental sections and the extracts obtained were analyzed by the GC method. Given the fatty acid contents of the mixtures, the proposed PLS approach was performed and the mixture composi-

Table 7
Chromatogram *F*-ratios for ternary and quaternary validation sets (a = soybean oil; b = sunflower oil; c = olive oil; d = cottonseed oil)

Sample number	Chromatogram *F*-ratio									
	Ternary mixtures[a]			Quaternary mixtures[b]				Quaternary mixtures[c]		
	a	b	c	a	b	c	d	a	b	c
1	0.03	0.01	0.02	0.01	0.03	0.04	0.07	0.59	0.86	0.88
2	0.03	0.04	0.03	0.03	0.03	0.05	0.01	3.33	6.11	8.50
3	0.03	0.01	0.03	0.01	0.03	0.04	0.02	0.44	0.59	0.21
4	0.02	0.01	0.01	0.08	0.07	0.06	0.02	14.86	29.88	32.09
5	0.02	0.01	0.03	0.01	0.03	0.05	0.02	0.44	0.56	0.22
6	0.06	0.03	0.02	0.02	0.03	0.05	0.04	19.29	40.37	39.00
7	0.04	0.04	0.08	0.06	0.06	0.03	0.02	13.85	27.89	29.75
8	0.07	0.12	0.10	0.06	0.05	0.05	0.02	4.67	7.27	9.60
9	0.05	0.03	0.02	0.13	0.09	0.06	0.03	1.21	0.68	0.78
10	0.02	0.03	0.04	0.13	0.10	0.06	0.02	1.20	0.79	0.51
11	0.07	0.13	0.19	0.08	0.08	0.03	0.04	1.28	0.82	0.39
12	0.04	0.04	0.09	0.08	0.07	0.05	0.02	1.43	1.86	0.38
13	0.04	0.02	0.04	0.01	0.03	0.05	0.04	18.64	37.69	38.09
14	0.03	0.01	0.05	0.05	0.04	0.03	0.03	38.97	86.59	83.19
15	0.09	0.03	0.14	0.02	0.03	0.06	0.01	0.63	0.98	0.06
16	0.03	0.02	0.02							
17	0.10	0.05	0.18							
18	0.07	0.04	0.15							
19	0.07	0.04	0.19							

[a] Chromatogram *F*-ratios for ternary validation sets predicted by ternary training set.
[b] Chromatogram *F*-ratios for quaternary validation sets predicted by quaternary training set.
[c] Chromatogram *F*-ratios for quaternary validation sets predicted by ternary training set.

Table 8
Simultaneous determination of oils by PLS models (a = soybean oil; b = sunflower oil; c = olive oil; d = cottonseed oil)

Sample	Content				Found with PLS method			
	a	b	c	d	a	b	c	d
1	10	0	90	0	10.72	1.66	88.23	1.57
2	20	20	60	0	19.86	21.24	61.78	0.38
3	0	40	50	10	−2.54	41.62	48.04	11.2
4	20	0	60	20	17.89	1.06	59.96	19.4
RSE					2.33	7.23	7.85	9.85

The values given are percentages of the total oil.

tions were predicted. The analytical results are listed in Table 8.

4. Discussion

Inclusion of high-quality vegetable oils is of great importance in a fully balanced diet. Hence, there is a need for analytical techniques to control the purity and authenticity of vegetable oils. The aim of this paper is to establish the level of different oils present in a mixture, based on the fatty acid profiles determined for these mixtures. The chromatograms of the fatty acid profiles obtained from the four types of commercial vegetable oils (Fig. 1) indicated that the fatty-acid fingerprints are different for the four oils and determining the FA composition might be useful for the fingerprinting of oils and for detecting their authenticity. A GC method without derivatization process in combination with PLS was evaluated to determine the levels of the different oils present in the mixtures. In this, work the quantitative analysis of two sets of three- and four-component combination of oils were performed. The PRESS values obtained by optimizing the calibration mixtures are listed in Table 5. It can be seen that generally, the optimum number of factors is greater than the number of components in the mixtures. High number of factors indicated obtained calibration models could be over fitted. In order to clarify this point, two independent set of validation samples were prepared (Tables 2 and 3) and the prediction of oils was carried out in these validation samples using calibration models evaluated. As can be seen in Tables 2 and 3, good recovery values (the last row) were obtained for the components of all mixtures under study. To compare precision in calibration and prediction steps, standards error of calibration and standard error of prediction were calculated and the obtained results are presented in Table 6. *F*-statistics test was applied to compare SEC and SEP values. It was found that for two ternary and quaternary mixtures there is no statistically significant difference between precision in calibration steps and in prediction step ($p < 0.05$), so the respective calibration models have satisfactory prediction ability with no over fitting problems. The above method can be applied to the determination of the olive, soybean, cottonseed and sunflower oil fractions mixtures. For the analyzing of unknown samples, it is important to check that the calibration samples are representative of the unknown samples to predict or not. During prediction, outliers in the unknown samples can be detected from the chromatogram *F*-ratio. In Table 7, chromatogram *F*-ratios for ternary and quaternary validation sets are shown. Chromatogram *F*-ratios for quaternary validation sets predicted by ternary training set (last three columns) are significantly higher in samples number 2, 4, 6, 7, 8, 13 and 14. As can be seen in Table 3, in these samples there are cottonseed oils and therefore these samples cannot be detected with ternary training set. Chromatogram *F*-ratios in samples with 60% of cottonseed (sample 14) and in samples with 40% of cottonseed (samples 4, 6, 7 and 13) indicate a highly significant outlier ($p < 0.01$). Chromatogram *F*-ratios for olive oil in samples with 20% of cottonseed (samples 2 and 8) also indicate a highly significant outlier ($p < 0.01$) but in sunflower and soybean oil chromatogram *F*-ratios are significantly high with $p < 0.05$ (Table 7). Therefore, it would be concluded that chromatogram *F*-ratios can be used to flag possible outlier samples and the above method can be applied to the determination of the olive, soybean, cottonseed and sunflower oil fractions in the mixtures. The relative standard error of each oil and the total prediction error for the mixtures are shown in Table 6. As can be seen, the relative standard error of each oil in ternary mixtures is lower than quaternary mixtures and this is evident especially for soybean and sunflower oil. Therefore, it would be better to analyze unknown samples with chromatogram *F*-ratios lower than ternary mixtures with ternary training set but these differences are not significant ($p < 0.05$). The relative standard error for olive oil in both ternary and quaternary mixtures is lower than the other oils and it can be explained with due attention to olive oil chromatogram and its high percentage of oleic acid. The analytical results of five synthetic samples (Table 8) indicate that the predicted concentration are comparable with the actual ones, therefore the analysis in real samples can be used to determine adulteration. This study is especially interesting since olive oils are frequently adulterated with other vegetable oils.

5. Conclusion

The present study demonstrated that the chemometric approach using PLS based on GC data of fatty-acid fingerprint

of oils and oil mixtures allows predicting the composition of unknown oil mixtures. This approach is very useful for an important aspect of vegetable oil quality control namely the detection of adulteration.

Acknowledgment

This study was supported by a grant from the Ministry of Health and Medical Education.

References

[1] W. Harwitz (Ed.), Official Methods of Analysis of AOAC International, 17th ed., AOAC International, Maryland, USA, 2000, p. 15 (Chapter 41).
[2] A. Townshend, Encyclopedia of Analytical Science, vol. 3, second ed., Academic Press, London, 1995, 1540 pp.
[3] E.P. Labrinea, N.S. Thomaidis, C.A. Georgiou, Anal. Chim. Acta 448 (2001) 201.
[4] S.A. Passaloglou-Emmanouillidou, Z. Lebensm Unters. Forsch. 191 (1990) 132.
[5] A.N. Davies, P. Mclintyre, E. Morgan, Appl. Spectrosc. 54 (2000) 1864.
[6] V. Baeten, P. Hourant, M.T. Morales, R. Aparicio, J. Agric. Food Chem. 46 (1998) 2638.
[7] G. Valhov, Prog. Nucl. Magn. Reson. Spectrosc. 35 (1999) 341.
[8] T. Sato, Lipid Technol. 9 (1997) 46.
[9] Y.W. Lai, E.K. Kemsley, R.H. Wilson, J. Agric. Food Chem. 42 (1994) 1154.
[10] G.F. Spencer, J. Am. Oil Chem. Soc. 53 (1976) 94.
[11] M. Hein, H.D. Isengard, Z. Lebensm Unters. Forsch. A 204 (1997) 420.
[12] G.M. Wood, P.T. Slack, J.B. Rossell, P.J. Mann, P.J. Famell, J. Agric. Food Chem. 42 (1994) 2525.
[13] J. Mannina, M. Patumi, P. Fiordiponti, M.C. Emanuele, A.L. Segre, Ital. J. Food Sci. 2 (1999) 139.
[14] S.E. Woodburg, R.P. Evershed, J.B. Rossell, J. Am. Oil Chem. Soc. 75 (1998) 371.
[15] P.R. Brown, J.M. Beebe, J. Turcotte, Anal. Chem. 21 (1989) 193.
[16] J.H. Ha, W.D. Hawer, J.B. Hwang, Korean J. Food Sci. Technol. 25 (1993) 345.
[17] J. Vicanova, E. Tvrzicka, K. Stulik, J. Chromatogr. B 656 (1994) 45.
[18] S.D. Brown, R.S. Bear, T.B. Blank, Anal. Chem. 64 (1992) 22R.
[19] S.Y. Bae, D.S. Lee, J. Korean Chem. Soc. 41 (1997) 105.
[20] D.S. Lee, B.S. Noh, S.Y. Bae, K. Kim, Anal. Chim. Acta 358 (1998) 163.
[21] S. Husain, K. Sita Devi, D. Krishna, P.J. Reddy, Chemom. Intell. Lab. Syst. 35 (1996) 117.
[22] Y.G. Martin, J.L.P. Pavon, B.M. Cordero, C.G. Pinto, Anal. Chim. Acta 384 (1999) 83.
[23] D.S. Lee, E.S. Lee, H.J. Kim, S.O. Kim, K. Kim, Anal. Chim. Acta 429 (2001) 321.
[24] M. Tsimidou, R. Macrae, I. Wilson, Food Chem. 25 (1987) 251.
[25] P. Kaufmann, Anal. Chim. Acta 277 (1993) 467.
[26] R. Marbach, H.M. Heise, Trends Anal. Chem. 11 (1992) 270.
[27] E.V. Thomas, D.M. Haaland, Anal. Chem. 62 (1990) 1091.
[28] G.F. Spencer, W.F. Kwolek, L.H. Princen, J. Am. Oil Chem. Soc. 56 (1979) 972.
[29] Joint FAO/WHO Report of the 17 Session of the Codex Alimentarius Committee on Fats and Oils, London, UK, February 2001.
[30] P. Geladi, B.R. Kowalski, Anal. Chim. Acta 185 (1986) 1.
[31] M.J.R. Stone, Stat. Soc. 36 (1974) 111.
[32] D.M. Haaland, E.V. Thomas, Anal. Chem. 60 (1988) 1193.
[33] M. Otto, W. Wegscheider, Anal. Chem. 57 (1985) 63.
[34] W. Lindberg, J.A. Persson, S. Wold, Anal. Chem. 55 (1983) 643.