Received: 4 November 2022 | Revised: 7 March 2023 | Accepted: 14 March 2023

DOI: 10.1002/aocs.12697

ORIGINAL ARTICLE

AOCS WILEY

Identification of Turkish extra virgin olive oils produced in different regions by using NMR (^{1}H and ^{13}C) and IRMS (^{13}C/^{12}C)

Didar Sevim[1] | **Oya Köseoğlu**[1] | **Hasan Ertaş**[2] | **Durmuş Özdemir**[3] | **Mehmet Ulaş**[1] | **Salih Günnaz**[4] | **Veysel Umut Çelenk**[5]

[1]Olive Research Institute, Ministry of Agriculture and Forestry, Izmir, Turkey

[2]Faculty of Science Department of Chemistry, Ege University, Izmir, Turkey

[3]Faculty of Science Department of Chemistry, Izmir Institute of Technology, Izmir, Turkey

[4]NMR Satellite Laboratory, Ege University Science and Technology Centre, Izmir, Turkey

[5]Research and Application Center of Drug Development and Pharmacokinetics, Ege University, Izmir, Turkey

Correspondence
Didar Sevim, Olive Research Institute, Ministry of Agriculture and Forestry, Izmir, Turkey.
Email: dcengeler@gmail.com

Funding information
General Directorate of Agricultural Research and Policies (TAGEM) of Turkey, Grant/Award Number: TAGEM/HSGYAD/15A05/P01/75

Abstract

Isotope ratio mass spectroscopy (IRMS) and nuclear magnetic resonance (NMR) spectroscopy techniques are two of the analytical methods that are used to characterize food products. The aim of this study is to classify extra virgin olive oil (EVOO) samples collected from different regions of Turkey based on ^{1}H and ^{13}C NMR spectra along with IRMS δ^{13}C carbon isotope ratio data by using chemometrics multivariate data analysis methods. A total of 175 EVOO samples were analyzed in 2014/15 and 2015/16 harvest seasons. Multivariate classification and clustering models were used to identify geographical and botanical origins of the EVOOs. IRMS results showed that there was no significant difference in terms of δ^{13}C values between the years in terms of harvest year ($p > 0.05$), only extraction phase and variety were statistically significant factors ($p < 0.05$). The interactions of the factors showed that the *harvest year* × *variety* interaction is important. The outcomes of this research clearly indicated that considering the partial least squares discriminant analysis result with NMR spectra, the percent success of the model in the South Marmara, North Aegean, and South Aegean region samples were 95%, 95.7%, and 96.4% in the model set, respectively. The results showed that by using classification and clustering models, geographic marking and labeling of these oils can be carried out regardless of differences in year and production systems (2 and 3 phase extraction system) according the NMR analysis.

KEYWORDS
classification, clustering, extra virgin olive oil, IRMS, NMR, δ^{13}C carbon isotope ratio

INTRODUCTION

EVOO, obtained solely through mechanical and physical processes from the fruit of *Olea europea* L., is a significant agricultural product of the Mediterranean regions. It is also responsible for the well-known health benefits in the Mediterranean regions (Martín-Peláez et al., 2013). Due to the increasing awareness about healthy food consumption, olive oil has become one of the most important product of the food industry. The olive oils obtained from different varieties are marketed by blending or as a single variety. Thus, products with different flavor and type are provided. This also enriches the production of varieties according to the region. Recently, olive oil with protected denominations of origin (PDO) and protected geographical indication (PGI) began to be sold. This requires a precise definition of many parameters such as variety, geographical origin, agronomic applications, production technology and sensory quality. The quality of these monovarietal oils relates to superior taste, consistency, color, and direct olive variety. Therefore, studies have focused on defining their origin which can be cumbersome for identification (Rabiei & Enferadi, 2013).

Nowadays, the definition of geographical origin in EVOO has been one of the frequently encountered questions. The objective of many studies has been to verify the identity of EVOO with objective analytical

parameters and to prove its exact geographical origin. The composition of EVOO is influenced by many parameters such as olive variety, environmental conditions, fruit maturity and extraction technology. Many factors therefore need to be considered to accurately define the origin based on its chemical composition. This makes it difficult to identify the required markers (Gumus et al., 2017; Merchak et al., 2018). Many methods are used to determine the authenticity of olive oil. The most commonly used techniques are chemical composition, stable isotopes and DNA (González-Casado et al., 2013). Isotopic methods are based on the determination of isotope ratios of naturally occurring C, H, and O atoms in foods. Isotope ratio of foods is mostly determined by IRMS and NMR techniques. IRMS and NMR techniques are used in the characterization and reality control of wine, olive oil and fruit juice products (Ogrinc et al., 2003). ^{1}H NMR fingerprint in olive oil is an analytical tool used in traceability (food authentication and food quality) of olive oils. ^{1}H NMR fingerprint analysis of olive oils for identification purposes allows the determination of geographical origin at the national, regional and/or PDO level, and fingerprint techniques such as NMR are used to determine the identity of foods. ^{1}H and ^{13}C NMR analyzes are used in bulk oils and unsaturated fractions of oils in order to distinguish certain geographical origins of EVOOs in combination with multivariate techniques (Alonso-Salces et al., 2012). Fauhl et al. (2000) reported that ^{1}H NMR analysis was used to identify the geographical origin of Italian olive oils and to classify olive tree varieties, while ^{13}C NMR analysis was also used to classifiy of olive oil. In addition, they studied the olive oil with the combined approach of NMR and IRMS techniques and, it was stated that ^{1}H NMR results provided qualitative and quantitative information about major and minor compounds in oil. The IRMS technique determines the ratio of two stable isotopes (^{13}C/^{12}C) of an element in a product in relation to agricultural applications, soil, and climate (Calò et al., 2022). Official methods for specific authentication of EVOOs, such as geographic origin assessment, are still lacking today. Among the useful analytical techniques for geographical origin certification and authentication, ^{1}H and ^{13}C NMR and IRMS, along with chemometry, are widely used (Calò et al., 2022; Jiménez-Morillo et al., 2020; Ogrinc et al., 2003). As known EVOOs show significant differences in sensory, nutritional and functional characteristics (aromas, polyphenols, and tocopherols) related to the olive variety, production technology and geographical origin. In determining the authenticity of olive oil, it is very important to establish a database that provides isotopic information that can be obtained by NMR and IRMS techniques.

Considering the increasing commercial importance of olive oil due to its high sensory and nutritional quality, fraud made in the form of a mixture of low quality foreign oils, esterified, and refined oils unfortunately draw attention. It is now necessary to apply isotopic methods for the purpose of authenticity control of foods and determination of geographical origin. There are limited number of studies about the characterization of olive oils in terms of geographical origin and variety in our country such as Ok (2014), Dağ (2016) and Özdemir et al. (2018). In order to protect similar products and imitations of products of a specific region, which are of typical quality, a reference database is available for the olive oil sector. In spectroscopic techniques such as NMR and IRMS, it is very important to establish reliable accurate calibration models because the analysis results are dependent on reference analysis.

It was shown that the isotopic fractionation of the elements for the geographical discrimination of olive oil is highly correlated with the geographical and climatic parameters (Nasr et al., 2022). NMR technique has many advantages such as; high reproducibility, profitable use for fingerprint analysis, fast measurement, minimal sample preparation, long-term sample storage, and suitable for untargeted and targeted analysis and many components and quality characteristics can be simultaneously analyzed in a few minutes at the same time (Emwas et al., 2019). IRMS has the advantage of being able to be used in old or degraded samples because they refer to element isotopes (typically C, O, and H) regardless of which compounds they are in and giving results independent of variety and production technique (Bontempo et al., 2019; Calò et al., 2022). In addition, it is an alternative to traditional methods in the classification of olive oils according to geographical origin. These techniques are frequently associated with chemometric methods that combine metabolomics with statistical analysis of spectroscopic chemical data. It is critical to use these methods and obtain reliable results by combining them with appropriate chemometric methods (Calò et al., 2022).

The aim of this study is to determine the geographical origin of EVOOs produced in different regions by δ^{13}C (carbon isotope-IRMS) and ^{1}H and ^{13}C NMR techniques during 2014/15–2015/16 harvest years. One hundred and seventy-five samples collected from four different geographic regions were analyzed. Multivariate classification and clustering models were used to identify geographical and botanical origins of the EVOOs.

MATERIALS AND METHODS

Extra virgin olive oil samples

A total of 175 EVOO samples were analyzed in this work (Table 1).

The samples were collected from the Aegean (Northland South), Marmara (South), Southeastern

TABLE 1 Distribution of collected samples by harvest years and regions

Regions	Variety	2014/15		2015/16		Total
		2 phase	3 phase	2 phase	3 phase	
Aegean (South)	Memecik	6	5	15	14	40
Aegean (North)	Ayvalık	22	22	12	13	69
Marmara (South)	Gemlik	–	19	–	9	28
Mediterranean (West)	Gemlik, Tavşan Yüreği	–	4	–	2	6
Southeastern Anatolia	Sarı Ulak, Ayvalık, Gemlik, Esek Zeytini, Nizip Yağlık, Sarı Haşebi, Kilis Yağlık, Saurani	3	3	4	12	22
Total		84		81		165
		Abencor		Abencor		
Aegean, Marmara, Mediterranean	Ayvalık, Gemlik, Memecik, Domat, Sarı Ulak	5		5		10
Total		89		86		175

FIGURE 1 Geographic regions where EVOO samples were collected (Marmara, Aegean, Mediterranean, and Southeastern Anatolia regions)

Anatolia, and Mediterranean (West) regions (Figure 1) in Turkey from the most common olive mills as follows: two and three phase extraction systems during 2014/15 and 2015/16 harvest seasons between October and December. The collection of samples from the regions was carried out in a controlled by Olive Research Institute of Ministry of Agriculture and Forestry of Turkey.

EVOOs samples (165 olive samples) were collected from the regions as listed at Table 1. Ayvalık EVOOs were collected from the North of Aegean region, Memecik EVOOs were collected from the South of Aegean region, Gemlik EVOOs were collected from the South of Marmara region, Sarı Ulak, Ayvalık, Gemlik, Eşek Zeytini, Sarı Haşebi, Nizip Yağlık, Kilis Yağlık, and Saurani EVOOs were collected from the Southeastern Anatolia region and Gemlik, Tavşan Yüreği EVOOs were collected from the Western Mediterranean region in Turkey. For each variety, 500 mL of EVOO was obtained from the factories. The samples were stored in amber-glass bottles, labeled with the code, without headspace and kept at 4°C in the dark until analysis.

In addition, as seen Table 1, a total of 10 olive varieties, including Ayvalık, Memecik, Gemlik, Domat and Sarı Ulak olive varieties, were harvested from their own regions of origin (five olive variety in the 2014/15 harvest year and five olive variety in the 2015/16 harvest year), and oils were obtained by Abencor system (MC2 Ingenieria y Sistemas, Sevilla, Spain). Olive samples (10 kg) were washed after being separated from their leaves and crushed in a crusher. It was subjected to mixing (malaxation) at 25°C for 30 min. After malaxation, a centrifuge was used to separate the oil from the olive paste. The obtained oils were stored under controlled conditions in amber colored bottles at 4°C until the analyses were done after filtering.

$\delta^{13}C/^{12}C$ stable isotope (IRMS) analysis

Determination of $\delta^{13}C$ isotope ratios in EVOOs was carried out according to the modified version of AOAC 998.12 (2005) method using an Isotope Ratio Mass Spectrometer (Micromass, IsoPrime) connected with Dumas combustion for an elemental analyzer (Euro Vector) for $\delta^{13}C$ isotope ratios. The results were expressed in δ‰ (parts per thousands) versus VPDB (Vienna Pee Dee Belemnite) $\delta^{13}C$. The samples were weighed (150 μg) into the capsules using a precision balance before being taken to the elemental analyzer. Then, isotope ratios of target sample are measured on the basis of acceptable universal standards. The samples were analyzed in two replicates (Wada, 2009).

$$\delta^{13}C\,(‰) = \left[\left({}^{13}C/{}^{12}C_{sample}/{}^{13}C/{}^{12}C_{std}\right) - 1\right] \times 1000.$$

Nuclear magnetic resonance (NMR) analysis

^{1}H and ^{13}C NMR analysis were performed on a liquid MERCURYplus-AS 400 model NMR (Agilent) spectrometer with 400 MHz operating frequency (Table 2)

TABLE 2 Working conditions of MERCURYplus-AS 400 model NMR spectrometer

	^{1}H NMR	^{13}C NMR
System frequency	399.88 MHz	100.56 MHz
Number of scans	8	128
Receiver gain	6	30
Relaxation Delay	1.000 s	1.000 s
Spectral width	6398.0 Hz	25125.6 Hz
Acquisition time	2.5608 s	1.3042 s

according to Sacchi et al. (1997). All experiments were conducted at 25 ± 1°C. For each sample, 80 μL of oil was placed in a 5-mm NMR tube, and 420 μL deuterated chloroform ($CDCl_3$) was added to obtain 20% oil samples. The oil samples were homogenized by vortexing for 30 s at 1000 rpm. Table 2 shows the working conditions of NMR spectrometer.

Statistical analysis

Multivariate classification methods allow development of mathematical and statistical models based on raw spectral or chromatographic data. The aim is to develop the best model which successfully assigns new samples to the correct classes. These methods can be divided into two sub-groups: supervised classification and unsupervised classification. In unsupervised classification, there is no prior information about which group each sample belongs to. Thus, as a rule of thumb, the initial data is subjected to various data processing such as principle component analysis (PCA) and hierarchical cluster analysis (HCA) to provide a general idea of how the samples are grouped based on the differences in their variance and/or means.

Unlike unsupervised classification methods, to obtain a model with a supervised classification method one must also provide which sample belongs to which class in a training data set. Then, in most cases, the classification problem becomes weighting of variables to provide best separation. Among these methods partial least squares discriminant analysis (PLS-DA) is one of the most common tool in chemometrics. PLS-DA is a projection based classification method and may provide noise reduction of spectral data as well as decreasing the number of variables significantly to overcome multicollinearity problem. The aim of this method is to find the best projection of the explanatory variables and the responses that maximizes their covariance. In this study, analysis of variance (ANOVA) was used to statistically evaluate data obtained from the IRMS analysis and chemometrics multivariate methods such as PCA and PLS-DA were applied for classification of oils according to geographical origin and harvest year based on the data obtained from ^{1}H and ^{13}C NMR analysis. ANOVA and PCA data analysis was carried out with Minitab statistical software (Minitab 16 Statistical Software, Minitab, Inc. State College, PA, USA) program whereas PLS-DA supervised classification was performed with customer developed algorithm in Matlab programming software (MATLAB and Statistics Toolbox Release 2018b, The MathWorks, Inc., Natick, MA, USA).

RESULTS AND DISCUSSION

$\delta^{13}C/^{12}C$ stable isotope (IRMS)

In this research, δ^{13}C results of a total of 175 EVOO samples were evaluated, 89 from the harvest year of 2014/15, 86 from the harvest year of 2015/16 (Table 2). $\delta^{13}C/^{12}C$ values of 175 samples are given in Table 3.

The δ^{13}C values of Ayvalık, Memecik, Gemlik, Domat, Sarı Ulak, Kilis Yağlık, Tavşan Yüreği, Eşek Zeytini, Nizip Yağlık, and Saurani were seen at the Table 3 for 2014/15 and 2015/16 harvest years. The δ^{13}C values for Ayvalık varied between −30.0‰ and −27.3‰ and between −29.0‰ and −28.15‰, for Memecik varied between 30.7‰ and −28.25‰ and between −30.55‰ and −28.30‰, for Gemlik varied between −31.25‰ and −29.15‰ and between −30.95‰ and −27.85‰ for the 2014/15 and 2015/16 harvest years, respectively. In the research conducted by Gumus et al. (2017) in Turkey stated that the δ^{13}C value of olive oils obtained from Ayvalık, Gemlik, Memecik and Uslu (49 olive oils) varieties ranged between −30.4‰ and −27.7‰. The δ^{13}C results of samples are in agreement with Gumus et al. (2017).

It was seen in the study that the samples other than Ayvalık, Gemlik and Memecik varieties were mostly composed of one sample each with different species and geographical origin. In this context, in order to make a healthy evaluation, ANOVA was performed with samples in three groups consisting of only Ayvalık, Gemlik, and Memecik varieties. Three-factor mixed ANOVA was applied as the harvest year (two levels as 2014/15 and 2015/16), variety (three levels as Ayvalık, Gemlik, and Memecik) and production systems (two levels, 2 and 3 phases). The *p*-values, showed that phase and variety were important factors at 95% confidence level among the three factors evaluated (Harvest year, Phase and Variety) (*p*-value <0.05). When we considered the results in terms of harvest year, it was seen that there is no significant difference between the δ^{13}C values between the 2 years (*p*-value>0.05). When examined the interactions of the factors, it was seen that *the harvest year*variety* interaction was important (*p*-value <0.05) and the others were not.

Portarena et al. (2015) stated that the isotope compositions in olive oil are determined by the variety and maturation stage, and these should be taken into account in traceability studies. Baum et al. (2010)

TABLE 3 $\delta^{13}C$ Values of EVOOs according to harvest years (‰) (2014/15 and 2015/16)

2014/15						2015/16					
EP	**V**	**$\delta^{13}C$**	**EP**	**V**	**$\delta^{13}C$**	**EP**	**V**	**$\delta^{13}C$**	**EP**	**V**	**$\delta^{13}C$**
3	G	−31.25	3	TY	−29.05	3	G	−30.95	2	M	−29.1
3	G	−31	3	TY	−29.05	2	M	−30.55	3	G	−29.05
3	G	−30.7	2	EZ	−29	2	M	−30.55	3	A	−29
3	M	−30.7	2	A	−29	3	G	−30.45	3	G	−29
3	G	−30.7	3	A	−29	2	M	−30.4	2	EZ	−29
3	G	−30.65	2	A	−28.9	3	M	−30.35	3	G	−29
AS	D	−30.5	2	A	−28.9	3	G	−30.3	3	A	−28.95
3	G	−30.45	2	A	−28.9	3	G	−30.3	3	A	−28.9
3	G	−30.4	2	A	−28.9	3	S	−30.3	2	SU	−28.9
3	G	−30.3	3	A	−28.8	3	S	−30.3	3	A	−28.85
3	G	−30.25	2	A	−28.8	2	M	−30.25	2	A	−28.8
3	G	−30.25	2	A	−28.8	3	M	−30.25	3	G	−28.8
3	G	−30.15	2	A	−28.8	3	M	−30.25	3	SU	−28.75
3	G	−30.15	2	A	−28.75	3	G	−30.2	3	A	−28.7
3	G	−30.1	2	A	−28.8	3	S	−30.1	2	A	−28.7
3	G	−30.1	2	A	−28.75	2	M	−30.05	3	KY	−28.7
3	G	−30.1	3	A	−28.75	2	M	−30	2	A	−28.7
2	M	−30.05	2	A	−28.7	3	M	−29.95	2	SU	−28.65
3	A	−30	2	A	−28.7	2	M	−29.9	AS	A	−28.6
2	M	−30	3	A	−28.6	2	M	−29.85	2	A	−28.6
3	G	−29.95	3	A	−28.5	3	M	−29.85	3	A	−28.55
3	G	−29.95	2	A	−28.5	3	M	−29.85	3	A	−28.55
3	G	−29.95	3	A	−28.5	3	SH	−29.85	2	A	−28.4
3	G	−29.95	2	A	−28.5	3	G	−29.8	3	A	−28.4
3	G	−29.95	2	A	−28.5	AS	D	−29.75	2	A	−28.4
3	A	−29.95	2	M	−28.45	3	M	−29.75	2	A	−28.4
AS	SU	−29.85	3	A	−28.4	AS	M	−29.7	3	A	−28.4
2	M	−29.8	3	A	−28.4	3	G	−29.7	2	A	−28.4
3	A	−29.7	2	A	−28.35	3	M	−29.7	3	A	−28.35
3	A	−29.55	2	A	−28.3	3	M	−29.65	3	A	−28.3
3	A	−29.55	3	A	−28.3	3	M	−29.55	2	A	−28.3
3	M	−29.55	2	M	−28.3	3	G	−29.5	3	KY	−28.3
3	M	−29.5	3	KY	−28.3	3	M	−29.5	2	M	−28.3
3	A	−29.35	2	A	−28.3	2	M	−29.45	2	M	−28.3
2	A	−29.3	2	M	−28.25	2	M	−29.45	AS	SU	−28.25
3	M	−29.3	2	SU	−28.25	2	M	−29.45	2	A	−28.25
3	M	−29.25	2	SU	−28.25	3	TY	−29.45	3	A	−28.25
2	A	−29.2	3	A	−28.2	3	M	−29.4	3	A	−28.25
3	G	−29.2	3	A	−28.15	3	M	−29.4	2	A	−28.25
3	G	−29.2	3	A	−28.1	3	M	−29.35	3	A	−28.2
AS	G	−29.15	3	A	−28	2	SU	−29.25	2	A	−28.15
3	A	−29.15	AS	A	−27.95	3	NY	−29.2	AS	G	−27.95
AS	M	−29.1	2	A	−27.3	2	M	−29.1	3	G	−27.85
3	A	−29.1	3	KY	−27.3						
3	A	−29.1									

Abbreviations: A:Ayvalık; AS: Abencor System; D: Domat; EP: extraction phase; EZ: Esek Zeytini; G:Gemlik; KY: Kilis Yağlık; M:Memecik; NY: Nizip Yağlık; S: Saurani; SH: Sarı Haşebi; SU: Sarı Ulak; TY: Tavşan Yüreği; V:Variety.

reported, it is possible and feasible to classify EVOOs of different origins with IRMS. In the results of this research, it was determined that the interaction of *harvest year*variety* was significant and the results were consistent with the study reported by the authors.

In Figure 2, the plots (upper right and lower right) of standardized residuals obtained from the ANOVA model against $\delta^{13}C$ values estimated by the model and the values given in the order of observation, as well as the normal probability and frequency distribution graphs (upper left and lower left of the standardized residues) are given.

The main effect and binary interaction plots of the three factors from ANOVA are shown in Figure 3a, b, respectively. It was determined that while there were significant changes in terms of phase and variety, there was no significant change in terms of harvest year. In this context, the average $\delta^{13}C$ values of the oils obtained from the 2-phase system are around −29.0‰, on the other hand, the average $\delta^{13}C$ value of the oils obtained from the 3-phase system was around −29.5‰. When we consider it in terms of variety, it was understood that the highest value of $\delta^{13}C$ was seen in Ayvalık variety and the lowest value was seen in Gemlik variety. It was also seen that $\delta^{13}C$ value of Memecik variety is around 0‰–29.5‰ on average.

Angerosa et al. (1999) applied $\delta^{13}C$ and $\delta^{18}O$ analyzes to 42 olive oil samples obtained from Italy, Greece, Morocco, Spain, Tunisia and Turkey. The results indicated that olive oil samples tended to cluster according to different climatic regions where olive fruits were grown, but some confusion was also observed for samples from neighboring countries with similar climates. They stated that the $\delta^{13}C$ value of Greek, Italian, Moroccan, Spanish Tunisian and Turkish oils ranged between −29.1‰ and −27.6‰, ranged between −31.3‰ and −27.0%, ranged between −29.2‰ and −28.4‰, ranged between −29.2 and −28.5‰, ranged between −29.6‰ and −28.2‰ and ranged between −29.1‰ and −27.7‰, respectively. Chiocchini et al. (2016) stated in a study conducted in Italy that there is a clear distinction between olive oils produced in the Northern regions and those from other Italian regions, and that high temperatures and low precipitation result in the enrichment of the $\delta^{13}C$ value of extra virgin olive oils. They stated that the $\delta^{13}C$ values of olive trees grown in dry conditions were higher than those grown in wet conditions. The $\delta^{13}C$ results obtained are in agreement with the values determined by Gumus et al. (2017), Camin et al. (2016) and Angerosa et al. (1999).

Nuclear magnetic resonance (NMR)

The typical ^{13}C NMR raw spectra and ^{1}H NMR spectrum of the analyzed EVOO samples is shown in Figure 4, respectively. The results ^{1}H NMR spectrum show the signals of fatty acids in different triglyceride combinations previously mentioned by Sacchi et al. (1997) and Fauhl et al. (2000).

PCA, which is a classification and clustering method that does not require any guidance, was applied to both

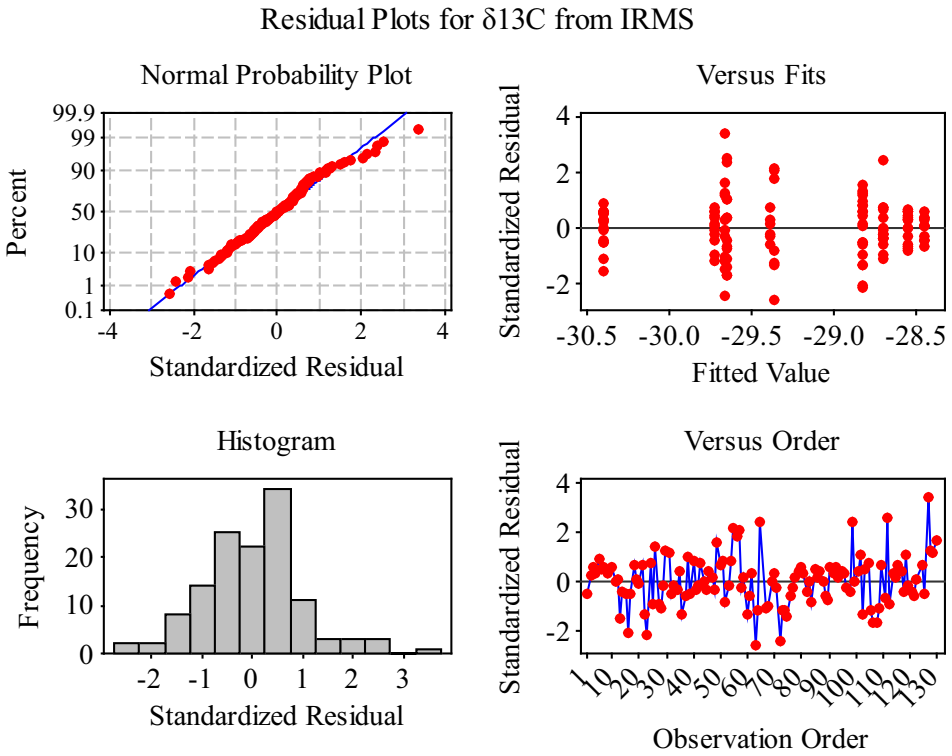

FIGURE 2 Residual plots of three factor ANOVA for $\delta^{13}C$ values EVOO samples for the three varieties (Ayvalık, Gemlik, and Memecik) from IRMS

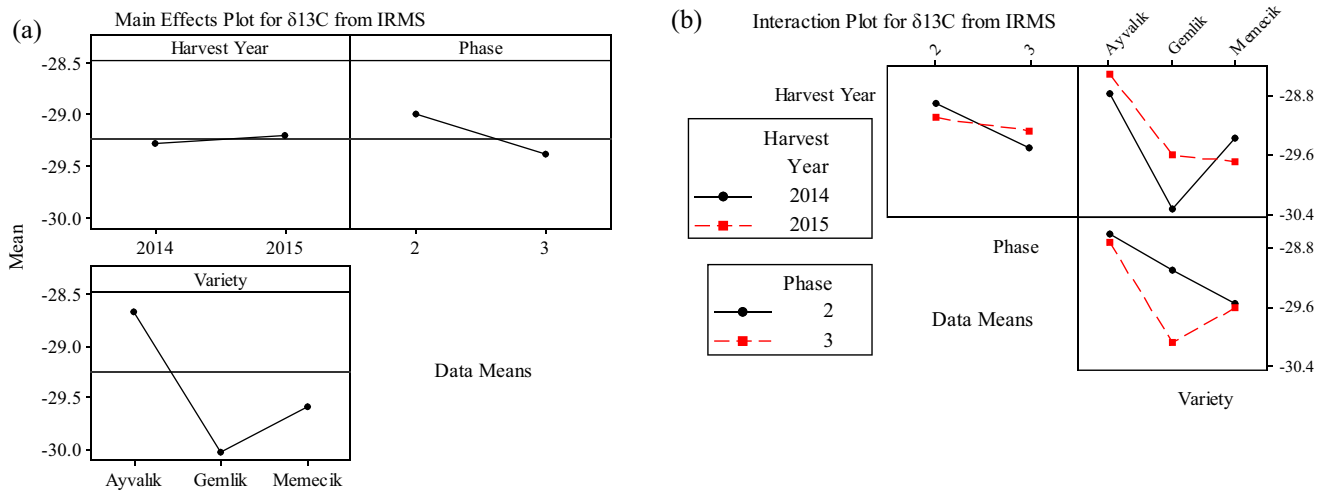

FIGURE 3 ANOVA main effect (a) plots of $\delta^{13}C$ values and (b) binary interaction plots of three factors (harvest year, variety, and phase)

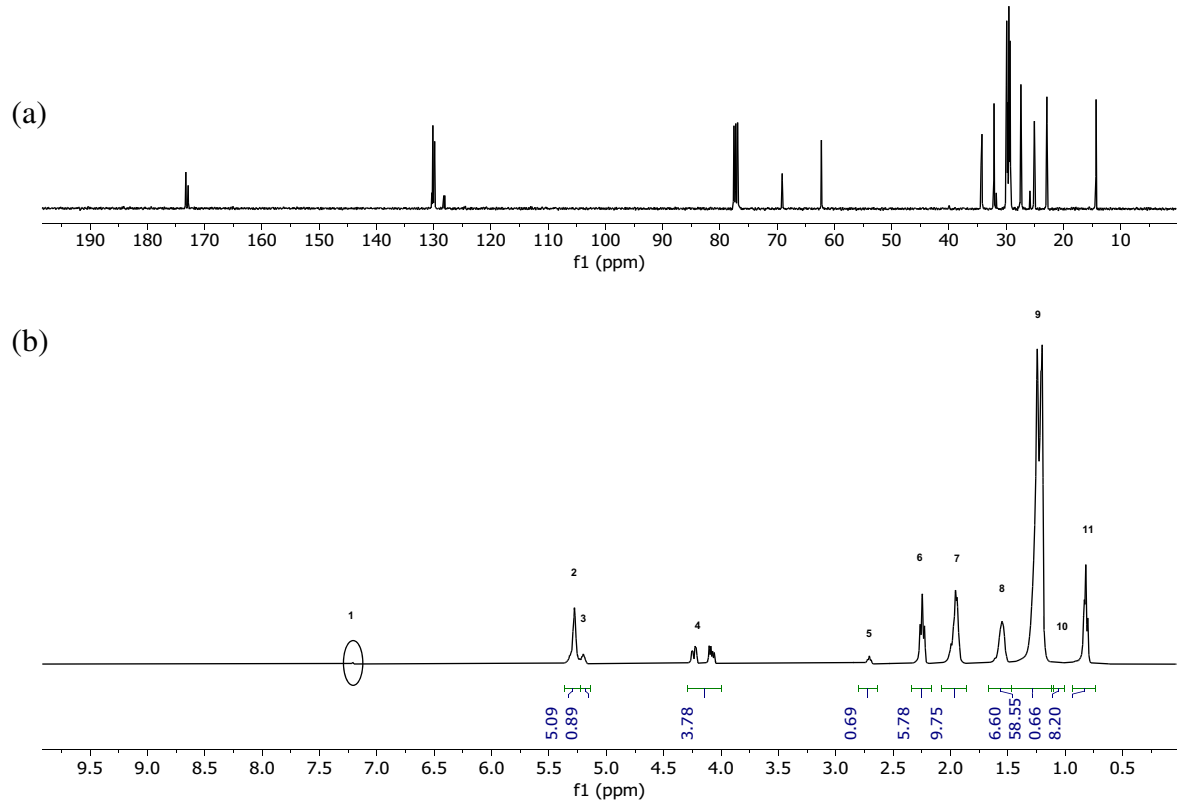

FIGURE 4 (a) One of the ^{13}C NMR raw spectrum and (b) one of the ^{1}H NMR spectrum of the EVOO samples

^{1}H and ^{13}C NMR spectra of the EVOO samples. First of all, ^{1}H NMR spectra were analyzed where spectral pre-processing is applied in a number of different ways (such as mean centering, scaling and standardization) and standardization of were chosen before PCA. Since Ayvalık (North Aegean; 69 samples), Memecik (South Aegean; 40 samples) and Gemlik (South Marmara; 28 samples) EVOO samples constitute the majority of the samples, only the samples from these three geographical region were chosen for PCA and PLS-DA analysis. The West Mediterranean and Southeastern Anatolia samples were not used since several varieties included from these regions and the number of samples from each cultivar were not sufficient to construct a training set from these samples for PLS-DA. Figure 5 shows the two and three dimensional score plots obtained from PCA analysis with ^{1}H NMR spectra where three different grouping (variety, phase and

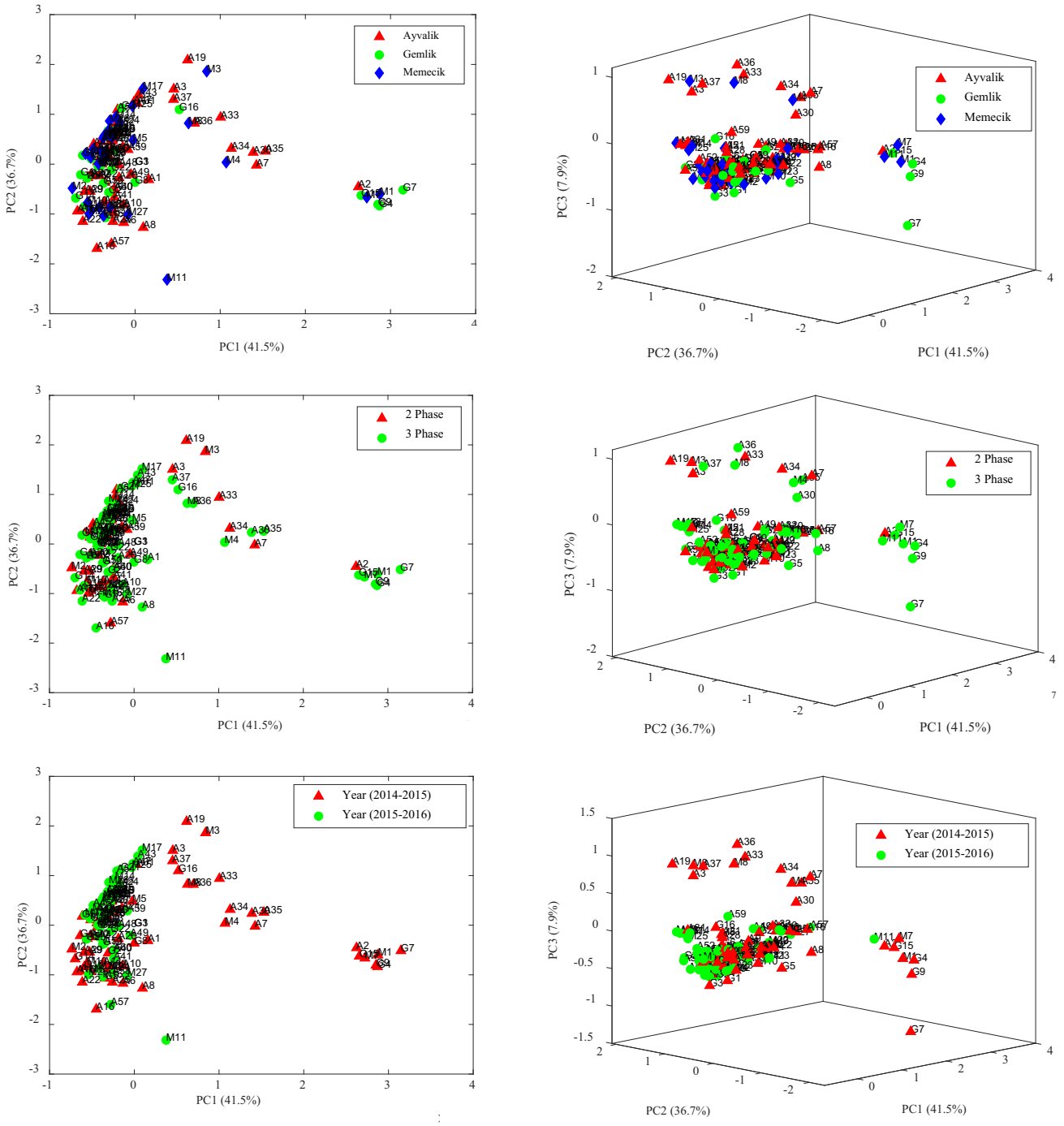

FIGURE 5 Two and three dimensional score plots obtained from PCA analysis with ^{1}H NMR spectra of the EVOO samples grouped according to variety (top), phase (middle) and harvest year (bottom)

harvest year) were applied in order to see any possible classification based on variety, phase and /or harvest year.

Even though the two PCA score vector explains the majority of the variability in the data (41% for PC1 and 37% for PC2), a clear distinction could not be achieved for all the three grouping categories (variety, phase and harvest year). On the other hand, samples from first harvest year (2014/15) were more scattered compared to 2015/16 harvest year. The score plots of the first two and three principal components (PC's) from the PCA analysis of ^{13}C NMR data are shown in Figure 6.

As in the case of ^{1}H NMR data, two and three dimensional score plots from PCA for ^{13}C NMR spectra were not able to show any clear separation among either the three varieties or two phases. This is also true for the two harvest years.

As mentioned above, there were 69 Ayvalık (North Aegean), 40 Memecik (South Aegean), and 28 Gemlik

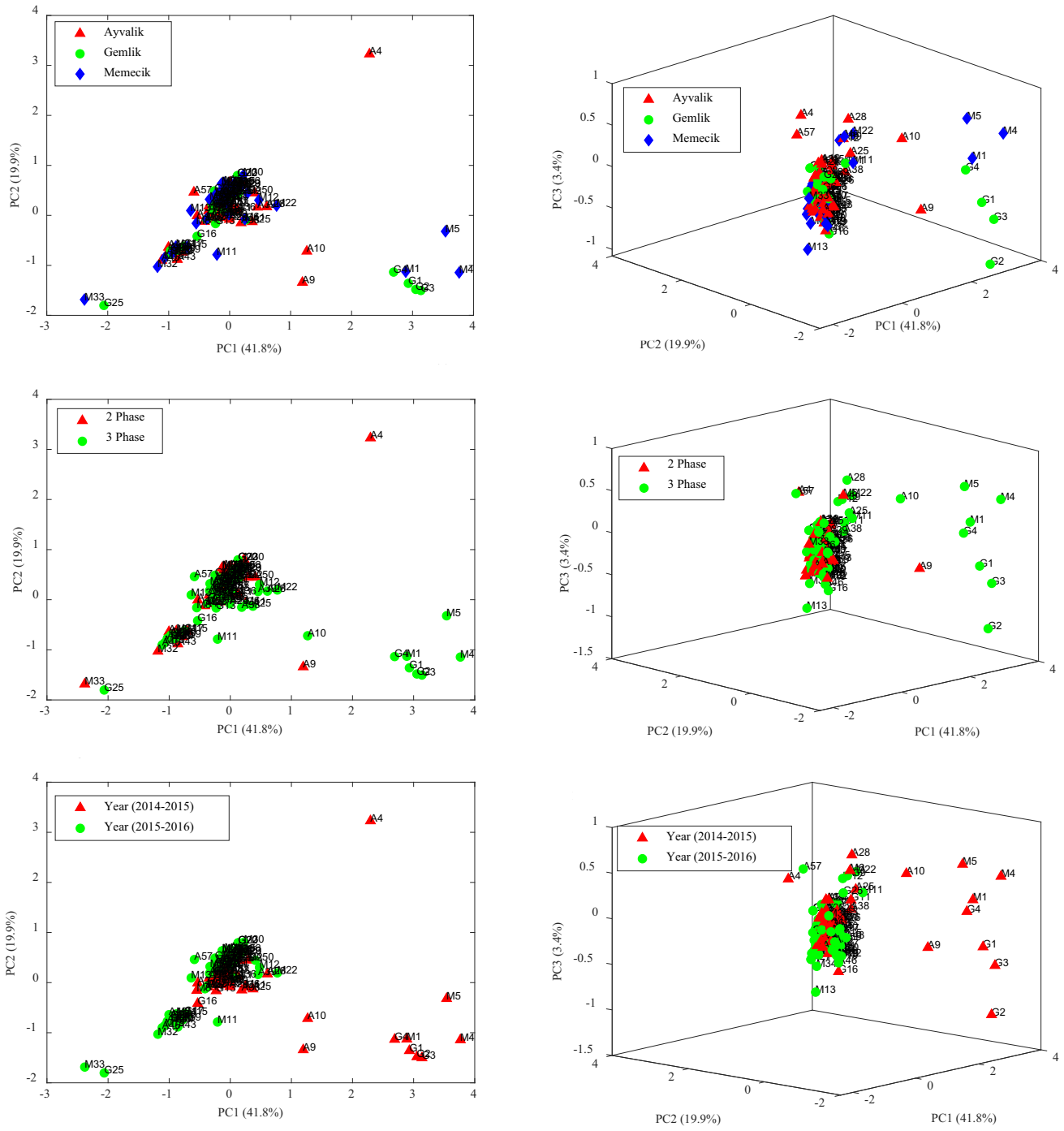

FIGURE 6 Two and three dimensional score plots obtained from PCA analysis with ^{13}C NMR spectra of the EVOO samples grouped according to variety (top), phase (middle) and harvest year (bottom)

(South Marmara) EVOO samples giving a total of 137 samples from these three cultivar or geographical regions. Among the 137 samples, two third of the samples were used for model (47 Ayvalık, 28 Memecik, and 20 Gemlik samples) construction and one third of them (22 Ayvalık, 12 Memecik, and 8 Geemlik) were reserved for independent test set in order to develop supervised classification models with PLS-DA which is one of the most widely used method. Although PLS-DA is set up to build models with leave one out cross validation in order to avoid any overfitting problem, an independent test set could be better for the sake of safe evaluation of the models. Leave one out cross validation was applied with PLS-DA in order to avoid any overfitting problem during the model building step. An independent test set was also used to evaluate the models. However, as the number of the samples from first and second harvest years were not uniform, the

number of sample in independent test were kept relatively small in order to generate a highest possible representative training set for the classification models. As in the case of the PCA analysis, different data proprocessing strategies were applied to raw NMR data and standardization were chosen for model construction in PLS-DA. Figure 7 shows three dimensional PLS-DA similarity plots obtained from (a) ^{1}H NMR and (b) ^{13}C NMR spectra of the EVOO samples from both harvest year (2014/15 and 2015/16).

As can be seen from the PLS-DA three-dimensional similarity plot of ^{1}H NMR spectra, EVOO samples from three different geographical regions are relatively well separated from each other by PLS-DA especially for the model data set but the same is not true for the independent test samples as some of the located in the center of the borders of the three model set groups. The optimum number of latent variables (number of PC) were found to be 14 which may explain the decrease in the success of the models for independent test samples. In fact, partial overfitting of the model is evident from the distance values of the test samples. On the other hand, PLS-DA similarity plots obtained from ^{13}C NMR spectra of the EVOO samples from both harvest year indicates somewhat poor classification of the model set compared to the ^{1}H NMR spectra. However, independent test set samples do not seem to be misclassified too much in different groups but a clear

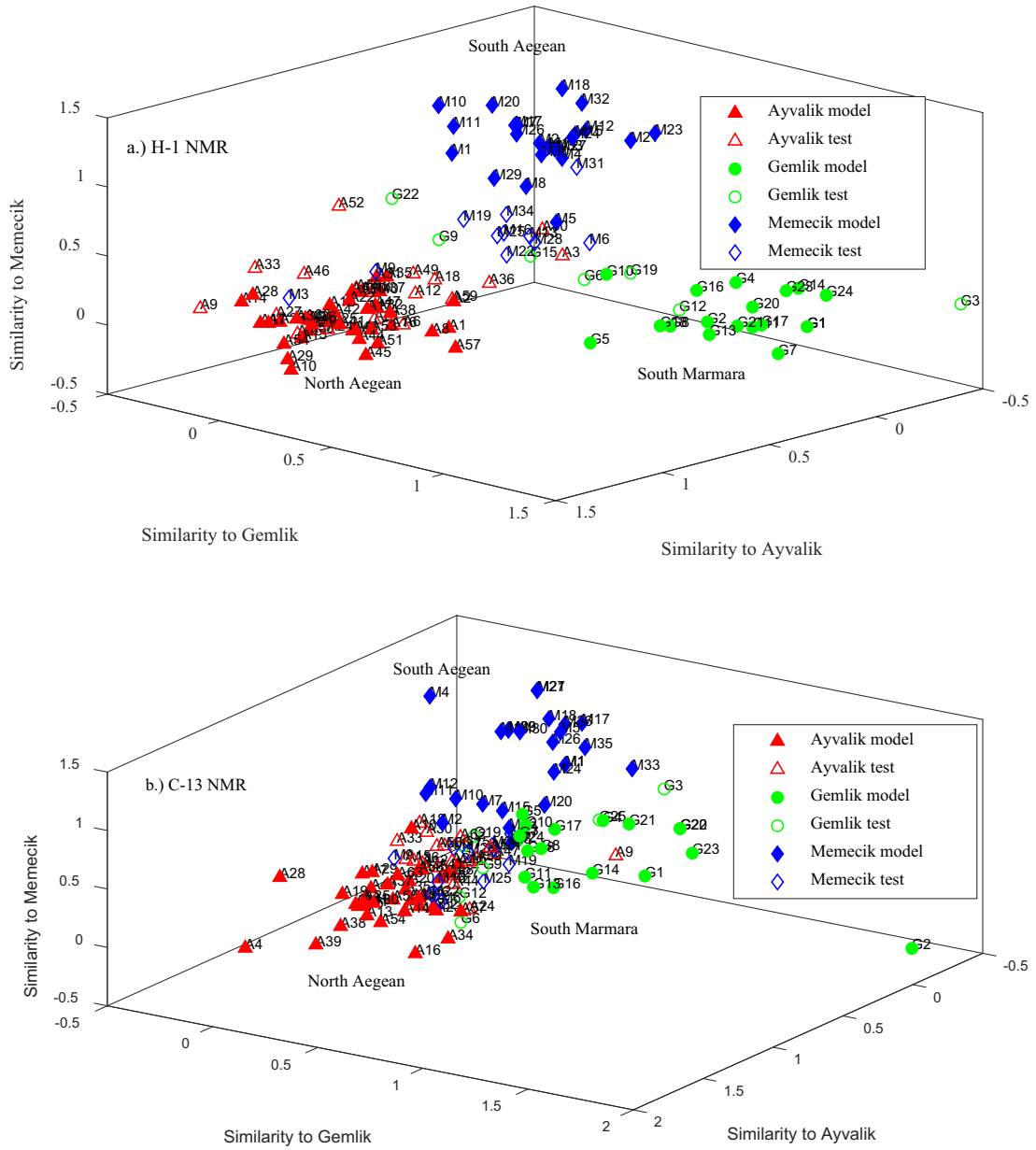

FIGURE 7 Three-dimensional PLS-DA similarity plots obtained from (a) ^{1}H NMR and (b) ^{13}C NMR of the EVOO samples from both harvest year (2014/15 and 2015/16)

separation from the three geographical regions is not observed. The number latent variable was nine in this case which probably explains somewhat overlap of the model set samples.

Since variety that dominates each geographical region is also unique, it can be said that the success of the model could be due to both geographical region and the variety. Table 4 shows percentage of success in the PLS-DA model set and the independent test set of 137 samples from ^{1}H NMR spectra.

In terms of statistical evaluation of the model obtained with the PLS-DA method, 19 of the 20 South Marmara samples in the model set were classified correctly, while one samples were classified incorrectly at 95% confidence level. Again, five of the eight independent test set samples belonging to this region were classified correctly, while three of them were incorrectly classified at the same confidence level. In this context, the percentage success of the model obtained was 95% in the model set, while this rate was 62.5% in the independent test set. Similarly, while 45 of the 47 North Aegean samples in the model set were classified correctly, only two samples were misclassified and the model success was around 95.7%. In addition, 18 of the 22 test set samples in the same region were classified correctly, while four of them was misclassified, so an accuracy of 81.8% was reached here as well. Finally, 27 of the 28 South Aegean samples in the model set were classified correctly and 1 were assigned to the wrong group. On the other hand, while eight test set samples belonging to this region were classified correctly, four of them were incorrectly classified. In this context, model set samples belonging to this region were classified with a success rate of 96.4% and test samples with a success rate of 66.7%. In summary, the success of the PLS-DA model for ^{1}H NMR spectra was around 95.8% for the training set, whereas it was seen that the same model achieved a success rate of 73.8% in the estimation of independent test set samples. Table 5 shows the percentage of success in the PLS-DA model set and the independent test set of the 137 samples from ^{13}C NMR spectra.

When compared to the PLS-DA model generated with ^{1}H NMR spectra, correct classification of the model set samples are relatively lower in ^{13}C NMR data. However, the results of the independent test set samples are comparable. As seen from Table 5, 15 of the 20 South Marmara samples in the model set were classified correctly, while five samples were classified incorrectly. Again, four of the eight test set samples belonging to this region were classified correctly, while four of them were incorrectly classified. In this context, the percentage success of the model obtained was 75% in the model set, while this rate was 50% in the independent test set. Similarly, while 38 of the 47 North Aegean samples in the model set were classified correctly, only 15 of the 22 samples were misclassified and the model success was around 81%. In addition, 20 of the 28 South Aegean (Memecik) samples in the model set were classified correctly, while 5 of the 12 independent test set samples were misclassified, so an accuracy of 71.4% and 58.3% was reached for model and test sets, respectively. In general, the success of the PLS-DA model based on ^{13}C NMR spectra, was around 76.8%, it was observed that the same model achieved a 61.9% success rate in estimating independent test set samples.

Alonso-Salces et al. (2010), reported that these spectral data contain useful information in the geographical characterization of olive oil. They have been also stated that the data obtained for the studied five countries (Spain, Italy, Turkey, Tunisia, and Syria) contain complementary information in the classification of olive oils according to geographical origin. Longobardi et al. (2012), stated in their research that NMR fingerprint was investigated with multivariate statistical analysis techniques, accurate predictions specific to the region were achieved between 53% and 100%, and correct predictions were made with a probability of 78%. In the study conducted by Dağ (2016), Edremit Yağlık, Nizip Yağlık and Memecik oils were successfully classified using the OPLS-DA method using NMR data. In addition, it was stated that OPLS-DA models obtained using NMR spectroscopy data were more successful in grouping according to olive varieties than

TABLE 4 Percentage of success in the PLS-DA model set and the independent test set of 137 samples from ^{1}H-NMR spectra

		Total sample	Correct classification	Wrong classification	Success in the model set (%)	Success in the independent test set (%)
South Marmara (Gemlik)	Model set	20	19	1	95	62.5
	Test set	8	5	3		
North Aegean (Ayvalık)	Model set	47	45	2	95.7	81.8
	Test set	22	18	4		
South Aegean (Memecik)	Model set	28	27	1	96.4	66.7
	Test set	12	8	4		
	Total	137	122	15	95.8	73.8

TABLE 5 Percentage of success in the PLS-DA model set and the independent test set of the 137 samples from ^{13}C NMR spectra

		Total sample	Correct classification	Wrong classification	Success in the model set (%)	Success in the independent test set (%)
South Marmara (Gemlik)	Model set	20	15	5	75	50
	Test set	8	4	4		
North Aegean (Ayvalık)	Model set	47	38	9	80.9	68.2
	Test set	22	15	7		
South Aegean (Memecik)	Model set	28	20	8	71.4	58.3
	Test set	12	7	5		
	Total	137	99	38	76.8	61.9

OPLS-DA models obtained using fatty acid or sterol composition profiles. Özdemir et al. (2018) stated that Turkish olive oils (Ayvalık, Memecik, Gemlik) and Slovenian olive oil (Bianchera) were effectively distinguished according to their NMR profiles. Camin et al. (2016) analyzed 177 Italian PDO olive oils and 86 imported Tunisian olive oils with IRMS ($\delta^{13}C$-carbon, $\delta^{18}O$-oxygen isotope and $\delta^{2}H$-hydrogen isotope) and NMR. The statistical model (a multivariate statistical approach) indicated that olive oils imported from Italy and Tunisia could be distinguished with an optimal differentiation ability of around 98%. In the research by Ok (2014), in which 38 olive oils collected from Turkey, Jordan, Palestine and Libya were screened, it was stated that the results of the quantitative analysis of NMR helped to distinguish the geographical origin of the olive oil samples. In addition, it was stated that NMR data distinguishes olive oils from the provinces where they are grown, based on regional origin. The results showed that by using the obtained classification and clustering models, geographical marking and labeling of these oils can be done with NMR analysis results regardless of year and production systems (2 and 3-phase continuous systems).

CONCLUSIONS

In this research, it is aimed to classify EVOOs according to their geographical origin by using $\delta^{13}C$ IRMS, ^{1}H and ^{13}C NMR analysis and chemometric data analysis methods. A total of 175 EVOO samples were studied in two harvest years. IRMS results showed that only phase and variety were statistically significant factors ($p < 0.05$). On the other hand, it was observed that there was no significant difference in terms of $\delta^{13}C$ values between the years in terms of harvest year ($p > 0.05$). The interactions of the factors showed that the *harvest year x variety* interaction is important. When the NMR analysis results were evaluated using the obtained classification and clustering models, the results showed that geographical marking and labeling of EVOOs could be done only with the NMR analysis results, regardless of the year and production systems (2 and 3-phase continuous systems).In conclusion, the outcomes of this research clearly indicated that considering the PLS-DA result (a total of 137 samples from South Marmara, North Aegean and South Aegean regions) with NMR spectra, the percent success of the model in the South Marmara region samples were 95% in the model set, while this rate 62.5% in the independent test set. Similarly, the model success was 95.7% in the North Aegean region samples. In addition, 18 of the 22 test set samples in the same region were classified correctly, while four of them were incorrectly classified, so an accuracy of 81% was achieved here as well. Finally, in the South Aegean region samples, the model set samples were classified with 96.4% success and the test samples were classified with 66.7% success. While the success of the PLS-DA model, which was created based on these results, was around 96%, it was seen that the same model achieved a 74% success in estimating independent test set samples. In addition, a comprehensive database was created with the data obtained from the study.

AUTHOR CONTRIBUTIONS

Didar Sevim: Methodology; investigation; oil extraction; writing, editing; supervision. **Oya Köseoğlu**: Investigation; oil extraction. **Hasan Ertaş**: Supervision; interpretation. **Durmuş Özdemir**: Statistic; software. **Mehmet Ulaş**: Field Observation; sample collection. **Salih Günnaz**: NMR analysis. **Veysel Umut Çelenk**: IRMS analysis.

ACKNOWLEDGMENTS

This project is supported by General Directorate of Agricultural Research and Policies (TAGEM) of Turkey under the grant No. TAGEM/ HSGYAD/15A05/P01/75.

FUNDING INFORMATION

This research is supported by General Directorate of Agricultural Research and Policies (TAGEM) of Turkey under the grant No. TAGEM/HSGYAD/15A05/P01/75. The authors wish to thank to TAGEM.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

ETHICS STATEMENT

This research did not involve any human or animal study and institutional ethical approval was not required.

ORCID

Didar Sevim https://orcid.org/0000-0003-0236-2294
Oya Köseoğlu https://orcid.org/0000-0002-6585-4221
Hasan Ertaş https://orcid.org/0000-0002-5539-3732
Durmuş Özdemir https://orcid.org/0000-0003-3297-8217
Mehmet Ulaş https://orcid.org/0000-0001-6777-4922
Salih Günnaz https://orcid.org/0000-0002-7422-6593
Veysel Umut Çelenk https://orcid.org/0000-0001-8536-6498

REFERENCES

Alonso-Salces RM, Héberger K, Holland MV, Moreno-Rojas JM, Mariani C, Belan G, et al. Multivariate analysis of NMR fingerprint of the unsaponifiable fraction of virgin olive oils for authentication purposes. Food Chem. 2010;118:956–65.

Alonso-Salces RM, Holland MV, Guillou C, Héberger K. Quality assessment of olive oil by ^{1}H-NMR fingerprinting, olive oil - constituents, quality, health properties and bioconversions. InTech. 2012;185–210. http://doi.org.10.5772/28701

Angerosa F, Bréas O, Contento S, Guillou C, Reniero F, Sada E. Application of stable isotope ratio analysis to the characterization of the geographical origin of olive oils. J Agri Food Chem. 1999;47:1013–7.

AOAC, 998.12. Official methods of analysis of the assocition of official analytical chemists. 18th Ed. 2005.

Baum A, Lu Y, Muccio Z, Jackson G, Harrington P. Differentiation between origins of extra virgin olive oils by GC-C-IRMS using principal component analysis, linear discriminant analysis, and hierarchical cluster analysis. Spectroscopy. 2010;25:40.

Bontempo L, Paolini M, Franceschi P, Ziller L, García-González DL, Camin F. Characterisation and attempted differentiation of European and extra-European olive oils using stable isotope ratio analysis. Food Chem. 2019;276:782–9.

Calò F, Girelli CR, Wang SC, Fanizzi FP. Geographical origin assessment of extra virgin olive oil via NMR and MS combined with chemometrics as analytical approaches. Foods. 2022;11(1):113.

Camin F, Pavone A, Bontempo L, Wehrens R, Paolini M, Faberi A, et al. The use of IRMS, ^{1}H NMR and chemical analysis to characterise Italian and imported Tunisian olive oils. Food Chem. 2016;196:98–105.

Chiocchini F, Portarena S, Ciolfi M, Brugnoli E, Lauteri M. Isoscapes of carbon and oxygen stable isotope compositions in tracing authenticity and geographical origin of Italian extra-virgin olive oils. Food Chem. 2016;202:291–301.

Dağ Ç. Türkiye Menşeli Sızma zeytinyağları için referenas NMR Yöntemleriyle Otantisite ve Orjin Belirleme Metotlarının Geliştirilmesi, Development of Authenticity and Origin Determination Methods for Turkish Origin Extra Virgin Olive Oils by Reference NMR Methods (in English) Doktora Tezi, Mugla Sıtkı Koçman University, Turkey. 2016.

Emwas AH, Roy R, McKay RT, Tenori L, Saccenti E, Gowda GAN, et al. NMR spectroscopy for metabolomics research. Metabolites. 2019;9:123.

Fauhl C, Reniero F, Guillou C. ^{1}H NMR as a tool for the analysis of mixtures of virgin olive oil with oils of different botanical origin. Magn Reson Chem. 2000;38:436–43.

González-Casado A, Cuadros-Rodriguez L, Ruiz-Samblás C, Pérez-Castaño E, Rodríguez-Garcia FP, Bagur-González MG. Chromatographic fingerprinting methodology for olive oil authentication.Work shop on olive oil authentication, IOC, Spain. 2013.

Gumus PZ, Celenk VU, Tekin S, Yurdakul O, Ertas H. Determination of trace elements and stable carbon isotope rations İn virgin olive oils from Westernturkey to authenticate geographical origin with a chemometric approach. Eur Food Res Technol. 2017;243:1719–27.

Jiménez-Morillo NT, Palma V, Garcia R, Dias CB, Cabrita MJ. Combination of stable isotope analysis and chemometrics to discriminate Geoclimatically and temporally the virgin olive oils from three Mediterranean countries. Foods. 2020;9:1855.

Longobardi F, Ventrella A, Napoli C, Humpfer E, Schütz B, Schäfer H, et al. Classification of olive oils according to geographical origin by using ^{1}H NMR fingerprinting combined with multivariate analysis. Food Chem. 2012;130:177–83.

Martín-Peláez S, Covas MI, Fitó M, Kušar A, Pravst I. Health effects of olive oil polyphenols: recent advances and possibilities for the use of health claims. Mol Nutr Food Res. 2013;57:760–71.

Merchak N, Rizk T, Silvestre V, Remaud GS, Bejjani J, Akoka S. Olive oil characterization and classification by ^{13}C NMR with a polarization transfer technique: a comparison with gas chromatography and ^{1}H NMR. Food Chem. 2018;245:717–23.

Nasr EG, Epova EN, Sebilo M, Larivière D, Hammami M, Souissi R, et al. Olive oil traceability studies using inorganic and isotopic signatures: a review. Molecules. 2022;27:2014. https://doi.org/10.3390/molecules27062014

Ogrinc N, Košir IJ, Spangenberg JE, Kıdrıč J. The application of NMR and MS methods for detection of adulteration of wine, fruit juices. And Olive Oil Anal Bioanal Chem. 2003;376:424–30.

Ok S. Fast screening of Turkish olive oil by NMR spectroscopy for geographical determination and discrimination purposes. Grasas Aceites. 2014;65(2):1–11.

Özdemir IS, Dag C, Makuc D, Ertas E, Plavec J, Bekiroglu S. Characterisation of the Turkish and Slovenian extra virgin olive oils by chemometric analysis of the presaturation ^{1}H NMR spectra. LWT-Food Sci Technol. 2018;92:10–5.

Portarena S, Farinelli D, Lauteri M, Famiani F, Esti M, Brugnoli E. Stable isotope and fatty acid compositions of monovarietal olive oils: implications of ripening stage and climate effects as determinants in traceability studies. Food Control. 2015;57:129–35.

Rabiei Z, Enferadi ST. Challenges for genetic identification of olive oil. Intech. 2013;201–18. https://doi.org.10.5772/53546

Sacchi R, Addeo F, Paolillo L. ^{1}H and ^{13}C NMR of virgin olive oil. Magn Reson Chem. 1997;37:133–45.

Wada E. Stable δ^{15}N and δ^{13}C İsotope ratios in aquatic ecosystems. Proc Jpn Acad Ser B Phys Biol Sci. 2009;85:98–107.

How to cite this article: Sevim D, Köseoğlu O, Ertaş H, Özdemir D, Ulaş M, Günnaz S, et al. Identification of Turkish extra virgin olive oils produced in different regions by using NMR (^{1}H and ^{13}C) and IRMS (^{13}C/^{12}C). J Am Oil Chem Soc. 2023;100(5):355–67. https://doi.org/10.1002/aocs.12697