Talanta 156-157 (2016) 180–190

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Chemometric study of Andalusian extra virgin olive oils Raman spectra: Qualitative and quantitative information

E. Sánchez-López [a], M.I. Sánchez-Rodríguez [b], A. Marinas [a,*], J.M. Marinas [a], F.J. Urbano [a], J.M. Caridad [b], M. Moalem [a]

[a] Organic Chemistry Department, Campus de Excelencia Internacional CeiA3, University of Córdoba, Campus de Rabanales, Marie Curie Building, E-14014 Córdoba, Spain
[b] Statistics and Business Department, University of Córdoba, Avda. Puerta Nueva, s/n, E-14071 Córdoba, Spain

ARTICLE INFO

Article history:
Received 24 January 2016
Received in revised form
19 April 2016
Accepted 2 May 2016
Available online 3 May 2016

Keywords:
Raman spectroscopy
Extra virgin olive oil
Andalusian PDO
Harvest Year
Geographical Origin
Olive Variety

ABSTRACT

Authentication of extra virgin olive oil (EVOO) is an important topic for olive oil industry. The fraudulent practices in this sector are a major problem affecting both producers and consumers. This study analyzes the capability of FT-Raman combined with chemometric treatments of prediction of the fatty acid contents (quantitative information), using gas chromatography as the reference technique, and classification of diverse EVOOs as a function of the harvest year, olive variety, geographical origin and Andalusian PDO (qualitative information). The optimal number of PLS components that summarizes the spectral information was introduced progressively. For the estimation of the fatty acid composition, the lowest error (both in fitting and prediction) corresponded to MUFA, followed by SAFA and PUFA though such errors were close to zero in all cases. As regards the qualitative variables, discriminant analysis allowed a correct classification of 94.3%, 84.0%, 89.0% and 86.6% of samples for harvest year, olive variety, geographical origin and PDO, respectively.

1. Introduction

Extra virgin olive oil (EVOO) is an edible oil very important in the Mediterranean diet. Spain is the first producer worldwide and Andalusia encompasses 80% of the national production. This oil is obtained by mechanical procedures only, so its characteristics are not affected and hence it is a top quality edible oil. It is constituted by two main families of chemical compounds: saponifiable and unsaponifiable fractions. Saponifiable fraction or major components is formed by fatty acids, which are classified in saturated fatty acids (SAFA) such as palmitic, stearic, margaric and arachidic acids, monounsaturated fatty acids (MUFA) such as palmitoleic, margaroleic, gadoleic and oleic acid, polyunsaturated fatty acids (PUFA) such as linoleic and linolenic acid, and free fatty acids. Unsaponifiable fraction or minor compounds is formed by a heterogeneous group of elements: sterols, fatty alcohols, pigments, hydrocarbons, volatile and phenolic compounds and others. EVOO is highly appreciated due to its flavor and beneficial health properties [1–3], so it is considered as a superior quality oil. Sometimes, this high quality could be affected with frauds in its marketing, such as adulterations with cheaper oils (such as palm, corn, hazelnut, soy or refined olive oil) or indication of a false geographical origin. Some other features whose traceability is of interest include precise varietal composition and harvest year. In order to ensure the preservation of olive oil quality, the European Union has introduced some legislation and Protected Designations of Origin (PDO) and Protected Geographical Indications (PGI) have been established.

Unlike other analytical techniques for authentication of olive oils such as chromatography (gas-GC- and liquid- HPLC-) coupled to mass spectrometry [4,5] or Nuclear Magnetic Resonance (^{1}H and ^{13}C NMR) [6,7], vibrational spectroscopy (Near Infrared-NIR-, Mid Infrared-MIR- and Fourier Transform Raman Spectroscopy -FT-Raman) techniques [8,9] do not require any sample manipulation though they are commonly used in combination with chemometric analysis. These techniques are fast, reliable and cost-effective and allow on-site testing using portable devices. Compared to Infrared (IR) which has been widely used in food quality [10,11], Raman spectroscopy has been less explored since traditionally it has been considered expensive and featuring significant fluorescence problems. Nowadays, both drawbacks have been solved and Raman is applied to several fields such as medicine, food, agriculture, chemistry and the environment among others [12,13]. Moreover, infrared and Raman spectroscopies can be considered as complementary techniques since they differ in the selection rules. Therefore, vibrational transitions in order to be

* Corresponding author.
E-mail address: alberto.marinas@uco.es (A. Marinas).

http://dx.doi.org/10.1016/j.talanta.2016.05.014
0039-9140/

active in IR spectroscopy require the change of dipole moment whereas for Raman polarizability should change. This results in Raman being preferable to IR when strong absorption by water and carbon dioxide can be a problem. The polar groups (e. g. C=O and O–H) exhibit strong bands in infrared, while the non-polar groups (C=C) show intense Raman bands, what makes Raman a suitable technique to distinguish between cis or trans double bonds in unsaturated fatty acids (bands at 1660 or 1670 cm^{-1}, respectively). Furthermore, the Raman spectrum is composed of isolated bands while IR spectrum can present overlapping peaks. Other advantage that could be considered in Raman is the enhancement of the signal when the excitation wavelength matches the electronic absorbance of a chemical system, due to the effect of resonance. So this technique could be a good method to provide some useful quantitative and qualitative information on olive oils.

There are different types of Raman devices namely dispersive Raman, FT-Raman, SERS (Surface-Enhanced Raman Spectroscopy) and SORS (Spatially Offset Raman Spectroscopy). FT-Raman is probably the most frequently utilized equipment, resorting to an interferometer and Fourier Transform.

This technique has been used in analysis, classification and authentication of oils and fats [14–18]. In the literature, there are several articles on the application of Raman spectroscopy to olive oil [19–31] (see some examples in Table 1).

The aim of this study is to determine some qualitative (such as the harvest year, olive variety, geographical origin or PDO) and quantitative information (content in saturated SAFA, MUFA and PUFA, using the values obtained by GC as a reference) of Andalusian extra virgin olive oil using Raman spectroscopy. The quantitative variable to study has been selected because fatty acids are the major compounds in olive oils and it is considered as a quality parameter. Regarding the qualitative information, it is important to emphasize that there are few or none published articles on some of the selected parameters such as harvest year, which could be of interest.

In this paper the extensive spectral information is summarized in a few factors or components, obtained by partial least squares (PLS) regression. Contrary to the cited works, the number of the PLS factors to retain is not fixed, but it is progressively increased to analyze the evolution in the goodness of the corresponding statistical models. In the quantitative study, some new dimensionless measures of calibration and prediction errors are proposed in order to compare different regression models and the different fit and validation data subsets are considered in cross-validation.

2. Materials and methods

2.1. Oil Samples

The samples studied in this work are 412 EVOO samples from six consecutive harvest years (from 2004–2005 to 2009–2010) and were supplied by Hojiblanca Cooperative (mainly from Córdoba and Málaga), Germplasm Bank of Córdoba (Spain) and different Andalusian PDOs (from Jaén, Cádiz, Córdoba, Málaga, Granada and Sevilla) see Fig. 1. 307 monovarietal samples corresponded to Frantoio (19 samples), Hojiblanca (142), Picual (128) and Gordal (18). The other 105 samples were mixtures of different varieties. These EVOOs were obtained through a two-phase centrifugation system. Once received, the samples were sorted and stored at 4 °C in the fridge to prevent deterioration due to heat and light.

For determination of fatty acids (quantitative study) and harvest year classification (qualitative study) all samples were studied whereas for the remaining qualitative studies, olive variety, geographical origin and PDO, three sets with 307, 145 and 67 samples, respectively, were selected.

2.2. Raman spectroscopy

Spectra were collected with a Perkin Elmer System 2000 Raman Spectrometer available at the Organic Chemistry Department (University of Córdoba). It is equipped with a Nd: YAG (9394 cm^{-1}) laser, operating at 300 mW. The analyses were carried out on a 2 mL sample vial, accumulating 16 scans in the 100–3100 cm^{-1} range at a 4 cm^{-1} resolution, using the Spectrum software (Perkin Elmer). The obtained spectra were given a Savitsky-Golay smoothing function over the 100–3100 cm^{-1} range to reduce the noise affecting the spectra quality and were normalized to reduce the external variables influences.

A typical Raman spectrum of an EVOO sample is depicted in Fig. 2, whereas the bands assignment can be found in Table 2.

2.3. GC fatty acid determination

Determination of fatty acid composition by Gas Chromatography-Flame Ionization Detector (GC-FID) was carried out in the Organic Chemistry Department (University of Córdoba), according to the official methods for olive and pomace oil control in the European Union [33] and the International Olive Council (IOC) [34,35], using an Agilent 7890 A gas chromatograph with a capillary column (SGE FORTE BPX-70 de 50 m × 220 μm × 0.25 μm).

The analytical conditions used were: 250 °C of injector temperature, 2 μL of injection volume, 260 °C of detector temperature. The oven temperature was programmed to remain at 180 °C for 15 min and then raised up to 240 °C at a rate of 4 °C/min and maintained at this temperature for 5 min

The triacylglycerol samples (olive oil samples), were initially submitted to a cold transesterefication procedure to convert the triacylglycerol into fatty acid methyl esters (FAMEs). This method is indicated for edible oils with acidity index lower than 3.3°. In this process, 0.1 g of olive oil is transferred into a 5 mL volumetric vial. Next, 2 mL n-heptane and 0.2 mL of a 2N KOH solution in methanol are added and reaction mixture is vigorously stirred. Finally, the methyl esters are extracted and subject to GC analyses.

Recommendations of IOC for GC analysis were taken into account for calculation of fatty acids [34]. Thus, SAFA comprises palmitic (C16:0), margaric (C17:0), stearic (C18:0) and arachidic (C20:0) acids, MUFA is composed by palmitoleic (C16:1), margaroleic (C17:1), oleic (C18:1) and gadoleic (C20:1) acids and PUFA by linoleic (C18:2) and linolenic (C18:3) acids.

2.4. Chemometric analysis

In the statistical modelling from Raman spectra, the number of explanatory variables (associated to the different Raman shifts) greatly exceeds the number of observations. This results in the appearance of multicollinearity, incompatible with the hypothesis of uncorrelation of general linear models, in general, and regression models, in particular. Therefore, to avoid the presence of multicollinearity, the information contained in Raman spectra is usually summarized in a few components or factors. In general lines, these factors can be determined by using two different optimization criteria: maximizing the correlation between the explanatory variables, in the principal components analysis (PCA), or maximizing the correlation between the explanatory variables and a quantitative dependent one, in the regression by partial least squares (PLS). In the present study, the spectral information is synthesized by using PLS instead of PCA factors, since some studies have shown that this approach provides better results [36–38]. The objective is not only to summarize the information in a few components but, fundamentally, to use these components to predict some quantitative (fatty acid content) or qualitative information (such as harvest year, olive variety, geographical origin or PDO).