Talanta 132 (2015) 829–835

Contents lists available at ScienceDirect

Talanta

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

Short wavelength Raman spectroscopy applied to the discrimination and characterization of three cultivars of extra virgin olive oils in different maturation stages

Irene Gouvinhas [a,*,1], Nelson Machado [a,1], Teresa Carvalho [b], José M.M.M. de Almeida [c,d,1], Ana I.R.N.A. Barros [a,1]

[a] CITAB – CITAB, University of Trás-os-Montes and Alto Douro, 5001-801 Vila Real, Portugal
[b] Instituto Nacional de Investigação Agrária, (INIAL.P.), Elvas, Portugal
[c] Department of Physics, University of Trás-os-Montes and Alto Douro, 5000-801 Vila Real, Portugal
[d] INESC-TEC, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

ARTICLE INFO

Article history:
Received 11 June 2014
Received in revised form
14 October 2014
Accepted 18 October 2014
Available online 28 October 2014

Keywords:
Raman spectroscopy
Extra virgin olive oil
Maturation stages
Cultivars
Peroxide value
Acidity

ABSTRACT

Extra virgin olive oils produced from three cultivars on different maturation stages were characterized using Raman spectroscopy. Chemometric methods (principal component analysis, discriminant analysis, principal component regression and partial least squares regression) applied to Raman spectral data were utilized to evaluate and quantify the statistical differences between cultivars and their ripening process.

The models for predicting the peroxide value and free acidity of olive oils showed good calibration and prediction values and presented high coefficients of determination (>0.933). Both the R^2, and the correlation equations between the measured chemical parameters, and the values predicted by each approach are presented; these comprehend both PCR and PLS, used to assess SNV normalized Raman data, as well as first and second derivative of the spectra.

This study demonstrates that a combination of Raman spectroscopy with multivariate analysis methods can be useful to predict rapidly olive oil chemical characteristics during the maturation process.

1. Introduction

The quality of olive oil is influenced by a great number of factors, namely the nature of the cultivar and geographic origin [1], the fruit ripening degree [2,3], where important chemical changes occur, the irrigation regimes [2], the oil extraction technology and the storage of the oil [4] and agricultural practices [5]. Nutritional and sensory properties of extra virgin olive oils (EVOO) were also affected by these factors. Thus, the European Community (EC) Council of Regulation established standards on olive oil production regarding labeling giving the origin for virgin and extra virgin olive oils to avoid consumers being misled about their true characteristics and origin [6]. Some of the important quality parameters consist on peroxide value (PV) and free acidity (FA) determination [7]. However, these conventional measurements are time consuming and require large amounts of reagents and solvents, which are often toxic, besides being very expensive.

Infrared (IR), mainly Fourier Transform Infrared (FTIR), and Raman, are vibrational spectroscopic techniques widely used for food and feed analysis [8,9]. These optical spectroscopic methods provide information about the chemical composition of various food and biological materials, and molecular structure, without requiring large amount of samples or sample preparation and pretreatments [10]. Furthermore, these techniques have a great potential due to their simplicity, rapidity, low-cost and reproducibility, besides being non-destructive [11].

Both techniques assess the same physical property – molecular vibrations – but present different selection rules. In one hand, the vibrational modes corresponding to a change in polarizability, such as the symmetrical vibrations of covalent bonds in the non-polar groups, (*e.g.* C=C bond stretching), give rise to intense Raman bands. While in the other hand, IR absorption is related to a change in the electrical dipole moment as the molecules vibrate, with the vibrational modes involving polar groups, such as C=O and O-H, showing intense infrared absorption bands [12]. Therefore, one of the major advantages of Raman spectroscopy is

* Corresponding author. Tel.: +351259350283; fax: +351259350480.
E-mail address: igouvinhas@utad.pt (I. Gouvinhas).
[1] www.utad.pt.

http://dx.doi.org/10.1016/j.talanta.2014.10.042
0039-9140/

the almost negligible influence of water and carbon dioxide on Raman spectra, as opposed to IR absorption spectroscopy [13]. Additionally, as the infrared peaks are often broad, they tend to overlap, being rather difficult to found isolated or well-defined peaks, and these might be also influenced from the aforementioned atmospheric components, which display also broad peaks. This problem is overcome in Raman, where a spectrum is normally composed of isolated bands, besides the previously mentioned advantage of the insignificant contributions from atmospheric gases and water in the Raman spectra.

Furthermore, in Raman, if the excitation wavelength matches the electronic absorbance of a chemical system, its Raman signal will be enhanced by resonance, making this technique suitable for monitoring particular compounds, or properties related to specific constituents of a sample. Thus, Raman spectroscopy can provide useful qualitative and quantitative information on chemical composition of olive oils. For instance, El-Abassy et al. (2009) took advantage of the resonance phenomenon in carotenoids at 514.5 nm excitation, and their correlation with free fatty acid (FFA) content, to assess the FFA percent values in distinct EVOO´s using Raman spectroscopy. Actually, the concentrations of both FFA, and carotenoids, change during the ripening process, with an inversely proportional trend, therefore, the changes occurring in those bands corresponding to carotenoids, can be (indirectly) related to the variations in the FFA content [14].

In fact, in the last years, Raman spectroscopy combined with chemometric data analysis methods allowed researchers to determine the total unsaturation in oils [15], the FFA contents [14,16], to discriminate and classify oils [17], to detect oil adulteration [18], and to distinguish the ripening stages of olive fruit [19]. Additionally, the rapid spreading, and developments occurred, concerning portable Raman devices, which are nowadays available with a variety of excitation wavelengths (*e.g.* 1064, 785, 671 and 532 nm), widen the potential of this technique, that can be now used *In situ*, a major advantage regarding this kind of work.

However, to the best of our knowledge, no comprehensive study has been undertaken involving chemometric analyses of the Raman spectra, combined with analytical parameters (peroxide value and free acidity), to discriminate varietal origin of olive oil and different maturation stages of cultivars 'Cobrançosa', 'Galega' and 'Picual' growing in Alentejo region.

Moreover, the vast majority of the aforementioned works have been undertaken with resort to excitation wavelengths within the NIR range (mainly the 1064 radiation from the Nd:YAG), while some authors have more recently used excitation wavelengths within the visible range for the study of olive oils, with optimal performance [14,20]. Therefore, the aim of this study was to use short wavelength Raman spectroscopy (488 nm excitation) associated with chemometrics in order to differentiate EVOO's produced with olives from three cultivars on three different maturation stages. Discrimination was achieved using an unsupervised method, principal component analysis (PCA).

Furthermore, quantitative models for the rapid prediction of the peroxide value and acidity of EVOO's based on Raman spectral data and on its 1st and 2nd derivatives were developed using the principal component regression (PCR) and the partial least square regression (PLS-R) methods.

2. Material and methods

2.1. Sampling

The present work was carried on monovarietal extra virgin olive oils from three cultivars (cv. 'Cobrançosa', 'Galega' and 'Picual'). The olive fruits were obtained from a certified olive orchard, at the National Plant Breeding Station, at Elvas (Portugal) during the crop season 2012/2013.

Only healthy olive fruits, without any kind of infection or physical damage, were collected from ten different trees of comparable age and vigor and located in distinct points of the same growing area. Thus, differences in climate conditions, agricultural practices and geographical were excluded. Olives were handpicked at three ripening stages, except Picual cultivar olives that were picked during two harvesting periods. The harvesting dates are presented in Table 1. For the classification of the maturity index, the olives were evaluated according to their skin and pulp color [21]. The ripeness index (RI) values range from 0 (100% intense green skin) to 7 (100% purple flesh and black skin).

After harvesting, the olive fruits were immediately transported to the laboratory mill where they were processed within 24 h. For the production of each olive oil, three kilos of fresh olive fruits were used using an Abencor system (INIA I.P., Elvas, Portugal) where olives were crushed with a hammer crusher and the past mixed at 25 °C for 30 min, centrifuged without addition of warm water and then filtered and transferred into dark glass bottles without headspace and stored in the dark at 4 °C until analysis. All samples were classified as extra virgin olive oils according to the EU regulations [7].

2.2. Quality chemical indices determination

For all the samples, the determination of the physicochemical quality indexes, free acidity (expressed as % of oleic acid) and peroxide value (expressed as meq O_2/kg), were performed according to the official method described in the EEC Reg. No. 2568/91 and subsequent amendments, which focus on the characteristics of olive oil and olive-residue oil and on relevant methods of analysis [7].

2.3. Raman spectra measurement

The Raman spectra were collected in the 250–3050 cm^{-1} spectral range, at room temperature, with a T64000 Jobin-Yvon triple spectrometer, coupled to a liquid-nitrogen-cooled CCD detector. The excitation line was the 488 nm of an Ar laser, and the incident laser power was kept below 100 mW on the sample, to avoid self-heating. A confocal Olympus microscope equipped with a 100X objective (NA=0.95) was used to focus the laser radiation on the sample and for collecting the Raman scattered signal in a back-scattering geometry. Under these conditions, which were maintained for all experiments, the spectral slit width was about 2.0 cm^{-1}. A substrate of crystalline silicon was used under the same conditions as an external standard for calibration, by recording the position of its well-defined Raman band at 520.0 cm^{-1}.

For the measurement, a drop of olive oil was placed directly on a special microscope slide with a small concavity by pipetting a

Table 1
Description of olive fruits at each sampling date.

Cultivar	Maturity stage	Sample code	Harvest date	RI[a]
Cobrançosa	Green	Cob G	02/10/2012	0.4
	Semi-ripe	Cob SR	12/10/2012	2.1
	Ripe	Cob R	08/11/2012	5.5
Galega	Green	Gal G	02/10/2012	0.4
	Semi-ripe	Gal SR	12/10/2012	2.1
	Ripe	Gal R	08/11/2012	5.5
Picual	Green	Pic G	02/10/2012	0.4
	Ripe	Pic R	08/11/2012	5.5

[a] Ripening Index.

few drops. The integration time was set to 20 s and each recorded Raman spectrum is the average of five accumulations. The time necessary for acquire one spectrum was about 4 minutes. For each olive oil sample, four replicates were measured in identical conditions, for each one the microscope was focused on two different points and the corresponding Raman spectra recorded, therefore, retrieving eight spectra for each sample.

2.4. Mathematical treatment of data set

The Raman spectral data can be considered as a multidimensional set of variables where each spectrum consists of hundreds of variables, corresponding to the Raman intensities (counts) recorded at the different Raman shifts (cm^{-1}). These data sets were then analyzed by a multivariate analysis method as described below.

A principal component analysis (PCA) was applied to inspect differences between samples. The PCA transforms a large number of potentially correlated factors into a smaller number of uncorrelated factors (principal components, PCs), and thus reduces the size of the data set.

For qualitative analysis, principal components contributing to the variance of the data set were subjected to discriminant analysis (DA) in an attempt to predict the likelihood of a sample belonging to a previously defined group. Since the raw spectral data could not be used because of the strong correlation between the wavenumbers, uncorrelated PCs resulting from PCA were employed. DA is a statistical method used to find a linear combination of structures, which characterizes or separates classes of objects or observations. The resulting arrangement may be used as a linear classifier or dimensionality reduction priori to classification.

Both PCR and PLS regression methods are utilized to model a response variable when a biological system is analyzed through a large number of predictor variables that are highly correlated or even collinear [22]. Both methods give rise to new predictor variables, usually known as principal components (PC) or latent variables that are linear combinations of the original predictor variables. Those components are calculated in different ways, PCR creates PC's to explain the observed variability in the predictor variables, without considering the response variable. In PLS-R the response variable is taken into account and, consequently, leads to models that can fit the response variable with fewer factors [23]. Whether or not it ultimately translates into a useful model depends on the specific biological system.

2.5. Treatment of spectral data

Pre-treatment of the raw spectral data included several steps. To subtract the fluorescence background the baseline of each spectrum was approximated by a fourth-order polynomial. After smoothing the spectra were normalized using the standard normal variate (SNV) method [24,25]. First and 2^{nd} derivatives were determined by the Savitzky–Golay method [26].

The Raman spectra were collected in the 250–3050 cm^{-1} spectral range. However, it was found that well-defined Raman bands associated with free fatty acids and with changes due to oxidation could be observed in specific regions. The most relevant spectral region that can be linked to free acidity and to the peroxide value is the region between 950–1800 cm^{-1}, in accordance with El-Abassy et al. (2009) [14]. The inclusion of irrelevant spectral information in the Chemometrics procedures yields an over-fitting model.

2.6. Model selection

For quantitative analysis of chemical parameters of olive oil samples, factors considerably contributing to the variance of the data set were regressed using principal component regression (PCR) and partial least squares regression (PLS-R) onto the referred variables. This multivariate calibration technique, sometimes called factor analysis, transform the original variables (Raman spectra) into the new ones (known as factors), which are linear combination of original variable.

The method relied on two steps, so-called calibration and validation. In the calibration step, a mathematical model was built to establish a correlation between the matrix of Raman spectra (predictor variables) and the concentration of analytes of interest (response variables) used a set of observations usually named calibration set. In the validation step, the developed calibration model was used to calculate the concentration of samples not used to set-up the model [26].

The quality of the fitting was scrutinized by the root mean square error of calibration (RMSEC), multiple coefficient of determination or regression coefficient (R^2, where R is the correlation factor) and by the root mean square error of cross validation (RMSECV). The optimum number of factors either for PLS-R or PCR models was determined using Leave-One-Out (LOO) cross-validation method. The optimal number of factors is the one that minimizes the RMSECV.

PCA, DA, PCR and PLS-R calculations were performed using the XLSTAT-v2006.06 package (Addinsoft, Inc).

3. Results and discussion

3.1. Chemical analysis

The regulated quality indices studied in this work (free acidity and peroxide value) are displayed in Table 2. In all samples of cv Cobrançosa, Galega and Picual olive oils, analytical parameters were widely within the limits established in the European Legislation for EVOO [27].

The PV is useful once it's an indicator of the initial stage of oxidation, where the primary oxidation products are measured. As can be seen, PV decreased as the maturation stage increased and were significantly lower in Galega and Picual olive oils in the first and latest ripening stage than in Cobrançosa olive oils. This behavior can be explained by a decrease in the activity of lipoxygenase enzyme [28]. Furthermore, for each cultivar studied,

Table 2
Means and standard deviations for peroxide value and free acidity of olive oil samples from cv's Cobrançosa, Galega and Picual in three maturation stages, for four replicates ($n=4$).

Cultivar		Peroxide value (meq O_2/kg)	Free acidity (% oleic acid)
Cobrançosa	Green	$8.33^{aA} \pm 0.42$	$0.29^{aA} \pm 0.01$
	Semi-ripe	$6.69^{bA} \pm 0.42$	$0.24^{bA} \pm 0.01$
	Ripe	$4.97^{cA} \pm 0.25$	$0.31^{aA} \pm 0.02$
	p	< 0.001	< 0.001
Galega	Green	$6.66^{aB} \pm 0.33$	$0.21^{aB} \pm 0.01$
	Semi-ripe	$7.50^{bB} \pm 0.37$	$0.17^{bB} \pm 0.01$
	Ripe	$3.34^{cB} \pm 0.17$	$0.24^{cB} \pm 0.01$
	p	< 0.001	< 0.001
Picual	Green	$8.27^{aA} \pm 0.41$	$0.31^{aA} \pm 0.02$
	Ripe	$3.33^{bB} \pm 0.17$	$0.17^{bC} \pm 0.01$
	p	< 0.001	< 0.001

For each cultivar, means with different lower case letters differ significantly. For each maturation stage, means with different capital letters differ significantly.

significant differences of PV were evident during the maturation process.

Regarding free acidity, this percentage didn't show a trend during fruit ripening of cultivars studied. For Cobrançosa and Galega samples, free acidity decreased from the first to the second maturation stage and then increased in the ripe stage. However, olive oil samples from Picual cultivar presented free acidity values ranged between 0.31 ± 0.02 to $0.17 \pm 0.01\%$ in green and ripe stages, respectively. Usually, olives at a later stage of ripening originate oils with higher levels of acidity due to a progressive enzymatic activity, especially by lipolytic enzymes, and are more sensitive to pathogenic infections and mechanical damage [28]. This increase has been observed for Cobrançosa and Galega cultivars, but not for Picual olive oils that revealed a significant decrease of free fatty acids during maturation, in agreement with previously reported data by Garcia et al. (1996) [29].

3.2. Raman spectra of olive oil

Representative Raman spectral data in the range of 950 and 1800 cm^{-1} together with its 1st and 2nd derivatives are shown in Fig. 1 for EVOO from cv Cobrançosa in three maturation stages.

The Raman spectrum includes bands that can provide information on biochemical changes occurring during the maturation of olives used to produce the olive oils [19]. Some of these bands are displayed in the aforementioned region and correspond to the peaks: at 1749 cm^{-1}, associated to the ester stretching mode υ (C=O), at 1651 cm^{-1}, due to stretching mode υ (C=C) of the cis double bond, at 1439 cm^{-1} and 1303 cm^{-1}, assigned to the methylene (CH_2) scissoring and twisting deformations, respectively, and at around 1267 cm^{-1}, related to the scissoring deformation in the cis unsaturated moiety δ (=C-H) [30]. These bands are assigned to fatty acids, both free and in triacylglycerols, therefore, being directly related to the fatty acid content of each sample. It must be mentioned that the band at 1267 cm^{-1} is very week under 488 nm excitation. This band has been pointed out by Baeten [30], and more recently by Machado [31], which related its intensity loss to the degradation of the *cis*-unsaturated moiety, occurring during the peroxidation process, therefore, being proposed for monitoring the oxidation status.

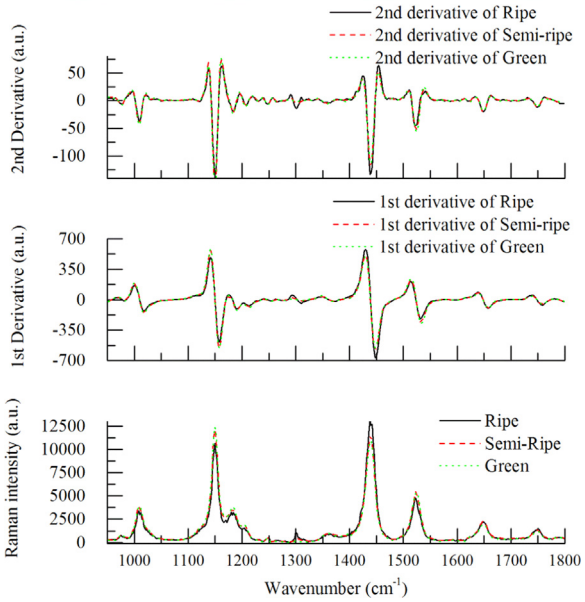

Fig. 1. Raman spectra between 950 and 1800 cm^{-1} of extra virgin olive oil from cv Cobrançosa in three maturation stages together with its 1st and 2nd derivatives.

Regarding the spectral changes due to compositional differences, there are three more bands to be pointed at around 1009 cm^{-1}, 1150 cm^{-1} and 1523 cm^{-1}, attributed to carotenoids, and assigned to the (C-CH_3) deformation, υ (C-C) and υ (C=C) mode, respectively [32]. These compounds are natural antioxidants, responsible for different characteristics of distinct EVOO's, and their quantity changes during the maturation process [19]. Their bands are not observed in investigations using laser excitation wavelengths within the Infrared range (*e.g.* 785, 1064 nm), however, they became visible under shorter excitation wavelengths, such as the 488 nm laser line presently used. Therefore, short wavelength excitation improves differentiation between olive oils produced using distinct cultivars, in different maturation stages. Similar results were reported in El-Abassy et al. (2009) using 514.5 nm excitation [14].

3.3. Supervised olive oil maturation monitoring using Raman spectral data

3.3.1. Preliminary analysis of the spectral data set

Standardized (SNV) Raman spectral data collected between 950 and 1800 cm^{-1} of the 64 spectra (8 spectra for each maturation stage, resulting in 24 spectra from Olive cv Cobrançosa and Galega, in three maturation stages, and 16 from Picual cv in two maturation stages) were subjected to PCA. It was found that the variance of the data set, based in 425 original variables (wavenumbers), is described by 47 principal components but 95% of the total variance is explained by the first 11 principal components.

In statistics, communality is defined as the sum of the squared factor loadings for all factors for a given original variable. It is the variance in that variable accounted for by all the factors. In other words, the communality measures the percentage of variance in a given variable explained by all the factors jointly and may be interpreted as the consistency of the indicator [33]. By definition, the initial value of the communality in PCA is 1. Small communalities values after extraction indicate variables that do not fit well the factor solution and should be dropped from the analysis [34]. According to Stevens (2002), a lower limit of 0.6 should be used [35].

Wavenumbers (variables) for which the communality value of each principal components out of the eleven was higher or equal to 0.6 were considered as meaningfully explaining the variance of the spectral data set and then were considered as prospective wavenumbers associated with the biochemical changes happening during maturation of olives used to produce the olive oil samples. These wavenumbers are in agreement with the above mentioned bands related to compositional differences, and belong to the following intervals: 997 to 1023 cm^{-1}, 1139 to 1161 cm^{-1}, 1295 to 1309 cm^{-1}, 1425 to 1455 cm^{-1}, 1511 to 1537 cm^{-1}, 1631 to 1665 cm^{-1} and 1737 to 1761 cm^{-1}. This new set of 86 variables, each one corresponding to a Raman intensity, collected from the above-mentioned intervals (one registered for each 2 cm^{-1}), was then selected for additional analyses. The band at 1267 cm^{-1}, used by other authors, has been left out after this selection procedure of intervals for DA. This is probably due to the low intensity of this band, under the presently used excitation of 488 nm, which led its variation to be insignificant for this analysis.

A new PCA performed on this reduced spectral data set showed that 89% of the variance could be explained by five principal components. These variables were then subjected to a discriminant analysis based on the known membership (maturation stage) of each sample analyzed, constituting the dependent variable.

The observation diagram represented in 2 is defined by discriminant factors F1 and F2, which explained the total variance. The cultivars of green ripening stage are in the lower left side of the plot, with exception of Picual, placed near the center. EVOO's

of Cobrançosa and Galega cultivars in semi-ripe stage are mostly located in positive part of F2 axis and negative part of F1 axis. EVOO' samples in ripe stage are located in the right side of the F1/F2 plane. The classification in Table 3, resulting from the discriminant analysis, provided 96.5% correct classification for the calibration set and 81.9% correct classification when cross-validated. A high degree of correct classification (91.7%) is achieved for ripe EVOO, as opposed to the semi-ripe group that has 25% probability of being classified as green. Thus, though not being suitable for the discrimination between green and semi-ripe EVOO's, discriminant analysis allows the possibility to distinguish EVOO samples obtained from ripe fruits.

3.3.2. PCR and PLS-R models for prediction of peroxide value and free acidity

PLS-R and PCR calibration models were carried out in order to determine a relationship between predictor variables (Raman scattering intensities) and the chemical characteristics of olive oils referred above. Specifically, the main goal was to develop a model able to predict the peroxide value and free acidity, considering the high relevance these parameters for consumers of EVOO.

Actually, as the olive oil Raman spectrum contains contributions from different species, each observable band might either correspond to a specific component, or to a sum of different contributions (bands) arising from distinct chemical systems. Thus, Raman is suitable for the assessment of these chemical parameters – peroxide value and free acidity – due to the different contributions of the samples' constituents, which vary, and also to the chemical modifications related to these parameters, which are reflected in the behavior of certain bands. For instance, some Raman signs are specifically related to the peroxide value, such as the vibrational modes due to the conjugated (C=C-C=C) system formed due to the peroxidation process (stretching ~1660 cm^{-1}),

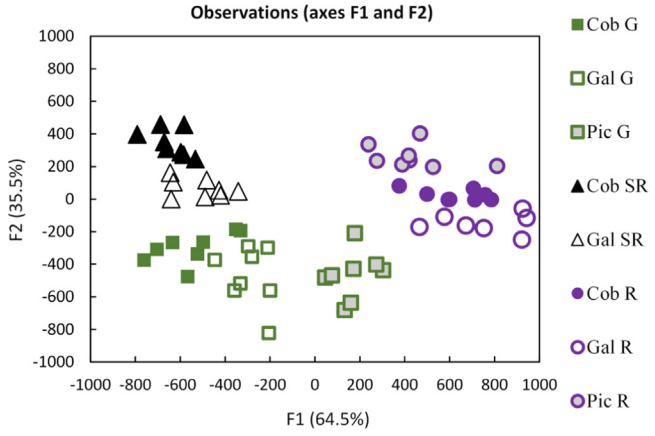

Fig. 2. Observations diagram from a principal component analysis of extra virgin olive oil samples from cv's Cobrançosa, Galega and Picual in three maturation stages, using the normalized (SNV) Raman spectral data.

or the loss of intensity in bands corresponding to the intact fatty acids, such as the band at 1267 cm^{-1} [30,31].

Furthermore, triacylglycerides represent different chemical systems, respecting free fatty acids or glycerol molecules. For example, the band at 1749 cm^{-1}, is raised by the carbonyl ν (C=O) mode, either from the free fatty acids (carboxylic group), or from the triglyceride bonded fatty acids (ester group), displaying slightly different frequencies. Thus, the shape of this band, and Raman features on the corresponding interval (1737–1761 cm^{-1}), retrieve information on the free acidity. Moreover, fatty acids might undergo structural changes when inserted in a triglyceride, these conformational changes are reflected in the frequency of the ν(C=C) mode [36]. Therefore, the band at 1651 cm^{-1}, corresponding to the ν(C=C) modes of the fatty acids, arises as meaningful for the assessment of free acidity. Additionally, in our case, the three bands of carotenoids (1009, 1150 and 1523 cm^{-1}), whose intensities are known to be inversely proportional to the FFA content [14], can be detected. Therefore, there is a rational relationship between Raman obtained data, and these chemical parameters, both free acidity and peroxide value, thus, this spectroscopic technique displays the potential for their assessment.

PLS regression method is able to collect information from large spectral intervals, correlating changes therein to the concentration of specific constituents, and concomitantly consider other possible contributions to these changes not related to the sample constitution [37]. On the other way, in PCR method the spectral and concentration information are incorporated into the model in one step [38].

The results achieved for the PLS-R and PCR calibrations of peroxide value and free acidity in terms of R^2, RMSEC and RMSECV, either for normal spectra, from 950 to 1800 cm^{-1}, and its 1st and 2nd derivatives are presented in Table 4. The high value of R^2 and the lowest of RMSEC and RMSECV indicate the good performance and precision of the models. Furthermore, to obtain a good calibration model, the number of regression factors used should be the lowest as possible [39]. The plot of the measured concentrations of the chemical parameters against the predicted values based on Raman scattering intensities reveals the quality of the models. To illustrate the quality of the models established in this work, PLS-R calibrations are presented in Fig. 3a) and b).

Moreover, though some authors undertook previous works resorting to spectral intervals bellow 1600 cm^{-1} [14], the usefulness of the bands in the 1600–1800 cm^{-1} interval, which reached a relative intensity high enough for its variations to be considered in the model, has been observed in the present study, also Korifi has previously considered these bands, in a study conducted with the 532 nm excitation [20] Therefore, it can be elated that, besides the enhanced intensity of these bands, the improved sensitivity due to the use of a lower excitation wavelength, allows the retrieval of useful information from the shape of these bands, which are raised by contributions from different chemical systems.

For both chemical parameters under analysis, the multivariate calibrations showed, in general, the highest value of R^2 and the lowest of RMSECV when using the 1st derivative of the spectral

Table 3
Confusion matrix for the training and for the cross-validation sets based on the maturation stage. Observed classifications in the rows. Predicted classifications in columns.

From/To	Calibration					Cross-validation				
	Green	Semi-ripe	Ripe	Total	% correct	Green	Semi-ripe	Ripe	Total	% correct
Green	23	1	0	24	95.8	19	5	0	24	79.2
Semi-ripe	1	15	0	16	93.8	4	12	0	16	75.0
Ripe	0	0	24	24	100	2	0	22	24	91.7
Total	24	16	24	64	96.5	25	17	22	64	81.9

Table 4
Partial least squares and principal component regressions based calibrations for quantification of peroxide value and acidity, using the Raman spectral data and its 1st and 2nd derivatives.

Chemical parameter	Regression method	Factor	Spectra	Equation		R^2		RMSE	
				Calibration	Validation	Calibration	Validation	Calibration	Validation
Peroxide value	PLS-R	4	Normal	y=0.83 x+1.23	y=0.82 x+1.27	0.979	0.966	0.80	0.85
		4	1st der	y=1.01 x−0.02	y=1.04 x+0.10	0.986	0.971	0.68	0.72
		5	2nd der	y=0.99 x+0.04	y=0.98 x+0.07	0.965	0.956	0.74	0.81
	PCR	5	Normal	y=1.01 x−0.08	y=1.04 x−0.17	0.975	0.968	0.84	0.89
		3	1st der	y=0.90 x+0.61	y=0.93 x+0.42	0.989	0.981	0.72	0.77
		5	2nd der	y=1.01 x−0.08	y=1.04 x−0.17	0.937	0.888	0.77	0.83
Free acidity	PLS-R	4	Normal	y=0.93 x+1.19	y=1.13 x−0.11	0.992	0.987	0.02	0.03
		3	1st der	y=1.01 x+0.04	y=1.04 x−0.01	0.994	0.988	0.01	0.02
		4	2nd der	y=0.98 x+0.59	y=1.11 x+0.31	0.981	0.952	0.02	0.03
	PCR	4	Normal	y=0.83 x+1.23	y=1.02 x−0.05	0.988	0.984	0.02	0.04
		4	1st der	y=0.97 x+0.03	y=1.05 x−0.32	0.991	0.990	0.02	0.02
		5	2nd der	y=0.88 x+0.61	y=0.97 x+0.24	0.969	0.933	0.03	0.04

RMSE of Peroxide value in meq O_2/kg; RMSE value of Acidity in %.

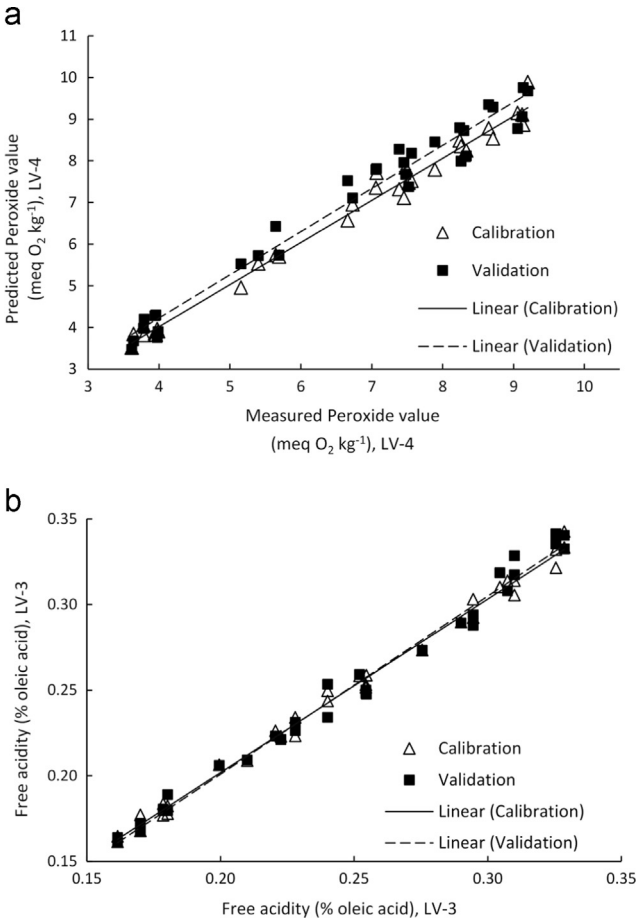

Fig. 3. PLS-R calibration models for the relationship between measured values and Raman predicted values of: a) peroxide values and of b) free acidity of olive oil.

data. In general, PLS-R offers enhanced results compared to others regression methods, such as PCR [40] or multiple linear regression (MLR) for quantitative analysis of chemical parameters [23]. In our study, despite both calibration models presented good prediction values, it was also PLS-R method that showed the best results, displaying good performances for both the prediction of peroxide value, and free acidity.

The relationship between measured values of the chemical parameters predicted values based on Raman scattering intensities showed R^2 values ranging from 0.937 to 0.994 using the calibration set and from 0.933 to 0.990 when cross validated. The RMSECV were as low as 0.72 meq O_2/kg for the peroxide value, when using the 1st derivative and a PLS model with 4 factors. However, a slightly higher value of 0.77 meq O_2/kg was attained for a simpler 3 factor PCR model. For the free acidity the lowest RMSECV value was 0.02% attained for both PCR and PLS-R models, the latter being simpler (3 factors) than the former.

4. Conclusion

In the present work, Raman spectroscopy, alongside with multivariate analysis methods, has been successfully used, in order to monitor free acidity and peroxide values in EVOO's, as well as to distinguish ripe from non-ripe oils. Since each Raman spectrum is constituted of several different contributions, arising from distinct constituents, these unique Raman features, corresponding to each species, render a possibility to the assessment of the samples' content through this methodology.

Furthermore, the use of a short wavelength excitation, 488 nm, allowed, not only, to observe bands due to carotenoids, which are important for these assessments, but also to register good quality spectra, so as to retrieve information from bands raised by distinct constituents.

In order to correlate the Raman spectra with the quality parameters assessed, as well as to evaluate the best approach to do so, the present work has been conducted with resort to two distinct approaches, PCR and PLS-R. Both have shown good results, and best performance when applied to the first derivative of the spectra, for the prediction of either peroxide value, or free acidity. In the latter case both approaches, PCR and PLS-R, retrieve similar results, presenting R^2 values of 0.990 and 0.988, respectively, and equal errors for cross-validation. In the case of peroxide value, PLS-R displays somewhat lower errors of validation in comparison to PCR (0.72 *vs* 0.77 meq O_2/kg) therefore, the former being proposed as the most suitable for these determinations.

Additionally, Raman shows some advantages over Infrared technique, which are exploited by the advent of the Raman portable devices, bringing the hot prospect of the *in situ* monitoring of these quality parameters. Therefore, this methodology represents an important trend, respecting the assessment of these quality parameters in field.

Acknowledgments

This work was supported by the Project 'PhD grant SFRH/BD/78013/2011 to IG from the Fundação para a Ciência e Tecnologia.

This work is financed by the ERDF – European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and by National Funds through the FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project «FCOMP-01-0124-FEDER-037281».

This work was also supported by the Project INNOFOOD - INNOvation in the FOOD sector through the valorization of food and agro-food by-products - NORTE-07-0124-FEDER-0000029, financed by the North Portugal Regional Operational Programme (ON.2 – O Novo Norte) under the National Strategic Reference Framework (QREN), through FEDER, as well as by PIDDAC through FCT/MEC, and by the project PEst-OE/AGR/UI4033/2014.

References

[1] A. Bakhouche, J. Sánchez, R. Beltrán-Debón, J. Joven, A. Segura-Carreter, A. Fernández-Gutiérrez, Food Res. Int. 50 (2013) 401–408.
[2] M. Machado, C. Felizardo, A.A. Fernandes-Silva, F.M. Nunes, A. Barros, Food Res. Int. 51 (2013) 412–421.
[3] C. Sousa, I. Gouvinhas, D. Barreira, M.T. Carvalho, A. Vilela, J. Lopes, P. Martins-Lopes, A.I. Barros, J. Am. Oil Chem. Soc. http://dx.doi.org/10.1007/s11746-013-2406-x.
[4] S. Dabbou, I. Gharbi, S. Dabbou, F. Brahmi, A. Nakbi, M. Hammami, Afr. J. Biotechnol. 10 (74) (2011) 16937–16947.
[5] J. Ayton, R.J. Mailer, A. Haigh, D. Tronson, D. Conlan, J. Food Lipids. 14 (2007) 138–156.
[6] European Commision. Commission Regulations (EC), Off. J. Eur. Comm., L155 (2002) 27–31.
[7] European Community Commission Regulation (EEC) Off. J. Eur. Comm., L248 (1991) 1–83.
[8] E.C.Y Li-Chan, P.R. Griffiths, J.M. Chalmers, Applications of Vibrational Spectroscopy in Food Science, Analysis of Food, Drink and Related Materials, Volume II, John Wiley and Sons, Chichester, UK, 2010.
[9] M. Holse, F.H. Larsen, A. Hansen, S.B. Engelsen, Food Res. Int. 44 (1) (2011) 373–384.
[10] B.J. Marquardt, J.P. Wold, Food Sci. Technol 37 (1) (2004) 1–8.
[11] L. Kyung-Min, J. Timothy, Y. Herrman, J. Cereal. Sci. 59 (2014) 70–78.
[12] A.M.A. Pistorius, Spectrosc. Eur 7 (1995) 8–15.
[13] A. Zaja, J. Hanuza, L. Dymin ska, Food Chem. 156 (2014) 333–338.
[14] R.M. El-Abassy, P. Donfack, A. Materny, J. Am. Oil. Chem. Soc. 86 (2009) 507–511.
[15] R.C. Barthus, R.J. Poppi, Vib. Spectr 26 (1) (2001) 99–105.
[16] B. Muik, B. Lendl, A. Molina-Díaz, M.J. Ayora-Cañada, Anal. Chim. Acta 487 (2003) 211–220.
[17] H. Yang, J. Irudayaraj, M.M. Paradkar, Food Chem. 93 (2005) 25–32.
[18] V. Baeten, J.A.F. Pierna, P. Dardenne, M. Meuren, D.L. García-González, R. Aparicio-Ruiz, J. Agr. Food Chem. 53 (2005) 6201–6206.
[19] M. López-Sanchez, M.J. Ayora-Cañada, A. Molina-Díaz, J. Agric. Food Chem. 58 (2010) 82–87.
[20] R. Korifi, Y. Le Dréau, J. Molinet, J.N. Artaud, Dupuy, J. Raman Spectrosc. 42 (2011) 1540–1547.
[21] M. Uceda, M. Hermoso, El Cultivo del Olivo, Junta de Andalucia Ediciones Mundi-Prensa (1998) 547–572.
[22] J.N. Miller, J.C. Miller (Eds.), Pearson Education Limited, Edinburgh Gate Harlow, 2005.
[23] P.D. Wentzell, L.V. Montoto, Chemometr. Intell. Lab. Syst. 65 (2003) 257–279.
[24] A. Nunes, J. Martins, A.S. Barros, A.C. Galvis-Sánchez, I. Delgadillo, Sens. Instrum. Food Qual 3 (2009) 187–191.
[25] A. Bendini, L. Cerretani, F. Virgilio, P. Belloni, G. Lercker, T.G. Toschi, Eur. J. Lipid Sci. Technol. 109 (2007) 498–504.
[26] M. De Luca, F. Oliverio, G. Loele, G. Ragno, Chem. Intel. Lab. Syst. 96 (2009) 14–21.
[27] European Community Commission Regulation (EEC), Off. J. Eur. Comm., L23 (2011) 1–13.
[28] F. Gutiérrez, B. Jiménez, A. Ruiz, M.A. Albi, J. Agric. Food Chem. 47 (1999) 121–127.
[29] J.M. Garcia, S. Seller, M.C. Pérez-Camino, J. Agr. Food Chem. 44 (1996) 3516–3520.
[30] V. Baeten, J. Agric. Food Chem. 46 (1998) 2638–2646.
[31] N.F.L. Machado, L.A.E. Batista de Carvalho, J.C. Otero, M.P.M. Marques, J. Raman Spectrosc. 43 (2012) 1991–2000.
[32] P.S. Bernstein, J. Pure Appl. Chem. 74 (2002) 1419–1425.
[33] H. Abdi, L.J. Williams, Comput. Stat. 2 (2010) 43–459.
[34] A.P. Field, Discovering Statistics using SPSS, 2nd Edition, Sage, London, 2005.
[35] J.P. Stevens, Applied Multivariate Statistics for the Social Sciences, 4th Edition, Erlbaum, Hillsdale, NJ, 2002.
[36] M. Motoyama, Bull. NARO Inst. Livest. Grassl. Sci. 12 (2012) 19–68.
[37] Y.B. Che Man, Z.A. Syahariza, M.E.S. Mirghani, S. Jinap, J. Bakar, Food Chem. 90 (2005) 815–819.
[38] B.C. Smith, Quantitative Spectroscopy: Theory and Practice, Academic Press, Amsterdam (2002) 125–179.
[39] L.U. Hui-Shan, X.U. Hui-Rong, Y. Yi-Bin, F.U. Xia-Ping, Y.U. Hai-Yan, T.J. Hai-qing, Zhejiang Uni. Sci 7 (2006) 794–799.
[40] Y. Liang, O.M.K. Valheim, Chemometr. Intell. Lab. Syst. 32 (1996) 1–10.