



# Discrimination and characterisation of extra virgin olive oils from three cultivars in different maturation stages using Fourier transform infrared spectroscopy in tandem with chemometrics



Irene Gouvinhas<sup>a,\*,1</sup>, José M.M. de Almeida<sup>b,c,1</sup>, Teresa Carvalho<sup>d</sup>, Nelson Machado<sup>a,1</sup>, Ana I.R.N.A. Barros<sup>a,1</sup>

<sup>a</sup> CITAB – CITAB, University of Trás-os-Montes and Alto Douro, 5001-801 Vila Real, Portugal

<sup>b</sup> Department of Physics, University of Trás-os-Montes and Alto Douro, 5000-801 Vila Real, Portugal

<sup>c</sup> INESC Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

<sup>d</sup> Instituto Nacional de Investigação Agrária (INIA I.P.), Elvas, Portugal

## ARTICLE INFO

### Article history:

Received 8 September 2013

Received in revised form 19 September 2014

Accepted 6 November 2014

Available online 14 November 2014

### Keywords:

Fourier transform infrared spectroscopy

Extra virgin olive oil

Maturation stages

Cultivars

Chemometrics

Chemical parameters

## ABSTRACT

A methodology based on Fourier transform infrared (FTIR) spectroscopy, combined with multivariate analysis methods, was applied in order to monitor extra virgin olive oils produced from three distinct cultivars on different maturation stages. For the first time, this kind of methodology is used for the simultaneous discrimination of the maturation stage, and different cultivars.

Principal component analysis and discriminant analysis were utilised to create a model for the discrimination of olive oil samples. Partial least squares regression was employed to design calibration models for the determination of chemical parameters. The performance of these models was based on the multiple coefficient of determination ( $R^2$ ), the root mean square error of calibration (RMSEC) and root mean square error of cross validation (RMSECV). The prediction models for the chemical parameters resulted in a  $R^2$  ranged from 0.93 to 0.99, a RMSEC ranged from 1% to 4% and a RMSECV from 2% to 5%.

It has been shown that this kind of approach allows to distinguish the different cultivars, and to clearly discern the different maturation stages, in each one of these distinct cultivars.

Furthermore, the results demonstrated that FTIR spectroscopy in tandem with chemometric techniques allows the creation of viable and accurate models, suitable for correlating the data collected by FTIR spectroscopy, with the chemical composition of the EVOOs, obtained by standard methods.

© 2014 Elsevier Ltd. All rights reserved.

## 1. Introduction

Extra virgin olive oil (EVOO) is a vegetable oil made from healthy and intact fruits of the olive tree (*Olea europaea* L.) only by mechanical means (crushing, malaxation and centrifugation) and can be directly consumed unrefined by humans. As no chemicals are used in this extraction process, the EVOO keeps

the original characteristics and constituents which are lost in refined oils (Nieto, Hodaifa, & Peña, 2010). EVOO is one of the most significant food products in Mediterranean countries, and the olive tree counts among the oldest and most important oil-producing crops after the oil palm (Baldoni & Belaj, 2009).

The high demand for olive oil is associated with the Mediterranean culture based on dietary habits correlated with health benefits (Allalout et al., 2011). This has been correlated with the presence of high content of monounsaturated fatty acids, specifically oleic acid (60–80%) and its richness in minor components, including tocopherols and phenolic compounds, that other seed oils lack (Cicerale, Conlan, Sinclair, & Keast, 2009). These phenolic compounds have a great importance in biological systems once they act as natural antioxidants (Bendini, Cerretani, Carrasco-Pancorbo, et al., 2007; Bendini, Cerretani, Di virgilio, et al., 2007). Furthermore, they are also responsible for the stability and flavour

**Abbreviations:** FTIR, Fourier transform infrared spectroscopy; ATR, Attenuated Total Reflectance; PCA, principal component analysis; DA, discriminant analysis; PLS-R, partial least squares regression; PCR, Principal Component Regression; RI, Ripening Index; GAE, gallic acid; ABTS, 2,2-azino-bis(3-ethylbenzothiazoline)-6 sulphonic acid;  $R^2$ , multiple coefficient of determination; RMSEC, root mean square error of calibration; RMSECV, root mean square error of cross validation.

\* Corresponding author. Tel.: +351 259350283; fax: +351 259350480.

E-mail address: [igouvinhas@utad.pt](mailto:igouvinhas@utad.pt) (I. Gouvinhas).

<sup>1</sup> <http://www.utad.pt>.

of olive oil, and are associated to “pharmacological” properties, since these compounds have demonstrated some positive effects on certain physiological parameters (Bouskou, Tsimmidou, & Blekas, 2006). However, the quality of olive oil is influenced by a great number of factors and its phenolic composition and concentration depends on two of the most important ones, namely the nature of the cultivar and geographic origin (Bakhouch et al., 2013), and of the fruit ripening degree (Machado, Felizardo, Fernandes-Silva, Nunes, & Barros, 2013), where important chemical changes occur. There are other factors that affect the phenolic composition of olive oils such as irrigation regimes (Machado et al., 2013), oil extraction technology and the storage of the oil (Dabbou et al., 2011) and agricultural techniques used to cultivate olive fruit (Ayton, Mailer, Haigh, Tronson, & Conlan, 2007).

In this work, we studied olive oils from some cultivars that are cultivated in the main Portugal region for olive oil production (Alentejo) (INE, 2012). ‘Cobrançosa’ is a characteristic cultivar of Trás-os-Montes region. However, due to its interesting characteristics, it has been spread to other regions, namely in Alentejo. Galega, as Cobrançosa cultivar, is one of the most important cultivar used in Portugal, but there is little information available related to the study of its chemical composition. The other cultivar studied is Picual which is one of the most important cultivar grown in Andalusia, Spain, and characterised by having one of the highest content of phenolic compounds (Nieto et al., 2010).

The determination of total phenolic compounds, including *ortho*-diphenols and flavonoids, and antioxidant activity by colourimetric methods involves a pre-treatment of samples and consequently the destruction of the sample. Furthermore, these analyses are time consuming and require large amounts of reagents and solvents, which are quite expensive, and often toxic. To overcome these hurdles, spectroscopic methods have been used, such as Fourier-transform infrared spectroscopy (FTIR), which is an analytical technique, rapid, direct and simple to perform. It is non-destructive and does not require any sample preparation, particularly when used in conjunction with Attenuated Total Reflectance (ATR).

For these reasons, the application of FTIR in the study of olive oils has increased recently, mainly to evaluate the composition of fatty acids (Inarejos-Carcía, Gómez-Alonso, Fregapane, & Salvador, 2013), oxidised fatty acids (Lerma-García, Simó-Alfonso, Bendini, & Cerretani, 2011), peroxide value (Bendini, Cerretani, Carrasco-Pancorbo, et al., 2007; Bendini, Cerretani, Di virgilio, et al., 2007), acidity (Lerma-García et al., 2011), adulterations (Rohman & Che Man, 2012), sensory characteristics, phenolic and volatile compounds (Lerma-García et al., 2011), freshness (Sinelli, Cosio, Gigliotti, & Casiraghi, 2007) and authenticity (Lerma-García, Ramis-Ramos, Herrero-Martínez, & Simó-Alfonso, 2010). Other authors described the use of FTIR–ATR for the simultaneous quantification of fatty acid composition, peroxide value and free acidity (Maggio et al., 2009). Despite some studies used FTIR–ATR–PLS tool to determine some analytical parameters (water content, phenolic content and antioxidant activity) in olive oils (Cerretani et al., 2010), the discrimination of varietal origin of olive oil and different maturation stages of the specific cultivars ‘Cobrançosa’, ‘Galega’ and ‘Picual’ growing in Alentejo region with resort to this technique, has not been undertaken yet.

The aim of this study was to use FTIR–ATR spectroscopy associated with chemometrics in order to differentiate EVOO's produced with olives from three cultivars on three different maturation stages. Discrimination was achieved using either an unsupervised method, principal component analysis (PCA), and a supervised method, factor discriminant analysis (FDA).

Furthermore, quantitative models to predict the chemical characteristics of EVOO's based on FTIR spectra measurements were developed using partial least square regression method (PLS-R)

based on the Non-linear Iterative Partial Least Squares algorithm (NIPALS) algorithms.

## 2. Material and methods

### 2.1. Chemicals

Folin–Ciocalteu's reagent, 3,4,5-trihydroxybenzoic acid (gallic acid) and acetic acid, both extra pure (>99%) were purchased from Panreac (Panreac Química S.L.U., Barcelona, Spain). Sodium nitrate, aluminium chloride and sodium carbonate, all extra pure (>99%), were purchased from Merck (Merck Darmstadt, Germany). 2, 2-azino-bis(3-ethylbenzothiazoline)-6 sulphonic acid (ABTS), Trolox, and catechin, all of extra pure grade (>99%), were purchased from Sigma–Aldrich (St. Louis, MO, USA). Sodium molybdate (99.5%) was purchased from Chem-Lab (Chem-Lab N.V., Zedelgem, Belgium). All other reagents used were of analytical grade. The water used in the experiments was deionised, obtained from a SGS water purification system.

### 2.2. 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 vigour 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 colour (Uceda & Hermoso, 1998). 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 (EEC n.º2568/1991, 1991).

### 2.3. Extraction of the phenolic fraction

Three distinct aliquots were collected from each olive oil. For the extraction of polar phenolic compounds of each one of these

**Table 1**  
Description of olive fruit at each sampling date.

Cultivar	Maturity stage	Sample code	Harvest date	RI <sup>a</sup>
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

<sup>a</sup> Ripening Index.

samples, 4 mL were weighed in a test tube, followed by the addition of 2.5 mL of hexane and 2.5 mL of MeOH/H<sub>2</sub>O (80:20) mixture. The mixture was centrifuged for 10 min at 6000 rpm. The lower phase was carefully discarded and reserved in a 10 mL volumetric flask. To the remaining mixture, 2.5 mL of hexane and 2.5 mL of MeOH/H<sub>2</sub>O (80:20) was added and centrifuged for 10 min at 6000 rpm and the lower phase collected into the same volumetric flask. This procedure was repeated once more.

#### 2.4. Measurement of olive oil chemical parameters

All the absorbance measurements were performed with resort to a SpecTronic Instruments – Genesys 2 PC spectrophotometer.

##### 2.4.1. Total phenolic content

The content of total phenolic compounds in olive oil samples was determined using Folin–Ciocalteu reagent, with gallic acid as standard (Barros et al., 2012). This method is based on the reduction of a phosphotungstate–phosphomolybdate complex by phenolics to blue reaction products. 1 mL of diluted samples was mixed with 500  $\mu$ L of Folin–Ciocalteu reagent, 2 mL of 7.5% sodium carbonate solution and 6.5 mL of water. The mixture was shaken and the absorbance of the standard solutions and samples was measured at 750 nm after 30 min reaction at 70 °C in relation to a gallic acid standard curve. All measurements were performed in triplicate. The results were expressed as milligrams of gallic acid equivalents per gram of sample (mg GAE g<sup>-1</sup>).

##### 2.4.2. Ortho-diphenol content

Ortho-diphenols were measured colourimetrically at 370 nm after adding 1 mL 5% (w/v) sodium molybdate in 50% ethanol to the aliquots of the extract (4 mL). Gallic acid served as a standard to prepare a calibration curve in the 5–200 mg range per gram of solution (Mateos et al., 2001). All measurements were performed in triplicate. The ortho-diphenol content in olive oils was expressed as milligrams of gallic acid equivalents per gram of sample (mg GAE g<sup>-1</sup>).

##### 2.4.3. Flavonoids content

Total flavonoids were measured by the colourimetric assay developed by Zhishen, Mengcheng, and Jiamming (1999). In a test tube 0.5 mL of the sample-working solution, resulting from the extraction procedure, and 150  $\mu$ L of NaNO<sub>2</sub> 5% were introduced. After 5 min, 150  $\mu$ L of AlCl<sub>3</sub> 10% were added and 6 min after 1 mL of 1 M NaOH was added. The mixture was shaken and the absorbance of the standards and samples was measured at 510 nm in relation to a catechin standard curve. All measurements were performed in triplicate. The results were expressed as milligrams of catechin per gram of sample (mg catechin g<sup>-1</sup>).

##### 2.4.4. Antioxidant activity determination

The radical-scavenging activity was determined soon after extraction by the 2,2-azino-bis(3-ethylbenzothiazoline)-6-sulphonic acid (ABTS) radical cation decolourisation assay (Barros, Nunes, Gonçalves, Bennett, & Silva, 2011). For the assay, ABTS<sup>+</sup> radical was prepared by mixing an ABTS stock solution (7 mM in water) with 2.45 mM potassium persulfate. This mixture was allowed to stand for 12–16 h at room temperature in the dark until reaching a stable oxidative state. The ABTS<sup>+</sup> solution was diluted with 20 mM sodium acetate buffer (pH 4.5) to an absorbance of 0.70  $\pm$  0.01 at 734 nm. The reaction was started by the addition of 25, 50, 100, 150 and 200  $\mu$ L of the methanolic extract of the olive oils diluted conveniently to 2 mL of the diluted ABTS<sup>+</sup> solution. ABTS<sup>+</sup> bleaching was monitored at 734 nm and 25 °C for at least

30 min and the percentage of discolouration after 15 min was used as the measure of antioxidant activity. The ABTS<sup>+</sup> bleaching was proportional to the concentration of the sample added to the medium. The antioxidant activity of the extract was calculated as Trolox Equivalent Antioxidant Capacity (TEAC) and was expressed as mmol of Trolox equivalents per kg of sample. All measurements were performed in triplicate. A standard curve of the percentage of ABTS<sup>+</sup> inhibition in function of Trolox concentration (0.11–0.014 mM) was used for the calculations.

#### 2.5. FTIR–ATR measurement

Infrared spectra were collected in a “Unicam Research Series” FTIR spectrometer equipped with a heated “Golden Gate” single reflection ATR module, a DLATGS detector and a KBr beamsplitter. The equipment is connected to computer and controlled by WinFirst Software – v1.1.

FTIR–ATR measurements were performed by pipetting a small drop (~1  $\mu$ L) of olive oil on top of the ATR baseplate, which was kept at 30 °C. We calculate that the evanescent field was probing a depth of approximately 1.0  $\mu$ m. All infrared spectra were recorded from 500 to 3000 cm<sup>-1</sup>, co-adding 128 interferograms at a resolution of 2 cm<sup>-1</sup>. The collection time for each sample spectrum was approximately 2 min. These spectra were subtracted against background air spectrum. After every scan, a new reference air background spectrum was taken. The ATR base was carefully cleaned in situ by scrubbing with ethanol and dried with soft tissue before measuring the next sample. The cleaning method was verified by collecting a background spectrum and compared to the previous one. These spectra were recorded as absorbance values at each data point. Each sample measurement was repeated two times and the spectra averaged.

#### 2.6. Mathematical treatment

Spectral data collected between 500 and 3000 cm<sup>-1</sup> were mean-centred and standardised (1/SD) and subjected to a principal component analysis (PCA) to inspect differences between samples. The PCA transforms the 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.

For quantitative analysis of chemical parameters of olive oil samples, factors considerably contributing to the variance of the data set were regressed using partial least squares regression (PLS-R) onto the referred variables. This multivariate calibration technique, sometimes called factor analysis, transform the original variables (FTIR spectra absorbances) 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 FTIR 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 (De Luca, Oliverio, Loele, & Ragno, 2009). PCA, DA and PLS-R calculations were performed using the XLSTAT-v2006.06 package (Addinsoft, Inc).

### 3. Results and discussion

#### 3.1. Chemical analysis

Chemometric analysis of FTIR spectra on extra virgin olive oils of three cultivars with different maturation stages were combined with analytical parameters, determined only by colourimetric methods, to develop calibration and validation models in order to be able to estimate the chemical parameters, namely, total phenolic content, *ortho*-diphenols, flavonoids and antioxidant activity during the ripening process. Table 2 shows the mean and the standard deviation of these analytical parameters. For Cobrançosa and Galega cultivars, the samples presented an increase of total phenolics, *ortho*-diphenols, flavonoids and antioxidant activity from the first to the second stage of ripening. Then a decrease was observed in the third maturation stage. These results are in agreement with the literature (Nieto et al., 2010). For Picual' EVOO's, the values of all the chemical parameters decreased between the first and the last maturation stages studied, green and ripe. Furthermore, this cultivar presented the biggest values of the analytical parameters in both stages.

ATR-FTIR spectra of all EVOO samples were obtained. Fig. 1 shows an example of the measured spectra collected from 500 to 3000  $\text{cm}^{-1}$ . These particular spectra correspond to cv Cobrançosa olive oil samples in the three maturation stages. The FTIR spectra of all cultivars presented spectral differences due to some peaks, namely at 721  $\text{cm}^{-1}$  associated with  $\text{CH}_2$  rocking modes, and at 1097, 1117, 1160 corresponding to C–O stretching vibrations, the last one displaying strong intensity, and constituting a feature characteristic of the aliphatic esters. Another peak related to the C–O stretching is found, at 1236  $\text{cm}^{-1}$ , in this case presenting some contribution from  $\text{CH}_2$  bending modes. In which respects to the alkylic chains, there are two additional peaks to be pointed in this frequency region ( $<1800 \text{ cm}^{-1}$ ), at 1374 and 1465  $\text{cm}^{-1}$ , the first due to the terminal ( $\text{CH}_3$ ) groups symmetric bending, while the last can be assigned to the scissoring vibration of the  $\text{CH}_2$  groups. Finally, in this frequency interval, the strong peak at 1740  $\text{cm}^{-1}$  arises from the carbonyl ( $\text{C}=\text{O}$ ) stretching vibration, in this case associated with the triglyceride ester linkage, as well as the carboxylic group of free fatty acids (Sinelli et al., 2010; Yang & Irudayaraj, 2007).

The  $\text{CH}_2$  and  $\text{CH}_3$  stretching modes (aliphatic moiety of the fatty acids) are observed in the high wavenumber region (2750–3000  $\text{cm}^{-1}$ ), where the symmetric and asymmetric stretching vibrations of the  $\text{CH}_2$  groups are responsible for the absorbances at 2853 and 2923  $\text{cm}^{-1}$ , respectively (Lerma-García et al., 2010). Additionally, a shoulder at 2953  $\text{cm}^{-1}$  is also visible in this interval, due to the asymmetric stretching of the terminal ( $\text{CH}_3$ ) group from the alkane chains. Though presenting mainly features related with the fatty acids content, namely triglycerols formed by trans-fatty acids, these spectra present fingerprint regions, where the differences allow to distinguish the samples, and can be related to the phenolic content of the olive oil samples, and antioxidant activity, by chemometric procedures.

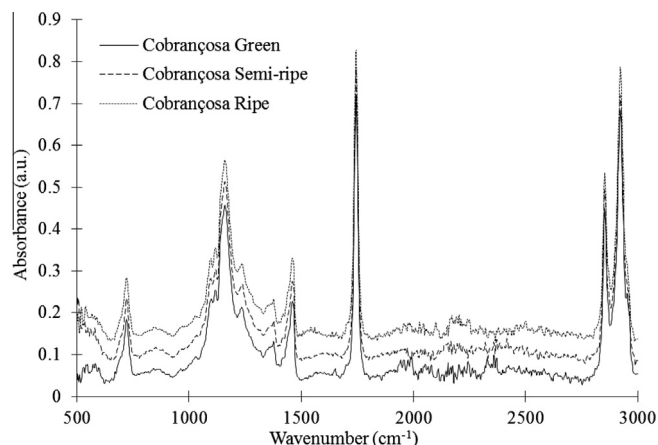


Fig. 1. FTIR-ATR spectra collected for Cobrançosa olive oils samples in three maturation stages.

Based on the aforementioned fact, the wavenumbers selected for chemometrics ranged from 700 to 740  $\text{cm}^{-1}$ , 950 to 1050  $\text{cm}^{-1}$ , 1100 to 1250  $\text{cm}^{-1}$ , 1350 to 1500  $\text{cm}^{-1}$ , 1700 to 1800  $\text{cm}^{-1}$  and 2750 to 3000  $\text{cm}^{-1}$ . The frequencies observed have not changed with the ripening process, while the absorbance value of each functional group varied according to the maturation stage, this fact showing that the variations are due to different quantities of each component, instead of the appearance of new compounds, during the maturation process. A similar behaviour has been observed between the distinct cultivars, where the differences found in the intensities are related to the compositional differences, in quantitative terms, between the EVOO varieties.

#### 3.2. Supervised olive oil maturation monitoring using FTIR spectral data

Mean-centred and standardised FTIR spectral data collected between 3000 and 500  $\text{cm}^{-1}$  of the 64 mean spectra (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. Prior to PCA each spectra was divided in two regions: a first region from 600 to 1800  $\text{cm}^{-1}$  and a second from 2750 to 3000  $\text{cm}^{-1}$ . It was observed that the total variance of the data set could be explained by 53 principal components, among which the first nine explain approximately 91% of the total variance.

For 468 wavenumbers (variables) the communality value of each principal component 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 were then selected for additional analyses.

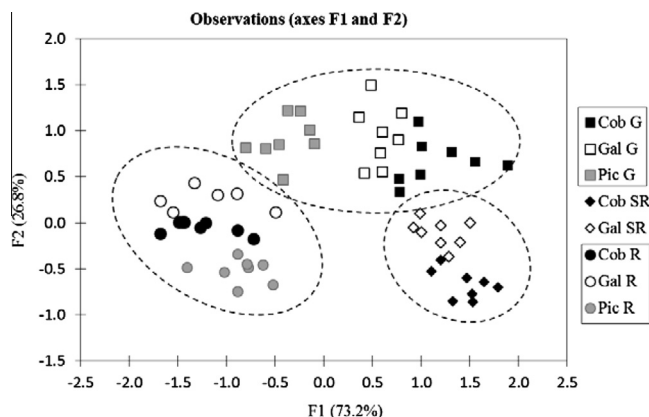
A new PCA was then performed on the spectral data corresponding to the above mentioned 468 variables, which showed

Table 2

Effects of picking date and cultivar on phenolic content, *ortho*-diphenols, flavonoids and antioxidant activity in olive oil samples.

Cultivar	Phenolic content (mg GAE g <sup>-1</sup> )	<i>Ortho</i> -diphenols (mg GAE g <sup>-1</sup> )	Flavonoids (mg catechin g <sup>-1</sup> )	Antioxidant activity (mmol trolox kg <sup>-1</sup> )
Cob green	0.408 ± 0.09	0.213 ± 0.04	0.443 ± 0.11	1.819 ± 0.37
Cob semi-ripe	0.928 ± 0.03	0.465 ± 0.02	1.107 ± 0.03	4.126 ± 0.14
Cob ripe	0.619 ± 0.03	0.315 ± 0.02	0.722 ± 0.04	2.809 ± 0.15
Gal green	0.448 ± 0.01	0.245 ± 0.00	0.528 ± 0.00	2.231 ± 0.03
Gal semi-ripe	0.468 ± 0.01	0.253 ± 0.01	0.573 ± 0.01	2.432 ± 0.01
Gal ripe	0.348 ± 0.03	0.187 ± 0.01	0.373 ± 0.02	1.836 ± 0.07
Pic green	0.960 ± 0.06	0.413 ± 0.02	0.952 ± 0.04	3.655 ± 0.15
Pic ripe	0.739 ± 0.01	0.339 ± 0.01	0.758 ± 0.02	3.057 ± 0.14





**Fig. 2.** Similarity map as determined by discriminant analysis using factors F1 and F2 for FTIR-ATR spectral data of olive oils from different maturation stages.

that the 85% of the variance could be explained by six principal components. These variables were then subjected to a discriminant analysis based on the known membership (maturation stage) of each sample analysed, constituting the dependent variable.

In Fig. 2 it is shown the observation diagram as defined by discriminant factors F1 and F2, which explained the total variance. We observed that all cultivars of green ripening stage are in the positive side of the plot. EVOO's of Cobrançosa and Galega cultivars in semi-ripe stage are mostly located in negative part of F2 axis and positive part of F1 axis. EVOO' samples in ripe stage are located in the left side of the F1 axis.

The classification in Table 3, resulting from the discriminant analysis, provided 100% correct classification for the calibration set and 73.6% correct classification when cross-validated. A high degree of correct classification (87.5%) is achieved for ripe EVOO, as opposed to the semi-ripe group that has ~35% probability of being classified as green. Thus, discriminant analysis allows the possibility to categorise EVOO samples according to its maturation stages.

### 3.3. PLS-R models for prediction of chemical parameters

PLS-R is a regression technique for multivariate data, principally applied for prediction, and is commonly used to quantify some chemical parameters in olive oils (Cerretani et al., 2010). Moreover, this calibration model offers enhanced results compared to others regression methods, such as PCR (Principal Component Regression)

(Liang & Kvalheim, 1996) or multiple linear regression (MLR) for quantitative analysis of chemical parameters (Wentzell & Montoto, 2003).

PLS-R calibration models were carried out in order to determine a relationship between predictor variables (absorbances) and the chemical characteristics of olive oils referred above. Specifically, the main goal was to predict the content of phenolic compounds, *ortho*-diphenols, flavonoids and the quantification of antioxidant activity, considering their high predominance in EVOO's and their importance in human health.

The PLS quantification was performed on all frequency regions used for DA.

The quality of the fitting was scrutinised 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). To validate the developed PLS model, Leave-One-Out (LOO) cross-validation method was used. In this technique, one sample at a time is excluded. Then, the removed sample was predicted with a model constructed with the remaining samples. This procedure was repeated until each sample was excluded once (Gurdeniz, Tokatli, & Ozen, 2007).

To illustrate the typical behaviour of the calculated RMSEC and RMSECV, Fig. 3 (supplementary material) represents its values as a function of the latent variable number for total phenolics and *ortho*-diphenols. Fig. 4 (supplementary material) reflects the accuracy and the performance of the models which correlate the actual and estimated values of (a) total phenolics, (b) *ortho*-diphenols, (c) flavonoids and (d) antioxidant activity obtained from FTIR spectra. The plot of the measured concentrations of the chemical components against the predicted values from FTIR measurements reveals the accountability of the models.

Table 4 resumes the performance of the multivariate calibrations in terms of the multiple coefficient of determination ( $R^2$ ), the root mean square error of calibration (RMSEC), and root mean square error of cross validation (RMSECV). The high value of  $R^2$  and the lowest of RMSEC and RMSECV indicate the good performance and precision of PLS model. Furthermore, to obtain a good calibration model, the number of regression factors used should be the lowest as possible (Hui-shan et al., 2006).

Interestingly, only three regression factors stood in phenolic compounds, *ortho*-diphenols and flavonoids prediction models and four regression factors for the antioxidant activity prediction model.

For *ortho*-diphenols and flavonoids, the slope  $R^2$  and the RMSE, indicators of the quality of fit, were higher than 0.98 and lower

**Table 3**  
Confusion matrix for the learning sample and the cross-validation results based on the maturation stage. Observed classifications in the rows. Predicted classifications in columns.

From/To	Learning sample					Cross-validation				
	Green	Semi-ripe	Ripe	Total	% Correct	Green	Semi-ripe	Ripe	Total	% Correct
Green	24	0	0	24	100	17	6	1	24	70.8
Semi-ripe	0	16	0	16	100	6	10	0	16	62.5
Ripe	0	0	24	24	100	2	1	21	24	87.5
Total	24	16	24	64	100	25	17	22	64	73.6

**Table 4**  
Parameters of the multivariate calibrations for quantification of phenolic content, *Ortho*-diphenols, flavonoids and antioxidant activity in extra virgin olive oil.

	Number of factors	$R^2$		RMSE	
		Calibration	Validation	Calibration	Validation
Phenolic content (mg GAE g <sup>-1</sup> )	3	0.94	0.91	0.02 (3.25%)	0.02 (3.25%)
<i>Ortho</i> -diphenols (mg GAE g <sup>-1</sup> )	3	0.99	0.99	0.01 (3.29%)	0.02 (6.58%)
Flavonoids (mg catechin g <sup>-1</sup> )	3	0.99	0.98	0.02 (2.93%)	0.03 (4.40%)
Antioxidant activity (mmol trolox kg <sup>-1</sup> )	4	0.93	0.88	0.04 (1.46%)	0.05 (1.82%)

than 0.03, respectively. The phenolic content showed lowest multiple coefficient of determination in the calibration model ( $R^2 = 0.94$ ), as well as in the cross validation ( $R^2 = 0.91$ ). The RMSE values were lowest and the same for calibration and cross validation (RMSE = 0.02). For the quantification of antioxidant activity, the slope  $R^2$  and the RMSE, were 0.93 and 0.04, respectively, which become worse when data were cross-validated ( $R^2 = 0.86$ , RMSECV = 0.05).

#### 4. Conclusions

According to the results reported, it can be concluded that FTIR–ATR spectroscopy followed by chemometric treatment of the data, namely principal component analysis, discriminant analysis and partial least squares regression, is an appropriate and powerful technique that can be useful in the indirect quantification of chemical parameters of extra virgin olive oils obtained from olives in different ripening stages.

Discriminant analysis allows the classification of EVOO's made from 'Cobrançosa', 'Galega' and 'Picual' cultivars according to their maturation stages using their FTIR spectra. High values of  $R^2$  and low values of RMSEC and RMSECV were obtained for all analytical parameters studied.

Thus, this spectroscopic methodology has high potential to distinguish the ripening stage of olive oils and to quantify simultaneously their chemical composition. This could be useful for the producers to determine the optimal stage for harvesting and also to avoid both fruit loss and possible fungus infections related to late harvesting.

#### Acknowledgements

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

This work is supported by national funds by FCT – Portuguese Foundation for Science and Technology, under the project PEst-OE/AGR/UI4033/2014 and 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.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.foodchem.2014.11.037>.

#### References

Allalout, A., Krichène, D., Methenni, K., Taamalli, A., Daoud, D., & Zarrouk, M. (2011). Behavior of super-intensive spanish and greek olive cultivars grown in northern Tunisia. *Journal of Food Biochemistry*, 35, 27–43.

Ayton, J., Mailer, R. J., Haigh, A., Tronson, D., & Conlan, D. (2007). Quality and oxidative stability of Australian olive oil according to harvest date and irrigation. *Journal of Food Lipids*, 14, 138–156.

Bakhouch, A., Sánchez, J., Beltrán-Debón, R., Joven, J., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2013). Phenolic characterization and geographical classification of commercial Arbequina extra-virgin olive oils produced in southern Catalonia. *Food Research International*, 50, 401–408.

Baldoni, L., & Belaj, A. (2009). In J. Vollmann & I. Rajcan (Eds.), *Oil crops-handbook of plant breeding*. Springer Science & Business Media, 4, 397.

Barros, A., Freire, I., Gonçalves, B., Bacelar, F., Gomes, S., Lopes, J., et al. (2012). Evaluation of chemical and phenotypic changes in Blanqueta, Cobrançosa and Galega during ripening. *CYTA – Journal of Food*. <http://dx.doi.org/10.1080/19476337.2012.705329>.

Barros, A., Nunes, F. M., Gonçalves, B., Bennett, R. N., & Silva, A. P. (2011). Effect of cooking on total vitamin C contents and antioxidant activity of sweet chestnuts (*Castanea sativa* Mill.). *Food Chemistry*, 128, 165–172.

Bendini, A., Cerretani, L., Carrasco-Pancorbo, A., Gómez-Caravaca, A. M., Segura-Carretero, A., Fernández-Gutiérrez, A., et al. (2007). Phenolic molecules in virgin olive oils: a survey of their sensory properties, health effects, antioxidant activity and analytical methods. An overview of the last decade. *Molecules*, 12, 1679–1719.

Bendini, A., Cerretani, L., Di virgilio, F., Belloni, P., Bonoli-Carbognin, M., & Lercher, G. (2007). Preliminary evaluation of the application of the FTIR spectroscopy to control the geographic origin and quality of virgin olive oils. *Journal of Food Quality*, 30, 424–437.

Bouskou, D., Tsimmidou, M., & Blekas, G. (2006). Olive oil composition. In D. Bouskou (Ed.), *Olive oil: Chemistry and technology* (pp. 1–33). The American Oil Chemists' Society.

Cerretani, L., Giuliani, A., Maggio, R. M., Bendini, A., Toschi, T. G., & Cichelli, A. (2010). Rapid FTIR determination of water, phenolics and antioxidant activity of olive oil. *European Journal of Lipid Science and Technology*, 112(10), 1150–1157.

Cicerale, S., Conlan, X. A., Sinclair, A. J., & Keast, R. S. J. (2009). Chemistry and health of olive oil phenolics. *Critical Reviews in Food Science and Nutrition*, 49, 218–236.

Dabbou, S., Gharbi, I., Dabbou, S., Brahmi, F., Nakbi, A., & Hammami, M. (2011). Impact of packaging material and storage time on olive oil quality. *African Journal of Biotechnology*, 10(74), 16937–16947.

De Luca, M., Oliverio, F., Loele, G., & Ragno, G. (2009). Multivariate calibration techniques applied to derivative spectroscopy data for analysis of pharmaceutical mixtures. *Chemometrics and Intelligent Laboratory Systems*, 96, 14–21.

EEC (1991). Characteristics of olive and olive pomace oils and their analytical methods. Regulation EEC/2568/1991. *Official Journal of the European Communities*, L128, 1–82.

Gurdeniz, G., Tokatli, F., & Ozen, B. (2007). Differentiation of mixtures of monovarietal olive oils by mid-infrared spectroscopy and chemometrics. *European Journal of the Lipid Science and Technology*, 109, 1194–1202.

Hui-shan, L. U., Hui-rong, X. U., Yi-bin, Y., Xia-ping, F. U., Hai-yan, Y. U., & Hai-qing, T. (2006). Application Fourier transform near infrared spectrometer in rapid estimation of soluble solids content of intact citrus fruits. *Journal of Zhejiang University Science*, 7, 794–799.

Inarejos-Carcía, A. M., Gómez-Alonso, A., Fregapane, G., & Salvador, M. D. (2013). Evaluation of minor components, sensory characteristics and quality of virgin olive oil by near infrared (NIR) spectroscopy. *Food Research International*, 50, 250–258.

INE (2012). Inquérito à produção de azeite. International Olive Oil, key-figures on the world market for olive oils, Data adopted at the 100th session of the IOOC (Madrid, Spain), 19–23 November 2012.

Lerma-García, M. J., Ramis-Ramos, G., Herrero-Martínez, J. M., & Simó-Alfonso, E. F. (2010). Authentication of extra virgin olive oils by Fourier transform infrared spectroscopy. *Food Chemistry*, 118, 78–83.

Lerma-García, M. J., Simó-Alfonso, E. F., Bendini, A., & Cerretani, L. (2011). Rapid evaluation of oxidised fatty acid concentration in virgin olive oil using Fourier-transform infrared spectroscopy and multiple linear regression. *Food Chemistry*, 124, 679–684.

Liang, Y., & Kvalheim, O. M. (1996). Robust methods for multivariate analysis – A tutorial analysis. *Chemometrics and Intelligent Laboratory Systems*, 32, 1–10.

Machado, M., Felizardo, C., Fernandes-Silva, A. A., Nunes, F. M., & Barros, A. (2013). Polyphenolic compounds, antioxidant activity and L-phenylalanine ammonia-lyase activity during ripening of olive cv. "cobrançosa" under different irrigation regimes. *Food Research International*, 51, 412–421.

Maggio, R., Kaufman, T., Del Carlo, M., Cerretani, L., Bendini, A., Cichelli, A., et al. (2009). Monitoring of fatty acid composition in virgin olive oil by Fourier transformed infrared spectroscopy coupled with partial least squares. *Food Chemistry*, 114, 1549–1554.

Mateos, R., Espartero, J. L., Trujillo, M., Ríos, J. J., León-Camacho, M., & Alcudia, F. (2001). Determination of phenols, flavones, and lignans in virgin olive oils by solid-phase extraction and high-performance liquid chromatography with diode array ultraviolet detection. *Journal of Agriculture and Food Chemistry*, 49(5), 2185–2192.

Nieto, L., Hodaifa, G., & Peña, J. (2010). Changes in phenolic compounds and Rancimat stability of olive oils from varieties of olives at different stages of ripeness. *Journal of the Science of Food and Agriculture*, 90, 2393–2398.

Rohman, A., & Che Man, Y. B. (2012). The chemometrics approach applied to FTIR spectral data for analysis of rice bran oil in extra virgin olive oil. *Chemometrics and Intelligent Laboratory Systems*, 110, 129–134.

Sinelli, N., Casale, M., Di Egidio, V., Oliveri, P., Bassi, D., Tura, D., et al. (2010). Varietal discrimination of extra virgin olive oils by near and mid infrared spectroscopy. *Food Research International*, 43, 2126–2131.

Sinelli, N., Cosio, M. S., Gigliotti, C., & Casiraghi, E. (2007). Preliminary study on application of mid infrared spectroscopy for the evaluation of the virgin olive oil "freshness". *Analytica Chimica Acta*, 598, 128–134.

Uceda, M., & Hermoso, M. (1998). La calidad del aceite de oliva. In D. Barranco, R. Fernández-Escobar, & L. Rallo (Eds.), *El Cultivo del Olivo* (pp. 547–572). Madrid, Spain: Junta de Andalucía Ediciones Mundi-Prensa.

- Wentzell, P. D., & Montoto, L. V. (2003). Comparison of principal components regression and partial least squares regression through generic simulations of complex mixtures. *Chemometrics and Intelligent Laboratory Systems*, 65, 257–279.
- Yang, H., & Irudayaraj, J. (2007). Comparison of near-infrared, Fourier transform-infrared, and Fourier transform-Raman methods for determining olive pomace oil adulteration in extra virgin olive oil. *Journal of American Oil Chemists Society*, 889–895.
- Zhishen, J., Mengcheng, T., & Jiaming, W. (1999). The determination of flavonoid contents in mulberry and their scavenging effects on superoxide radicals. *Food Chemistry*, 64, 555–559.