Combining PLS regression with portable NIR spectroscopy to on-line monitor quality parameters in intact olives for determining optimal harvesting time

Antonio J. Fernández-Espinosa*

Department of Analytical Chemistry, Faculty of Chemistry University of Sevilla, C. Profesor García González 1, Campus of Reina Mercedes, E-41012 Sevilla,Spain

ABSTRACT

This study presents a systematized method for predicting water content, fat content and free acidity in olive fruits by on-line NIR Spectroscopy combined with chemometric techniques (PCA, LDA and PLSR). Three cultivar varieties of Olea europaea – *Hojiblanca cv.*, *Picual cv.* and *Arbequina cv.* – were monitored. Five olive cultivation areas of Southern Spain (Andalucia) and Southern Portugal (Alentejo) were studied in 2011 and 2012. 465 olive samples were collected during the ripening process (non-mature olives) and compared with other 203 samples of mature olives collected at the final ripening stage. NIR spectra were measured directly in the olive fruits in the wavelength region from 1000 to 2300 nm in reflectance mode. The reference analyses were performed on the olive paste by oven drying for the moisture, by mini- Soxhlet extraction for the fat content and by acid titration of the oil extracted from the olive paste. Calibrations and predictive models were developed by Partial Least Square Regression (PLSR) previous Principal Component and Linear Discriminant analyses (PCA and LDA) were employed as exploratory and clean-up tools of data sets. The final models obtained for the total samples showed acceptable statistics of prediction with R^2¼ 0.88, RMSEV% ¼ 4.88 and RMSEP% ¼ 4.98 for water content, R^2¼ 0.76, RMSECV% ¼ 19.5 and RMSEP% ¼ 20.0 for fat content and R^2¼ 0.83, RMSECV% ¼ 36.8 and RMSEP% ¼ 38.8 for free acidity. Regression coefficients were better for only one maturity state (ripe period) than for olive fruit with different composition (ripening period). All models obtained were applied to predict LQPs on a new set of samples with satisfactory results, a good prediction potential of the models.

Keywords: PLS, AOTF-NIR, *Olea europaea* L., Moisture, Fat, Acidity

1. Introduction

The main constituents of the fruit of the olive tree (*Olea euro- paea*) are water and oil. During the ripening period, the oil begins to form with the synthesis of lipids and all the fatty acids con- stitute the acidity of the oil, which comprises mainly oleic acid. Ata certain stage, the lipids lose their capacity to synthesize. This is the optimum moment for harvesting, when the oil contained in the olives is of the highest quality [1].The free acidity of olive oils is one of the key parameters for classifying them into different commercial qualities [2]. Values of water and fat content in olives vary during the ripening process. These variables are determined in the laboratory by standard methods and they determine the beginning of harvesting periodfor olives (decision parameters). The free acidity is determined in olives by titration of the olive oil, which is the more usual method[3,4]. The fat content is determined by Soxhlet extraction [3,5], butalso by other techniques such as by NMR [5,6,7]. The water contentor moisture is normally determined by drying oven [3,6,7] and Karl Fisher titration [8].

Determination of other quality parameters in olive oils, such asperoxide value, specific extinctions at 232, 270, 225 nm, etc., areusually carried out according to the analytical methods describedin Regulation EEC/2568/91 of the European Union Commission [9]or using others, such as the AOAC, IUPAC and AOCS methods [3,10].The values usually take one week to be obtained and the de- cision to harvest can therefore be delayed. Predicting these para-meters, among others, in intact olives is very useful in the olive processing industry.

The values can be obtained by automated means directly in the fruit on the tree with the use of a portable instrument, whose usefulness justifies the search for an effective analytical metho- dology. A Near-Infrared Spectrometer with an appropriately se- lected sensor and conveniently optimized experimental para- meters constitutes the aim of the new instrumental methodology. Additionally, the combination of this instrument with multivariate statistics constitutes an objective of the present investigation. NIRSoffers the advantages of simplicity, noninvasive and rapid analysis,and low cost [11].

It is well known that NIR Spectroscopy is based on the ab-sorption of electromagnetic radiation between 750 and 2500 nm (13,330 to 4000 cm^{-1}) and it provides qualitative information on composition of samples measured. However at short wavelengths (1600 nm) signals are less intense with poor spectral resolution, while signals at longer wavelength are more intense and require short path length for better resolutions [12]. In addition, because absorption bands are usually wide and overlapping, it cannot perform a correct quantitative analysis. Instead, multivariate sta- tistics can be used to correlate spectral data with analytical data, allowing quantitative results of samples and eliminating the need of physical or chemical treatments.

The application of near infrared spectroscopy for determining parameters in fruits was already employed in the last decade [13]. In intact fruit, NIRS has been applied to measure fruit quality parameters in apple [14,15], citrus fruits [16], cherries [17] and melons [18]. Determination of quality parameters in fresh and intact olives by spectroscopic techniques is less frequent than other intact fruits: the parameters fat content, oil content and moisture in olive paste have been measured by NIR Spectroscopy [19]. Also, studies on quality of the most similar fruit – wild olives, *Olea europaea* subsp. *europaea var. Sylvestris* – are practically non-existent [20].

On the other hand, ATR-MIR (4000–400 cm^{-1}) is also a suitable method for the fast quantitative analysis of natural products, al- though, in general, NIR demonstrated advantages over ATR-IR spectroscopy. Overtones and combinations can be found in the NIRregion containing a manifold of information compared to MIR [21,22], although NIR-intensities are 10–1000 times lower than forthe MIR [23]. Besides, ATR-MIR has disadvantages for in- homogeneous natural products because of the low penetration depth and the small scanning area [24]. Therefore NIR showed better suitability for many applications in plant production.

The present paper provides an improved application based on the use of portable analytical instruments and selected rapid sensors combined with chemometric tools in order to on-line control the composition of olives throughout the ripening period in agricultural practices. The present research therefore attempts to optimize mathematical models from NIR spectra to predict decisive quality parameters in olives in order to establish the ideal harvesting time to be applied to a wide range of olive varieties, different geoclimatic zones and different maturity indexes.

2. Material and methods

2.1. Olive varieties. Characteristics

Approximately 262 varieties of olives are grown on the IberianPeninsula, only 24 of which are used regularly for oil production[25]. Of all the varieties produced, the most important and re-presentative are *Arbequina, Cornicabra*, *Empeltre, Hojiblanca*,*Lechín*, *Picual*, *Picudo* and *Verdial*. Andalucía, which accounts forover 75% of Spanish production, cultivates the biggest amount ofolives in the world. The remaining 25% of Spanish production is distributed throughout different areas of the Mediterranean Basin. In the present research, *Arbequina*, *Hojiblanca* and *Picual* were the three differentiated varieties of olives studied by means of aportable NIR Spectrometer. In the *Arbequina* variety, the olive issmall, around 2 g, but the tree produces a large amount of fruitwith a relatively high oil yield: 20.5%. The oil produced has a high content of palmitic and linoleic acid [26,27].

In the *Hojiblanca* variety, the olive is big, over 4.8 g, and has a relatively low oil yield: 17–19%. The oil has a high content of oleic acid and linoleic acid. In the *Picual* variety, the olive fruit is usually medium to large insize, weighing from 2 to 4 g, oil productivity is high, with an oil yield of 27%. It contains around 80% of oleic acid and 4% of linoleicacid. As for oil formation, in the *Picual* and *Hojiblanca* cultivar vari- eties this phase coincides with the coloring of fruits; some of these are black, and some are green [28]. The *Arbequina* variety presentsa different behavior pattern, and is the first to be harvested when the olives become green.

2.2. Design of the experiment. Sampling campaigns

In an attempt to validate the complete method, we designed a systematic experiment which we divided into three sampling campaigns. The main objective involved establishing good corre- lations between the spectral and the analytical data. The objectives of the campaigns were different: to differentiate between mature and non-mature fruits and to differentiate among varieties. The final mathematical prediction models should show a high quality for this range of conditions.

In sampling *Campaign 1* we selected olives of the *Picual cv.*, in order to validate the on-line instrumental measurements. In *Campaign 2* we selected olives of the *Hojiblanca*, *Picual* and *Arbe- quina* cultivar varieties to optimize the instrumental parameters and to obtain the calibration equations for olives at the mature stage (Model 1). In *Campaign 3* we selected olives of the three varieties to obtain the calibration equations for olives in any state of maturation (Model 2). Finally, we applied the calibration equations for all three campaigns (Model 3) to olives collectedduring the ripening period (Campaign 3). This enabled us to select the best model of prediction for a wide range of conditions, de- pending on the number of samples collected.

2.2.1. On-line measurements: evaluation of suitability. Campaign 1

In *Campaign 1* the objective involved estimating the maximumtime that olives are not subjected to non-biological degradationafter they have been picked. Thus, the experiment consisted ofcomparing the spectral and analytical measurements at differentdelay times after harvesting (reference values). References are thespectra captured on-line in the trees just before picking, as well asthe laboratory analyses initiated 20 min prior to picking (Series 1).

To this end, olives of the *Picual cv.* were collected in the botanical park "Parque de María Luisa" of the city of Seville (37°22′ 27.8″N 5°59′8.6″W, http://goo.gl/maps/0Bekh), close to our la- boratory, in January 2011 (one zone, one variety). The experiment was divided into six series of measurements of different delay times, and six olive samples were collected:

– *Series 1* (*on-line spectra and rapid analyses, the references*): a series of 5 spectra were captured individually in one sample comprising 50 olives and averaged for 50 5 250 scans. During acquisition of the spectra, the olives were always on the tree prior to picking (intact olives). Olives were selected from different trees and different zones (sun-shade, external-internal), and branches and leaves were ink-marked. Chemical analyses were initiated.

– *Series 2* (*in-line spectra and rapid analyses*): immediately after Series 1, a second series of spectra was captured in another 50 olives (2nd sample), just after they were picked from the same trees and ink-marked zones. We captured the five spectra on the ground under the tree (*in-line*). Chemical analyses were also in- itiated before 20 min have passed after harvesting.

– *Series 3* (*spectra and analyses in the laboratory*): after Series 2,a third series of spectra was captured in 50 similar olives (3rd sample). This sample was transported to the laboratory at 3–4 °C in a portable refrigerator. Spectra were captured in the laboratory

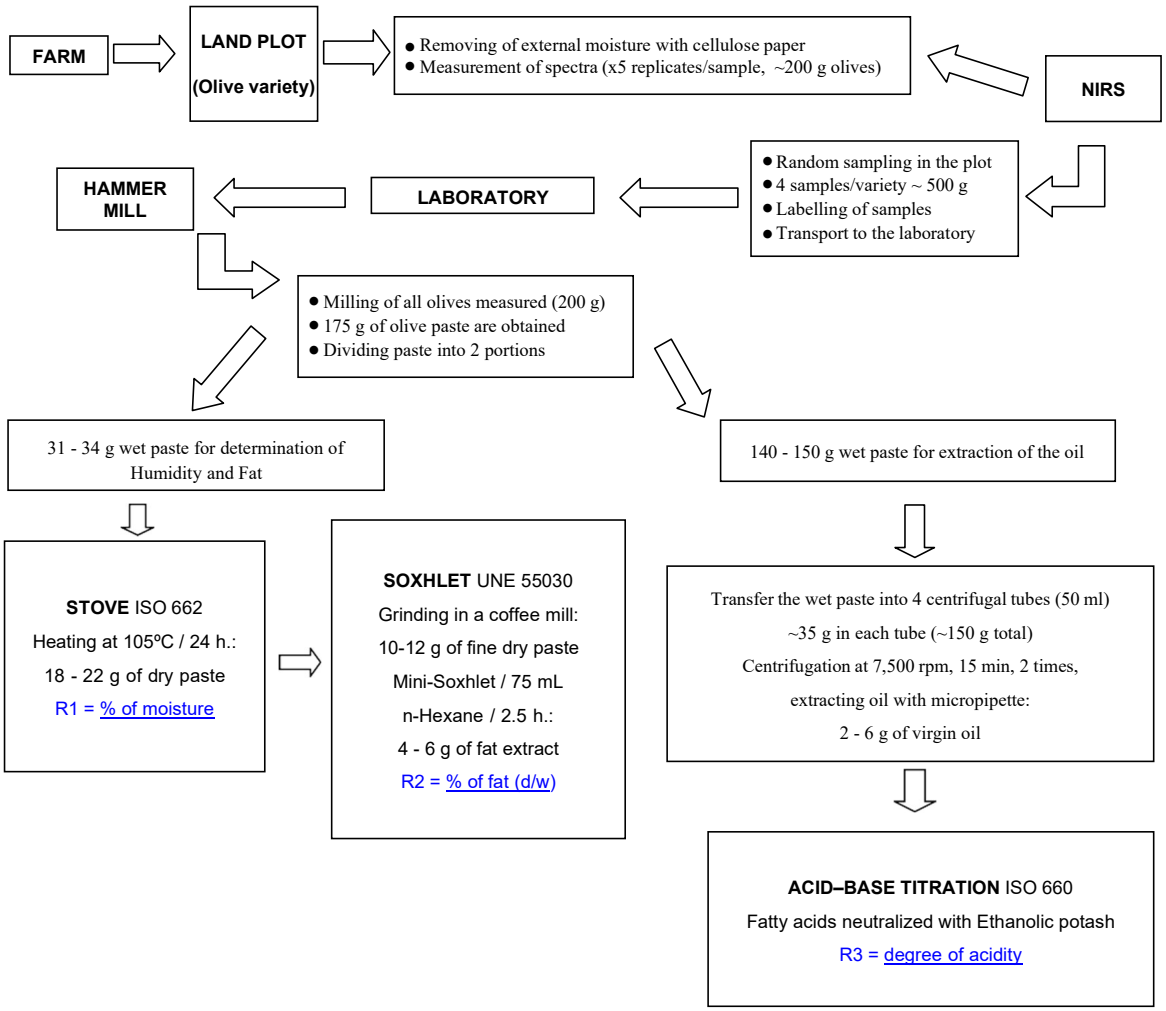


Fig. 1. Complete analytical methodology applied to determine the quality parameters of olives and oils.

4 h after picking and chemical analyses were immediately initiated.

– Finally, *Series 4–6* (*delayed spectra and analyses in the la-boratory*): Another three series of 50 similar olives were cold- stored at 3–4 °C during transport and in the laboratory re- frigerator. For these series (4th, 5th and 6th sample), spectra werecaptured in the laboratory at 24, 48 h and one week after picking, and chemical analyses were immediately initiated.

This experiment, based on six series of measurements, was replicated 5 times (5 ~ 6 ¼ 30 samples) in order to corroborate the results and their reproducibility. To detect significant differences (p o0.05) between series for spectral and analytical data, we performed an ANOVA using STATISTICA® Version 8 software, from StatSoft Inc., Tulsa, OK, USA.

Tukey's HSD post-hoc test was used for comparison of several means, such as the most significant spectral bands or analytical parameters. Additionally, a paired t-test was applied to each batch to detect significant differences ($t$calc 4 $t$crit) for comparison be- tween the means values of predicted and measured data.

2.2.2. Calibration equations of the ripe period: Campaign 2

The second campaign was conducted from February 2011 (week no. 5) to March 2011 (week no. 12), and 203 samples of olives of the three varieties were collected from different olive farms. The olive farms were always located close to an agro-cli- matic station belonging the Andalusia government (EAMs). The objective of the campaign was to obtain the best calibration equations for the three quality parameters – moisture, fat and acidity – when all olive samples collected were at the mature stage(Model 1).

Five sampling zones were selected for *Campaign 2*:

1. *Trasmulas Farm* in Pinos Puente (Province of Granada) cultivates the *Picual* variety. (37°11′35.9″N 3°51′45.0″W/https://goo.gl/maps/FFrJA).
2. *Caserio del Conde Farm* in Mollina, the area of Antequera (Pro- vince of Malaga) cultivates the *Hojiblanca* variety. (37°11′07.4″N 4°37′36.9″W/https://goo.gl/maps/p4Z4s).
3. *El Cañaveral Farm* in Adamuz, the area of Córdoba province cultivates the *Arbequina* and *Picual* varieties. (38°00′12.8″N 4°26′07.9″W/https://goo.gl/maps/2RYs1).
4. *La Barranca Farm* in Palma del Rio (Córdoba) cultivates the three varieties. (37°41′12.4″N 5°22′40.5″W/https://goo.gl/maps/ lYuEL).
5. *Cegonha Farm* in Vidigueira, in the Portuguese area of Alentejo (Province of Beja) cultivates also the three varieties of olives. (38°07′13.4″N 7°46′31.9″W/https://goo.gl/maps/76PDw).

2.2.3. Calibration equations of the ripening period: Campaign 3

Campaign 3 was conducted from 19 September 2011 (week no. 38) to 10 February 2012 (week no. 06), when 465 samples of the three varieties were collected weekly from the five sampling zones. In this case, both ripe and unripe fruits were collected (4 week no 5), because the olives are still immature from September to January.The objective of this campaign was to obtain the best calibra- tion equations for the three quality parameters in the case of olivesat different stages of maturity (Model 2). We will also verify that Model 2 (different maturity stages) predicts the analytical valuesof the olives in Campaign 2 (only mature fruits, external validation).

2.2.4. Final validation. Calibration equations for the complete period: Campaigns 2 + 3

The final validation consisted of testing models based both on Campaigns 2 and 3, to obtain the best calibration equations for both mature and non-mature olives (Model 3). In this case a total of 668 samples was processed. The samples represented the per- iod from January 2011 to February 2012, for the three olive vari- eties and for the five sampling zones. This enabled us to compare and to establish the best models of prediction for a wide range of conditions, such as different varieties, zones and maturation grade.

2.3. Reference methods. Olive quality parameters

The reference quality parameters (QP) determined for the olivesamples were analyzed in the laboratory (LQP) in accordance withofficial methods. The methods used and the complete analytical methodology are shown in the flow diagram of Fig. 1. We determined water content (humidity or moisture) accord- ing to international standard ISO662 [29] by heating 31–34 g of wet olive paste in a stove. Fat content was determined following Spanish standard UNE55030 [30] by means of extraction of fat from 10 to 12 g of dry olive paste with n-hexane in a mini-Soxhlet 4-battery extractor (Selecta, JP Selecta, Barcelona, Spain). We determined the free acidity of the olive oil, according to interna- tional standard ISO660 [31] by titration of 1–3 g of oil with etha- nolic potash in a 5 mL-microburette. We obtained wet olive paste by crushing 200 g of olives with a laboratory hammer mill (Pieralisi FP HP-15, Jesi, Italy) at 1000 rpm.The olive oil was obtained by centrifugation of 140–150 g of wet olive paste, at 7500 rpm for 15 min in a 3–15 tabletop Sigma centrifuge (Osterode am Harz, Germany). Efficiency of oil extrac- tion was 5–15 mL of oil per 100 g of wet olive paste.

2.4. NIR spectrometer. Photodiode detectors, instrument parameters and performance

On-line acquisition of spectra was performed with the use of a portable Miniature AOTF-NIR Spectrometer (Brimrose Solid-state Luminar 5030, Brimrose Corp., Maryland, USA) equipped with a Reflectance post-dispersive optical configuration. The main ad- vantage of AOTF (Acousto-Optic Tunable Filter) technology in- volves its fast scanning speed (16,000 wavelengths per second) and a high spectral resolution (2–10 nm), which is higher than those obtained with other technologies [32,33]. The additional advantage of the Hand-held Luminar 5030 is that it is more por- table for travel and field measurements.

After careful selection of detectors, the spectrometer acquires the Dual Beam IR signal by employing a detector based on indium–gallium–arsenide (InGaAs) photodiodes (PDs). A photodetector is a solid-state sensor that converts light energy into electrical energy. High-speed photodetectors are required for fiber-optic commu- nications (800–1600 nm), spectroscopy (400–6000 nm) and pho- ton counting (400–1800 nm), and the InGaAs detector presents this main advantage. Furthermore, InGaAs PD it appears to be the best detector material for high-temperature operation in the 1000–3000 nm spectrum, in comparison with other detectors studied (InSb, InAs, PbS, HgCdTe). Additionally, InGaAs and silicon PDs are more sensitive than germanium PDs due to their lower dark current and lower multiplication noise [34]. Other types of detectors studied, such as Diode Array Detector (DAD) provide a low spectral resolution.

We measured spectra by Diffusive Reflectance (R), at intervals (wavelength increment) of 1 nm, performing 5 scans on each of the 50 olive fruits. According to the first assays, we considered it appropriate to perform 5 scans along the circumference of each olive fruit for good representativeness of the composition. Using 1 nm rather than 2 nm increases the measurement time of spectra to 30 s per 5 individual spectra. However, measuring at 1 nm, both spectrum resolution and correlation coefficients show an increase. Spectra were measured directly in the intact olive fruit in a non-destructive manner at a sampling distance of 6–10 mm from the olives along their diameter. The portable spectrometer uses two batteries, which enabled over 8 h of autonomy in the field, along with a netbook computer that controls the measurement unit via Ethernet interface. The instrument is controlled by *SNAP-32!* Brimrose Analytical Software.

2.5. Methodology of chemometric calibrations. PCA, LDA and PLSR

In the chemometric calibrations, we determined the relation- ships between the instrumental and the laboratory measurements, and employed a set of numerous samples: 203 samples of mature olives and 465 samples of olives at different stages of maturity. Thus, these data were processed with Unscrambler® v9.5 software (CAMO-Computer Aided Modeling PROCESS AS, Trondheim, Norway). Partial Least Squares Regression (PLSR) was the main tool applied to the full spectrum (1100–2300 nm, 1201 variables). The quality of the PLS model was tested by means of two types of validation: cross-validation and test-set validation. In test-set va- lidation two different sets of samples were used, and the cross- validation uses the same set of samples for calibration and vali- dation. A principal component analysis (PCA) and a linear dis- criminant analysis (LDA) were previously employed as un- supervised exploratory analyses of the data structure and as classification techniques.

PLRS is a two-block linear regression method based on esti- mated latent variables; we applied this to the two datasets (spectra and analytical data) referring to the same objects [35] to determine the optimal prediction. Moreover, light scattering af- fects path length and spectral corrections must be applied: several pretreatments of the spectra were investigated. Multiplicative signal correction (MSC) was tested for variations in light scatter- ing. First and second derivatives were calculated because they are useful for extracting band-shift and band-shape features and in eliminating baseline effects. The influence of smoothing using Savitzky–Golay (SG) and baseline correction was also tested. Fi- nally, prior to calibration, we used the Savitzky–Golay (SG) method first and second derivatives to smooth the noisy spectra of Reflectance. The SG method is a least-square fit of a third-order polynomial with equidistant spacing. The smoothed data are the moving averages of the data weighted by the Savitzky–Golay coefficients. The raw absorbance spectra, processed as the $\log[1/R]$ (R reflectance), were also subjected to SG filtering.

PLS was focused toward a calibration procedure based on the study of certain parameters such as the SEC (standard error of calibration) or the SECV (Standard Error of Cross-Validation) to extract information from the system, such as the optimum number of factors of the Principal Component Analysis (PCA) to be in-cluded in the calibration models.

The Principal Component Analysis (PCA) is a multivariate che- mometric method which is optimal for handling co-linearity and as such, PCA and NIR spectroscopic data provide a perfect match. The PCA is superior for handling the highly co-linear data often found in spectroscopy, and constitutes an excellent tool for ex-ploratory data analysis.

With PCA, the previous qualitative analysis of the olive samples $SECV = \sqrt{\frac{\sum_{i=1}^{N}(C_i - C_i')^2}{N - 1}}$ dimensionally reduced the number of variables and transformed the wavelengths (original variables) into new axes (principal components). The PCA breaks the spectral data down into a structure component and a noise component. Thus, the PCA en- abled us to recognize trends, highly correlated groups of variables and outliers. When it found spectra that differed from the rest of the population (outliers) these were excluded and removed forequation development and testing. The first principal component (PC1) is a new coordinate axis representing the direction of maximum variation through the data. After PC1, the following PCs lie along a direction orthogonal to PC1. The PCs are uncorrelated with each other (orthogonal). The PCA was performed with a Varimax rotation, extracting the factors or principal components (PCs). Each selected factor should have an eigenvalue over 1 or 2, explaining at least 1–5% of the total variance of the dataset, and together they should account for over 70–75% of the accumulated variance.

In relation to the PCA, a classification model based on LDA was proposed with the aim of establishing whether the three olive cultivar varieties could be discriminated prior to PLSR application.When there is a high correlation among the original spectral data, direct application of LDA is not possible. For this reason, a PCA is necessary as a variable reduction tool before application of the LDA. The main objective of the LDA consists of constructing a base of vectors providing the finest discrimination among the classes, inan attempt to exploit the between-class differences and to mini- mize the within-class differences by using spread matrices. Re- garding the quantitative aspect PLS can be seen as two inter- connected PCA analyses.

The study of the best calibration models was based on a few statistical parameters. In PLSR the standard error of calibration (SEC) is the standard deviation for the residuals resulting from differences between the laboratory-measured values and the NIR- predicted values for samples within the calibration set [15,36], andmathematically it is calculated as the root mean square (RMS) of the residuals

$$SEC = \sqrt{\frac{\sum_{i=1}^{N}(C_i - C_i')^2}{N - 1 - p}}$$

where C_i is the predicted concentration in the sample and C_i' is the measured value of the concentration in the sample, for N samples and p independent variables of the validation set in Eq. (1).

Another parameter, the Standard Error of Prediction (SEP) was also assessed [37,38]. SEP is the standard deviation for the re- siduals resulting from differences between the laboratory-mea- sured values and the NIR-predicted values for M samples outside the calibration set (prediction set) based upon a specific calibra- tion equation, Eq. (2):

$$\sqrt{\frac{\sum_{i=1}^{M}(C_i - C_i')^2}{M}}$$

In addition, the residual predictive deviation (RPD¼SD/SECV), defined as the ratio between the standard deviation (SD) and the SECV for the NIR predictions, is a useful statistic often applied to evaluate how well a calibration model can predict chemical data [15,39]. An RPD value greater than three (range 3.1–4.9) is con- sidered to be fair and recommended for screening purposes, andan RPD value greater than 5 (range 5–6.4) is considered to be goodfor quality control.

The best regression equations for an olive quality parameter were those presenting a maximum value of coefficients of multiple determination (R^2) in the calibration set and minimum values of SEC (I), SEP (II) and SECV (III).

In the assessment of the models, we divided the overall set of samples into groups, using one of them to construct the calibrationmodel (cross-validation or calibration set) and the remaining ones for testing the model (prediction or validation set). The criteria for selection of samples for the calibration step covered a wide range of free acidity levels, from low to high fat and moisture values and for the three varieties, in order to improve the predictive capacity of the model. For the validation step all samples were selected randomly. We selected the distribution of samples in the calibra- tion and validation sets using the Kennard–Stone algorithm [40].

3. Results and discussion

3.1. Campaign 1: statistical suitability of on-line measurements

When the ANOVA is performed on Campaign 1, it indicated that there were no significant differences among Series 2–4, on one hand, and Series 1 (the reference), on the other, for both the spectra and the quality parameters: The values of sum of squares are low and the values of F_{exp} are lower than the values of F_{tab}.From Series 2 to 4, values of the sum of squares between and within groups become higher, until they reach the major difference in Series 5 and 6 ($F_{exp}4F_{tab}$). This fact suggests a major degree of degradation of olives over time for more than 24 h,therefore spectra of olives can be captured directly in the field, aswell as in the laboratory (N.B. before 24 h). Similarly, analyticaldeterminations can be performed just after picking or before 24 h.Despite that NIRS measurements are more easily made in the laboratory than in the field, all spectra were finally captured di- rectly in the trees during the early hours of the morning and la- boratory determinations were initiated before the following 5 h. No temperature effects there are prior to measurement, becauseonly from midday until afternoon ambient temperatures are high in Andalucía.

3.2. Campaigns 2 and 3: spectral characteristics and varieties

The NIR Reflectance spectra of 203 olive samples (Campaign 2) of the three varieties are accumulated in Fig. 2a. Following spectroscopic transformation to Absorbance as log[1/*R*], bands ofspectra were better observed (Fig. 2b): spectra shows a strong IR.

The accuracy of the equations for NIRS predictions was assessed by using the Standard Error of Cross-Validation (SECV) as a method for determining the 'best' number of independent vari- ables to use in building a calibration equation to avoid over-fitting. The SECV is an estimate of the SEP and is calculated as the square root of the mean square of the residuals for $N-1$ degrees of freedom, where the residual equals the actual value minus thepredicted value in Eq. (3) band close to 1460 nm, which is characteristic of the water con- tent, and another band near 1920 nm also characteristic of thepresence of water, corresponding to the harmonic and combina- tion bands, respectively, of O–H bonds in hydroxyl groups.

Regarding oil absorptions, bands can be observed at 1210 nm and 1245 nm (second overtones of C–H stretching vibrations –alkyl groups and alkenes), 1145 nm and 1160 nm (stretching of theC–O bonds of aliphatic esters), and 1175 nm, 1185 nm, 1260 nm.

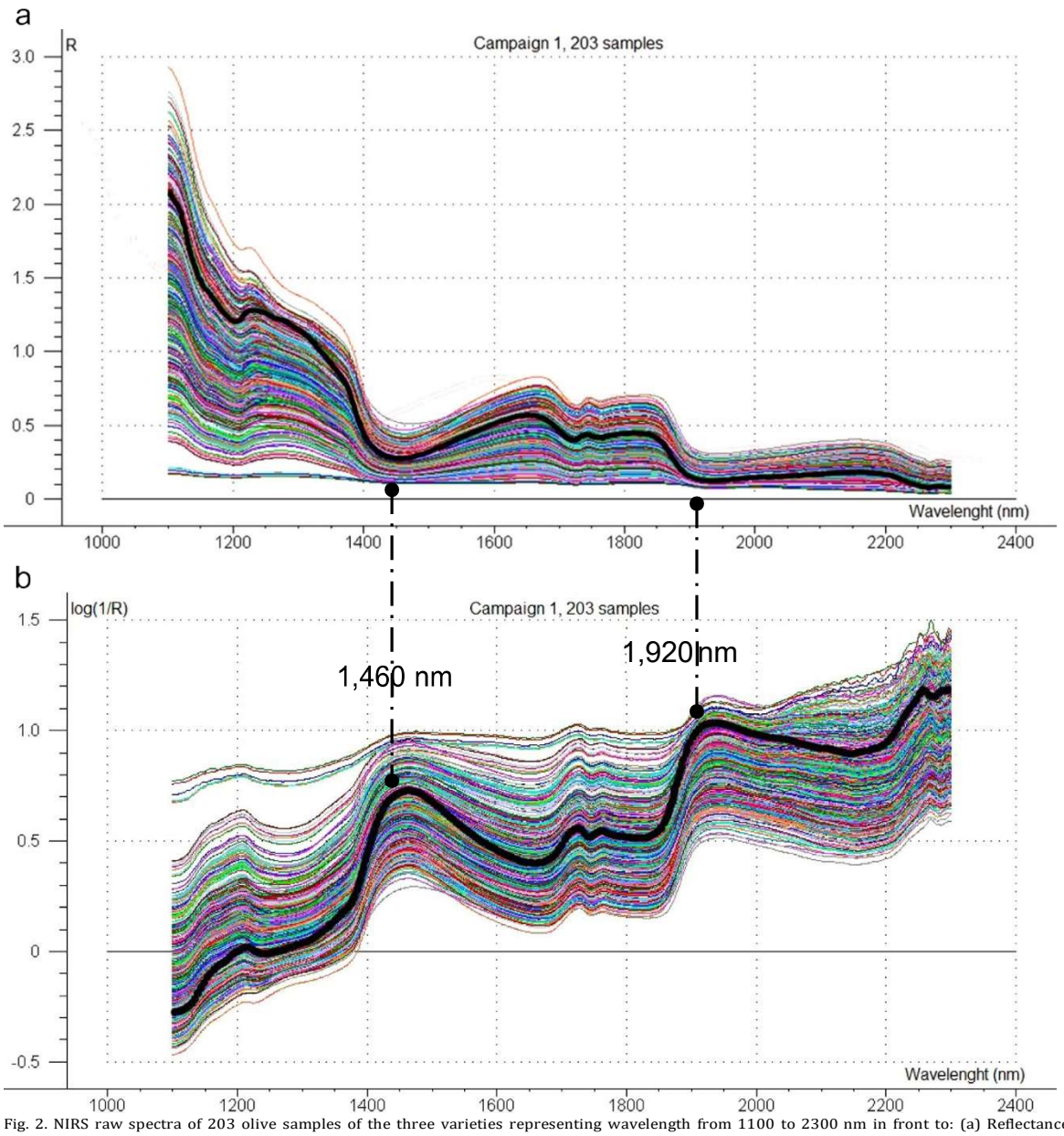


Fig. 2. NIRS raw spectra of 203 olive samples of the three varieties representing wavelength from 1100 to 2300 nm in front to: (a) Reflectance, *R*; (b) absorbance, log(1/*R*). Bold black line is the averaged spectra.

and 1275 nm (alkyl groups and alkenes) [41] (Fig. 3a). Other bandscorresponding to triglycerides are at around 1745 nm (carbonyl ofthe ester-linkage) and at 1693 nm (minor bands); others are at 1662 nm (cis-olefins), 1728 nm (oleic acid) and 1763 nm and 1819 nm (first overtone of C–H stretching vibrations of methyl, methylene and ethylene groups) (Fig. 3b). The NIR Reflectancespectra of Campaign 3 (465 olives samples) were similar to the spectra of Campaign 2 (similar figures not shown).

No differences were graphically observed among the three different varieties in the Absorbance spectra. When the ANOVA was applied to the spectral data both in Campaigns 2 and 3, the values of sum of squares and F_{exp} showed that no significant dif- ferences were found among varieties.

3.3. Campaigns 2 and 3: multivariate statistics of spectral data

Prior to PLSR, a PCA and an LDA were applied to the spectral data. The PCA studied the structure of the data in Absorbance (*log [1/R]*) mode and after applying to them the Savitzky–Golay (SG). In the PCA results (Fig. 4a, Campaign 3) the PC1 appears to have differentiated the wavelength range 1890–2300 nm of water content from the range 1100–1375 nm of aliphatic esters, alkyl groups and alkenes of PC2, both negatively correlated. The range 1890–2300 nm in PC1 was more significant for the *Arbequina* variety, but mixed with *Picual cv*. and *Hojiblanca cv*. in a proportionof 3:2:1 (Table 1). The range 1100–1375 nm in PC2 was moresignificant for the *Hojiblanca cv*. although mixed with *Picual* and *Arbequina cv*. at 9:4:4 and the range 1860–1900 nm of water content in PC3 for the *Arbequina cv* mixed with *Hojiblanca* and *Picual cv*. at 4:3:2. Thus, the structure found – similar for Cam- paigns 2 and 3 – was formed by one swarm of variables/cases withsome extreme outliers (Fig. 4b and c), and any wavelength region was therefore significantly important for a specific variety of olive samples.

For the discriminant analysis, if we keep in mind the high correlations found in the PCA between the wavelengths in, theLDA was performed for possible discrimination of cases taking the matrix of scores as a new matrix of data [42]. The LDA was applied to the Reflectance and Absorbance modes and similarly, no dis- crimination of samples by varieties was found for the complete wavelength range. Other authors find it difficult to discriminate among olive varieties or among pure olive oils [43], although they have greater success discriminating between oils and non-oil samples. Consequently, these results will enable us to apply the equations of PLS regression models to the three varieties of olives simultaneously, thus preventing the need to find one model for each cultivar variety.

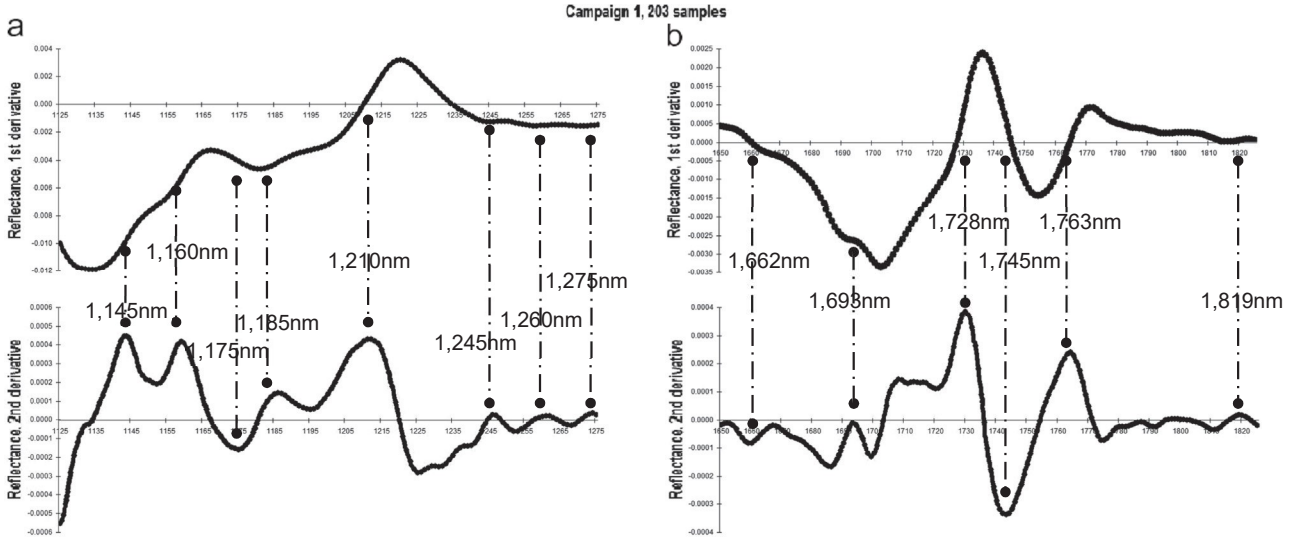


Fig. 3. NIRS averaged spectra of 203 olive samples of the three varieties representing the first derivative and the second derivative of *R*: (a) Wavelengths from 1125 to 1275 nm and (b) wavelengths from 1650 to 1825 nm.

3.4. Campaigns 2 and 3: multivariate statistics of analytical data

The analytical parameters habitually employed to decide the optimum time of harvesting (decision parameters) are water content (moisture), fat content and free acidity (oleic acid) (Ta-ble 2). However, other standard parameters for olive oils (quality parameters) were determined in the oil extracted from olive paste,which were analyzed in half of the oil samples (164 samples in Campaign 3). Thus, a PCA and an LDA were applied as exploratory tools previous to PLSR to the set of variables of Table 2, including olive variety (VAR) and geo-climatic zone (GEO).

The PCA showed a clear differentiation for *Hojiblanca cv.* and also for *Arbequina cv.* in two of the PCs (Fig. 5a, Table 3). The *Ho- jiblanca cv.* was the one showing the greatest sensitivity to changesin water and fat content during the ripening process (PC1, 45% global variance). Many quality parameters of the oils were groupedtogether in PC2 (36%), indicating the advanced oxidation state of mature olives and showing the most representative values for the *Arbequina* and *Hojiblanca* varieties. The most discriminating LQPs for the variety (VAR) are the specific extinctions at 232 and 225 nm (PC3, 19%), indicating the initial oxidation state of non- mature and bitter olives. Other authors applied a PCA to three cultivar varieties of extra virgin oils also through analytical vari- ables, obtaining that the "Leccino" cv. was separated from the other two varieties and, on the contrary, some overlap was ob- served between the "Casaliva" and "Frantoio" cv. [44].

LDA results demonstrated more clearly how through LQ para- meters the three cultivar varieties are discriminated better than through wavelengths (Fig. 5b vs. Fig. 4c). Thus, exploratory studieshave confirmed at this moment that the on-line monitoring sys- tem is prepared for PLSR performance, which will predict LQPvalues by calibrations from wavelengths.

3.5. Campaigns 2 and 3: evaluation of LQP parameters

Regarding Campaign 2, the values obtained for moisture and fatwere in accordance with normal levels in mature olives [9]. Values of free acidity in both campaigns were very low and agree with virgin olive oils – less than 0.8°, excellent quality oils –. In relation to the standard deviations, values for Campaign 3 were higher than those recorded in Campaign 2, because the olives in Cam- paign 2 were collected only at a mature stage. The olives in Campaign 3 were collected during the ripening period, which gave rise to significant differences between both campaigns. Besides, the coefficient of variation of free acidity was higher than those of moisture and fat content, suggesting that free acidity depends more upon the changing stage of biological ripening than fat and moisture. Indeed, the number of replicates (N¼ 356) was lower than for moisture and fat (N¼465), owing to the fact that the oil cannot be extracted from the olive fruit until the 45th-47th week (mid-November).

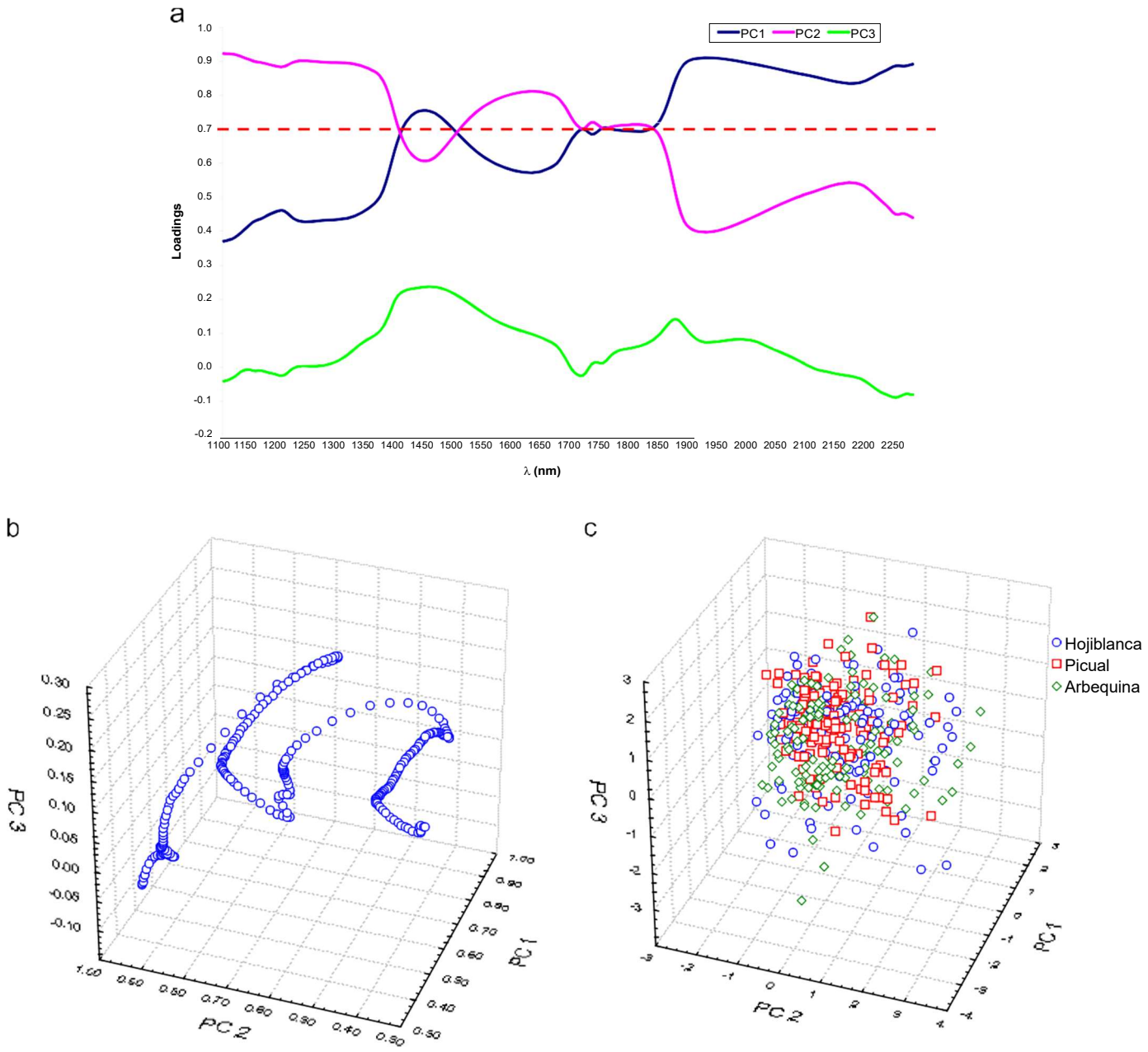


Fig. 4. Distribution of the three PCs. (a) Distribution in the plane of variables (wavelengths), (b) distribution in the space of variables, and (c) distribution in the space of cases (cultivar varieties).

Inter-variety differences evaluated in Section 3.2 in relation to spectral data showed that no significant difference was found in Campaigns 2 and 3. However, the ANOVA performed on the ana- lytical data shows significant differences between *Picual cv.* and the others, and no differences were found between *Arbequina* and *Hojiblanca cv.* (Table 2). Specifically, the Tukey's HSD post-hoc test revealed that there were no significant differences among varieties for moisture (p ¼ 0.305) and fat (p ¼ 0.413) in Campaign 2 and Campaign 3 (p ¼ 0.231 and p ¼ 0.184). However, a significant dif-ference was obtained for free acidity in Campaign 2 (p ¼ 0.006) and also in Campaign 3 (p 0.012). Consequently, discrimination inrelation to spectral data must be achieved with more advanced statistical tools.

3.6. Campaigns 2 and 3: chemometric calibrations, internalcalibrations

In Campaign 2, analysis with PLS of the regression coefficients showed that it was impossible to identify any specific interval of significantly different wavelengths, hence statistics were per- formed on the full spectral range. The most favorable statistics in the PLSR were obtained following Savitzky–Golay (SG) smoothing for the second derivative. Some authors use first or second deri- vative for one or another type of analysis [38,45]. On observing the statistics of calibration set (Table 4), the best calibration results were achieved for the determination of moisture and free acidity. The RPD values for moisture and free acidity were higher than three (5.01 and 6.17), demonstrating the robustness of the two calibration models [46,47]. The RPD value for the determination of fat content was 2.84, thus indicating that this model (Model 1) was almost robust. Additionally, for moisture and fat low values of SECV (o1) and *RMSECV* (1.33%, 5.14%) were obtained, and a high R_{cal}^2 coefficient of 0.95. For free acidity *RMSECV* was 11.8% and a high R_{cal}^2 of 0.96.

The calibration models were tested with the validation set (Table 4, internal validation). In this case, the best results were achieved for free acidity, with R_{val}^2 40.95 and SEP o0.1. The performance of moisture and fat was acceptable with good R_{val}^2 (0.95 and 0.94) and low SEP (0.7 and 1.2). Fig. 6a–c shows the correla- tions found between the values of reference analysis and the NIRS predicted values in Campaign 2. Good correlations demonstrate that NIRS can be used to predict the decision parameters in mature olives of the three varieties.

In Campaign 3 statistics were also applied to the full spectral range. The global results (Table 1) show wide ranges of dispersion in the parameters measured, due to the new factor introduced – ripening. Consequently, our objective involved performing cali- brations for the entire working range, including different varieties, areas and maturity state. The statistics of calibration set (Table 4) show that the best calibration results were also achieved for the determination of moisture and free acidity. The RPD values for moisture and free acidity were higher than three (3.29 and 5.63 respectively).

Additionally, a low value of SECV o2 was obtained for moisture, with a low *RMSECV* value of 2.76%, and a high R_{cal}^2 coefficient of 0.90, and an *RMSECV* value of 9.12% and an R_{cal}^2 of 0.97 for free acidity. In the case of fat content values were also good (R_{cal}^2 of 0.90 and *RMSECV* of 15.3%). However, its RPD value was only 1.79, which indicates that the calibration model for fat was not very robust. In

Table 1
Result of the PCA from matrix comprising 1201 variables (wavelengths) and 420 ~ 5 ¼ 2100 cases (campaign 3).

PC	Eigenvalue	%Variance/Cum./ Global	Variables (wavelengths)	Interpretation	Cases (varieties)
PC1	217.8	91.1 / 91.1 / 91.6	λ_{1890} to λ_{2300} (0.88–0.91), Máx. 1930 nm λ_{1400} to λ_{1500} (0.68–0.75), Máx. 1450 nm	Water content –OH bonds (combination and 2nd overtone) Alkenes C–H bonds (2nd overtone)	*Arbequina cv.* (*51%*) *Picual cv.* (*33%*) *Hojiblanca cv.* (*16%*)
PC2	17.4	7.3 / 98.4 / 7.3	λ_{1100} to λ_{1375} (0.88–0.94), Máx. 1135 nm Máx. 1240 nm λ_{1525} to λ_{1700} (0.74–0.84), Máx. 1650 nm	Aliphatic esters C–O bonds Alkyl groups and alkenes C–H bonds (2nd overtone) cis-olefins C¼C bonds (1st overtone)	*Hojiblanca cv.* (*53%*) *Picual cv.* (*24%*) *Arbequina cv.* (*24%*)
PC3	2.3	1.0 / 99.4 / 1.0	λ_{1860} to λ_{1900} (0.66–0.70), Máx. 1880 nm λ_{1370} to λ_{1500} (0.66–0.72), Máx. 1450 nm	Water content (combination and 1st overtone) Alkyl groups and Fatty acids	*Arbequina cv.* (*46%*) *Hojiblanca cv.* (*32%*) *Picual cv.* (*21%*)

Column Variables: range of coefficients in brackets are loadings, maximum of bands were indicated; column Cases: percentages in brackets are abundances.

Table 2
Results of the decision parameters of olive fruits and the quality parameters of oi obtained for the three varieties in Campaign 2 and 3.

Parameter	N	Median	Range (P5-95)	SD	RSD (%)
Campaign 2					
Arbequina cv.					
Moisture (%)	69	46.2	43.2–47.6	2.3	5.0
Fat content (%)	69	23.2	19.6–27.3	2.0	8.6
Free acidity (º)	68	0.42	0.08–0.55	0.13	31.0
Hojiblanca cv.					
Moisture (%)	68	44.8	40.6–48.9	4.1	9.2
Fat content (%)	68	21.0	16.7–26.3	2.3	11.0
Free acidity (º)	68	0.23	0.09–0.80	0.14	60.9
Picual cv.					
Moisture (%)	66	46.8	32.5–53.0	2.2	4.7
Fat content (%)	66	21.3	16.3–22.7	1.6	7.5
Free acidity (º)	66	0.20	0.09–0.45	0.08	40.0
Campaign 3					
Arbequina cv.					
Moisture (%)	158	58.5	46.0–77.7	4.2	7.2
Fat content (%)	158	17.8	11.9–23.6	2.3	12.9
Free acidity (º)	126	0.28	0.09–1.1	0.15	53.6
Oil yield (%)	48	2.48	0.49–5.05	1.68	67.6
Peroxide value (mEq O_2 kg^{-1})	47	22.3	14.1–39.1	9.5	42.7
Specific extinction at 225 nm	48	2.58	2.26–2.91	0.26	10.1
Specific extinction at 232 nm	48	2.26	1.96–2.73	0.30	13.4
Specific extinction at 270 nm	48	0.121	0.050–0.287	0.14	120.2
Variation near 270 nm	48	0.0015	0–0.0090	0.0033	220.0
Hojiblanca cv.					
Moisture (%)	155	61.7	39.1–71.6	7.5	12.2
Fat content (%)	155	15.1	9.3–27.0	2.6	17.2
Free acidity (º)	110	0.19	0.09–1.4	0.17	89.5
Oil yield (%)	61	2.42	0.20–4.69	1.82	75.1
Peroxide value (mEq O_2 kg^{-1})	61	25.9	14.9–40.8	8.9	34.6
Specific extinction at 225 nm	61	2.35	2.16–2.71	0.17	7.4
Specific extinction at 232 nm	61	1.69	1.43–2.46	0.31	18.6
Specific extinction at 270 nm	61	0.146	0.083–0.244	0.062	42.4
Variation near 270 nm	61	0.0030	0–0.0068	0.0023	78.1
Picual cv.					
Moisture (%)	152	58.3	46.2–65.8	3.3	5.7
Fat content (%)	152	17.4	9.9–23.1	2.2	12.6
Free acidity (º)	120	0.14	0.07–0.47	0.09	64.3
Oil yield (%)	55	1.39	0.50–2.77	0.79	56.8
Peroxide value (mEq O_2 kg^{-1})	55	14.8	6.10–25.9	6.3	42.6
Specific extinction at 225 nm	55	2.31	2.11–2.70	0.17	7.4
Specific extinction at 232 nm	55	1.68	1.45–2.12	0.20	12.2
Specific extinction at 270 nm	55	0.128	0.074–0.197	0.041	31.7
Variation near 270 nm	55	0.0010	0.0020–0.0030	0.0013	130.8

Decision parameters: Moisture (WC), Fat content (FC), Free acidity (FA). Quality parameters: Oil yield (OY), peroxide value (PV), specific extinctions (E225, E232, E270), variation (ΔE).

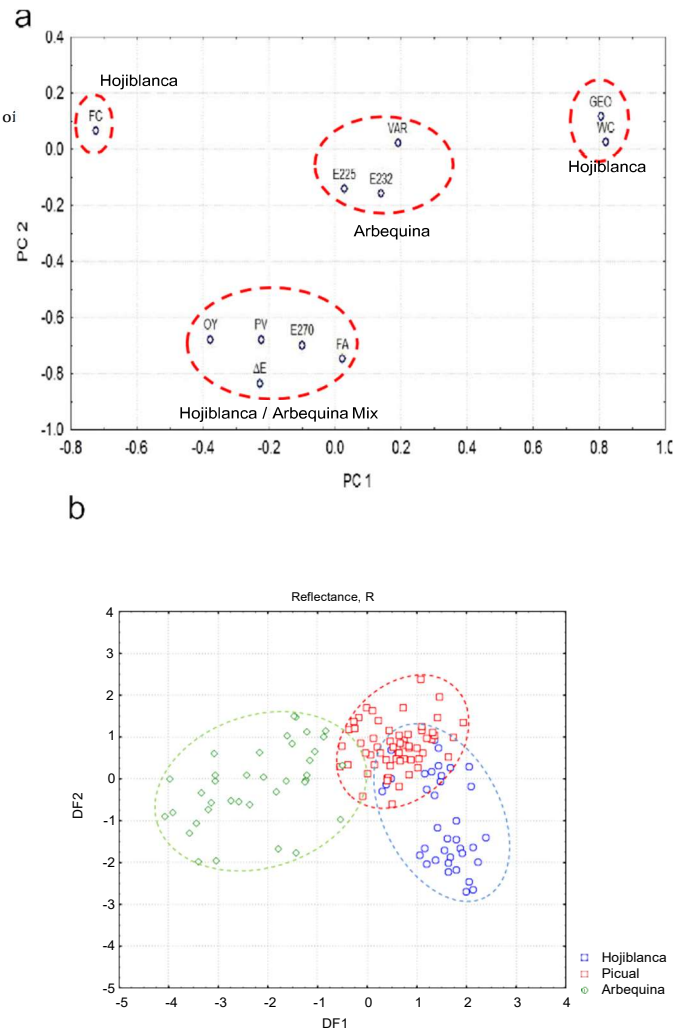


Fig. 5. Sample distribution in the plane of: (a) oil variables (analytical parameters) and cultivar varieties and (b) the linear discriminant functions of analytical parameters for the three varieties in Reflectance mode.

comparison with Campaign 2, the calibration model in Campaign 3 was less robust for moisture and fat content, but much more robust for the free acidity, the main decision parameter.

As for the statistics of internal validation, it can be seen that the best results were achieved again for free acidity, with an R^2_{val} value higher 0.96 and SEPo0.1. The performance of moisture and fat content was acceptable (R_{val}^2 values of 0.9 for both). Fig. 6d–f shows the correlations of measured/predicted values in Campaign 3, and good results were seen once again for these on-line

Table 3
Result of the PCA from matrix comprising 11 variables and 163 ~ 5¼815 cases (campaign 3).

PC	Eigenvalue	Variance/cumulated/global (%)	Variables (LQPs)	Cases (olives by variety)	Interpretation
PC1	2.3	33/33/45	þWC (0.82), GEO (0.81)/—FC (0.72)	*Hojiblanca cv.* (*86%*)	Change in water and fat composition during ripening
PC2	1.8	26/60/36	Δ*E* (0.84), FA (0.73), E270 (0.70), OY (0.69), PV (0.68)	*Arbequina cv.* (*54%*) *Hojiblanca cv.* (*40%*)	Advanced oxidation state in mature olives
PC3	1.0	14/73/19	E232 (0.94), E225 (0.90), VAR (0.62)	*Arbequina cv.* (*80%*)	Initial oxidation state of non-mature and bitter olives

Column Variables: Coefficients in brackets are loadings; column Cases: percentages in brackets are abundances.

predictions (Model 2). The NIR technique can therefore be used to predict the decision parameters, particularly for free acidity, also in the case of olives in process of ripening. Similar good correlations for calibration and validation sets (e.g., R_{val}^2 0.93–0.98) were ob- tained for different seed oils in Turkey in adulteration studies [48].The results of both campaigns encourage the use of portable NIR spectroscopy for monitoring olive fruits during ripening and to establish optimal harvesting time.

3.7. Campaigns 2 and 3: chemometric calibrations, external validations

Further prediction analyses were performed on new samples, different from the sample set of Campaigns 2 and 3. Thus, we will verify again the robustness of the best calibration models de- termined in the previous section, as a necessary stage for total validation of models [49]. For example, Fig. 7 shows the measured/predicted values of decision parameters and seasonal evolutions, observing excellent grades of prediction. Besides, measured vs.predicted comparisons with the Student's *t*-test for paired values showed significance *p*-levels40.05 for the LQPs, confirming the good quality of the models for monitoring the ripening process [39].

Additionally, we tested the samples of Campaign 2 with the calibration model from Campaign 3 (Model 2), obtaining sa- tisfactory results for the three LQPs. Thus, Model 2 for olives in theripening phase is acceptable for application to mature olives.

Comparing the two sampling strategies both can be seen to present some advantages and disadvantages. Whereas in Cam- paign 2 the calibration models provided better results for moistureand fat content, in Campaign 3 the results for free acidity showed

Table 4
Calibration and internal validation statistics obtained for NIRS quality control of all olive samples of Campaigns 2 and 3.

		Calibration set/measured						Validation set/predicted			
Campaign	LQP	*N*	No. PCs	R_{cal}^2	SECV	RMSECV%	RPD	*N*	R_{val}^2	SEP	RMSEP%
Campaign 2	Moisture	170	15	0.957	0.605	1.33	5.01	32	0.949	0.657	1.44
	Fat content	170	15	0.948	1.147	5.14	2.84	32	0.940	1.253	5.58
	Free acidity	169	16	0.963	0.078	11.8	6.17	32	0.956	0.085	12.9
Campaign 3	Moisture	420	14	0.903	1.629	2.76	3.29	38	0.896	1.686	2.85
	Fat content	416	19	0.900	2.607	15.3	1.79	38	0.891	2.728	16.0
	Free acidity	318	14	0.968	0.017	9.12	5.63	38	0.962	0.019	10.0
Campaign 2 and 3	Moisture	590	11	0.880	2.684	4.88	2.92	70	0.876	2.734	4.98
	Fat content	586	17	0.761	3.625	19.5	1.36	70	0.748	3.726	20.0
	Free acidity	487	20	0.832	0.192	36.8	2.44	70	0.814	0.202	38.8

RMSEC¼root mean square relative error of calibration, SECV¼absolute standard error of cross-calibration, RMSEP¼root mean square relative error of prediction, SEP¼absolute standard error of prediction, and RPD¼Predictive deviation ratio.

great improvement. Consequently, we concluded that, although the number of samples treated in Campaign 2 was sufficient for acceptable results in moisture and fat (N¼202), this number was insufficiently high for free acidity (N¼201). Thus, in Campaign 2 free acidity achieved better results for almost twice as many samples (N ¼ 356) and a much higher number of samples shouldtherefore be included in the PLSR for the best results for all parameters. Consequently, the sum of all samples collected in Campaigns 2 and 3 are assessed in the next section.

3.8. Global Campaign-2 and 3. Relationship between quality of models and number of samples, chemometric calibrations and ex-ternal validations

When PLSR is applied to 203 þ465 ¼ 668 olive, it can observe that the calibration statistical parameters show poorer values than the statistical values for individual campaigns. The best calibrationresult was achieved for the determination of moisture, the only with an RPD equal to three (2.92), similar to the value of Campaign 3 (3.29). The value of R_{cal}^2 0.88 is also close to 0.90, but the *RMSECV* was greater (4.88%) than in Campaign 3 (2.76%) and Campaign 2 (1.33%). With regard to the other parameters, the values of RPD, RMSECV and R_{cal}^2 for fat content and free acidity were poorer than those of Campaign 2 and Campaign 3.

Therefore, observing carefully the results of the three sampling strategies it can be seen (Fig. 8) that the increase in the number of samples caused a linear decrease in the quality of the calibration models for all LQP parameters. In the case of moisture and fat, the quality of models based on SECV decreases linearly with a high correlation coefficient (r ¼ 0.997 and 0.9992) and for free acidity the quality decreases with a higher linearity (r¼0.9995) and also based on RMSECV%. We can therefore affirm that an excessive number of samples (N) employed in the PLSR technique di- minishes the quality of the calibration models for the three LQP, particularly for free acidity. Regarding the internal validation, the increase in N was also associated with the decrease in the RMSEP% parameter – r¼0.999 for free acidity – and mainly in the SEP values – r 0.999 for free acidity and fat content –. As in the calibration set, the best correlation was obtained for free acidity for SEP and also for RMSEP%.

4. Conclusions

The present study has demonstrated that NIR Spectroscopy provided us important advantages to determine water and organic functional groups in olive fruits in a qualitative and non-destructive manner and in an extremely short time. The combi- nation of NIRS with PLS Regression, which has become quantita- tive the determinations, presented good prediction indicators. These excellent results have been possible because NIRS has been performed in a portable way. When different factors of variability are introduced, such as different maturity stage, in addition to the mayor coefficient of variation, the quality of calibration models shows a general decrease. Furthermore, the main factor influen- cing the diminishing quality of the models is the number of olive samples collected (N). Observing that for moisture the best sta- tistics were obtained for N ¼ 202–458, for fat content N ¼ 202, and for free acidity N ¼ 201–356, then the optimum number of samples for the best quality of models for the three LQPs should be between 202 and 356 samples, for instance a mean of 280 samples. Finally, application of on-line portable NIR reflectance spectro- scopy to ‘intact’ olives enabled us to detect, in real-time, changes in the composition of olive fruits during the ripening process.

.

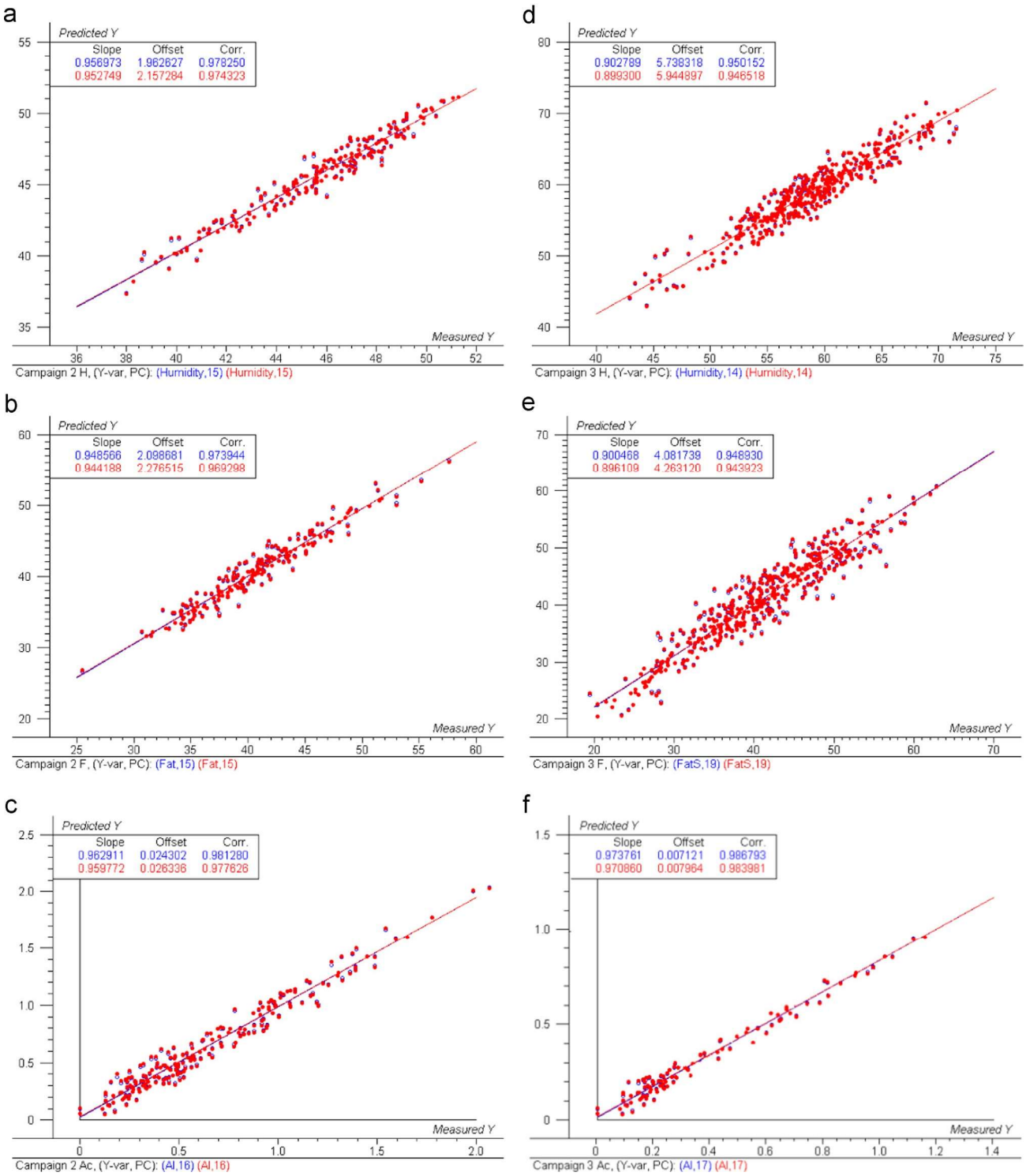
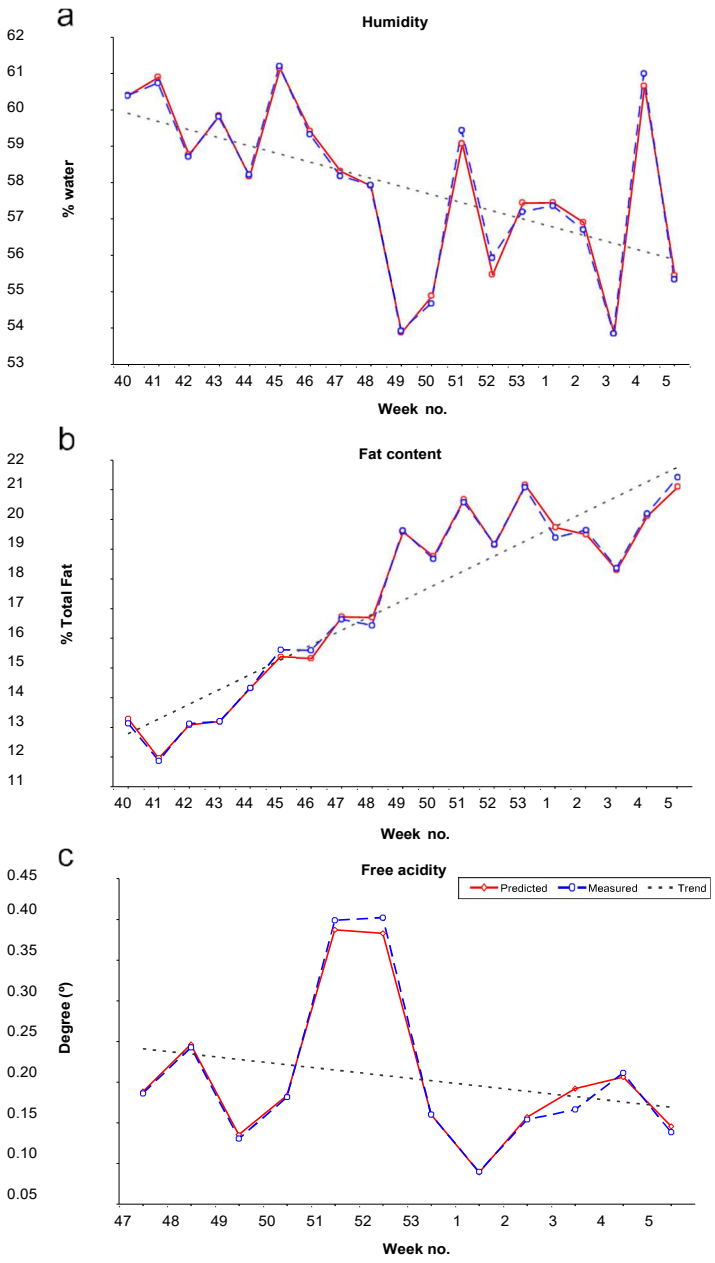


Fig. 6. Comparison between NIRS predicted and reference measured values for moisture – % (a), fat content – % (b) and free acidity – ° (c) in Campaign 2, and in Campaign 3 (d–f), using PLSR. Calibration values in blue and validation values in red. No. of PCs was also indicated. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



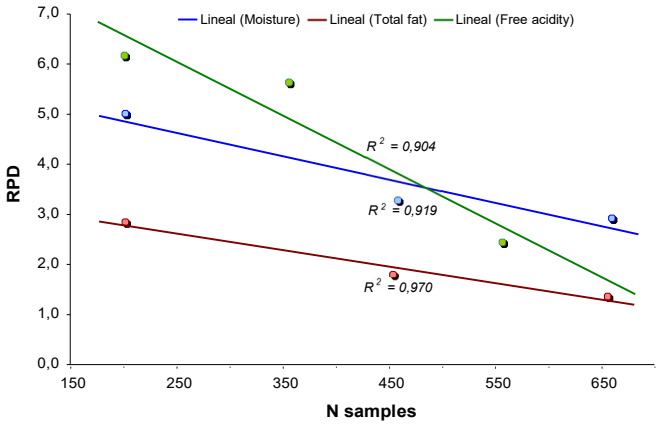


Fig. 8. Linear relationship between the number of olive samples selected for the PLSR study and the quality of models for moisture, total fat and free acidity.

Fig. 7. Evolution of the measured and estimated values during the ripening period of Campaign 3 for moisture – % (a), fat content – % (b) and free acidity – ° (c) using PLSR. Solid curve: prediction; dotted curve: laboratory analysis.

Acknowledgments

This work was carried out thanks to funding a research project (FIUS-0169/0338) supported by the Spanish Company Tecnilab. Likewise the author is grateful to the agriculture engineer Miguel Ybarra and to Jesús Espejo from the Laboratorios Espejo for pro- viding most of the olive samples. The author also thank the Edu- cational Department of the "Junta de Andalucia" for the Colla- borative Academic Agreement-FCT signed with the University of Seville, which standard chemical analyses were performed by four trained student from the IES Politécnico and supervised by F. Joaquín Jiménez and A. Gustavo Peñas.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version.

References

[1] F.D. Gunstone, Vegetable Oils in Food Technology: Composition, Properties and Uses, Wiley-Blackwell, Osford, UK, 2011
[2] D. Boskou, Characteristics of the Olive Tree and Olive Fruit, Ch. 2, Olive Oil:Chemistry and Technology, 2nd Ed., AOCS Press, Illinois, USA, 2006 (con- tributor American Oil Chemists Society).
[3] I.M. Desouky, L.F. Haggag, M.M.M. Abd El-Migeed, E.S. El-Hady, World J. Agric.Sci. 5 (2009) 760–765.
[4] A. Jimenez, A. Molina, M.I. Pascual, Sens. Actuator B 107 (2005) 64–68.
[5] B. Muik, B. Lendl, A. Molina-Díaz, L. Pérez-Villarejo, M.J. Ayora-Cañada, Anal.Bioanal. Chem. 379 (2004) 35–41.
[6] L. León-Moreno, Span. J. Agric. Res. 10 (2012) 141–148.
[7] A. Jimenez, G. Beltran, M.P. Aguilera, M. Uceda, Sens. Actuator B 129 (2008)985–990.
[8] L. Ragni, A. Berardinelli, C. Cevoli, E. Valli, J. Food Eng. 111 (2012) 66–72.
[9] European Communities, Regulation 1989/2003 of 6 November 2003 amendingRegulation (EEC) No. 2568/91 on the Characteristics of Olive Oil and Olive-pomace Oil and on the Relevant Methods of Analysis.
[10] Y. Yang, Q. Li, X. Yu, X. Chen, Y. Wang, Food Control 39 (2014) 198–203.
[11] I. Murray, Forage analysis by near infrared spectroscopy, in: A. Davies, R.
D. Baker, S.A. Grant, A.S. Laidlaw (Eds.), Sward Management Handbook, BritishGrassland Society, Cirencester, UK, 1993, pp. 285–312.
[12] P.C. Williams, K.H. Norris (Eds.), Near infrared technology in the agriculturaland food industries, 2nd, American Association of Cereal Chemist, St. Paul, Minnesota, USA, 2001.
[13] B.M. Nicolaï, K. Beullens, E. Bobelyn, A. Peirs, W. Saeys, K. Theron,
J. Lammertyn, Postharvest Biol. Technol. 46 (2007) 99–118.
[14] M. Zude, B. Herold, J.M. Roger, V. Bellon-Maurel, S. Landahl, J. Food Eng. 77

[15] and Uses, Wiley-Blackwell, Osford, UK, 2011.
[16] D. Boskou, Characteristics of the Olive Tree and Olive Fruit, Ch. 2, Olive Oil: Chemistry and Technology, 2nd Ed., AOCS Press, Illinois, USA, 2006 (con- tributor American Oil Chemists Society).
[17] I.M. Desouky, L.F. Haggag, M.M.M. Abd El-Migeed, E.S. El-Hady, World J. Agric.Sci. 5 (2009) 760–765.
[18] A. Jimenez, A. Molina, M.I. Pascual, Sens. Actuator B 107 (2005) 64–68.
[19] B. Muik, B. Lendl, A. Molina-Díaz, L. Pérez-Villarejo, M.J. Ayora-Cañada, Anal.Bioanal. Chem. 379 (2004) 35–41.
[20] L. León-Moreno, Span. J. Agric. Res. 10 (2012) 141–148.
[21] A. Jimenez, G. Beltran, M.P. Aguilera, M. Uceda, Sens. Actuator B 129 (2008)985–990.
[22] L. Ragni, A. Berardinelli, C. Cevoli, E. Valli, J. Food Eng. 111 (2012) 66–72.
[23] European Communities, Regulation 1989/2003 of 6 November 2003 amendingRegulation (EEC) No. 2568/91 on the Characteristics of Olive Oil and Olive-pomace Oil and on the Relevant Methods of Analysis.
[24] Y. Yang, Q. Li, X. Yu, X. Chen, Y. Wang, Food Control 39 (2014) 198–203.
[25] I. Murray, Forage analysis by near infrared spectroscopy, in: A. Davies, R.
D. Baker, S.A. Grant, A.S. Laidlaw (Eds.), Sward Management Handbook, BritishGrassland Society, Cirencester, UK, 1993, pp. 285–312.
[26] P.C. Williams, K.H. Norris (Eds.), Near infrared technology in the agriculturaland food industries, 2nd, American Association of Cereal Chemist, St. Paul, Minnesota, USA, 2001.
[27] B.M. Nicolaï, K. Beullens, E. Bobelyn, A. Peirs, W. Saeys, K. Theron,
J. Lammertyn, Postharvest Biol. Technol. 46 (2007) 99–118.

[28] M. Zude, B. Herold, J.M. Roger, V. Bellon-Maurel, S. Landahl, J. Food Eng. 77 (2009) 254–260.
[29] M. Schmutzler, C.W. Huck, Vib. Spectrsc. 72 (2014) 97–104.
[30] J.A. Guthrie, C.J. Liebensberg, K.B. Walsh, aust. J. Agric. Res. 57 (2006) 411–418.[17] R. Lu, Trans. ASAE 44 (2001) 1265–1271.
[18] H. Ito, S. Morimoto, R. Yamuchi, K. Ippoushi, K. Azuma, H. Higashio, ActaHortic. 588 (2002) 353–356.
[19] A. Bendini, L. Cerretani, F. Di Virgilio, P. Belloni, G. Lercker, T. Gallina, Eur. J.Lipid Sci. Technol. 109 (2007) 498–504.
[20] J. Espejo, Comparative Analytical Study Between Wild Olive Oil and VirginOlive Oil (Doctoral thesis), Sevilla, Spain, 2005.
[21] C.W. Huck, Phytochem. Lett. 11 (2015) 384–393.
[22] F. Barton, Spectrosc. Eur. 14 (2002) 12–18.
[23] W.F. Mcclure, J. Infrared Spectrosc. 11 (2003) 487–518.
[24] H. Schulz, S. Pfeffer, R. Quilitzsch, B. Steuer, K. Reif, Planta Med. 68 (2002)926–929.
[25] G. Luna, M.T. Morales, R. Aparicio, Food Chem. 98 (2006) 243–252.
[26] I.M. Desouky, L.F. Haggag, M.M.M. Abd El-Migeed, E.S. El-Hady, World J. Agric.Sci. 5 (2009) 760–765.
[27] M. Chtourou, B. Gargouri, H. Jaber, R. Abdelhedi, M. Bouaziz, Eur. J. Lipid Sci.Technol. 115 (2013) 631–640.
[28] R. Mailer, D. Conlan, J. Ayton, Olive Harvest: Harvest Timing for Optimal OliveOil Quality, in: RIRDC (Ed.), Kingston, Australia, 2005.
[29] ISO 662:1998. Animal and vegetable fats and oils – determination of moistureand volatile matter content. UNE-EN ISO 662-September 2001.
[30] UNE 55030:1961. Determination of the Total Fat Content in Olives.
[31] ISO 660:2000/A1. Animal and vegetable fats and oils – determination of acidvalue and acidity. UNE-EN ISO 660-February 2010.
[32] D. Tran Chieu, Talanta 45 (1997) 237–248.

[33] W.F. McClure, J. Infrared Spectrosc. 11 (2003) 487–518. Fundamentals, techniques and design. Part 5. Optical detectors. Optical Societyof America, McGraw-Hill, New York, USA, 2010.
[34] L. Nérgaard, A. Saudland, J. Wagner, J.P. Nielsen, L. Munck, S.B. Engelsen, Appl.Spectrosc. 54 (2000) 413–419.
[35] D.A. Burns, E.W. Ciurczak, Handbook of Near-Infrared Analysis, Third Edition,CRC Press, Florida, USA, 2008.
[36] W. Guggenbichler, C.W. Huck, A. Kobler, M. Popp, G.K. Bonn, J. Food Agric. Env.4 (2006) 98–106.
[37] M. Navarro Escamilla, F. Rodenas Sanz, H. Li, S.A. Schönbichler, B. Yang, G.K. Bonn, C.W. Huck, Talanta 114 (2013) 304–310.
[38] G. Wells Moncada, M.I. González Martín, O. Escuredo, S. Fischer, M. Míguez,Talanta 116 (2013) 65–70.
[39] O.Y. Rodionovaa, A.L. Pomerantsev, J. Chemometr. 22 (2008) 674–685.
[40] F. Westad, A. Schmidt, M. Kermit, J. Infrared Spectrosc. 16 (2008) 265–273.
[41] A. Palacios-Morillo, A. Alcázar, F. De Pablos, J.M. Jurado, Spectrochim. Acta PartA 103 (2013) 79–83.
[42] P. De la Mata, A. Dominguez-Vidal, J.M. Bosque-Sendra, A. Ruiz-Medina, L. Cuadros-Rodríguez, M.J. Ayora-Cañada, Food Control 23 (2012) 449–455.
[43] N. Sinelli, M. Casale, V. Di Egidio, P. Oliveri, D. Bassi, D. Tura, E. Casiraghi, FoodRes. Int. 43 (2010) 2126–2131.
[44] I. González-Martín, N. Álvarez-García, J.M. González-Cabrera, Talanta 69(2006) 706–710.
[45] D. Cozzolino, M.J. Kwiatkowski, R.G. Dambergs, W.U. Cynkar, L.J. Janik, G. Skouroumounis, M. Gishen, Talanta 74 (2008) 711–716.
[46] H. Yu, Y. Ying, X. Fu, H. Lu, J. Infrared Spectrosc. 14 (2006) 37–44.
[47] G. Gurdeniz, B. Ozen, Food Chem. 116 (2009) 519–525.
[48] M.A. Cantarelli, I.G. Funesc, E.J. Marchevsky, J.M. Camiña, Talanta 80 (2009)489–492.