Prediction of quality of intact olives by near infrared spectroscopy

José Antonio Cayuela*¹, María del Carmen Pérez Camino²

¹Department of Physiology and Technology of Plant Products. ²Department of Food Quality and Characterization
Instituto de la Grasa -CSIC. Avda. Padre García Tejero, 4 41012 Sevilla, Spain

Running title: Prediction of intact olives’ quality by NIRS

Correspondence should be sent to:
José Antonio Cayuela Sánchez
Instituto de la Grasa – CSIC, Avda. Padre García Tejero, 4. 41012 Sevilla. Spain
Tel. 0034 954611550 ext. 259
Email: jacayuela@ig.csic.es

Key words: acidity/dry matter/moisture/oil/olive/NIR

© 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Received: September 22, 2009 / Revised: June 4, 2010 / Accepted: August 16, 2010

DOI:10.1002/ejlt.201000372
Abbreviations used

DM: dry matter
FCV: full cross validation
FMI: fruit maturity index
FM: fruit moisture
OC_{DM}: oil content in dry matter
OC_{FW}: oil content in fresh weight
OFA: oil free acidity
PLS: partial least squares
PCs: principal components
R_{C}: correlation coefficient of calibration
R_{CV}: correlation coefficient of cross validation
RCV: random cross validation
RMSECV: root mean square error of cross validation
RMSEP: root mean square error of prediction
RPD: residual predictive deviation
SECV: standard error of cross validation
SEP: standard error of performance
Vis/NIRS: visible/near infrared spectroscopy

Prediction of quality parameters of intact olives by near infrared spectroscopy

José Antonio Cayuela*1, María del Carmen Pérez Camino2
Department of Physiology and Technology of Plant Products1. Department of Food Quality and Characterization2
Instituto de la Grasa - CSIC. Avda. Padre García Tejero, 4 - 41012 Sevilla, Spain
(Corresponding author: jacayuela@ig.csic.es)

Abstract
In this work we investigated the intact olive fruit quality prediction parameters measured directly by visible and near infrared spectroscopy (Vis/NIRS); the usefulness of a portable spectrometer is also assessed. The analysed parameters of the olive fruits were moisture, dry matter, oil content, oil free acidity and fruit maturity index. It was also studied whether NIR prediction of dry matter on olives may be more useful than NIR moisture measurement. Likewise, the results from the NIR prediction of olive oil contents related to dry matter as well as to fresh weight, were compared. Models for oil
content were developed using Soxhlet extraction from dried olive paste as the reference analysis. Results indicate a good prediction potential of the models for the olive quality parameters analysed, with RPD ratios from 2.51 to 3.18. The successful NIR prediction of these quality parameters are reported for the first time.

**Keywords**: acidity/dry matter/ moisture/oil/olive/NIR

**Practical Applications**

The technique presented here can expedite the milling procedure by allowing early detection of the olives’ quality and, by following the parameters, a quick calculation of the economic returns to the producers becomes possible. Furthermore, since oil quality depends largely on the optimal harvesting date when the olives should be taken to the mill, techniques that enable the monitoring of the oil content in olive fruit at different stages of maturity, even while still on the tree, become a useful and practical tool.

In the future, this technique will allow monitoring the quality attributes of large amounts of the olive fruit entering the mill without the need of laboratory analysis that can only be conducted on a small number of olive fruit samples.
1 Introduction

Non-destructive techniques can improve handling at the olive mills because the quality characteristics of the olive fruit would be determined beforehand. As a result, this could allow for quick calculation of payment to producers according to the quality parameters measured. In addition, these techniques could allow monitor quality attributes of large amounts of the olive fruit that enters the mill without the need of laboratory analysis, analysis that can only be conducted on a small number of olive fruit samples. Besides, oil quality depends largely on the optimal harvesting date when the olives should be taken to the mill [1]. Hence, techniques that enable the monitoring of the oil content in the olive fruit at different stages of maturity would be useful.

Although the use of Near Infrared Spectroscopy (NIRS) for determining internal quality of intact fruit was already investigated in the eighties, its industrial application has not started until now in most developed countries [2]. In intact fruit, NIRS has been used mainly to determine non-destructively the soluble solids content (SSC) of several commodities such as apples [3-5], citrus fruits [6-9], peaches [10-11], cherries [12] and melons [13-16].

Regarding olive products, oil content and humidity determinations by NIRS in olive paste are habitually applied for routine analysis [17-19], with comparable efficacy against already well established methodologies such as nuclear magnetic resonance and the Soxhlet extraction [17-18, 20-22]. The analysis of various other quality parameters of vegetable oils by NIRS is also possible [23-25].

Quantitative or qualitative analysis of intact olives through NIRS is very scarcely found in scientific literature. León et al. [26] reported NIRS as a useful tool in olive breeding programmes, using NMR as reference method and obtained regression models by partial least squares (PLS) and NIRS with prediction errors sufficiently small, for analysis of oil content (SEC=1.33), moisture (SEC=1.88) and fatty acid composition (SEC from 2.91 to 4.73 for oleic and linoleic) in intact olive. The above mentioned work was conducted using a DA-7000 diode array Vis/NIR spectrometer (Perten Instruments) with spectral range 400-1700 nm.

Cayuela et al. [27] reported interesting results using an original methodology of calibration based upon individual olive fruits using NIRS and solvent extraction as reference olive oil content analysis. In the same study, a small test was carried out with fourteen samples that allowed the development of a 'pre-calibration' procedure using the same methodology as described in the present work. Encouraging results were reported, although no validation was conducted. Farther that article, the development of calibrations for olive moisture, oil content, fruit maturity index and oil
free acidity predictions using Soxhlet method directly as reference analysis have not been reported in scientific literature.

The oil free acidity can be determined according to UNE 55070 standard [28] on olive oil samples extracted by Soxhlet. Significant differences about 0.3 degrees higher to that of the acidity in the original oil content of the olive fruit, has been reported using this method in comparison with the olive oil obtained by the Abencor system, which is considered better [29]. An important consideration regarding the use of NIRS prediction models is that there exists a huge variety of tools, mainly regarding techniques for selecting the measurement wavelength [30]. A description of the various technologies and their respective advantages has been reported by Nicolaï et al [2]. Thus, the properties of models using a NIR spectrometer with a specific methodology of calibration are not exportable to different operational configurations.

The objective of the present research was to test the usefulness of a portable Vis/NIR spectrometer using PLS calibrations for the prediction of fruit moisture, oil content, fruit maturity index and oil free acidity. We have studied whether NIR predicting dry matter on olives can be a better alternative to NIR moisture measurement. Likewise, there are comparisons of NIR predictions based on olive oil contents in dry matter and in fresh weight. Olive oil contents were determined using Soxhlet as the reference method for oil extraction, NIR calibrations were developed from intact fruit spectra. The successful NIR predictions of the oil content of the olives using Soxhlet as reference and of the fruit maturity index, based on the spectra of intact olives, are now reported for first time.

2 Materials and Methods

2.1 Olives

The calibrations were carried out with olives (Olea europaea L.) from two varieties, Picual and Arbequina, both harvested at an experimental olive grove belonging to the ‘Instituto de la Grasa’ (CSIC) in Dos Hermanas (Seville). These two olive varieties have morphological and physiological differences, particularly regarding fruit size and date of veraison. Therefore, these two varieties provide a wide range of analysed parameters in conjunction with calibration robustness. Since the strength of calibrations depends largely on the statistical range of the parameters examined, it is desirable to include into them a wide variety of stages of maturity. For this purpose, three batches of each olive variety were taken in 2008 December 3, 10 and 22 that were within a wide range of maturity stages, including each of the maturity levels considered by the fruit maturity
index (FMI) defined by Uceda [31]. Once in the lab, olives were selected and classified according to seven levels of maturity defined by the FMI and samples were formed with approximately 100 g of olives which had a uniform maturity index, from 0 to 7, as defined in the section 2.7. Each olive fruit was selected and each sample was formed, therefore, by olives in the same maturity stage. The fruit size mean was 25.2 mm length by 19.0 mm wide in Picual and 20.4 length by 15.4 mm wide in Arbequina. The total number of 100 g olive samples analysed were 183 being 104 and 79 corresponding respectively to Arbequina and to Picual. Once in the lab and the samples formed, their spectra were acquired.

2.2 NIRS instrumentation

A portable Vis/NIR Labspec (Analytical Spectral Devices Inc., Boulder) spectrophotometer equipped with three detectors was used. The detector for the range 350-1000 nm is a fixed reflective holographic diode array with 512 pixels sensitivity. The wavelength range 1000-1800 nm is covered by a holographic fast scanner InGaAs detector cooled at -25°C. The same device coupled with a high order blocking filter operates for the 1801-2500 nm interval. The instrument is equipped with internal shutters and automatic offset correction, the scanning speed being 100 ms.

2.3 Spectral acquisition

The acquisition of spectra was performed using the ‘Sample Turn Table’ accessory of the spectrometer (Figure 1) with standard SMA 905 optical fibre connectors. A spectrum average was acquired in reflectance mode from each 100 g olive fruit sample contained in a quartz Petri dish, whilst the ‘turn table’ was turning. For this purpose, the configuration for continuous acquisition of 100 scans of spectra, that were averaged by the acquisition software to form a single spectrum from each sample, was used. The resolution of the spectra corresponded to an interval of 1 nm. Acquisition was carried out using Indico Pro software (Analytical Spectral Devices Inc., Boulder). Total acquisition time for each sample was approximately 15 s.

Figure 1

2.4 Dry Matter and Moisture Content

After spectral acquisition, each olive sample of approx. 100 g olives was ground in a MM-100 mill (mc2 Ingeniería y Sistemas S.L., Sevilla), the paste placed in capsules and dried in a stove at 110 °C for 24 h. The resulting dried material was weighed to determine FM and DM. The olive dry matter
(DM) was determined by gravimetry [32] as a percentage of fresh weight that dried weight represents. Olive moisture (FM) was determined similarly, considering the difference between fresh and dried weight.

2.5 Oil Content

The oil content from the dry material of each sample was extracted by Soxhlet using hexane as solvent. For this purpose, the dry material was ground with a manual mortar, preparing with it paper cartridges that were introduced into the Soxhlet extraction units. After a 6 h extraction process, flasks containing the oil of each sample were introduced into a oven at 110 °C for 2 h to ensure the absence of hexane traces, cooled at room temperature and then gravimetrically determined [33]. The oil content taken for both dry matter (OC_{DM}) and fresh weight (OC_{FW}), is expressed as percentages. The weight must be taken four times to determine OC_{DM} and one more additional time for OC_{FW}, because the process requires the fresh weight. Exhaustive oil extraction by hexane can be assumed. Hence, when a 0.01 g precision balance is used, a 4% error margin can be expected at least for OC_{DM} and 5% for OC_{FW}. However, it is clear that these error margins are an estimation based on gravimetric measurement accuracy. Jiménez et al. [34] reported similar errors determined by the statistical procedure described by Grant [35].

2.6 Oil Free Acidity

Oil free acidity was determined on olive oil samples extracted by Soxhlet, according to the UNE 55070 standard [28]. Despite the differences with the acidity of olive oil obtained by the Abencor system, as indicated in the Introduction, this methodology is useful to test the ability of a NIR technology in the prediction of this parameter, and is used within that concept in this study. Free acidity (OFA) of olive oils Soxhlet extracted was analysed according to Official Methods of Analysis of the EC [36] and expressed as oleic acid. Briefly, 4 to 6 g of olive oil were weighed in 250 mL wide mouth Erlenmeyer flasks, 50-mL ethyl alcohol: ethyl ether 1:1 with a few drops phenolphthalein was added, and then neutralized with NaOH 0.1 N until it turned pink in colour.

2.7 Fruit Maturity Index

Not only chemical but also physical information is contained in visible/NIR spectra. In fact, colour is one of the properties reflected that keeps a close relationship with the ripeness of the olives. As previously described, each sample formed by approximately 100 g of olives, was selected and classified according to one of the seven levels of maturity as defined by Uceda et al. [31], from 0 (olive skin intense green colour) to 7 (olive skin black colour and flesh completely purple to the
bone). Each olive fruit was selected and each sample was formed with olives in the same maturity stage and classified according to the value of its maturity index. These values were used as reference for constructing NIR prediction models for FMI.

2.8 Chemometry and Calibration Procedure

Partial Least Squares [37] regression generalises and combines features from principal component analysis and multiple regression. It is particularly useful when we need to predict a set of dependent variables from a very large set of independent variables. In the development of prediction models for olive fruit quality parameters from its NIR spectrum, a regression of reference values of these parameters over a very large number of spectral variables, defined by the different wavelengths, is carried out.

Partial Least Squares models were developed using Unscrumbler 9.7 (CAMO Software AS, Norway). The calibrations for each parameter analysed were developed with Arbequina and Picual varieties together. For this purpose 4/5 of the total available samples were used, the 1/5 remaining being forest aside to be used as the external validation set. The set for validation was formed by Unscrambler’s specific application, taken every 5 samples. Before calibrations the reflectance data were transformed to absorbance and then mean normalized. Tests were conducted with Multiplicative Scatter Correction (MSC) and first order Savitzky-Golay differentiation to ascertain the best pre-treatment of that data. Standard Normal Variate Transformation (SNV) was also used as indicated in the text. The spectral data pre-treatment was made using Unscrambler 9.7.

Analysis with the ‘spectral important variables’ application of the same software was conducted on each model and calibration tests were made with the corresponding wavelength intervals. The spectral interval 350-430 nm was eliminated in all the calibrations, to avoid this noise region at the beginning of spectra. Calibration tests were also conducted with different numbers of principal components in order to determine the number of optimum PCs.

Exceptionally, where indicated in the text, points identified as outliers were investigated using the specific application of Unscrambler 9.7.

An important aspect in PLS is the need of prevention against over-fitting, using proper methodology of calibration [38]. For this purpose, the incidence of ‘over-fitting’ was tested comparing full and randomised cross-internal validation (FCV and RCV), and assessed in terms of RMSECV considering also the comparison between the values of $R_C$ and $R_{CV}$.

2.9 Model validation procedure
External validations were conducted for each analysed parameter, the model performance assessed in terms of the ratio residual predictive deviation (RPD), described [39] as the ratio of the standard deviation of the reference data from the validation set to the SEP. The root mean square error of prediction (RMSEP) and the percentage it represents with respect to the mean of the analyzed parameter in the validation set, were also considered.

Paired samples test T for dependent samples were also conducted to verify the results from RPD and RMSEP analysis. For this purpose, data pairs of the reference value and the resulting prediction corresponding to the external validation sets were compared. The paired test T is a parametric procedure, useful for testing whether the means of two groups are different, when the samples are drawn in pairs. Since it is not established as specific statistic for assessing PLS model performance, here the T test was applied referring exclusively to the data pairs included in the external validation exercise. The T test was carried out using SPSS Statistics software (SPSS Inc., Chicago). The compliance with the null hypothesis in this test (*P>0.05) indicates that the measure NIR has at least the same accuracy as the reference method.

3 Results

3.1 Olive fruit NIR spectrum

Near-infrared spectra show various overlapped bands, result of first and second overtones and combination of fundamental vibrations, mainly carbon-hydrogen. The assignment of the major near-infrared absorption bands of agricultural products have been described in various publications, such as Shenk et al. [40] or Harwood and Aparicio [41] among others. Typical diffuse absorbance spectra from two samples each of intact olives Picual and Arbequina acquired with the Labspec are shown in Figure 2. The spectra showed characteristics similar to those previously described [26] beyond differences regarding the wavelengths acquired.

There are bands of high intensity, related to the strong water absorbance that exist from its first overtone at 1400 to 1500 nm and combination band at 1880-2100 nm. A broad absorbance band exists around 1220 nm, probably from oil and due to seconds overtones of C-H and CH=CH-stretching vibrations [42]. The visible wavelengths show bands of 600-650 nm on which absorbance rise is seen; perhaps due to the presence in the samples of anthocyanin pigments, which contribute to reddish, as well as green and yellow colours that appear at 550-625 nm. The purple colouring matches a noise zone at the beginning of the spectrum.

Figure 2

3.2 Population characterization
The values from the parameters studied resulting from the reference analysis corresponding to the calibration and external validation sets are included in Table 1. As can be seen, wide variation ranges of the olive quality parameters studied were integrated in the calibrations. The descriptive statistics corresponding to Picual and Arbequina samples are also detailed (Table 1). Arbequina showed averaged maturity stage corresponding to a FMI 3.00, whereas the same index for Picual was 3.98, what indicates a more advanced maturity stage of the Picual olive. Oil free acidity was the quality parameter studied showing lowest variation in the different samples of both varieties, according to its statistics. The correlation coefficient between FM and DM was 1.00, which agrees with the standard deviation as presented by these parameters, that is related to the results of the calibrations obtained as discussed below. For OC\textsubscript{DM}/OC\textsubscript{FW} it was 0.99, that also agrees with the closed outcomes of the calibrations developed. The correlation between FMI and OFA showed coefficient 0.32, which reflects that factors different from fruit maturity index are affecting oil acidity.

Table 1

3.3 Spectral Variable Analysis and Chemometry

Absorbance data mean normalized provided the best fits in all the prediction models. None of the tests performed using different spectral data pre-calibration treatments such as MSC, differentiation or SNV, have improved the results.

The analysis of important variables conducted for the calibrations, using for this purpose the specific application of Unscrambler 9.7, revealed the most relevant spectral intervals, those shown in Table 2. Calibration and external validation tests were conducted with these sets of spectral variables. The results of these tests have not shown in the analyzed parameters the yield better than that obtained with the whole spectrum. Some calibration tests showed RMSEP values in the external validation exercises clearly lower and some others very close to those (data not shown). Therefore, for all the olive quality parameters analysed, calibrations developed from the spectrum falling within from the range of 431nm to 2500 nm provided the best performance. The PCs number best fit for each calibration reached are indicated in Table 3.

Table 2 - Table 3
3.4 Cross Validation Tests

Random cross-internal validation always provided lower bias for the different parameters analysed than the full-cross procedure, for the same number of PCs. Since over-fitting implicates model performance degradation in the prediction stage [36], comparing SEP/SECV ratios from different calibrations (data not shown) allows to understand the over-fitting degree. This ratio was less than one except in oil content referred to fresh weight, for which it showed the value 1.33, in the calibrations built using RCV. Full-cross validated models, on the contrary, showed SEP/SECV ratios greater than one in all cases, although $R_C$ and $R_{CV}$ were generally better with this procedure. These data suggests that FCV built models are suffering from 'over-fitting', indicating that their prediction potential was worse. Therefore, the RCV procedure was selected for model development.

3.5 Fruit Moisture and Dry Matter

The calibration for FM showed RMSECV=1.74, that represents 4.3% of the calibration set mean. The statistical coefficients of calibration and external validation are included in Table 3. The prediction error obtained in the external validation exercise, expressed by the RMSEP, the 3.7% moisture mean of the validation set, RPD ratio being 2.51. Figure 3 represents fruit moisture predictions compared to the values analysed.

The same figure also includes the dispersion of the prediction with the model obtained with both olive varieties for DM. The RMSECV=1.71 represents 2.9% of the calibration set mean. An external validation exercise using samples different from those used in the calibration was carried out. For this purpose, DM values were predicted with an independent validation set, RMSEP value being 1.47 and 2.4% of mean from this set, RPD ratio being 2.51.

Figure 3

3.6 Oil Content

Separate calibrations were developed for predicting the oil content in olive fruit fresh weight and to the same fruit’s dry matter, using for this purpose the corresponding reference values calculated beforehand.

Five samples from OC$_{FW}$ calibration were identified as outliers using the specific Unscrambler 9.7 ‘mark outliers only’ application, the analytical features investigated and eliminated from the calibration, since its analysis procedure presented potential defects. This calibration was the only one, among the studied, showing outliers. The RMSECV in OC$_{FW}$ model reached the value 0.90
(3.6%). The dispersion of the prediction obtained compared with the corresponding analytical values is shown in Figure 3. The statistical coefficients of calibration are included in Table 3. External validation was conducted, RPD ratio being 2.77 and RMSEP value 0.96, which represented 3.9% of $OC_{FW}$ mean in the validation set.

Table 3 includes the statistical coefficients of $OC_{DM}$ calibration and those from the external validation exercise, RMSEP representing 3.8% and RPD ratio 3.18, therefore, slightly better than the same coefficient for $OC_{FW}$ calibration. Figure 3 provides the dispersion of prediction with $OC_{DM}$ calibration.

### 3.7 Fruit Maturity Index

The calibration conducted for FMI prediction was based upon reference values with discrete variation [29], unlike the rest of the parameters analysed, whose variation are continuous. The results of calibration are expressed graphically in Figure 3. The external validation exercise showed RPD value 3.00, indicating a good performance of the model. The RMSEP was 0.51, representing 15.7%. This result should be assessed considering the methodological characteristics of FMI reference determination, since it is a visual method to classify olives, therefore, susceptible to error.

### 3.8 Oil Free Acidity

The calibration for OFA showed RMSECV 0.03, representing 14.3% of OFA mean. The external validation output showed RPD ratio 2.67 and RMSEP 0.03. Although RPD indicates good model performance, the relatively high RMSEP percentage (14.28) revealed poor prediction potential. The corresponding dispersion plot for OFA prediction is included in Figure 3, Table 3 showing the statistics.

### 4 Discussion

It is not possible to establish a clear advantage from using DM or alternatively FM NIR calibrations, since the RPD ratios were the same (2.51) in the validation of both olive quality parameters, which reveals a very close model performance in both parameters. Nevertheless, DM validation showed a lower RMSEP percentage than that for FM, with values at 2.4% and 3.7% respectively.

The prediction of olive oil content in the dry matter provided better results, with RPD 3.18, than oil content results related to fresh weight, that showed RPD 2.77. This can be attributed to the fact that
lower analytical error in the reference analysis of oil content in dry matter is possible, because the calculation is more straightforward, hence, it fits the model better.

Percentages which represented the RMSEP from the external validations regarding the mean of each parameter DM, FM, OC_{DM} and OC_{FW} ranged from the value 2.31 to 3.90, RPD values fell between 2.51 and 3.18, which reveals sufficient good performance of the prediction models for practical application.

The external validation for FMI showed RPD 3.0, of the same order as that of OC_{DM} (Table 3), thus revealing good model performance. In the case of OFA external validation, RPD 2.67 was slightly better than the same coefficient from FM and DM calibrations, but its RMSEP was worse, revealing poor prediction potential.

Dependent paired samples test T were conducted to compare data pairs of the NIR predictions and the reference values corresponding to the external validation sets and to confirm the conclusions obtained from the analysis of the statistics RPD and RMSEP. The differences were not significant (*P > 0.05) in all the olive quality parameters. The result of the T-test reveals that the measures NIR has at least the same accuracy as the reference methods.

The T-test for FMI deserves a special comment, since the *P value was 0.087, although implied differences also not significant, it is closer to the limit of significance than the other parameters, which agrees with a higher RMSEP.

5 Conclusions

Predictions of fruit dry matter, fruit moisture and oil content carried out with the prediction models constructed, using intact olives, were relatively highly accurate according to RPD ratios from the corresponding external validations. These coefficients were very similar in all the olive quality parameters analysed, and slightly higher for OC_{DM} and FMI, that agrees with the high correlations between FM and DM and between OC_{DM} and OC_{FW} reference values.

Nevertheless, the RMSEP and its percentages showed more important differences between the various parameters. The lowest RMSEP among all the models was shown in the DM model, which percentage 2.42 emphasises the good functioning of the prediction model for this parameter. This result could be attributed to the fact that olive fruit dry matter was more abundant, hence higher
than moisture as Table 1 shows. This statistic was higher in those parameters which reference analysis method has lower accuracy.

Considering RPD and RMSEP together reveals that NIR predictions for all the olive quality parameters studied showed high prediction potential. The successful prediction of the olives’s maturity index using non-destructive technology, as shown in this work, is reported for first time. The results from the T test carried out indicate, in agreement with RPD and RMSEP statistics, that NIR models were at least as accurate as the reference methods.

The good results obtained using calibration comprising two olive varieties show that valid NIR models for predicting quality attributes of fruits of different varieties can be developed. However, more research is necessary for the construction and validation of specific models.

5. Acknowledgements
This work is a result of a research contract funded by Bonsai Advanced Technologies, S.L. (Madrid, Spain).

6. Conflict of Interest
The authors have declared no conflict of interests.

7. References


[42] P. Hourant, V. Baeten, M.T. Morales, M. Meurens, R. Aparicio. Oil and fat classification by selected band

**Figure captions**

Figure 1. ‘Sample Turn Table’ accessory

Figure 2. Example of olives spectra acquired with the Labspec.

Figure 3. Dispersion graphs of the calibrations. Fruit moisture (%), fruit dry matter (%), oil content in fresh weight (%), oil content in dry matter (%), oil free acidity (g oleic acid) and fruit maturity index
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Samples</th>
<th>Range</th>
<th>σ</th>
<th>$\bar{X}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Arbequina</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM</td>
<td>104</td>
<td>30.71-42.43</td>
<td>2.2</td>
<td>38.51</td>
</tr>
<tr>
<td>DM</td>
<td>104</td>
<td>57.57-69.29</td>
<td>2.2</td>
<td>61.49</td>
</tr>
<tr>
<td>OC&lt;sub&gt;FW&lt;/sub&gt;</td>
<td>104</td>
<td>21.82-27.15</td>
<td>1.4</td>
<td>24.16</td>
</tr>
<tr>
<td>OC&lt;sub&gt;DM&lt;/sub&gt;</td>
<td>104</td>
<td>34.15-44.38</td>
<td>2.6</td>
<td>39.32</td>
</tr>
<tr>
<td>FMI</td>
<td>104</td>
<td>1.00-6.00</td>
<td>1.6</td>
<td>3.00</td>
</tr>
<tr>
<td>OFA</td>
<td>96</td>
<td>0.08-0.71</td>
<td>0.0</td>
<td>0.16</td>
</tr>
<tr>
<td><strong>Picual</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM</td>
<td>79</td>
<td>31.52-48.71</td>
<td>2.82</td>
<td>41.91</td>
</tr>
<tr>
<td>DM</td>
<td>79</td>
<td>51.29-68.48</td>
<td>2.82</td>
<td>58.08</td>
</tr>
<tr>
<td>OC&lt;sub&gt;FW&lt;/sub&gt;</td>
<td>79</td>
<td>20.38-36.14</td>
<td>2.25</td>
<td>25.59</td>
</tr>
<tr>
<td>OC&lt;sub&gt;DM&lt;/sub&gt;</td>
<td>79</td>
<td>35.70-58.39</td>
<td>3.24</td>
<td>44.09</td>
</tr>
<tr>
<td>FMI</td>
<td>79</td>
<td>1.00-5.00</td>
<td>0.83</td>
<td>3.98</td>
</tr>
<tr>
<td>OFA</td>
<td>66</td>
<td>0.08-0.79</td>
<td>0.14</td>
<td>0.17</td>
</tr>
<tr>
<td><strong>Calibration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM</td>
<td>145</td>
<td>30.72-48.71</td>
<td>2.85</td>
<td>40.14</td>
</tr>
<tr>
<td>DM</td>
<td>145</td>
<td>51.29-69.29</td>
<td>2.85</td>
<td>59.68</td>
</tr>
<tr>
<td>OC&lt;sub&gt;FW&lt;/sub&gt;</td>
<td>145</td>
<td>20.38-29.94</td>
<td>1.76</td>
<td>24.74</td>
</tr>
<tr>
<td>OC&lt;sub&gt;DM&lt;/sub&gt;</td>
<td>145</td>
<td>35.00-48.55</td>
<td>3.47</td>
<td>41.41</td>
</tr>
<tr>
<td>FMI</td>
<td>145</td>
<td>1.00-6.00</td>
<td>1.41</td>
<td>3.48</td>
</tr>
<tr>
<td>OFA</td>
<td>129</td>
<td>0.12-0.49</td>
<td>0.09</td>
<td>0.23</td>
</tr>
</tbody>
</table>

**External validation**

| FM        | 38      | 31.52-47.46 | 3.52| 39.40     |
| DM        | 38      | 52.54-68.48 | 3.52| 60.60     |
| OC<sub>FW</sub> | 38     | 21.82-36.14 | 2.69| 24.86     |
| OC<sub>DM</sub> | 38     | 34.15-58.39 | 4.71| 41.13     |
| FMI       | 38      | 1.00-6.00   | 1.47| 3.24      |
| OFA       | 33      | 0.12-0.49   | 0.08| 0.21      |

Table 1. Statistical data. Sets of samples of the calibrations and external validations.

FM, fruit moisture; DM, dry matter; OC<sub>FW</sub>, oil content referred to fresh weight; OC<sub>DM</sub>, oil content referred to dry matter; FMI, fruit maturity index; OFA, oil free acidity; $\sigma$, standard deviation; $\bar{X}$, mean.
<table>
<thead>
<tr>
<th>Calibration</th>
<th>Spectral intervals (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FM</td>
<td>549-726, 1015-1700</td>
</tr>
<tr>
<td>DM</td>
<td>425-750, 960-1869</td>
</tr>
<tr>
<td>OC&lt;sub&gt;FW&lt;/sub&gt;</td>
<td>none</td>
</tr>
<tr>
<td>OC&lt;sub&gt;DM&lt;/sub&gt;</td>
<td>780-1125</td>
</tr>
<tr>
<td>OFA</td>
<td>520-571, 704-755, 847-940, 1055-1337, 1626-2300</td>
</tr>
<tr>
<td>FMI</td>
<td>940-1135, 1368-2300</td>
</tr>
</tbody>
</table>

Table 2. Spectral intervals more relevant from the analysis of important variables.

FM, Fruit moisture; DM, Dry matter; OC<sub>FW</sub>, Oil content referred to fresh weight; OC<sub>DM</sub>, Oil content referred to dry matter; OFA, Oil free acidity; FMI, Fruit maturity index
Table 3. Statistical coefficients of calibration and validation.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>PCs</th>
<th>$R_c$</th>
<th>$R_{CV}$</th>
<th>RMSECV</th>
<th>Bias</th>
<th>SEP</th>
<th>RMSEP</th>
<th>RPD</th>
<th>T *p value from paired samples test</th>
</tr>
</thead>
<tbody>
<tr>
<td>FM</td>
<td>8</td>
<td>0.792</td>
<td>0.759</td>
<td>1.741</td>
<td>2.63 $10^{-7}$</td>
<td>1.13</td>
<td>1.47</td>
<td>3.73</td>
<td>2.51</td>
</tr>
<tr>
<td>DM</td>
<td>7</td>
<td>0.799</td>
<td>0.766</td>
<td>1.710</td>
<td>8.68 $10^{-7}$</td>
<td>1.13</td>
<td>1.47</td>
<td>2.42</td>
<td>2.51</td>
</tr>
<tr>
<td>OCFW</td>
<td>10</td>
<td>0.848</td>
<td>0.783</td>
<td>0.901</td>
<td>8.72 $10^{-7}$</td>
<td>0.97</td>
<td>0.96</td>
<td>3.86</td>
<td>2.77</td>
</tr>
<tr>
<td>OCFDM</td>
<td>9</td>
<td>0.887</td>
<td>0.852</td>
<td>1.619</td>
<td>3.89 $10^{-7}$</td>
<td>1.48</td>
<td>1.56</td>
<td>3.79</td>
<td>3.18</td>
</tr>
<tr>
<td>FMI</td>
<td>5</td>
<td>0.944</td>
<td>0.937</td>
<td>0.465</td>
<td>9.08 $10^{-8}$</td>
<td>0.49</td>
<td>0.51</td>
<td>15.74</td>
<td>3.00</td>
</tr>
<tr>
<td>OFA</td>
<td>8</td>
<td>0.851</td>
<td>0.799</td>
<td>0.030</td>
<td>-6.99 $10^{-9}$</td>
<td>0.03</td>
<td>0.03</td>
<td>14.28</td>
<td>2.67</td>
</tr>
</tbody>
</table>

FM, fruit moisture; DM, dry matter; OCFW, oil content referred to fresh weight; OCFDM, oil content referred to dry matter; FMI, fruit maturity index; OFA, oil free acidity; $\sigma$, standard deviation; $\overline{X}$, mean; PCs, number of principal components; $R_c$, calibration coefficient; $R_{CV}$, cross-validation coefficient; RMSECV, root mean square error of cross validation; Bias, bias; SEP, standard error of performance; RMSEP, root mean square error of performance; RPD, residual predictive deviation; $T$, *p value from paired samples test.