

Elsevier Editorial System(tm) for Food Chemistry  
Manuscript Draft

Manuscript Number:

Title: CHEMOMETRIC ANALYSIS FOR DISCRIMINATION OF EXTRA VIRGIN OLIVE OILS FROM WHOLE AND STONED OLIVE PASTES

Article Type: Analytical Methods Article

Keywords: Virgin olive oil; Chemometrics; Partial least square discriminant analysis; Principal component analysis; Fourier Transform infrared; Stoned olive pastes.

Corresponding Author: Dr. Michele De Luca,

Corresponding Author's Institution: University of Calabria

First Author: Michele De Luca

Order of Authors: Michele De Luca; Donatella Restuccia; Maria Lisa Clodoveo; Francesco Puoci; Gaetano Ragno

1           **CHEMOMETRIC ANALYSIS FOR DISCRIMINATION OF EXTRA VIRGIN OLIVE**  
2                           **OILS FROM WHOLE AND STONED OLIVE PASTES**

3  
4   Michele De Luca<sup>a\*</sup>, Donatella Restuccia<sup>a</sup>, Maria Lisa Clodoveo<sup>b</sup>, Francesco Puoci<sup>a</sup>, Gaetano Ragno<sup>a</sup>

5  
6   <sup>a</sup>Department of Pharmacy, Health and Nutritional Sciences, University of Calabria, Rende, Italy

7   <sup>b</sup>Department of Agricultural and Environmental Science, University of Bari, Italy

8  
9   \*Corresponding author: Tel/Fax +390984493201; Email: [michele.deluca@unical.it](mailto:michele.deluca@unical.it)

10  
11   **ABSTRACT**

12           Chemometric discrimination of extra virgin olive oils (EVOO) from whole and stoned olive  
13   pastes was carried out by using Fourier transform infrared (FTIR) data and Partial Least Squares –  
14   Discriminant Analysis (PLS-DA) approach. Four Italian commercial EVOO brands, all in both  
15   whole and stoned version, were considered in this study. The adopted chemometric methodologies  
16   were able to describe the different chemical features in phenolic and volatile compounds contained  
17   in the two types of oil by using unspecific IR spectral information. The PLS-DA algorithm was  
18   used as supervised discriminant analysis to classify the different EVOO matrices. Discriminant  
19   analysis was extended to the evaluation of possible adulteration by addition of aliquots of oil from  
20   whole paste to the most valuable oil from stoned olives. The statistical parameters from external  
21   validation of all the PLS models were very satisfactory, with low root mean square error of  
22   prediction (RMSEP) and relative error (RE%).

23  
24   **Keywords:** *Virgin olive oil; Chemometrics; Partial least square discriminant analysis; Principal*  
25   *component analysis; Fourier Transform infrared; Stoned olive pastes.*

## 1. INTRODUCTION

A large increase in demand for high-quality virgin olive oil during the last years can be attributed not only to its particular organoleptic properties but also to its potential health benefits (Clodoveo, Camposeo, De Gennaro, Pascuzzi & Roselli, 2014). Usually, the quality of an extra virgin olive oil (EVOO) is strictly related with the quality of fruits, with the harvesting systems and above all with the oil extraction process. In this regard, the improvement of the quality standards for EVOO is continuously stimulating the search for new extraction technologies. In particular, it has been reported that the type of crusher is critical in determining the quantity and the quality of EVOO (Clodoveo, 2013; Clodoveo, 2012; Clodoveo, Durante & La Notte, 2013; Clodoveo & Hbaieb, 2013).

Currently, the de-stoner technique is the only mechanical system that allows a selective crushing of the fruits excluding the stone from the paste. The stone exclusion has the undoubted advantages to increase both nutritional quality of the product and sustainability of the process. When the olive paste is made exclusively from the fleshy part of the olive (mesocarp), the bio-phenol content is protected from oxidation because most of the oxidative enzymes, as polyphenol oxidase and peroxidase, mainly localized in the stone (endocarp), is removed (Amirante, Clodoveo, Dugo, Leone & Tamborrino, 2006; Clodoveo, Dipalmo, Schiano, La Notte & Pati, 2014; Restuccia, Spizzirri, Chiricosta, Puoci, Altimari & Picci, 2011; Servili, Taticchi, Esposto, Urbani, Selvaggini & Montedoro, 2007). On the contrary, when the “integral crushing” is applied, the presence of olive kernel reduces the work capacity of the whole mill plant. Moreover, the crushing of the kernel requires a large dissipation of mechanical energy partially converted into thermal energy which heats the olive paste before the triggering of thermolabile enzymes, such as lipoxygenase and hydroperoxide lyase (Clodoveo, Hbaieb, Kotti, Mugnozza & Gargouri, 2014). The heating of the fragments constitutes again a by-product, absorbing in vain thermal energy, with the burden of a heavy economic and energetic cost. Secondary advantages derive from the olive seed that can be source of valuable compounds for the cosmetics and pharmaceutical industries. The de-stoned

53 pomace is easier to use as an animal feed, while the kernel can be employed for the production of  
54 coal for food industries (Spizzirri et al., 2011). Despite these benefits, this technology is not  
55 widespread and the main reason lies in the fact that the de-stoning process produces higher quality  
56 EVOO but in lower yield. At the time, the commercial market does not acknowledge in EVOO  
57 stoned an extra charge sufficient to compensate for the loss of yield. The justification for a higher  
58 price should also be supported by security on product quality. Therefore, it seems necessary to have  
59 analytical techniques able to ascertain that the oil from olives pitted is intact and not adulterated  
60 with oil of different origin.

61 In recent years, numerous analytical studies using physical–chemical techniques combined with  
62 chemometric methodologies have been proposed for the characterization of food genuineness  
63 (Arvanitoyannis & Vlachos, 2007; Bassbasi, De Luca, Ioele, Ragno, & Oussama, 2014). Clustering  
64 of the samples (objects) can be performed by unsupervised methods, which identify the natural  
65 pattern, and group objects based on similarities among the samples. Information stored in the  
66 experimental data elaborated in models, used to predict new unknown samples. Application of  
67 spectroscopic techniques in the study of the origin and differentiation of food products has  
68 considerably developed in the recent years (Bassbasi, De Luca, Ioele, Ragno & Oussama, 2014;  
69 Bassbasi et al., 2014; De Luca et al., 2011a; De Luca et al., 2011b; Forina, Oliveri, Lanteri &  
70 Casale, 2008; Poulli, Mousdis & Georgiou, 2005). In particular, FTIR has been successfully used to  
71 characterize many olive oil parameters (Tapp, Defernez & Kemsley, 2003). This technique is  
72 simple to perform, not expensive in terms of time and money and usually does not require sample  
73 pre-treatment (Lerma-García, Ramis-Ramos, Herrero-Martínez & Simó-Alfonso, 2010; Ozen &  
74 Mauer, 2002). FTIR has been proposed to authenticate extra virgin olive oils or to detect  
75 adulteration and some attempts in using FTIR to distinguish olive oils from different geographical  
76 origin and different genetic varieties have been published (De Luca et al., 2012; De Luca et al.,  
77 2011b; Gurdeniz, Tokatli & Ozen, 2007; Sinelli, Cerretani, Di Egidio, Bendini & Casiraghi, 2010;  
78 Tapp, Defernez, & Kemsley, 2003).

79 Multivariate analysis provides a series of tools for handling a very large number of data and  
80 variables from different analytical methods.

81 The aim of this paper is to develop a multivariate analytical method able to certify that an olive oil  
82 has been extracted solely from stoned olives in order to offer the maximum guarantee to consumers  
83 and a proper profit for producers. In addition, a further PLS modeling on the ATR-FTIR data has  
84 been defined, allowing to assess any adulteration due to the addition of oil from whole paste to oil  
85 from stoned olives.

86

## 87 **2. MATERIAL AND METHODS**

### 88 **2.1.Samples data set**

89 EVOO samples of four different commercial Italian brands, in both whole (WO) and stoned (SO)  
90 version, named Cerasoli-Cuscinà (CER), De Finis (DEF), Sannicandrese (SAN) and Trisole (TRI),  
91 were considered in this work. Two different types of training sample sets were arranged (Figure 1).  
92 The first sample set (set A) consisted of 32 samples, including 4 samples of both WO and SO for  
93 each brand (4 samples x 2 types x 4 brands = 32). 24 samples from set A were used in PLS-DA  
94 calibration step and the remaining 8 samples in the external validation step.

95 Other samples sets were used to build and validate the PLS models for the detection of possible oil  
96 adulteration. For each oil brand (CER-B, DEF-B, SAN-B and TRI-B), a set of 15 mixtures (in  
97 triplicate,  $15 \times 3 = 45$ ) was prepared with different percentage composition of the two kinds of oil,  
98 with addition of WO to SO from 0 to 95%. 35 samples from each set were randomly selected for  
99 PLS modeling and the remaining 10 samples were used to validate the prediction performance of  
100 the models.

101

102 Figure 1 near here

103

### 104 **2.2.Instrumentation**

105 Infrared spectra were acquired by a Spectrum Two Fourier transform infrared (FTIR) spectrometer  
106 (Perkin Elmer) which was equipped with an attenuated total reflection (ATR) accessory. ATR  
107 measures the variations of an infrared beam when it takes contact with a sample. The IR beam is  
108 directed onto an optically dense crystal with a high refractive index, giving an evanescent wave that  
109 penetrates a few microns on the sample surface and will be attenuated or altered if the sample  
110 absorbs energy. This energy alteration is collected for all the selected wavelengths, furnishing the  
111 corresponding infrared spectrum.

112 EVOO samples were placed on the ATR surface and infrared spectra were recorded between 4000  
113 and 450  $\text{cm}^{-1}$ . Scan number and resolution were optimized at 16 scans and 4  $\text{cm}^{-1}$ , respectively. The  
114 spectrum of the ATR element (after cleaning and drying) versus room air was used as background.  
115 Before each analysis, the ATR plate was cleaned in situ by scrubbing it with ethanol solution,  
116 making possible to dry the ATR surface. Cleanliness was verified by collecting the background  
117 spectrum and comparing it to the previous background spectra in order to evaluate the instrumental  
118 conditions and laboratory interferences from  $\text{H}_2\text{O}$  and  $\text{CO}_2$ .

119 The Unscrambler software version 10.3 from CAMO (Computer Aided Modelling, Trondheim,  
120 Norway) was used for the chemometric treatment of ATR-FTIR spectral data.

121

### 122 **2.3. Chemometric methods**

123 The supervised classification and calibration modeling adopted in this work was based on the PLS  
124 algorithm. The PLS method performs a PCA on X and Y matrices and each principal component  
125 (PC) is extracted from the independent variables and correlated with the variance of the dependent  
126 variable. PCA procedure provides to reduce the amount of original variables to a few latent  
127 variables or PCs that represent the main information from the experimental data set (Wise, Ricker,  
128 Veltkamp & Kowalski, 1990; Wold & Sjöström, 1998). In PLS-DA, Y-variable results 1 for the  
129 objects belonging to the class and 0 for those not belonging to the class, while in PLS calibration, Y  
130 values are corresponding to the standard mixtures composition (De Luca et al., 2011b; Lerna-

131 García, Ramis-Ramos, Herrero-Martínez & Simó-Alfonso, 2010). The PLS models were validated  
132 by full cross-validation procedure, which provides to give a direct estimate of the error rate incurred  
133 by the model. The number of PCs was chosen by evaluating the parameter root mean square error of  
134 cross validation (RMSECV) and the correlation coefficient  $R^2$ . In order to test the real ability in  
135 predicting new samples, a further validation was performed for all models by applying them on  
136 external validation sets.

137

### 138 3. RESULTS AND DISCUSSION

139

140 Figure 2 near here

141

142 The FTIR spectra of the oil brands in both versions, SO and WO, are shown in Figure 2. The  
143 spectral variations were not enough to highlight differences in the oil samples. All the spectra were  
144 characterized by two peaks at 2860 and 2850  $\text{cm}^{-1}$  from asymmetrical and symmetrical stretching  
145 vibrations of methylene groups, a band at around 1725  $\text{cm}^{-1}$  from the stretching vibration of the  
146 triglycerides (ester carbonyl groups) and bands from 1480 to 1450  $\text{cm}^{-1}$  arising from the bending  
147 vibrations of the  $-\text{CH}_2$  and  $-\text{CH}_3$  aliphatic groups. The use of single peaks or limited wavelength  
148 ranges to distinguish the different extraction procedure seemed very hard, making it necessary to  
149 use multivariate analysis approaches.

150 The spectra were firstly subjected to mathematical pre-processing, applying derivatization by  
151 Savitzky–Golay algorithm, standard normal variate (SNV) and multiple scatter correction (MSC)  
152 (Iñón, Garrigues, Garrigues, Molina & de la Guardia, 2003; Savitzky & Golay, 1964). PCA  
153 clustering was performed on the transformed data in order to select the best pre-processing  
154 technique, by adopting the best separation of the classes in the score plot as a criterion. Savitzky–  
155 Golay derivative elaboration showed the most effective results (Fig. 3), after optimizing the  
156 parameters at the following values: 1st order, number of smoothing points 6, polynomial order 1.

157

158 Figure 3 near here

159

### 160 **3.1.PLS-DA modeling**

161 FTIR spectral data of 24 oil samples (from set A) were used in the PLS-DA multivariate modeling.

162 Two wavelength ranges, 4000-3100 and 2800-1900  $\text{cm}^{-1}$ , were not considered because rich in

163 instrumental noise and bringing useless information (as can be seen in Fig. 2) (Maggio et al., 2009).

164 The PLS-DA model was so built by considering only the ranges 3100-2800 and 1900-450  $\text{cm}^{-1}$  as

165 X-descriptor variables. The Y variable was associated with the two oil extraction procedures, fixing

166 its value to 1 for oils obtained by malaxation of stoned olives (SO class), and to 0 for those obtained

167 by malaxation of whole olives (WO class). The classification model was validated by full cross

168 validation and its performance was evaluated in terms of correlation coefficient  $R^2$  and RMSECV.

169 Both statistical parameters resulted acceptable by considering just 4 PCs, obtaining RMSECV value

170 of 0.120 and  $R^2$  of 0.946.

171 The model was applied on the external prediction set (8 samples from set A) and the results

172 obtained were listed in Table 1. For each sample, a Y predicted value higher than 0.55 accounted

173 the sample as belonging to the class SO, while the samples showing a value under 0.45 were

174 considered as belonging to the class WO. The results in the range 0.45 - 0.55 could be considered

175 doubtful because these limits expressed  $\pm 5\%$  of error in prediction. 100% of the tested samples

176 were well class-assigned and no sample detected as suspicious origin.

177

178 Table 1 near here

179

### 180 **3.2.PLS modeling to assess adulteration of stoned olive oil**

181 With the aim to develop a method to identify possible adulteration of SO oil samples, PLS

182 calibration models for all studied brands were calculated on the FTIR data matrices of the oil



183 samples in set B (CER set B, DEF set B, SAN set B and TRI set B). Full cross validation permitted  
184 to select the optimal PCs number for all the models by using correlation  $R^2$  and RMSECV. The  
185 values of these statistical parameters for all the matrices are listed in Table 2.

186 In the SAN calibration model, only one PC was enough (Y explained variance = 98,81%) to detect  
187 the composition of the oil mixtures and then to well predict the addition of oil from whole olives,  
188 with a  $R^2$  value of 0.988 and RMSECV of 3.395. Modeling of the brands CER and TRI needed two  
189 PCs (Y explained variance = 94.22 % for CER and 97.91% for TRI), with  $R^2$  above 0.94 and  
190 RMSECV under 6. The two oil versions of the brand DEF was characterized from high similarity  
191 and four PCs were necessary in PLS modeling (Y explained variance = 95.40 %,  $R^2 = 0.973$ ,  
192 RMSECV = 4.998).

193

194 Table 2 near here

195

196 The four PLS models built on the calibration data-sets were submitted to external validation in  
197 order to verify their prediction ability in terms of accuracy and precision. Four independent sets of  
198 10 reference mixtures (from sets B), as previously described, were used for this aim. The data of  
199 these matrices were previously processed by derivative transformation. Figure 4 shows the  
200 comparison between nominal (Reference) and predicted values obtaining in all the cases correlation  
201 above 0.97. Table 3 summarizes the statistical results in terms of  $R^2$ , Root mean square error of  
202 prediction (RMSEP) and percentage relative error (RE %) carried out from the external validation.  
203 Prediction showed satisfactory results for all EVOO brands, with RE % between 4.382 and 6.914  
204 %.

205

206 Figure 4 and table 3 near here

207

208 **4. CONCLUSIONS**

209 The very small significant differences in the FTIR spectra from oil samples obtained by two  
210 different extraction procedures on stoned and whole olives, were amplified by mathematical  
211 derivative and the transformed data elaborated by PLS procedure. The PLS-DA model gave a high  
212 percentage of correct classification and proved to be robust in distinguishing the different oil  
213 manufacture. Multivariate analysis was able to handle the unspecific information stored in the FTIR  
214 data and allowed to empathize the different chemical composition due to the different extraction  
215 processes. PLS was also applied to detect possible adulteration in stoned EVOO, showing very  
216 good predictive performance. This study demonstrated the high potential of the combination of  
217 FTIR spectroscopy and chemometrics in oil analysis, providing reliable results and mathematical  
218 models of simple application. The method is also inexpensive and requires reduced running times  
219 requiring any pretreatment of the samples

220 **References**

221

222 Amirante, P., Clodoveo, M. L., Dugo, G., Leone, A. & Tamborrino, A. (2006). Advance technology  
223 in virgin olive oil production from traditional and de-stoned pastes: Influence of the introduction  
224 of a heat exchanger on oil quality. *Food Chemistry*, 98, 797-805.

225 Arvanitoyannis, I. S. & Vlachos, A. (2007). Implementation of physicochemical and sensory  
226 analysis in conjunction with multivariate analysis towards assessing olive oil  
227 authentication/adulteration. *Critical Reviews in Food Science and Nutrition*, 47, 441-498.

228 Bassbasi, M., De Luca, M., Ioele, G., Ragno, G. & Oussama, A. (2014). Prediction of the  
229 geographical origin of butters by Partial Least Square Discriminant Analysis (PLS-DA) applied  
230 to Infrared Spectroscopy (FTIR) data. *Journal of Food Composition and Analysis*, 33, 243-249.

231 Bassbasi, M., De Luca, M., Souhassou, S., Hirri, A., Berkani, M., Kzaiber, F., Ioele, G., Ragno, G.  
232 & Oussama, A. (2014). Determination of Milk Adulteration by Sucrose Using FT-MIR  
233 Spectroscopy and Chemometrics Methods. *Agricultural Science Research Journal*, 4, 175- 180

234 Clodoveo, M. L. (2013). New advances in the development of innovative virgin olive oil extraction  
235 plants: Looking back to see the future. *Food Research International*, 54(1), 726-729.

236 Clodoveo, M. L., Camposeo, S., De Gennaro, B., Pascuzzi, S., & Roselli, L. (2014). In the ancient  
237 world, virgin olive oil was called “liquid gold” by Homer and “the great healer” by Hippocrates.  
238 Why has this mythic image been forgotten?. *Food Research International*, 62, 1062-1068.

239 Clodoveo, M. L., Dipalmo, T., Schiano, C., La Notte, D., & Pati, S. (2014). What’s now, what’s  
240 new and what’s next in virgin olive oil elaboration systems? A perspective on current knowledge  
241 and future trends. *Journal of Agricultural Engineering*, 45(2), 49-59.

242 Clodoveo, M. L., Durante, V., & La Notte, D. (2013). Working towards the development of  
243 innovative ultrasound equipment for the extraction of virgin olive oil. *Ultrasonics  
244 sonochemistry*, 20(5), 1261-1270.

245 Clodoveo, M. L., & Hbaieb, R. H. (2013). Beyond the traditional virgin olive oil extraction  
246 systems: Searching innovative and sustainable plant engineering solutions. *Food Research*  
247 *International*, 54(2), 1926-1933.

248 Clodoveo, M. L., Hbaieb, R. H., Kotti, F., Mugnozza, G. S. & Gargouri, M. (2014). Mechanical  
249 strategies to increase nutritional and sensory quality of virgin olive oil by modulating the  
250 endogenous enzyme activities. *Comprehensive Reviews in Food Science and Food Safety*, 13,  
251 135-154.

252 Clodoveo, M. L. (2012). Malaxation: Influence on virgin olive oil quality. Past, present and future-  
253 An overview. *Trends in Food Science & Technology*, 25, 13-23.

254 De Luca, M., Terouzi, W., Ioele, G., Kzaiber, F., Oussama, A., Oliverio, F., Tauler, R. & Ragno G.  
255 (2011a). Derivative FTIR spectroscopy for cluster analysis and classification of Morocco olive  
256 oils. *Food Chemistry*, 124, 1113–1118.

257 De Luca, M., Terouzi, W., Bolli, A., Oussama, A., Patumi, M., Ioele, G. & Ragno G. (2011b). A  
258 discriminant method for classification of Moroccan olive varieties by using direct FT-IR analysis  
259 of the mesocarp section. *Vibrational Spectroscopy*, 56, 123–128.

260 De Luca, M., Terouzi, W., Kzaiber, F., Ioele, G., Oussama, A. & Ragno, G. (2012). Classification  
261 of moroccan olive cultivars by linear discriminant analysis applied to atr-ftir spectra of  
262 endocarps. *Food Science and Technology International*, 47, 1286-1292.

263 Forina, M., Oliveri, P., Lanteri, S. & Casale, M. (2008). Class-modeling techniques, classic and  
264 new, for old and new problems. *Chemometrics and Intelligent Laboratory Systems*, 93, 132-148.

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

268 Iñón, F. A., Garrigues, J. M., Garrigues, S., Molina, A. & de la Guardia, M. (2003). Selection of  
269 calibration set samples in determination of olive oil acidity by partial least squares–attenuated  
270 total reflectance-Fourier transform infrared spectroscopy. *Analytical Chimica Acta*. 489, 59–75.

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

274 Maggio, R. M., Kaufman, T. S., Del Carlo, M., Cerretani, L., Bendini, A., Cichelli, A. &  
275 Compagnone, D. (2009). Monitoring of fatty acid composition in virgin olive oil by Fourier  
276 transformed infrared spectroscopy coupled with partial least squares. *Food Chemistry*, 114, 1549-  
277 1554.

278 Ozen, B. F. & Mauer, L. J. (2002). Detection of hazelnut oil adulteration using FT-IR spectroscopy.  
279 *Journal of Agricultural and Food Chemistry*, 50, 3898-3901.

280 Poulli, K. I., Mousdis, G. A. & Georgiou, C. A. (2005). Classification of edible and lampante virgin  
281 olive oil based on synchronous fluorescence and total luminescence spectroscopy. *Analytica*  
282 *Chimica Acta*, 542, 151-156.

283 Restuccia, D., Spizzirri, U. G., Chiricosta, S., Puoci, F., Altimari, I. & Picci, N. (2011). Antioxidant  
284 Properties of EVOO from Cerasuola cv Olive Fruit: Effect of Stone Removal. *Italian Journal of*  
285 *Food Science*, 23, 62-71.

286 Savitzky, A. & Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least  
287 Squares Procedures. *Analytical Chemistry*, 36, 1627–1639.

288 Servili, M., Taticchi, A., Esposto, S., Urbani, S., Selvaggini, R. & Montedoro, G. (2007). Effect of  
289 olive stoning on the volatile and phenolic composition of virgin olive oil. *Journal of Agricultural*  
290 *and Food Chemistry*. 55, 7028-7035.

291 Sinelli, N., Cerretani, L., Di Egidio, V., Bendini A. & Casiraghi, E. (2010). Application of near  
292 (NIR) infrared and mid (MIR) infrared spectroscopy as a rapid tool to classify extra virgin olive  
293 oil on the basis of fruity attribute intensity. *Food Research International*, 43, 369–375.

294 Spizzirri, U. G., Restuccia, D., Parisi, O. I., Cirillo, G., Curcio M., Iemma, F., Puoci, F., Chiricosta,  
295 S. & Picci, N. (2011). Olive stones as source of antioxidants for food industry”, *Journal of Food*  
296 *and Nutrition Research*, 50, 57-67.

297 Tapp, H. S., Defernez, M. & Kemsley, E. K. (2003). FTIR spectroscopy and multivariate analysis  
298 can distinguish the geographic origin of extra virgin olive oils. *Journal of Agricultural and Food*  
299 *Chemistry*, 51, 6110-6115.

300 Wise, B. M., Ricker, N. L., Veltkamp, D. F. & Kowalski, B. R. (1990). A theoretical basis for the  
301 use of principal component models for monitoring multivariate processes. *Process Control and*  
302 *Quality*, 1, 41-51.

303 Wold, S. & Sjöström, M. (1998). Chemometrics, present and future success. *Chemometrics and*  
304 *Intelligent Laboratory Systems*, 44, 3-14.

305

306 **Figure captions**

307

308 Figure. 1. Sample set scheme.

309

310 Figure. 2. Mean FT-IR spectra for each oil brand in SO and WO versions.

311

312 Figure. 3. Score plot (PCs 1 vs 2) by PCA analysis on the calibration set A after Savitzky–Golay  
313 derivative elaboration.

314

315 Fig. 4. Predicted Vs References plots from external validation on the four PLS models.

316

317

Table 1. Prediction of oil extraction procedure by PLS-DA elaboration of FTIR spectra (External Validation)

<i>Oil from stoned olive</i>		
<b>Samples</b>	<b>Reference</b>	<b>Predicted</b>
CER SO	1	0,909
CER WO	0	0,103
DEF SO	1	0,903
DEF WO	0	0,093
SAN SO	1	0,985
SAN WO	0	-0,061
TRI SO	1	0,942
TRI WO	0	0,249

318

319

Table 2. Statistical parameters from leave-1-out cross validation of the PLS models calculated on the calibration sets CER-B, DEF-B, SAN-B and TRI-B

	<i>EVOO brands</i>			
	<b>CER</b>	<b>DEF</b>	<b>SAN</b>	<b>TRI</b>
PCs	2	4	1	2
Slope	0.921	0.925	0.983	0.957
R <sup>2</sup>	0.942	0.973	0.988	0.979
RMSECV	6.043	4.998	3.395	4.206

320

321

322

323

324

325

326



Table 3. Statistical parameters from external validation of the PLS models for CER, DEF, SAN and TRI for prediction of adulteration amount

	$R^2$	<i>RMSEP</i>	<i>RE %</i>
<b>CER</b>	0.981	4.449	5.797
<b>DEF</b>	0.970	4.903	6.914
<b>SAN</b>	0.991	3.351	4.382
<b>TRI</b>	0.986	4.680	5.164

Figure 1  
[Click here to download high resolution image](#)

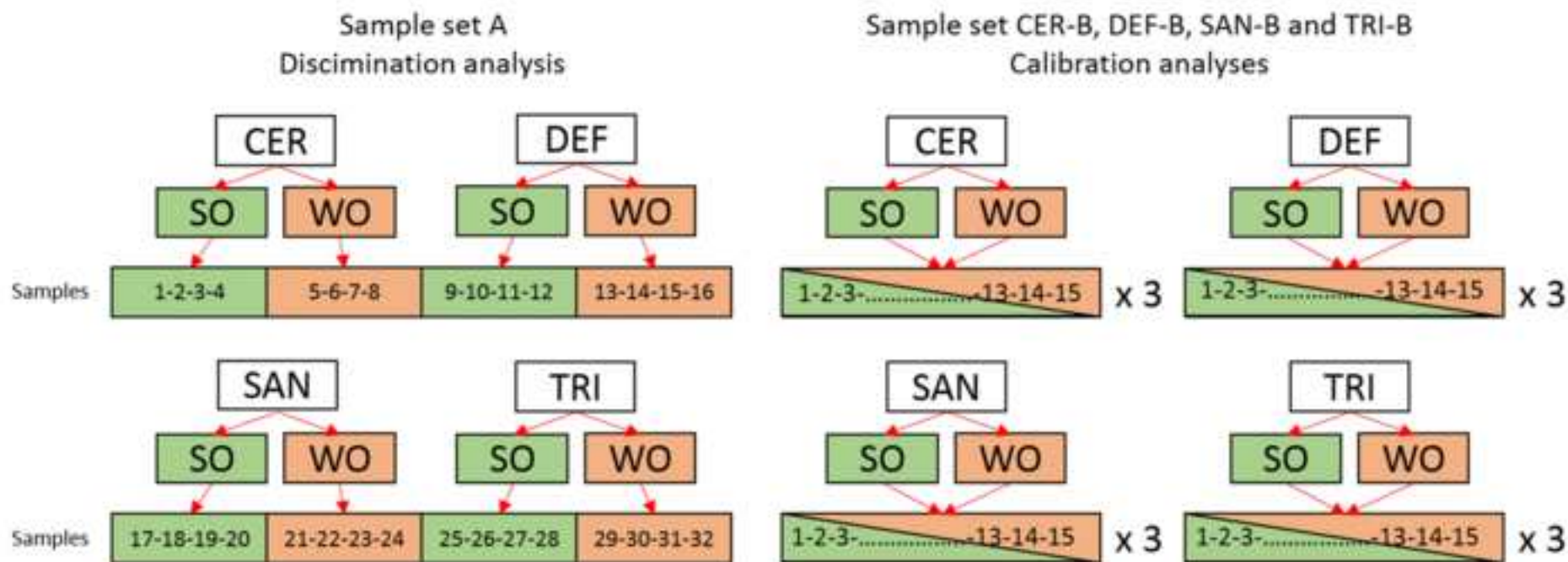


Figure 2  
[Click here to download high resolution image](#)

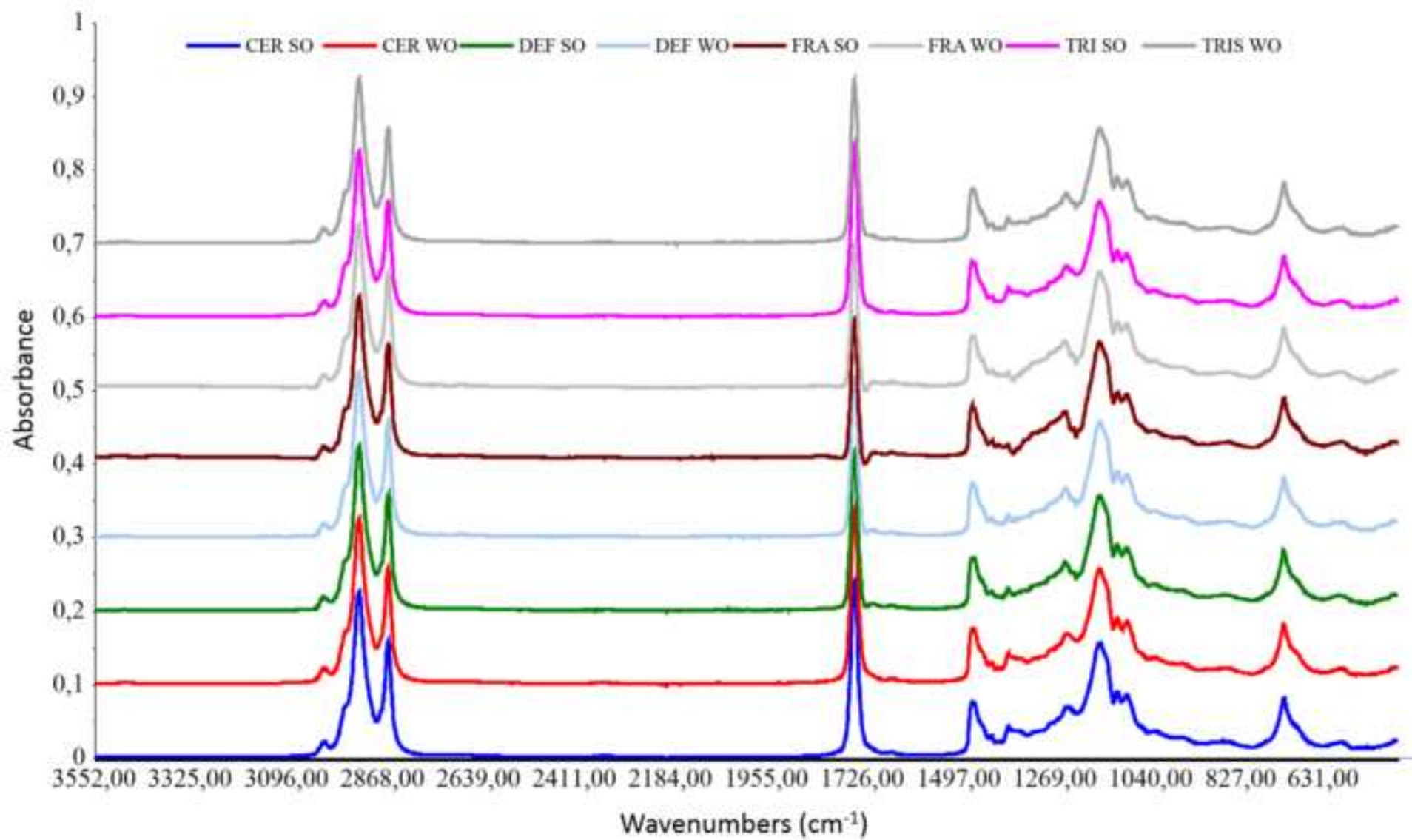


Figure 3  
[Click here to download high resolution image](#)

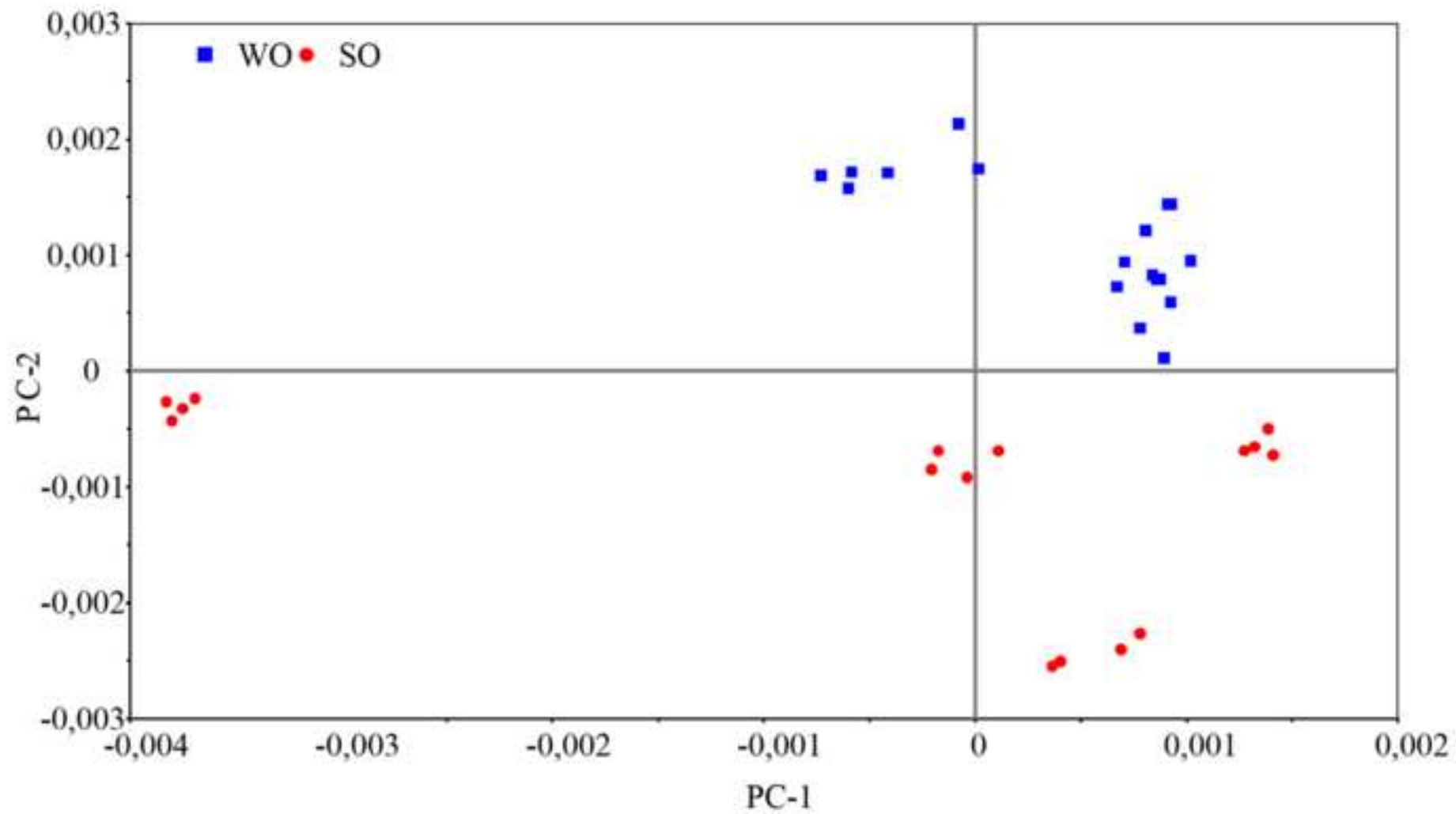


Figure 4  
[Click here to download high resolution image](#)

