Application of NIR and MIR spectroscopy to extra virgin olive oil classification

Casiraghi E.[a], Bendini A.[b], Alamprese C.[a]

[a] Department of Food, Environmental and Nutritional Sciences (DeFENS); Università degli Studi di Milano; Italy.

[b] Department of Agricultural and Food Sciences (DISTAL), Alma Mater Studiorum-Università di Bologna; Cesena, Italy.

Corresponding author: ernestina.casiraghi@unimi.it

Introduction

Extra virgin olive oil (EVOO) plays an important role amongst the Mediterranean typical food products. Its typicality is strongly affected by the origin of raw material and the manufacturing technique. The European Union has recognized many EVOOs as Protected Designation of Origin (PDO) or Protected Geographical Indication (PGI). These typical oils are mainly known for their sensory properties (Caporale et al., 2006), which are deeply influenced by the genetic matrix, that is by the cultivar of the olives used in production (Rotondi et al., 2011). Sensory and chemical properties of EVOO are also affected by the soil type on which the olives are grown, as well as by agronomic practices and climatic conditions (Tovar et al., 2002).

Since EVOO is one of the more expensive vegetable oils, notably in presence of a PDO or PGI label, it is subjected to several commercial frauds. To protect consumers against fraudulent claims, rapid and easy analytical methods are necessary, at least for screening. Spectroscopic techniques, associated with chemometric methods, have been shown to be useful for olive oil classifications (Galtier et al., 2007; Casale et al., 2010; Sinelli et al., 2010; De Luca et al., 2011; Casale et al., 2012).

The aim of this work was to verify the possibility of using NIR and MIR spectroscopy for the classification of EVOOs on the basis of their fruity flavor intensity, geographical origin, and varietal composition.

Materials and Methods

The 245 oil samples used in this study belong to the "Innovative technologies, quality and typicality of Italian virgin oils from seeds and fruits" project (PRIN 2009, granted by the Italian Ministry for Education, University and Research). They were collected during two different Italian national competitions (named 1 and 2), in which oils of different geographical origin and different olive cultivars were judged for their sensory properties. Sensory analysis was carried out according to the method developed by the International Olive Council (IOC, 2007), recognized by the European Community legislation (Commission Regulation (EC) No 1019/2002). In particular, the following levels of the fruity attribute were assigned:

- light: fruitiness intensity median less than 3 in a scale of 10 cm
- medium: fruitiness intensity median between 3 and 6 in a scale of 10 cm
- intense: fruitiness intensity median higher than 6 in a scale of 10 cm

On the basis of the fruitiness intensity and the other information available, samples were grouped in different classes, as reported in Table 1. Only samples derived from competition 1 were used for classification on the basis of the varietal composition (214 samples).

The oil samples were stored in the dark at 4°C until the analyses. Before spectral acquisition, they were conditioned at 25°C for at least 45 min.

Table 1. Class distribution of extra virgin olive oil samples, on the basis of different criteria

Criterion	Fruitiness intensity			Geographical origin				Varietal composition	
Class	LF	MF	IF	N	S	C	I	B	M
Sample no.	88	139	18	24	115	69	37	100	114

LF, light fruitiness intensity; MF, medium fruitiness intensity; IF, intense fruitiness intensity; N, North Italy; S, South Italy; C, Central Italy; I, Italian Isles; B, blend of different cultivars; M, monovarietal oil.

NIR spectra were recorded in transmission mode using vials of 8 mm path length, by means of an FT-NIR spectrometer (MPA, Bruker Optics, Ettlingen, Germany), working in the range 12500-4000 cm^{-1}, with a resolution of 8 cm^{-1} and 16 scans both for background and samples.

MIR analyses were carried out by means of an FT-IR spectrometer (VERTEX 70, Bruker Optics) equipped with a germanium crystal ATR cell. Spectra were acquired in the range 4000-700 cm^{-1}, with 4 cm^{-1} resolution and 16 scans, both for background and samples.

All spectra were collected in duplicate, using the software OPUS v. 6.5 (Bruker Optics). Replicates of spectral data were

averaged, standardized by different pretreatments (MSC or SNV alone or coupled with first or second derivatives), and processed with Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). LDA was applied working with Principal Components (PCs) or with the wavenumbers selected by the algorithm SELECT. Classification models were validated using both cross-validation (5CV) and three different external test sets, randomly created, each containing about 30% of the spectra used for the analysis. Data processing was performed by the software The Unscrambler X (v. 10.2, Camo Software AS, Oslo, Norway) and V-PARVUS (Forina et al., 2008).

Results and discussion

FT-NIR and FT-IR raw spectra of the EVOO samples analyzed are shown in Fig. 1. In order to eliminate non- informative regions, spectral ranges were reduced as follows: 11251-4520 cm^{-1} for FT-NIR; 3700-2600 and 2000-1200 cm^{-1} for FT-IR.
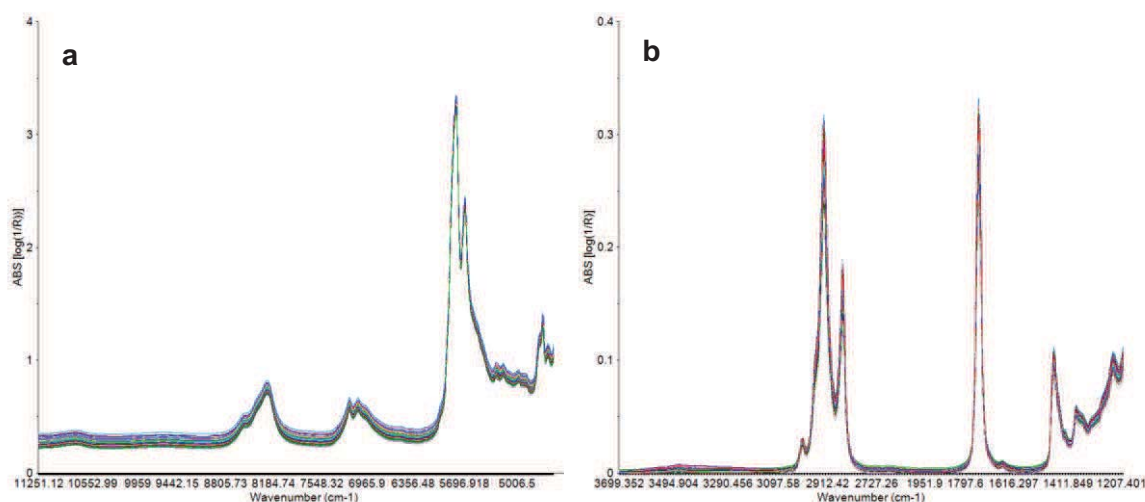

Figure 1. FT-NIR (a) and FT-IR (b) reduced raw spectra of extra virgin olive oil samples.

FT-NIR and FT-IR spectra were characterized by the typical absorption bands of lipids. In particular, in the NIR region, the peaks at 5790 and 5670 cm^{-1} arose from the first overtone of C-H stretching. Bands around 7180 and 7075 cm^{-1} were due to the combination band of C-H, whereas those around 8250 cm^{-1} corresponded to second overtones of C-H stretching vibrations. In the MIR spectra, absorbance around 2924 and 2852 cm^{-1} arose from CH_2 stretching vibrations, asymmetric and symmetric, respectively. The major peak at 1743 cm^{-1} was due to C=O stretching vibrations. The bands around 1463 and 1377 cm^{-1} came from CH2 and CH3 scissoring vibrations (Sinelli et al., 2010).

Principal component analysis (PCA) was used as an exploratory tool, in order to visualize possible data structures. The score plots obtained did not show a good sample distribution according to the criteria of interest. As an example, the score plots obtained from the FT-NIR data pre-treated by SNV are shown in Fig. 2. In the different plots, the sample identification was made on the basis of the three criteria used for class distribution. Similar results were obtained also with FT-IR data.
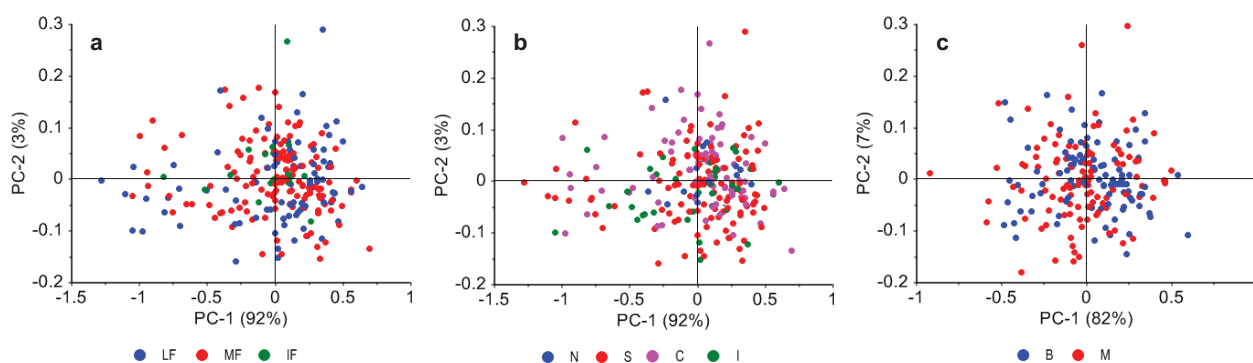

Figure 2. Score plots obtained from FT-NIR data, pre-treated by SNV. Oil sample identification was made on the basis of fruitiness intensity (a), geographical origin (b), and varietal composition (c).

NIR2013 Proceedings, 2-7 June, La Grande-Motte, France. **E1** – Extra Papers
Bellon-Maurel V., Williams P., Downey G., Eds

764

Notwithstanding the poor results obtained by PCA, the use of the Linear Discriminant Analysis (LDA) gave satisfactory classification results. Since for LDA the number of samples must be higher than the number of variables, the analysis was performed both by working in Principal Components (PCs) and by using the 30 variables with the largest classification weight. Feature selection was carried out by means of a stepwise de-correlation technique (algorithm SELECT), implemented in the V-Parvus package (Kowalski & Bender, 1976; Forina et al., 2007). The LDA results were validated using both cross-validation (5CV) and three different external test sets, each composed of 30% of the spectra used for the analysis, randomly selected. For all of the classification criteria, the best LDA models were obtained using the 30 selected wavenumbers of the SNV pre-treated spectra.

The oil classification on the basis of fruitiness was quite successful: the three classes of samples (light, medium, and intense fruity flavor) were correctly classified in prediction with average percentages higher than 67 and 61 for NIR and MIR spectroscopy, respectively, as reported in Table 2.

Table 2. Fruitiness: results of LDA performed on the 30 selected features of SNV pre-treated IR spectra.

	FT-NIR				FT-IR			
	LF	MF	IF	Average	LF	MF	IF	Average
Cross-validation								
Calibration	74.1	80.0	90.0	78.7	72.3	80.0	97.8	78.5
5CV	64.7	71.2	72.2	69.0	64.8	69.1	88.9	69.0
External Set 1								
Calibration	75.0	85.9	95.4	83.2	71.9	78.5	100	77.5
Prediction	75.8	68.3	40.0	69.6	79.0	70.2	60.0	71.8
External Set 2								
Calibration	80.0	82.5	100	82.8	76.7	77.9	96.9	78.9
Prediction	69.6	80.6	83.3	76.9	55.6	76.7	100	70.7
External Set 3								
Calibration	79.1	82.1	94.6	81.8	78.3	82.0	96.9	81.9
Prediction	72.7	64.3	71.4	67.6	46.4	68.0	80.0	61.5

LF, light fruitiness intensity; MF, medium fruitiness intensity; IF, intense fruitiness intensity.

LDA results are also graphically displayed in Fig. 3, showing the sample projection on the first two canonical variables plane. The class that was best distinguished was that of the samples with intense fruitiness, but also the other two classes were quite well separated.
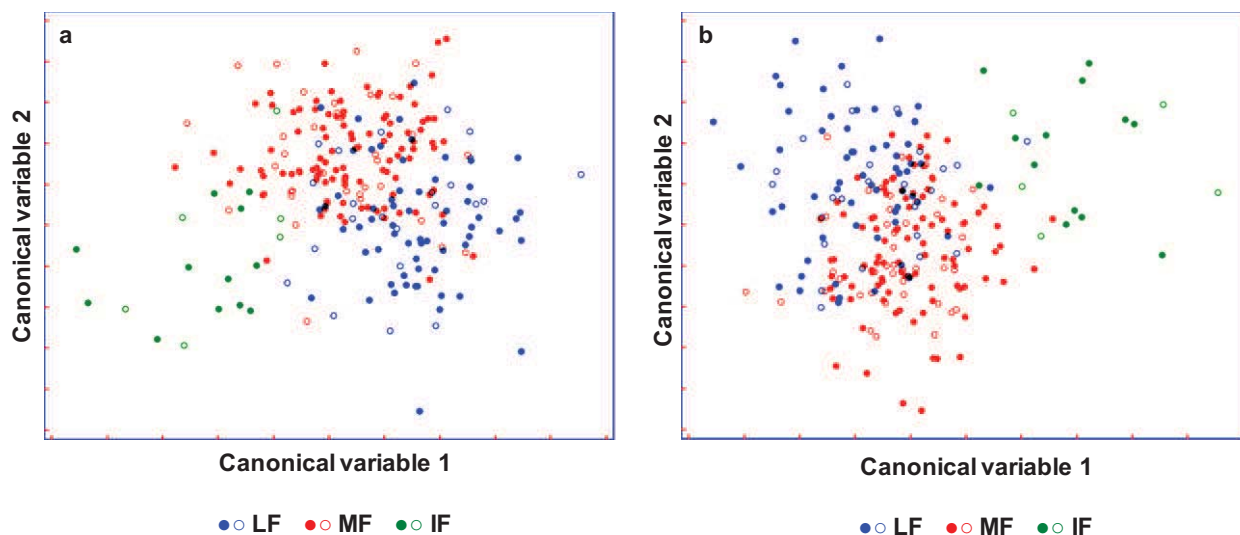

●○ LF ●○ MF ●○ IF

Figure 3. Sample projection on the first two canonical variables of LDA performed on the 30 selected features of SNV pre-treated FT-NIR (a) and FT-IR (b) spectra. Full dots refer to calibration set samples; empty dots refer to external test set samples. LF, light fruitiness intensity; MF, medium fruitiness intensity; IF, intense fruitiness intensity.

With regard to geographical origin, average percentages of correctly classified samples higher than 55 and 64 were obtained in prediction with FT-NIR and FT-IR data, respectively (Table 3).

Table 3. *Geographical origin: results of LDA performed on the 30 selected features of SNV pre-treated IR spectra.*

| | \multicolumn{5}{c}{FT-NIR} | \multicolumn{5}{c}{FT-IR} |
	N	S	C	I	Average	N	S	C	I	Average
Cross-validation										
Calibration	84.2	78.0	79.1	89.2	80.7	79.2	77.0	73.9	91.9	78.6
5CV	62.5	65.2	65.2	83.8	67.8	50.0	64.4	65.2	83.8	66.1
External Set 1										
Calibration	91.8	82.1	84.1	95.8	85.8	87.1	82.0	78.9	90.0	82.6
Prediction	42.9	63.0	40.0	76.9	59.3	28.6	54.6	82.4	84.6	64.3
External Set 2										
Calibration	90.5	78.1	86.5	89.7	83.5	83.2	77.3	77.3	93.8	80.4
Prediction	40.0	60.7	45.0	87.5	57.3	40.0	77.8	58.8	87.5	70.2
External Set 3										
Calibration	81.3	84.8	82.3	95.6	85.4	94.7	73.5	80.4	94.0	81.1
Prediction	75.0	46.0	58.8	70.0	55.6	55.6	68.3	62.5	85.7	67.1

N, North Italy; S, South Italy; C, Central Italy; I, Italian Isles.

As it can be observed also in Fig. 4, the best classified EVOO samples belonged to the Italian isles, whereas the most overlapped classes were those containing samples from the North and the Center of Italy. In this last case, the misclassification could be due to the similarity of some EVOOs coming from areas with analogous pedoclimatic conditions.
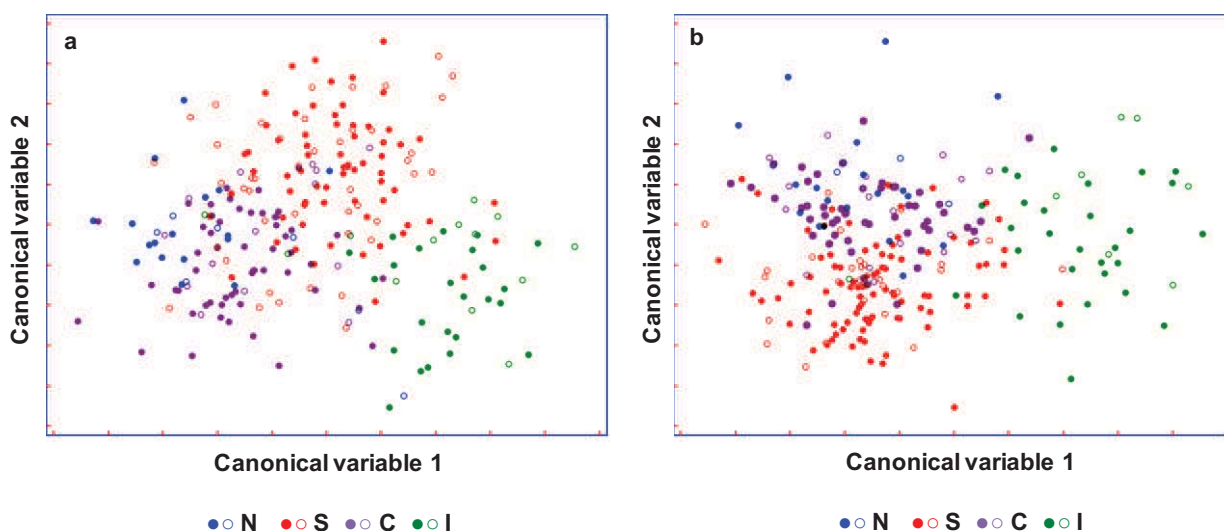

●○ N   ●○ S   ●○ C   ●○ I   ●○ N   ●○ S   ●○ C   ●○ I

Figure 4. Sample projection on the first two canonical variables of LDA performed on the 30 selected features of SNV pre-treated FT-NIR (a) and FT-IR (b) spectra. Full dots refer to calibration set samples; empty dots refer to external test set samples. N, North Italy; S, South Italy; C, Central Italy; I, Italian Isles.

Both the spectroscopy techniques correctly classified the 2 classes of blended and monovarietal oils, with average percentages in prediction ranging from 67 to 75 for FT-NIR and from 66 to 73 for FT-IR (Table 4).

Table 4. Varietal composition: results of LDA performed on the 30 selected features of SNV pre-treated IR spectra.

	FT-NIR			FT-IR		
	B	M	Average	B	M	Average
Cross-validation						
Calibration	80.2	84.1	82.3	82.0	82.0	82.0
5CV	66.0	73.5	70.0	66.0	71.7	69.0
External Set 1						
Calibration	84.4	85.7	85.1	87.7	86.4	87.0
Prediction	66.7	74.4	70.7	60.0	71.4	66.2
External Set 2						
Calibration	84.4	88.1	86.4	85.4	85.7	85.6
Prediction	60.7	87.1	74.6	75.8	58.6	67.7
External Set 3						
Calibration	86.5	87.9	87.2	83.5	85.0	84.2
Prediction	79.3	56.3	67.2	73.7	72.4	72.9

B, blend of different cultivars; M, monovarietal oil.

Conclusion

FT-NIR and FT-IR spectroscopic methods, combined with suitable chemometric strategies, could represent a reliable, cheap and fast classification tool, able to describe EVOO on the basis of intrinsic quality attributes. These attributes, related to the fruitiness intensity, the geographical origin and the varietal composition, are often reported in the product label. Thus IR spectroscopy could be used as an initial screening procedure in EVOO authentication, to protect consumers from misleading claims. Results were encouraging, despite the fact that the data set used was not well balanced in terms of the number of samples of each category.

Acknowledgements

References

Caporale G., Policastro S., Carlucci A., Monteleone E. (2006). Consumer expectations for sensory properties in virgin olive oils. Food Quality and Preference, 17, 116-125.

Casale M., Sinelli N., Oliveri P., Di Egidio V., Lanteri S. (2010). Chemometrical strategies for feature selection and data compression applied to NIR and MIR spectra of extra virgin olive oils for cultivar identification. Talanta, 80, 1832-1837.

Casale M., Oliveri P., Casolino C., Sinelli N., Zunin P., Armanino C., Forina M., Lanteri S. (2012). Characterisation of PDO olive oil Chianti Classico by non-selective (UV–visible, NIR and MIR spectroscopy) and selective (fatty acid composition) analytical techniques. Analytica Chimica Acta, 712, 56-63.

Commission Regulation (EC) No 1019/2002 of 13 June 2002 on marketing standards for olive oil. Official Journal of the European Communities, L155, 27-31.

De Luca M., Terouzi W., Ioele G., Kzaiber F., Oussama A., Oliverio F., Tauler R., Ragno G. (2011). Derivative FTIR spectroscopy for cluster analysis and classification of morocco olive oils. Food Chemistry, 124, 1113-1118.

Forina M., Lanteri S., Armanino C., Casolino C., Casale M., Oliveri P. (2008). V-PARVUS 2008, University of Genova, http://www.parvus.unige.it

Forina M., Lanteri S., Casale M., Cerrato Oliveros M.C. (2007). Stepwise orthogonalization of predictors in classification and regression techniques: an "old" technique revisited. Chemometrics and Intelligent Laboratory Systems, 87, 252-261.

Galtier O., Dupuy N., Le Dreau Y., Ollivier D., Pinatel C., Kister J., Artaud J. (2007). Geographic origins and compositions of virgin olive oils determinate by chemometric analysis of NIR spectra. Analytica Chimica Acta, 595, 136-144.

IOC, 2007. Sensory analysis of olive oil – Method – Organoleptic assessment of virgin olive oil. COI/T.20/Doc. No. 15/2nd Review. International Olive Council, Madrid.

Kowalski B.R., Bender C.F. (1976). An orthogonal feature selection method. Pattern Recognition, 8, 1-4.

Rotondi A., Beghè D., Fabbri A., Ganino T. (2011). Olive oil traceability by means of chemical and sensory analyses: A comparison with SSR biomolecular profiles. Food Chemistry, 129, 1825-1831.

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

Tovar M.J., Romero M.P., Alegre S., Girona J., Motilva M.J. (2002). Composition and organoleptic characteristics of Arbequina olive oil from olive trees (Olea Europaea L.) under deficitary irrigation. Journal of the Science of Food and Agriculture, 82, 1755-1763.

NIR2013 Proceedings, 2-7 June, La Grande-Motte, France. **E1** – Extra Papers
Bellon-Maurel V., Williams P., Downey G., Eds

767