JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Confirmation of Declared Provenance of European Extra Virgin Olive Oil Samples by NIR Spectroscopy

TONY WOODCOCK,*,†,§ GERARD DOWNEY,§ AND COLM P. O'DONNELL†

Biosystems Engineering, School of Agriculture, Food Science and Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland, and Teagasc, Ashtown Food Research Centre, Ashtown, Dublin 15, Ireland

The potential of near-infrared transflectance spectroscopy (1100−2498 nm) combined with chemometric techniques to confirm the geographical origin of European olive oil samples was evaluated. In total, 913 extra virgin olive oil samples (210 Ligurian and 703 non-Ligurian) were collected over three consecutive harvests (2005, 2006, and 2007). A multivariate spectral fingerprint for Ligurian olive oil was developed and deployed to confirm or refute a claim that any given sample was Ligurian. Samples were pseudorandomly split into calibration ($n = 280$) and validation sets ($n = 633$); the only selection constraint applied was to insist on equal numbers of Ligurian and non-Ligurian samples in the calibration set. Following preliminary examination by principal component analysis, the full spectrum modeling method applied to the spectral data set was discriminant partial least-squares regression; various data pretreatments were also investigated. The best models correctly predicted the origins of samples in the prediction set up to 92.8 and 81.5% for Ligurian and non-Ligurian olive oil samples, respectively, using a first-derivative data pretreatment. The potential of this approach in commercial traceability and quality assurance schemes is noted.

KEYWORDS: Near-infrared; spectroscopy; olive oil; authenticity; geographical origin; classification

INTRODUCTION

Olive oil is obtained from the fruit of the olive tree (*Olea europaea* L.) and is a genuine fruit juice with excellent nutritional, sensory, and functional qualities (*1*). Consumer interest in olive oil as a healthy food source has increased due to its polyunsaturated fatty acid composition and content of other functional food components (*2*). The main phenolics in olive oil include hydroxytyrosol, tyrosol, and oleuropein, which occur in the highest levels in virgin olive oil and have demonstrated antioxidant activity (*3*). Oleic acid, a monounsaturated fatty acid, and squalene have been identified as having anticancer effects; olive oil consumption has also been linked with benefits for colon and breast cancer prevention and with its ability to reduce blood pressure and low-density lipoprotein (LDL) cholesterol (*3*). Virgin olive oils are the oils obtained solely by mechanical or other physical means under conditions, particularly thermal conditions, that do not lead to alterations in the oil and which have not undergone any treatment other than washing, decanting, centrifugation, and filtration (*3*). Extra virgin olive oil is virgin olive oil that has a free acidity, expressed as oleic acid, of not more than 0.8 g/100 g, and the other characteristics of which correspond to those specified for this category (*4*). This type of oil is of limited production and viewed as high quality; as a result, extra virgin olive oil is one of the more expensive vegetable oils (*5*). An additional quality factor associated with such oils in the mind of the consumer is the geographical area of production, or provenance; oils from certain regions are viewed as superior in quality to others (*6*). To clearly label high-quality food products, the European Union (EU) has created specific terms governing provenance claims such as protected designation of origin (PDO); PDO is a term used to describe foodstuffs that are produced, processed, and prepared in specific, controlled, and limited geographic areas using recognized methods (*7*). Under EU law, only 85 geographical areas are permitted to use a PDO label for their olive oil—15 Greek, 19 Spanish, 7 French, 37 Italian, 1 Slovenian, and 6 Portuguese (*8*). Ligurian olive oil (Riviera Ligure) is one such PDO category. These facts coupled with the quality aspects associated with extra virgin olive oils in general mean that such oils may be particularly susceptible to economic fraud. Specifically with regard to provenance, this would involve false claims of origin on product labels. The consumer and food industry in general need protection from such fraudulent labeling.

The combination of near-infrared (NIR) spectroscopy and chemometric data analysis techniques provides a powerful set of tools for quantitative and qualitative modeling of a wide variety of foodstuffs and food production processes (*9*). Recent reported applications of NIR spectroscopy in edible oil analysis include adulteration detection (*10, 11*), geographical origin prediction (*12*), quality parameter determination (*13, 14*), classification (*15, 16*), and online monitoring of carotenoid and chlorophyll pigments (*17*). NIR spectroscopy facilitates real-

* Corresponding author (e-mail tony.woodcock@teagasc.ie).
† University College Dublin.
§ Teagasc.

10.1021/jf802792d CCC: $40.75
Published on Web 11/11/2008

Table 1. Sources of Olive Oil Samples

	Italy							
harvest	Liguria	non-Liguria	Greece	Spain	France	Cyprus	Turkey	total
2005	79	173	46	38	10	6	0	352
2006	63	163	25	42	9	0	14	316
2007	68	116	7	34	20	0	0	245
total	210	452	78	114	39	6	14	913

time measurements at all stages of production from raw material analysis to ingredient and finished product verification; it offers a fast, nondestructive, and cost-effective method of food analysis (*18*).

The objective of this paper is to report investigations on the potential of NIR spectroscopy and chemometric techniques to confirm a specific provenance claim relating to olive oils, that is, that a sample claimed to originate in Liguria actually does so. The approach used involved collection of authentic olive oils from this and other important olive oil producing regions of Europe over a number of production harvests.

MATERIALS AND METHODS

Sample Collection and Preparation. Olive oil samples were collected from a number of different areas in Europe as part of the EU-funded TRACE project (*19*) over a period spanning three harvests—2005, 2006, and 2007; all olive oil samples were analyzed shortly after harvest. Numbers and sources of samples in each of these years are shown in **Table 1**. All oils were transported to a single laboratory (Joint Research Centre, Ispra, Italy) for subsampling and delivery by air to Ashtown Food Research Centre. Oils from harvest 2 (2006) were collected and distributed in two separate batches. All olive oils were stored in a refrigerated room (4 °C) between delivery and spectral acquisition (<2 weeks), minimizing the chance of any significant change occurring during this period.

Spectral Collection. Olive oil samples (approximately 50 mL) were placed in screw-capped glass vials in a water bath maintained at 30 °C and allowed to equilibrate for 30 min before spectral acquisition. Transflectance spectra (1100−2498 nm) were collected using a camlock cell and a gold-plated reflector (0.1 mm sample thickness; part 99213) on a scanning spectrophotometer (NIR Systems 6500, NIR Systems Inc., Silver Spring MD). Between samples, this cell was cleaned using tepid distilled water, then Triton X-100 solution (0.1% w/w), and finally rinsed with tepid distilled water before drying with a paper tissue; all washing liquids were equilibrated at 30 °C. Spectra were recorded in triplicate for each sample, and the mean of these replicates was used in subsequent calculations. WINISI software (v 1.05; ISI International, Port Matilda, PA) was used for spectrophotometer control and spectral file manipulation. Spectra were exported from WINISI in JCAMP.DX format (*20*) and imported into The Unscrambler (v. 9.2; CAMO A/S, Oslo, Norway) for statistical analysis.

Statistical Analysis. Principal component analysis (PCA) was initially carried out on the raw spectral data to (a) assist with the detection of any outlying or unusual samples and (b) investigate any possible clustering of samples on the basis of their provenance. Development of class models to identify samples as originating in Liguria or not was performed by partial least-squares discriminant analysis (PLS-DA). For this purpose, a dummy *Y*-variable was assigned to each oil sample, 1 for Ligurian and 0 for non-Ligurian olive oils. To determine the univariate specification, a cutoff point of 0.5 was assigned to the resulting PLS predictions; samples predicted as having a value of ≥0.5 were classified by the model as being Ligurian, and all other samples were classified as being non-Ligurian. PLS1 calibration models were initially developed using the complete data set with full, that is, leave-one-out cross-validation. Following this, the sample set was split up into two groups with two-thirds of the total samples being used as a calibration set and the remaining one-third of samples acting as a validation or test set, for the calibration models. To remove or at least minimize unwanted spectral contributions arising, for example, from light scatter (*9*), first- and second-derivative pretreatments using the Savitzky−Golay (*21, 22*) method and the standard normal variate (SNV) transform (*23*) were investigated for all models. In this way, 10 models (raw, first derivative with 5, 9, 13, and 21 points, second derivative with 5, 9, 13, and 21 points, and SNV) were created for each technique, as can be seen in the tables.

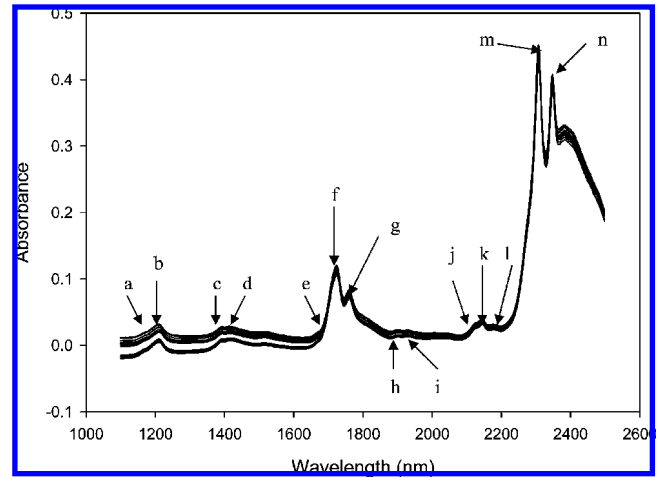

Figure 1. NIR (1100−2498 nm) transflectance spectra of a random selection of olive oils from Liguria and non-Ligurian regions over three harvests (see **Table 2** for explanation of symbols).

Given the imbalance between Ligurian and non-Ligurian samples in the olive oil collection, calibration models developed as stated above are, in statistical terms, unbalanced. Balanced models were then developed using equal numbers of Ligurian and non-Ligurian olive oil samples in the calibration set; two-thirds of all the Ligurian samples were chosen at random as was an equal number of non-Ligurian samples. Calibrations thus developed were validated on the remaining oil samples. The imbalance between Ligurian and non-Ligurian oils in this sample set poses no statistical difficulty and may be considered to be representative of sample testing in the future.

Finally, in an effort to enhance model robustness and specificity, a variable selection algorithm, the jack-knife uncertainty test (*24*), was applied both to the original data set and to the calibration set containing equal numbers of Ligurian and non-Ligurian samples. This test was used to avoid misinterpreting spurious effects, to identify the dominating sources of instability in the modeling, and to allow more or less automatic optimization of the models (*24*). It aims to predict the most important *X*-variables (wavelengths) for model development, and after their identification, models were recreated using only the wavelength regions indicated. An analysis of misclassified Ligurian olive oil samples was carried out at this point in an effort to compare models created using the original data and models created after application of this variable selection algorithm.

Classification results are presented in terms of percent correct classification, percent misclassification, sensitivity, and specificity. Sensitivity is the probability that a given model will classify a test sample as positive given that it is known to be positive, that is, the probability that an authentic Ligurian olive oil sample will be correctly identified as originating in Liguria. Sensitivity is the ratio number of predicted positive classifications to total number of actual positives (*25*). Model specificity is the probability that a model will classify a test sample as not belonging to the model given that it is known not to belong. In this case, it is the probability that an olive oil sample from a region outside Liguria is classified correctly as being non-Ligurian in origin. Specificity is calculated as the ratio of the number of predicted negative classifications to the number of known negatives (*25*).

RESULTS AND DISCUSSION

Data Examination. Spectra of randomly selected samples, both Ligurian and non-Ligurian, representing each of the three harvests are shown in **Figure 1**. The main peak locations are marked in **Figure 1**, and the chemical parameters to which they

Table 2. Functional Group Assignment[a] in Olive Oil Spectra (See **Figure 1**)

notation	wavelength (nm)	functional group	assignment
a	1168	CH_3-	C—H stretch second overtone
b	1211	$-CH_2-$	C—H stretch second overtone
c	1391	CH_3-	2C—H str + C—Hdef
d	1414	$-CH_2-$	2C—H str + C—Hdef
e	1664	*cis* $R_1CH{=}CHR_2CH_3-$	*cis* CH
f	1727	$-CH_2-$	C—H first overtone
g	1761		C—H first overtone
h	1901	C=O str	second overtone
i	1931	C=O str	second overtone ester
j	2124	−COOR	C—H str + C=O str
k	2145	$-HC{=}CH-$	=CH str+ C=O str
l	2176	$-HC{=}CH-$	CH assym str + C=C str
m	2310		CH combinations and deformation
n	2350		CH combinations and deformation

[a] Functional groups assigned per refs 26−29.

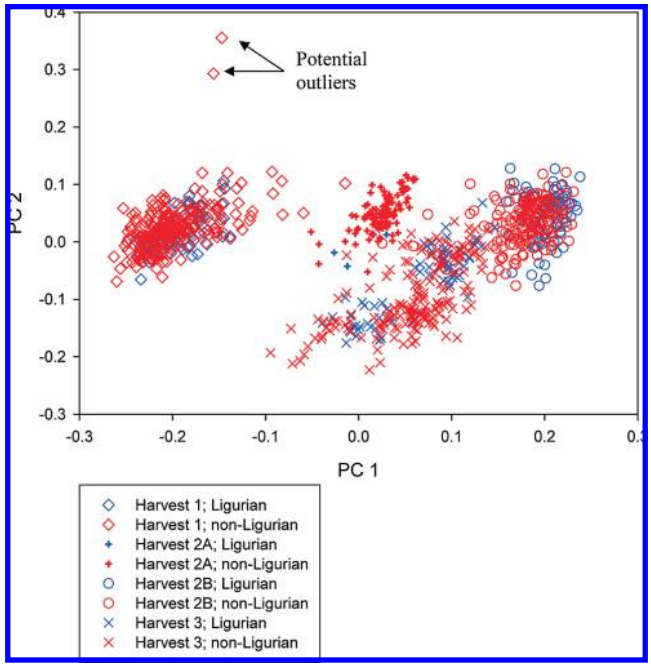

Figure 2. PCA scores plot (complete sample data set; raw data; 1100−2498 nm).

relate are outlined in **Table 2** (*26, 27*). The most important absorption maxima are clearly evident at 1211, 1727, 1761, 2310, and 2350 nm. Bands around 1211 nm arise from second overtones of C−H stretching vibrations (*28*), whereas those at 1727 and 1761 nm are attributed to the first overtone of C−H stretching vibrations of methyl, methylene, and ethylene groups (*10*). Absorbances at 2310 and 2350 nm arise from combination bands arising from C−H stretching vibrations and other vibrational modes (*28*).

As part of initial data examination, PCA was performed on the complete, raw spectral data set. **Figure 2** is the most informative scores plot available, showing a good deal of clustering; PC1 (accounting for 82% of the variation) is plotted against PC2 (accounting for 16% of the variation), revealing some clustering behavior that relates to production harvest or scanning time. The former phenomenon may be seen in the separation of samples from harvests 1 and 3, whereas the separation of samples from harvest 2 into two groups may be related to the latter. It is unclear what underlies this split distribution of harvest 2 samples—it could involve issues related to sampling, storage, climate, chemical composition, or analysis.

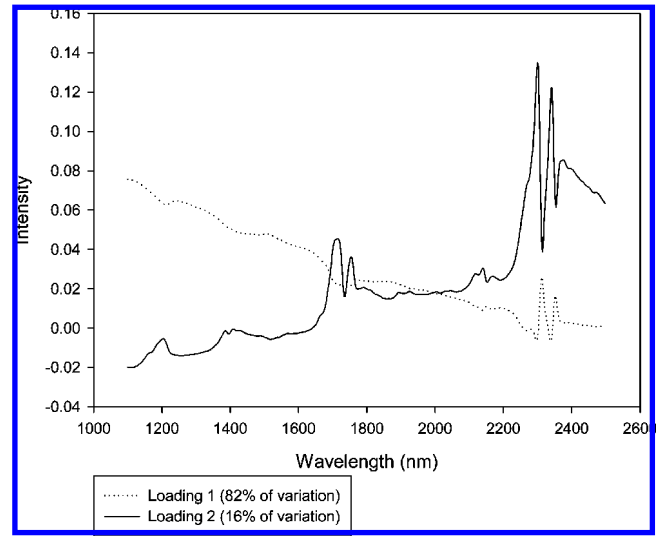

Figure 3. PCA loadings plot.

In these two dimensions, the main separation (along PC1) relates to spectral differences between oils from harvest 1 and those from harvest 2. Such seasonal differences often originate in weather conditions during growth and harvesting, and anecdotal evidence of major differences in these conditions in 2005, 2006, and 2007 was supplied by oil sample collectors. A plot showing the PCA loadings for PC1 and PC2 is shown in **Figure 3**. Two potential outlier samples (data points that are numerically distant from the rest of the data but for which there is no other supporting information to merit their deletion) that originated in harvest 1 are marked in **Figure 2**; these samples also show a very high leverage in the residual *X*-variance plot (not shown). Particular attention was paid to these possible outliers in subsequent work, but they were not excluded from PLS-DA regression analysis as their removal (a) did not affect the spatial distribution of the other samples and (b) did not adversely affect the corresponding prediction models. The scores plot showing PC3 against PC1 revealed the existence of the same two outliers, this time along PC3; they do not show up as being obvious outliers along any other PC. Regions of interest in the *X*-loadings plots for PC2 and PC3 are 1724, 1766, 2304, and 2342 nm, which suggests that a combination of the functional groups associated with these spectral locations, all of which were discussed in relation to **Figure 1**, could be responsible for these outliers. However, all of these spectral locations are also shown to be important in the *X*-loading plot of PC1, which means that, although they may be responsible for the location of the outliers, they are also mainly responsible for the separation shown along PC1 between olive oil samples from different harvests.

Further examination of **Figure 2** reveals some possible clustering within the samples from harvest 3 in this two-dimensional principal component space. It is not possible to advance any explanation for this behavior on the basis of the sample information supplied. The color scheme in **Figure 2** shows Ligurian samples in blue and non-Ligurian samples in red; it is apparent that there is complete overlap between Ligurian and non-Ligurian samples in these two principal component dimensions.

PLS-DA Regression. PLS-DA classification models were initially developed on the complete sample data set ($n = 913$ olive oils) using complete (1100−2498 nm) spectral data and full (leave-one-out) cross-validation; summary results are presented in **Table 3** as percent correct classifications (percent correct classification of Ligurian samples relating to sensitivity results and percent correct classification of non-Ligurian samples

Table 3. PLS-DA Carried out on the Full Dataset ($n = 913$) Using Full (Leave-One-Out) Cross-Validation

		% correct classification	
data pretreatment	#L	Ligurian	non-Ligurian
raw data	8	32.9	95.0
first deriv, 5 points	8	42.9	95.7
first deriv, 9 points	9	44.3	95.6
first deriv, 13 points	8	36.2	95.7
first deriv, 21 points	9	40	95.3
second deriv, 5 points	7	49.5	95.6
second deriv, 9 points	7	43.8	96.4
second deriv, 13 points	8	47.1	96.0
second deriv, 21 points	6	42.4	95.9
SNV	7	31.9	94.2

Table 4. PLS-DA Results Using Equal Numbers of Ligurian (140) and Non-Ligurian (140) Olive Oil Samples in the Calibration Set and Subsequent Validation Models Using the Remaining Samples (70 Ligurian and 563 Non-Ligurian)

		% correct classification	
data pretreatment	#L	Ligurian	non-Ligurian
raw data	8	85.5	78.7
first deriv, 5 points	8	88.4	80.6
first deriv, 9 points	8	88.4	80.6
first deriv, 13 points	8	92.8	81.5
first deriv, 21 points	9	91.3	77.4
second deriv, 5 points	9	92.8	79.4
second deriv, 9 points	9	91.3	79.8
second deriv, 13 points	8	88.4	79.4
second deriv, 21 points	7	87.0	81.2
SNV	6	87.0	77.4

relating to specificity results). A number of observations may be made on these results: (a) all discriminant models developed involved between 6 and 9 PLS loadings with 7 and 8 being the most common numbers; (b) percentage correct classification results for Ligurian were low, ranging from a minimum of 31.9% obtained with SNV-treated data up to a maximum of 49.5% in the case of a second-derivative transform (20 data point gap). Therefore, in most cases, model sensitivity was low, whereas specificity was high, >0.90. This low sensitivity is likely to arise from the obvious imbalance in the numbers of Ligurian ($n = 210$) and non-Ligurian ($n = 703$) samples in the oil collection.

Separate Validation Sample Set. In an effort to improve the performance of models described above, a different strategy was adopted using equal numbers of Ligurian and non-Ligurian samples in the calibration set ($n = 280$), as described under Statistical Analysis. All of the remaining samples (70 Ligurian and 563 non-Ligurian) were used as a validation sample set. The results of this modeling strategy for the samples in the validation set are shown in **Table 4** and, when the average correct classification results for both the Ligurian and non-Ligurian samples are considered, demonstrate a significant improvement over those obtained initially (**Table 3**). Most models involve eight PLS loadings, whereas the percent correct classification for Ligurian and non-Ligurian samples span the ranges of 88.4−92.8 and 77.4−81.5 respectively. The most accurate model overall involved a first-derivative spectral pretreatment (13 data point gap) and produced correct classification rates of 92.8 and 81.5% for Ligurian and non-Ligurian oils, respectively.

Reduced Wavelength Numbers. The jack-knife uncertainty test was applied to the regression models reported above with the aim of removing uninformative *X*-variables and potentially

Table 5. PLS-DA Results Using Equal Numbers of Ligurian (140) and Non-Ligurian (140) Olive Oil Samples in the Calibration Set and Subsequent Validation Models Using the Remaining Samples (70 Ligurian and 563 Non-Ligurian) Having Subjected the Original Data Set to the Uncertainty Variable Selection Algorithm

		% correct classification	
data pretreatment	#L	Ligurian	non-Ligurian
raw data	6	88.4	77.6
first deriv, 5 points	6	87	80.5
first deriv, 9 points	8	88.4	79.9
first deriv, 13 points	7	92.8	78.7
first deriv, 21 points	7	91.3	80.3
second deriv, 5 points	7	91.3	75.3
second deriv, 9 points	7	92.8	79.4
second deriv, 13 points	8	85.5	77.3
second deriv, 21 points	6	91.3	78.2
SNV	6	85.5	77.6

improving model sensitivity and specificity. Classification results are shown in **Table 5**; the main observations which may be made about this summary table are that sensitivity and specificity are very close to those obtained using the full *X*-variable data set, whereas a small reduction in the number of loadings required by each model is apparent. Therefore, there has been little or no significant loss or degradation of model performance and model simplicity has improved. Although very small variations may be observed, predicted *Y*-values remain almost the same for each individual olive oil sample when the jack-knifing is carried out as had been prior to application of the uncertainty test. This procedure reduced the number of variables used in model development from 700 in the original full spectrum to an average of 230 *X*-variables for the 10 attenuated models. As a result, these latter models may have enhanced robustness over the full spectrum variants given the reduction in the required number of loadings, an advantage that could have commercial significance.

Misclassified Olive Oil Samples. A method to analyze the misclassified samples was developed; of the 10 models that were created for each technique, samples which were misclassified by 7 or more of these models were selected as being misclassified. This analysis was carried out for the models created using equal numbers of Ligurian and non-Ligurian samples for the calibration set and validated using a separated sample set that consisted of 70 Ligurian samples and 563 non-Ligurian samples. Of the 70 Ligurian samples (23 from harvest 1, 24 from harvest 2, and 23 from harvest 3), 5 samples were misclassified by 7 or more models before the uncertainty test was applied and the same 5 samples were similarly misclassified by 7 or more models after the application of the uncertainty test. These 5 samples contained 3 samples from harvest 1 and 2 samples from harvest 2; no Ligurian samples from harvest 3 were misclassified by 7 or more models. The 3 samples that originated in the Liguria region from harvest 1 were non-PDO olive oils. More information is known about the samples from harvest 2; both are PDO olive oils, one is simply labeled "extra virgin" and the other, "Riviera Ligure". Both of these oils originated in the Savona province, the first "extra virgin" oil from Andora and the second "Riviera Ligure" oil from Balestrino. It is difficult to claim any pattern exists in the data with any degree of certainty because the number of misclassified samples is so small.

Of the 563 non-Ligurian samples in the validation set (204 from harvest 1, 220 from harvest 2, and 139 from harvest 3, representing each of the countries shown in **Table 1**) 93 samples were misclassified by 7 or more models prior to the

application of the uncertainty test: 20 from harvest 1, 35 from harvest 2, and 38 from harvest 3. The majority of these 93 misclassified samples (76 in all) were Italian samples, which is to be expected both because there are far more Italian samples in the sample set, because the regions are geographically more similar to Liguria, and because they may use the same or related olives in the production of these oils. There were also 8 Spanish, 7 Greek, and 2 French samples misclassified. There is no clearly discernible pattern in the misclassified Italian samples, with the oils originating in Puglia ($n = 25$), Campania ($n = 10$), Veneto ($n = 8$), Lazio ($n = 6$), Toscana ($n = 6$), Umbria ($n = 5$), Calabria ($n = 3$), Trentino Alto Adige ($n = 3$), Marche ($n = 3$), Molise ($n = 2$), Sicilia ($n = 2$), Emilia Romagna ($n = 2$), and Abruzzo ($n = 1$). Expressed as a percentage of the relevant validation set for each region, these are 41.7, 41.7, 30.8, 13.0, 54.6, 12.2, 15.0, 50.0, 37.5, 8.0, 5.1, 33.3, and 4.8%, respectively. The physical distance between regions does not seem to be the factor causing confusion for the models; within Italy, the regions Toscana, Trentino Alto Adige, Emilia Romagna, and Veneto are considerably closer geographically to Liguria than Puglia or Campagnia. However, the former group does not as a rule have higher misclassification rates than the latter. One fact which is clear is that, as was the case for the Ligurian olive oils, the application of the uncertainty test did not alter greatly the samples which were misclassified; following the application of the uncertainty test 90 non-Ligurian samples were misclassified by 7 or more models, of which 85 overlapped with the 93 misclassified prior to the application of the uncertainty test. This suggests that whether the uncertainty test was applied or not, it is roughly the same samples which are being misclassified by the models. It can thus be concluded that applying the uncertainty test does not cause the models to lose much, if any, important information.

Overall, it may be concluded that NIR spectroscopy coupled with chemometric analysis of data provides a promising tool for geographical classification of olive oils. Best results of 0.93 sensitivity and 0.82 specificity were achieved for PLS-DA models developed using equal numbers of Ligurian and non-Ligurian samples in calibration sets and validated using a separate sample set. These levels of accuracy may be sufficient for sample screening purposes. Models created using considerably more non-Ligurian samples than Ligurian samples were biased toward the non-Ligurian samples, and the predictive capacities of these models were consequently relatively poor.

The application of the uncertainty variable selection algorithm made no improvement for the PLS-DA models developed using equal numbers of Ligurian and non-Ligurian olive oil samples in terms of predictive capacity. The misclassified rate of Ligurian samples was roughly the same for models created with or without the uncertainty test, proving that the elimination of sections of the NIR spectra did not decrease the prediction capacity of the models developed and that no important information was lost due to the application of this technique.

Transition of this technique into an industrial setting would require the establishment of a larger database that would take into account greater variability in factors such as weather conditions at the time of harvest. Given that NIR spectroscopy is a fast and relatively cost-effective method, it would be an appropriate technique for deployment as a screening method for dealing with large numbers of samples quickly and economically. For 100% accuracy, samples that are misclassified by NIR techniques would require further confirmatory analysis using established wet chemistry techniques.

LITERATURE CITED

(1) Matos, L. C.; Cunha, S. C.; Amaral, J. S.; Pereira, J. A.; Andrade, P. B.; Seabra, R. M.; Oliveira, B. P. P. Chemometric characterisation of three varietal olive oils (Cvs. Cobrançosa, Madural and Verdeal Transmontana) extracted from olives with different maturation indices. *Food Chem.* **2007**, *102*, 406–414.

(2) Temime, S. B.; Manai, H.; Methenni, K.; Baccouri, B.; Abaza, L.; Daoud, D.; Casas, J. S.; Bueno, E. O.; Zarrouk, M. Sterolic composition of Chétoui virgin olive oil: influence of geographical origin. *Food Chem.* **2008**, *110*, 368–374.

(3) Waterman, E.; Lockwood, B. Active components and clinical applications of olive oil. *Alternative Med. Rev.* **2007**, *12* (4), 331–342.

(4) Trade standard applying to olive oils and olive-pomace oils, International Olive Council, COI/T.15/NC no. 3/Rev. 2, 24 Nov 2006.

(5) Galtier, O.; Dupuy, N.; Le Dréau, Y.; Ollivier, D.; Pinatel, C.; Kister, J.; Artaud, J. Geographic origins and compositions of virgin olive oils determinated by chemometric analysis of NIR spectra. *Anal. Chim. Acta* **2007**, *595*, 136–144.

(6) Tovar, M. J.; Romero, M. P.; Alegre, S.; Girona, J.; Motilva, M. J. Composition and organoleptic characteristics of oil from Arbequina olive (*Olea europaea* L.) trees under deficit irrigation. *J. Sci. Food Agric.* **2002**, *82* (15), 1755–1763.

(7) Luykx, D. M. A. M.; van Ruth, S. M. An overview of analytical methods for determining the geographical origin of food products. *Anal. Methods* **2008**, *107*, 897–911.

(8) http://ec.europa.eu/agriculture/qual/en/pgi_11en.htm.

(9) Blanco, M.; Pagès, J. Classification and quantification of finishing oils by near infrared spectroscopy. *Anal. Chim. Acta* **2002**, *463*, 295–303.

(10) Downey, G.; McIntyre, P.; Davies, A. N. Detecting and quantifying sunflower oil adulteration in extra virgin olive oils from the eastern Mediterranean by visible and near-infrared spectroscopy. *J. Agric. Food Chem.* **2002**, *50*, 5520–5525.

(11) Wang, L.; Lee, F. S. C.; Wang, X.; He, Y. Feasibility study of quantifying and discriminating soybean oil adulteration in camellia oils by attenuated total reflectance MIR and fibre optic diffuse reflectance NIR. *Food Chem.* **2006**, *95*, 529–536.

(12) Galtier, O.; Dupuy, N.; Le Dréau, Y.; Ollivier, D.; Pinatel, C.; Kister, J.; Artaud, J. Geographic origins and compositions of virgin olive oils determinated by chemometric analysis of NIR spectra. *Anal. Chim. Acta* **2007**, *595*, 136–144.

(13) Armenta, S.; Garrigues, S.; de la Guardia, M. Determination of edible oil parameters by near infrared spectrometry. *Anal. Chim. Acta* **2007**, *596*, 330–337.

(14) Jiménez Marquez, A.; Molina Diaz, A.; Pascual Reguera, M. I. Using optical NIR sensor for on-line virgin olive oils characterization. *Sensors Actuators B* **2005**, *107*, 64–68.

(15) Hourant, P.; Baeten, V.; Morales, M. T.; Meurens, M.; Aparicio, R. Oil and fat classification by selected bands of near-infrared spectroscopy. *Appl. Spectrosc.* **2000**, *54*, 1168–1174.

(16) Yang, H.; Irudayaraj, J.; Paradkar, M. M. Discriminant analysis of edible oils and fats by FTIR, FT-NIR and FT-Raman spectroscopy. *Food Chem.* **2005**, *93*, 25–32.

(17) Jiméne Marquez, A. Monitoring carotenoid and chlorophyll pigments in virgin olive oil by visible-near infrared transmittance spectroscopy. On-line application. *J. Near Infrared Spectrosc.* **2003**, *11*, 219–226.

(18) Woodcock, T.; Downey, G.; O'Donnell, C. P. Better quality food and beverages: the role of near infrared spectroscopy. *J. Near Infrared Spectrosc.* **2008**, *16*, 1–29.

(19) EU-funded TRACE project: http://www.trace.eu.org/.

(20) Rutledge, D. N.; McIntyre, P. A proposed European implementation of the JCAMP-DX data file transfer format. *Chemom. Intell. Lab. Syst.* **1992**, *16*, 95–101.

(21) Alciaturi, C. E.; Escobar, M. E.; De La Cruz, C. A numerical procedure for curve fitting of noisy infrared spectra. *Anal. Chim. Acta* **1998**, *376*, 169–181.
(22) Vaiphasa, C. Consideration of smoothing techniques for hyperspectral remote sensing. *ISPRS J. Photogrammetry Remote Sensing* **2006**, *60*, 91–99.
(23) Barnes, R. J.; Dhanoa, M. S.; Lister, S. J. Standard normal variate transformation and de-trending of near infrared diffuse reflectance spectra. *Appl. Spectrosc.* **1989**, *43*, 772–777.
(24) Martens, H.; Martens, M. Modified jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Qual. Pref.* **2000**, *11*, 5–16.
(25) Massart, D. L.; Vandeginste, B. G. M.; Buydens, L. M. C.; DeJong, S.; Lewi, P. J.; Smeyers-Verbeke, J. In *Handbook of Chemometrics and Qualimetrics: Part A: Data Handling in Science and Technology*; Elsevier: Amsterdam, The Netherlands, 1997; Vol. 20A, pp 436.
(26) Christy, A. A.; Kasemsumran, S.; Du, Y.; Ozaki, Y. The detection and quantification of adulteration in olive oil by near-infrared spectroscopy and chemometrics. *Anal. Sci.* **2004**, *20*, 935–940.
(27) Cozzolino, D.; Murray, I.; Chree, A.; Scaife, J. R. Multivariate determination of free fatty acids and moisture in fish oils by partial least-squares regression and near-infrared spectroscopy. *LWT* **2005**, *38*, 821–828.
(28) Hourant, P.; Baeten, V.; Morales, M. T.; Meurens, M.; Aparicio, R. Oil and fat classification by selected bands of near-infrared spectroscopy. *Appl. Spectrosc.* **2000**, *54*, 1168–1174.

Received for review September 10, 2008. Revised manuscript received October 1, 2008. Accepted October 13, 2008.

JF802792D