Measurement of Adulteration of Olive Oils by Near-Infrared Spectroscopy

I.J. Wesley[a,*], R.J. Barnes[b] and A.E.J. McGill[a]

[a]Seale-Hayne, Department of Agriculture and Food Studies, University of Plymouth, Newton Abbot, Devon, United Kingdom TQ12 6NQ and [b]Perstorp/NIRSystems, Highfield House, Foundation Park, Roxborough Way, Maidenhead, Berkshire, United Kingdom SL6 3UD

ABSTRACT: Authentication of olive oils is of great importance, not only because they command a high price but also because of the health implications of adulteration with seed oils. A method for predicting the level of adulteration in a set of virgin and extra-virgin olive oils adulterated with corn oil, sunflower oil, and raw olive residue oil by near-infrared spectroscopy is presented. The best result was a correct prediction for 98% of the samples. Principal component analysis was used to predict the type of adulterant. The best result was a 75% prediction rate. From these results, it is concluded that it is possible to design a quality control system, which uses near-infrared technology to measure the level of adulteration. In the case where the only test is whether the sample is adulterated or not, a simple calibration for adulteration can be used. The results suggest that principal component analysis may offer a means of identifying the adulterant, although more work is required to give an acceptable level of accuracy.
JAOCS 72, 289–292 (1995).

KEY WORDS: Near-infrared spectroscopy, olive oil, principal component analysis.

The dramatic increase in demand for olive oil over the past few years can be attributed not only to flavor but also to reports of potential health benefits. Hence, olive oil commands a high price on the market, and authentication of purity is therefore desirable. Olive oils are classified according to purity, which varies from extra-virgin through semifine to lampante (which is unfit for human consumption without further refining), and method of production by European Union (EEC) Commission Regulations. Virgin olive oil is defined in these regulations as follows: "Virgin olive oil means oils derived solely from olives using mechanical or other physical means under conditions, and particularly thermal conditions, that do not lead to deterioration of the oil, and which have undergone no treatment other than washing, decantation, centrifugation or filtration, but excluding oils extracted from olives using solvents" (1).

The adulteration of olive oils with cheaper low-grade oils is potentially a great commercial problem in countries that manufacture seed oils and import olive oils. The most common adulterants found in virgin olive oil include refined olive oil, raw olive residue oil, synthetic olive oil–glycerol products, and seed oils (such as rapeseed oil). The adulteration of extra-virgin olive oil with virgin olive oil is not such a serious problem. The health implications of adulteration with seed oil have been made apparent by cases such as "Spanish Toxic Syndrome," an outbreak of food poisoning in Spain in 1981 in which hundreds of deaths and thousands of cases of serious illnesses were attributed to consumption of low-quality olive oils which had been adulterated with rapeseed oil (2).

Various chemical techniques exist that can be used to detect adulteration of olive oil by low-grade olive oils and seed oils. These include determination of iodine value, saponification value, density, viscosity, ultraviolet absorbance, fluorescence, refractive index, and colorimetric reactions. These have been extensively reviewed elsewhere (3) and form the backbone of EEC legislation on the characteristics of olive oils (1). It is, however, relatively easy to adulterate pure olive oil to quite high levels, such that the physical and chemical properties still fall within the accepted limits for the pure sample. Other techniques that have come to prominence in the last few years include modified high-performance liquid chromatography and gas chromatography in which the separated components are identified quantitatively by traditional methods. Although this method of analysis offers the possibility of high-resolution separation, it is time-consuming and requires a high degree of technical knowledge when interpreting the data. Hence, it is expensive in the commercial environment. The ^{13}C nuclear magnetic resonance, mass spectrometry, and infrared and Raman spectroscopy have also all been applied to the problem with varying degrees of success (4).

Near-infrared (NIR) spectroscopy has many applications in the food industry and is of particular use in the quality control of raw materials and final products. There are several references in the literature to studies of fats and oils, particularly in the areas of fatty acid composition (5) and level of unsaturation (6,7), and there are also reports of detection of foreign fat adulteration of milk fat (8). To date, there appears to be little data on the specific problem of the adulteration of oils.

*To whom correspondence should be addressed.

EXPERIMENTAL PROCEDURES

The following samples were obtained: extra-virgin olive oil (four samples); virgin olive oil (three samples); refined olive oil (three samples); raw olive residue oil; corn oil; and sunflower oil. The extra-virgin and virgin olive oils came from different geographical regions, and the refined olive oils came from different blends. In addition, commercially available extra-virgin and virgin olive oils were obtained from a local retailer. Samples of various mixtures of virgin olive oil or extra-virgin olive oil adulterated with either corn oil, sunflower oil, or refined olive residue oil were prepared for analysis. In total, 319 samples were prepared for calibration purposes that covered the range 0–30% w/w adulteration in 5% steps. Samples were scanned in an NIRSystems model 6500 scanning NIR spectrometer (Perstorp Analytical, Maidenhead, Devon, United Kingdom). This instrument covers the NIR spectral region from 400–2500 nm and was configured for direct transmission measurements with a standard 1-mm quartz cuvette. Spectra were recorded as log 1/Transmittance at 2 nm intervals from 800–2500 nm. The scan speed was 1.8 scans/s, and 4-point Fourier smoothing was applied. Statistical analysis was carried out with standard ISI Systems (Perstorp Analytical) software. The NIR spectrum of a typical extra-virgin olive oil is shown in Figure 1.

Calibration and statistical analysis. Before calibration, outliers were removed, and then the data set was split in two to provide an initial calibration set of 250 spectra and a validation set of 60 spectra. A variety of different mathematical treatments were tried, and it was found that the best results were obtained with a modified partial least squares regression with a first-derivative math treatment applied over 12 data points and an 8-point smoothing function over the range 1108–2200 nm. Scatter correction was not used, and outliers were downweighted, not eliminated. The final equation calculated from the calibration set had 14 factors and was checked by predicting the validation set and comparing the results with the known laboratory values. Figure 2 shows the graph obtained when the laboratory value is plotted against the predicted value for the validation set. The line y = x is marked, and the error bars on the data points are within the limits of the data point markers. On the basis of these results, a final calibration was prepared by the entire data set by the same mathematical treatment. A wild validation set of twenty samples was prepared, each sample being analyzed twice to give a total of 40 spectra. This validation set contained some samples of olive oil that were not represented in the calibration data and were expected to be outliers. When these data were predicted with the new equation, the four outlier spectra were easily identified. The data are plotted in Figure 3, together with the line y = x.

The cross-validation statistics for the two validation sets are shown in Table 1. For the final validation set, the four samples (8 spectra) known to be outliers have been removed from these statistics. The standard error of calibration (SEC) is the standard deviation for the residuals due to differences between actual and the NIR-predicted values for samples within the calibration set. The standard error of prediction (SEP) is the standard deviation for the residuals due to differences between actual and the NIR-predicted values for samples outside of the calibration set by using a specific calibration equation. The standard error of cross validation (SECV)

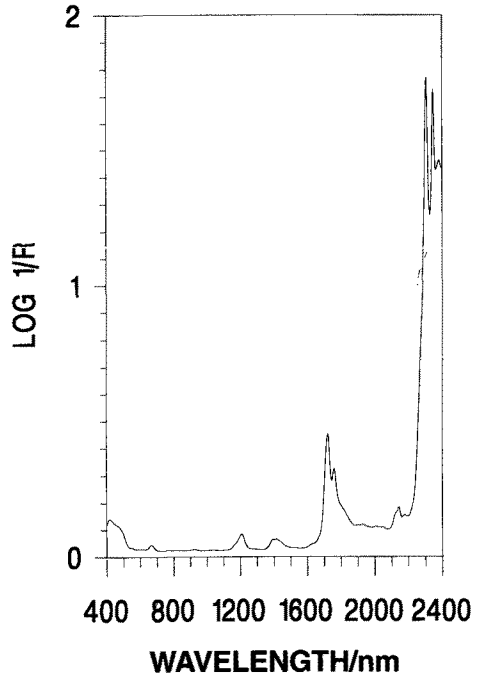

FIG. 1. The near-infrared spectrum of an extra-virgin olive oil.

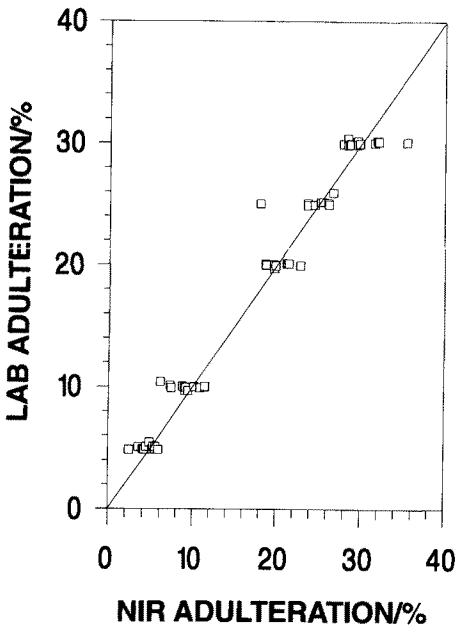

FIG. 2. Plot of actual (lab) adulteration against predicted near-infrared (NIR) adulteration for the initial validation set predicted with the initial calibration equation. The line y = x is also plotted.

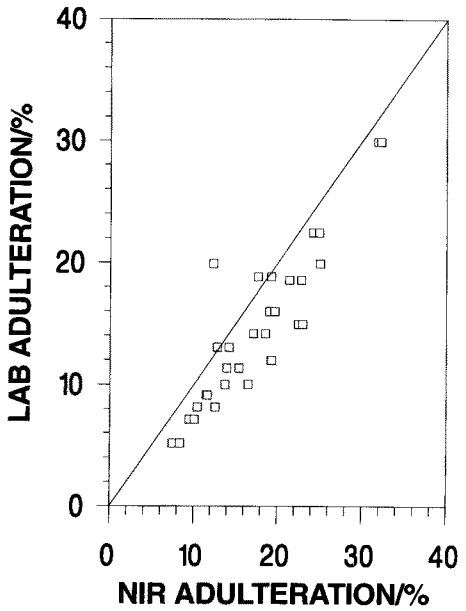

FIG. 3. Plot of actual (lab) adulteration against predicted near-infrared (NIR) adulteration for the wild validation set predicted with the final calibration equation. The line y = x is also plotted.

is calculated during the calibration procedure and is used for determining the "best" number of independent variables to use in building an equation. The cross validation method is a repetitive algorithm, which selects samples from the calibration set population to develop the calibration equation and then uses the remaining samples as a prediction set. The SECV is an estimate of the SEP and is always higher than the SEC. The ISI software allows the user to set the number of cross validations to perform and selects the equation with the lowest SECV as the best calibration. A measure of the amount of variation in the data, which is modelled by the calibration equation as a total fraction of 1.0 is r^2, e.g., if $r^2 = 0.50$, then 50% of the variation in the differences between the actual values for the data points and the predicted values for the points are explained by the calibration equation.

The results of the calibration were sufficiently encouraging that it was decided to try to predict the type of adulterant as well as the level of adulteration. The initial calibration set of 250 spectra was split into four subsets—pure, adulterated with corn oil, adulterated with sunflower oil, and adulterated with raw olive residue oil. The pure subset was the smallest, containing only 15 spectra, so each of the other subsets was reduced in size to 15 spectra by using a "Select" subroutine, which ensures that the selected samples accurately reflect the whole population. The principal components were derived for each subset and used with the original calibration equation to predict the two validation sets. It was thought that the small size of the pure subset might adversely affect the performance of the prediction, so the prediction was repeated with only the adulterated subsets. From the external validation set, the level of adulteration was correctly identified for 93% of the spectra when all four principal component files were used. This improved to 98% of the spectra when the pure spectra were removed. The adulterant was successfully identified for 58% of the spectra from four principal component files, rising to 63% of the spectra when the pure spectra were ignored. Similarly, for the wild validation set, the level of adulteration was identified for 45% of the spectra when all four principal component files were used, improving to 55% of the spectra when the pure samples were ignored. The identification of the adulterant was correct for 75% of the samples, and it was not improved by removing pure samples.

TABLE 1
Cross Validation Statistics for Calibrations

	Initial calibration	Final calibration
Standard error of calibration (%)	1.06	1.23
Standard error of cross validation (%)	1.31	1.45
Number of samples	57	32
Sample range (%)	4.86–30.39	5.17–29.95
Mean of lab values (%)	17.52	14.42
Standard error of prediction (%)	1.78	4.15
r^2	0.97	0.80
Standard deviation (%)	9.85	6.32
Slope of line of best fit	0.95	0.89

RESULTS AND DISCUSSION

Two recent studies (9,10) have shown that principal components analysis can be used to identify unknown vegetable oils. From the results shown in Figure 2, it is quite clear that it is possible to produce a calibration for adulteration from a variety of base olive oils and a variety of adulterants. The results of the wild validation set (Fig. 3) shows a similar relationship, although there is a bias. This is not unexpected as it is well known that the best calibrations are achieved when the calibration data set is a fair representation of the data to be analyzed. For the first calibration, this is clearly the case because both the calibration and validation sets are a subset of the entire data set. However, in the second case, many of the levels of adulteration used were not represented in the calibration set, i.e., the validation data set is truly wild. It is, however, relatively easy to adjust the bias to bring the predicted values closer to the real values. The results from the principal component analysis indicate that the adulterant can be predicted at least three times out of five. Although this may not seem particularly high, this strike rate can be significantly improved by combining the sunflower and corn oil adulterants and predicting seed oil against raw olive residue oil. In this case, the success rate is approximately 75%. It is probable that the prediction statistics can be improved by increasing the size of the calibration set, notably filling the gaps in the data set. It should also be remembered that the variation in the spectra caused by adulteration could equally be caused by natural variation in the olive oil itself. Consequently, many more pure samples should be recorded to account for this variation. In addition, the spectrum of each sample should be

recorded a number of times to account for variation in the NIR spectra.

In many applications, it is only necessary to identify whether a sample is adulterated, and the type of adulterant may be of secondary importance. In addition, the sample set used contained a wide range of olive oils with different characteristics. In an industrial application, this may be much more restricted, which makes calibration easier. In that case, the number of samples and spectra could be safely reduced, and a simple calibration for adulteration could be used.

There is obviously some interest in whether any of the factors used in the equations can be related to specific adulterant type. It is sometimes possible to identify such features by plotting the equation loadings. In this case, it was not possible to assign specific features to the different adulterants, although interesting features were noted in the C-H overtone regions. Work is in progress to ascertain whether these features can be used as a means of identifying oil type, and hence improve the ability of this method to predict adulterant type.

ACKNOWLEDGMENTS

This work was supported by the University of Plymouth. M. Manley provided technical assistance.

REFERENCES

1. Commission Regulation (EEC) No 2568/91, OJ No L 248, 5.9.1991, 1–83.
2. World Health Organization, *Toxic Oil Syndrome. Current Knowledge and Future Perspectives,* World Health Organization, Geneva, Switzerland 1992.
3. Gracian Tous, in *Analysis and Characterization of Oils, Fats and Fat Products, Vol. 2,* edited by H.A. Boekenoogen, Interscience, 1968, pp. 315–606.
4. Li-Chan, E., *Trends Food Sci. Tech. 5*:3 (1994).
5. Sato, T., S. Kawano and M. Iwamoto, *J. Am. Oil Chem. Soc. 68*:827 (1991).
6. Roy, R.B., and L. McDermott, *Ibid. 65*:501 (1988).
7. Ingner, N.M., and D.S. Kalka, *Ibid. 61*:652 (1984).
8. Sato, T., S. Kawano and M. Iwamoto, *J. Dairy Sci. 73*:3408 (1990).
9. Sato, T., *J. Am. Oil Chem. Soc. 71*:293 (1994).
10. Bewig, K.M., A.D. Clarke, C. Roberts and N. Unklesbay, *Ibid. 71*:195 (1994).

[Received August 9, 1994; accepted December 8, 1994]