Discrimination of Edible Oil Products and Quantitative Determination of Their Iodine Value by Fourier Transform Near-Infrared Spectroscopy

Hui Li[a], F.R. van de Voort[a,*], A.A. Ismail[a], J. Sedman[a], R. Cox[b], C. Simard[b], and H. Buijs[b]

[a]McGill IR Group, Department of Food Science and Agricultural Chemistry, Macdonald Campus , McGill University, Ste. Anne de Bellevue, Québec, Canada H9X 3V9, and [b]Bomem, Québec, Québec, Canada G2E 5S5

ABSTRACT: This work demonstrates the application of partial least squares (PLS) analysis as a discriminant as well as a quantitative tool in the analysis of edible fats and oils by Fourier transform near-infrared (FT-NIR) spectroscopy. Edible fats and oils provided by a processor were used to calibrate a FT-NIR spectrometer to discriminate between four oil formulations and to determine iodine value (IV). Samples were premelted and analyzed in glass vials maintained at 75°C to ensure that the samples remained liquid. PLS calibrations for the prediction of IV were derived for each oil type by using a subset of the samples provided as the PLS training set. For each oil formulation (type), discrimination criteria were established based on the IV range, spectral residual, and PLS factor scores output from the PLS calibration model. It was found that all four oil types could be clearly differentiated from each other, and all the validation samples, including a set of blind validation samples provided by the processor, were correctly classified. The PLS-predicted IV for the validation samples were in good agreement with the gas chromatography IV values provided by the processor. Comparable predictive accuracy was obtained from a calibration derived by combining samples of all four oil types in the training set as well as a global IV calibration supplied by the instrument manufacturer. The results of this study demonstrate that by combining the rapid and convenient analytical capabilities of FT-NIR spectroscopy with the discriminant and predictive power of PLS, one can both identify oil type as well as predict IV with a high degree of confidence. These combined capabilities provide processors with better control over their process.

Paper no. J9230 in *JAOCS 77,* 29–36 (January 2000)

KEY WORDS: Discriminant analysis, fats and oils, Fourier transform near-infrared spectroscopy, iodine value, partial least squares, quality control.

The application of near-infrared (NIR) spectroscopy in edible oil analysis has predominantly involved its use for the rapid quantitative determination of the oil content in oilseeds, with relatively little work being carried out on the analysis of oils *per se*. Most work on NIR oil analysis development has focused on classifying and/or discriminating between oil types as well as detecting adulteration, particularly of olive oil. Bewig *et al.* (1) used a filter-based NIR instrument to differentiate between four types of oils (cottonseed, canola, soybean, and peanut) using discriminant analysis based on Mahalanobis distance principles. Sato (2) used principal component analysis (PCA) to classify vegetable oils using second-derivative NIR spectra, with PCA providing the benefit of using all the spectral data collected rather than only the data at selected wavelengths. Wesley *et al.* (3,4) worked on olive oil adulteration and demonstrated that it is possible to effectively use NIR spectroscopy in conjunction with PCA to predict both the purity of olive oil and the type of adulterant as well as to quantitate the adulterant using partial least squares (PLS) regression. These studies indicate that discriminant analysis, PCA, and PLS are potentially powerful tools for qualitatively characterizing oils as well as detecting adulteration and estimating the levels of adulterants.

The McGill IR Group has focused on the development of Fourier transform infrared (FTIR) spectroscopic methods for the rapid quantitative analysis of edible oils based on measurements in the mid-IR region of the spectrum (5–9). Recently, two FT-NIR oil analysis methods have been developed, one for the determination of peroxide value (PV) and the other for the simultaneous determination of *cis* and *trans* content, iodine value (IV), and saponification number of edible oils (10,11). This work has elicited positive feedback from industry sources who cite the ease of sample handling and amenability to at-line and on-line implementation as important attributes of FT-NIR oil analysis methods. For NIR oil analysis, disposable glass vials are simply filled with neat sample and discarded after measurement. Thus, NIR analysis is more suitable for industrial applications than mid-IR analysis, which employs transmission IR cells with narrow path lengths (typically 0.025 mm) and salt windows. NIR analysis also can be performed remotely with the use of low-cost NIR-transmitting fiber optics. Conventional dispersive NIR instruments can be unreliable in terms of maintaining calibration stability, but this problem has been largely overcome with a

---

*To whom correspondence should be addressed at McGill IR Group, Dept. of Food Science and Agricultural Chemistry, Macdonald Campus, McGill University, Box 187, Ste. Anne de Bellevue, Québec, Canada H9X 3V9. E-mail: vandevoort@macdonald.mcgill.ca new generation of FT-NIR instruments that have superior wavelength reproducibility and stability, thereby minimizing calibration drift and reducing the need to recalibrate.

In recent years, several groups have used NIR spectroscopy to characterize and classify different types of fats and oils. This work has largely been based on the application of qualitative analysis techniques such as discriminant analysis based on measurements at selected wavelengths and PCA. However, quantitative analysis techniques such as PLS can also serve as a basis for classification through the setting of discrimination criteria based on the output obtained from the calibration models. There is substantial interest in the edible oils sector in instrumental methods for both the rapid identification of the type of oil and the quantitation of specific oil parameters such as IV and *trans* content as well as discriminating between oils. This work reports the protocol and results of an investigation of the capabilities of FT-NIR spectroscopy as a practical at-line process control tool for discriminating between various formulated oil products as well as determining their IV.

MATERIALS AND METHODS

Oil samples. For this work, four sets of fats and oils (A–D) were sent to the McGill IR Group by an oil processor. These samples had all been preanalyzed for IV by gas chromatography (GC), and the GC IV data were provided with the samples. The IV of the samples ranged from 133.3–134.8 for oil A, 91.3–96.3 for oil B, 117.1–118.8 for oil C, and 113.7–117.0 for oil D. In addition, 35 unknowns also were provided as a blind validation set to be used in evaluating the accuracy of product classification and IV determination by FT-NIR spectroscopy. The FT-NIR prediction results obtained for these samples were subsequently sent to the processor, who then made the classification and GC IV data for these samples available for statistical analysis.

Instrumentation and sample handling. The instrument used in this study was a Bomem FT-NIR analyzer (Bomem Inc., Québec, Canada) equipped with a deuterated triglycine sulfate (DTGS) detector capable of scanning the spectral range of 12,000–2,000 cm^{-1}. The spectrometer was controlled by an IBM-compatible 486 DX 66 MHz PC running under Windows-based Bomem-Grams/386 software (Galactic Industries Co., Salem, NH) and AIRS, a specific quality assurance program produced by DHC Analysis (Cleveland, OH). The sample handling accessory used in this study was a temperature-controllable multivial-holding block capable of accepting 8-mm (o.d.) transparent glass vials (Kimble Glass Inc., Vineland, NJ) with a volume of ~1 mL. Figure 1 illustrates the sample handling accessory installed in the spectrometer. The temperature of the sample handling accessory was held at 75 ± 0.2°C. For sample analysis, vials were filled with 0.5–0.7 mL of oil or premelted fat and scanned over the range of 12,000–4,500 cm^{-1}.

All sample and background spectra were recorded by coadding 128 scans at a resolution of 16 cm^{-1}. Air background

FIG. 1. The Bomem (Québec, Canada) MB series Fourier transform near-infrared (FT-NIR) spectrometer with temperature-controlled multivial-holding block.

spectra were collected every 30 min with the vial holder in the IR beam, and sample spectra were ratioed against the most recently collected air background. The ratioed spectra were subsequently normalized to account for inherent variations in the vial pathlength by using a normalization routine in the AIRS software.

Standards, software, and calibration. From each oil product category provided, 20 samples were used as calibration standards and 5 samples were employed as validation samples. The calibration standards were scanned twice and the glass vial containing the sample was rotated 90° in the vial holder between the scans. The validation standards were run in duplicate. The spectra of the calibration standards together with the GC IV data obtained from the processor were combined in the PLSPlus chemometrics program (Galactic Industries Co.) to develop PLS-NIR IV calibrations. The predicted residual error sum of squares (PRESS) test and the root mean square deviation (RMSD) associated with the cross validation of the calibrations tested were used to select optimal calibrations. The performance of the calibrations was evaluated by linear regression of the cross-validation predictions and the predictions for the validation samples against the GC data provided by the oil processor, with uncertainty and reproducibility assessed using mean differences (MD) and standard deviations of the differences (SDD) according to the method of Youden and Steiner (12). In addition to the PLS calibrations developed for IV using the samples provided, a global IV calibration provided by Bomem was also assessed. This calibration was developed using a broad selection of edible oils and fats that were collected from over 15 facilities worldwide (13). These oils and fats varied in type and degree of hydrogenation. Oil classification was carried out using AIRS, which provides a means to discriminate between sample types or categories based on (i) the PLS-predicted value of a parameter being measured; (ii) the spectral residual, representing the spectral data not fit by the PLS calibration model; or (iii) the PLS factor scores associated with the spectrum, or any combination of the three, as explained below.

Concepts and principles of the FT-NIR PLS method. PLS is a powerful multivariate analysis technique that has largely been pioneered for NIR applications and has played a major role in the recent resurgence of quantitative mid-IR spectroscopy. PLS develops a calibration model by compressing the spectral data for a set of calibration standards into a series of orthonormal basis vectors, known as loading spectra or factors. The basis vectors selected to model the spectra of the calibration standards emphasize spectral variations due to differences in concentration and do not model variations due to random noise, hence the name PLS. A PLS calibration can, in principle, be based on the whole spectrum, although in practice, the analysis tends to be restricted to regions of the spectrum that exhibit the strongest variations with changes in the concentrations of the components of interest. PLS decomposes the spectrum of each calibration standard into a weighted sum of the loading spectra, and the weights given to each loading spectrum, known as "scores," are regressed against the concentration data for the standards. When the spectrum of an unknown is analyzed, PLS reconstructs the spectrum from the loading spectra. The amounts of each loading spectrum employed in reconstructing the spectrum, i.e., the "scores," are then used to predict the concentration of the unknown. PLS also generates a spectral residual, which corresponds to the difference between the actual and the reconstructed spectrum.

The scores and spectral residual provide useful information for the detection of outliers, i.e., samples whose spectra differ significantly from those of the training set. Outlier detection is valuable in relation to assessing quantitative accuracy in the prediction of unknowns since, if an unknown is an outlier, the PLS-predicted value for the parameter of interest cannot be considered reliable. It also provides a means of classifying an unknown as being part of the population modeled by the training set, likely being part of this population or not being part of this population on the basis of selected criteria. These criteria may be based on the value of the parameter being predicted, the magnitude of the spectral residual, and the factor scores determined for the unknown. The underlying basis for each of these criteria is as follows: (i) Value within range. If an unknown is part of the population represented by the training set, then the value of the predicted parameter should be within the range of values spanned by the training set. (ii) Residual. When an unknown is predicted, a residual spectrum is computed by subtracting the actual spectrum of the sample from the synthetic spectrum generated by PLS. In order to obtain a numerical value for the magnitude of the residual, the absorbance values at each data point of the residual spectrum are squared and then summed, and then the square root of the sum is taken. If an unknown is part of the population represented by the training set, then the value of the spectral residual should not exceed the average residual encountered in the training set by a large factor. The lower limit for the spectral residual is necessarily set to zero because a zero residual would represent a perfect fit of the spectral data for the unknown. (iii) Factor scores. When a sample is analyzed using PLS, the output includes the contribution (factor score) of each loading spectrum or factor to the synthetic spectrum generated by PLS. When the factor scores for an unknown are within the range of the scores obtained for the training set, it is a very strong indication that the sample truly belongs to the population modeled by the training set. Thus, the criterion based on factor scores is probably the most important classification tool of the three.

The concepts, up to this point, consider only a single population, but can be expanded to consider several sample populations simultaneously. In our work, four oil types (populations) were considered with the objective being to classify unknowns as belonging to one of these types as well as to determine IV. In principle, provided the oil types are sufficiently distinct from each other from a spectral standpoint, this can be achieved by developing a PLS calibration for the prediction of IV for each oil type and then classifying the unknowns on the basis of the output from PLS. This type of classification can be implemented through the AIRS software package, which allows one to develop a method to carry out the following sequence of events: (i) When a sample is analyzed by AIRS, the IV will be predicted from each of the PLS calibration models developed for the different oil types and the discrimination criteria selected for each model (based on range, residual and factor scores) will be applied. (ii) Provided that the sample belongs to one of the oil types and that there are sufficient differences between the oil types, the sample analyzed will pass the discrimination criteria for only one of the models and will accordingly be classified as belonging to the corresponding oil type. (iii) The IV prediction from the selected calibration model will then be reported.

The creation of a method of this type by the AIRS software extends PLS beyond its basic predictive capabilities, providing a basis for classification of a sample within a defined class or among defined classes of oil types. This ability can be particularly useful in quality control situations where one may want not only to predict the value of a parameter but also to confirm to which oil type the sample belongs.

RESULTS AND DISCUSSION

The concepts and principles discussed above have been applied in the development of a FT-NIR method for discrimination among four defined oil types provided by a processor. PLS calibrations for the prediction of IV from the FT-NIR spectra in the region between 9100 and 7560 cm^{-1} were first individually developed for each oil type using 20 oil samples of that type as the training set. This spectral region was chosen as it contains on-scale overtone bands containing information related to the degree of unsaturation. Table 1 presents the mean, SD, and range of the GC IV data of the standards used for the development of the PLS calibrations, as well as the cross-validation statistics obtained from those calibrations in terms of mean difference (MD_a) and standard deviation of the differences (SDD_a). The MD_a and SDD_a reflect, respectively, the bias and random error of the PLS predictions. The cross-validation statistics indicate that the IV theoretically can be predicted within a standard deviation of 0.20 IV units or less for oils A, C, and D, whereas for oil B, the SDD_a is ~0.50.

Figure 2 presents a composite plot of the individual cross-validation predictions for each of the oil types, showing each oil grouping in a tight cluster except for oil B, which appears to be composed of two subgroups. This somewhat broader spread for oil B may explain the substantially higher SDD_a obtained for this oil. In relation to the feasibility of discriminating among the four oil types on the basis of IV only, oils C and D could potentially pose problems. Although fairly similar in their IV values, careful inspection of their respective SD and ranges indicates that their IV values have no overlap. On the other hand, on the basis of the PLS cross-validation predictions, the situation is less clear-cut in terms of discriminating between oils C and D on the basis of predicted IV alone because the SDD_a values for both oils exceed the difference between their ranges. Therefore, a more elaborate discrimination strategy was employed.

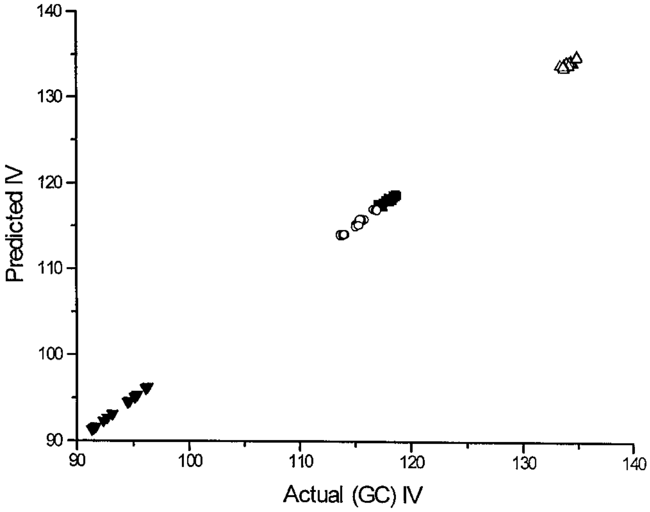

FIG. 2. Composite plot of the cross-validation results for the four calibrations, showing the predicted iodine value (IV) for the individual training sets vs. the gas chromatographic (GC) IV data. △, oil A; ▼, oil B; □, oil C; ○, oil D.

Three discrimination criteria based on predicted IV, the spectral residual, and the factor scores were used for classification of samples among the four oil types. In the case of factor scores, the score ranges for only the first two factors in the calibration model were found to be sufficient for discrimination among the four oil types. Figure 3 presents a schematic representation of the three discrimination criteria set up in the AIRS program to determine whether a validation sample belongs to oil type A. The diagram indicates how the program renders its decisions as to whether a sample belongs, likely belongs, or does not belong to the oil type A population using the three criteria. As shown, each criterion is defined by two sets of limits. The inner region (I) represents the region where

TABLE 1
GC IV Data and PLS Cross-Validation Statistics for Oil Types A–D[a]

Oil	Mean GC IV	SD GC IV	Range GC IV	PLS MD_a	PLS SDD_a	R^2	n
A	134.01	0.39	133.3–134.8	0.0005	0.2164	0.999	20
B	93.90	1.79	91.3–96.2	0.0001	0.4827	0.999	20
C	117.85	0.51	117.1–118.8	0.0020	0.1940	0.999	20
D	115.36	1.11	113.7–117.0	–0.0047	0.1468	0.999	20

[a] GC, gas chromatographic; IV, iodine value; PLS, partial least squares; SD, standard deviation; SDD_a, standard deviation of differences; MD_a, mean difference.

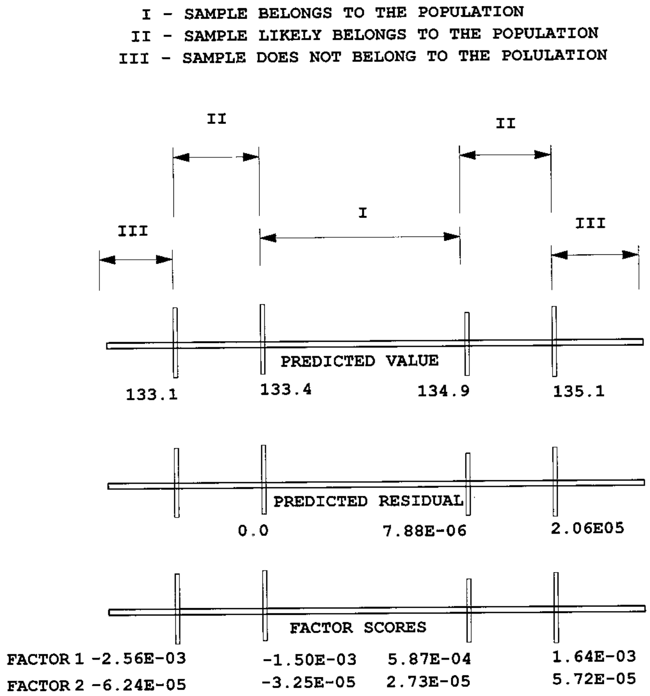

FIG. 3. Schematic representation of the partial least squares (PLS)-based discrimination criteria (predicted IV, spectral residual, and factor scores) used to accept (region I), accept with qualification (regions II) and reject (regions III) the oil samples as belonging to oil type A. See Figure 2 for abbreviation.

a sample is considered to be part of the population, while samples falling in regions II are likely part of the population, and samples falling into regions III are definitely not part of the population. As noted earlier, region I is rigorously defined on the basis of PLS results obtained for the training set, and the limits of this region should in theory be the basis for the acceptance or rejection of a sample to or from a population. The calculated rate of incorrect classification (i.e., a false positive) based on these rigorously defined discrimination criteria is frequently less than 1 part in 1 million. It is much more likely, however, that a sample belonging to a certain population will not be classified as such (i.e., a false negative), owing to minor deviations in the spectra, and thus a degree of tolerance is built into the discrimination criteria based on the variability of the spectra in the training set. Hence, the additional set of limits (regions II) is included for each discrimination criterion, which can be adjusted to attain a degree of discrimination commensurate with the characteristics of the samples one is working with. In our work, the limits defining regions II were set as follows: range, limits of region I ± root mean square error obtained from cross validation of the calibration; residual, lower limit is zero by default; upper limit is twice the upper limit of region I; and factor scores, limits of region I ± half the range spanned by region I.

Table 2 presents the classifications obtained for 20 validation samples, five from each oil type, using the three discrimination criteria based on these limits as well as each discrimination criterion individually. For the majority of the samples, the use of the inner region (I) yields the correct classification with each of the individual discrimination criteria. The results for oil A illustrate the need to build some tolerance into the discrimination criteria. Because the value of the spectral

TABLE 2
Classification of 20 Validation Samples Among Oil Types A–D Using PLS-Based Discrimination Criteria

Sample code	Actual (GC) IV	Range, residual, and factor scores[a,b]	Range only[a]	Residual only[a]	Factor scores only[a,b]
A1073	134.2	A*	A	A*, B*	A
A1078	134.5	A	A	A, B*	A
A1083	133.9	A*	A	A*	A
A1088	134.2	A	A	A, B*	A
A1093	133.9	A*	A	A*	A
B1566	95.2	B	B	B	B
B1571	94.3	B	B	B	B
B1576	94.2	B	B	B	B
B1581	94.6	B	B	B	B
B1586	94.8	B*	B	B	B*
C2216	118.7	C*	C	C, D*, B*	C*
C2221	118.0	C	C	C, D*, B*	C
C2226	118.5	C*	C*	C, D*, B*	C
C2231	118.2	C	C	C*, D*, B*	C
C2236	118.1	C	C	C, D*, B*	C
D2556	115.3	D	D	D, C*, B*	D, C*
D2561	115.2	D	D	D, C*, B*	D
D2566	115.4	D	D	D, C*, B*	D
D2571	115.1	D	D	D, C*, B*	D
D2576	115.1	D	D	D, C*, B*	D

[a] Classification of samples without an asterisk is based on region I; those with an asterisk are based on regions II.

[b] Factor scores were those for the first two factors in the calibration model. See Table 1 for abbreviations.

residual from the PLS calibration for oil A is very small, three of the five validation samples belonging to this oil type would be rejected on the basis of the spectral residual discrimination criterion with the use of region I only. However, when the spectral residual discrimination criterion is extended to include regions II, the sample identification is inconclusive in several cases, indicating the need to base the classification on more than a single discrimination criterion. The results from the classification when the three discrimination criteria were combined in the AIRS program were 100% correct; the samples marked with an asterisk in Table 2 were flagged by the AIRS program as possible outliers without any indication of which criteria were marginal.

For circumstances in which classification is not of interest *per se*, we investigated the predictive performance of a calibration developed for the four oil types combined, rather than calibrating for each oil type individually. In addition, we evaluated the relative predictive accuracy of a "global calibration" developed and supplied by Bomem (13). Table 3 presents the MD and SDD for uncertainty and reproducibility for the pooled data for the validation samples. As can be seen, the individual and combined calibrations perform comparably, while the global calibration has a small bias (<1.0 IV unit) and a somewhat larger error associated with it. The global calibration, which was developed using an extensive selection of oil types (13), actually performs very well in tracking IV, but requires a bias adjustment in order to match the agreement obtained with the calibrations developed for the specific oil types. As seen from the data in Table 3, it does not matter whether one calibrates on oil types individually or together, except that the capability for classification is clearly lost in the latter case. However, the combined calibration has the advantage that it can be applied over a wider IV range with the possibility of interpolation for blends of the different types of oils, which have IV values outside the ranges of the individual oil types.

Although additional information was not provided by the processor about the samples supplied for evaluation in this study, one can surmise, based on the IV data provided, that oils C and D are two different blends of oils A and B. If this were the case, quality control would be interested in knowing which oil or blend is being processed and verifying that the IV is within specifications. By using the individual calibrations, one is capable of both classifying a sample as one of the four oil types and obtaining an accurate value for its IV. If a formulation error has been made (e.g., blending oils A and B in the wrong ratio), the sample would be rejected as an outlier by all four individual calibrations, and no IV prediction would be obtained. In this situation, having a combined calibration would be useful. On the basis of the discrimination criteria for the combined calibration, it would be possible to ascertain whether the sample is in fact some blend of oils A and B and, if it is, to obtain a prediction for its IV. Taking this a step further, if the sample were completely unrelated to oils A and B and their blends, the combined calibration would reject the sample and, by resorting to the global calibration, the IV value could be predicted, possibly providing a clue as to what type of oil the sample may be. The scenario described above illustrates how the AIRS program would allow one to analyze and troubleshoot a sample by FT-NIR spectroscopy and make practical use of a PLS-based classification method in a quality control situation.

TABLE 3
MD_a and SDD_a of the IV Predictions for 20 Validation Samples from the Individual, Combined, and Global Calibrations Relative to the GC IV Reference Values

Statistic[a]	Individual calibrations	Combined calibration	Global calibration
MD_a	0.15	0.02	0.89
SDD_a	0.32	0.33	0.49
MD_r	0.01	0.03	–0.08
SDD_r	0.15	0.15	0.34

[a] *a* = uncertainty, *r* = reproducibility. See Table 1 for other abbreviations.

Analysis of unknowns. Although the results presented in Table 3 were derived using proper validation techniques and samples, an additional blind validation study was conducted to provide an added degree of confidence in the method. Table 4 presents the results obtained for classification of the blind unknowns as well as their IV predictions from the individual, combined, and global calibrations. This table also includes the information provided to us after completion of the validation study by the processor supplying these samples. Twenty of the 35 samples were identified by the processor as belonging to oil types A–D, and all but one of these had been correctly classified by the FT-NIR method. Approximately half of these correct classifications were based on the less rigorously defined discrimination criteria (regions II), again illustrating the need to build some tolerance into the discrimination criteria. It is particularly noteworthy that despite the similarity of oil types C and D in terms of their IV, the FT-NIR method was successful in distinguishing between these two oil types.

Five unknowns were incorrectly blended samples of oil type B that had been deliberately included in the validation set to test the capability of the FT-NIR method to detect formulation errors. All five of these samples were rejected by the calibration model for oil type B and only passed the discrimination criteria for the combined calibration model. The IV predictions obtained for these samples from the combined calibration model indicated that the IV of these samples were about 5 IV units above the IV range for oil type B. This example provides an excellent illustration of the point made previously regarding the utility of the combined calibration model in cases of formulation errors.

The 10 remaining unknowns were identified by the processor as being of types E and F, and thus did not belong to any of the oil types on which the FT-NIR method was calibrated. The five samples of type F were all correctly rejected by each of the individual calibration models as well as the combined calibration model. However, the FT-NIR method classified four samples of type E as oil type C and the fifth one as oil type D. In addition, the FT-NIR-predicted IV for the samples

TABLE 4
PLS-based Classification and IV Predictions for a Blind Sample Set

Code	PLS classification[a]	PLS IV individual	PLS IV combination	PLS IV global	Actual classification	Actual IV
U2696b	B	95.0	94.6	94.9	B	95.4
U2701b	B[f]	95.3	95.0	95.3	B	94.8
U2706b	B	94.9	94.7	95.1	B	94.3
U2711b	B	95.2	94.7	95.3	B	94.5
U2716b	A[r]	134.4	134.2	135.0	A	134.1
U2721b	B[r]	95.7	94.62	95.5	B	94.37
U2726b	A	134.3	134.2	135.1	A	133.9
U2731b	A[r]	134.5	134.3	134.9	A	133.7
U2736b	A[r]	134.4	134.2	135.4	A	133.8
U2741b	A	134.0	134.0	134.8	A	134.0
U2746b	—[b]	—	101.4	102.5	B[c]	102.4
U2751b	—	—	101.2	102.2	B[c]	102.2
U2756b	—	—	101.3	102.3	B[c]	102.5
U2761b	—	—	101.5	102.2	B[c]	102.4
U2766b	—	—	101.5	102.4	B[c]	102.4
U2771b	C[Rr]	117.2	117.2	116.9	C	117.4
U2776b	C[r]	117.3	117.2	117.3	C	117.5
U2781b	—	—	116.5	117.5	C	117.3
U2786b	C	117.3	117.2	117.4	C	117.2
U2791b	D	115.6	115.4	117.4	D	115.7
U2796b	C[r]	117.3	117.3	117.4	C	117.7
U2801b	D[rf]	115.6	115.4	116.8	D	113.9
U2806b	D	115.7	115.6	116.7	D	115.6
U2811b	D[rf]	115.6	115.2	117.4	D	115.5
U2816b	D	115.6	115.6	116.9	D	115.6
U2821b	C	117.4	117.1	117.4	E	122.0
U2826b	D[rf]	116.9	116.8	116.8	E	122.0
U2831b	C[Rr]	117.2	117.0	116.8	E	122.0
U2836b	C[Rr]	117.2	116.9	116.9	E	122.7
U2841b	C[r]	117.4	117.0	117.1	E	122.7
U2846b	—	—	—	64.7	F	70.6
U2851b	—	—	—	65.0	F	67.6
U2856b	—	—	—	64.6	F	67.8
U2861b	—	—	—	64.6	F	67.9
U2871b	—	—	—	64.7	F	67.7

[a] Classification based on region I, except for those marked with a superscript, indicating that the discrimination criteria defined in terms of range (R), spectral residual (r), or factor scores (f) were only met when classification was based on regions II. See Table 1 for other abbreviations.

[b] Sample rejected by the calibration model as not belonging to the population represented by the training set.

[c] Sample reported by the processor to have been incorrectly blended.

of oil type E are 5 IV units lower than the GC-determined IV reported by the processor, in contrast to the generally good agreement overall between the FT-NIR and GC values shown in Table 4 and examined in more detail below. To investigate the reasons for this discrepancy, we determined the IV of these samples, as well as those of nine samples of oil types A–D, using a Fourier transform mid-IR (FTIR) method previously developed in our laboratory (9). We also analyzed two samples of oil type E and one sample of oil type C by the AOCS iodometric method (14). The results of these various analyses corroborated the FT-NIR predictions, indicating that the GC IV for the samples identified as belonging to oil type E are incorrect. Accordingly, it may be concluded that the samples provided to us may not have been the same as the ones analyzed by the processor, accounting for the seemingly incorrect classification of these five samples by the FT-NIR method.

For the 20 samples belonging to oil types A–D, the predicted IV obtained from the individual calibration for the assigned oil type, the combined calibration, and the global calibration were compared to the IV obtained by GC analysis. The results showed that the individual, combined, and global calibrations yielded MD of 0.32, 0.04, and 0.83 IV units, respectively, and SDD of 0.56, 0.51, and 0.87, respectively, relative to the GC data. These results are similar to those shown in Table 3 for the initial validation set and again illustrate that slightly better matching of the GC data is achieved using the oil-specific calibrations as opposed to the global calibration.

In our previous work (11), we demonstrated the utility of PLS as a quantitative analysis tool for the prediction of IV as well as *cis* and *trans* content of fats and oils by FT-NIR spectroscopy. In the present work, we have demonstrated that PLS calibrations also can be employed to classify samples among defined oil types. We also have shown that when IV calibrations are specifically developed for defined oil types, the agreement between the NIR predictions for unknowns and the data obtained by the reference method is better than that obtained from a global calibration based on a broad range of calibration standards. The global calibration was able to predict all the samples provided for this study without flagging them as outliers. For the analytical performance of the global calibration to be comparable to that of the specific calibrations, a validation process is required to determine the bias and correct it. This study has demonstrated that FT-NIR oil analysis, when fully exploiting the combined predictive and discriminant capabilities of PLS, is a powerful and practical analytical quality control tool.

ACKNOWLEDGMENTS

The authors would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for financial support of this work through the Collaborative Research and Development Program (CRD).

REFERENCES

1. Bewig, K.M., A.D. Clarke, C. Roberts, and N. Unklesbay, Discriminant Analysis of Vegetable Oils by Near-Infrared Reflectance Spectroscopy, *J. Am. Oil Chem. Soc. 71*:195–200 (1994).
2. Sato, T., Application of Principal-Component Analysis on Near-Infrared Spectroscopic Data of Vegetable Oils for Their Classification, *Ibid*. *71*:293–298 (1994).
3. Wesley, I.J., F. Pacheco, and A.E.J. McGill, Identification of Adulterants in Olive Oils, *Ibid*. *73*:515–518 (1996).
4. Wesley, I.J., R.J. Barnes, and A.E.J. McGill, Measurement of Adulteration of Olive Oils by Near-Infrared Spectroscopy, *Ibid*. *72*:289–292 (1995).
5. van de Voort, F.R., J. Sedman, G. Emo, and A.A. Ismail, Rapid and Direct Iodine Value and Saponification Number Determination of Fats and Oils by Attenuated Total Reflectance/Fourier Transform Infrared Spectroscopy, *Ibid*. *69*:1118–1123 (1992).
6. Ismail A.A., F.R. van de Voort, J. Sedman, and G. Emo, Rapid Quantitative Determination of Free Fatty Acids in Fats and Oils by FTIR Spectroscopy, *Ibid*. *70*:335–341 (1993).
7. van de Voort, F.R., A.A. Ismail, J. Sedman, J. Dubois, and T. Nicodemo, The Determination of Peroxide Value by Fourier Transform Infrared Spectroscopy, *Ibid. 71*:921–926 (1994).
8. van de Voort, F.R., A.A. Ismail, J. Sedman, and G. Emo, Monitoring the Oxidation of Edible Oils by FTIR Spectroscopy, *Ibid*. *71*:243–253 (1994).
9. van de Voort, F.R., A.A. Ismail, and J. Sedman, A Rapid Determination of *Cis* and *Trans* Content of Fats and Oils by FTIR Spectroscopy, *Ibid*. *72*:873–880 (1995).
10. Dong, J., K. Ma, F.R. van de Voort, and A.A. Ismail, Stoichiometric Determination of Hydroperoxides in Oils by Fourier Transform Near-Infrared Spectroscopy, *JAOAC Int. 80*:345–353 (1997).
11. Li, H., F.R. van de Voort, J. Sedman, and A.A. Ismail, Rapid Determination of *cis* and *trans* Content, Iodine Value and Saponification Number of Edible Oils by Fourier Transform Near-Infrared Spectroscopy, *J. Am. Oil Chem. Soc. 76*:491–497 (1999).
12. Youden, W.J., and E.H. Steiner, *Statistical Manual of the AOAC*, Association of Official Analytical Chemists, Arlington, 1975, pp. 18–19.
13. A Global Iodine Value Calibration for the *Oil and Fat Analyzer*, Application Note, Bomem Inc., Quebec, Canada.
14. *Official Methods and Recommended Practices of the American Oil Chemists' Society*, 4th edn., American Oil Chemists' Society, Champaign, 1989; Method Cd 1-25.

[Received May 3, 1999; accepted September 4, 1999]