

# NIR-FT/Raman Spectroscopy for Nutritional Classification of Cereal Foods

Miryeong Sohn,<sup>1,2</sup> David S. Himmelsbach,<sup>1</sup> Sandra E. Kays,<sup>1</sup> Douglas D. Archibald,<sup>3</sup> and Franklin E. Barton, II<sup>1</sup>

## ABSTRACT

Cereal Chem. 82(6):660–665

The classification of cereals using near-infrared Fourier transform Raman (NIR-FT/Raman) spectroscopy was accomplished. Cereal-based food samples ( $n = 120$ ) were utilized in the study. Ground samples were scanned in low-iron NMR tubes with a 1064 nm (NIR) excitation laser using 500 mW of power. Raman scatter was collected using a Ge (LN<sub>2</sub>) detector over the Raman shift range of 202.45–3399.89 cm<sup>-1</sup>. Samples were classified based on their primary nutritional components (total dietary fiber [TDF], fat, protein, and sugar) using principle component analysis (PCA) to extract the main information. Samples were classified according to high and low content of each component using the spectral variables. Both soft independent modeling of class analogy (SIMCA) and partial least squares (PLS) regression based classification were investigated to determine which technique was the most appropriate. PCA

results suggested that the classification of a target component is subject to interference by other components in cereal. The Raman shifts that were most responsible for classification of each component were 1600–1630 cm<sup>-1</sup> for TDF, 1440 and 2853 cm<sup>-1</sup> for fat, 2910 and 1660 cm<sup>-1</sup> for protein, and 401 and 848 cm<sup>-1</sup> for sugar. The use of the selected spectral region (frequency region) for each component produced better results than the use of the entire region in both SIMCA and PLS-based classifications. PLS-based classification performed better than SIMCA for all four components, resulting in correct classification of samples 85–95% of the time. NIR-FT/Raman spectroscopy represents a rapid and reliable method by which to classify cereal foods based on their nutritional components.

Near-infrared (NIR) spectroscopy has gained wide acceptance in the field of food analysis. This has been mainly due to its ease of use and the ready availability of chemometric programs for the development of models. Like NIR spectroscopy, Raman yields information about molecular structure through measurement of the energy of molecular vibrational-state transitions. However, Raman spectroscopy has been, until recently, a less accessible and more instrumentally demanding technique in the field of food analysis. Yet near-infrared Fourier transform Raman (NIR-FT/Raman) spectroscopy has some unique advantages. Due to the fact the Raman scattering effect is based on polarizability of bonds and not their dipoles (Colthup et al 1990), it is relatively insensitive to the presence of water. In addition, Raman spectroscopy provides structural details on the level of the mid-infrared region but with even fewer overlapping bands. The use of NIR monochromatic light excitation reduces the interference from fluorescence, which is often experienced with its use in the analysis of pigmented materials. As a Fourier-transform method, it has the advantage of excellent frequency precision. These attributes, plus the ease of sample preparation, make it amenable for consideration as a routine method of analysis in food technology.

Determining of constituents simultaneously using spectroscopic technique could have a significant impact on the nutrient evaluation of foods as regulated by the U.S. Nutrition Labeling and Education Act. In the previous work, our team described the development of NIR-based calibration models for quantitative analysis of nutritional components such as total dietary fiber (TDF), protein, and fat in cereal food products (Kays et al 1996, 1997, 1998, 2000, 2005).

In this study, we investigated whether FT/Raman spectroscopy would lend itself to classification of cereal foods according to percent level of each component. The classification of samples used as a sample selection before quantitative prediction helps to understand how food constituents alter the spectral properties of cereal foods and should be able to produce models of greater

accuracy. Here we only demonstrate results for which we have quantitative information to ensure valid results for predicting relative levels of components. However, the capability exists to predict the level of or existence of nearly 40 different components in cereal products from a single Raman spectrum. We focused on the problem of improving the performance of linear calibration algorithms in this study.

## MATERIALS AND METHODS

### Samples

The samples ( $n = 120$ ) were utilized in two previous reports (Archibald et al 1998a,b) that focused on the prediction of total dietary fiber (TDF) of cereal foods. These samples included cereal grain products such as breakfast cereals, crackers, brans, flours, cake mixes, and muffin mixes as found in retail stores. They also included three synthetic mixture samples containing fiber additives used in cereal product formations (Archibald et al 1998a). Samples were milled to pass a 1-mm screen using a cyclone mill (Udy Corp., St. Louis, MO). Samples that were extremely high in sugar were frozen with liquid nitrogen before milling. Samples with >10% fat were ground with a coffee grinder. The sample set was divided into calibration and validation data sets by ranking the samples in increasing order of each component and assigning each second sample to the validation data set. The percent of contribution of samples for each component and statistical values of the calibration and validation data sets are summarized in Table I.

### Chemical Analysis

Fiber was determined as TDF according to AOAC official method 991.43 (AOAC 1992; Lee et al 1992) and as previously described (Kays et al 1998). Protein (N × 5.7 for samples with >50% wheat and N × 6.25 for all others) (Jones 1931) values were determined by combustion analysis according to AOAC official method 992.23 (AOAC 1995) using a nitrogen analyzer (model FP-2000, Leco Corp., St. Joseph, MI) on duplicate 0.5-g samples of ground cereal. Fat was determined by Soxhlet extraction analysis with petroleum ether as the solvent according to AOAC official method 945.16 (AOAC 1995). Sugar values were taken from the products' food labels, as separate analysis for this component was not available in our laboratory. Sugar is added during processing. For purposes of classification each sample was assigned a "dummy" classification variable; these are based on the average value of reference (≈12% for TDF, protein, sugar, and ≈5% for fat). Samples with >12% fiber, protein, sugar,

<sup>1</sup> United States Department of Agriculture, Agricultural Research Service, R. B. Russell Agricultural Research Center, Athens, GA 30605.

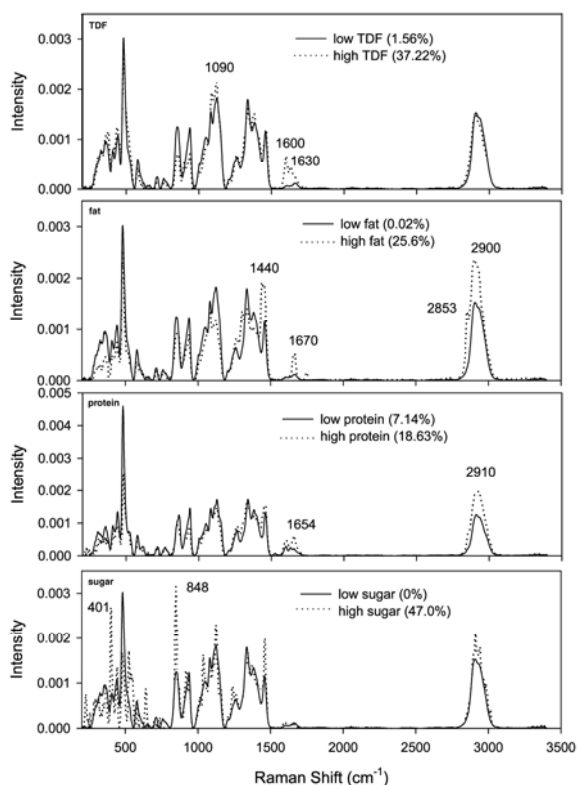
<sup>2</sup> Corresponding author. Phone: (706) 546-3374. Fax: (706) 546-3607. E-mail: msohn@qaru.ars.usda.gov

<sup>3</sup> Department of Crop and Soil Sciences, Penn State University, University Park, PA 16802.

and >5% fat were considered high level class and assigned a number 1, whereas samples with <12% fiber, protein, sugar, and <5% fat were considered low level class and assigned a number -1. (See code in Table I.)

### Spectroscopic Analysis

Raman spectroscopy was conducted on a Nicolet 950 Raman (Thermo Nicolet Instruments, Madison, WI) bench using a 1064 nm NIR laser source, a CaF<sub>2</sub> beam splitter and a liquid N<sub>2</sub> cooled Ge detector. Sample was loaded into a NMR tube (4 in. long, 5 mm diameter) and filled at least 1 in. in height (≈1 g). Individual sample tubes were pretested and selected to minimize fluorescence interference from the container. Raman spectra were collected at



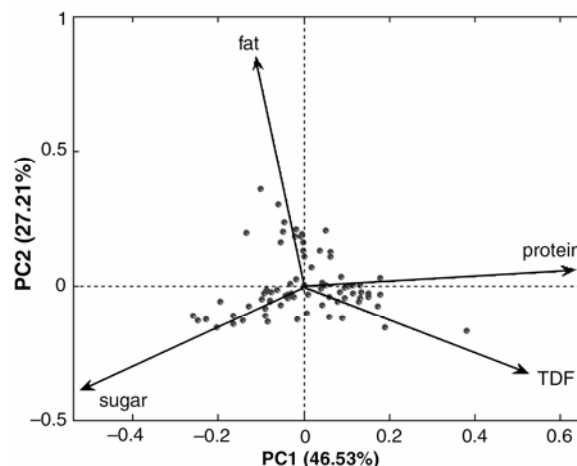
**Fig. 1.** NIR-FT/Raman spectra of cereal samples with low and high percent for TDF, fat, protein, and sugar.

4 cm<sup>-1</sup> resolution and 256 scans using Nicolet Omnic software (v. 3.1) and corrected with a white-light source reflecting from powdered KBr. Spectra were used in the Raman shift range from 202 to 3400 cm<sup>-1</sup> with a ≈2 cm<sup>-1</sup> data interval (1659 data points per spectrum). Duplicate spectra were collected from two separate tubes on each sample and averaged to produce a single spectral data file for each sample.

### Data Processing and Chemometric Analysis

The averaged Raman spectral data files were imported into GRAMS/32, v. 4.10 (Galactic Industries Corp., Salem, NH). Baseline removal was conducted by way of a locally developed Array Basic macro that automatically found the lowest points within specified spectral ranges and leveled them to zero intensity. Spectral normalization on this data set was accomplished with another locally developed Array Basic macro that automatically measured the total integrated area of the spectrum and normalized it. Chemometric analysis was performed using a Matlab software package (v. 7.01, The MathWorks, Natick, MA) with PLS\_Toolbox (Eigenvector, v. 3.5). Spectral data were all mean-centered with no additional pretreatments.

For SIMCA models, separate PCA models of high and low classes for each component were developed using the calibration set, thus producing two groups for each component, then the model developed was applied to the separate sample set to test the



**Fig. 2.** Biplot of PCA scores and loadings on y-variables of cereal samples.

**TABLE I**  
Summary of Cereal Samples Used in This Study

Components	Sample Set	Sample Group	Code <sup>c</sup>	Min (%)	Max (%)	Mean (%)	SD (%)
TDF <sup>a</sup> (n = 120)	Cal (n = 60)	High (n = 13)	1	12.29	52.14	26.03	13.05
		Low (n = 47)	-1	0.55	11.93	6.56	3.23
	Val (n = 60)	High (n = 13)	1	13.34	49.17	27.08	13.59
		Low (n = 47)	-1	0.97	11.94	6.67	3.22
Fat <sup>a</sup> (n = 90)	Cal (n = 46)	High (n = 15)	1	5.27	25.6	12.60	6.45
		Low (n = 31)	-1	0.02	4.26	1.51	1.11
	Val (n = 44)	High (n = 13)	1	6.38	23.25	12.47	5.81
		Low (n = 31)	-1	0.17	3.81	1.47	1.02
Protein <sup>a</sup> (n = 118)	Cal (n = 59)	High (n = 27)	1	12.00	23.67	14.84	2.60
		Low (n = 32)	-1	4.04	11.95	9.28	1.99
	Val (n = 59)	High (n = 28)	1	12.02	21.69	14.91	2.41
		Low (n = 31)	-1	4.70	11.86	9.33	1.87
Sugar <sup>b</sup> (n = 120)	Cal (n = 60)	High (n = 24)	1	16.66	58.06	36.81	11.21
		Low (n = 36)	-1	0.00	10.71	3.66	3.39
	Val (n = 60)	High (n = 21)	1	12.9	36.01	20.97	7.32
		Low (n = 39)	-1	0.00	10.71	2.49	3.66

<sup>a</sup> Measured by analysis.

<sup>b</sup> Derived from product label.

<sup>c</sup> Code 1 means high percent of component (>12% for TDF, sugar, protein; >5% for fat) and code -1 means low percent of component (<12% for TDF, sugar, protein; <5% for fat).

accuracy of sample recognition. In the development of SIMCA and PLS regression models, separate calibration and validation sample sets were used; these included 60 TDF, 46 fat, 59 protein, and 60 sugar samples (calibration set) and 60 TDF, 44 fat, 59 protein, and 60 sugar samples (validation set). For PLS regression models, the high percent samples were arbitrarily ascribed a value of 1.0 for the dummy variables and the low percent samples were ascribed a value of -1.0. After model development using a calibration set and prediction on a validation set, those samples with a predicted value of  $\geq 0$  were identified as being in the high percent sample. All samples with a predicted value  $< 0$  were identified as being in the low percent sample. A separate PLS regression model was developed for each component.

## RESULTS AND DISCUSSION

### FT-Raman Spectra

Figure 1 displays the Raman spectra of cereal samples. The spectra are free of fluorescence effects (sloping baseline) and baseline corrections and normalization performed. The figures show the differences between samples with high and low percent for each component. For high TDF sample (Fig. 1A), strong band at  $\approx 1600\text{--}1630\text{ cm}^{-1}$  that corresponds to the aromatic ring quadrants stretching (Lin-Vien et al 1997) were observed. The bands due to aromatics arise because lignins and other phenolics are a part of dietary fiber in brans and synthetic mixes included in the calibration to obtain higher fiber samples. The higher intensity band at  $\approx 1090\text{ cm}^{-1}$  is related to the cellulose. For the high fat sample (Fig. 1B), Raman shift in the regions of  $1265\text{--}1750\text{ cm}^{-1}$  and  $2850\text{--}2900\text{ cm}^{-1}$  were prominent. The bands correspond to the major Raman scattered bands of an edible oil spectrum (Baeten et al 1998). The bands at  $\approx 1265$  and  $1300\text{ cm}^{-1}$  are associated with C-H vibration of *cis* RCH=CHR and  $-\text{CH}_2$  group. Bands at  $\approx 1440$  and  $2853\text{--}2900\text{ cm}^{-1}$  correspond to C-H vibration of  $-\text{CH}_2$  or  $-\text{CH}_3$  group. Raman shift at  $\approx 1670\text{--}1750\text{ cm}^{-1}$  is assigned to C=C or C=O vibration of *cis/trans* RCH=CHR or RC=OOR group of fatty acid. The high protein sample (Fig. 1C) showed high intensities at  $\approx 1654\text{ cm}^{-1}$  due to amide I and at  $2910\text{ cm}^{-1}$  due to  $\text{CH}_2$  stretching mode (Lin-Vien et al 1997). The  $480\text{ cm}^{-1}$  band shown in both the high fat and high protein samples is due to a C-O vibration of starch (Parker 1983). At  $\approx 202\text{--}1500\text{ cm}^{-1}$ , the Raman scattering intensity is higher for the high sugar samples, particularly at  $\approx 401$  and  $848\text{ cm}^{-1}$ , which are clearly correlated with C-O or C-C vibrations of sucrose. Actually, many of the bands have contributions from multiple functional groups, thus making assignments very difficult in most cases.

TABLE II  
Correlation Coefficients Between Nutritional Components  
in Cereal Products

	TDF	Fat	Protein	Sugar
TDF	1	-0.24	0.11	-0.24
Fat		1	-0.07	-0.08
Protein			1	-0.63
Sugar				1

TABLE III  
Five Different Sample Groups Selected from Original Sample Set

	TDF	Fat	Protein	Sugar
Group A ( $n = 8$ )	0.9 ~ 4.3% (-1) <sup>a</sup>	0.1 ~ 2.5% (-1)	7.1 ~ 11.0% (-1)	0.0 ~ 9.0% (-1)
Group B ( $n = 6$ )	31.2 ~ 49.2% (1)	0.4 ~ 3.2% (-1)	5.7 ~ 11.6% (-1)	0.0 ~ 9.0% (-1)
Group C ( $n = 7$ )	2.0 ~ 10.9% (-1)	11.4 ~ 25.6% (1)	9.0 ~ 10.8% (-1)	0.0 ~ 10.7% (-1)
Group D ( $n = 14$ )	2.7 ~ 10.7% (-1)	0.3 ~ 3.5% (-1)	13.0 ~ 18.6% (1)	0.0 ~ 8.8% (-1)
Group E ( $n = 15$ )	0.5 ~ 8.7% (-1)	0.2 ~ 3.8% (-1)	4.0 ~ 10.2% (-1)	21.8 ~ 53.3% (1)

<sup>a</sup> Codes (in parentheses) as defined in Table I.

### Principal Component Analysis

Principal component analysis (PCA) was conducted using the calibration data sets to extract the primary information in the spectral data. Figure 2 shows a biplot of PCA scores and loadings on  $y$ -variables of the samples. Most of the samples were located near the center, except for few high percent samples. The directions of the loadings showed no intercorrelation between components. The correlation coefficients between the parameters shown in Table II confirmed this result.

Figure 3 shows the PCA scores on  $x$ -variables of the calibration samples on the plane formed by the first two principal components, in which the entire  $x$ -variables of  $202\text{--}3400\text{ cm}^{-1}$  were used for analysis. For TDF, fat, and protein, separation between two groups of class 1 and -1 was not clear. The first two PC explained at  $\approx 63\text{--}64\%$  of the  $x$ -variables. The score plot for sugar samples was slightly better than those of the other components and the  $x$ -variables explained at  $\approx 76\%$  by the first two PC.

Kays et al (1996, 1997, 1998) reported interference from fat and sugar in the development of the NIR model to predict fiber in cereal products. To investigate the interference by other components in this study, we attempted to perform PCA using samples that vary only in the amount of one target component. For this experiment, five different sample groups were selected from the original data set (Table III). Group A consists of samples with lower percents in all four components, whereas other groups (B, C, D, E) consist of samples with higher percents in only one component relative to the others; high TDF (group B), high fat (group C), high protein (group D), and high sugar (group E). Figure 4 shows how the first two PC are able to separate the samples into two groups of low and high percents (A-B, A-C, A-D, A-E). The variables most closely related to the first PC for characterizing the group of samples were also shown in each score plot. The variables displayed large loading values on this component and were the variables related to the existing variability in the direction of the first PC. For group A and B with a TDF content difference, two groups were clearly separated from each other and the most responsible  $x$ -variables for PC1 were at  $1600$  and  $1630\text{ cm}^{-1}$  (Fig. 4A). High fat samples (group C) were well classified from the low fat samples (group A) and the variables of  $1440$  and  $2853\text{ cm}^{-1}$  had high loading values (Fig. 4B). Groups D and A with a protein content difference were well distinguished, even though some of samples in group D were close to those in group A (Fig. 4C). The variables of  $2910$  and  $1660\text{ cm}^{-1}$  were relevant for characterizing group D. Groups A and E that have a sugar content difference were clearly separated and  $x$ -variables most highly correlated to sample grouping were  $401$  and  $848\text{ cm}^{-1}$  (Fig. 4D). The  $x$ -variables with high variation in loadings for each component corresponded well to the absorption bands related to the chemical composition in Fig. 1. In each frame, group A showed a high loading value at  $476\text{ cm}^{-1}$ , which is associated with starch (Parker 1983). This is due to the fact that the group A has a lower percent in TDF, fat, protein, and sugar but a higher percent in starch compared with the other groups. The PCA results suggest that classification of nutritional components in cereal products is subject to interference from other components in the sample.

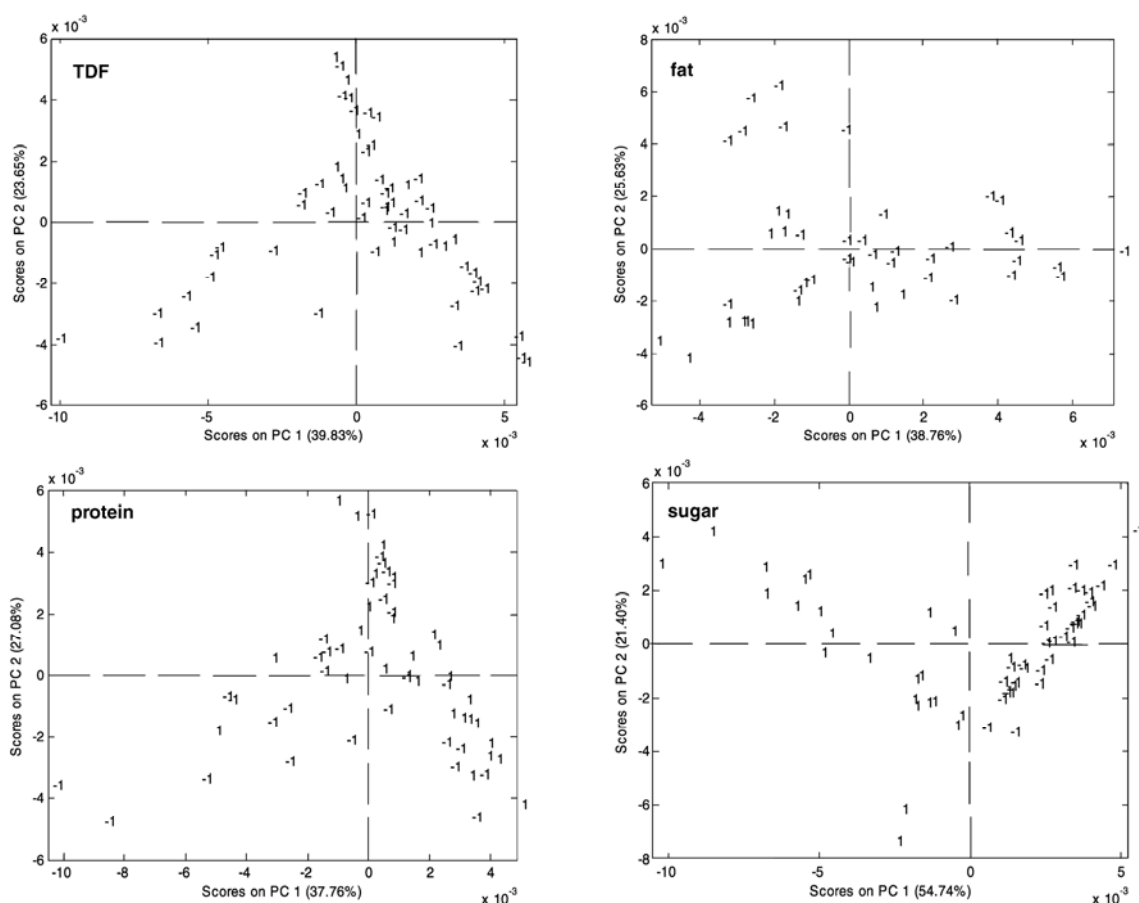
### SIMCA Classification

The above observations suggest that selection of  $x$ -variables for each component and its use may be helpful for classification. Therefore, it was of interest to examine whether the SIMCA result for classifying the cereal samples according to higher and lower percents of each component could be improved by the use of the selected regions instead of the entire region. Selection of  $x$ -variables was accomplished through examination of PCA loading plots and normalized spectral plots. The selected ranges are 1070–1140 and 1500–1720  $\text{cm}^{-1}$  for TDF, 1200–1800 and 2700–3000  $\text{cm}^{-1}$  for fat, 1550–1720 and 2770–3090  $\text{cm}^{-1}$  for protein, and 202–970  $\text{cm}^{-1}$  for sugar. The summary results for SIMCA models using the entire region and the selected regions are shown in Table IV. The performance of all models showed improvement by the use of the selected regions as would be expected, except for the TDF model that produced the same prediction accuracy with the same number of PC. For the protein and sugar data, one outlier that was based on high  $T^2$  (Hotelling's

$T^2$  statistic) and high  $Q$  residuals was detected and excluded for the analysis. The  $T^2$  means square of distance of sample from model mean that can be explained by normal variations within the group. The  $Q$  residuals mean distance of sample from the model that can be explained by random error. The correct classification of validation samples showed >90% for TDF, fat, and protein. The result for sugar was far less satisfactory (59%) compared with those of the other components.

### PLS Regression-Based Classification

Table V and Fig. 5 show the results from the PLS regression for each component. The use of the selected  $x$ -variables produced equal or higher accuracy than using the entire region for all four components, resulting in a reduced number of PC or an increased accuracy or both. For the TDF model, the number of PC was decreased to one, but there was no change in predictive accuracy, showing 92%. The wrongly identified samples resulting from this model were five high-percent samples classified as low-percent

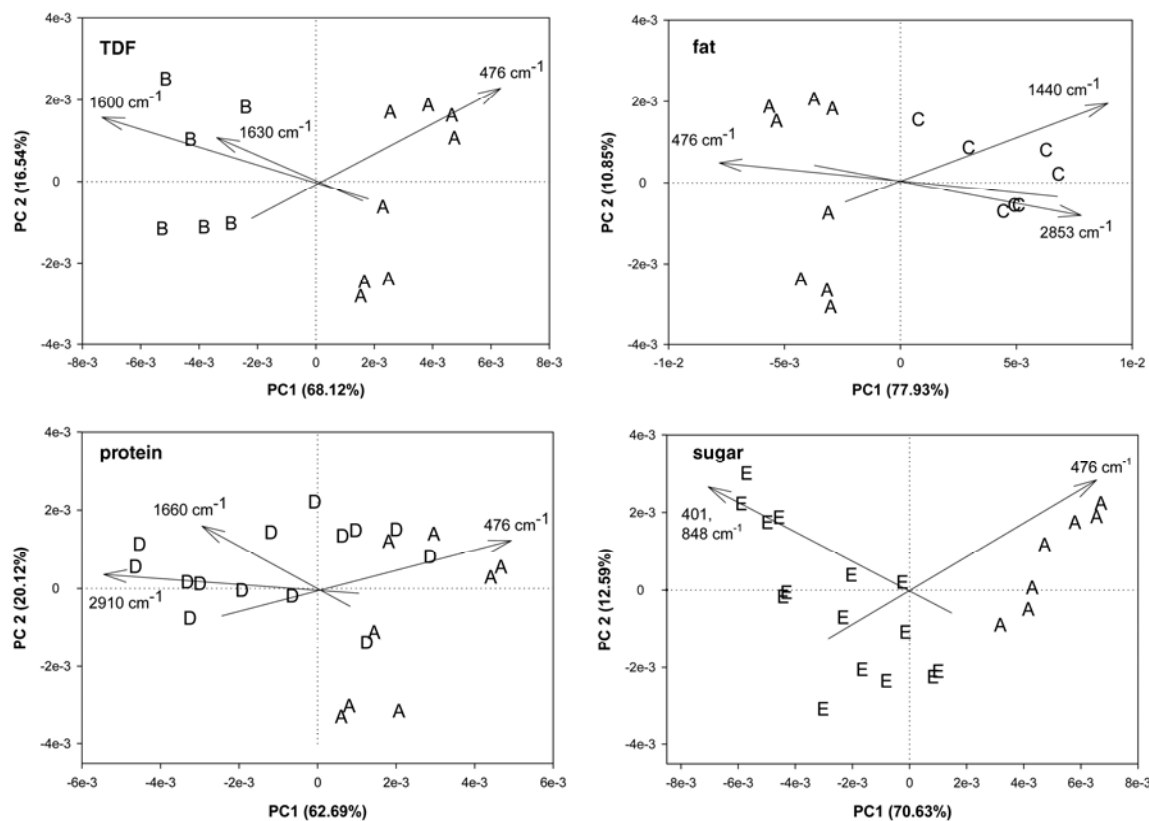


**Fig. 3.** PCA score plots for classification of TDF, fat, protein, and sugar in cereal products by NIR-FT/Raman spectroscopy. Results were obtained using a calibration sample set with the entire spectral region. Values of -1 and 1 mean low and high percent samples, respectively.

**TABLE IV**  
Summary Results of SIMCA Models for Nutritional Components of Cereal Samples

Components	Raman Shifts Used ( $\text{cm}^{-1}$ )	PC	Validation Samples	Correct Samples	Incorrect Samples	Recognition Accuracy (%)
TDF	202–3,399	5	60	54	6	90
Fat	202–3,399	2	44	39	5	89
Protein	202–3,399	4	58 <sup>a</sup>	44	14	76
Sugar	202–3,399	3	59 <sup>a</sup>	33	26	56
TDF	1,070–1,140, 1,500–1,720	5	60	54	6	90
Fat	1,200–1,800, 2,700–3,000	2	44	40	4	91
Protein	1,550–1,720, 2,770–3,090	3	58 <sup>a</sup>	52	6	90
Sugar	202–970	2	59 <sup>a</sup>	35	24	59

<sup>a</sup> One outlier based on high  $T^2$  and high  $Q$  residuals was detected and excluded from the data set.



**Fig. 4.** PCA score plots for classification of TDF, fat, protein, and sugar in cereal products by NIR-FT/Raman spectroscopy. Results were obtained using the selected sample sets in Table III with the entire region. Arrows toward B, C, D, E indicate the most highly correlated  $x$ -variables for each component and arrow toward A in each frame indicates starch related  $x$ -variable.

**TABLE V**  
Summary Results of PLS Regression Models for Nutritional Components of Cereal Samples

Components	Raman Shifts Used ( $\text{cm}^{-1}$ )	PC	Validation Samples	Correct Samples	Incorrect Samples	Recognition Accuracy (%)
TDF	202–3,399	2	60	55	5	92
Fat	202–3,399	3	44	42	2	95
Protein	202–3,399	4	58 <sup>a</sup>	48	10	83
Sugar	202–3,399	4	59 <sup>a</sup>	44	15	75
TDF	1,070–1,140, 1,500–1,720	1	60	55	5	92
Fat	1,300–1,760, 2,800–2,900	3	44	42	2	95
Protein	1,550–1,720, 2,770–3,090	4	58 <sup>a</sup>	54	4	93
Sugar	202–970	3	59 <sup>a</sup>	50	9	85

<sup>a</sup> One outlier based on high  $T^2$  and high  $Q$  residuals was detected and excluded from the data set.

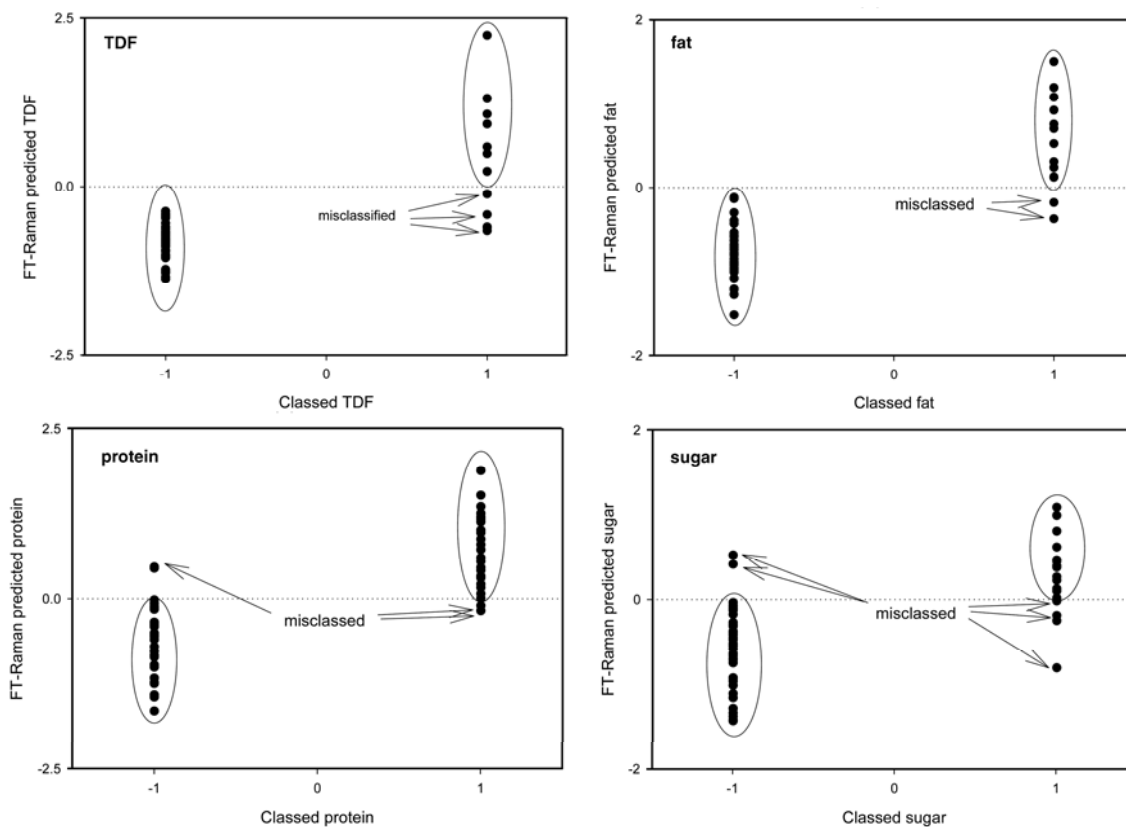
(Fig. 5A); their actual percents were 13–17%. In Fig. 5, the ovals indicate the correctly identified samples, whereas the arrows indicate the wrongly identified samples. The prediction accuracy of the fat model was 95% using three PC. Two wrongly identified samples were found from the high fat sample set (Fig. 5B) and their actual fat contents were 5.3 and 5.6%, respectively. These values are very close to the cut off value of 5%. For the protein and sugar, the improvement of model performance by the use of the selected region was more prominent than other components. The recognition accuracy was 93% for protein using a four PC model and four samples were misclassified (Fig. 5C), their actual fat percents were 11–13%, near to the cut off value of 12%. The classification accuracy for sugar was 85% using three PC and nine misclassified samples were found (Fig. 5D).

Consequently, PLS-based classification gave better results than SIMCA for classification of the four nutritional components. This result corresponds to the report by McElhinney et al (1999), where the PLS discrimination also produced better results than SIMCA in classification performance for species identification in selected raw homogenized meats.

The fact that sugar classification was not as good as that for the other components might be related to the use of values obtained from product labels rather than more accurate laboratory analysis. Further study using an accurate direct assay could likely result in better classification of sugar.

## CONCLUSIONS

FT-Raman spectroscopy has the potential as a viable method to classify the cereal foods into two groups (high or low) according to the percent level of each component with a prediction accuracy of 85–95%. Classification performance of linear calibration algorithms was successful in the presence of high nonlinearity due to product type. In general, classification performance was improved by the use of the selected regions. PLS-based classification produced better results than SIMCA for all four components. NIR-FT/Raman spectroscopy provides, with 5-min total sample preparation and analysis time, a more rapid means of classification for additional (or new) products based on spectra than time-consuming laboratory assays ( $\approx 1\text{--}2$  days) for fiber, protein, fat, and sugar.



**Fig. 5.** Prediction results of cereal samples with PLS regression-based classification model developed using a calibration set with the selected regions. Ovals indicate correctly identified samples and arrows indicate wrongly identified samples.

#### ACKNOWLEDGMENTS

We are grateful to F. L. Poole for technical assistance in data acquisition and generation of Array-Basic programs.

#### LITERATURE CITED

- AOAC 1992. Method 991.43. Official Methods of Analysis, 15th Ed. 3rd Supplement. AOAC International: Arlington, VA.
- AOAC 1995. Methods 992.23 and 945.16. Official Methods of Analysis, 16th Ed. AOAC International: Arlington, VA.
- Archibald, D. D., Kays, S. E., Himmelsbach, D. S., and Barton, II, F. E. 1998a. Raman and NIR spectroscopic methods for determination of total dietary fiber in cereal foods: A comparative study. *Appl. Spectrosc.* 52:22-31.
- Archibald, D. D., Kays, S. E., Himmelsbach, D. S., and Barton, II, F. E. 1998b. Raman and NIR spectroscopic determination of total dietary fiber in cereal foods: Data preprocessing and calibration. *Appl. Spectrosc.* 52:32-41.
- Baeten, V., Hourant, P., Morales, M. T., and Aparicio, R. 1998. Oil and fat classification by FT-Raman spectroscopy. *J. Agric. Food Chem.* 46:2638-2646.
- Colthup, N. B., Daly, L. H., and Wiberley, S. E. 1990. Introduction to Infrared and Raman Spectroscopy, 3rd Ed. Academic Press: San Diego, CA.
- Jones, D. B. 1931. Factors for converting percentages of nitrogen in foods and feeds into percentages of proteins. Circular 183. USDA: Washington, DC.
- Kays, S. E., Windham, W. R., and Barton, II, F. E. 1996. Prediction of total dietary fiber in cereal products using near-infrared reflectance spectroscopy. *J. Agric. Food Chem.* 44:2266-2271.
- Kays, S. E., Windham, W. R., and Barton, II, F. E. 1997. Prediction of total dietary fiber by near-infrared reflectance spectroscopy in cereal products containing high sugar and crystalline sugar. *J. Agric. Food Chem.* 45:3944-3951.
- Kays, S. E., Windham, W. R., and Barton, II, F. E. 1998. Prediction of total dietary fiber by near-infrared reflectance spectroscopy in high-fat- and high-sugar-containing cereal products. *J. Agric. Food Chem.* 46:854-861.
- Kays, S. E., Barton, II, F. E., and Windham, W. R. 2000. Predicting protein content by near infrared reflectance spectroscopy in diverse cereal food products. *J. Near Infrared Spectrosc.* 8:35-43.
- Kays, S. E., Archibald, D. D., and Sohn, M. 2005. Prediction of fat content in intact cereal food products using NIR reflectance spectroscopy. *J. Sci. Food Agric.* 85:1596-1602.
- Lee, S. C., Prosky, L., and De Vries, J. W. 1992. Determination of total, soluble, and insoluble dietary fiber in foods enzymatic-gravimetric methods, MES-TRIS buffer: Collaborative study. *J. AOAC Int.* 75:395-416.
- Lin-Vien, D., Colthup, N. B., Fateley, W. G., and Grasselli, J. G. 1997. The Handbook of Infrared and Raman Characteristic Frequencies of Organic Molecules. Pages 477-479 in Appendix 3. Academic Press: San Diego, CA.
- McElhinney, J., Downey, G., and Fearn, T. 1999. Chemometric processing of visible and near infrared reflectance spectra for species identification in selected raw homogenized meats. *J. Near Infrared Spectrosc.* 7:145-154.
- Parker, F. S. 1983. Carbohydrates. Pages 315-347 in: Applications of Infrared, Raman and Resonance Raman in Biochemistry. F. S. Parker, ed. Plenum Press: New York.

[Received March 23, 2005. Accepted July 20, 2005.]