Determination of Total Dietary Fiber of Intact Cereal Food Products by Near-Infrared Reflectance

Douglas D. Archibald*,† and Sandra E. Kays

Quality Assessment Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Russell Research Center, Athens, Georgia 30604-5677

Near-infrared reflectance spectra of cereal food products were acquired with a commercial dual-diode-array (Si, InGaAs) spectrometer customized to allow rapid acquisition of scans of intact breakfast cereals, snack foods, whole grains, and milled products. Substantial gains in the performance of multivariate calibration models generated from these data were obtained by a computational strategy that systematically analyzed the performance of various spectral windows. The calibration model based on 137 cereal food products determined the total dietary fiber (TDF) content of a test set of 45 intact diverse cereal food products with root-mean-squared error of cross-validation of between 1.8 and 2.0% TDF, relative to the laborious enzymatic−gravimetric reference method. The calibration performance is adequate to estimate TDF over the range of values found in diverse types of cereal food products (0.7−50.1%). The method requires no sample preparation and is relatively unaffected by specimen moisture content.

Keywords: *Total dietary fiber; cereal food products; nondestructive analysis; near-infrared reflectance; spectroscopy; InGaAs diode array; multivariate calibration; spectral window*

INTRODUCTION

The dietary fiber component of foods is important for the health of individuals and the public at large (Gorman and Bowman, 1993; Marlett and Slavin, 1997), and fiber content often is used in the marketing of products. As a consequence, the Nutrition Labeling and Education Act of 1990 requires food producers to report total dietary fiber (TDF) on consumer product labels (United States Congress, 1995). In the United States, the current accepted definition for TDF is the sum of lignin and polysaccharides not digestible by human alimentary enzymes as estimated by AOAC method 991.43, which approximates the enzymatic aspects of human digestion by gravimetric analysis of the residue after a sequence of timed enzymatic treatments (AOAC, 1990; Lee et al., 1992; AOAC, 1992).

Fiber assays are time-consuming and often produce a considerable amount of chemical waste, so it is not surprising that efforts have been made to develop instrumental methods. Technology for near-infrared reflectance (NIRR) analysis of dietary fiber was reviewed by Kays et al. (1999). Briefly, Baker published the first reports on estimation of neutral detergent fiber (NDF) in human foods in the early 1980s (Baker, 1983; Baker, 1985). However, the near-infrared (NIR) analysis of fiber in animal forage began in the mid-1970s (Norris et al., 1976; Marten and Templeton, Jr., 1989) and culminated in a handbook in 1985 (USDA, 1985; USDA, 1989), and an official method for acid-detergent fiber (ADF) in 1988 (Barton, II, and Windham, 1988). This laboratory recently has developed spectroscopic methods for dietary fiber in human foods; these methods apply to a broad range of cereal food products using milled specimens (Barton, II, et al., 1995; Kays et al., 1996; Windham et al., 1997; Kays et al., 1997; Archibald et al., 1998a; Archibald et al., 1998b; Kays et al., 1998; Kays and Barton, II, 1998; Kays et al., 2000).

The work presented in this manuscript is the first effort to determine TDF from spectral properties of intact cereal food products taken directly from the box or bag. The study started with the proposal that it would be possible to create a single TDF calibration from NIRR spectra of diverse intact cereal food products. One objective of the analysis was to determine if there are classes of products that are more difficult to model. NIRR measurements were made with a commercial spectrometer that incorporates silicon and InGaAs diode arrays spanning the range from 400 to 1700 nm. Due to the high sensitivity of InGaAs diode arrays, these devices are finding use in applications of NIR spectroscopy that require rapid remote reflectance measurements (Huthfehre et al., 1995). In the study presented here, the rapid measurement capability was used to obtain multiple optical samples representing NIRR from a large surface area of each cereal food product, because many of the products were highly heterogeneous. The NIR portion of the spectral range measured by the silicon and InGaAs arrays (800−1700 nm) is suited to this application because it has more penetrating ability than the longer wavelength NIR. The geometry of the measurement also may be appropriate for at-line or on-line process measurements for a variety of constituents.

MATERIALS AND METHODS

Samples and Sample Treatments. One-hundred thirty-six cereal food products, including breakfast products, snack foods, flours, and baking mixes, were obtained from local retail stores. Few retail products

* E-mail: dda10@psu.edu. Telephone: (814) 865−8449. Fax: (814) 863−7043.

† Current address: Agronomy Department, 116 Agricultural Sciences and Industries Building, The Pennsylvania State University, University Park, PA 16802.

10.1021/jf000206j
Published on Web 08/31/2000

Table 1. Cereal Food Product Groups Represented in the Study after Removal of Five Outliers

	product class	calibr. set	test set	full set
I	[a]breakfast cereals without fruit or nuts	44	13	57
II	[a]breakfast cereals with fruit or nuts	13	5	18
III	crackers, chips and biscuits	11	10	21
IV	grain flour and baking mixes	7	3	10
V	whole and crushed grains	10	10	20
VI	predominantly bran	7	4	11
	total	92	45	137

[a] Not including breakfast products found in other classes.

Table 2. Composition and Property Variation of Untreated Cereal Food Products in the Study

	moisture %[a]	TDF %[b]	protein %[c]	sugar %[d]	fat %[d]	density g/mL[e]
minimum	2.3	0.7	4.0	0.0	0.0	0.056
maximum	13.7	50.1	22.3	55.6	25.0	0.842
range	11.3	49.4	18.3	55.6	25.0	0.786
mean	6.4	10.2	11.8	13.5	4.9	0.319
median	5.1	7.6	11.8	8.8	3.2	0.227
# samples	137	137	137	137	137	113

[a] By oven method. [b] Total dietary fiber by AOAC method as described in Materials and Methods. [c] Combustion nitrogen analysis × 6.25. [d] As determined from the product nutrition labels. [e] Bulk product density by mass/volume measurements.

were found in the middle to upper range of TDF. Therefore, the sample set was supplemented by creating 6 simulated cereal products from 6 real products by binary mixing, resulting in a total of 142 untreated intact cereal food products. Five of the 142 were identified as outliers, as described later. As outlined in Table 1, approximately one-third of the 137 untreated samples was reserved for testing, and the remainder was used to optimize the parameters of the calibration model. Sample composition was diverse (Table 2). Therefore, available untreated samples were distributed between calibration and test sets to achieve similar sample histograms for the constituents TDF, protein, sugar, and fat, based on either reference measurements or the values obtained from product nutrition labels.

Because the condition of the sample may not be controllable by the analyst in some applications, calibration development involved specimens that were treated to simulate varying moisture content levels, and degrees of breakage of products. The spectra of treated specimens were used to characterize and stabilize the model response. Treated specimens were created by reducing the particle size of products (by crushing in heavy plastic bags) or by altering the moisture content by desiccation or contact with moist air. The particle size and moisture treatments expanded the number of specimens from 137 to 361, and increased the moisture content range substantially to 2.2−21.4%, with the mean and median moisture contents each increasing by 1.1% over the untreated set described in Table 2. Although the character of many products was altered substantially by crushing, the range, mean, and median of product bulk density did not change significantly after particle size treatments, because breakfast cereal products were the primary products treated, and these had low-to-intermediate bulk densities. When treated samples were used to stabilize or test models, they were assigned to the same set, calibration or test, as their untreated counterpart.

NIRR Sampling Strategy and Data Collection. Accurate subsampling was required to split intact cereal food specimens for treatments or reference analyses. To do this, typically three to six boxes or bags of a product were mixed thoroughly and then poured out at a constant rate on a rotating platform. The resulting doughnut-shaped pile was split along a diameter to form two subsamples. To obtain a representative optical subsample for a product, the product was packed onto the measuring window eight individual times. For each "repack", the instrument was set to collect four rescans of the stationary specimen. In most instances, the regression models were generated from the average spectrum for each specimen (eight repacks by four rescans). The repeat data was used to analyze the variance, within the calibration model, that was associated with rescans or repacks.

The NIRR spectra were collected with a Perten Instruments model DA7000 spectrometer (Springfield, IL). This instrument has silicon and InGaAs diode arrays and an intense broadband light source, making it possible to measure reflectance from a large area of the specimen surface (about 10 cm diameter). The diodes are centered at 10-nm intervals, but software is used to spline-interpolate spectra to a data interval of 5 nm. After acquiring a reference scan of Spectralon (Labsphere, Inc., North Sutton, New Hampshire), the instrument returned the $-\log[R_{sample}/R_{reference}]$ for the sample, with the two spectral regions spliced at 950 nm to cover the range from 400 to 1700 nm. Although the instrument internally averages 30 spectra per second of acquisition time, for the purposes of this study, a spectral scan was defined as the spectrum generated by the instrument after a 1-s acquisition. To facilitate the loading and unloading of specimens, the spectrometer was operated in side-view mode with the addition of a custom-built hopper with a drop hatch and an adjustable sample thickness (2.5−10 cm) (Figure 1). With this design, the Spectralon reference disk also can be inserted in place of the sample. The manufacturer recommends taking a reference scan every hour, or sooner if suggested by the automatic instrument diagnostics. In the present study, reference scans were collected more frequently, typically after every second sample (two by eight repacks).

Reference Analysis of Total Dietary Fiber. The TDF of the specimens was determined by AOAC method 991.43 as described previously (Lee et al., 1992; AOAC, 1992; Kays et al., 1998). Not every sample was subjected to an identical analysis procedure: for products high in fat (>10%) or sugar (>20%), solvent extraction procedures were used to either defat or desugar the sample prior to application of the enzymatic−gravimetric method. An AOAC validation study for method 991.43 yielded a mean repeatability (standard deviation of blind duplicates) of 0.71% TDF and a mean interlaboratory reproducibility of 1.28% for three cereal products analyzed by 11 laboratories (Lee et al., 1992). The standard error of the laboratory (SEL) for application of the AOAC TDF reference method in this study was 0.39% when calculated from 46 of the cereal food product specimens with replicates performed on different days. SEL is the pooled standard deviation of the repeatability (ASTM, 1995).

Calibration Model Development. Data processing was performed in the MATLAB computing environment (The MathWorks, Natick, MA) using custom subroutines that incorporated de Jong's PLS algorithm (de Jong, 1993) and the Savitzky−Golay derivative (Sav-

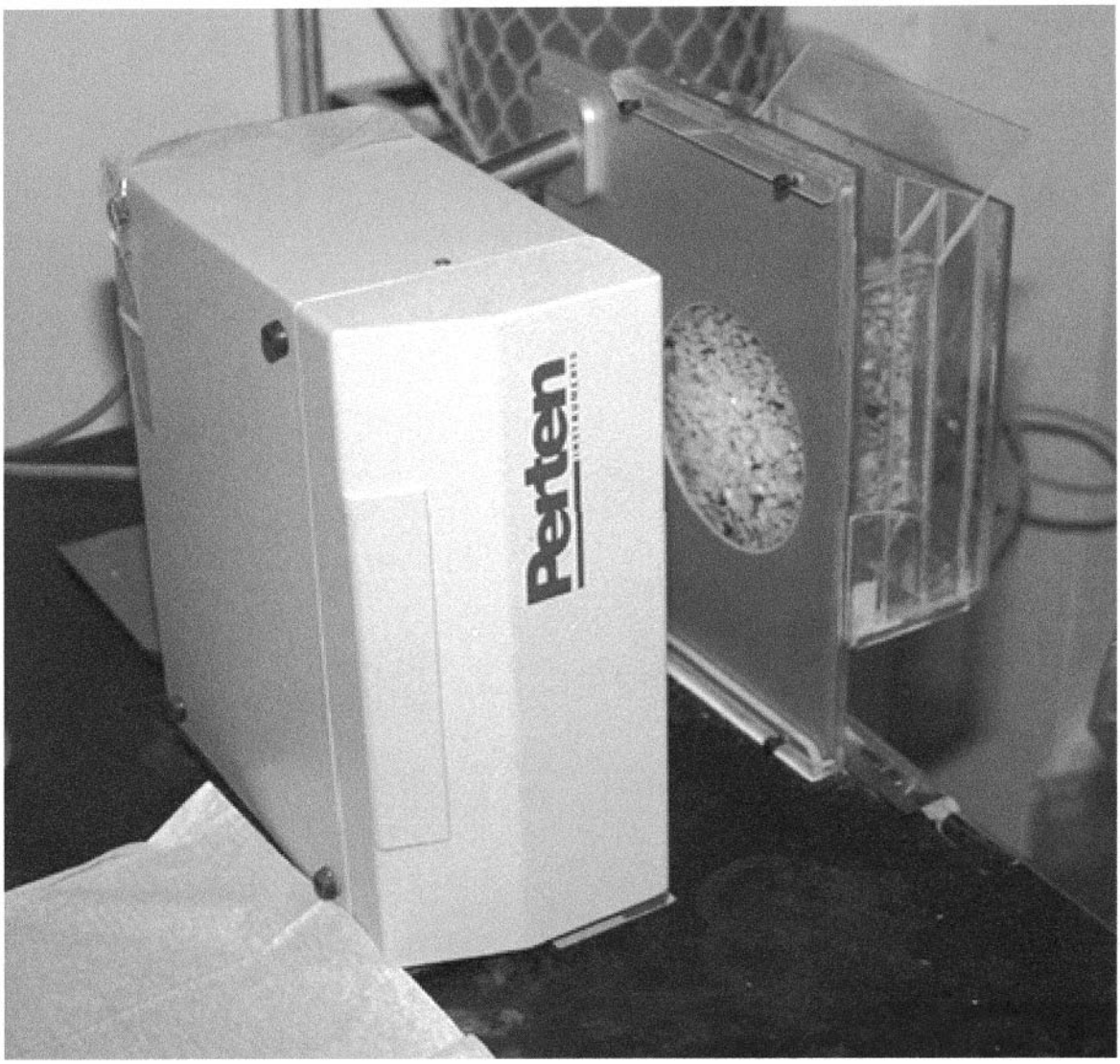

Figure 1. InGaAs/silicon dual-diode-array NIR spectrometer and hopper accessory for rapid loading and measurement of the reflectance of a large surface area of each cereal food product.

itzky and Golay, 1964; Madden, 1978) MATLAB functions available in the PLS Toolbox 2.0 (Eigenvector Research, Manson, WA). Preliminary analysis of the data led to the conclusion that interference from color, hydration effects, or irrelevant chemical variations was larger than that due to instrument noise or scattering differences of the various intact products. Furthermore, because interference was not evenly distributed across the spectrum, it was found that selecting subsets of wavelength variables substantially improved the calibration performance. Establishing which variables to retain or delete, however, was not trivial because of extensive band overlap and mixing, the insufficient knowledge about the spectral properties of the various components of dietary fiber, and the presence of a highly variable background.

To make variable selection more systematic, a computational strategy called spectral window preprocessing was developed. This procedure extensively tested the data set to determine which spectral windows to include and which windows to delete (Archibald and Akin, 2000). In this study, the spectral window preprocessing technique was implemented as described below. The calibration set of untreated samples (n = 92) was evaluated first by conventional full-spectral partial least-squares regression (PLSR) (Martens and Næs, 1989) to remove obvious outliers and to establish the optimal spectral transformation. The latter was determined by comparison of the calibration-set root-mean-squared error of cross-validation (RMSECV) (Martens and Næs, 1989) produced by various types of spectral preprocessing. Second, a spectral window deletion experiment was performed where model leave-one-out cross-validation statistics were compiled for a triangular array of spectral window widths (5, 10, 15, ..., 400 nm) and window positions (beginning at 800, 805, ..., 1695 nm). The optimal range to delete was determined from the window parameters of the model with the lowest RMSECV. The next step was to compute a spectral window addition experiment, with widths of 400, 405, ..., 900 nm and beginning positions of 800, 805, ..., 1300 nm, with all of the data except for the variables of the deleted window. Again, the global minimum RMSECV was used to determine the best limits, which in this case established the spectral range. These exhaustive computational testing experiments involved 92 samples and 181 spectral variables and may be calculated in a few days with currently available desktop computer technology, although the computation time can be reduced easily by changing the spectral

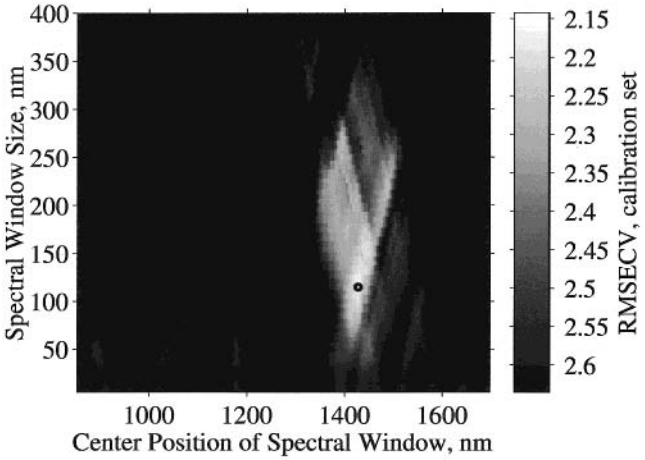

Figure 2. Response image for a spectral window deletion experiment for the prediction of TDF of freshly opened, intact cereal food products from SNV-transformed second derivative NIRR calibration spectra ($n = 92$). The gray level of each pixel represents the RMSECV of the calibration set for a 13-factor PLSR model when a spectral window of a given size (ordinate) and position (abscissa) was deleted from the matrix of calibration data. The highest error of the color bar (darkest shade) was set to the value of RMSECV obtained using the full spectral range (2.63% TDF for 850−1700 nm). The bullet symbol indicates the optimal region to exclude.

window limits, the window step size, or the number of cross-validation segments. Predictive models were generated from the optimized set of PLSR model parameters using untreated and treated calibration samples and were evaluated with the test set.

RESULTS AND DISCUSSION

Development and Testing of Calibration Models with Untreated Samples. Initial attempts to develop a TDF calibration model with the use of either underivatized or derivatized −log(R) spectra identified 3 of the 95 original untreated calibration samples that consistently were outliers. One specimen (toasted wheat germ) was removed because its protein content was far higher than any other product. A second was removed because of its extremely different optical properties (whole popcorn kernels). A third (buckwheat groats) was omitted because the TDF predictions were always extremely high compared to the measured value. Of the preliminary models evaluated with the untreated calibration set, digital second derivative (D2) spectral processing followed by standard normal variate transformation (Barnes et al., 1989) (SNV) and mean centering yielded the best PLSR cross-validation error for the NIR region from 800 to 1700 nm. The derivative was calculated by the Savitzky−Golay convolution method (Savitzky and Golay, 1964; Madden, 1978) with a width of nine data points (45 nm) and a third-order polynomial fit. To improve these results, two spectral window computational experiments were conducted to optimize the selection of the spectral range. First, a PLSR window deletion experiment on D2-SNV data was calculated, where the SNV normalization was performed for each tested set of wavelength variables (Figure 2). This showed that substantial improvement in RMSECV, from 2.6 to 2.1% TDF, was obtained by deleting the spectral range represented by the bullet symbol in Figure 2. Those spectral window parameters produced the lowest calibration set RMSECV among all of the

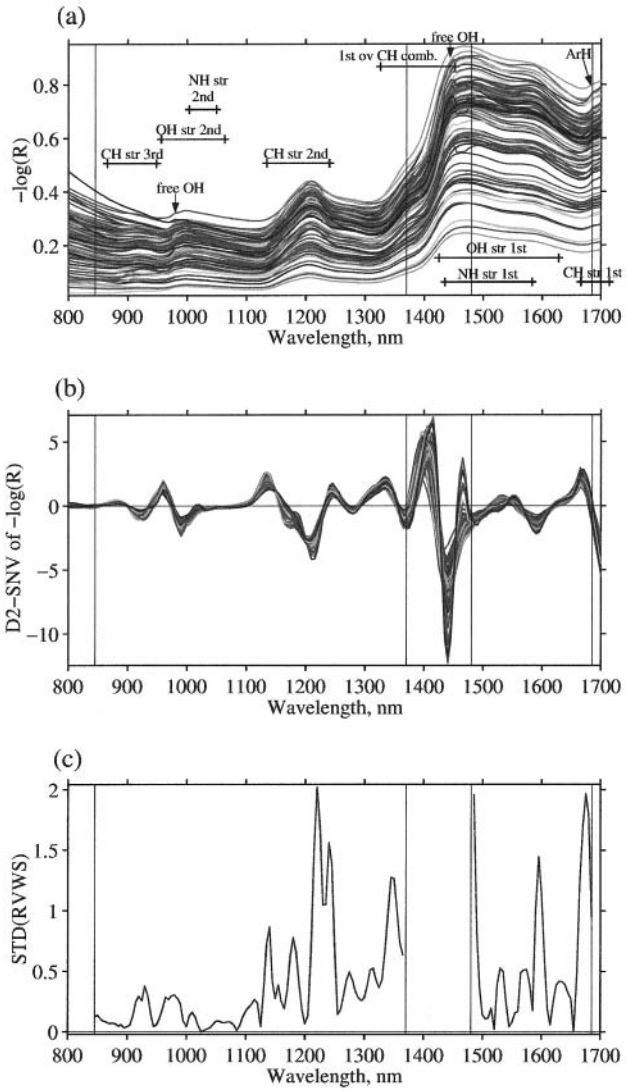

Figure 3. Plots illustrating the transformation of intact cereal food product NIRR spectra for optimization of the linear calibration for TDF: (a) unprocessed calibration spectra ($n = 92$), where R is [$R_{sample}/R_{reference}$]; (b) D2-SNV preprocessed calibration spectra with vertical dotted lines indicating the separation between included and excluded regions; and (c) spectral activity within the model expressed as the standard deviation across samples of the regression-vector-weighted spectra (STD(RVWS)). RVWS are computed by element-by-element multiplication of each preprocessed calibration spectrum by the optimized regression coefficients. Abbreviations: ov = overtone; comb. = combination band; str = overtone of bond stretching vibration.

possible window deletions that formed the grid of the computation. The other computational experiment evaluated which spectral window to include, and this procedure established the best spectral limits for the preprocessing (data not shown). The latter step produced only a slight improvement in error but used two fewer factors. The effect of the preprocessing parameters optimized by the spectral window experiments is shown in Figure 3a,b. The spectral range from 845 to 1685 nm was selected, except for the window from 1370 to 1480 nm. Second derivative preprocessing alone effectively extracted the structural bands from optical artifacts. However, the process of normalizing the relevant bands and extraction of signal with the PLSR algorithm was improved by deleting the highly variable spectral region between 1370 and 1480 nm.

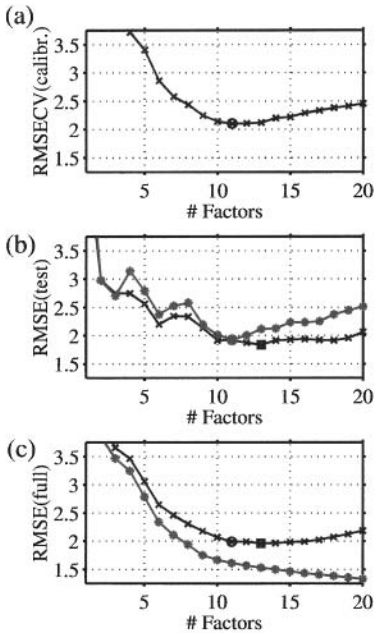

Figure 4. Estimates of the uncertainty of PLSR models for the determination of TDF from NIRR of a diverse set of freshly opened, intact cereal food products: (a) RMSECV for the untreated calibration set ($n_{cal} = 92$) with D2-SNV preprocessing over the optimal spectral range; (b) test set error ($n_{tst} = 45$) estimated by prediction of the test set by the calibration set alone (*), or by cross-validation on the specimens in the test set combined with the calibration set (×); (c) full set error ($n_{cal} + n_{tst}$) estimated as RMSEC (*) or RMSECV (×). The circles mark the error estimate for the optimal model as determined from part a, whereas the square symbol indicates the optimal model by the cross-validation procedure when both the calibration and test samples were included.

The calibration set RMSECV versus the number of factors for the optimal preprocessing parameters is shown in Figure 4a. An 11-factor model was optimal for the 92-member calibration set (RMSECV = 2.10% TDF, $R^2 = 0.96$). On the basis of the residual standard deviation of the spectral residuals, two test set samples were determined to be outliers of the model and so were removed, leaving 45 test samples. One sample was "Low Fat Apple Cinnamon Muffin Mix" and the other was "Cinnamon Toast Crunch" breakfast cereal. The fact that both contain cinnamon is probably not a factor, because several other products contained cinnamon but were not spectral outliers. The muffin mix may be spectrally unique because it contained an unusual component, artificial fruit particles made of dextrose. Similarly, the breakfast cereal may be a spectral outlier as a result of the composition of the crystalline sugar coating. Since the spectral window technique evaluated a large number of wavelength sets, it could be argued that the calibration set cross-validation error estimates are likely to be too low. Therefore, final conclusions about model performance were based on analysis of a large test set that was not used for optimization of the spectral window ranges. An estimate of the model error was generated from the test set by two methods (Figure 4b). By simple prediction of the test set samples using the calibration model based on 92 samples, the root-mean-squared error of prediction (RMSEP) (Martens and Naes, 1989) was 1.92% TDF for an 11-factor model ($R^2 = 0.96$). By leave-one-out cross-validation of each of the 45 test samples (while always retaining the 92 calibration samples), the RMSECV was 1.92% TDF for an 11-factor model and 1.84% for a 13-factor model (R^2 = 0.96). Unlike the aforementioned RMSEP statistic, the test-set cross-validation method estimated the performance of a PLSR model produced from 137 samples (instead of 92). More independent samples generally yields more modeling power, and hence, the data set could support a model with two additional factors to achieve slightly better estimates of the same test samples. The number of factors added is reasonable, as ASTM guidelines recommend an infrared multivariate calibration set contain at least six spectra per regression factor (ASTM, 1995). The number of samples in this application probably was critical because the products have widely varying ranges of composition and optical properties. By use of all of the calibration and test data for untreated samples, the error was estimated to be 1.99% or 1.96% TDF by the RMSECV method for 11- or 13-factor models, respectively (both with R^2 of 0.96) (Figure 4c). The root-mean-squared error of calibration (RMSEC) on all samples was considerably lower (1.61 and 1.53% TDF, Figure 4c). The ratio of TDF range to test set RMSECV was over 25, indicating good discriminating power for the TDF content of intact cereal food products. This also can be judged visually in Figure 5. The model testing and stabilization studies described in the remainder of this report use the preprocessing and spectral windows determined from the 92 untreated calibration samples, and estimate model errors by performing cross-validation on the test set with 13 PLSR factors.

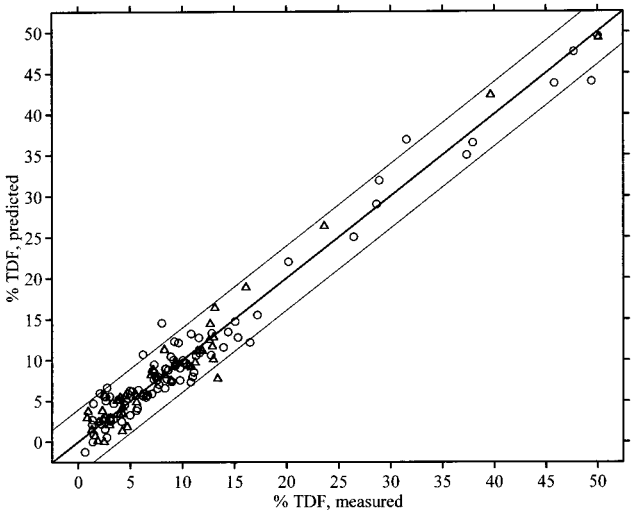

Figure 5. Plot of predicted versus measured TDF values for the 92 untreated calibration samples (circles) and 45 untreated test samples (triangles) as estimated by full cross-validation on all 137 samples. The PLSR predictions were generated with 13 factors using the spectral preprocessing parameters established by computational experiments performed on the untreated calibration samples. The thick solid line is the ideal correlation with an intercept of 0 and slope of 1. The thin lines are parallel to the ideal correlation at plus or minus two times the full-set RMSECV.

Analysis of Repeated Spectral Measurements for Untreated Samples. Repeated measurements of each specimen were taken in two ways: (1) the specimen was scanned four times without removal (a rescan), and (2) the specimen was loaded into the hopper eight times (a repack). Each of the individual scans was preprocessed with the optimized parameters (D2-SNV), and the rescan or repack variances in the model for each specimen were pooled (Figure 6a,b). The noise due to rescans arose primarily from the instrument and should be extremely low because of negligible changes in the

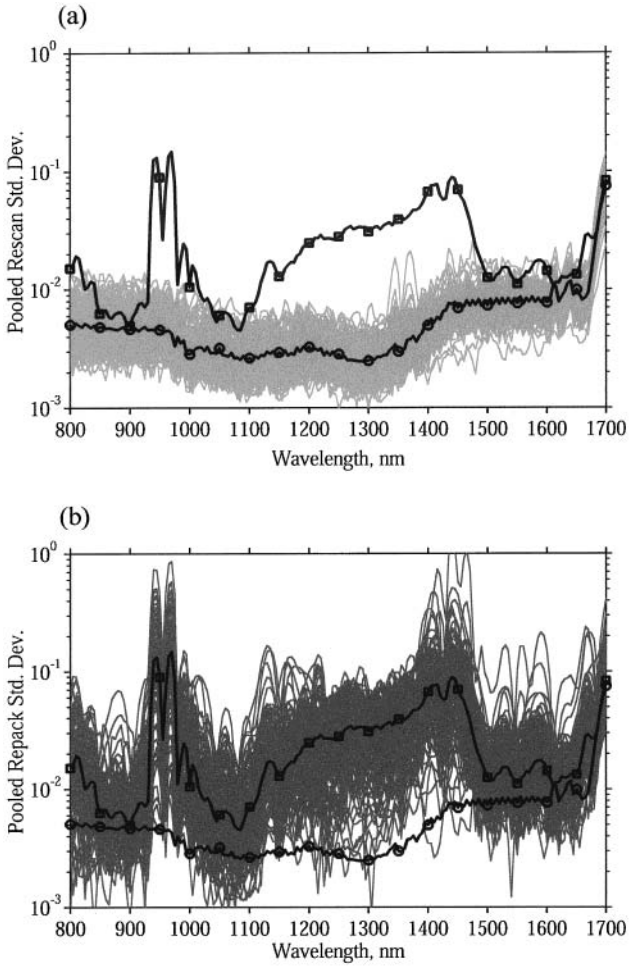

Figure 6. Pooled repeat standard deviation per specimen at each wavelength as calculated from the four rescans of the eight repacks obtained for each untreated cereal food product (n = 137, both calibration and test samples): (a) 137 pooled rescan standard deviation spectra create the shaded area, the line with the square symbols marks the mean repack standard deviation, and the line with the circle symbols marks the mean rescan standard deviation; (b) similar to part a, except that 137 pooled repack standard deviation spectra create the shaded region. Each scan was preprocessed with the second derivative and SNV applied over the optimal range.

sample condition, instrument condition, and elapsed time between measurements. However, the rescan noise varied with wavelength, having many of the features of the average cereal product spectrum. Rescan noise was the same magnitude as the repack noise above 1680 nm, which probably accounted for the optimal upper limit of the spectral range established by the spectral window experiments. The rescan noise was only marginally smaller than the repack noise in the range 1500−1680 nm, a region that contains at least two important spectral signals. These spectral features can be seen in the spectral plot and the plot of spectral activity within the model, as described later (Figure 3a,c). Rescan noise increased at wavelengths shorter than 950 nm, where the silicon array operates. Rescan noise was approximately the same as the repack noise from 845 to 930 nm, a range that contained one important spectral peak (see below). However, repack noise created by the tails of absorption by chromophores was the main cause of the preferred deletion of the lower region of the spectral range (800−845 nm). Increased repack noise also was observed in the deleted window

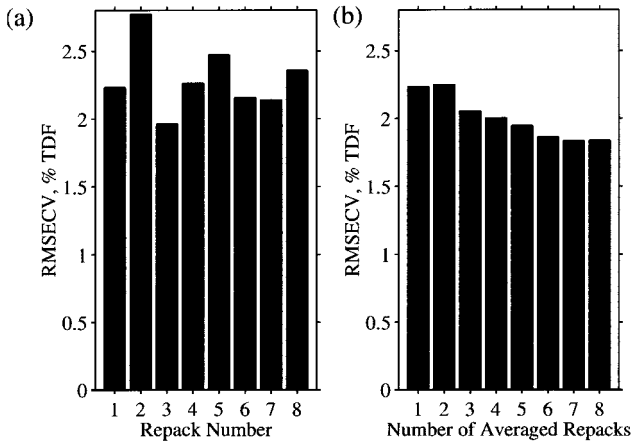

Figure 7. Variation in TDF prediction error due to the fluctuation in the spectral properties of the untreated subsample loaded into the sample hopper: (a) model error when individual repacks were used for calculating and testing the model; and (b) decrease in sampling error as increasing numbers of repack scans were averaged. The error was estimated from the 45-element test set by cross-validation using 92 untreated calibration samples and 13 PLSR factors.

from 1370 to 1480 nm, primarily as a result of variation in absorbance bands attributed to the free O−H stretch of crystalline sugars, which arose from the nonuniform distribution of added sugar in many cereal food products. There were some similar free O−H bands at 976 nm, but these bands were small enough that they did not degrade the calibration model appreciably (compare Figure 6 with Figure 3b). The repack noise increased well above the background at 942 and 967 nm. These wavelengths are near the splice point between the silicon and InGaAs spectral ranges. For cereal foods, the spectral region of the splice was a relatively featureless valley between two absorbance bands. Differences in reflected signal of each repack generated a mismatch in the slope of each spectrum meeting at the splice point. In the region between 1100 and 1365 nm, the repack noise was much larger than the rescan noise (Figure 6). This range contained many spectral features that were significant to the model (Figure 3c and see below).

When each repack scan was analyzed separately by the PLSR model, substantially different TDF predictions were produced, causing the RMSECV for the test set to vary widely (Figure 7a). This is thought to be a result of the heterogeneous nature of many of the products. The fluctuation was minimized by averaging multiple optical samples to represent a larger area of each specimen. When six or seven repacks were averaged, the model error was relatively constant (Figure 7b). An additional repack yielded no improvement. Moreover, no correlation was found between the pooled repack noise for a spectrum and the residual error in predicting TDF when using the eight-repack average spectra (data not shown). Since the diameter of the illumination area was approximately 10 cm, approximately 550 cm^2 of product surface must be measured with the NIR beam. In contrast to repacks, use of two or more rescans had no effect on the model performance (data not shown).

Spectral Basis of the Calibration Model Developed from Untreated Samples. Exact spectral interpretation of the NIRR model for TDF was difficult, as molecular vibration absorbance bands were highly overlapped. Also, the indigestible fraction of food products consists of several very different components, further confusing the identification of specific bands.

Table 3. Statistical Analysis of Test Set TDF Prediction Error and Bias for the 6 Product Classes

product class[a]	# of samples (within class)	SECV[b] (within class)	class bias[c]	# of samples (outside class)	SECV[b] (outside class)	bias (outside class)	T-test on bias (class vs outside class)[d]	F-test on SECV (class vs outside class)[e]	F-test on SECV (class vs class I)[f]
I	13	1.40	−0.08	32	1.99	0.05	58%	72%	
II	5	1.81	−0.05	40	1.84	0.02	53%	62%	70%
III	10	1.28	−0.06	35	1.97	0.03	55%	72%	53%
IV	3	2.31	0.04	42	1.80	0.01	51%	78%	81%
V	10	2.43	−0.07	35	1.62	0.03	56%	83%	82%
VI	4	2.00	0.73	41	1.80	−0.06	79%	70%	76%

[a] Described in Table 1. [b] Standard error of cross-validation (ASTM, 1995) calculated for the 13-factor PLSR D2-SNV model with optimized spectral range, and using only untreated calibration samples. [c] Bias is the difference between the average predicted values and the average measured TDF values. [d] Probability that the bias is different between the samples in the product class and samples outside the product class. [e] Probability that the SECV is different between the samples in the product class and samples outside the product class. [f] Probability that the SECV is different between the samples in the product class and samples in product class I.

However, the relative importance of different spectral ranges can be demonstrated by calculating the standard deviation across samples of the set of preprocessed calibration spectra whose variables have been weighted by the regression vector (STD(RVWS), Figure 3c). The STD(RVWS) spectrum indicates spectral bands that were weighted heavily in calculation of the predictions. The full-width at half-maximum of each of these bands fall in the range 20−35 nm, as would be expected for solid-state vibrational bands. However, the band positions are only approximate, because the D2-SNV preprocessing could cause signal variance to shift from the peaks of absorbance bands to nearby spectral features such as valleys. Correlation of individual bands to TDF was low (typically no greater than 0.5). One plausible explanation is that even if a band measures an individual component of the food product, that component may only indirectly be related to TDF. For example, low-lipid-content products did not always have a high fiber content, but high-fiber-content products often had low lipid content. Another possibility is that an active band could be sensitive to more than one component, and consequently, its value is only significant in the context of other bands. For example, one band might be sensitive to both digestible and indigestible polysaccharides, while another was sensitive to lipid and digestible polysaccharides, and a third to lipid alone.

On the basis of published correlation tables (Murray and Williams, 1987; Osborne et al., 1993), the molecular vibrations of various groups of bands were identified in the plot in Figure 3a. Two sharp bands were assigned as the first and second overtone of O−H stretching of free hydroxyl in a crystalline sugar (mainly sucrose). A sharp feature due to the first overtone of aromatic C−H of bran was present at 1685 nm, which was an important analytical band for TDF. Various C−H stretch second overtones and C−H combination bands were used by the calibration model, as were the C−H stretch third overtones. The series of C−H stretching bands suggests that the model was sensitive to a wide range of functional groups in different types of molecules (aromatic, methyl, methylene, or methine; lipids, proteins, polysaccharides, and lignins). First and second overtones of O−H stretch vibrations also were important for the calibration model and could arise from multiple components of the products (starch, cellulose, hemicellulose, simple sugars). The absorption of protein N−H stretch first and second overtones also could contribute to the model.

Effect of Product Class and Composition on Model Performance. Cereal food products differ in their morphology, structure, and composition, and any one of these characteristics could create bias or error in the NIRR prediction model. The effect of membership in the product classes of Table 1 on test set prediction error and bias was evaluated statistically (Table 3). The bias-corrected standard error of cross-validation (ASTM, 1995) (SECV) was used in the table because test-set RMSECV was found to be the same as SECV except for type VI products (predominantly bran), where it was slightly higher because of positive bias (2.13 vs 2.00). The T-test, on the bias between a particular class and all other data in the test set, indicated that type VI products had the only product-class-related bias problem (columns 4, 7, and 8 of Table 3), suggesting that type VI products were different from the others and were not well represented in the data set. Product class SECVs in column 3 of Table 3 were evaluated in two ways: (1) by comparison of the SECVs produced for the class against the remainder of the samples in the test set (columns 6 and 9 of Table 3) and (2) comparison between the class SECV and the SECV of class I (column 10 of Table 3). The lowest SECVs were for breakfast products of class I, and snack products (class III), and these SECVs were not found to be statistically different from each other. The breakfast products with fruit or nuts (class II) had a higher SECV, which was only marginally statistically significant even by direct comparison to class I. Classes IV−VI had higher SECVs that were significant by either statistical comparison. The most extreme was the whole grain class (V), perhaps because of relatively less accurate optical sampling caused by differences in the light penetration through the surface layers of these products.

Because the cereal food product samples were very diverse, there may have been spectral interference between the various constituents of the products. To examine the magnitude of this potential problem, the TDF prediction-residuals, or the absolute value of those prediction residuals, from the untreated-sample test set were plotted versus the six major sample attributes in Table 2 (data not shown). Though relatively minor, the largest observed effect was that elevated sugar content was correlated with higher bias, but without any correlation to the amplitude of the error. By contrast, as fat content increased, the bias did not change, but the error magnitude decreased significantly. As moisture content increased, within the range found in untreated samples, bias became slightly more negative, and error magnitude increased slightly. Among the samples of the data set, densities and moisture contents were slightly inversely correlated to sugar content, so the slight negative bias attributed to moisture may in fact have been a result of the sugar content.

Stabilization of Models to Moisture and Particle-Size Variation. The variation of coating thickness,

Table 4. Variation in Prediction Error Due to Particle Size or Moisture Treated Specimens in Either the Calibration or Test Sets

				RMSECV[a] difference between treated and untreated pairs test set treatment & MC[b] range			
	# calibr.	untreated products		desiccated	moistened	moistened	moistened
calibration set description	samples	RMSECV[a]	R^2	(MC < 5.25%)	(MC < 10.7%)	(MC < 16.0%)	(MC < 21.5%)
untreated products	137	1.83	0.96	−0.28	0.70	0.27	3.11
untreated and particle size treated	219	1.80	0.96	−0.38	0.76	0.26	3.10
untreated, particle size, and moisture treated (MC[b] < 10.7%)	276	2.01	0.96	−0.33	***−1.06***[c]	0.35	1.79
untreated, particle size, and moisture treated (MC < 16.0%)	341	1.80	0.97	−0.73	−0.56	***−0.10***	0.87
untreated, particle size, and moisture treated (MC < 21.5%)	361	1.87	0.96	−0.76	−0.56	−0.04	***0.33***
number of test samples		45		5	7	13	17

[a] Calculated for the 13-factor PLSR D2-SNV model with optimized spectral range. [b] MC is moisture content by oven method. [c] Bolded italicized numbers indicate the best errors achieved for each treated test set.

particle size, and moisture content (MC) of a product can create bias and error in the predictions of NIRR calibration models if the models do not sufficiently account for these effects. However, if the calibration samples have wide variation in these attributes, model error may needlessly be increased for future samples to be analyzed. To evaluate these effects, the data for treated samples were analyzed by observing the change in model error between pairs of treated and untreated specimens in the test set, as predicted by calibration sets containing different sets of treated specimens (Table 4). Note that only 17 of the 45 test samples were treated, so this subset of the test set did not represent very accurately all sample types found in the study. Therefore, the arguments about the effects of treatments are based on differences observed with the same (small) sets of samples. Although some of the calibration sets contained treatments, the calculations were implemented such that the calibration sets for each cross-validation segment did not contain any treated versions of the sample being predicted by that cross-validation segment.

By use of a calibration model developed from untreated products (row 2 of Table 4), the prediction error for a set of five samples improved after desiccation (−0.28 in column 5 of Table 4). In fact, the data suggested that the NIRR models produced better results with drier samples, no matter which set of treatments was included in the calibration. Likewise, moistening products at the highest levels studied (up to 21.5% MC) always created poorer predictions. For example, in column 8 of Table 4, RMSECV for 17 test samples was 3.11 larger using spectra collected after the samples received moisture treatments. The increase in RMSECV was primarily due to a large positive bias for the wet samples, in contrast to the aforementioned small negative bias associated with the inherent moisture content of untreated samples. The highest level of moisture content was extreme for many of these products, resulting in friable saturated products and cereal glazes that were sticky or syrup-like. In the more realistic lower ranges of test-sample moisture content, the error increases due to moisture treatment were not nearly as large (maximum of 0.70). Furthermore, including moisture treatments in the calibration dramatically improved the error for the moist samples but did not greatly alter the RMSECV values for the untreated-sample test set, which fell in the range 1.8−2.0% TDF for the 13-factor models built from various calibration sets. In contrast, including only particle size treatments in the calibration did not yield substantial improvement in the ability to predict wet samples. In the three cases where the maximum moisture content of the calibration set could be matched to that of the test set, the best error was obtained for that test set. This effect was largest for the drier test set (MC < 10.7%). For the test set with MC < 16.0%, there was only a small level of degraded performance of the models as a result of adding calibration samples with moisture contents exceeding 16%. Surprisingly, in many cases the moistened products (except for the highest moisture level) were predicted better than their untreated counterparts. It is speculated that regression performance is improved because moderate levels of hydration convert various types of crystalline sugars to their amorphous forms, which are more spectrally similar to one another. Reeves has previously demonstrated this effect in model studies (Reeves, 1995).

Reducing the particle size of a subset of test samples ($n = 20$) by crushing treatments improved the RMSECV by 0.09, and furthermore, a slightly better result (0.13 improvement) was obtained when particle size treatments were included in the calibration set (data not shown). However, in regard to predicting the test set of intact cereal food products, the benefits of including particle size treatments in the calibration were marginal, though positive (Table 4). Moisture content variation among untreated products appeared to have a more profound effect on model performance than the effects of particle size or structure. When the particle size and inherent moisture content of cereal food product specimens cannot be adjusted prior to making an NIRR determination, it is recommended that the calibration set include all levels of crushing treatments and moisture treatments up to 16.0% MC. With this calibration set, the RMSECV was 1.80% for the 45-sample untreated test set, and slightly lower residual errors were obtained for 13 test specimens moistened to a moderate level (<16.0% MC).

Comparison of Spectral Methods for Fiber Determination. In this study, and those described in the literature, the accuracy errors for spectral determination of dietary fiber are greater than twice the reference method repeatability (Baker, 1983; Baker, 1985; Williams et al., 1991; Kays et al., 1996; Windham et al., 1997; Kays et al., 1997; Archibald et al., 1998a; Archibald et al., 1998b; Kays et al., 1998; Kays and Barton, II, 1998). The AOAC method for NIRR analysis of fiber in animal forage performed better relative to the reference methods, but this is because the reference methods were reported to be less repeatable. The AOAC validation study for the NIRR acid-detergent-fiber method

demonstrated that the ADF values of multiple cultivars of bermudagrass could be determined, using multiple linear regression, with a median standard error of NIRR analysis of 1.5%, and R^2 of 0.89 for the ADF range 22.7−45.5% (Barton, II, and Windham, 1988). Neutral detergent fiber (NDF) has been determined to a similar degree of accuracy by NIRR of milled forage samples, and both reference methods have a repeatability error of about 1.0% (Barton, II, 1989). Cereal food products are much more varied than the aforementioned forage products. However, Kays, Windham, and Barton's studies on NIRR determination of TDF of milled cereal food products, using the full 1100−2500-nm spectral range and PLSR (Kays et al., 1996; Kays et al., 1997; Windham et al., 1997; Kays et al., 1998), show that TDF prediction errors can be reduced to nearly the same level of accuracy as that for fiber in forages. Initial publications developed calibrations for various cereal food product subtypes (based on levels of constituents such as sugar and moisture), where the errors in predicting reference data ranged from 1.40 to 1.88% TDF, with R^2 from 0.98 to 0.99 (Kays et al., 1996; Kays et al., 1997; Windham et al., 1997). A recently published study, utilizing cereal food products of diversity similar to that reported here, produced calibration set RMSECV of 1.73% ($n = 117$) and independent validation set RMSEP of 1.34% ($n = 54$) (Kays et al., 1998). That model produced an RMSEP of 1.63% for an independent test set that included samples treated to expand the moisture content to cover the range from 1.19 to 16.21% (personal communication, Kays, 1999). A study on the combined use of NIRR and Raman spectroscopy to determine the TDF of diverse untreated milled specimens reported RMSEP of 1.5% for a 63-element test set predicted by a 63-element calibration set (Archibald et al., 1998a; Archibald et al., 1998b). The simultaneous use of two spectral regions, however, would add substantial expense to a spectroscopic system for prediction of TDF.

Unlike previous reports, the study presented here developed calibration models to predict TDF from NIRR of intact cereal food products using the 800−1700-nm spectral region. The model errors appear to be 0.17−0.37% higher than the 1.63% that has been achieved by NIRR analysis of milled specimens using the 1100−2500-nm range. Despite a narrower spectral range, the presented models require a greater number of PLSR factors. It may be that additional factors are needed to model the more complex reflectance properties of intact products, or alternatively that the narrower spectral range and greater bond vibration anharmonicity accentuate the problem of spectral interference between constituents. Some sample types were underrepresented in the data set used in the current study, and consequently, model error will probably improve after addition of appropriate samples. The current results have demonstrated a useful level of performance considering the 49.4% range of TDF in diverse cereal food products. Moreover, the method has some unique characteristics: no sample preparation is required; less than a minute is needed to repack a sample enough times to obtain a representative optical sample; and there is no damage to the physical integrity of the cereal food products.

ABBREVIATIONS USED

TDF, total dietary fiber; NIRR, near-infrared reflectance; NDF, neutral-detergent fiber, ADF, acid-detergent fiber; NIR, near-infrared; SEL, standard error of laboratory values; PLSR, partial least-squares regression; RMSECV, root-mean-squared error of cross validation; SNV, standard normal variate spectral transformation; D2, digital second derivative spectral transformation; RMSEP, root-mean-squared error of prediction; RMSEC, root-mean-squared error of calibration; STD(RVWS), standard deviation across samples of the regression-vector-weighted spectra; SECV, standard error of cross-validation; MC, moisture content.

ACKNOWLEDGMENT

W. C. Munday and A. Khettry (now at Bayer) of Perten Instruments are thanked for arranging the loan of the spectrometer used in this trial, as well as for providing technical information regarding its use and characteristics. The late R. G. Leffler designed and constructed the sample hopper. M. E. Mitchell and A. C. Loper assisted with processing of the specimens. F. E. Barton, II, is thanked for helpful discussions regarding the application of NIR reflectance spectroscopy.

LITERATURE CITED

AOAC. *Total Dietary Fiber in Foods: Enzymatic−Gravimetric Method*; AOAC: Arlington, VA, 1990.

AOAC. *991.43 Total, Soluble, and Insoluble Dietary Fiber in Foods;* AOAC: Arlington, VA, 1992.

Archibald, D. D.; Akin, D. E. Use of spectral window preprocessing for selecting near-infrared reflectance wavelength sets for determination of the degree of enzymatic retting of intact flax stems. *Vib. Spectrosc.* **2000**, *23*, 169−180.

Archibald, D. D.; Kays, S. E.; Himmelsbach, D. S.; Barton, F. E., II. Raman and NIR spectroscopic methods for determination of total dietary fiber in cereal foods: A comparative study. *Appl. Spectrosc.* **1998a**, *52*, 22−31.

Archibald, D. D.; Kays, S. E.; Himmelsbach, D. S.; Barton, F. E., II. Raman and NIR spectroscopic methods for determination of total dietary fiber in cereal foods: Utilizing model differences. *Appl. Spectrosc.* **1998b**, *52*, 32−41.

ASTM. E 1655−94 Standard Practices for Infrared, Multivariate, Quantitative Analysis. In *Annual Book of ASTM Standards*; American Society for Testing and Materials: Conshohocken, PA, 1995.

Baker, D. The determination of fiber in processed cereal foods by near-infrared reflectance spectroscopy. *Cereal Chem.* **1983**, *60*, 217−219.

Baker, D. The determination of fiber, starch, and total carbohydrate in snack foods by near-infrared reflectance spectroscopy. *Cereal Foods World* **1985**, *30*, 389−392.

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.

Barton, F. E., II. Considerations of chemical analyses. In *Near Infrared Reflectance Spectroscopy (NIRS): Analysis of Forage Quality. U.S. Department of Agriculture, Agricultural Handbook No. 643 (Revised With Supplements)*; Marten, G. C., Shenk, J. S., Barton, F. E., II, Eds.; National Technical Information Service: Springfield, VA, 1989.

Barton, F. E., II; Himmelsbach, D. S.; Windham, W. R.; Lyon, B. G.; Kays, S. E. Quality assessment of agricultural and food products by spectroscopic methods. In *Proceedings of the United States/Japan Cooperative Program in Natural Resources (UJNR) Protein Resources Panel: 24th Annual Meeting: Tsukuba: October 16−20, 1995*; National Food Research Institute of Japan: Tsukuba Science City, Japan, 1995.

Barton, F. E., II; Windham, W. R. Determination of acid-detergent fiber and crude protein in forages by near-infrared reflectance spectroscopy: collaborative study. *J. Assoc. Off. Anal. Chem.* **1988**, *71*, 1162−1167.

de Jong, S. SIMPLS: an alternative approach to partial least-squares regression. *Chem. Intell. Lab. Syst.* **1993**, *18*, 251−263.

Gorman, M. A.; Bowman, C. Position of the American dietetic association − health implications of dietary fiber. *J. Am. Dietetic Assoc.* **1993**, *93*, 1446−1447.

Huthfehre, T.; Feldhoff, R.; Kantimm, T.; Quick, L.; Winter, F.; Cammann, K.; Vandenbroek, W.; Wienke, D.; Melssen, W.; Buydens, L. NIR − remote-sensing and artificial neural networks for rapid identification of post consumer plastics. *J. Molecular Struct.* **1995**, *348*, 143−146.

Kays, S. E., Personal communication, Dec. 5, 1999.

Kays, S. E.; Barton, F. E.; Windham, W. R.; Himmelsbach, D. S. Prediction of total dietary fiber by near-infrared reflectance spectroscopy in cereal products containing high sugar and crystalline sugar. *J. Agric. Food Chem.* **1997**, *45*, 3944−3951.

Kays, S. E.; Barton, F. E., II. The use of near-infrared reflectance spectroscopy to predict the insoluble dietary fiber fraction of cereal products. *J. NIR Spectrosc.* **1998**, *6*, 221−227.

Kays, S. E.; Barton, F. E., II; Windham, W. R. NIR analysis of dietary fiber. In *Complex Carbohydrates in Foods*; Cho, S. S.; Prosky, L.; Dreher, M., Eds.; Marcell Decker: New York, 1999.

Kays, S. E.; Barton, F. E., II; Windham, W. R. Analysis of nutritional components in cereal foods with NIR reflectance spectroscopy. In *Proceedings of NIR-99: Verona, Italy: 13−18 June, 1999*; Michael, I.; Davies, A. M. C., Eds.; NIR Publications: Charleton, Chichester, U. K., 2000.

Kays, S. E.; Windham, W. R.; Barton, F. E. Prediction of total dietary fiber in cereal products using near-infrared reflectance spectroscopy. *J. Agric. Food Chem.* **1996**, *44*, 2266−2271.

Kays, S. E.; Windham, W. R.; Barton, F. E. Prediction of total dietary fiber by near-infrared reflectance spectroscopy in high-fat- and high-sugar-containing cereal products. *J. Agric. Food Chem.* **1998**, *46*, 854−861.

Lee, S. C.; Prosky, L.; De Vries, J. W. Determination of total, soluble, and insoluble dietary fiber in foods − enzymatic−gravimetric method, MES-TRIS buffer: collaborative study. *J. Assoc. Off. Anal. Chem. Int.* **1992**, *75*, 395−416.

Madden, H. H. Comments on Savitzky−Golay Convolution Method for Least-Squares Fit Smoothing and Differentiation of Digital Data. *Anal. Chem.* **1978**, *50*, 1383−1386.

Marlett, J. A.; Slavin, J. L. Position of the American dietetic association − health implications of dietary fiber. *J. Am. Dietetic Assoc.* **1997**, *97*, 1157−1159.

Marten, G. C.; Templeton, W. C., Jr. History of the national NIRS forage research project network. In *Near Infrared Reflectance Spectroscopy (NIRS): Analysis of Forage Quality. U.S. Department of Agriculture, Agricultural Handbook No. 643 (Revised With Supplements)*; Marten, G. C., Shenk, J. S., Barton, F. E., II, Eds.; National Technical Information Service: Springfield, VA, 1989.

Martens, H.; Næs, T. *Multivariate Calibration*; John Wiley & Sons: New York, 1989.

Murray, I.; Williams, P. C. Chemical principals of near-infrared technology. In *Near-Infrared Technology in the Agricultural and Food Industries*; Williams, P., Norris, K., Eds.; American Association of Cereal Chemists, Inc.: St. Paul, Minnesota, 1987.

Norris, K. H.; Barnes, R. F.; Moore, J. E.; Shenk, J. S. Predicting forage quality by infrared reflectance spectroscopy. *J. Anim. Sci.* **1976**, *43*, 889−897.

Osborne, B. G.; Fearn, T.; Hindle, P. H. Theory of near-infrared spectrophotometry. In *Practical NIR Spectroscopy With Applications in Food and Beverage Analysis*; Longman Scientific & Technical: London, 1993.

Reeves, J. B. Efforts to quantify changes in near-infrared spectra caused by the influence of water, pH, ionic-strength, and differences in physical state. *Appl. Spectrosc.* **1995**, *49*, 181−187.

Savitzky, A.; Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, *36*, 1627−1639.

United States Congress. *Code of Federal Regulations*, 21, Food and Drugs, Part 101.9. Office of the Federal Register, National Archives and Records Service, General Services Administration: Washington, D. C., 1995.

USDA. *Near Infrared Reflectance Spectroscopy (NIRS): Analysis of Forage Quality. U.S. Department of Agriculture, Agricultural Handbook No. 643*, National Technical Information Service: Springfield, VA, 1985.

USDA. *Near Infrared Reflectance Spectroscopy (NIRS): Analysis of Forage Quality. U.S. Department of Agriculture, Agricultural Handbook No. 643 (Revised With Supplements)*, National Technical Information Service: Springfield, VA, 1989.

Williams, P. C.; Cordeiro, H. M.; Harnden, M. F. T. Analysis of oat bran products by near-infrared reflectance spectroscopy. *Cereal Foods World* **1991**, *36*, 571−574.

Windham, W. R.; Kays, S. E.; Barton, F. E. Effect of cereal product residual moisture content on total dietary fiber determined by near-infrared reflectance spectroscopy. *J. Agric. Food Chem.* **1997**, *45*, 140−144.

Received for review February 16, 2000. Revised manuscript received July 10, 2000. Accepted July 10, 2000. Mention of products or trade names is for descriptive purposes only and does not imply the endorsement of the U.S. Department of Agriculture.

JF000206J