Effect of Cereal Product Residual Moisture Content on Total Dietary Fiber Determined by Near-Infrared Reflectance Spectroscopy

W. R. Windham,* Sandra E. Kays, and Franklin E. Barton, II

Agricultural Research Service, Richard. B. Russell Agricultural Research Center, United States Department of Agriculture, P.O. Box 5677, Athens, Georgia 30613

Published studies on the applications of near-infrared reflectance spectroscopy (NIRS) to the analysis of fiber in forages, feeds, grains, and cereal products indicate the presence of O-H absorbance, due to sample moisture content, in the calibration models. The objective of this study was to determine the extent to which residual moisture in samples interferes with the ability of NIRS to predict total dietary fiber (TDF) in cereal products and grains. Milled cereal products and grains were stored in 20%, 60%, and 80% experimental relative humidity (rh) environments and a vacuum oven. Samples (N = 143) were analyzed for moisture and predicted for TDF. Results showed significant differences between laboratory reference TDF and predicted TDF for samples that were either very low or very high in moisture. Cereal products and grain samples stored under ambient conditions (N = 90)were combined with selected samples stored under different rh environments (N = 53) to develop a new calibration using partial least squares regression. The standard error of cross validation and multiple coefficient of determination (R^2) were 1.85% and 0.98, respectively. The model was validated with an independent set of cereal products (N = 29) stored under ambient and rh environments. Samples stored under ambient and rh environments were predicted with standard errors of performance of 1.70% and 1.86%, respectively. The study shows that NIRS can be used to predict TDF in cereal products and grains with a wide range of residual moistures when calibrations include the range of residual moisture expected.

Keywords: Dietary fiber; near-infrared spectroscopy; moisture

INTRODUCTION

Near-infrared reflectance spectroscopy (NIRS) has been an invaluable tool for the analysis of fiber in forages and feeds for many years (Murray, 1993). More recently, NIRS has been reported to be an accurate method for the determination of total dietary fiber (TDF) in cereal products and grains (Kays et al., 1996) and in oat bran products (Williams et al., 1991). In the majority of these applications, O-H absorption due to sample residual moisture content is present in the calibration models. Baker (1985) reported the presence of a moisture band (1936 nm) in a stepwise multiple linear regression model to predict fiber in snack foods. The influence of moisture was also apparent in the development of a partial least squares (PLS) calibration for soluble dietary fiber, neutral detergent fiber, and TDF in oat bran products (Williams et al., 1991). Kays et al. (1996) reported the presence of O-H absorbance in PLS loadings for the three most important factors in the TDF regression model. It was not known whether the influence of O-H absorption was due to residual moisture or carbohydrate.

Moisture in forages and cereal grains is present as a "free" fraction and as one or more "bound" fractions (Walter and Hope, 1971). Free moisture comprises most of the moisture in a sample. Bound moisture is held by a combination of van der Waals forces and H bonding and is not easily removed or added (Windham et al., 1987). Milled cereal products may equilibrate to different residual moisture contents, by a change in the free moisture fraction, depending on handling, storage,

* Corresponding author. Phone: 706-546-3513. FAX: 706-546-3607.

and laboratory temperature and humidity (Windham et al., 1993). Moisture is detected in the first overtone and combination wavelength absorption bands for O-H groups at 1450 and 1930 nm (Hoffmann, 1963). Due to the strong absorbance of moisture in the near-infrared (NIR) spectrum, sample residual moisture content may have a significant effect on prediction with NIRS. Fales and Cummins (1982) reported significant differences between laboratory and predicted acid detergent fiber for forage samples that were stored under medium or high relative humidities. The purpose of the present study was to determine the extent to which residual moisture present in samples stored under very dry and humid conditions interferes with the ability of NIRS to predict TDF in cereal products and grains.

MATERIALS AND METHODS

Cereal Samples and Sample Preparation, Reference Laboratory Method for Total Dietary Fiber, and NIRS Instrumentation. Cereal and grain products (n = 119), including breakfast cereals, crackers, brans, flours, and commercial oat and wheat fibers (Kays et al., 1996), were assayed chemically and spectroscopically. Samples were dry milled to <500 μ m in a cyclone mill (Cylotec 1093 Sample Mill, Perstorp Analytical, Silver Spring, MD). Samples were assayed for total dietary fiber by AOAC method 991.43 (AOAC, 1990b, modified by Kays et al., 1996). Dry milled cereal samples were scanned with the NIRSystems 6500 monochromator (MIRSystems Silver Spring, MD) as described by Kays et al. (1996) to obtain reflectance spectra.

Relative Humidity Storage Conditions. Storage conditions and procedures were as described by Windham et al. (1993). The relative humidity (rh) in each storage cabinet was controlled by saturated salt solutions. Potassium acetate, sodium bromide, and ammonium sulfate solutions established

Cereal and Grain Products



Figure 1. Flowchart of the procedure for storing cereals and grain products in experimental relative humidity (rh) environments, vacuum oven (VO), and ambient (AMB) laboratory conditions. N = number of samples, * = groups with the same samples.

the rh at 20%, 60%, and 80%, respectively (Greenspan, 1977). Selected samples were stored in a vacuum oven at ambient temperature with a 4 kPa laboratory vacuum as described by Windham et al. (1987).

Cereal Samples. Calibration samples (N = 90 minus 4 samples not available) of cereal and grain products, described by Kays et al. (1996), were stored in experimental relative humidity environments and vacuum oven (Figure 1). Twenty samples were randomly selected and each sample stored in each rh environment and the vacuum oven (N = 80). The remaining 66 samples, were randomly assigned to four groups. One group was stored under ambient conditions in the laboratory (N = 15). The three remaining groups were stored in the three experimental controlled rh environments (N = 17 in each rh environment).

NIRS Validation Samples. Validation samples (N = 29) of cereal and grain products, described by Kays et al. (1996), were divided into four groups such that each group had a similar TDF mean and standard deviation. The groups were randomly assigned to the experimental rh environments or the vacuum oven (Figure 1) and stored as previously described.

NIRS and Moisture Measurements. After storage and equilibration, cereal samples and NIRS validation samples were removed and scanned with a NIRSystems 6500 monochromator (NIRSytems, Silver Spring, MD) as described by Kays et al. (1996). Samples were analyzed for moisture by drying in a forced air oven (AOAC, 1990a).

Multivariate Calibration. A commercial spectral analysis program (NIRS3, Infrasoft International, Inc., Port Matilda, PA) was used to process the data and develop chemometric models. Using an algorithm called SELECT (Shenk and Westerhaus, 1991a), selected samples (N = 143 out of 146, three samples were lost) from the rh environments and vacuum oven were used to recalibrate the original TDF NIR model to account for any effects of residual water. The SELECT algorithm identifies samples within and outside the neighborhoods which had previously been defined by the original calibration data set of Kays et al. (1996). Twelve principal component factors were used by SELECT and were based on counting the number of eigenvalues greater than the average eigenvalue divided by the square root of the number of samples (Shenk and Westerhaus, 1991b). Using the scores in 12-dimensional space, the neighborhood H Mahalanobis distance was calculated among all spectral pairs in the original TDF calibration data set and the rh and vacuum oven stored samples. Any stored sample whose neighborhood H value was less than 0.6 H (Shenk and Westerhaus, 1991b) from any sample in the original calibration data set was eliminated. This process was repeated with the remaining samples in the pool, and so on, until every sample was in either the calibration update (N = 53) or the eliminated set (N = 90). The samples in the eliminated set were subsequently used to validate the recalibrated TDF NIR model.

Calibration samples (N = 90) of cereal products, described by Kays et al. (1996), were combined with the selected calibration update samples (N = 53) for development of a new TDF calibration model. The multivariate data technique partial least squares regression (PLS1 and PLS2) was used for calibrations (Martens and Naes, 1989). When only one

Table 1. Mean and Standard Deviation (SD) of MoistureContent (%) in Cereal Products and Grains Stored inDifferent Relative Humidity (rh) Environments

	calibr	ation	validation		
treatment	mean	SD	mean	SD	
ambient	9.05	2.04			
vacuum oven	4.61	1.22	2.29	0.35	
20% rh	5.44	2.50	6.62	0.50	
60% rh	11.25	1.00	11.02	1.13	
80% rh	15.26	1.92	15.10	1.13	

dependent variable is modeled (e.g., TDF alone), the PLS algorithm is noniterative and is termed PLS1. When several variables are modeled simultaneously (e.g., TDF and moisture), the algorithm is iterative and called PLS2 (Martens and Martens, 1986). Prior to the PLS procedure, $log_{10}(1/R)$ spectra were mean centered, transformed with standard normal variate and detrending procedures (Barnes et al., 1989), to remove multiplicative interferences of scatter, and then transformed with a second-derivative processing (gap = 8 nm), smoothing interval = 4 nm) to enhance absorption peaks. The transformations improved the standard error of calibration compared to PLS analysis with untransformed data. A predictive cross-validation method was used to determine the optimum number of PLS factors and to guard against overfitting (Martens and Naes, 1989). One-fifth of the samples were removed from calibration and used for prediction. This was repeated five times, each time with a different fifth of the samples removed. Performance statistics were accumulated for each group of samples removed. The optimal number of factors for the model was that which produced the first minimum in error between modeled and reference values (standard error of cross validation) for the samples removed during cross validation.

The recalibrated TDF equation was first validated using the samples not chosen by the SELECT algorithm (N= 90), second using the independent set of cereal samples (N= 29) described by Kays et al. (1996), and third using the independent set of cereal samples after storage in the experimental rh environments and vacuum oven (N = 29). Model performance was reported as the coefficient of determination (r^2), the standard error of performance (SEP), and the average difference between reference method and modeled values (bias) (formulas in Hruschka, 1987).

RESULTS AND DISCUSSION

Moisture contents of cereal products and grains stored in different rh environments are shown in Table 1. Across storage environments, the values for moisture in the cereal samples ranged from 1.29% to 20.57%. This is in agreement with Windham et al. (1993) for moisture content of wheat across similar storage environments. The calibration and validation samples with no special humidity storage treatment had average moisture contents of 9.05% and 8.20%, respectively. The moisture content of calibration samples placed in the vacuum oven and in the 20% rh environment decreased by 4.44%



Figure 2. Total dietary fiber residuals (AOAC-determined TDF – NIRS-predicted TDF) versus descending moisture content for cereal products (N = 143) stored in different relative humidity environments. TDF was predicted using the equation developed by Kays et al. (1996) with samples stored at ambient conditions.

and 3.61% and validation samples decreased by 5.91% and 1.58%, respectively. Conversely, in the 60% and 80% rh environments, moisture content of the calibration samples increased 2.20% and 6.21% and validation samples increased by 2.82% and 6.90%, respectively.

When the original TDF NIR equation described by Kays et al. (1996) was used to predict the samples after storage in the five different rh environments (N= 143), the SEP and bias were 2.63% and -1.65% age units, respectively. These performance statistics are higher than reported by both Kays et al. (1996) and Williams et al. (1991) for NIRS prediction of TDF in cereals, grains, and oat bran products stored under ambient conditions. In addition, 27% of the samples stored under the rh regimes were spectrally extreme samples (Mahalanobis distance > 3.0), indicating that overall they were not represented by the calibration set derived from samples stored under ambient conditions only.

Figure 2 shows the relationship between TDF residuals (AOAC-determined TDF – NIR-predicted TDF) versus the descending moisture content of the samples from the five different storage conditions (N = 143). Sorting the samples by moisture content revealed the errors due to varying residual moisture content. The curved distribution of the residuals and negative values indicated that the samples stored in the 80% and 20% rh and vacuum oven environments were overpredicted. Samples with moisture concentrations of >12% had an SEP and a bias of 3.43% and -2.63, respectively. Samples with moisture concentrations of <6.0% had an SEP and bias of 2.66% and -1.87, respectively,

The SELECT algorithm selected $5\overline{3}$ samples to be added to the calibration data set (N = 90) described by Kays et al. (1996). Moisture and TDF content were modeled simultaneously (PLS2) to explore the relationship between these variables and the NIRS measurement. The PLS2 algorithm produces a set of factors which relate to all the variables used in their derivation. The resulting factors describe the variation in the NIR data that are relevant for modeling the variations in moisture and TDF. Cross-validated PLS2 revealed that 96% of the total variation in moisture and TDF could be predicted by NIR using the first four factors. The main variation (75%) was described by the first two factors.



Figure 3. Two vector PLS2 score plots of the regression between NIRS spectra and TDF and moisture of (A) high fiber, >20%; (B) intermediate fiber, 9%–20%; and (C) low fiber, <9%. Labels for moisture content are denoted as follows: •, high moisture, >12%; \bigtriangledown , intermediate moisture, 5–12%; •, low moisture, <5%.

The PLS2 algorithm also creates scores, which represent the position of samples relative to the factors and are the basis for regression models that predict composition. Scores are derived by taking, for each factor, the sum across the spectrum of weights times centered reflectance values. Sample scores from the PLS2 model were divided into three groups corresponding to high (>20%), intermediate (9%-20%), and low (<9%) TDF content. The three groups were further subdivided into high (>12%), intermediate (5-12%), and low (<5%) moisture content. The scores for these groups of samples are shown in Figure 3. Scores from the first PLS2 factor separated the samples into high, intermediate, and low moisture categories regardless of fiber level (Figure 3, panels A, B, and C). High- and intermediate-moisture samples had negative scores in intermediate- and lowfiber groups, whereas low-moisture samples had positive scores (Figure 3, panels B and C). Intermediatemoisture samples also had positive scores for high-fiber samples (Figure 3, panel A). Factor 2 scores separated the samples into high-, intermediate-, and low-TDF levels. High-fiber samples had higher positive scores than intermediate-fiber samples, whereas low-fiber samples generally had negative scores. Scores from the first two PLS2 factors were correlated (by Pearsons correlation coefficients) to both constituents (e.g., TDF and moisture). The first factor had a positive correlation of 0.72 with TDF but a negative correlation of -0.50



Figure 4. PLS1 regression coefficients (loadings) for factors 1 (A) and 2 (B) as a function of wavelength for TDF.

with moisture. The negative correlation indicates that factor 1 scores decreased as moisture content increased (Figure 3). Scores from the second factor had a positive correlation with oven moisture (r = 0.79) and TDF (r = 0.48). These results suggest that prediction of either constituent was based on information relating to both constituents and that residual moisture can be modeled by NIRS and hence eliminated as an interfering substance.

A separate PLS1 NIRS calibration equation was obtained for TDF in cereal products varying in residual moisture content. The standard error of cross validation (SECV), using five cross-validation groups, was 1.85%, with a multiple coefficient of determination (R^2) of 0.98. The SECV is slightly larger than that reported by Kays et al. (1996) (SECV of 1.58%) for a TDF calibration derived from cereal products stored under ambient conditions with a mean moisture content and standard deviation of 7.80% and 2.50%, respectively. The equation contained seven factors, with scores from factors 1, 2, and 3 having correlations of 0.77, 0.43, and 0.39, respectively, with TDF. The seven factors explained 98.4% of the spectral variation and 98.3% of the total variation in TDF data.

The PLS1 regression coefficients (loadings) for factors 1 and 2 as a function of wavelength for TDF are shown in Figure 4, panels A and B, respectively. Plots of



Figure 5. Total dietary fiber residuals (AOAC-determined TDF – NIRS-predicted TDF) versus descending moisture content for the validation set (N= 90) of cereal products stored in different relative humidity environments. TDF was predicted using the PLS1 equation developed with samples from the original equation (Kays et al., 1996) plus samples stored under different relative humidity environments.

loadings often resemble the spectra of samples and the spectra of constituents and thus offer scope for interpretation of maximum weighting matching known absorbance bands. The shape of the plot for the first factor (Figure 4A) showed effects relating to moisture at 1416 and 1932 nm (Cowe and McNicol, 1985); oil at 1722, 1764, 2304, and 2346 nm (Williams and Norris, 1987); and carbohydrate at 2286 and 2322 nm (Kays et al., 1996). The first factor had a correlation of 0.77 with TDF and a smaller negative correlation with oven moisture (r = -0.36). The low correlation with moisture is surprising, as the effect of the absorbance of water can be detected at 1416 and 1932 nm (Figure 4A). The low and negative correlation with oven moisture would indicate that the moisture detected in this factor must be tightly bound to other constituents and is not released during oven drying (Cowe and McNicol, 1985).

The second factor had a correlation of only 0.43 with TDF, but a high positive correlation of 0.82 with moisture. The presence of O-H absorbance due to moisture (Figure 4B, 1410 and 1920 nm) and the high positive correlation to oven moisture content suggests that the second factor is dependent upon the variation in residual moisture concentration. Loading 2 also had regression coefficients related to C-H absorption in carbohydrate at 2268 and 2346 nm and C=O absorption in protein at 2052 nm (Williams and Norris, 1987). Scores from factor 3 had a correlation of 0.39 with TDF and -0.18 with oven moisture and a loading plot similar to that described by Kays et al. (1996).

Equation Validation. The PLS1 TDF calibration was first validated on the samples not chosen by the SELECT algorithm (N = 90). Samples were predicted with an SEP of 1.47% and a bias of 0.14% age units. Linear regression of AOAC-determined TDF against NIR-predicted TDF (Y = -0.15 + 1.02X) gave an intercept and slope not significantly different from 0.0 and 1.0, respectively (p > 0.05). The relationship between TDF residuals versus the descending moisture content of the validation samples is shown in Figure 5. The residuals are evenly distributed and do not exhibit a curved distribution. Samples stored in very dry (vacuum oven and 20% rh) and very humid (80% rh) conditions had SEPs of 1.37% and 1.71%, respectively. The

Table 2. Validation Statistics for NIRS Prediction of TDF in Cereal Products Stored in Different Relative Humidity (rh) Environments: Mean, Standard Deviation (SD), Bias, Standard Error of Performance (SEP), and Coefficient of Determination (r^2)

	mean		SD				
model	AOAC	NIRS	AOAC	NIRS	bias	SEP	I^2
original ^a recalibrated ^b	15.54 15.54	19.32 15.83	12.54 12.54	10.96 11.36	$-3.78 \\ -0.29$	4.75 1.86	0.96 0.99

^{*a*} Original equation developed from samples at ambient rh. ^{*b*} Recalibrated equation developed from the original calibration samples plus samples stored in different rh environments.

original equation reported by Kays et al. (1996) resulted in an overall less accurate prediction of rh treated and vacuum oven samples with an SEP and bias of 2.31% and -1.45, respectively.

The TDF models (e.g., original and recalibrated) were validated using the independent validation samples described by Kays et al. (1996) and the independent validation samples after storage in controlled relative humidity environments and the vacuum oven. Prediction of the independent validation samples described by Kays et al. (1996) with the recalibrated model produced a SEP and bias of 1.70% and -0.45%, respectively. The SEP and bias were similar to those reported by Kays et al. (1996) (SEP = 1.51% and bias = -0.38). These data indicate that the calibration population can be expanded to include samples which vary in residual water concentration with only a slight loss in prediction accuracy. The original TDF model resulted in a high SEP and bias of 4.75% and -3.78, respectively (Table 2), when predicting the rh treated and vacuum oven samples. The negative bias was a result of higher predicted TDF values from samples stored in very dry (vacuum oven and 20% rh) and very humid (80% rh) conditions. Samples with moisture concentrations of >12% had a mean bias of -3.93% age units vs a -6.39% age unit bias for samples with moisture concentrations of < 5.0%. Upon prediction with the recalibrated model, a small negative bias (-0.29) and an SEP of 1.86% were found (Table 2.).

CONCLUSIONS

A TDF prediction model, developed using nearinfrared reflectance spectroscopy of cereal products varying in residual moisture, can be used to measure TDF in samples with a wide range of residual moisture. Though slightly less accurate than an NIRS TDF model based on a narrower range of residual moisture, the model based on a wide range of residual moisture was sufficiently accurate to provide an acceptable method for fiber analysis. Additionally, PLS1 factors produced O-H absorbance that was dependent on the variation in residual moisture concentration. Factors for TDF from PLS1 produced scores which were correlated to both TDF and moisture, indicating that prediction of TDF was based on information related to both constituents.

LITERATURE CITED

- AOAC Official Methods of Analysis, 15th ed.; Moisture in Cereal Adjuncts: Air Oven Method (103-104); AOAC: Arlington, VA, 1990a.
- AOAC Official Methods of Analysis, 15th ed.; 3rd Supplement; Total, Soluble, and Insoluble Dietary Fiber in Foods; AOAC: Arlington, VA, 1990b.

- 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 and de-trending of of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.* **1989**, *43*, 772–777.
- Cowe, I. A.; McNicol, J. W. The use of principal components in the analysis of near-infrared spectra. *Appl. Spectrosc.* **1985**, *39*, 257–266.
- Fales, S. L.; Cummins, D. G. Reducing moisture-induced error associated with measuring forage quality using near infrared reflectance. *Agron. J.* **1982**, *74*, 585–588.
- Greenspan, L. Humidity fixed points of binary saturated aqueous solutions. J. Res. Natl. Bur. Stand., Sect. A. 1977, 81A, 89–96.
- Hoffmann, K. Feuchtemessung durch infrarotreflection. *Chem. Ing. Technol.* **1963**, *35*, 55–60.
- Hruschka, W. R. Data Analysis: Wavelength selection methods. In *Near-infrared Technology in the Agricultural and Food Industries*, Williams, P. C., Norris, K. H., Eds.; American Association of Cereal Chemists: St. Paul, MN, 1987; pp 53–54.
- Kays, S. E.; Windham, W. R.; Barton, F. E., II. Prediction of Total Dietary Fiber in Cereal Products using Near-Infrared Reflectance Spectroscopy. J. Agric. Food Chem. 1996, 44, 2266–2271.
- Martens, M.; Martens, H. Near-infrared reflectance determination of sensory quality of peas. *Appl. Spectrosc.* **1986**, *40*, 303–310.
- Martens, M.; Naes, T. Methods for calibration. In *Multivariate Calibration*; Martens, H., Naes, T., Eds.; John Wiley & Sons Ltd.: New York, 1989; pp 73–163.
- Murray, I. Forage analysis by near infr-red spectroscopy. In *Sward Measurement Handbook*, 2nd ed.; Davies, A., Baker, R. D., Grant, S. A., Laidlaw, A. S., Eds.; British Grassland Soc.: Reading, U.K., 1993; pp 285–312.
- Shenk, J. S.; Westerhaus, M. O. Population definition, sample selection, and calibration procedures for near infrared reflectance spectroscopy. *Crop Sci.* **1991a**, *31*, 469–474.
- Shenk, J. S.; Westerhaus, M. O. Population structuring of near infrared spectra and modified partial least squares regression. *Crop Sci.* **1991b**, *31*, 1548–1555.
- Walter, J. A.; Hope, A. B. Nuclear magnetic resconance and the state of water in cells. *Prog. Biophys. Mol. Biol.* **1971**, *23*, 1–20.
- Williams, P. C.; Norris, K. H. Qualitative applications of nearinfrared reflectance spectroscopy. In *Near-infrared Technol*ogy in the Agricultural and Food Industries; Williams, P. C., Norris, K. H., Eds.; American Association Cereal Chemists: St. Paul, MN, 1987; pp 246.
- 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.; Robertson, J. A.; Leffler, R. G. A comparison of methods for moisture determination of forages for near infrared reflectance spectroscopy calibration and validation. *Crop Sci.* **1987**, *27*, 777–783.
- Windham, W. R.; Gaines, S. G.; Leffler, R. G. Effect of wheat moisture on hardness scores determined by near-infrared reflectance and on hardness score standardization. *Cereal Chem.* 1993, 70, 662–666.

Received for review June 3, 1996. Accepted October 18, 1996. $^{\circ}$ Mention of company or trade names is for a purpose of description only and does not imply endorsement by the U.S. Department of Agriculture.

JF960385Y

 $^{\otimes}$ Abstract published in Advance ACS Abstracts, December 1, 1996.