

Concatenation of Near- and Mid-Infrared Spectra To Improve Calibrations for Determining Forage Composition

James B. Reeves III

Nutrient Conservation and Metabolism Laboratory, Livestock and Poultry Sciences Institute, ARS, USDA, Beltsville, Maryland 20705

The objective of this study was to determine the usefulness of combining near- and mid-infrared diffuse reflectance spectra to improve partial least squares calibrations. A set of 174 chlorite-treated feedstuffs were assayed chemically and spectroscopically for fiber, crude protein, lignin, digestibility, and lignin composition. All samples were scanned by diffuse reflectance in the near- and mid-infrared using a Fourier transform spectrometer equipped with a sample transport (to increase the amount of sample scanned) under various instrument configurations. Partial least squares analysis showed little or no improvement in calibration results by combining near- and mid-infrared spectra, although results using both regions together were at times better than those obtained using one of the two regions. In conclusion, results indicated that combining data from the near- and mid-infrared spectral regions offers little or no advantage over the use of a single spectral range.

Keywords: NIRS; mid-infrared; forage composition

INTRODUCTION

Over the last 20 years, the near-infrared (NIR) spectral region, roughly defined as the region from 1100 to 2500 nm (9091–4000 cm^{-1}), has been investigated and used extensively for the determination of material composition. It has been applied to the compositional determination of everything from animal feedstuffs (Marten et al., 1985; Barton, 1987) and foods (Osborne and Fearn, 1986) to medical diagnostics (Ciurczak, 1992) and industrial process monitoring (Kemeny, 1992). The NIR has been successfully applied to such a wide range of interests due to the nature of the spectral absorptions involved. Because NIR spectra are based on the overtone and combination band absorptions of CH, NH, and OH groups, almost any organic material can be investigated (Murray and Williams, 1987). Also, due to the weak nature of NIR absorptions, long path lengths and deep sample penetration are possible (Kemeny, 1992). On the other hand, inorganic materials, such as salt, cannot be determined by NIR spectroscopy unless they influence the spectra of other materials, such as water, which do absorb in the NIR (Hirschfeld, 1985). Finally, a wide range of sampling methods from diffuse reflectance and transmission (Williams, 1987) to fiber optic cables (Kemeny, 1992) has allowed the NIR to be applied to many problems in many different environments.

While NIR spectroscopy (NIRS) has enjoyed a great deal of success, until recently, the mid-infrared (MIR) has been limited to aiding in spectral interpretation (Barton et al., 1992) when it comes to many applications for which NIRS has proved so useful. Recent efforts have shown, however, that MIR in the reflectance mode can determine the composition of dried forages and byproducts as well as or even better than NIRS (Reeves, 1994a,b, 1996a). The question then arises as to whether any advantage can be gained by combining the two spectral regions. The objective of this study was to investigate that question using a set of chlorite-treated feedstuffs.

MATERIALS AND METHODS

Sample Generation and Chemical Analysis. One hundred seventy-four samples (Reeves, 1987) consisting of six hays

(alfalfa, tall fescue, orchardgrass, red clover, timothy, and a grass-legume mix of mainly orchardgrass and clover, with some timothy), two straws (barley and wheat), corn cobs, four stovers (two corn and two soybean), and three hulls (peanut, rice, and soybean) treated at one of 11 different levels of sodium chlorite were assayed chemically and spectroscopically. Samples were assayed (Reeves, 1987) in triplicate for neutral and acid detergent fiber (NDF and ADF, respectively), permanganate lignin, total crude protein (CP), cell wall and dry matter digestibility (CWDG and DMDG, respectively), and seven nitrobenzene oxidation products of lignin: *p*-hydroxybenzaldehyde (pHB), vanillin (VAN), a mixed product of acetovanillone and 4-allyl-2,6-dimethoxyphenol (MIX), vanillic acid (VA), syringaldehyde (SYAL), an unknown (UNK), and syringic acid (SA).

Spectra. Samples, ground using a cyclone grinder (20 mesh), were scanned by diffuse reflectance on a Digi-Lab FTS-65 Fourier transform spectrometer (Bio-Rad, Cambridge, MA) equipped with an MIR bench (KBr beamsplitter, ceramic source, TGS detector, and dry air purge), an NIR bench (quartz beamsplitter, halogen source, PbSe detector, and dry air purge), and a sample transport device (Reeves, 1996a). The transport device consisted of a linear motion device that allowed approximately a 50 mm by 2–3 mm sample path to be scanned. Neat samples were scanned in both spectral regions (NIR and MIR). For background spectra, sulfur for the NIR (sublimed, 100 mesh) and KBr for the MIR region were used. Near-infrared spectra were taken from 10 000 cm^{-1} (1000 nm) to 4000 cm^{-1} (2500 nm) at nominal resolutions (aperture open) of 4 cm^{-1} (actual resolutions: 10 cm^{-1} at 10 000 cm^{-1} to 4 cm^{-1} at 4000 cm^{-1} , corresponding to 1.0 nm at 1000 nm to 2.5 nm at 2500 nm) and 16 cm^{-1} (actual resolutions: 40 cm^{-1} at 10 000 cm^{-1} to 16 cm^{-1} at 4000 cm^{-1} , corresponding to 4 nm at 1000 nm to 10 nm at 2500 nm). Mid-infrared spectra were taken from 4000 to 400 cm^{-1} at actual resolutions of 4 and 16 cm^{-1} .

To produce a single spectrum from the NIR to the MIR (10 000 to 400 cm^{-1}), the data points in the region of 4000 cm^{-1} , common to both spectral ranges, were used to match the spectra. Briefly, the average (to reduce the effects due to random noise in any single data point) of the three overlapping data points was taken, and the ratio of the MIR to the NIR average was used to scale all the remaining points in the NIR spectrum. The MIR spectrum and the NIR spectrum (minus the three common spectral points) were then simply concatenated to produce one continuous spectrum at the resolution in question (4 or 16 cm^{-1}).

Statistical Analysis. Data were analyzed by SAS to calculate the chemical results and to produce means. All

Table 1. Summary of Chemical Values (% DM or % of Total Nitrobenzene Products, $N = 174$) Used for Calibrations (Reeves, 1987)

assay ^a	min ^b	max ^c	mean	SD
NDF	33.4	91.6	68.3	14.3
ADF	26.6	80.3	47.3	10.6
lignin	2.47	26.2	9.35	5.39
CWDG	1.52	100.2	65.9	27.2
DMDG	11.5	100.2	75.0	22.0
CP	1.55	16.8	7.87	5.21
Nitrobenzene Oxidation Products				
pHB	0.0	26.5	6.36	6.44
VAN	10.2	82.8	36.1	13.5
MIX	0.0	76.6	11.4	15.0
VA	0.0	37.9	10.3	7.59
SYAL	0.0	33.8	13.5	11.3
UNK	0.0	47.9	17.8	12.2
SA	0.0	24.6	4.53	4.04

^a NDF, neutral detergent fiber; ADF, acid detergent fiber; lignin, permanganate lignin; CWDG, cell wall digestibility; DMDG, dry matter digestibility; CP, total crude protein; pHB, *p*-hydroxybenzaldehyde; VAN, vanillin; MIX, mixed product of acetovanillone and 4-allyl-2,6-dimethoxyphenol; VA, vanillic acid; SYAL, syringaldehyde; UNK, unknown; SA, syringic acid. ^b Minimum. ^c Maximum.

spectroscopic data were analyzed using partial least squares regression or PLS (PLSplus V2.1G, Galactic Industries). Because of previous data analysis efforts, the data were analyzed using a series of 36 pretreatment/data point combinations found useful (Reeves, 1994a,b). Briefly spectra were pretreated by either "mean centering and variance scaling" or "multiplicative scatter correction" followed by either a first- or second-derivative treatment. In addition, calibrations using every n th spectral data point, where $n = 1, 2,$ or 4 , were examined (Reeves, 1994a,b). Where all data points were not used, the data points were averaged or smoothed to yield the final data (Galactic, 1992). A complete list of all pretreatments used may be found in Reeves (1994a). Finally, the data sets were randomized, and the same random order was used for all reference data (NDF, ADF, etc.) and all calibrations.

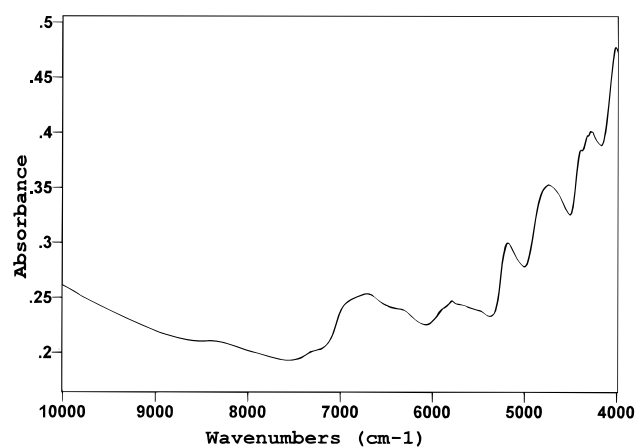
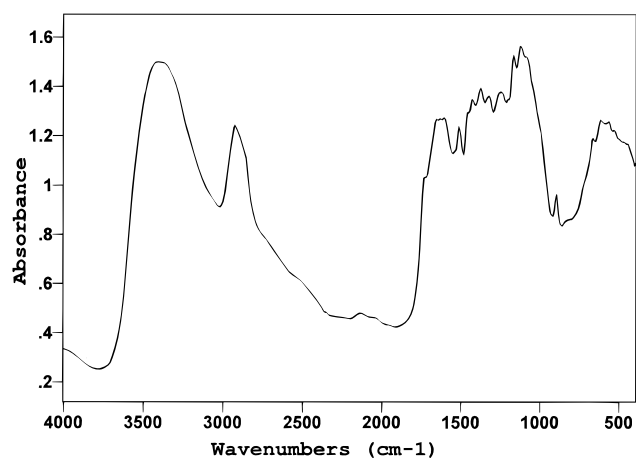
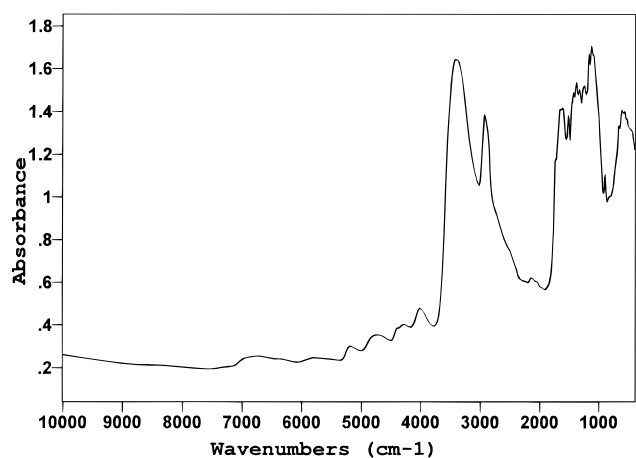
The best results (based on number of factors producing the lowest PRESS statistic) from a one-third out cross validation analysis [highest R^2 , lowest relative mean square deviation (RMSD) between actual and predicted values] are reported for each calibration discussed. In essence, the results presented represent the average results from developing three calibrations in which two-thirds of the samples are used to develop a calibration that is then applied to the remaining one-third. The procedure is automatically repeated three times so that each sample is in a prediction set once and only once (Galactic, 1992).

RESULTS AND DISCUSSION

Chemical Analysis. Table 1 shows a summary of the chemical values for the chlorite-treated materials (Reeves, 1987). These values were then used for the various spectroscopic calibrations. As can be seen, both a wide variety of constituents and values for each constituent were available for testing. In addition, the data presented came from 16 different materials representing a wide range of forages and byproducts ranging from high-quality hays to low-quality straws, hulls, and stovers. More detailed information on the composition of the samples may be found in Reeves (1987).

Spectra. The NIR, MIR, and combined spectra (16 cm^{-1} resolution, 256 scans) of untreated orchardgrass hay are shown in Figures 1–3 and for wheat straw in Figures 4–6. The combined spectra (Figures 3 and 6) are presented to demonstrate that the procedure used to concatenate the spectra did not result in any spectral aberrations, such as discontinuities at the concatenation point.

PLS Results for Fiber-Based Assays. In Table 2, the calibration results for NDF, ADF, lignin CWDG,

**Figure 1.** Near-infrared spectrum at 16 cm^{-1} resolution (256 scans) of untreated orchardgrass hay.**Figure 2.** Mid-infrared spectrum at 16 cm^{-1} resolution (256 scans) of untreated orchardgrass hay.**Figure 3.** Combined near- and mid-infrared spectrum at 16 cm^{-1} resolution (256 scans) of untreated orchardgrass hay.

DMDG, and CP, using various spectroscopic scans and spectra, are presented. Based on the R^2 or RMSD produced for the six assays, no apparent differences exist between the results using NIR, MIR, or the combined spectra. In general, only slight differences in R^2 or RMSD were found, with slightly better results (higher R^2 and lower RMSD) for ADF and CP found using NIR spectra, for lignin and CWDG using MIR spectra, and for NDF and DMDG using the combined spectra.

Examination of the calibration parameters (scans, resolution, math pretreatment, number of factors) used also did not indicate any strong pattern, with the possible exception of the number of scans and resolution

Table 2. Best Partial Least Squares Analysis Results for Analysis for Fibers, Permanganate Lignin, Digestibility, and Crude Protein Obtained Using Various Spectral Regions^a

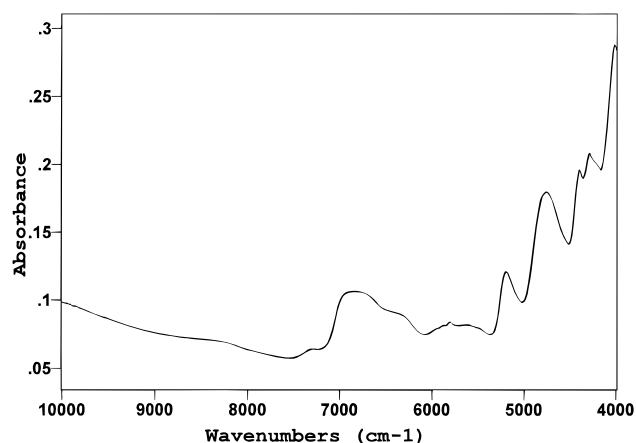
assay ^b	near-IR					mid-IR					combined				
	R/S ^c	math ^d	no. of F ^e	R ²	RMSD ^f	R/S	math	no. of F	R ²	RMSD	R/S	math	no. of F	R ²	RMSD
NDF	16/256	MSC-1-16-A	16	0.972	2.40	16/256	MSC-2-4-H	29	0.974	2.28	16/256	MSC-1-16-H	23	0.977	2.19
ADF	16/256	MSC-1-16-A	21	0.980	1.50	16/256	MSC-2-4-H	28	0.972	1.80	4/64	MSC-2-16-F	28	0.972	1.79
LIGNIN	16/256	MSC-1-16-A	31	0.939	1.33	4/64	MSC-1-4-F	31	0.956	1.14	16/256	M&V-1-4-A	31	0.953	1.18
CWDG	16/256	MSC-1-16-A	22	0.941	6.65	16/256	M&V-2-16-F	33	0.965	5.15	4/64	MSC-1-4-A	19	0.960	5.46
DMDG	16/256	MSC-1-16-A	25	0.961	4.31	4/64	M&V-2-8-A	28	0.972	3.69	4/64	MSC-2-8-H	33	0.973	3.61
CP	16/256	MSC-1-16-A	31	0.994	0.388	16/256	M&V-1-8-H	30	0.991	0.484	16/256	M&V-1-8-A	33	0.994	0.421

^a Near-IR, 10 000–4000 cm⁻¹; mid-IR, 4000–400 cm⁻¹; combined, 10 000–400 cm⁻¹. ^b NDF, neutral detergent fiber; ADF, acid detergent fiber; CWDG, cell wall digestibility; DMDG, dry matter digestibility; CP, total crude protein. ^c R/S, resolution of spectra in cm⁻¹ and number of co-added scans per spectrum. ^d Math, data pretreatment used: MSC and M&V = multiplicative scatter or mean & variance corrected, respectively, followed by derivative used: 1 = 1st and 2 = 2nd, gap in data points used and fraction of total available data points used: A = all, H = half, and F = one-fourth (data points averaged to produce the final number). ^e No. of F, no. of PLS factors used (based on minimization of PRESS statistic). ^f RMSD, relative means square difference.

Table 3. Best Partial Least Squares Analysis Results for Analysis for Lignin Nitrobenzene Oxidation Products Obtained Using Various Spectral Regions^a

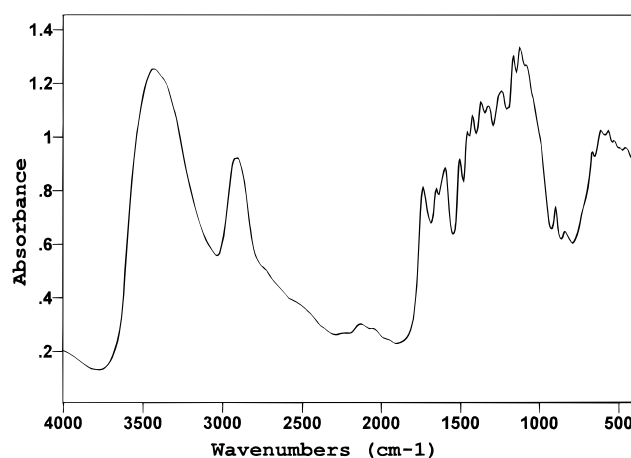
assay ^b	near-IR					mid-IR					combined				
	R/S ^c	math ^d	no. of F ^e	R ²	RMSD ^f	R/S	math	no. of F	R ²	RMSD	R/S	math	no. of F	R ²	RMSD
pHB	16/256	M&V-1-8-H	16	0.885	2.18	4/256	M&V-2-4-A	12	0.940	1.59	16/256	MSC-2-8-F	26	0.927	1.76
VAN	4/256	MSC-2-8-F	14	0.927	3.77	16/256	MSC-2-4-A	29	0.955	2.98	16/256	MSC-2-4-A	34	0.962	2.76
MIX	16/256	M&V-1-16-F	23	0.915	4.59	16/64	M&V-2-4-F	23	0.945	3.57	4/256	MSC-2-4-H	10	0.938	3.79
VA	16/256	MSC-2-4-H	22	0.837	3.07	4/64	M&V-1-8-F	25	0.899	2.45	4/64	MSC-2-8-F	21	0.889	2.60
SYAL	16/256	MSC-2-4-A	15	0.878	4.05	16/256	M&V-1-16-A	30	0.967	2.07	4/256	MSC-2-4-F	20	0.964	2.14
UNK	16/256	MSC-2-4-A	17	0.847	4.79	4/256	MSC-2-4-A	10	0.887	4.10	4/256	MSC-2-4-H	12	0.883	4.18
SA	16/256	M&V-1-16-A	13	0.575	2.67	4/64	M&V-2-16-F	10	0.734	2.11	16/256	M&V-1-8-A	21	0.713	2.25

^a Near-IR, 10 000–4000 cm⁻¹; mid-IR, 4000–400 cm⁻¹; combined, 10 000–400 cm⁻¹. ^b pHB, *p*-hydroxybenzaldehyde; VAN, vanillin; MIX, mixed product of acetovanillone and 4-allyl-2,6-dimethoxyphenol; VA, vanillic acid; SYAL, syringaldehyde; UNK, unknown; SA, syringic acid. ^c R/S, resolution of spectra in cm⁻¹ and number of co-added scans per spectrum. ^d Math, data pretreatment used: MSC and M&V = multiplicative scatter or mean & variance corrected, respectively, followed by derivative used: 1 = 1st and 2 = 2nd, gap in data points used and fraction of total available data points used: A = all, H = half, and F = one-fourth (data points averaged to produce the final number). ^e No. of F, no. of PLS factors used (based on minimization of PRESS statistic). ^f RMSD, relative means square difference.

**Figure 4.** Near-infrared spectrum at 16 cm⁻¹ resolution (256 scans) of untreated wheat straw.

of the spectra used. With the NIR spectra, the best results were always obtained using 256 scans at 16 cm⁻¹ resolution, while for the MIR and combined spectra a mix of 4 and 16 cm⁻¹ resolution spectra produced the best results. The fact that, at 4 cm⁻¹ resolution, 64 scans was preferred to 256 indicates that some sample heating or damage may still be occurring even with the use of the sample transport device, since more scans would normally yield lower noise and therefore better spectra. Finally, examination of the math pretreatments and number of factors used shows that, except for the consistency found for the NIR spectra, no clear pattern emerges.

In summary, for the fiber-based measures examined, combining NIR and MIR spectra does not appear to offer any advantages or result in any dramatic changes in the calibration parameters found to yield the best

**Figure 5.** Mid-infrared spectrum at 16 cm⁻¹ resolution (256 scans) of untreated wheat straw.

calibration. The results thus indicate that the information needed to obtain a calibration was present in either spectral range (NIR or MIR).

PLS Results for Lignin Oxidation Products. In Table 3, the calibration results are presented for various spectroscopic scans and spectra variations for seven nitrobenzene oxidation products of lignin. For these assays, the best results were obtained using the MIR spectra alone for all but one of the assays (VAN), although for all seven products the results using the combined spectra were better than those obtained using the NIR spectra. However, overall, the differences between the various results were again quite small. Also, as with the results with the various fiber-based measures, no pattern was apparent with respect to the various calibration parameters, with a wide variety of math treatments, calibration factors, etc. being used.

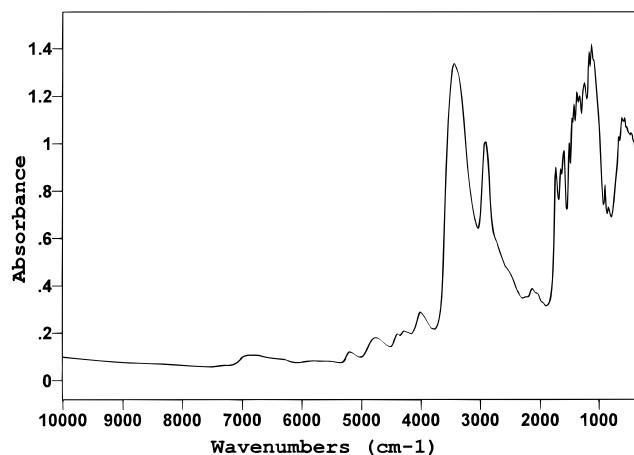


Figure 6. Combined near- and mid-infrared spectrum at 16 cm^{-1} resolution (256 scans) of untreated wheat straw.

However, one difference can be seen in that, at 4 cm^{-1} resolution, 256 scans was preferred in several cases, regardless of the spectral range used. This was somewhat surprising in that the same spectra were used for all calibrations and with the fiber-based assays 64 scans gave better results than 256. The reason for this difference is not known, but indicates that the sample heating or damage that occurs with increased scans and/or higher resolution spectra damages or alters specific portions of the spectra. If these spectral regions are important to a specific determination (i.e., NDF, ADF, etc.) then better results are obtained with a decrease in the time the sample is exposed to the MIR source. If the region is not important, then increasing the number of scans improves the calibration by producing lower noise spectra.

In summary, it would appear that combining NIR and MIR spectra offers little or no benefit over MIR spectra alone for analysis of lignin oxidation products, although the results for combined spectra were an improvement over the results using NIR spectra alone.

CONCLUSIONS

Results have shown that for spectra obtained using a Fourier transform spectrometer equipped with a sample transport device, the combining of MIR and NIR spectra results in little or no improvement in PLS calibration results for a variety of compositional quantities. This is in contrast to earlier results, obtained using a stationary sample cup, where results indicated some improvement when spectra were concatenated (Reeves, 1996b). Since with a static cell less sample would be scanned (2–3 mm diameter area as contrast with a 2–3 \times 50 mm area using the transport cell) and the sample would be subject to more heating, the two sets of results indicate that concatenating spectra may be somewhat useful where lower quality spectra are concerned. It might also be speculated that using data from the two spectral regions might be helpful if only low-resolution spectra were available.

ABBREVIATIONS USED

ADF, acid detergent fiber; CP, total crude protein; CWDG, cell wall digestibility; DMDG, dry matter digestibility; MIR, mid-infrared; MIX, mixed product of acetovanillone and 4-allyl-2,6-dimethoxyphenol; NDF, neutral detergent fiber; NIR, near-infrared; NIRS, near-infrared spectroscopy; pHB, *p*-hydroxybenzaldehyde; PLS, partial least squares regression; RMSD, relative

mean square deviation; SA, syringic acid; SYAL, syringaldehyde; UNK, unknown; VA, vanillic acid; VAN, vanillin.

LITERATURE CITED

- Barton, F. E., II. Analytical Application to Fibrous Foods and Commodities. In *Near-Infrared Technology in the Agricultural and Food Industries*; Williams, P. C., Norris, K., Eds.; American Association of Cereal Chemists, Inc.: St. Paul, MN, 1987; p 169.
- Barton, F. E., II; Himmelsbach, D. S.; Duckworth, J. H.; Smith, M. J. Two-Dimensional Vibration Spectroscopy: Correlation of Mid- and Near-Infrared Regions. *Appl. Spectrosc.* **1992**, *46*, 420–429.
- Ciurczak, E. W. NIR Analysis of Pharmaceuticals. In *Handbook of Near-Infrared Analysis*; Burns, D. A., Ciurczak, E. W., Eds.; Marcel Dekker, Inc.: New York, 1992; pp 549–564.
- Galactic Industries Corp. *PLSplus V2.1G*; Galactic Industries Corp., Salem, NH, 1992.
- Hirschfeld, T. Salinity determination using NIRA. *Appl. Spectrosc.* **1985**, *39*, 740–741.
- Kemeny, G. J. Process Analysis. In *Handbook of Near-Infrared Analysis*; Burns, D. A., Ciurczak, E. W., Eds.; Marcel Dekker, Inc.: New York, 1992; pp 53–106.
- Marten, G. C., Shenk, J. S., Barton, F. E., II, Eds. *Near infrared reflectance spectroscopy (NIRS): Analysis of forage quality*; USDA Agricultural Handbook 643; U.S. Government Printing Office: Washington, DC, 1985.
- Murray, I.; Williams, P. C. Chemical Principles of Near-Infrared Technology. In *Near-Infrared Technology in the Agricultural and Food Industries*; Williams, P. C., Norris, K., Eds.; American Association of Cereal Chemists, Inc.: St. Paul, MN, 1987; p 17.
- Osborne, B. G.; Fearn, T. *Near Infrared Spectroscopy in Food Analysis*; John Wiley and Co.: New York, 1986.
- Reeves, J. B., III. Sodium chlorite treatment of plant materials: Fiber and lignin composition, digestibility, and their interrelationships. *J. Dairy Sci.* **1987**, *70*, 2534–2549.
- Reeves, J. B., III. Near- versus mid-infrared diffuse reflectance spectroscopy for the quantitative determination of the composition of forages and by-products. *J. Near Infrared Spectrosc.* **1994a**, *2*, 49–57.
- Reeves, J. B., III. Near- Versus Mid-Infrared Spectroscopy for Quantitative Analysis of Chlorite Treated Forages and By-products. *J. Near Infrared Spectrosc.* **1994b**, *2*, 153–162.
- Reeves, J. B., III. Improvement in Fourier near- and mid-infrared diffuse reflectance spectroscopic calibrations through the use of a sample transport device. *J. Appl. Spectrosc.* **1996a**, *50*, 965–969.
- Reeves, J. B., III. Combining near and mid-infrared reflectance spectra: Any spectral advantage? In *Near Infrared Spectroscopy: The Future Waves*; Davies, A. M. C., Williams, P., Eds.; NIR Publications: West Sussex, United Kingdom, 1996b; p 14.
- SAS. *SAS User's Guide: Statistics*, version 6.03; SAS Institute: Cary, NC, 1988.
- Williams, P. C. Commercial Near-Infrared Reflectance Analyzers. In *Near-Infrared Technology in the Agricultural and Food Industries*; Williams, P. C., Norris, K., Eds.; American Association of Cereal Chemists, Inc.: St. Paul, MN, 1987; pp 107–142.

Received for review June 13, 1996. Accepted February 19, 1997. [⊗] Mention of a trade name, proprietary product, or specific equipment does not constitute a guarantee or warranty by the U.S. Department of Agriculture and does not imply its approval to the exclusion of other products that may be suitable.

JF9604246

[⊗] Abstract published in *Advance ACS Abstracts*, April 15, 1997.