What Three Questions Should I Ask Before Using NIRS?

F. E. Barton, II
USDA-Agricultural Research Service
Richard B. Russell Agricultural Research Center
Athens, GA 30604-5677

ABSTRACT

Three components are required for any spectroscopic method: 1) an instrument to make the measurement, 2) a reference value and 3) a mathematical algorithm to relate them. It also can be said that three things are required for any spectroscopic chemometric technique to be successful. First, the reference data must have a real relationship to the spectral data so that the appropriate statistical data treatment can be employed. The second requirement is a well behaved instrument, i.e. high Signal-to-Noise ratio **(S/N)** and wavelength precision. Third, the spectra must be obtained in an optimal geometry to permit the first two criteria to achieve the best results. There are also three things that are required for any analytical method to be accepted as a *good* method. First, it must be accurate. Second, it must be rapid. Third, it must be inexpensive. The above *trios* provide the framework for answering the questions about the use of spectroscopic techniques for the analysis of forages, feeds, foods and fibers. Examples herein show both the utility of these chemometric methods and the folly. Near infrared spectroscopy started in agriculture and in recent years has exploded into a major analytical realm for analysis in just about every venue.

INTRODUCTION

The use of near infrared spectroscopy **(NIRS)** in ruminant nutrition goes back to the mid 1970's (Norris et al., 1976). The National Near Infrared Research Project sponsored by the Agricultural Research Service **(ARS)** was begun in 1979 (Marten et al., 1985). There are three questions which one must ask before employing NIRS as a means of analysis. These questions are listed below:

1. Do you *truly* understand the reference methods, quality indices, or properties you want to measure?
2. Why do you want a spectroscopic analytical procedure and is near infrared the correct spectral region?
3. How accurate/precise must the measurement be and do you *truly* understand the chemometric statistics?

If you can answer these three questions then you know the requirements of an analytical method, realize the limitations of the instrumentation, and understand the rudiments of chemometrics. If you cannot answer all of them, then you are like most of us and realize there is risk associated with the models we develop. The use of NIRS will be traced using the analysis of forages, grains, and food as examples. The early work was principally on forages and provided some unique challenges. Forage structure and its relation to quality was not always matched to the quality index used to measure composition.

RESULTS AND DISCUSSION

Question 1. How well do we know the reference assay?

Fibrous materials traditionally have been analyzed by the Weend Proximate Analysis Procedures as a means of estimating total digestible nutrients. In the proximate analysis procedures, % dry matter is determined by oven drying, % crude protein is expressed as 6.25 X % nitrogen from the Kjeldahl analysis, % fat by ether extraction, % crude fiber by alternate base and acid treatments, and % ash by incineration. These procedures continue to be the standard methods in use by many state testing

TABLE 1: Percent compositional analysis of grasses[a]

Grass	IVDMD[b]	Protein	Ash	NDF[c]	ADF[d]	Hemi-cellulose	Holo-cellulose	PML[e]
Coastal: 4 week	66.1	19.2	7.8	61.0	29.1	31.8	61.2	4.1
Coastal: 8 week	50.4	11.0	5.4	71.2	40.0	31.2	66.8	6.0
Coastcross-1: 4 week	66.1	18.7	7.2	60.0	31.9	28.1	53.6	3.5
Coastcross-1: 8 week	54.9	13.9	7.2	62.9	39.0	23.9	55.6	5.5
Bahia: 4 week	59.6	15.7	6.2	71.0	35.7	35.3	60.8	3.4
Bahia: 8 week	53.2	9.2	5.9	67.5	35.0	32.5	76.0	5.3
Pangola: 4 week	54.5	7.0	4.7	69.4	41.7	22.7	57.3	6.3
Pangola: 8 week	48.6	5.9	5.1	67.0	29.5	37.6	47.4	4.8
Average tropical	57.8	12.6	6.2	66.2	35.2	30.4	59.8	4.9
Kenhy: 4 week	65.6	13.2	8.3	58.2	33.6	24.7	41.6	3.2
Ken-Blue: 4 week	58.1	15.5	7.1	54.0	30.6	23.5	43.5	4.3
Brome: 4 week	64.2	14.3	8.8	56.2	34.3	21.9	51.1	4.8
Orchard: 4 week	62.8	14.8	8.2	57.9	33.3	24.6	43.9	4.1
Kentucky-31: 4 week	62.7	14.2	8.8	58.4	31.4	25.5	45.2	3.4
Timothy: 4 week	66.8	13.4	8.8	55.6	34.7	20.9	42.0	4.1
Kentucky-31: 4 week (fall)	55.0	12.6	8.4	59.8	31.6	28.2	46.0	5.6
Kenhy: 4 week (fall)	60.1	12.6	8.0	57.3	30.7	26.6	43.8	4.2
Orchard: 4 week (fall)	58.8	17.8	9.7	54.0	29.4	24.6	45.5	4.8
Ken-Blue: 4 week (fall)	61.0	17.3	6.3	57.6	27.5	30.1	40.3	4.1
Average temperate	61.5	14.6	8.2	56.9	31.7	25.1	44.3	4.3
Average standard deviation	2.62	0.13	0.13	1.08	1.00	1.04	1.00	0.30

[a] Source: Barton et al. (1976); used by permission.
[b] In vitro dry matter digestibility.
[c] Neutral detergent fiber.
[d] Acid detergent fiber.
[e] Permanganate lignin.

laboratories. These analyses are empirical. The assumption is made that the reagents or experimental conditions affect each sample in an identical manner. They are all gravimetric procedures and the calculated results are relative percentages. Moore and Mott (1973) and Martens and Russwurm (1981) have published excellent reviews which detail the status of gravimetric forage analyses. Since the molecular weight of a forage sample or any constituent therein cannot be determined, these percentages are the only way to express compositional properties quantitatively. The analyses are very dependent on sampling techniques, technician experience, and the environment in which the sample is analyzed. Finally, all of the procedures are destructive. The literature values for many forages are comparable to those in Table 1 which were obtained by the detergent analyses procedures. However, the assumption is made that the reagents are acting on all samples equally regardless of species, environment of growth and agricultural management practices.

One way to examine the question of what is neutral detergent fiber **(NDF)** or acid detergent fiber **(ADF)** is with microscopy, i.e. evaluation of leaf sections before and after microbial digestion. Akin et al. (1975) examined the tissues that comprise the residues of NDF and ADF of leaf sections from a warm-season, Coastal bermudagrass **(CBG)** and cool-season, Kentucky-31 **(KY-31)** grass by scanning electron microscopy. In these experiments, 5 mm sections of the leaf blades were treated with the boiling reagents, prepared for microscopy, and viewed with scanning electron microscopy. The mild treatment with neutral detergent reagent left the cell

TABLE 2: Percent residue of neutral detergent fiber and acid detergent fiber from whole, Wiley-Milled forage and intact leaf samples of Coastal bermudagrass and Kentucky-31 Tall Fescue [a]

	Neutral Detergent Fiber		Acid Detergent Fiber	
Grass	**Whole [b]**	**Leaf [c]**	**Whole [b]**	**Leaf [c]**
Coastal bermudagrass	59.4 ± 0.3	78.3 ± 1.6	29.1 ± 0.8	25.3 ± 0.7
Kentucky-31 tall fescue	50.7 ± 0.6	79.1 ± 2.3	28.6 ± 0.2	27.8 ± 1.3

[a] Source: Akin et al. (1975); used by permission.

[b] Average of 12 determinations plus standard deviation for whole, ground samples.

[c] Average of 3 determinations plus standard deviation for leaf samples.

walls virtually intact in CBG and slightly distorted the mesophyll in some KY-31 samples. The amount of tissue removed from the leaf section was determined gravimetrically. It was found that much less tissue was removed from the sections than from Wiley-Milled, ground leaf blades. The NDF conditions were such that the fragile cell wall membranes were not ruptured and cell contents not removed unless the cell was opened by the knife when the sections were cut (Table 2). Treatment with acid detergent reagents revealed differences both between species and for all species when compared to digestion. For the warm-season CBG, the residue contained portions of the parenchyma bundle sheath. This tissue, which resisted the acidic treatment, is slowly degraded by rumen microorganisms. The opposite is true for KY-31. The only tissues remaining after 1 hr treatment were cuticle, sclerenchyma patches, and pieces of vascular tissue. This far exceeds the digestion of KY-31 by rumen microorganisms. Thus, as a measure of extent of digestion (Rohweder et al., 1978), ADF would overestimate the digestion of KY-31 and underestimate the digestion or nutritive value of CBG. Direct comparisons of quality estimated from ADF values between temperate and tropical (cool- and warm-season) grasses must be made with caution. The differential response of the plant cell wall to these analytical reagents reflects differences in their availability to rumen microorganisms and a linear response suitable for all species should not be expected.

Warm-season grasses are recognized to have higher fiber contents than cool-season grasses. The values in Table 1 taken from Barton et al. (1976) reflect average increases of ten percentage units in NDF and four percentage units in ADF for the warm season grasses. These differences persist when only the four week warm-seasons are compared (NDF ave. 65.0, ADF ave. 34.6). The lignin data are of particular interest. The higher fiber content, less digestible, warm-season grasses also have a higher lignin content (4.9 versus 4.3). If one considers only the four week regrowth samples, the higher fiber/less digestible warm season grasses are identical in lignin (4.3%) and virtually identical in digestibility (61.6 to 61.5%) to the cool-season grasses, while maintaining an average of nine percentage units NDF and four percentage units ADF higher fiber content. Clearly compositional differences do not answer the question of quality and animal performance differences for warm-season versus cool-season forages.

The analysis of highly fibrous feeds with NIRS by diffuse reflectance is different from that of grains in several respects. The components of the plant matrix are more complex and involve numerous discrete interactions. The work by Hruschka and Norris (1982) showed that for ground wheat, the summation of spectra of the chemical components (i.e., protein, starch, cellulose, moisture, and simple sugars) did not adequately reflect the total spectral composition when curve-fitting techniques were applied to the spectra. When the complexities of a forage sample are considered, it becomes obvious that the interaction of protein with lignin and carbohydrate along with minor constituents would make analyses by *pure components* impossible. Alternatively, it is possible to consider analyses based solely on functionality present in the spectrum if their relationship to some measure of quality was known. This requires a much better understanding of

TABLE 3: Effect of Laboratory data on calibration[a]

Analysis	Run[b]	Mean + SD	λ's[c]	R^2[d]	SEC[e]	Repeat[f]
Dry matter	I	93.3 ± 1.82	3	0.63	1.11	0.03
	R	95.3 ± 1.16		0.73	0.61	0.06
	C	94.7 ± 1.06		0.84	0.43	0.02
Protein	I	12.2 ± 1.96	3	0.84	0.80	0.03
	R	12.4 ± 2.00		0.87	0.71	0.04
	C	12.3 ± 2.01		0.94	0.49	0.05
Neutral detergent fiber	I	67.7 ± 3.12	5	0.65	1.86	0.16
	R	68.2 ± 2.91		0.73	1.52	0.08
	C	67.6 ± 2.71		0.82	1.15	0.27
Acid detergent fiber	I	38.8 ± 2.55	3	0.63	1.55	0.10
	R	38.6 ± 2.95		0.81	1.27	0.25
	C	38.8 ± 2.90		0.87	1.04	0.09
Permanganate lignin	I	4.9 ± 1.14	3	0.38	0.90	0.05
	R	3.7 ± 1.23		0.66	0.71	0.01
	C	3.7 ± 0.87		0.61	0.54	0.01
IVDMD[g]	I	60.4 ± 4.10	3	0.68	2.33	0.21
	R	60.2 ± 5.78		0.65	3.40	0.33
	C	62.5 ± 3.34		0.83	1.36	0.11

[a] Source: Coleman and Barton (1982); used by permission.

[b] I = file data were several years old. R = samples were all reanalyzed routinely. C = samples from R which were statistically outliers were reanalyzed and the new data incorporated into the file.

[c] Number of wavelengths used in equation.

[d] Coefficient of determination.

[e] Standard error of calibration.

[f] Repeatability error.

[g] In vitro dry matter digestibility.

both the spectra of forages and what constitutes quality than we currently have. Therefore, the analyses must be made on the basis of the spectra correlated to empirical results. Again, the amount of fiber is much greater, i.e. the fiber is the matrix as opposed to being a component, as in most foods and feeds. There has been very little work completed on fiber content and fiber structure in foods and grains. The study by Lund and Smoot (1982) concerned the dietary fiber content of tropical fruits and vegetables. These authors found that not only was the amount of fiber small and variable, but the fibrous fraction differed between species, and from that of forage plants. Baker (1978) and Baker and Holden (1981) examined cereals and grains for fiber content and evaluated several methods of fiber analysis to see which one would be most suitable for cereals. Baker (1978) found that a buffered ADF determination apparently improved recovery of cellulose. The general viewpoint of these authors was that no one method seemed to be best as no one method gave a number that corresponded to an identical fraction in all cereals and grains. The new enzymatic procedures of Lee et al. (1992) may provide both an assay and a means of characterizing fiber for monogastrics and ruminants.

Tables 2-4 contain data which illustrates several examples of how NIRS improves reference data and serves as a validation and in most cases a more reliable number. The experiment in Table 3 shows that it is important to analyze samples at the same time the spectra are taken, and that only *good* replicated values build models with low standard errors of calibration. Table 4 illustrates how much of an improvement taking the analyst out of the procedure improves results. The semi-automated FiberTec system is one example of how to accomplish this. Table 5 shows a laboratory error in in vitro dry matter digestibility **(IVDMD)** that would go undetected without NIRS. In this case three racks of tubes contained residual detergent from washing

TABLE 4: Effect of laboratory data on calibration

Analysis	Type[a]	Range	$\overline{SD}$	λ's[b]	R^2 [c]	SEC[d]	Repeat[e]
Protein	A	7-21	0.62	1	0.94	1.07	0.1676
	B	8-23	0.18	1[f]	0.95	1.05	0.1434
	R	7-22	0.15	1	0.94	1.10	0.1789
Neutral detergent fiber	A	43-76	1.99	3	0.97	1.45	0.2750
	B	43-75	0.45	3	0.96	1.76	0.2693
	F	44-75	0.42	3	0.98	1.24	0.2458
Acid detergent fiber	A	30-44	2.03	3	0.93	1.18	0.2627
	B	29-45	0.25[g]	3	0.85	1.84	0.2994
	F	28-43	0.43	3	0.87	1.49	0.2684
Permanganate lignin	A	3-12	0.51	3	0.92	0.71	0.1662
	B	2-12	0.21	2	0.71	1.50	0.1029
	F	----	----	----	-----	----	----
Acid detergent lignin	A	3-10	0.56	3	0.87	0.70	0.1552
	B	3-10	0.11	3	0.90	0.67	0.1201
	F	1.5-10	0.30	3	0.94	0.57	0.1386
In vitro dry matter digestibility	A	42-76	2.57	4	0.95	2.00	0.6426
	B	47-69	1.04	4	0.95	1.38	0.5251
	F	----	----	----	----	----	----

[a] A = average of four best labs (conventional method); B = conventional method data from Athens laboratory; R = rerun percentage protein in Athens laboratory; F = Fibertec analysis data.

[b] Number of wavelengths used in equation.

[c] Coefficient of determination.

[d] Standard error or calibration.

[e] Repeatability error.

[f] Using 3 degrees $R^2 = 0.98$, SEC = 0.64.

[g] Six replications per sample.

that lowered the IVDMD results. The triplicate tubes were all consistent. Table 6 describes a result on the *easiest* laboratory assay, oven dry matter. Here 20 samples were weighed out early in the morning (1-20), after the mid-morning break (21-40) and after lunch (41-60). This consistent bias is the result of opening the desiccator door 180 times in one day. Moisture collected on the crucibles and was counted as moisture when in fact it was not part of the sample. Oven dry matter is a very precise way to measure a wrong number. The NIRS caught all these errors that the laboratory could not.

Question 2. Why do you want an NIRS or spectroscopic method?

In 1994 the ***Nutritional Labeling and Educational Act*** became law. Earlier in 1993 the Federal Grain Inspection Service **(FGIS)** was to institute a classification system to classify grain by *end use*. At the same time the U.S. Environmental Protection Agency **(EPA)** enacted rules to limit waste chemicals from research laboratories. These three pieces of legislation have created a regulatory dilemma. How does a regulatory agency comply with increased requirements for analysis and reduce the level of chemical waste generated in the laboratory? The U.S.A. is not alone in the world with these requirements. The European Community **(EC)**, Australia and some Pacific Rim Nations have similar rules. The answer in all cases has been to employ spectroscopic analysis with the aid of chemometric models. The major obstacle is the lack of *certified* or *official* methods. In order for quality parameters to be regulated, their accuracy and precision must be able to stand the test of litigation. We have been involved with the FGIS and Food Safety Inspection Service (**FSIS**) on studies to develop NIR methods to measure quality parameters for foods.

TABLE 5: Actual versus predicted IVDMD[a]

Sample No.	Average Residual Size	Bias
1-26	2.5	Positive and negative
27-36	10.0	All negative
37-40	4.4	All negative
41-60	1.7	Positive and negative

[a] In vitro dry matter digestibility.

TABLE 6: Actual versus predicted dry matter

Sample No.	Average Residual Size	Bias
1-20	1.74	All negative
21-40	0.42	0.0
41-60	1.78	All positive

Question 3. Do we know what the NIRS statistics mean?

Chemometrics as a discipline within chemistry can be defined as the development and application of mathematical and statistical methods to extract useful chemical information from chemical measurements (Kowalski, 1977). The extraction of compositional information from spectral curves can be considered as a chemometric method (Norris, 1983a, 1983b). The basic impetus for the development of *predictive analyses* is the increasing cost of performing laboratory analyses and the time required to obtain the results. Chemometric methods have been used for decades. Whenever a standard curve is constructed from a series of standard solutions assuming linearity from Beer's Law and used to read the concentration of unknowns directly from the scale knowing only its absorbance, a chemometric method has been used. When I last attended a Texas A&M University Ruminant Nutrition Conference in Bryan-College Station 16 years ago, I would have answered yes to this question and been wrong. Today multiple linear regression has largely been replaced by partial least squares and principal component regression and there are hundreds of applications of chemometrics in the literature each year. It is possible to find chemometrics and NIRS taught in graduate schools as part of physical science and agricultural curricula. One example of understanding the model is the classification of wheat. Figure 1 shows a simple classification of wheat using three principal components. While this seems to be a very successful model, it only works on small sets. When the number of samples is large the relationships which create the differences in the variance described by the three components falls apart. That is, the differences between individual samples becomes less, so statistically there is no basis on which to adequately classify them. Last year we showed that the spectral differences between wheats (hard and soft) were very slight. In Figure 2 the NIR, Mid-Infrared **(MIR)**, and Raman spectra of a single wheat sample is shown. The NIR spectra is the usual broad curve with prominent bands for water, carbohydrate, and protein. The MIR and Raman spectra are complementary and similar in their information content. The principle differences are the absence of O-H stretch in water and the Amide II band (1530 cm^{-1}). These characteristics should enable us to classify the wheat for its use based on spectra alone. However, the statistics will be more complex because we will need spectral data from multiple spectral regions.

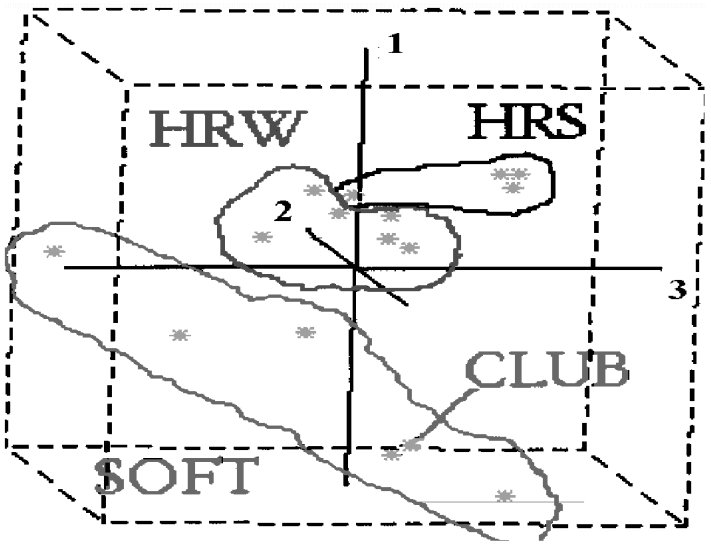

FIGURE 1: Classification of wheat by NIRS and principal component analysis

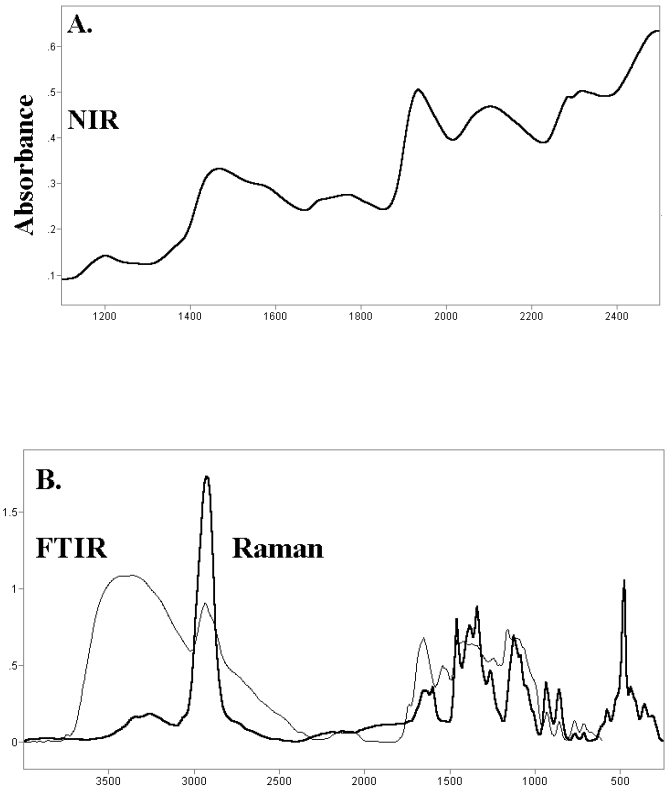

FIGURE 2: The (A.) NIR, (B.) FTIR and Raman Spectra of a wheat sample.

NIRS spectrometers have changed considerably since the days of the National NIRS Research Project. The sampling devices and software have gone through 4-5 generations. On-line and process control are the *hot* topics at PITCON and Eastern Analytical Symposiums. There are now more NIRS papers and sessions at major chemistry meetings than Fourier Transform Infrared **(FTIR)**. The technology is rapidly changing the way food, pharmaceutical, clinical and agricultural industries do business. Chemometric procedures and NIRS play a bigger role in the regulation of commerce and are used as *official* results. So as we head into the 21st century the question may really be, how can I apply NIRS and chemometrics to make my job more productive.

CONCLUSIONS

Question 1. If anything has been learned because of NIRS it is that the errors in our reference methods are much larger than we imagined, and that statistical relationships we felt were on solid ground were tenuous.

Question 2. Costs and the regulatory requirements alone will push us to use spectroscopic methods of analysis. We must choose the best region for specific measurements. We must accept the advantages of process control along with the initial cost of implementation.

Question 3. If we do not understand what the chemometric statistics mean and how to interpret them there are many short courses offered which can help. This technology is not a case where you can *leave the driving to someone else*.

LITERATURE CITED

Akin, D.E., F.E. Barton, II and D. Burdick. 1975. Scanning electron microscopy of Coastal bermuda and Kentucky-31 tall fescue extracted with neutral and acid detergents. J. Agric. Food Chem. 23: 924.

Baker, D. 1978. Fiber in wheat foods. Cereal Foods World 23: 557.

Baker, D. and J.M. Holden. 1981. Fiber in breakfast cereals. J. Food Sci. 46: 396.

Barton II, F.E., H.E. Amos, D. Burdick and R.L. Wilson. 1976. Relation of chemical analysis to in vitro digestibility for selected tropical and temperate grasses. J. Anim. Sci. 43: 504.

Coleman, S.W. and Barton II, F.E. 1982. Okla. Agr. Expt. Sta. Misc. Publ. 109.

Hruschka, W.R. and K.H. Norris. 1982. Least-squares curve fitting of near infrared spectra predicts protein and moisture content of ground wheat. Appl. Spectrosc. 36(3): 261.

Kowalski, B.R. 1977. Chemometrics: Theory and Applications. ACS Symposium Series No. 52. American Chemical Society, Washington, DC.

Lee,S.C., Prosky, L. And DeVries, J.W. 1992. Determination of total, soluble and insoluble fiber in foods-enzymatic-gravimetric method, MES-TRIS buffer: collaborative study. J.AOAC Int. 75:395.

Lund, E.D. and J.M. Smoot. 1982. Dietary fiber content of some tropical fruits and vegetables. J. Agric. Food Chem. 30: 1123.

Marten, G.C., J.S. Shenk and F.E. Barton, II, Editors. 1988. NIRS analysis of forage quality. USDA Handbook #643.

Martens, H. and H. Russwurm Jr., Editors. 1983. Food Research and Data Analysis. Applied Science Publishers. New York, New York.

Moore, J.E. and G.O. Mott. 1973. Structural inhibitors of quality in tropical grasses in anti-quality components of forages. A.G. Matches, ed. *In* Anti-quality components of forage crops. Crop Science Society of America, Madison, Wisconsin, p. 155.

Norris, K.H. 1983a. Extracting information from spectrophotometric curves. Predicting chemical composition from visible and near-infrared spectra. *In* Food Research and Data Analysis. H. Martens and H. Russwurm Jr., eds. Applied Science Publishers. New York, New York. p. 95.

Norris, K.H. 1983b. Multivariate analysis of raw materials. *In* Chemistry and World Food Supplies: The New Frontiers, CHEMRAWN II (Chemical Research Applied to World Needs). L.W. Shemilt, ed. Pergamon Press, New York, New York. p. 527.

Norris, K.H., R.F. Barnes, J.E. Moore and J.S. Shenk. 1976. Predicting forage quality by infrared reflectance spectroscopy. J. Anim. Sci. 43(4): 889.

Rohweder, D.A., R.F. Barnes and N. Jorgensen. 1978. Proposed hay grading standards based on laboratory analyses for evaluating quality. J. Anim. Sci. 47: 747.