

# Sampling Requirements for Forage Quality Characterization of Rectangular Hay Bales

Craig C. Sheaffer,\* Neal P. Martin, Jane Grimsbo Jewett, James Halgerson, Roger D. Moon, and Greg R. Cuomo

## ABSTRACT

Commercial lots of alfalfa (*Medicago sativa* L.) hay are often bought and sold on the basis of forage quality. Proper sampling is essential to obtain accurate forage quality results for pricing of alfalfa hay, but information about sampling is limited to small, 20- to 40-kg rectangular bales. Our objectives were to determine the within-bale variation in 400-kg rectangular bales and to determine the number and distribution of core samples required to represent the crude protein (CP), acid detergent fiber (ADF), neutral detergent fiber (NDF), and dry matter (DM) concentration in commercial lots of alfalfa hay. Four bales were selected from each of three hay lots and core sampled nine times per side for a total of 54 cores per bale. There was no consistent pattern of forage quality variation within bales. Averaged across lots, any portion of a bale was highly correlated with bale grand means for CP, ADF, NDF, and DM. Three lots of hay were probed six times per bale, one core per bale side from 55, 14, and 14 bales per lot. For determination of CP, ADF, NDF, and DM concentration, total core numbers required to achieve an acceptable standard error (SE) were minimized by sampling once per bale. Bootstrap analysis of data from the most variable hay lot suggested that forage quality of any lot of 400-kg alfalfa hay bales should be adequately represented by 12 bales sampled once per bale.

ALFALFA hay is often bought and sold based on its forage quality. During 12 yr of quality-tested hay auctions in Minnesota, hay price has been associated with relative feed value (RFV), which is calculated from ADF and NDF concentration (Linn and Martin, 1989; Martin and Schriever, 1997). Hay is often sold by the lot, defined as a set of bales from a single location at a single cutting. Variability within a lot might lead to inaccurate forage quality results if sampling is not done properly. Studies have been conducted on sampling procedures for small (20 to 40 kg) rectangular bales (Martin et al., 1992), but there is little information about sampling larger (400 kg) rectangular bales. Currier et al. (1984) suggested that small rectangular bales tied with two twines had a distinct pattern of high leaf concentration on one side and high stem concentration on the other side of a diagonal drawn across the small face of the bale and extending the entire length. This variability was confirmed by Martin et al. (1992) who found significant differences among sampling locations on small bales. Sampling at least 20 bales from a lot of small bales was recommended to characterize the forage quality of the lot, but this recommendation was based on a single lot of alfalfa hay (Martin et al., 1992). No information

was available on sampling strategies for 400-kg bales; therefore, our objectives were to determine the within-bale variation in 400-kg rectangular bales and to determine the number and distribution of core samples required to represent CP, ADF, NDF and DM concentration in commercial lots of alfalfa hay.

## MATERIALS AND METHODS

### Within-Bale Variation

Three lots of Minnesota-grown pure alfalfa hay were sampled to study within-bale variation in forage quality. A lot consisted of about 50 stored bales from the same field and same cutting. Alfalfa maturity for each lot at harvest was first flower, but handling conditions after cutting were unknown. A rectangular bale measured approximately 2.4 by 0.9 by 0.9 m and weighed about 410 kg. Four bales were randomly selected from each lot of bales. Lot 1996 was made using a Case IH<sup>1</sup> model 8575 baler (Racine, WI) and sampled at the Minnesota Valley Alfalfa Producers processing plant in Priam, MN. Lots 1-1998 and 2-1998 were made using a New Holland<sup>1</sup> model D-2000 baler (New Holland, PA) and were sampled at the West Central Experiment Station at Morris, MN. The position of knots in the baling twine was used as a reference to keep front, back, top, bottom, left, and right faces consistent from bale to bale. Each of the six faces of a bale was divided into a three-by-three grid (Fig. 1). There were nine cells per side and 54 cells per bale. The order and location of coring was consistent from bale to bale so that cores numbered 1 to 54 on one bale could be compared with cores numbered 1 to 54 on another bale. The center of each grid cell was probed to a 35-cm depth using a 1.5-cm internal diameter Forageurs<sup>1</sup> hay probe (Lakeville, MN). About 15 g dry weight of hay was obtained with each probe.

The cores were ground to pass a 1-mm screen using a Foss-Tecator<sup>1</sup> mill (Eden Prairie, MN). Near infrared reflectance spectroscopy (NIRS) was used to analyze cores for Kjeldahl CP, DM, ADF, and NDF (Goering and Van Soest, 1970). Spectra for NIRS analysis were collected with a NIRSystem<sup>1</sup> model 6500 scanning monochromator (Silver Spring, MD) with a range of 400 to 2500 nm. Equations for NIRS were developed using the Infrasoft International NIRS 3 ver. 4.0 software program "Calibrate" with the modified partial least squares regression option (Shenk and Westerhaus, 1991). Details of the equations used to predict CP, ADF, NDF, and DM are shown in Table 1.

Forage quality data from bales in the three lots were subjected to a randomized complete block (RCB) analysis of variance (ANOVA) procedure (SAS Institute, 1996) with

C.C. Sheaffer, J. Grimsbo Jewett, J. Halgerson, Dep. of Agronomy & Plant Genetics, 1991 U. Buford Circle, Univ. of Minnesota, St. Paul 55108; N.P. Martin, USDA-ARS Dairy Forage Res. Center, Madison, WI; R.D. Moon, Dep. of Entomology, Univ. of Minnesota, St. Paul; and G.R. Cuomo, West Central Research and Outreach Center, Morris, MN. Minnesota Agric. Exp. Stn. Journal Series Paper No. 98-1-13-0076. Received 10 May 1999. \*Corresponding author (sheaf001@maroon.tc.umn.edu).

<sup>1</sup> Mention of a proprietary product does not constitute a recommendation or warranty of the product by the University of Minnesota and does not imply approval to the exclusion of other suitable products.

**Abbreviations:** ADF, acid detergent fiber; ANOVA, analysis of variance; CP, crude protein; DM, dry matter; NDF, neutral detergent fiber; NIRS, near infrared reflectance spectroscopy; RFV, relative feed value; RCB, randomized complete block design; SE, standard error; SEL, standard error of laboratory analysis.

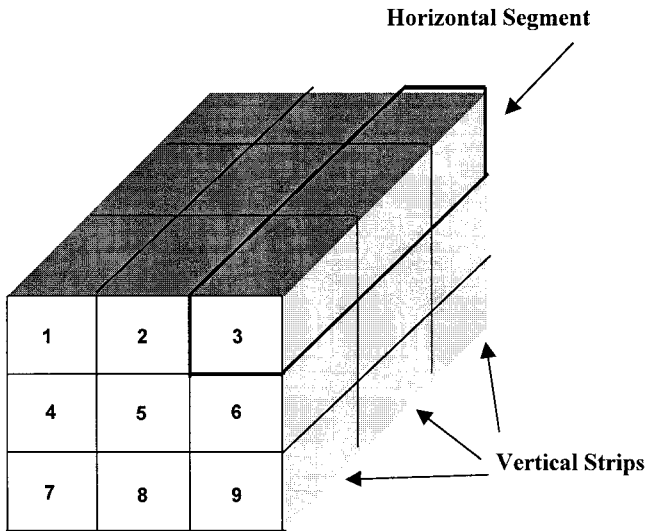


Fig. 1. Division of a rectangular bale into 54 cells, three vertical strips, or nine horizontal segments.

bales as blocks, faces as treatments, and the nine cores per side as sampling error. The UNIVARIATE procedure was used to verify that the forage quality data followed a normal distribution. A second ANOVA was done using a strip-plot design (Gomez and Gomez, 1984) with bales as blocks and faces and vertical thirds of the bale as strips (Fig. 1). Each face by vertical third intersection plot contained three cores. The front and back faces of the bales were excluded from the second analysis.

A third ANOVA used a RCB design with bales as blocks. Treatments were eight horizontal bale segments divided along the grid lines on the front bale face, down the length of the bale. The center segment, corresponding to Cell 5 on the front face of the bale (Fig. 1), was excluded because there were no data for the length of the bale. Average values for segments were used in the ANOVA because segments differed in number of cores. All ANOVAs were conducted within the three lots and also with the averaged data across lots.

Correlation analysis was done within lots and combined across the three lots to determine whether any part of a bale adequately represented the entire bale. The 54 cores per bale were divided into ten regions, with some overlap among regions. The first region included six cores, one from the center of each face. The second, third, and fourth regions were horizontal thirds of the bale: top, middle, and bottom, respectively. The fifth through seventh regions were vertical thirds along the length of the bale: front, middle, and back, respectively. The eighth, ninth, and tenth regions were vertical thirds across the width of the bale: left, middle, and right, respectively. Correlations were done between region means and the bale grand means.

Table 1. Calibration statistics for NIR equations to predict crude protein, acid detergent fiber, neutral detergent fiber, and dry matter percentage of medium and large rectangular bales.

Forage quality variable	<i>n</i>	Mean	SEC†	<i>R</i> <sup>2</sup>	SECV‡	1 - VR§
		$\text{g kg}^{-1}$				
Crude Protein	755	182	0.787	0.96	0.837	0.96
Acid Detergent Fiber	755	355	1.447	0.93	1.500	0.93
Neutral Detergent Fiber	441	491	2.247	0.95	2.383	0.95
Dry Matter	526	929	0.365	0.97	0.403	0.96

† Standard error of calibration.

‡ Standard error of cross validation.

§ One minus the variance ratio (similar interpretation as *R*<sup>2</sup> value).

Table 2. Average crude protein (CP), acid detergent fiber (ADF), neutral detergent fiber (NDF), and dry matter (DM) concentration in bale faces of Lots 1-1998, 2-1998, and 1996.

Bale face	CP	ADF	NDF	DM
	$\text{g kg}^{-1}$			
Top	215	328	416	858
Bottom	215	323	408	852
Left	216	319	401	854
Right	215	320	402	854
Front	213	316	398	849
Back	219	320	401	857
LSD <sub>0.05</sub> †	NS‡	8	10	NS

† Least significant difference ( $\alpha = 0.05$ ) for comparing bale faces.

‡ Not a significant difference.

### Within-Lot Variation

Three lots of Minnesota-grown pure alfalfa hay were used to study within-lot variation. A lot consisted of about 50 bales weighing 410 kg of pure, first flower alfalfa hay from a common source as described above. One lot of 55 bales was made using a Case IH model 8575 baler. It was sampled at the Minnesota Valley Alfalfa Producers processing plant and designated Lot 1997. Fourteen bales were sampled from each of the two lots that also were used for the within-bale sampling, Lots 1-1998 and 2-1998.

The Forageurs hay probe was used for coring. One core was taken from the center of each bale face for a total of six cores per bale. The cores were analyzed via NIRS for CP, ADF, NDF, and DM, as described above. Data normality was checked using the UNIVARIATE procedure of SAS as described above, and normality assumptions were not violated. Total variance in CP, ADF, NDF, and DM was calculated for each lot. A RCB analysis of variance was conducted for each lot with bales as blocks and cores as treatments. The mean squares from the analysis of variance were entered into a two-stage sampling formula developed by Snedecor and Cochran (1989) to calculate estimates of the SE associated with sampling from various combinations of number of bales and cores per bale. A sampling scheme was considered adequate for determining a forage quality predictor if the calculated SE was equal to or less than the standard error of laboratory analysis (SEL). The SEL is for CP 6  $\text{g kg}^{-1}$ , for ADF 12  $\text{g kg}^{-1}$ , for NDF 15  $\text{g kg}^{-1}$ , and for DM 3  $\text{g kg}^{-1}$  (Marten et al., 1989).

Data from Lot 1997 were used to conduct a computer-simulated sampling procedure similar to bootstrap analysis (Efron and Tibshirani, 1996). The bootstrap analysis involved repeated sampling with replacement from all the cores sampled from a lot. We modified a SAS (SAS Institute, 1991) program for random sampling with replacement to randomly select 1000 sets of a given sample size ( $n = 4, 6, 8, 10, 12, \text{ up to } 20$  cores per set).

Each core from a lot had an equal chance of being drawn

**Table 3. Average crude protein (CP), acid detergent fiber (ADF), neutral detergent fiber (NDF), and dry matter (DM) concentration in front, back, and middle thirds of bales from Lots 1-1998, 2-1998, and 1996.**

Bale third	CP	ADF	NDF	DM
	g kg <sup>-1</sup>			
Front	214	321	405	852
Middle	215	316	401	859
Back	216	330	414	852
LSD <sub>0.05</sub> †	NS‡	NS	NS	5

† Least significant difference ( $\alpha = 0.05$ ) for comparing bale faces.

‡ Not a significant difference.

each time a selection was made. The variance was calculated for each set and an average variance for a given sample size was calculated from the 1000 set variances. The SE associated with a sample size was then calculated from the average of 1000 set variances. This bootstrap SE was compared with the SEL (Marten et al., 1989), and a sample size was declared adequate if its bootstrap SE was equal to or less than the SEL.

## RESULTS AND DISCUSSION

### Within-Bale Variation

#### Bale Faces

Bale faces did not differ in CP and DM concentration either within the three lots or when values were averaged for the lots. On average, bale tops had higher ADF and NDF concentrations than other bale faces (Table 2), but analysis within lots showed that only Lot 1-1998 had a higher NDF concentration in bale tops (data not shown) than the other faces. This difference did not occur in Lots 2-1998 or 1996, or for ADF concentration in any lot.

#### Vertical Thirds

The front, back, and middle thirds of bales were similar for CP, ADF, or NDF concentration when values were averaged for the three lots (Table 3) or when analysis was done within lots (data not shown). The middle third of bales on average had slightly higher DM concentration than the front or back thirds, but analysis within lots showed that the middle third of bales was not consistently higher in DM concentration.

**Table 4. Average crude protein (CP), acid detergent fiber (ADF), neutral detergent fiber (NDF), and dry matter (DM) concentration in eight horizontal segments running the length of bales in Lots 1-1998, 2-1998, and 1996.**

Segment	CP	ADF	NDF	DM
	g kg <sup>-1</sup>			
1	214	327	414	857
2	218	319	405	856
3	215	323	408	860
4	215	318	398	854
6	217	314	394	852
7	215	323	405	852
8	216	319	403	851
9	214	322	404	852
LSD <sub>0.05</sub> †	NS‡	NS	11	NS

† Least significant difference ( $\alpha = 0.05$ ) for comparing bale faces.

‡ Not a significant difference.

**Table 5. Pearson's correlation coefficients between bale regions and bale grand means† for crude protein (CP), acid detergent fiber (ADF), neutral detergent fiber (NDF), and dry matter (DM) concentration.**

Bale region	CP	ADF	NDF	DM
Face centers	0.95**	0.98**	0.98**	0.95**
Top third	0.96**	0.99**	0.99**	0.99**
Middle third	0.98**	0.99**	0.99**	0.98**
Bottom third	0.97**	0.99**	0.99**	0.96**
Front third	0.95**	0.97**	0.97**	0.99**
Middle third	0.56	0.97**	0.98**	0.99**
Back third	0.69*	0.98**	0.98**	0.99**
Left third	0.95**	0.98**	0.98**	0.98**
Middle third	0.89**	0.99**	0.99**	0.99**
Right third	0.96**	0.99**	0.98**	0.99**

\*\* Significant at  $\alpha = 0.01$ .

† Average data from four bales in each of three alfalfa hay lots.

### Horizontal Segments

When values were averaged for the three lots, Segment 6 had lower ADF and NDF concentrations than the other cells (Table 4). Analysis within lots showed that Lot 1-1998 also had lower ADF and NDF concentration in Segment 6 than in other segments, but Lots 2-1998 and 1996 did not show this difference (data not shown). Crude protein and DM concentration were similar among segments both within lots and averaged across the three lots.

Our results from intensive sampling of individual bales in three alfalfa hay lots suggest that on average the bale ends may have lower dry matter concentrations than the middle. On average, bale tops may have higher fiber concentration than other bale faces, but no pattern of forage quality or dry matter concentration was consistent across all hay lots. Correlation analysis revealed that, on average, values for any portion of a bale were highly correlated with the bale grand means for CP, ADF, NDF, and DM concentration (Table 5). Thus, cores may be taken from any location on a 400-kg bale for forage quality analysis.

### Within-Lot Variation

Calculations of variances for Lots 1-1998, 2-1998, and 1997 revealed that Lot 1997 was more variable for CP, ADF, and NDF than the other two lots (Table 6). The most variable data will produce the most broadly applicable results, so we report only the results from Lot 1997 for CP, ADF, and NDF. Lots 1-1998 and 2-1998 had more variability for DM concentration than Lot 1997, so we include bootstrap analysis results from these two lots for DM.

Standard error estimates calculated from a formula for two-stage sampling (Snedecor and Cochran, 1989)

**Table 6. Variances for crude protein (CP), acid detergent fiber (ADF), neutral detergent fiber (NDF), and dry matter (DM) concentration in Lots 1-1998, 2-1998, and 1997.**

Lot†	CP	ADF	NDF	DM
1-1998	3.61	8.69	12.09	4.12
2-1998	0.90	3.60	7.27	9.77
1997	4.45	9.08	15.70	0.67

† There were 14 bales sampled from Lot 1-1998 and Lot 2-1998; all 55 bales of Lot 1997 were sampled.

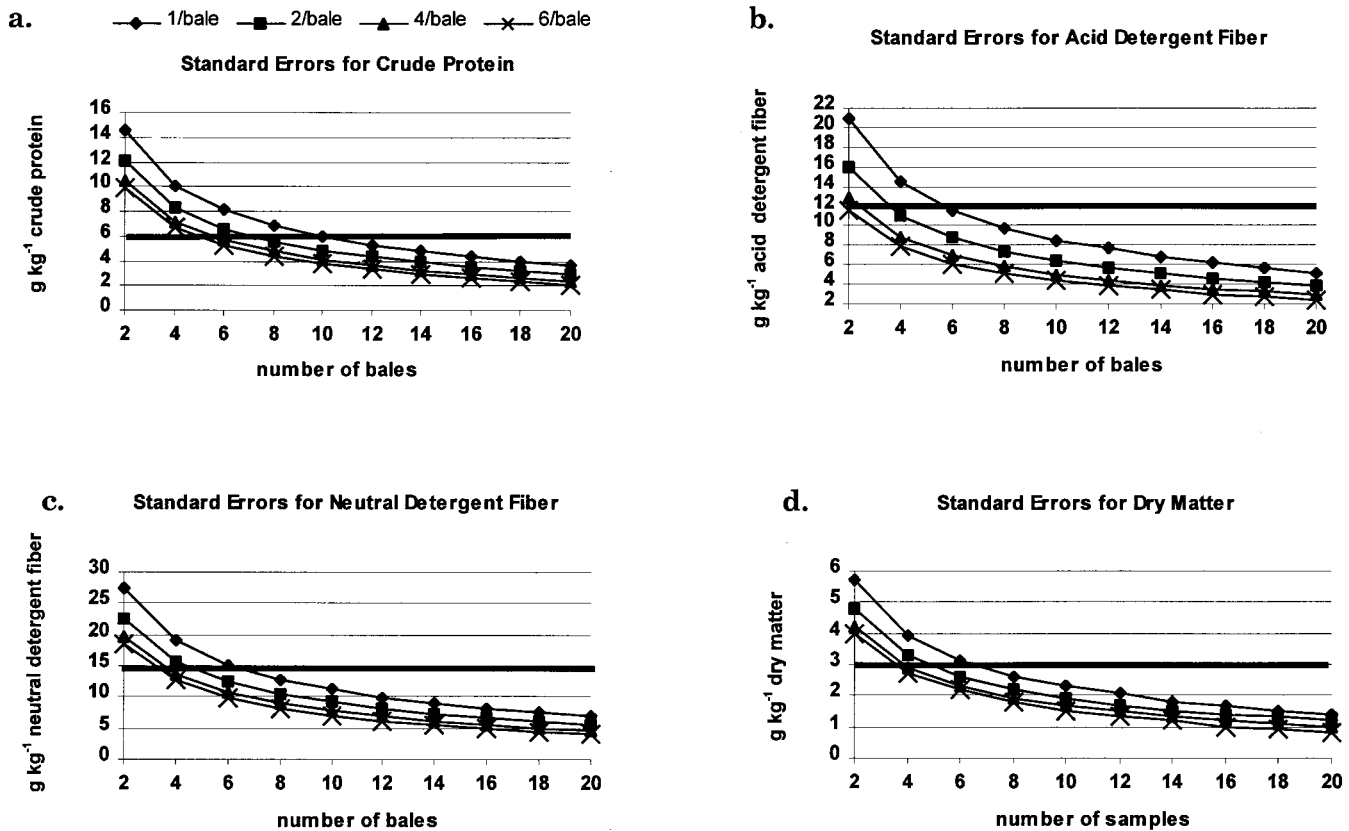


Fig. 2. Standard errors associated with sampling one, two, four, or six cores from 2 to 20 alfalfa bales for: a) crude protein, standard error of laboratory analysis (SEL) = 6 g kg<sup>-1</sup>; b) acid detergent fiber, SEL = 12 g kg<sup>-1</sup>; c) neutral detergent fiber, SEL = 15 g kg<sup>-1</sup>; and d) dry matter, SEL = 3.0 g kg<sup>-1</sup>.

showed that probing once per bale would minimize the total number of cores necessary to achieve acceptable SE. In Lot 1997, to achieve a SE for CP equal to the 6 g kg<sup>-1</sup> SEL, 10 bales would be needed if bales were probed once; but if each bale were probed twice, eight bales would be needed for a total of 16 cores (Fig. 2a). Probing four times per bale would require six bales for a total of 24 cores, and probing six times per bale would require five bales for a total of 30 cores. Similar relationships between total number of samples and cores per

bale are shown in Fig. 2b to Fig. 2d for ADF, NDF, and DM, respectively.

We used the bootstrap analysis to broaden the applicability of our results from the two-stage calculations of SE. The bootstrap analysis was a systematic way to simulate the variability of additional hay lots without actually taking additional samples. Thus, the number of required samples predicted by the bootstrap analysis should give accurate results over a broader range of hay lots than would the smaller numbers predicted by the

Table 7. Standard errors for crude protein (CP), acid detergent fiber (ADF), neutral detergent fiber (NDF), and dry matter (DM) concentration for variable bale numbers calculated from analyses of variance (ANOVAs)<sup>†</sup> and from bootstrap analyses<sup>‡</sup> on Lot 1997 (N = 55 bales).

Number of bales	Standard error of sampling							
	CP		ADF		NDF		DM	
	ANOVA	Bootstrap	ANOVA	Bootstrap	ANOVA	Bootstrap	ANOVA	Bootstrap
	g kg <sup>-1</sup> dry matter							
4	10	11	14	15	19	20	4	4
6	8	9	12	12	15	16	3	3
8	7	7	10	11	13	14	3	3
10	6	7	8	10	11	13	2	3
12	5	6	8	9	10	12	2	2
14	5	6	7	8	9	11	2	2
16	4	5	6	8	8	10	2	2
18	4	5	6	7	7	9	2	2
20	4	5	5	7	7	9	1	2
SEL <sup>§</sup>	6		12		15		3	

<sup>†</sup> Standard errors were calculated from ANOVA mean squares by the two-stage sampling formula of Snedecor and Cochran (1989).

<sup>‡</sup> Standard errors calculated from the bootstrap analysis were the square root of (mean of 1000 set variances divided by the number of samples per set).

<sup>§</sup> SEL = standard error of laboratory analysis. Standard errors associated with sample-to-sample variation should be equal to or less than the SEL.

ANOVA and two-stage calculation. Standard errors calculated from the bootstrap analysis using Lot 1997 indicate that with one core per bale, 12 bales would be needed to achieve the acceptable  $6 \text{ g kg}^{-1}$  error for CP. The  $12 \text{ g kg}^{-1}$  acceptable error for ADF could be reached by probing six bales. For NDF, seven bales would be required to reach the  $15 \text{ g kg}^{-1}$  acceptable error (Table 7).

The acceptable error of  $3.0 \text{ g kg}^{-1}$  for DM could be reached with six bales, according to bootstrap analysis of 1997 data. When the more-variable DM data from Lots 1-1998 and 2-1998 were used in the bootstrap analysis, the calculations showed that sampling six bales would give a SEL for DM of 8 to  $13 \text{ g kg}^{-1}$ , and at least 40 bales per lot would be needed to reach the  $3.0 \text{ g kg}^{-1}$  SEL for DM. These results suggest that in the case of DM it may be more cost-effective to accept an error larger than the SEL, rather than sample most of the bales in a lot.

### SUMMARY

On average, tops of 400-kg bales may have higher fiber concentration than other bale faces; and the middle section of bales may have higher DM concentration than the front or back ends. Averaged across lots, forage quality in any portion of a bale was highly correlated with the bale grand mean. Hay lots differed in their amount of variability for CP, ADF, NDF, and DM. Probing once per bale was the most efficient way to obtain the required number of cores for accurate forage quality determination in a hay lot. Bootstrap analysis of the most variable of three hay lots indicated that CP, ADF, and NDF in any lot of 400-kg bales should be adequately represented by 12 bales when sampling is done once per bale. Dry matter may be less-well characterized by only 12 bales, but additional sampling is costly. We recommend sampling 12 randomly chosen bales once per bale from any location on the bale for forage quality characterization of an alfalfa hay lot.

### ACKNOWLEDGMENTS

The Minnesota Valley Alfalfa Producers and the University of Minnesota are conducting a comprehensive cost-shared research and development program in support of the Minnesota Agri-Power Project. This project is supported by the Minnesota Valley Alfalfa Producers, Minnesota farmers, U.S. Departments of Energy and Agriculture, University of Minnesota, Oak Ridge National Laboratory, National Renewable Energy Laboratory, Minnesota Department of Agriculture, Legislative Commission on Minnesota Resources, Agriculture Utilization Research Institute, and Minnesota Forage and Grassland Council.

### REFERENCES

- Currier, C.G., B.A. Malton, J.S. Arledge, and V.L. Marble. 1984. Sampling alfalfa hay. p. 8-16. *In Proc. National Alfalfa Hay Quality Testing Workshop*, Chicago, IL. 22-23 March 1984. Univ. of California Cooperative Extension Service, Berkeley.
- Efron, B., and R. Tibshirani. 1996. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Sci.* 1:54-77.
- Goering, H.K., and P.J. Van Soest. 1970. Forage fiber analysis: Apparatus, reagents, procedures, and some applications. *USDA Agric. Handb.* 379. U.S. Gov. Print. Office, Washington, DC.
- Gomez, K.A., and A.A. Gomez. 1984. *Statistical procedures for agricultural research*. 2nd ed. John Wiley & Sons, New York.
- Linn, J.G., and N.P. Martin. 1989. Forage quality tests and interpretation. *Bull. AG-FO-2637*. University of Minnesota Extension Service, St. Paul.
- Marten, G.C., J.S. Schenk, and F.E. Barton II (ed.). 1989. Near infrared reflectance spectroscopy (NIRS): Analysis of forage quality. *Agric. Handb.* 643. USDA-ARS. Washington, DC.
- Martin, N.P., R.L. Ellingboe, P.R. Peterson, M. Matteson, R.E. Stucker, and J.G. Linn. 1992. Sampling evaluation of alfalfa hay. *Proc. Am. Forage Grassl. Conf.* 1:66-71.
- Martin, N.P., and D.L. Schriever. 1997. Quality-tested hay auction, summary, 1996-1997. *Minnesota Forage Update* 22:7. Minnesota Forage and Grassl. Council, St. Paul, MN.
- SAS Institute. 1991. *SAS Language and Procedures: Usage 2*, Version 6, First Edition. SAS Inst., Cary, NC.
- SAS Institute. 1996. *SAS System for Windows*. Release 6.12. SAS Inst., Cary, N.C.
- Shenk, J.S., and M.O. Westerhaus. 1991. Population structuring of near infrared spectra and modified partial least squares regression. *Crop Sci.* 31:1548-1555.
- Snedecor, G.W., and W.G. Cochran. 1989. *Statistical Methods*. 8th ed. Iowa State University Press, Ames.