VARIATION IN FEED ANALYSIS AND APPROACHES TO MINIMIZE ITS MAGNITUDE AND IMPACT

David R. Mertens
USDA-Agricultural Research Service
U.S. Dairy Forage Research Center
Madison, WI 53706

Abstract

Accurate and precise analytical results are the cornerstones for economic evaluation of feeds, formulation of optimal rations, and diagnosing dietary problems associated with the health and productivity of dairy cows. Variation in analytical results occurs and can be troublesome. Several principles of analytical variation provide the basis for understanding variation and minimizing its detrimental effects on nutritional decision-making. Principle 1 – All analytical results are estimates of feed composition and nutritional value. Analyses are done on samples of the feed and results from a sample only provide an estimate of the true average composition of the feed. Principle 2 – Variation is natural and unavoidable. We may be able to control or minimize variation, but we can never eliminate the random variation in sampling or analysis. Principle 3 – Variation has two completely independent components: accuracy and precision. Accuracy measures the closeness of the result to the true value; whereas, precision measures the variation in repeated results. Precise repeatability of results tells nothing about the accuracy of the result. Principle 4 – Replication is key to minimizing the effects of variation. The average of 'n' replicated observations reduces variation in proportion to 1/n. Principle 5 – A representative sample is the crucial starting point for analytical results. Results are no better than the sample that is analyzed. The sample should consist of numerous, randomly selected sub-samples. Analytical variation can be reduced when the sources of imprecision and inaccuracy are identifed and minimized. Imprecision arises from heterogeneity of the sample, complexity of the procedure, and the carelessness of the technician. Inaccuracy occurs because methods differ among laboratories due to modifications by the method developer and in-house modifications by laboratories. When results are adjusted to a dry matter basis, inaccuracies in dry matter analysis add variation. *In vitro* results are variable because methods are biological assays with no standard method to define time of fermentation, inoculum source and preparation, grind size of the feed, media buffers, fermentation vessel, or anaerobic technique. Sample variation is often as large or larger than analytical variation and is a major source of variation in analytical

results. To minimize the effects of analytical variation on dairy rations, replicate samples and analyses, and increase the number of ingredients in the ration.

Introduction

Variation in feed analysis is important because analytical results are used to establish value and formulate rations. Both of these outcomes affect the profitability of forage producers and dairy farmers; therefore it is important that true error in analysis is eliminated and unnecessary variability in analysis is minimized. Unfortunately there can be many intentional and unintentional changes in the ration between formulation and feeding. There are at least four rations on every farm: the one formulated by the nutritionist, the one given to the feeder by the dairy farmer, the one mixed by the feeder, and finally the one eaten by the cow. The ration eaten is the most crucial to the health and productivity of the cow, but the accuracy of the ration that is formulated sets the basis for all the others and is critically dependent on laboratory analyses.

Accurate chemical and biological evaluation of feeds is the foundation for setting prices and designing rations. The objectives of this discussion will be to: (1) define analytical variation, (2) estimate the expected variation for routine analyses, (3) discuss the sources of variation in analytical results and (4) propose approaches for minimizing the effects of analytical variation.

Principles of Analytical Variation

1. <u>All analytical results are estimates of feed composition and nutritional value.</u> The only way to truly determine the nutritional composition of a lot of feed would be to analyze every pound of it. Because all of the feed would be used for analysis this would defeat the purpose of analysis. Instead, we analyze <u>samples</u> of the feed, both the composite sample taken from the lot and submitted to the laboratory and the test-sample analyzed by the lab. Given that the sample submitted or the test-sample analyzed are often less than $1/1{,}000{,}000^{th}$ of the mass of feed in the lot, it seems highly unlikely that the analytical result from any one sample will exactly match any other sample taken. No matter how hard we try, no two samples of a feed will be exactly alike. This brings us to the next principle of variation.

2. <u>Variation is natural and unavoidable</u>. We may be able to control or minimize variation, but we can never eliminate it. The hard feelings and accusations that have occurred because one lab reported that a sample of hay was 42% NDF and other lab reported that a different sample of the same hay was 40% NDF

is beyond comprehension and demonstrates the lack of understanding that variation is real and unavoidable. It is not uncommon in our laboratory for duplicate samples from the same finely ground and thoroughly mixed sample to differ by 1% NDF. This variation is not due to variation in analytical technique from samples analyzed one after the other, but is due to the fact that no two test samples are exactly alike.

It is amusing and somewhat ludicrous to expect no difference in results between two analyses or to expect that minor differences are real. Using guidelines to report prices are one thing, but using guidelines to set prices is quite another. Using discrete, arbitrary cutoffs to define the pricing of hay is a simple way of categorizing observations in a concise table. However, using discrete cutoffs for pricing hay (e.g. 36.0% NDF is dairy hay, but 36.1% is not) is ridiculous because it assumes that we can actually measure real differences in NDF to within 0.1% and that the sample submitted and analyzed exactly represents the NDF in the lot of hay.

3. Variation has two completely independent components: accuracy and precision. Accuracy measures the closeness of the result to the true value; whereas, precision measures the variation in repeated results. Because they are independent components of variation, every set of results has both attributes (Figure 1). The difference between the result and the true value is often called a bias, and if the bias is consistent across samples it is called a systematic bias. Thus, accuracy measures the systematic bias of analytical results. The variation between repeated or replicated results on the same sample using the same technique is measured as precision. Precision variation is primarily random variation (or bias) that we cannot eliminate or control because we cannot explain or predict its source. However, some imprecision is due to carelessness in replicating procedures during a method, and this source of imprecision can and should be minimized.

   Earlier, I stated that variation cannot be eliminated and only minimized, but after we divide total variation into accuracy (systematic bias) and precision (random bias) we discover that the systematic bias portion of total variation can be eliminated when the source of inaccuracy is determined. However, the random bias portion of total variation, which is related to precision, can never be eliminated, only controlled or minimized.

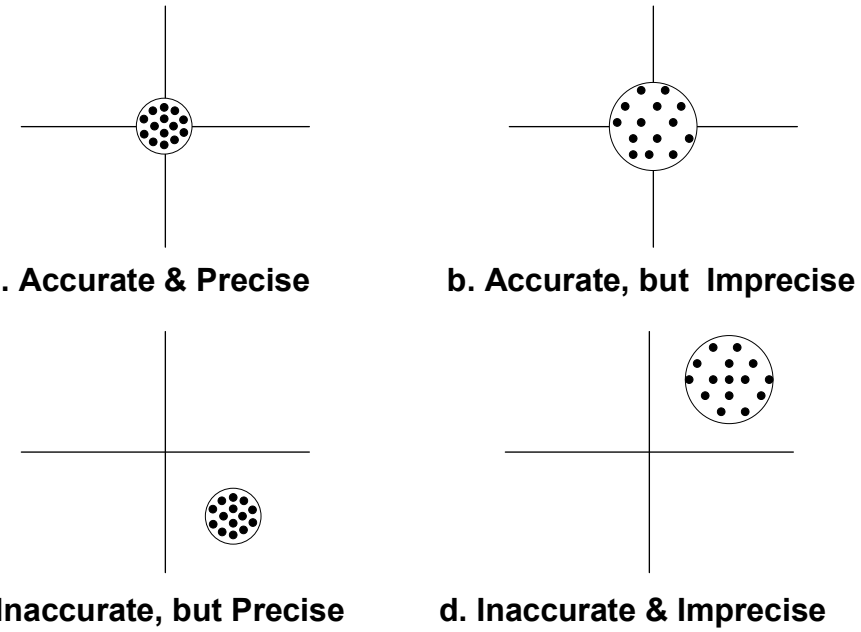

Figure 1. Illustration of the components of variation related to accuracy (systematic bias) and precision (random bias).

As shown in Figure 1, methods can be (in order of preference) accurate and precise, accurate and imprecise, inaccurate and precise, or worst of all inaccurate and imprecise. For a single analysis of a single sample, the error in the reported result can be equally different from the true value for the sample whether the method is 'accurate and imprecise' or inaccurate and precise.' However, 'accurate and imprecise' methods are better than 'inaccurate and precise' ones because imprecision can be overcome by repeated analysis of the material, but inaccuracy can only be corrected by improving the method or deriving an average bias correction. For imprecise methods, the addition of each replicated result moves the average of all results closer to the true or target value for that lot of feed.

Although inaccuracy can only be corrected by improving the method, it too can be partially overcome by replication. However, the replication has to be across feeds to calculate an average bias correction. This is a less satisfactory adjustment for accuracy than replication is for precision, because the systematic bias correction is general and across all samples, whereas replication is specific because it is for the same sample. Systematic bias correction is analogous to "Kentucky windage" correction for long rifles with

nonadjustable sights in which the shooter aims to a point different from the center of the target in order to hit it.

Inaccuracy should be the focus of our efforts to minimize variation. For most analytical results, inaccuracy is associated with differences in methods among laboratories (When Lab A and Lab B consistently differ in results) or with differences in the adaptability of methods for materials analyzed within laboratory (alfalfa is measured more accurately than corn silage). The sources of inaccuracy in analytical result will be discussed later. No matter how hard we try, no two samples we take will be exactly alike because random variation cannot be eliminated. Our only recourse is to minimize the effects of random variation by replication.

4. Replication is the key to minimizing the effects of variation. Replication is needed at all stages of the analytical process: replication in the collection of site-samples to generate a representative composite sample; replication in composite samples to more accurately reflect the true value for the lot of feed; replication in analysis to minimize the effects of analytical variation; and even replication in the number of feed ingredients in a ration to minimize nutrient variation in cow diets.

   It is a statistical axiom that replication is the key to the measurement of variation, minimizing the relative influences of random variation, and using variation to determine if differences between results or treatments are real (statistical inference). In fact, the relationship between the number of replications in a composite sample and the variation of the average or mean of the sample is known. The expected variance (σ_X^2) of the averages of samples with n replications is a function of the standard deviation (σ) of the population and the replication (n) used to determine the sample average: $\sigma_X^2 = \sigma^2/n$. We often refer to the 'standard error' of the mean to distinguish it from the 'standard deviation' of the population. However, the standard error (SE) refers to variation and not 'error' as the result of mistakes or incorrect technique. As the number of observations included in the mean goes up the SE of the mean goes down by $1/\sqrt{n}$ (Table 1). One result of the expected variance equation is as we average of more and more samples together the sample mean will become closer and closer to the true population mean because the variation among samples goes to zero.

   Notice in Table 1 that the first few replicates provide the most reduction in SE. It takes 4 samples to reduce the SE by half, but takes an additional 12 samples (16 in total) to half the SE again. Because the test-sample is finely

Table 1. Standard error of the averages of composite samples with n observations obtained from populations of results with various standard deviations.

Number of	Standard deviation among individual replications			
Observations (n)	1.00	2.00	3.00	4.00
1	1.00	2.00	3.00	4.00
2	0.71	1.41	2.12	2.83
4	0.50	1.00	1.50	2.00
8	0.35	0.71	1.06	1.41
16	0.25	0.50	0.75	1.00
32	0.18	0.35	0.53	0.71
64	0.13	0.25	0.38	0.50
128	0.09	0.18	0.27	0.35

ground and mixed, only a few replications (1 to 4) are needed to provide information about analytical variation in a feed. Duplicate samples provide an average estimate of the true average and, when the difference between duplicates is greater than four times the expected SD, they can indicate that one result is suspect. Triplicate or quadruplicate analyses are needed to identify and remove a suspected result.

More replication allows us to detect smaller differences between treatments because it reduces the SE among treatment means, and more replication gets us closer to the true composition or biological value of a lot of feed. Table 1 demonstrates that if the population is more heterogeneous (its standard deviation is larger), it takes more replications to obtain the same SE among composite samples. For example, it takes a composite with 16 site-samples when the population has a standard deviation of 4.0 to have the same SE as a composite with 4 site-samples when the population is less variable with a standard deviation of 2.0. This explains why we want more replication in the factor that is most variable. If the variability among *in vitro* results is greater across runs than within runs, we need more replications across runs to estimate the true biological availability of the feed. If there is more variation among labs than within labs, we need replication across labs rather than within labs. This last conclusion seems counter-intuitive to conventional wisdom, which suggests that we should pick a lab and use them exclusively.

Picking one laboratory for analysis is a useful approach if only relative differences are important or if experience allows the adjustment of results to make them acceptable for use. In the first instance, we are not concerned about the accuracy of the laboratory's results and are using the difference in

results among feeds to make decisions about selection or use in rations. In the latter case, we are, in effect, making a bias correction to the results or system for using them to make the results and system compatible. Both of these approaches can be successful in the right circumstance. However, they are inadequate for estimating the true nutritive value of a feed.

If we want to know the true analytical result for a lot of feed, it is better to replicate composite samples and laboratories. The average of these results will be our best estimate of the true analysis of the lot of hay. Too often there are contentious arguments because the results of a seller of a lot of hay who submitted a sample to one laboratory do not agree exactly with the results obtained by the buyer who submitted another sample to a different laboratory (later I will present the magnitude of analytical differences to expect). Everyone would be better served if they agreed that their average result was the best estimate of the value of the lot of feed, instead of an acrimonious debate about whose laboratory or sample was correct. Given that variation cannot be eliminated it is highly unlikely that the results would agree exactly. Given that replication is our main recourse to reduce the effects of variation, the average is always our best estimate of the true value.

5. A representative sample is the key starting point for analytical results. We should try to do everything to make sure that the sample we submit for analysis is representative. To be representative, we want to randomly select our sampling sites so it is an accurate reflection of the lot of feed (we don't select the best 10 bales or the worst 5 spots in bales or on the face of silages). In addition to randomness, we want the sampling sites to be numerous.

   An example easily illustrates how the number of site-samples in our composite sample affects our ability to make decisions about the population. If we put 20 white and 20 black marbles in a pail and ask a disinterested person to blindly (randomly) pick one marble out of the pail, how would they describe the marbles in the pail? They would say it was either black OR white marbles, even though we would know it was a mixture. What happens if they take a bigger sample size of 2 or 4 or 20 marbles? Without a doubt, the bigger their composite sample size the better their chance of saying the pail contained a 50:50 mixture of black and white marbles. Imagine what happens when we use five different colored marbles (we describe this as a heterogeneous mixture). Even more samples would be needed to describe the population of marbles in the pail.

Forages are very heterogeneous, which raises the question, "When do we stop taking more site-samples to make a composite for heterogeneous forages?" The answer depends on the heterogeneity of the forage. For a uniform lot of hay harvested from the same field on the same day only 8-12 site-samples may be needed to obtain a representative sample. For variable lots of feed, 16-20 site-samples may be needed to obtain a similarly representative sample.

One of the factors affecting heterogeneity of forages is the variation in crop growth and harvesting conditions (e.g., different fields, different moisture at harvest, etc.). Another factor is the diversity in plant components (e.g., leaves versus stems, grain versus forage). Particle size of the material also plays a role in heterogeneity, which is crucial in trying to split samples for analysis. Coarse materials are difficult to samples because large pieces easily segregate from small ones and mixing actually sifts the smaller particles to the bottom of the sample.

Factors Affecting Analytical Variation

Variation due to imprecision. Imprecision in analytical results arises from three primary sources: heterogeneity of the sample, complexity of the procedure, and carelessness of the technician. The AOAC International defines imprecision as repeatability within a laboratory using a defined method. Even when samples are finely ground and thoroughly mixed, no two test-samples are exactly alike. The more heterogeneous the sample and the more coarsely it is ground the lower the precision or repeatability of the analysis. Although it is impossible to prove that much of the repeatability of analysis is related to test-sample differences when the analysis destroys the sample (making impossible to replicate the analysis on the same test-sample), non-destructive methods such as near infrared reflectance spectroscopy supports the contention that much of the variation in repeatability or precision is related to test-sample differences. If separate test-samples are packed into cells, NIRS results indicate that repeated scans of the same test-sample give nearly identical results, but scans of a second test-sample give different results with variation similar to good chemical methods. Variation due to differences in test-samples is random and cannot be eliminated.

Complex methods with multiple steps offer greater opportunity for variation in technique within and among technicians. When analysis requires complex methods, the additional variation in repeatability cannot be removed. However, when techniques within a method are carelessly and inconsistently followed within and among technicians, variation in repeatability is increased

unnecessarily. This latter source of variation can be detected by replicated analysis of samples (preferably an in-house or external standard or reference sample) and eliminated by training and suitable quality assurance programs within the laboratory.

Variation due to inaccuracy. Systematic bias among laboratories exists. Inaccuracy or systematic bias is detected by submitting split samples to laboratories and comparing the differences among laboratories to those expected for the analytical method. To adequately detect differences among laboratories requires comparisons using multiple samples. Splitting a single sample between two labs confounds sample and laboratory, i.e., one cannot determine if the laboratories or samples were different. However, if more than one sample is split, it is highly unlikely that the same lab would get the high or low samples in all splits. The National Forage Testing Association uses six finely ground and thoroughly mixed samples per year to detect difference among laboratories in dry matter, crude protein, neutral detergent fiber (NDF), and acid detergent fiber. Given that systematic biases should be eliminated, it is important to identify their possible sources. Because NDF and *in vitro* NDF digestibility (IVNDFD) are two of the most useful and most variable analyses used in current forage evaluation and ration formulation, I will focus on the factors that affect their variability.

Probably the greatest source of variation in both NDF and IVNDFD measurements is differences among methods across laboratories. The NDF method has evolved from the original method developed in the late 1960s that used sodium sulfite to reduce protein contamination from fiber (Van Soest and Wine, 1967; Goering and Van Soest, 1970), to the neutral detergent residue (NDR) method of the 1980s that removed sodium sulfite but added heat-stable amylase to remove starch contamination from fiber (Robertson and Van Soest, 1981; Van Soest et al, 1991), and finally to the amylase-treated NDF (aNDF) that includes both sodium sulfite and amylase to minimize contamination of insoluble fiber (Mertens, 2002). Unfortunately, the results of all three methods are often called NDF even though there are demonstrative differences in the values of NDF, NDR and aNDF (Table 2). Although the difference in NDF among methods is small for forages, the differences in NDF among methods can be large for heated byproduct feeds (Table 2). Therefore, it is important to know which "NDF" was analyzed and reported, and to understand that most of the discrepancies among laboratories in NDF results may be due to differences in methods. This variability in methods has led to the misconception that NDF is difficult to measure when its variability is comparable to other chemical assays.

Table 2. Results obtained using various methods to measure NDF (% of DM).

Feed description	NDF[a]	NDR[b]	aNDF[c]	ANDF/NDR
Wheat straw[d]	83.9	86.0	82.8	96.3
Timothy[d]	67.2	68.0	65.1	95.7
Corn silage[d]	55.9	55.0	52.6	95.6
Alfalfa hay[d]	47.2	50.4	46.3	91.9
Alfalfa silage[e]		43.6	42.2	96.8
Citrus pulp[e]		21.3	20.2	94.8
Corn grain[e]		11.4	10.1	88.6
Brewer's grains[e]		52.3	40.9	78.2
Distiller's grains[e]		38.6	27.9	72.3
Soybean meal[e]		18.5	12.4	67.0

[a] Neutral detergent fiber - with sulfite, but no amylase (Van Soest and Wine, 1967).
[b] Neutral detergent residue - no sulfite, but with amylase (Robertson and Van Soest, 1980).
[c] Amylase-treated neutral detergent fiber - with sulfite and amylase (Mertens, 2002).
[d] R.B. Robertson, personal communication (1988).
[e] (Hintz et al., 1996).

In addition to differences in method that create variability in NDF among laboratories, there are also variations related to fiber extraction method (refluxing beakers, refluxing in crucible, or refluxing in bags inside of pressurized chambers). In-house changes to methods are made to accommodate unique characteristics of personnel, extraction equipment, fiber collection, and weighing techniques that can alter results. Too often these in-house modifications are not thoroughly tested to determine their general applicability across all types of feeds. Proficiency certification by the National Forage Testing Association is an excellent way to validate in-house modifications because certification is based on comparisons of a laboratory's results to those of laboratories using the reference method for aNDF.

The greatest source of variation among *in vitro* results is related to fermentation time. Digestion curves rapidly increase for the first 24 to 36 h and then plateau. This suggests that variability will be highest at early fermentation times when digestion is rapidly increasing (18-30 h). Typically a 48 h fermentation time is selected for measuring fiber digestion in the expectation that this result will coincide with *in vivo* digestibility at maintenance levels of intake. Sometimes, fermentation times of 30 h are used to represent digestibility in dairy cows, and fermentation times of ≥72 h are used to assess the maximum extent of digestion.

Other factors affect *in vitro* disappearance, such as animal and diet differences among inoculum donors, daily differences in inoculum, and grinding particle size of the substrate. In addition, *in vitro* results are affected by the type and efficacy of the buffer, supplementation of the media with proteins and trace minerals, ratio of substrate to buffer to inoculum, preparation technique for the inoculum, type of fermentation vessel, and anaerobicity of the inoculation and fermentation environment (Grant and Mertens, 1992). Typically, the variation among *in vitro* results is greater among laboratories, and is greater among-runs than within-run in a laboratory.

A source of variation in analytical results often overlooked is the adjustment of results to a DM basis. It is possible to create an apparent difference in results because the DM analysis differs between laboratories. For example, assume that laboratories A and B had similar CP and NDF analyses on an as-is basis, but they obtained different results for DM determination. When the CP and NDF data are adjusted for different DM determinations an apparent difference in the CP and NDF results is created (Table 3). The effect is greater for NDF than CP because the magnitude of the NDF value is larger. Analytical results should always be compared on an as-is or as-received basis so the difference in the analyte, but not DM, is evaluated.

Table 3. Effect of dry matter (DM) adjustment on comparisons of laboratory results for crude protein (CP) and neutral detergent fiber (NDF).

Nutrient	Labs A&B	Lab A 88% DM	Lab B 92% DM	Lab diff
	As-Is %	DM basis	DM basis	DM basis
CP	20	22.7	21.7	1.0
NDF	50	56.8	54.3	2.5

Expected Variation in Analytical Results

Horwitz (1982) evaluated hundreds of collaborative studies used to evaluate AOAC International official methods and observed that method reproducibility was related to the mean concentration of the analyte expressed as a fraction (C). He derived an equation to estimate the expected Horwitz's coefficient of variation of reproducibility among laboratories (HCV_R) from C: $HCV_R = C*2*e^{(1-.5*\log[C])}$. The Horwitz HCV_R provides a target for the variation among single analyses that is expected for all types of acceptable methods (Table 4). Although the coefficient

of variation of reproducibility ($HCV_R = HSD_R/C$) decreases as the concentration of the analyte increases (Table 4), the Horwitz standard deviation of reproducibility (HSD_R) of analytical results increases with concentration. This suggests that analytes with high concentrations, such as fiber, will be measured less accurately than analytes with low concentrations, such a protein.

Table 4. Expected analytical variation based on Horwitz's equation (Horwitz, 1982).

Concentration (%)	HCV_R[a] (%)	HSD_R[b] (%)	Typical Method Range
10	2.83	0.28	Ash, moisture
20	2.55	0.51	CP
30	2.40	0.72	ADF
40	2.30	0.92	ADF, aNDF (legumes)
50	2.22	1.11	aNDF (legumes)
60	2.16	1.30	aNDF (grasses)
70	2.11	1.48	aNDF (straws)

[a]Horwitz's expected coefficient of variation of reproducibility for a single analysis.
[b]Horwitz's expected standard deviation of reproducibility for single analytical results.

The standard deviation of repeatability within laboratories (SD_r) is typically 30 to 50% of the standard deviation of reproducibility among laboratories (SD_R) for most analytical methods (Horwitz, 1982). The SD_R contains the variation within laboratories (SD_r) as well as differences among laboratories in methods, techniques, training of the analyst, and the equipment and reagents used.

Variation in NDF results. The performance parameters for the collaborative study to evaluate aNDF (Mertens, 2002) match the target values expected by the Horwitz equations for forages and are slightly higher than the target for concentrates (Table 5). The R-value in Table 5 estimates the maximum variation in aNDF results expected among laboratories that analyze a single sample of the same material. Nineteen out of 20 laboratories should be within R of one another when analyzing a single sample of a well-mixed material. Thus, we should expect 19 of 20 laboratories to be within 3 to 4 %-units of NDF on an as-is or as-received basis when analyzing a split sample that is mixed well.

Table 5. Performance parameters for the aNDF method (Mertens, 2002).

Feed	No.[a]	Fiber	Mean	SD_r[b]	SD_R[c]	R[d]

			(%)			
Forages	5	aNDF[e]	52.2	0.84	1.10	3.08
Forages	5	aNDFom[f]	50.4	0.93	1.18	3.31
Concentrates <10% fat	3	aNDF	33.5	1.18	1.42	3.98
Concentrates <10% fat	3	aNDFom	32.5	1.16	1.46	4.08
Concentrates >10% fat	3	aNDF	8.7	1.01	1.61	4.50
Concentrates >10% fat	3	aNDFom	8.8	0.58	1.20	3.37
All materials	11	aNDF	38.7	1.05	1.33	3.72
All materials	11	aNDFom	37.7	1.02	1.28	3.58

[a] Number of materials.
[b] Standard deviation of repeatability within laboratories.
[c] Standard deviation of reproducibility within and among laboratories.
[d] $2.8*SD_R$ = approximate 95% confidence interval for a single analysis among laboratories.
[e] amylase-treated neutral detergent fiber.
[f] amylase-treated neutral detergent fiber organic matter (ash-free fiber).

Variation in IVNDFD. The expected variation for IVNDFD is larger than expected based on the Horwitz equation for several reasons (Table 6). *In vitro* digestibility is a biological assay that depends on a source of inoculum that is more variable than chemical reagents. Variation in inocula and how they are handled and prepared prior to its use probably are the greatest sources of variation in the *in vitro* method. Digestion is a dynamic process and the rate of digestion constantly changes over time. At early fermentation times (<30 h), the changes in digestion with time are large and variable, thereby increasing the variation in the measurement of IVNDFD compared to times after 48 h.

Table 6. Variation in *in vitro* NDF digestibility (IVNDFD) at various times (estimated from Mertens, unpublished and Barnes, 1967).

	Number of replications		
In vitro time and source	1	2	4
In vitro NDFD 30h, within lab between runs	5.0	3.5	2.5
In vitro NDFD 30h, among labs	12.0	8.4	6.0
In vitro NDFD 48h, within lab between runs	4.0	2.8	2.0
In vitro NDFD 48h, among labs	10.0	7.5	5.0

In vitro digestibility is also more variable than chemical methods because it requires the measurement of both NDF and *in vitro* NDF residue. Each of these

measurements has variation and the measurement of *in vitro* digestibility is a function of both measurements. Variability in IVNDFD is also a statistical/mathematical artifact. Digestibility has greater variation because the denominator in its calculation (NDF in dry matter) is a fraction and the variance of the results is directly proportional to the reciprocal of the fraction. Mathematically, when any group of numbers is divided by a constant, their standard deviation is divided by the same constant. When this constant is a fraction, then variation is increased.

Comparison of analytical and sample variation. Although analytical variation can be significant, it must be put into perspective relative to sample variation. As indicated in Table 1, the number of site-samples that are taken and composited to obtain the submitted sample is a crucial factor in reducing variation and insuring that the sample is representative of the stack or lot of feed. Table 7 provides

Table 7. Comparison standard errors (SE) of the mean related to analytical and sampling sources in estimating the composition of a lot of feed.

Analytical	SE	Hay sampling	SE	Silage sampling	SE
Avg 'A' Lab, Single analysis	0.60	Cores from 5 bales	2.15	Samples from 5 harvest loads	2.46
Avg 'A' Lab, duplicate analysis	0.54	Cores from 10 bales	1.52	Samples from 10 harvest loads	1.74
Two avg 'A' Labs, single analysis	0.48	Cores from 15 bales	1.24	Samples from 15 harvest loads	1.42
Avg 'B' Lab, Single analysis	1.50	Cores from 20 bales	1.07	Samples from 20 harvest loads	1.23
Avg 'B' Lab, duplicate analysis	1.35	5 grab samples from chopped hay	3.00	5 daily or site grab samples	2.37
Two avg 'B' Labs, single analysis	1.21	10 grab samples from chopped hay	2.12	10 daily or site grab samples	1.68
Avg 'C' Lab, Single analysis	2.50	15 grab samples from chopped hay	1.73	15 daily or site grab samples	1.37
Avg 'C' Lab, duplicate analysis	2.27	20 grab samples from chopped hay	1.50	20 daily or site grab samples	1.19
Two avg 'C' Labs, single analysis	2.00				

estimates of the variation among samples taken by various methods, and is based on published results and the methods we use during research trials. They are compared to the variation among laboratories with various levels of proficiency as

observed by National Forage Testing Association. Selecting an accurate and precise laboratory is important, and so is collecting an excellent, representative sample. Minimizing analytical variation cannot compensate for a poorly collected and mixed sample.

Minimizing the Effects of Analytical Variation

Average the results from multiple samples. The most important way to minimize the effects of analytical variation is to average the results of multiple, representative samples analyzed by laboratories that have demonstrated an ability to generate accurate and precise results. It is amazing that the cost of repeated analysis ($10 to 50) is a major concern in establishing the nutritional value of $20,000 worth of feed or in formulating more accurate rations that may save several cents in feed costs per cow per day. The cost of analysis is inconsequential in relation to the cost of a daily loss of a pound or less of milk per cow because the ration was not accurately formulated due to a poor estimate of nutritional value.

For most large lots of feed, at least two independent samples (not taken on the same day or location in the silo or stack of hay) should be analyzed. Heterogeneous feeds will require more than two results to determine composition adequately. The goal is not to obtain results that agree, but to obtain independent results that can be averaged to give a more accurate estimate of what is actually in the feed. The variation reported in Tables 4, 5, 6, and 7 can be used to determine if replicated results are within normal expectations. If not, obtain additional analysis to detect and remove any suspect data before averaging results.

Use multiple feed ingredients in the ration. After multiple results are averaged to estimate the true nutritional value of feeds, nutrient variation in the total mixed ration can be reduced by using multiple feed ingredients, when each feed is less than 30% of the ration. Each feed ingredient has variation in composition. By combining multiple feeds, it is extremely unlikely that they will vary in the same direction at the same time (e.g., they will not all be higher in fiber on the same day). Thus, complex rations have less variation than simple rations where one or two feeds comprise a majority of the ration.

Eliminate ration variation due to daily dry matter fluctuations. Change in dry matter (DM) of feed ingredients from day-to-day is the greatest source of variation in the final ration. Diets are formulated on a DM basis, but must be fed on an as-is basis. Changes in the DM of an ingredient, whether due to changes in the forage or a precipitation event, can have a dramatic effect on the actual

amount of ingredient DM that is fed on a given day and the nutrient composition of the ration.

Conclusions – Implications

- All analytical results are estimates of the actual average nutritional value in a lot of feed. The closeness of the analytical result to the actual average depends on averaging results from multiple representative samples that are analyzed by competent laboratories.
- Although variation can be controlled or minimized, some random variation is natural and unavoidable in analytical results.
- Accuracy (closeness to the correct value) and precision (consistency in replicated values) are independent sources of variation. Precise results may not be accurate.
- Replicated analyses not only provide a more accurate average estimate of actual composition, but also provide information that can be used to evaluate differences between results.
- Horwitz (1982) observed that reproducibility of analytical results was related to the mean concentration of the analyte across a broad range of methods.
- *In vitro* digestibility will have more variability than chemical analysis because of variability in biological reagents (inocula) and greater complexity of *in vitro* methods.
- Obtaining a representative sample to be submitted for analysis is often the largest source of variation among analytical results.
- Replicating sampling and analysis reduces variation in the average result.

References

Barnes, R.F. 1967. Collaborative *in vitro* rumen fermentation studies on forage substrates. J. Anim. Sci. 26:1120-1130.

Goering, H.K., and Van Soest, P.J. 1970. Forage fiber analyses. USDA Agric. Handbook No. 379. U.S. Government Printing Office, Washington, D.C. 20 pp.

Grant, R.J. and Mertens, D.R. 1992. Impact of *in vitro* fermentation techniques upon kinetics of fiber digestion. J. Dairy Sci. 75:1263-1272.

Hintz, R.W., Mertens, D.R. and Albrecht, K.A. 1996. Effects of sodium sulfite on recovery and composition of detergent fiber and lignin. J.A.O.A.C. Inter. 79:16-22.

Horwitz, W. 1982. Evaluation of analytical methods used for regulation of foods and drugs. J.A.O.A.C. 65:525-530.

Mertens, D.R. 2002. Gravimetric determination of amylase-treated neutral detergent fiber in feeds using refluxing in beakers or crucibles: Collaborative study. J. A.O.A.C. 85:1217-1240.

Robertson, J.B., and Van Soest. P.J. 1981. The detergent system of analysis and its application to human food. In The Analysis of Dietary Fiber in Food. W.P.T. James and O. Theander (ed.). Marcel Dekker, Inc., NY, p. 123-158.

Van Soest, P.J., and Wine, R.H. 1967. The use of detergents in analysis of fibrous feeds: IV. Determination of plant cell-wall constituents. J. AOAC. 50:50-55.

Van Soest, P.J., Robertson, J.B. and Lewis, B.A. 1991. Methods for dietary fiber, neutral detergent fiber and nonstarch polysaccharides in relation to animal nutrition. J. Dairy Sci. 74, 3583-3597.