Alfalfa Hay Testing: What do the results really mean?

Dan Putnam[1]

ABSTRACT

Growers, brokers, and buyers of alfalfa hay increasingly rely on laboratory analysis of alfalfa hay to predict animal performance and to determine both value and price. Chemical analysis through wet chemistry or Near Infrared Spectrophotometry (NIRS) have been increasingly used to provide an objective measure of forage quality, to be used along with visual inspection. The average value of high compared with low quality alfalfa hay has been 30$/ton or approximately $200 million annually in California for the past five years. However, many in the alfalfa industry complain about variation between laboratories results, and the abuse of these results in the marketplace. There are several sources of variation for hay analysis: sampling error, subsampling error, analytical error, errors in calculation, and the relationship of a calculated value such as TDN to animal performance. The greatest of these from a growers perspective is usually sampling error or the ability of a small sample to represent a large stack. Analytical errors tend to be smaller than sampling errors, but a high level of variation from lab to lab has been observed in split sample comparisons. Steps to help standardize laboratory practices across the region could be taken. Both the limitations and the value of hay testing must be recognized.

INTRODUCTION

According to USDA figures, California is now number one in total milk production, producing 1.93 billion pounds of milk per year as of January, 1994. Wisconsin produced 1.81 billion pounds during the same period (CDFA, 1994). The rankings were reversed in the previous year. What is more striking is a trend which shows California with an 6% increase in milk production over this period, and Wisconsin with an 8% decline. Average yield per cow in California (1,619 lb/cow) increased 4% over the previous year, and is significantly higher than the national average (1,319 lb/cow). High-yielding cows require premium nutrition. All of these factors place pressure on alfalfa growers in this region to not only maximize yields but maximize quality as well.

Lab analysis has been increasingly used to help evaluate the feeding value of hay in California and other western states. However, laboratory testing of alfalfa hay can be quite frustrating. Growers complain of differing results from lab to lab or that certain labs report consistently higher or lower numbers than others. Some labs rely on Near Infrared Spectrophotometry (NIRS), whereas others only use wet chemistry methods. Are NIRS results as reliable as wet chemistry? How can lab results differ so greatly? This paper is examines the value of forage quality to alfalfa growers, and the sources of variation and biases which may occur in the testing process.

PERCEPTIONS OF FORAGE QUALITY

Alfalfa hay quality may be generally defined as the ability of a given quantity of hay to produce a

---

[1]Extension Agronomist, Department of Agronomy and Range Science, University of California, Davis, CA 95616.

desired effect in the consuming animal. This "true feeding value" of animal performance is conceived of as the nutrients used per unit time (Marten et al., 1988). However, definitions of alfalfa hay quality will differ greatly depending upon which type of animal is being considered: e.g. high vs. low dairy producers, or beef vs. horse production. This very fact implies that there are animal genetic and management factors involved with "true feeding value". What we really mean when we consider alfalfa forage quality is the potential feeding value for a given application. This latter concept encompasses 1. Potential nutritive value, 2. Antiquality factors, and 3. Potential intake.

Table 1 California Alfalfa Hay Quality Designations.*

Designation	Verbal Description:	TDN (90% dm)	ADF (100% dm)
Premium	Prebud or prebloom, low fiber with soft stems, high energy and protein content. Good color, very good leaf attachment, mostly free of grasses and weeds, no noxious weeds, well cured	54 or greater	29 or less
Good	Prebloom to early maturity, low to medium fiber with soft stems, fairly high energy and protein content, good color, fairly good leaf attachment, fairly free of grasses and weeds, no noxious weeds, well cured.	52 to 54	29 to 32
Fair	Early to late maturity, medium to high fiber with coarse stems, low to moderate energy and protein content, fair color, fair leaf attachment, low to moderate grass and weed content, no noxious weeds, well cured.	49 to 52	32 to 37
Low	Hay with a serious fault or faults	< 49	> 37

*CDFA, 1994. Designations effective March 1, 1985. Also see Bath & Marble, 1989.

The California system for designating hay quality has been in place for some time, and the most recent designations (enacted in 1985) are provided in Table 1. This table illustrates the importance and role of objective evaluations of alfalfa hay. Certainly many feel that they have a good idea of what "high energy, good color, mostly free of grasses" means in judging a hay load. However, it is likely that those opinions will differ from the thousands of others who must also judge hay! Hay pricing based in part upon objective chemical analysis has been welcomed by most in the industry as a method to remove subjectivity from hay transactions. However, though growers have looked to lab testing as the "rock of Gibraltar", in recent years that too seems to be fraught with uncertainty and the possibility of error.

VISUAL ESTIMATION VS. LAB ESTIMATION OF HAY QUALITY:

Hay quality has been visually rated for centuries as animal husbanders learned of the differential responses of animals to different feeds. Visual rating of alfalfa includes the factors of growth stage, degree of leafiness, foreign material or weediness, color, texture (softness), pathogens or mold, and odor. Experienced buyers can often do a good job of judging hay quality visually, but examples abound of hay which looks good but isn't, or superficially poorer hay which may actually have a superior feeding value. It is clear that a more analytical approach is needed.

Several recent anecdotes illustrate this problem. In 1993, I was asked to help judge hay for the FFA Crops Judging Contest, which is facilitated by the University of California, Davis. The students were asked to judge four bales of hay. Two were dead ringers: one sample was soft, dark green, pure fine alfalfa with plenty of retained leaves, and an excellent odor. Another sample was very stemmy, rough texture, discolored, weedy, and slightly moldy. These were our obvious choices for rankings 1 and 4. However the two middle ones were tough. One was a

Table 2. Reliability of Visual and Chemical methods for hay judging.*

Quality Factor	Judgement by visual inspection	Judgement by Chemical Analysis
Growth Stage	Poor	Excellent
Leafiness	Fair	Excellent
Fiber Content	Poor	Excellent
Foreign Material	Excellent	Poor
Color/Odor	Excellent	Poor
Texture/condition/mold	Excellent	Poor

*Adapted from Bath & Marble, 1989.

deep green color, but less leafy than number 1, pure alfalfa with no weeds but stemmy. Another bale contain some decent alfalfa but had considerable grassy weed content. After considerable debate between myself and the other judges (more experienced than I), we chose the alfalfa bale with no weeds as number 2, and the weedy bale as number 3. Students were equally split on these two samples. However, when tests were returned shortly before the student contest, it showed the weedy sample with an Acid Detergent Fiber (ADF) level almost 5 percentage points lower than the stemmy sample, indicating that the weedy sample had better feeding value.

A similar experience was encountered in a hay-judging contest in McArthuur, CA in 1993. Three judges (myself, a hay broker, and a hay-testing lab representative) were able to generally identify the better samples, but failed to identify the top sample in TDN (we rated it number 4 of 16). The reverse was also true. A bale rated as high quality by lab analysis turned out to be quite moldy, of questionable feeding value. It is clear that visual analysis alone cannot accurately or consistently describe alfalfa forage quality. Conversely, chemical analysis should always be used in addition to visual inspection. The reliability of visual vs. chemical analysis for various quality factors of alfalfa hay is shown in Table 2.

THE VALUE OF HAY QUALITY.

What has high quality been worth in the California hay markets? From 1989 to 1993, the average penalty for low quality hay (or reward for production of premium quality hay) was approximately $30 per ton, averaged over all California markets. It ranged from less than $10 to over 50$/ton. There were some differences between sites; for example the differences between premium and fair hay in Imperial and northern California were not as large as the differences in the Central Valley. There were also slightly smaller differences in years when the price was higher. If all alfalfa hay in California received this premium, the difference between high and low quality hay would be about $200 million each year to California growers. Of course, it is much more complicated than that; nature does not always allow growers to produce high quality hay, and the price paid for quality depends upon a lot of factors, including availability of hay, cow

Table 3. Price differential between premium quality (TDN 54 or greater) and fair quality (TDN 49-52) hay at different California Market locations over a 5 year period.*

Reporting Area	1989	1990	1991	1992	1993	Ave.
Chino Valley - L.A.	$19.37	$21.76	$25.11	$30.56	$22.63	**$23.88**
Tulare-Visalia-Hanford	$23.38	$25.29	$42.42	$42.01	$29.62	**$32.54**
Escalow-Modesto-Turlock	$30.70	$38.41	$48.36	$40.02	$31.32	**$37.76**
Petaluma	$28.65	$26.36	$30.73	$33.07	$33.78	**$30.52**
Imperial Valley	$26.36	$20.14	$25.87	$28.74	$26.88	**$25.60**
Kern	$13.34	$23.21	$24.88	$40.55	$21.35	**$24.67**
Hanford-Corcoran-Tulare	-	-	$42.28	$43.15	$30.66	**$38.67**
W. Fresno-Madera	-	-	$42.83	$50.27	-	**$46.55**
Los Banos-Dos Palos	25.93$	$26.99	$39.90	$44.14	-	**$34.24**
Tracy-Patterson	$23.86	$26.84	$31.72	$37.03	$30.97	**$30.08**
Stockton	$26.59	$24.80	$32.34	$34.47	$28.00	**$29.24**
Sacramento Valley	$25.60	$21.79	$22.71	$27.67	$17.67	**$23.09**
Northeastern California	-	$25.71	$28.28	$30.02	$32.49	**$29.12**
Average	$24.38	$25.57	$33.65	$37.05	27.76	$30.16

* Data calculated from CDFA, 1994. Yearly averages of monthly averages of weekly quotations. Data not adjusted for missing entries in monthly reports.

numbers and milk price. If all hay were premium quality, there would be no price differential! But this analysis provides a baseline to answer the question in the title of this paper: the results of quality testing could mean about $30/t to alfalfa growers or about $200 million/year to the industry in California.

THE TESTING PROCESS

What we are attempting to do with forage testing is at first glance a little preposterous. We collect less than a half pound sample, and expect it to be representative of tons of a heterogeneous mass of plant material consisting of alfalfa stems, leaves, flowers, various weeds and foreign material. Then we perform some chemical tests, and hope that these have some relationship to the performance (usually milk production) of an equally varied population of cows. But actually the process is not that bad if we understand that forage testing is innately <u>statistical</u>. That is, that it gives an approximate answer within a certain level of variation. Chemical constituents <u>do</u> have a relationship to animal productivity. The only thing we can do is minimize the variation, and understand that the variation will always be there, so that we can make intelligent use of the numbers that emerge.

The chemical constituents which have the most value for animal nutrition are crude protein and energy. Crude protein is generally measured by standard Kjeldahl procedures that estimate percent N. Protein is then calculated from N content. However energy is not measured directly. It is estimated indirectly by measuring fiber (either Modified Crude Fiber or MCF, or Acid Detergent Fiber, ADF, or Neutral Detergent Fiber, NDF). Fiber is inversely related to energy value and feed intake. As fiber goes up, energy and intake are reduced. Although it would be undesirable to completely eliminate fiber from forages (it has important physical effects on animal nutrition), high levels of fiber in forages are the primary cause of low forage quality.

Total Digestible Nutrients (TDN) is a concept used to approximate energy content. It is simply calculated from the lab value for fiber, either ADF or MCF. Although both ADF and MCF can be used to calculate energy, ADF has more national acceptance, and is now more widely accepted in California. Unfortunately, not all labs use the same equations to calculate TDN from ADF (there are several published equations), a significant source of variation. Other concepts such as Net Energy for Lactation (NEl) may be more useful concepts as nutritionists improve their concepts of forage quality. Many nutritionist are now recommending obtaining NDF values, which have a strong relationship to forage intake.

RANDOM ERROR AND BIAS

The hay testing process consists of:

- Sampling from the stack or field,
- Preparing and grinding samples,
- Subsampling,
- Analyzing the sample,
- Calculating a chemical value,
- Predicting TDN or energy, and using that value to
- Predict animal performance.

Each of these steps contain some error and the possibility of bias. None of these steps is perfectly repeatable. Any time you go to properly sample a stack of hay, you will probably get a slightly different lab value. The same is true if the same ground sample is subsampled for lab analysis (subsampling error), etc. However, bias occurs when the methodology is not adequate, or when multiple methodologies are used. For example, a probe of 3" depth in a bale are likely to bias the results towards the stemmier outside of the bale compare with a complete sample of the entire bale. Random error is easier to deal with than bias, since even if it is large, on the average over a large range of tests it should give the correct answer. If significant bias is present in the sampling,

Table 4. Sampling error (standard deviations) and errors for NIRS prediction for protein, Acid Detergent Fiber and Neutral Detergent Fiber.

Source of Variation	Protein	ADF	NDF
Variation within hay loads	.54	2.41	4.17
Variation within individual bales	0.65	.4	.68
NIRS Analytical Error	0.66	1.02	1.97

Data from Pennsylvania (Schenk, 1981).

analysis, or prediction, it is difficult to get an accurate sample no matter how many samples are run.

Of the errors associated with each of these processes, the largest from a growers perspective is without a doubt sampling error (Table 4). One cannot overestimate the significance of sampling error. In most cases it will be higher than analytical error. Shenk (1981) estimated the sampling errors to be 2.4 percentage points for ADF. This translates into an even higher error for TDN. To minimize sampling error, Bath and Marble (1989) recommend a minimum of 20 cores to be obtained per hay lot. Considerable attention should be paid to sampling methodology . The more variable the lot, the greater number of samples required.

CONSISTENCY BETWEEN LABS.

Although sampling variation may to a large extent determine the variation in lab results, there may also be substantial lab variation. How successful are laboratories at arriving at reliable values for forage quality? Two recent surveys of various labs in California show that there is reason for concern. In one test of laboratory consistency, 4 bales were sampled, ground, properly divided, and sent to 5 different labs (Table 5). (These were the same four bales that were used in the FFA Hay judging contest.) This survey revealed an high level of variation in both crude protein and ADF. Protein levels for sample 1 ranged from 20.9% to 28.2% protein, and ADF values ranged from 30.8% to 39.9% ADF. These labs reported TDN values from 51.6 to 57.5 for this sample (Table 6). This hay bale could have been classified as either "premium" or "low" depending upon the lab one chooses to believe.

A similar test of laboratory consistency was performed by the American Registry of Professional Animal Scientists (ARPAS). Their data showed somewhat less, but still significant variation between labs (Table 7). In this test, samples were split after coring but before grinding (lab error is confounded with subsampling error). Nevertheless, both tests indicate a range of variation that is disturbing. For example, sample 1 in the ARPAS test would be classified as either premium quality (56.8 TDN adjusted to 90% DM) or low quality hay (TDN = 49.6) depending upon which lab was chosen (Table 7). Many labs are using different equations to predict TDN. When one equation is used consistently, then the standard errors are usually reduced (Table 6 and 7).

MISUSE OF LAB RESULTS

It is clear from these data that lab results are often misused in the marketplace. Often, growers and brokers have a definition of "test hay", or a "quality cutoff". If 55% TDN is the cutoff, a sample which comes back at 54.6 is considered unacceptable. Under current conditions given sampling and lab error, we clearly cannot test to this level of precision. Bath and Marble (1989) recommend an incremental pricing structure, which allocates a price increment for each unit increase in ADF or TDN rather than an abrupt change at 54 or 55 TDN which currently occurs. An approach such as this would be a definite step in the right direction. Even if we were able to control sample and lab variation to a greater degree, forage quality should be viewed as a continuum, with many small levels of changes in feeding value. There is no such thing as "test" and "no test" hay. There will be a range of error associated with any result, no matter how careful the methodology.

Table 5. Crude Protein and ADF results from four hay bales* sampled and split after grinding, and sent to 5 labs. Results are on 100% DM basis.

	Crude Protein				Acid Detergent Fiber			
	Hay Sample Number							
	1	2	3	4	1	2	3	4
Wet Chemistry Method								
Lab 1	20.9	16.9	27.0	21.2	39.9	41.7	26.3	32.1
Lab 2	22.0	17.0	27.0	21.4	36.0	40.2	24.7	30.3
Lab 3	23.5	19.3	22.1	21.8	31.6	37.0	21.9	28.9
Lab 4 (run 1)	23.7	18.8	29.8	21.4	33.8	37.4	23.0	29.2
Lab 4 (run 2)	25.6	20.6	29.4	23.2	34.0	38.0	22.5	28.9
NIRS Method								
Lab 4 (run 1)	25.6	20.6	29.4	23.2	34.0	38.0	22.5	28.9
Lab 4 (run 2)	28.2	22.0	31.6	24.8	30.8	35.7	20.1	28.3
Lab 5	24.5	19.4	28.6	22.6	33.7	37.8	22.4	29.0
Mean	**23.9**	**19.1**	**28.0**	**22.3**	**34.3**	**38.2**	**22.9**	**29.5**
Stnd Dev	2.25	1.71	2.82	1.22	2.83	1.88	1.85	1.16

Table 6. Effect of calculation method for prediction of Total Digestible Nutrients (TDN) from Acid Detergent Fiber (ADF) from several laboratories from a sample from a single bale (split after grinding). ADF values from these labs are given in Table 5 (Hay Sample 1).

	TDN Reported	TDNa*	TDNb*	TDNc* (recommended)
Lab 1	57.4	51.6	45.4	47.2
Lab 2	49.8	54.4	49.5	49.8
Lab 3	52.5**	57.5	54	52.8
Lab 4 (run 1)	50.4	55.1	50.4	50.4
Lab 5	***	56.5	52.5	51.8
Mean(8 obs.)	**52.2**	**55.6**	**51.2**	**51**
Stnd Dev	2.81	1.88	2.74	1.79

*Eq-a: %TDN = 88.9-(0.79 x ADF%) ('legume & grasses' - Kautz et al., 1990)
Eq-b: %TDN = 96.35 - (1.15 Xx ADF%) ('Alfalfa' - Kautz et al., 1990)
Eq-c %TDN = 82.38 - (0.7515 x ADF%) (Alfalfa in Western States - Bath & Marble, 1989)
This lab reported TDN calculated from Modified Crude Fiber. * Not predicted by this lab.

Table 7. Results of two alfalfa hay samples sent to 9 laboratories, summer, 1994. The letter "n" indicates the number of laboratories participating in each test. Samples were split before grinding. TDN in this comparison was not adjusted to 90% dry matter (to adjust, multiply values by .9).

	Crude Protein	ADF	TDN Reported	TDNc*
		Sample 1		
NIRS Method				
Range	17.1 to 20.1	31.3 to 37.3	56.4 to 63.1	54.3 to 58.8
Mean	18.2 (n=7)	34.1(n=6)	58.3(n=6)	56.7(n=6)
Std. Dev.	0.93	1.86	2.47	1.39
Wet Chemistry Method				
Range	16.7 to 19.1	33.2 to 40.5	55.1 to 59	51.9 to 57.4
Mean	17.7(n=6)	35.7(n=5)	57(n=5)	55.5(n=5)
Std. Dev.	0.81	2.74	1.31	2.06
		Sample 2		
NIRS Method				
Range	23.2 to 26.1	23.3 to 26.6	62.4 to 69.4	62.3 to 64.9
Mean	2 (n=7)	24.4(n=6)	65.0 (n=6)	64.0(n=6)
Std. Dev.	0.98	1.05	2.16	0.79
Wet Chemistry Method				
Range	23.4 to 24.9	22.0 to 25.9	63.6 to 69.6	62.9 to 65.8
Mean	24.2 (n=6)	24.1(n=5)	65.4 (n=5)	64.2 (n=5)
Std. Dev.	0.44	1.29	2.24	0.97

* TDNc = TDN calculated from reported ADF using a single equation recommended by University of California. Data courtesy of John Kennedy, California/Nevada Chapter, American Registry of Professional Animal Scientists.

VALUE OF NIRS

The data presented here show a disturbing amount of variation among laboratories. This variation occurred whether wet chemistry or Near Infrared Spectrophotometry (NIRS) methods were used. Current use of NIRS technology does not appear to be any more variable than wet chemistry methods, in fact the standard deviations for NIRS were often lower than for wet chemistry. In theory, NIRS might be less variable, since it is more mechanized, and machines can be networked so that all machines in a region will provide the same reading. Repacked NIRS scans (indicates subsampling error) of sample 1 (Table 7), showed Standard deviations of 0. 76, for ADF , far less than the lab to lab variation. Four scans of the same sample (machine error) produced standard deviations of <0.01 (Tony Pietrantonio, pers. comm.)

However, NIRS must continually be updated with wet chemistry calibrations, and machines correctly maintained . NIRS has become widely accepted in the forage industry nationally, and with proper attention, there is no reason it should not be useful here. With the increasing pressure for rapid results (an advantage of NIRS), and environmental restrictions on chemical reagents (a disadvantage of wet chemistry), NIRS is likely to be more widely used in the future. Accurate prediction of forage quality from an NIR spectrum or from wet chemistry will depend upon the testing industrys' willingness to cooperate to iron out the controllable sources of variation, such as lab methodology, reliable calibration sets, and uniform calculations of energy value.

CONCLUSIONS

The value of hay quality in California is estimated to be an average of about $200 million annually to alfalfa growers, or about $30/ton. Hay testing provides a valuable objective estimate of hay feeding value which will be of increasing importance to the alfalfa industry in the future. Test values should always be viewed as relative not absolute values, with an error range surrounding each number. Sampling error is often the greatest source of variation, but lab to lab variation was also found be quite high. This creates considerable difficulties in reliable quality estimation.

RECOMMENDATIONS

There are several steps that the alfalfa industry can take to improve the problems with hay testing:

Insist upon vigorous coring and sampling methods, following published recommendations (eg. Bath & Marble, 1989).
Encourage incremental pricing methodologies rather than "cutoff" levels.
Encourage labs to organize and work together to improve methodologies.
Networks of NIR instruments, for example, vastly improve precision of prediction.
Encourage labs to participate in the National Hay Testing Association program.
Calculate TDN from ADF, not MCF. Both can be used, but some labs are using ADF and others MCF, creating added bias. ADF is now more widely accepted.
Insist upon uniform calculations of TDN from ADF across the state, or base pricing primarily upon ADF values.
The alfalfa industry should work with dairy nutritionists and labs to better understand the definitions of hay quality and agree on testing methodology.

REFERENCES

Bath, D.L., V.L. Marble. 1989. Testing Alfalfa for its feeding value. University of California Cooperative Extension and Division of Agric. and Natural Resources. Leaflet 21457 Western Regional Extension Publication WREP 109.

California Department of Food and Agriculture, 1994. Alfalfa Hay California Market Summary 1993. Federal-State Market News Service, CDFA, Market News Branch. Sacramento, CA. USDA Cooperating.

Kautz, W.P., E.J. Lazowski, W.C. Mahanna, R. Reinhart. 1990. Pioneer Forage Manual - A nutritional guide. C. Holland and W. Kezar (eds.). Pioneer Hi-Bred International, Inc.
Des Moines, IA.

Shenk, J.S., 1981. Using the near infrared analysis system to predict hay quality. Proc. 11th California Alfalfa Symposium, Dec. 8-9, Fresno, CA. Univ. of California Cooperative Extension.