Accuracy and Precision of Near Infra-red Spectroscopy (NIRS) versus Wet Chemistry in Forage Analysis

Saha, U.K.*; Kern-Lunbery, R.**
*Agricultural and Environmental Services Laboratories, University of Georgia
** WARD Laboratories, Inc.

Keywords: Accuracy, Forage, Near Infra-red Spectroscopy, Precision, Wet Chemistry

Abstract. Near Infra-red Spectroscopy (NIRS) is an attractive option for forage analysis. NIRS is less labor intensive, nondestructive, rapid, environmentally friendly and provides accurate and precise results. However, many nutritionists are quick to brush off NIRS, citing 'poor accuracy'. We evaluated the accuracy and precision of dry matter (DM), crude protein (CP), acid detergent fiber (ADF), and neutral detergent fiber (NDF) of 33 National Forage Testing Association (NFTA) proficiency test (PT) alfalfa hay samples analyzed by NIRS in 7 NIRS Forage and Feed Testing Consortium (NIRSC) member laboratories. The reference method averages (RMA), used to evaluate the NIRS results, were based on the wet chemistry results reported by numerous laboratories participating in the corresponding NFTA proficiency testing rounds. Thus, this study is a robust comparison of NIRS determined results with the corresponding wet chemistry results, which is still a "gold standard" to many nutritionists. These results demonstrate that when NIRS calibrations are developed using good science and applied properly, NIRS is as accurate as wet chemistry in forage nutritional analysis. Further, both intra-laboratory and inter-laboratory precision of NIRS methods are superior to wet chemistry methods.

Introduction

The two pillars of quality assurance are accuracy and precision. Precision is the magnitude of the errors associated with repeated measures of the same analyte on a specific sample. When an analyte is measured on a given sample in replicates in multiple laboratories, there are two different types of errors associated with precision. First, repeatability error, which measures the error associated with replicated measurements done within a given laboratory. Second, reproducibility error, which is the sum of the repeatability error plus the inter-laboratory error associated with repeated measurements done on the same sample in multiple laboratories. Accuracy simply measures the closeness of the measured value to the true value, known value, or consensus value.

Near Infra-red Spectroscopy (NIRS) is an attractive option for forage analysis because it is less labor intensive, low cost, nondestructive, rapid, and environmentally friendly. However, many nutritionists are skeptical about NIRS, citing its poor accuracy compared to wet chemistry. That was probably true during the early days of NIRS. However, enormous advancement in NIRS hardware along with sophisticated mathematical/statistical tools/software, allow NIRS analysis to produce results equivalent to wet chemistry methods.

NIRS Forage and Feed Testing Consortium (NIRSC, https://nirsconsortium.org/) member laboratories regularly participate in the proficiency testing (PT) administered by the National Forage Testing Association (NFTA, https://www.foragetesting.org/). NFTA evaluates the reported NIRS results against the wet chemistry reference method averages (RMA). This study evaluates the intra- and inter-laboratory precision of forage analysis results reported by 7 NIRSC laboratories as well as overall accuracy of NIRS compared to wet chemistry methods, which are still considered the "Gold Standard" in forage analysis.

Methods

We used the dry-matter (DM), crude protein (CP), acid detergent fiber (ADF), and neutral detergent fiber (NDF) of 33 standard alfalfa hay samples analyzed in triplicates by NIRS in 7 NIRSC member laboratories and reported as PT participants to NFTA during 2013-19. For accuracy, NFTA evaluates the NIRS-predicted results against corresponding RMAs calculated using the results reported by the participating wet chemistry laboratories. Therefore, wet chemistry is the reference method for evaluation of proficiency of the NIRS method.

We evaluated the precision and accuracy of forage analysis by NIRS using the well-known Horwitz function (Horwitz et al., 1980). The authors pointed out, "An examination of the results of over 50 inter-laboratory collaborative studies conducted by AOAC on various commodities for numerous analytes shows a relationship between the mean coefficient of variation (CV), expressed as powers of 2, with the mean concentration measured, expressed as powers of 10, independent of the determinative method".

The above statement was originally expressed by the statistician Jung Keun Lee in the following form:

$\mathbf{RSD_R, \% = 2^{(1 - 0.5 \text{ Log } C)}}$

Where C, is the concentration of analyte expressed as dimensionless mass fraction (numerator and denominator have the same units); and $\mathbf{RSD_R}$ is the coefficient of variation CV under reproducibility conditions. Michael Thompson (1999) transformed the original equation into an easier equivalent form:

$\mathbf{RSD_R, \% = 2\ C^{-0.15}}$ or as a standard deviation: $\mathbf{SR = 2\ C^{0.85}}$

The Horwitz Standard Deviation or **HSD** expressed as % on the NFTA PT report is:

$\mathbf{HSD = S_R \times 100}$

C is indeed the RMA on the PT-NFTA reports, but expressed as mass fraction (e.g., for 95%, DM; C = 0.95). Later, Michael Thompson (2000) found that precision was overestimated at the extreme values of C. As a result, the Horwitz Equation was further adjusted as:

$\mathbf{S_R = 0.22C}$; if C is $<1.2 \times 10^{-7}$

$\mathbf{S_R = 0.02C^{0.8495}}$; if $1.2 \times 10^{-7} \leq C \leq 0.138$

$\mathbf{S_R = 0.01C^{0.5}}$; $C > 0.138$

Evaluation of Precision

Equations used to evaluate intra- and inter-laboratory precisions or repeatability and reproducibility, respectively are:

For repeatability S,

$$\boldsymbol{S_r} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_k - 1)s_k^2}{(n_1 + n_2 + \cdots + n_k - k)}}$$

For repeatability RSD,

$$\boldsymbol{RSD_r} = \sqrt{\frac{(n_1 - 1)RSD_1^2 + (n_2 - 1)RSD_2^2 + \cdots + (n_k - 1)RSD_k^2}{(n_1 + n_2 + \cdots + n_k - k)}}$$

For inter-laboratory S,

$$\boldsymbol{S_L} = \sqrt{\frac{1}{k - 1}\sum_{i=1}^{k}(Xmean_i - Grand\ Mean)^2}$$

For reproducibility S,

$$\boldsymbol{S_r} = \sqrt{S_r^2 + S_L^2}$$

For reproducibility RSD_R,

$$\boldsymbol{RSD_R} = \frac{S_R}{Grand\ Mean} \times 100$$

Where, **k** is the number of labs, **n** is the number of replications in each lab, $\mathbf{Xmean_i}$ is the mean values of individual labs and **Grand mean** is the overall mean derived from **n x k** values (or data points).

The **HorRat** is the "Horwitz Ratios" for precision were calculated as:

$$Repeatablility\ HorRat_r = \frac{RSD_r}{PRSD_r}$$

Acceptable $HorRat_r$: 0.3 – 1.3

$$Reproducibility\ HorRat_R = \frac{RSD_R}{PRSD_R}$$

Acceptable $HorRat_R$: 0.5 – 2.0

Where, **RSD** is calculated from the reported concentrations by the participating labs in replicates **PRSD** the **RSD** predicted from Horwitz Equation.

Evaluation of Accuracy

Evaluation of accuracy using "Z-Score" based on Horwitz function:
The Z-score was calculated as follows:

$$Z = \frac{X_{meanL} - RMA}{HSD}$$

Where, $\mathbf{X_{meanL}}$ is the mean value of replicated measurements reported by a given laboratory, **HSD** = $0.02C^{0.8495}$; if $1.2 \times 10^{-7} \leq C \leq 0.138$ (or 13.8%), **HSD** = $0.01C^{0.5}$; C > 0.138 and **Z-score** ≤3.0: Satisfactory Accuracy, which is used by NFTA as the "Passing Grades".

Results and Discussion

All DM (Fig. 1), CP (Fig. 1), ADF and NDF values determined by NIRS in triplicates reported by the 7 NIR laboratories during 33 PT rounds were within the acceptance windows assigned by NFTA based on the wet chemistry method. This is a robust demonstration that NIR is as accurate as wet chemistry.

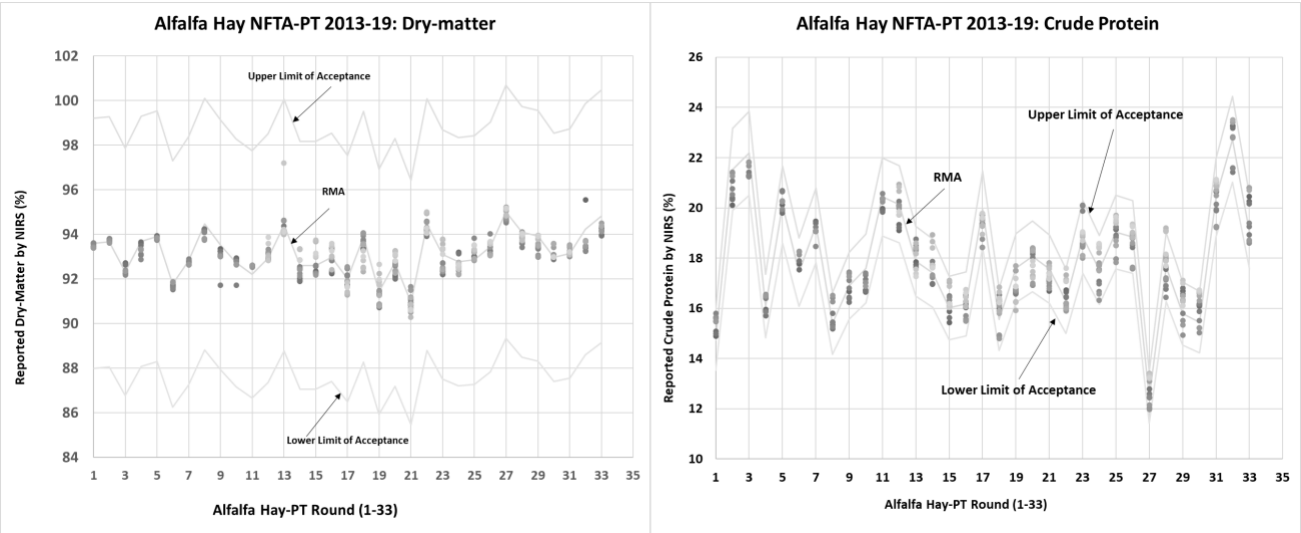

Figure 1. DM and CP contents in triplicates reported by 7 NIR laboratories along with RMA, upper and lower limits of acceptance in 33 Alfalfa Hay PT rounds conducted NFTA during 2013-19.

Figure 2 shows the plots of repeatability ($HorRat_r$) and reproducibility ($HorRat_R$) Horwitz ratios for CP analyzed by NIRS versus wet chemistry. Both repeatability and reproducibility Horwitz ratios for NIRS are much lower than the corresponding upper limits, suggesting that both repeatability and reproducibility of NIRS are excellent. In contrast, wet chemistry had excellent repeatability but relatively inferior reproducibility because the $HorRat_R$ exceeded the upper limit 2.0 in around one-third of the total 33 cases. The deviations, i.e., wet chemistry minus NIRS, were positive in most cases for both repeatability ($HorRat_r$) and reproducibility ($HorRat_R$) Horwitz ratios, demonstrating that in forage CP testing both repeatability and reproducibility of NIRS are better than wet chemistry. The plots for DM, ADF, and NDF are not shown for the sake of brevity.

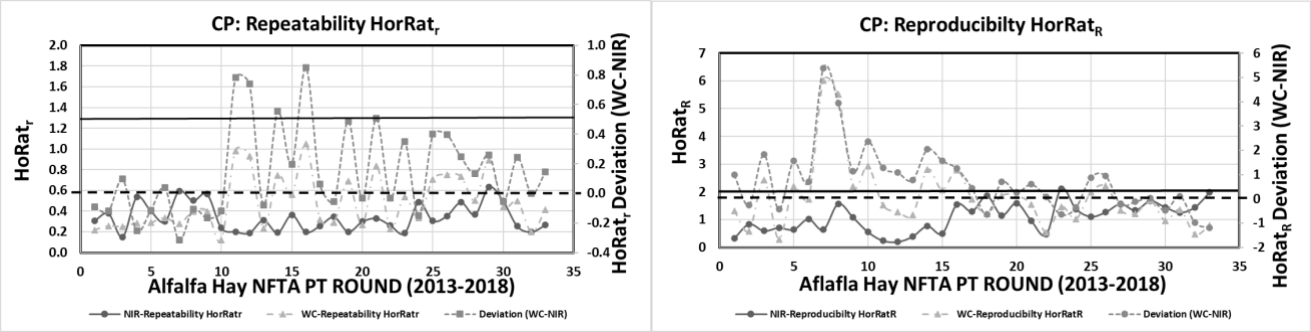

Figure 2. Plots of repeatability ($HorRat_r$) and reproducibility ($HorRat_R$) Horwitz ratios in forage CP testing by NIRS versus wet chemistry. The plots for DM, ADF, and NDF are not shown for the sake of brevity.

The plotted Z-scores for CP are lower than 1.5 in most cases (Fig.3), suggesting that all 4 labs achieved great accuracy in CP analysis by both wet chemistry and NIRS methods. There is not a clear-cut difference between the Z-scores of the two methods for most of the 33 alfalfa hay NFTA-PT rounds conducted from 2013-2019. Thus, there is hardly any difference in accuracy of CP between the two methods.

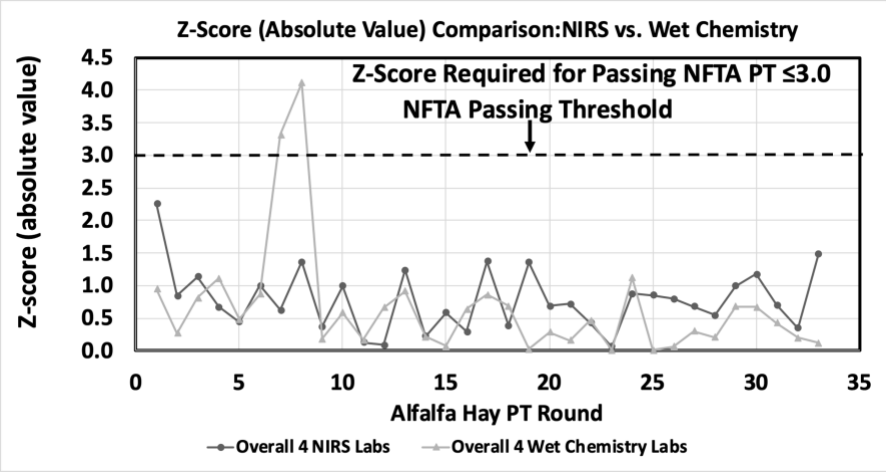

Figure 3. Plots of Z-scores showing accuracy of NIRS versus wet chemistry in forage crude protein analysis. The plots for DM, ADF, and NDF are not shown for the sake of brevity.

Conclusions

When NIRS calibrations are developed using good science and applied properly, NIRS is as accurate as wet chemistry in forage nutritional analysis. Both intra-laboratory and inter-laboratory precisions of NIRS method are superior to the wet chemistry methods. This is a robust demonstration of similar accuracy of NIRS and wet chemistry and even better precision of NIRS over wet chemistry in forage analysis.

References

Horwitz, W. and Albert, R.J. 2006. The Horwitz Ratio (HorRat): A useful index of method performance with respect to precision. *Journal of the Association of Official Analytical Chemist.*, 89: 1095-1109.

Horwitz, W., Kamps, L.R. and Boyer, R.W. 1980. Quality assurance in the analysis of foods for trace constituents. *Journal of the Association of Official Analytical Chemist.*, 63: 1344-1354.

Thompson, M. 1999. A natural history of analytical methods. *Analyst.*, 124: 991.

Thompson, M. 2000. Recent trends in inter-laboratory precision at ppb and sub-ppb concentrations in relation to fitness for purpose criteria in proficiency testing. Analyst., 125: 385–386.