Title: Evaluating Hand-Held NIRS Units for Measuring Forage Quality of Fresh-Chopped Alfalfa and Alfalfa Hay and Haylage.

PI: J.H. Cherney
Address: Section of Soil & Crop Sciences, School of Integrative Plant Science, Cornell Univ., Ithaca, NY 14853
Phone: 607-255-0945
Fax: 607-255-2644
Email: jhc5@cornell.edu

CoPI: M. Digman
Address: Dept. of Biological Systems Engineering, University of Wisconsin – Madison, Madison, WI 53706
Phone: 608-890-1941
Email: digman@wisc.edu

CoPI: D.J. Cherney
Address: Dept. of Animal Science, Cornell Univ., Ithaca, NY 14853
Phone: 607-255-2882
Email: djc6@cornell.edu

Abstract: (Limit 200-300 words)

Handheld near infrared reflectance (NIR) instruments are now available for non-scientist operators for on-farm evaluation of forages. There are a variety of instruments with different detector technologies and greatly different NIR scanning ranges. However, there is little public data or independent evaluation of their relative effectiveness in the field. Our objectives were to assess the magnitude of day-to-day alfalfa haylage variation, and evaluate several hand-held NIR instruments for precision and accuracy of currently available calibrations for dry matter (DM) compared to a state-of-the-art electronic moisture meter, and to evaluate forage nutritive value calibrations available on one instrument. The AuroraNir, NIR4, and SCiO handheld NIR instruments were evaluated, along with the HST-1 electronic impedance moisture probe. Alfalfa and alfalfa-grass haylage and total mixed ration samples were collected from farms in NY and WI. After scanning, two subsamples of each sample were dried to determine oven DM. Samples from three farms were measured 10 times each, with remixing between measurements, to evaluate precision. The Aurora and NIR4 instruments, with a similar NIR scanning range, were similar in precision for estimating oven DM, while the SCiO and HST-1 instruments were less precise. Using currently available calibrations, accuracy of DM estimations was in the order of AuroraNir > NIR4 >> SCiO > HST-1, although current calibrations were not sufficiently robust for a range of mixed haylage and TMR samples from NY and WI. Calibrations for the AuroraNir did not produce consistently accurate results for nutritive value determinations, particularly when dealing with mixed species haylage.
Introduction:

A majority of the diet DM fed to lactating dairy cows is typically comprised of forages, and alfalfa or alfalfa-grass are commonly used. Daily variation in forage nutritive value can result in transient nutrient deficiencies, and this risk is often mitigated by formulating rations with excess nutrients (Sniffen et al., 1993). Relatively large transient changes in silage dry matter, resulting in dairy rations of 49% and 55% forage, did not affect milk production if cows were offered excess feed (McBeth et al., 2013). Including safety factors in dairy rations by providing excess nutrients, however, is rapidly becoming economically and environmentally unacceptable. Moisture content can vary considerably from day to day, leading to the development of several options for real-time measurement of silage moisture. Within-farm variation in nutritive value of forages can be large, particularly when alfalfa-grass mixtures are included in the ration (Yoder et al., 2013). Most NY dairy farmers include a mixed haylage in rations, as over 80% of the alfalfa acreage in NY is sown with perennial grass (Cherney et al., 2020). A few NY dairy farmers re-balance rations daily based on forage DM determinations, while many farmers re-balance weekly based on forage DM as well as forage composition.

Near infrared reflectance (NIR) spectroscopy is currently used extensively for routine forage analysis in commercial laboratories. Recently portable spectroscopy has evolved such that successful handheld instruments can be designed for non-scientist operators (Crocombe, 2018). The drastic reduction in spectrometer size and weight coupled with good performance and high-volume manufacturability has resulted in practical analytical applications in a variety of fields (Yan and Siesler, 2018). Interest in the use of miniaturized hand-held spectrometers for on-farm analysis applications is rapidly increasing (Bec et al., 2020). Several companies currently produce hand-held near infrared reflectance (NIR) instruments calibrated for various forage nutritive value parameters, enabling the possibility of on-farm forage analysis and daily rebalancing of dairy rations. Portable spectrometers are now routinely installed on forage harvesting equipment and have been adapted to other field equipment such as liquid manure applicators (Bedord, 2019). A key component to any successful NIR application is a robust calibration.

The goal of handheld NIRs for forage use is to analyze samples without any preprocessing, and calibrations are developed using wet, chopped forage. In addition to high moisture content and large particle size forage samples, other challenges for handheld NIRs include signal-to-noise ratio issues, a variable portable power supply, and operation in unfavorable environments. In general, there is a paucity of refereed information available on the precision and accuracy of available calibrations for hand-held NIR units. To our knowledge, there have been no published refereed evaluations of hand-held NIR calibrations for forage analysis. Our objective was to evaluate several hand-held NIR instruments for precision and accuracy of currently available calibrations for dry matter (DM) and forage nutritive value.

Materials and Methods:

Initially, a set of over 600 alfalfa samples were used to evaluate the relative success of handheld NIR instruments compared to a laboratory NIR, using dry, ground samples. The sample sets
were compiled from eight alfalfa experiments in 2015, 2016 or 2017 at sites in central, western and northern New York State. Ten alfalfa cultivars were included. Depending on the goals of each particular study, alfalfa was in either pure or mixed stands, and samples were collected from early spring to late fall. Forage samples were hand-harvested at a 10-cm stubble height from all plots for nutritive value determinations. Samples from mixed alfalfa-grass plots were separated into pure alfalfa for analysis.

All samples were dried in forced-air ovens to a constant weight at 60°C, and ground in a Wiley mill. Samples were analyzed using wet chemistry procedures described in Valentine et al. (2019), using sodium sulfite in the neutral detergent solution. Forages were weighed into ANKOM F57 filter bags (ANKOM Technology, Macedon, New York, USA) for aNDF, ADF, ADL and 48 h in vitro digestibility analyses. Neutral detergent fiber digestibility (NDFD) was calculated as the proportion of the total fiber digested, expressed on an NDF basis. Total ash content was determined by heating samples to 510°C for four hours. Nitrogen was determined using a combustion process (LECO CN628 analyzer, DairyOne, Ithaca, NY) and CP was calculated as N x 6.25 (AOAC, 1995). All analyses were conducted in duplicate, with the exception of nitrogen, which was determined in duplicate on a subset of samples to calculate a standard error of the laboratory for CP. The standard error of the laboratory (SEL) was determined according to the equation SEL = \sqrt{(\text{sum}(d1-d2)^2)/2n}, where d1 = duplicate 1, d2 = duplicate 2, and n = number of samples.

Alfalfa haylage and total mixed ration (TMR) samples (2 to 5 kg) were collected from 19 farms in NY and WI during 2019 and 2020. A total of 375 samples were collected for this analysis. The mixed species haylage ranged from mostly alfalfa to mostly grass. Besides random sampling of bunkers on NY and WI farms, bunker silo samples alfalfa and alfalfa-grass haylage were collected daily from seven dairies located in central and western NY. Daily samples of alfalfa haylage and TMR were collected for two two-week periods on four of the farms, starting Jan. 27, 2020 and Feb. 20, 2020. Also, on Jun 30, 2020, four separate samples were collected of haylage and TMR at these four farms. Samples of haylage were collected daily for one two-week period on three additional farms, starting Dec. 3, 2019, Jan. 22, 2020, and Feb. 13, 2020. The goal of daily sampling was to determine variability in haylage, to assess the potential benefits of daily forage analysis.

The TMR samples were scanned the day of collection, some of the haylage samples were vacuum packed in oxygen-limiting polyethylene bags using a commercial vacuum packing machine (Cherney et al., 2007) for scanning at a later date. All samples were mixed well in a large plastic bin, prior to scanning and subsampling. After scanning with NIR units and measuring moisture with the moisture probe, two subsamples (approx. 250g each) for each sample were oven dried (60°C) for 48h. One of the two subsamples was ground in a Wiley mill (Thomas Scientific, Swedesboro, NJ) to pass a 1-mm sieve. Three handheld NIR instruments were compared to a state-of-the-art electronic moisture meter. We decided not to attempt studies with dry hay until we had confidence that an instrument would work with silage. Dry hay is considerably more problematic to validate NIR readings.
AuroraNir (Grainit s.r.l., Italy). Separate calibrations for haylage and TMRs were available. Prior to scanning every sample, the lens was covered to prevent external light from entering the device and an internal standard was scanned. Sample was compressed in a rectangular box (4 x 16 inches, 5 inches deep) and the instrument was moved along the forage surface during scanning, maintaining continuous close contact with the surface. Four scans per sample were averaged by the instrument. The abbreviation ‘Aurora’ is used throughout this discussion.

NIR4 (AB Vista, Marlborough, Wiltshire, UK). The only calibration available for testing was a universal moisture calibration (DairyOne, Ithaca, NY), applied to haylage and TMR samples. Prior to scanning every sample, the probe window was covered and an internal standard was scanned. Sample was placed in a round container approximately 10 inches in diameter and 6 inches deep, compressed, and the NIR probe was pressed to the surface, sliding along the surface while maintaining continuous close contact with the sample. Five scans per sample were averaged by the instrument.

SCiO (Consumer Physics, Israel). Moisture calibrations for mixed haylage were provided by REVEAL (Cargill Inc. Minneapolis, MN). Prior to using the SCiO each day, the internal calibration program was run with the lens covered. Sample was compressed in a rectangular box. A plastic cowling was placed over the end of the instrument, maintaining approximately a 2 cm distance between the sample surface and the sensor window, and excluding external light from the sensor. The SCiO did not have a TMR calibration and TMR samples were evaluated using a mixed haylage calibration. Seven scans per sample were averaged by the instrument.

HST-1 (AgraTronix LLC, Streetsboro, OH) is an electronic impedance 20-inch probe with hay silage calibrations and includes both density and temperature compensation. Sample was compressed in a container, and the probe was inserted (minimum of 8 inches) and internal forage temperature was measured prior to moisture determinations to compensate for temperature. Forage density was automatically compensated for based on the force required to insert the probe. HST-1 did not have a TMR calibration and TMR samples were evaluated using the “hay silage” calibration. Three probe readings per sample were averaged by the instrument. The abbreviation ‘HST’ is used throughout this discussion.

Calibration issues. In order to make a reasonable assessment of calibrations, we attempted to avoid automatic updating of calibrations if possible. The Aurora and NIR4 were not connected to the internet during the study. The NIR4 automatically updates calibrations when connected to the internet. Initially we had included the X-NIR instrument in the study, but it required regular us to regularly send samples to DairyLand Labs for expensive analysis and updating the calibrations, which made it difficult to compare the X-NIR with other instruments. The SCiO only operates by connection to the internet, so we cannot be certain that calibrations were not updated during the experiment.

Precision estimates for dry matter. Haylage and TMR samples were collected from three farms and all samples were mixed thoroughly in a large plastic container. A representative subsample was transferred to a separate container for each scan or probe, and the subsample was returned to the plastic container afterwards for remixing. This process was repeated 10 times for each
instrument, generating 10 separate moisture determinations for each sample/instrument combination. Two subsamples from each sample were dried at 60º C for 48 h to determine oven dry matter (DM).

**Accuracy estimates.** All samples were mixed well in a large plastic bin, prior to scanning and subsampling for laboratory analysis. The success of NIR predictions were assessed based on the coefficient of determination ($R^2$), root mean square error of prediction (RMSEP), residual prediction deviation (RPD) and the ratio of RMSEP to the range of the reference values (RER) (Foster et al., 2013; Ward et al., 2011). RMSEP = $\sqrt{\text{sum}(R1-P1)^2/n}$, where $R1 =$ reference data set, $P1 =$ predicted data set, and $n =$ number of samples. RPD = reference data set standard deviation (SD) / RMSEP. RER = Range of the reference data set / RMSEP.

**Project Objectives**

1. Evaluate accuracy of handheld NIR instruments on dry, ground alfalfa.
2. Evaluate day-to-day variability in alfalfa and alfalfa-grass DM and composition in bunkers.
3. Evaluate precision and accuracy of handheld NIR instruments on alfalfa haylage and on total mixed rations which include alfalfa.

**Corresponding Results**

1. Spectrometers likely need to scan at least up to 1650nm of the near infrared region for forage quality estimates.
2. Some portable spectrometers can produce accurate results using dry, ground alfalfa samples.
3. None of the handheld NIR instruments evaluated generated accurate DM estimates of wet chopped alfalfa haylage.

**Results and Discussion:**

**Objective 1.** This study compared the performance of prediction of three different portable instruments compared to a typical laboratory NIR instrument, using a wide range of dried, ground alfalfa samples. NIR spectra scanned are shown in Fig. 1. Laboratory instrument scans were replicated with a reduced spectral range to match the range of each portable instrument. Portable instruments evaluated did not scan the upper portion of the spectral range (1652 – 2498mn), and the upper portion did impact forage calibration.

The SCiO instrument scanned a very narrow range (740 – 1070nm) and although it had comparable results to the laboratory instrument constrained to same wavelength range, most major peaks related to forage quality traits are outside this range. The expensive laboratory instrument had the best performance as expected, while the very inexpensive SCiO instrument had much greater error of predictions to the point that, for most traits, the prediction would not be considered reliable. However, the AuroraNir and instruments with a similar scanning range may provide an alternative to expensive laboratory equipment, while still providing sufficiently accurate predictions.
Objective 2. The seven participating dairy farms were relatively consistent in their use of haylage in TMR rations, however, the alfalfa% in haylage mixtures varied widely, and likely contributed to more day-to-day variability compared to haylage containing a single forage species. Within-field variation in grass% in mixtures is often large, and grass% in alfalfa-grass mixtures was estimated by producers in this study. The size of the seven dairies sampled daily ranged from 300 to over 3000 cows, and size of the bunker silos varied proportionally to the number of cows. Although there was a wide range in alfalfa% in haylage mixtures across farms, none of the farms stood out as exceptionally more consistent or more variable in haylage or TMRs day-to-day. The standard error of the laboratory (SEL) was determined through duplicate analyses for all parameters measured, with results as good or better than other published SEL values. Haylage SEL was considerably larger than for TMRs. Variability in haylage analysis may have been influenced by the fact that all haylages were mixtures of legumes and grasses.

For the purposes of this analysis, we assumed that a typical NY dairy farmer would rebalance their TMR weekly. Therefore, daily sampling data was divided up into 7-day increments, to assess the weekly variation in nutritive parameters. With 24 weeks of haylage samples and 16 weeks of TMR samples we achieved a relatively uniform distribution in the weekly range for all parameters. The weekly range distribution for alfalfa haylage, sorted from smallest to largest is shown in Fig. 2.

With consistent sampling and duplicate laboratory analyses, the day-to-day variation in parameters reported here was relatively independent of sampling and analysis variation. Day-to-day variation during a 7-d period ranged from minimal to substantial (Fig. 2). Haylage was more variable on a day-to-day basis than TMRs for all parameters measured. Haylage is more difficult to consistently ensile at the appropriate moisture content than corn silage, and haylage also had the most day-to-day variability in DM, similar to results of Stone et al. (2008).

Focusing on DM, our results indicate that haylage is considerably more variable than TMRs. In our study some weeks would likely benefit from daily rebalancing, while some weeks would not. Based on multiple sample collection at some sites, DM determinations can vary up to +/- two% units due to sampling alone. If we assume that a 5% unit DM weekly range in haylage is great enough to benefit from daily rebalancing of rations, that threshold should probably be increased to 7% units to account for variability in DM determination. A 7% unit threshold for weekly range in DM was exceeded 42% of weeks for alfalfa haylage and 25% of weeks for TMR. Understanding the range of day-to-day variability across weeks is essential to evaluate whether daily NIR measurements followed by re-balancing rations is practical.

Objective 3. Multiple alfalfa haylage DM estimates were consistent for the Aurora and NIR4 (Fig. 3). Aurora haylage DM estimates were slightly above or below oven DM, while NIR4 DM estimates were consistently underestimating oven DM for this set of samples. SCiO haylage DM estimates were less consistent than Aurora and NIR4 for Farm C. The HST probe estimates of haylage DM tended to be more variable than repeated NIR estimates (Fig. 3).

The Aurora was consistent and relatively accurate in estimating TMR oven DM for this small set of samples (Fig. 4). The NIR4 DM estimates for TMR also were very consistent but several percentage units below that of oven DM. Neither SCiO nor HST were calibrated for TMRs, so
“mixed haylage” (SCiO) and “hay silage” (HST) calibrations were applied to TMR samples (Fig. 4). Both SCiO and HST DM estimates for TMRs were variable and well below that of oven DM. A bias was not surprising when using haylage calibrations for TMR samples, but this would not necessarily have reduced precision, which was considerable.

Averaging standard deviations (SD) for DM for three farms, the Aurora and NIR4 were similar in precision across haylage and TMR samples (Fig. 5). Estimates of DM with SCiO and HST, however, lacked precision with large SD that were variable from farm to farm. Using haylage calibrations for the TMRs for SCiO and HST resulted in SD somewhat greater than for haylage.

The NIR region with the strongest absorption of water is around 1930 nm which is outside the range of all three NIR instruments evaluated. Secondary water NIR absorption peaks between 1400 and 1650 nm are captured by the Aurora and NIR4 but not the SCiO. The Aurora and NIR4 instruments were similar in precision, but the Aurora was somewhat more accurate (Fig. 6). The data pattern was similar for the Aurora and NIR4, and any sample groups that tended to be outliers did so for both instruments. The SCiO was not as accurate as the Aurora and NIR4 for estimating alfalfa haylage DM, and the HST electronic moisture meter readings had almost no relationship to oven DM (Fig. 7).

The Aurora was the only instrument tested with calibrations for NDF, ADF, CP and Ash. Both NDF and ADF were similar in their estimates, compared to wet chemistry values (Fig. 8). Crude protein (CP) and Ash were also similar (Fig. 9), the Aurora overestimated smaller lab values and underestimated larger lab values.

Performance criteria for calibrations have been suggested (Malley et al., 2005; Ward et al., 2011) (Table 1). The coefficient of determination ($R^2$) is generally not a good measure of model performance because it is highly dependent on the range of the data set (Foster et al., 2013). The RPD and RER both relate RMSEP to the range of reference data, using either the standard deviation or the range of values for the reference data. The Aurora DM calibration was considered moderately useful based on RPD and RER statistics (Table 2), but none of the other instruments met any of the RPD or RER thresholds for haylage DM. The Aurora calibration was considered moderately useful for NDF based on RPD, but ADF, CP and Ash calibrations were considered unsuccessful.

**Acknowledgements:** Funding for this study was provided by the U.S. Alfalfa Farmer Research Initiative of the National Alfalfa & Forage Alliance, and by the NY Farm Viability Institute, State of NY. Thanks to Paolo Berzaghi and Mike Casler for assisting with the evaluation of hand-held NIR instruments using dry ground alfalfa samples.

**References:**


http://dx.doi.org/10.1177/0960336020916815.


**Keywords:** forage variability, total mixed ration, silage, feed sampling, feed composition
Fig. 1. Average alfalfa spectra with three instruments. LAB = FOSS 6500, DA = Aurora, SW = SCiO. The NIR4 spectrum (not shown) is very similar to Aurora.
Fig. 2. Weekly range in alfalfa haylage composition over 24 weeks of daily bunker sampling, sorted from smallest to largest weekly range.
Fig. 3. Ten measurements of alfalfa haylage DM content per sample from three dairy farms. Aurora, NIR4, and SCiO NIR units, compared to HST-1 electronic moisture probe. Oven DM (blue lines) was measured in duplicate for each sample.
Fig. 4. Ten measurements of total mixed rations (TMR) DM content per sample from three dairy farms. Aurora, NIR4, and SCiO NIR units, compared to HST-1 electronic moisture probe. Oven DM (blue lines) was measured in duplicate for each sample.
Fig. 5. Precision of 10 DM measurements of alfalfa haylage and total mixed ration (TMR) samples from three farms. Vertical bars are standard deviations of values.
Fig. 6. Aurora and NIR4 DM estimates compared to oven DM determinations for alfalfa haylage (red) and total mixed rations (TMR, green).
Fig. 7. SCiO and HST estimates compared to oven DM determinations for alfalfa haylage (red).
Fig. 8. Aurora NDF and ADF estimates compared to laboratory determinations for alfalfa haylage (red) and total mixed rations (TMR, green).
Fig. 9. Aurora CP and Ash estimates compared to laboratory determinations for alfalfa haylage (red) and total mixed rations (TMR, green).
Table 1. Validation performance criteria according to Malley et al. (2005).

<table>
<thead>
<tr>
<th>Level of success</th>
<th>R²</th>
<th>RPD</th>
<th>RER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>&gt; 0.95</td>
<td>&gt; 4.0</td>
<td>&gt; 20</td>
</tr>
<tr>
<td>Successful</td>
<td>&gt; 0.90 to 0.95</td>
<td>&gt; 3.0 to 4.0</td>
<td>&gt; 15 to 20</td>
</tr>
<tr>
<td>Moderately successful</td>
<td>&gt; 0.80 to 0.90</td>
<td>&gt; 2.25 to 3.0</td>
<td>&gt; 10 to 15</td>
</tr>
<tr>
<td>Moderately useful</td>
<td>0.70 to 0.80</td>
<td>1.75 to 2.25</td>
<td>8 to 10</td>
</tr>
</tbody>
</table>

a R², coefficient of determination; RPD, residual prediction variation; RER, range of the reference data divided by root mean square error of prediction (RMSEP).

Table 2. Success of calibration predictions for alfalfa haylage and total mixed rations (TMR).

<table>
<thead>
<tr>
<th></th>
<th>Aurora</th>
<th>NIR4</th>
<th>SCiO</th>
<th>HST</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>aNDF</td>
<td>ADF</td>
<td>CP</td>
<td>Ash</td>
</tr>
<tr>
<td>Haylage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.81c</td>
<td>0.50</td>
<td>0.55</td>
<td>0.31</td>
</tr>
<tr>
<td>RMSEP</td>
<td>5.06</td>
<td>3.74</td>
<td>3.08</td>
<td>1.39</td>
</tr>
<tr>
<td>RPD</td>
<td>1.82c</td>
<td>1.35</td>
<td>1.25</td>
<td>1.21</td>
</tr>
<tr>
<td>RER</td>
<td>1.08</td>
<td>5.61</td>
<td>5.59</td>
<td>5.76</td>
</tr>
<tr>
<td>n</td>
<td>263</td>
<td>263</td>
<td>263</td>
<td>263</td>
</tr>
<tr>
<td>TMR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.03</td>
<td>0.04</td>
<td>0.19</td>
<td>0.01</td>
</tr>
<tr>
<td>RMSEP</td>
<td>3.87</td>
<td>2.93</td>
<td>1.03</td>
<td>1.31</td>
</tr>
<tr>
<td>RPD</td>
<td>0.28</td>
<td>0.27</td>
<td>0.86</td>
<td>0.25</td>
</tr>
<tr>
<td>RER</td>
<td>1.37</td>
<td>1.42</td>
<td>3.98</td>
<td>1.63</td>
</tr>
<tr>
<td>n</td>
<td>112</td>
<td>112</td>
<td>112</td>
<td>112</td>
</tr>
</tbody>
</table>

a aNDF, neutral detergent fiber; ADF, acid detergent fiber; CP, crude protein; DM, dry matter.
b R², coefficient of determination; RMSEP, root mean square error of prediction; RPD, residual prediction deviation; RER, range of the reference data divided by RMSEP.
c Meets one of the validation performance criteria listed in Table 1.