Misconceptions in near-infrared calibration development: Understanding crop year and regional variability

Nicole Schlau, Phillip Goldblatt, and Kyle Taysom

Dairyland Laboratories, Inc, Arcadia, WI, USA

1 BACKGROUND

- Early commercial near-infrared reflectance spectroscopy (NIR) calibrations relied on localized datasets (Mahanna, 2008) which raised concerns about transferability across regions
- Current best practices emphasize diverse datasets across crop years, regions, and hybrids to improve calibration reliability and applicability (Shaver, 2005; Abrams et al., 1987)
- Larger commercial laboratories now develop and license calibrations built from diverse datasets to smaller, regional laboratories
- Despite this, concerns about cross-regional transferability persist in the industry (Husmoen and Kern, 2023)
- This study evaluated the accuracy of NIR calibrations for aNDFom and starch in corn silage across crop years and regions.

2 MATERIALS AND METHODS

Sample collection

- A dataset of 2,127 corn silage samples submitted during the 2021–2023 crop years (August 2021 – July 2024) was obtained from Dairyland Laboratories, Inc (Arcadia, WI, USA).
- Samples were previously analyzed for aNDFom (AOAC method 2002.04; AOAC International, 2023) and starch (acetate buffer method; Hall, 2015).
- Spectra were collected every 2 nm over 1100 – 2500 nm to predict aNDFom and starch using NIR calibrations developed by Dairyland Laboratories, Inc.
- Predictive performance of the NIR calibrations was evaluated across crop years using a subset of 50 samples selected per year from the Midwestern region of the United States (Illinois, Indiana, Michigan, Minnesota, Ohio, and Wisconsin) and across regions using a subset of 20 samples from China, Europe, the Middle East, and North and South America.
- Samples were selected to evenly represent normal ranges for aNDFom (25 – 50% of DM) and starch in corn silage (18 – 48% of DM) and were not previously used for NIR calibration development.

Statistical analysis

- Standard error of predictions (SEP) were compared across crop years and regions using F-test (Snedecor & Cochran, 1989).
- Slopes of reference chemistry values vs. NIR predictions were tested for deviations from 1 and for differences by crop year or region using the 'lstrends' function from the 'emmeans' package (Lenth, 2022) in R 4.2.2.
- Residuals (predictions - reference) were analyzed with 'lm' to assess biases by crop year or region. Type III tests were calculated with the 'Anova' function of the 'car' package (Fox et al., 2021) and group differences were evaluated using 'emmeans'.
- Bias within each group was also tested for deviation from 0.
- Significance was declared at $p \leq 0.05$ and tendencies are discussed when $0.05 < p \leq 0.10$.

3 RESULTS

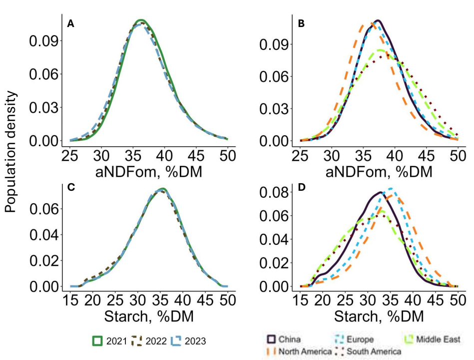

Figure 1. Distributions of aNDFom (A, B) and starch (C, D) predicted by NIR for corn silage across crop years (A, C) or global regions (B, D)

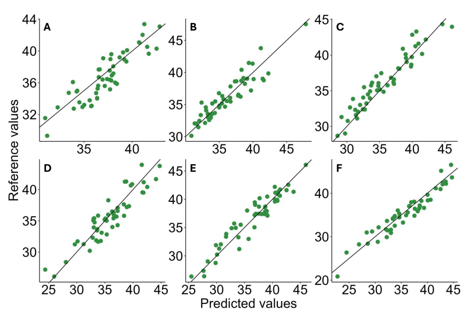

Figure 2. aNDFom (A, B, C) and starch (D, E, F) predicted by NIR for corn silage across the 2021 (A, D), 2022 (B, E), or 2023 (C, F) crop years. The black line is a 1:1 bisector where reference values = predicted values.

Table 1. Validation statistics by crop year (n = 50 per year; August to July each year) for corn silage

	Crop year			SE[1]	P-Value[2]
	2021	2022	2023		
aNDFom					
SEP	1.49	1.31	1.38	-	0.19
Slope	1.01	0.90	0.90	0.06	0.27
Bias	-0.50[‡a]	-0.01[a]	0.59[b,*]	0.16	< 0.001
R^2	0.79	0.85	0.91	-	-
Starch					
SEP	1.70	1.45	1.48	-	0.13
Slope	0.86	0.92	0.93	0.04	0.38
Bias	-0.38	-0.19	0.13	0.19	0.26
R^2	0.84	0.90	0.92	-	-

[1,2] Lowest standard errors (SE) and P-Values are presented within a comparison

[a,b] Different values within a row differ ($p < 0.05$) or [‡] tend to differ ($0.05 < p \leq 0.10$)

[‡] Tend to differ from 0 ($0.05 < p \leq 0.10$) or [*] differ from 0 ($p < 0.05$)

Table 2. Validation statistics by region (n = 20 per region) for corn silage

	Region[1]					SE[2]	P-Value[3]
	China	EUR	ME	NA	SA		
aNDFom							
SEP	1.52	1.37	1.37	1.65	1.60	-	0.21
Slope	0.96	1.03	0.98	0.95	0.97	0.08	0.99
Bias	1.17[*]	0.64	0.41	0.26	0.42	0.28	0.09
R^2	0.93	0.91	0.91	0.82	0.90	-	-
Starch							
SEP	1.22	1.58	1.51	1.61	1.31	-	0.12
Slope	0.89	1.04	0.93	0.93	1.04	0.07	0.45
Bias	-0.17	0.28	-0.54	0.46	-0.06	0.30	0.23
R^2	0.93	0.91	0.91	0.94	0.96	-	-

[1] Europe (EUR), Middle East (ME), North America (NA), South America (SA)

[2,3] Lowest standard errors (SE) and P-Values (P) are presented within a comparison

[*] Bias differed significantly from 0 ($p < 0.05$)

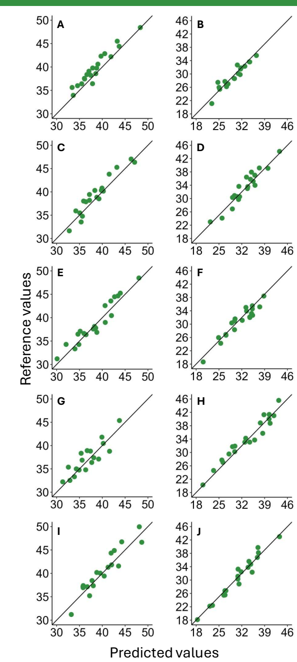

Figure 3. aNDFom (A, C, E, G, I) or starch (B, D, F, H, J) predictions by NIR for corn silage across China (A, B), Europe (C, D), the Middle East (E, F) North America (G, H), and South America (I, J). The black line is a 1:1 bisector where reference values = predicted values.

4 SUMMARY

- For corn silage, variation for aNDFom and starch across crop years within the Midwestern United States was minimal, while differences across global regions were more pronounced (Figure 1)
- Slopes for aNDFom did not differ among crop years (Figure 2) or regions (Figure 3), and none differed significantly from 1
- Bias for aNDFom was highest for 2023 and tended to be lowest for 2021 for Midwestern corn silage ($p = 0.08$, Table 1). For 2023, bias was greater than 0, while for 2021, bias tended to be less than 0 ($p = 0.08$).
- aNDFom bias for China was greater than 0 ($p < 0.05$, Table 2) but only tended to be higher for China than the other regions
- For starch, slopes did not differ among crop years or regions, but the slope for 2021 Midwestern corn silage was less than 1 ($p < 0.05$)
- Biases for starch did not differ among crop years or regions, and none differed significantly from 0
- For regional comparisons, alternative drying or grinding methods adopted based on local resource availability may have introduced variability at affiliate laboratories, potentially contributing to the perception of regional bias.
- Concerns about cross-regional transferability of NIR calibrations may stem from differences in reference methods, similar but distinct nutrient definitions (ex. aNDF vs. aNDFom), or vague sample descriptions that lead to misapplied calibrations

5 CONCLUSIONS

- NIR prediction errors did not vary across crop years or regions for corn silage
- Apparent regional variation in NIR prediction accuracy is more likely due to inconsistencies in reference methods or improper sample preparation (drying or grinding) than to limitations in calibration transferability
- Consistent reference methods and proper sample preparation are just as critical for assessing NIR transferability as they are for comparing chemistry results directly

REFERENCES

Abrams, S. M., Westerhaus, M. O., Shenk, J. S., & Barton, F. E. II. (1987). Determination of forage quality by near infrared reflectance spectroscopy: Efficacy of broad-based calibration equations. Journal of Dairy Science, 70(4), 806–813.

Association of Official Analytical Chemists. (2023). Official methods of analysis (22nd ed.). Arlington, VA, USA.

Fox, J., S. Weisberg, and B. Price. 2021. Companion to applied regression. R package version 3.0-12.

Hall, M. B. (2015). Determination of dietary starch in animal feeds and pet food by an enzymatic-colorimetric method: Collaborative study. Journal of AOAC International, 98(2), 397–409.

Kern-Lubey, R., & Anderson Husmoen, B. J. (2023). Hesitation about NIRS evaluation of grass forages. Progressive Forage. https://www.agproud.com/articles/58270-hesitation-about-nirs-evaluation-of-grass-forages

Lenth, R. 2022. Emmeans: Estimated marginal means, aka least squares means. R package version 1.7.2.

Mahanna, B. (2008). NIRS Analysis Has Long and Credible History. Pioneer. https://www.pioneer.com/us/agronomy/nirs.html

Shaver, R. D. (2020). Evaluating corn silage quality for dairy cattle. University of Wisconsin—Extension. https://fyi.extension.wisc.edu/forage/files/2016/10/EvaluatingCornSilageforDairy-1.pdf

Snedecor, G. W., & Cochran, W. G. (1989). A test of the equality of two variances. In Statistical methods (8th ed.). Ames, IA: Iowa State Press.

ACKNOWLEDGEMENTS

The authors extend their gratitude to the technicians at Dairyland Laboratories for their assistance analyzing samples, Nikolaus Lettner for his contributions to NIR spectral collection, and to Dr. David Mertens for technical review.