

“Is near infrared spectroscopy only as good as the laboratory reference values?” An empirical approach

D.B. Coates

CSIRO Sustainable Ecosystems, Davies Laboratory, PMB PO Aitkenvale, Qld 4814, Australia

Editor's Note

I was very fortunate to be invited to the 10th Australian National Near Infrared Users Group Conference, which was held at Coffs Harbour, NSW, Australia, 7–8 May 2002. Among several very interesting papers, this paper by David Coates particularly caught my attention and I am very pleased that David agreed to it being published in *Spectroscopy Europe*. The idea that Near Infrared (NIR) spectroscopic results can be better than the reference chemistry is statistically sound but as the reference chemistry is normally the arbiter in deciding how well the NIR method is performing it is difficult to persuade users that this can be true. This original approach by David Coates should persuade even the most hardened of the “doubting Thomases”.

Do I need to assure other authors that the selection of this paper was not connected with the author's address? These coincidences just happen!

Tony (A.M.C.) Davies

Introduction

Accuracy in the determination of laboratory reference values for use in the development of NIR spectroscopic calibration equations has been accepted as a critical component of useful NIR technology and conventional wisdom dictates that the accuracy of NIR predictions can only be as good as the laboratory reference values used for calibration. Because absolute certainty concerning the determination of an error-free analyte value is an impossibility, confirmation of such a dictum presents real difficulties. This paper, however, describes an empirical approach to provide some insight into the question. The basis of the approach is the deliberate addition of error to the laboratory reference values.

Methods

A calibration set of 100 forage samples was selected from approximately 700 samples that had been previously analysed for nitrogen (N) concentration. Laboratory reference values determined by the primary analytical method (N_{E0}) were regarded in this exercise as containing zero error (nominally error-free). Nitrogen concentration ranged from 0.22 to 4.50%. Error sets were then derived by adding random error to the N_{E0} values. Within each set, error was normally distributed with a mean of zero and standard deviations (*SD*) of 0.10 (N_{E1}), 0.15 (N_{E2}) and 0.20 (N_{E3}). Three error-sets (a), (b) and (c) were developed within each *SD* category. NIR spectra were obtained by scanning the samples with a Foss NIRSystems 6500 spectrometer (Foss NIRSystems, Silver Spring, MD, USA) fitted with a spinning sample cup holder.

Calibration equations (modified partial least squares regression (MPLS);¹ wavelengths 700–2500 nm; math treatment 1,4,4,1; standard normal variates (SNV) with detrend;² maximum of seven terms) were developed for the error-free reference value set (N_{E0} -EQA) and the nine sets with added error ($N_{E1(a)}$ -EQA; $N_{E1(b)}$ -EQA; $N_{E3(c)}$ -EQA). Linear regression statistics were determined for relationships between:

- lab reference sets containing added error ($N_{E1(a)}$; $N_{E1(b)}$; $N_{E3(c)}$) and the lab reference set without error (N_{E0})
- NIR spectroscopic predictions derived from the various calibration equations and the N_{E0} reference values and
- NIR spectroscopic predictions derived from the error-affected calibration equations $N_{E1(a)}$ -EQA –

$N_{E3(c)}$ -EQA and predicted values derived from N_{E0} -EQA.

Regression standard errors (*SE*) and coefficients of determination (*RSQ*) were used as the measures of accuracy.

Results and discussion

Table 1 (Column A) presents statistics for the regression of lab reference values with added error on the error-free reference values (N_{E0}). The regression standard errors are, of course, identical with the *SD* of the added error in each set while the *RSQ* values are dependent on the range of analyte values, which is the same for each set, and the level of added error.

The *RSQ* and standard error of calibration (*SEC* where $SEC = SE$) for calibration equation N_{E0} -EQA are acceptable but as the level of added error increased these calibration statistics rapidly deteriorated [Table 1 (Column B)].

Some of these results are shown graphically in Figures 1–3.

A comparison of the regression statistics of predicted values regressed on the error-free reference values (N_{E0}) [Table 1 (Column C)] with the comparable statistics in Table 1 (Column A) and Table 1 (Column B) is enlightening. At the higher levels of added error (levels E2 and E3), the correlations between the values predicted by the relevant error-affected calibration equations and the nominally error-free reference values were closer than the corresponding correlations between the error reference sets and the error-free reference set. This was particularly noticeable at the highest level of added error. Thus, on average, the predicted values were closer to the error-free reference values than were the error-loaded values used in developing the

Table 1. Linear regression coefficients of determination (*RSQ*) and standard errors (*SE*) for: (A) Error reference sets regressed on the error-free reference values. (B) Predicted values regressed on reference value for each calibration equation. (C) Predicted values from each calibration equation regressed on error-free reference values. (D) Predicted values from each calibration equation regressed on predictions from N_{E0} .EQA.

Reference values	(A) Reference values against N_{E0}		(B) Calibration statistics		(C) Predicted values against N_{E0}		(D) Predicted values against predictions from N_{E0} .EQA	
	<i>RSQ</i>	<i>SE</i>	<i>RSQ</i>	<i>SE</i>	<i>RSQ</i>	<i>SE</i>	<i>RSQ</i>	<i>SE</i>
N_{E0}			0.988	0.106	0.988	0.102	1.000	0.000
$N_{E1(a)}$	0.990	0.097	0.974	0.153	0.984	0.120	0.995	0.070
$N_{E1(b)}$	0.989	0.099	0.972	0.160	0.984	0.120	0.995	0.067
$N_{E1(c)}$	0.989	0.105	0.973	0.160	0.986	0.113	0.996	0.043
$N_{E1(mean)}$	0.989	0.100	0.973	0.158	0.985	0.118	0.996	0.060
$N_{E2(a)}$	0.976	0.147	0.975	0.149	0.985	0.113	0.996	0.057
$N_{E2(b)}$	0.976	0.152	0.966	0.180	0.984	0.123	0.996	0.058
$N_{E2(c)}$	0.975	0.149	0.963	0.180	0.979	0.133	0.991	0.089
$N_{E2(mean)}$	0.976	0.149	0.968	0.170	0.983	0.123	0.994	0.068
$N_{E3(a)}$	0.959	0.202	0.938	0.235	0.974	0.155	0.984	0.122
$N_{E3(b)}$	0.945	0.226	0.931	0.252	0.973	0.152	0.986	0.111
$N_{E3(c)}$	0.953	0.207	0.932	0.239	0.968	0.162	0.978	0.135
$N_{E3(mean)}$	0.952	0.212	0.934	0.242	0.972	0.156	0.983	0.123

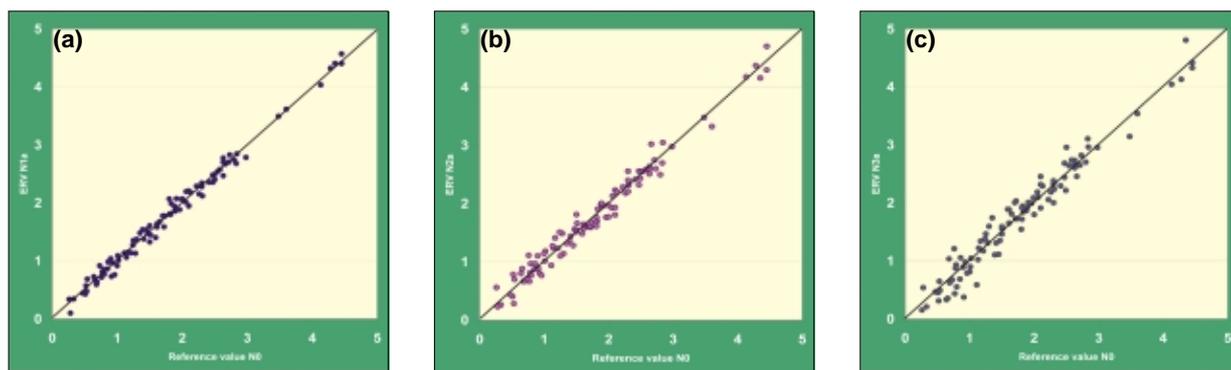


Figure 1. Plots of noise-free reference values plotted against the values with increasing added noise. (a) *SE* 0.10; (b) *SE* 0.15; (c) *SE* 0.20.

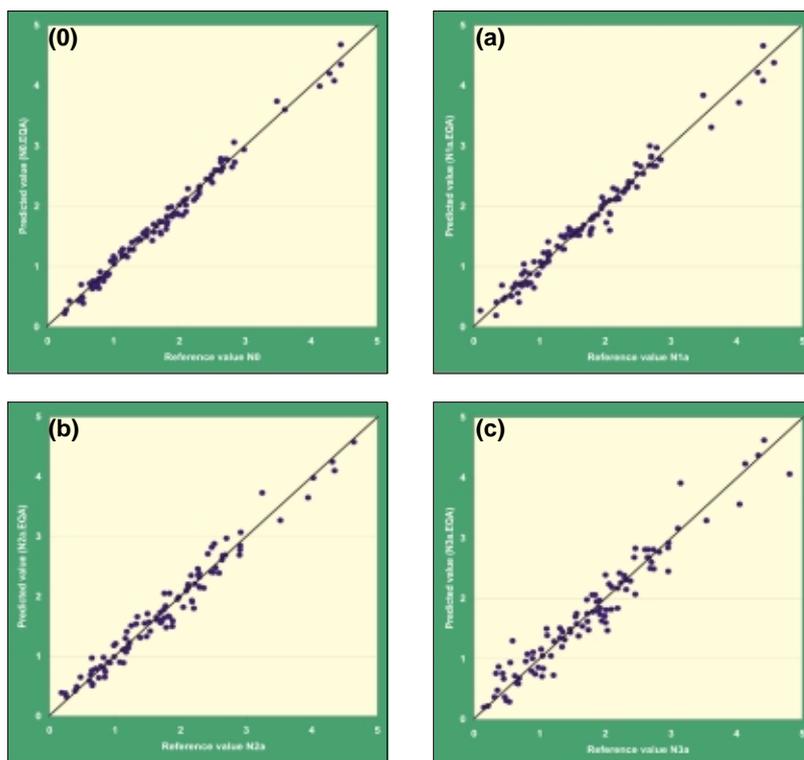


Figure 2. Plots for the calibrations using sets with increasing amount of added noise. (0) original data; (a) 0.10; (b) 0.15 (c) 0.20. SE results are: (0) 0.106; (a) 0.158; (b) 0.170; (c) 0.242.

Table 2. A comparison of prediction errors (reference value N_{E0} – predicted value) with added error for a sub-set of N_{E3} .

Sample rank according to magnitude of added error	Mean of added error (%N)	Mean error of predictions using $N_{E3(a)}$ -EQA (%N)
81–90	0.312	0.150
91–100	0.429	0.136

calibration equations. At all three levels of added error, regression statistics were superior for predicted against error-free reference value (Column C) than for the corresponding calibration equation statistics (predicted values against error reference set, Column B). It seems clear, therefore, that the degree to which the MPLS regression process could be “pulled off course” by erro-

neous reference values was very much limited.

The results in Table 1 (Column D) illustrate how the predictions using the error-affected calibration equations compared with the predictions using the equation developed from the error-free reference values, N_{E0} -EQA. It can be seen that the deterioration in SE and RSQ as the level of added error

increased was far less in the predicted values (Column D) than in the error reference sets themselves (Column A). For want of a better description, this indicated that the NIR spectroscopic calibration equations tended to “pull” the predicted values back towards the error-free reference values. This is illustrated in a different way in Table 2 where added errors are compared with prediction errors at the E3 level of added error.

Summary

The results derived from this exercise clearly illustrate that it is possible for NIR spectroscopic calibration equations to produce predictions that are more accurate than the laboratory reference values used in the calibration set.

For this analyte at least, the closeness of fit between predicted and reference values was clearly correlated with the accuracy of the reference values. It is reassuring to know that the MPLS regression process could not fit a good regression equation to error laden reference data.

The best calibration statistics and the most accurate predictions were aligned with the most accurate reference values. Therefore, although it is possible for NIR spectroscopic predictions to be more accurate than error-laden reference values, it still holds that the more accurate the reference values, the more accurate will be the predictions.

References

1. J.S. Shenk and M.O. Westerhaus, *Crop Sci.* **31**, 469–474 (1991). For a discussion of PLS regression, see Chapter 5 in Reference 2.
2. T. Næs, T. Isaksson, T. Fearn and T. Davies, *A User-Friendly Guide to Multivariate Calibration and Classification*. NIR Publications, Chichester, p. 124 (2002).

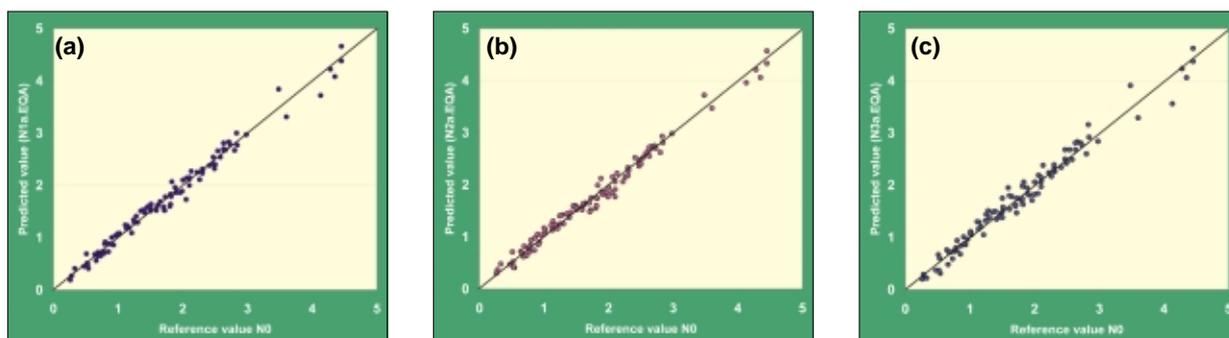


Figure 3. Plots for predictions using calibrations produced with added noise (or error) regressed against error-free reference values (N_{E0}); SE results are: (a) 0.118; (b) 0.123; (c) 0.156. Compared to 0.106 for the calibration with no added noise shown in Figure 2(0).