Instrumentation considerations for robust near infrared applications

David B. Funk

USDA Grain Inspection, Packers and Stockyards Administration, 10383 N. Executive Hills Blvd, Kansas City, MO 64153, USA.

Introduction

Robustness is a universal, almost mystical, goal in near infrared (NIR) work. A robust NIR application would provide acceptably accurate results in the face of the worst conditions the real world could conceive. Every NIR application falls short of that pinnacle of confidence. Calibration development and calibration transfer are two major limitations on NIR application robustness. NIR instruments, though carefully standardised, have residual differences among units—and within any unit over time. A calibration that is accurate for all samples on the master instrument, but extremely sensitive to residual instrumentation differences, is worthless for implementation in a network of instruments. Therefore, calibration development must consider how to minimise sensitivity to instrumentation differences.

The purpose of this paper is to propose an effective strategy for achieving robustness. This route to robustness involves a simple geometric relationship between NIR calibrations and spectra and some mathematical models relating instrumentation characteristics to spectral measurements. Taken together, these concepts link instrumentation effects to their impact on NIR calibrations, allowing the interactions to be assessed and minimised.

Visualising calibrations

Linear calibrations

The first important step is to "see" NIR calibrations from a different perspective. Linear NIR calibrations give the predicted value *(%C)* as sums of products of calibration coefficients (*K*) and the associated (*n*) spectral measurements (*L*).

$$\%C = K_0 + K_1 L_1 + \cdots + K_n L_n \tag{1}$$

Analytical geometry teaches that an equation of the form

$$0 = B + Mx + Ny \tag{2}$$

is the equation of a line in two dimensions.[1] Consider a two-wavelength NIR calibration with Equation 1 slightly rewritten.

$$0 = (K_0 - \%C) + K_1 L_1 + K_2 L_2 \tag{3}$$

Equation 3 is in the general form of a line in two-dimensional spectral space. All points (L_1, L_2) in spectral space that the NIR calibration (K_0, K_1, K_2) assigns to a given predicted value of $\%C$ (i.e. 12.0 %) must lie on a single line in spectral space. Every point in spectral space that is predicted as

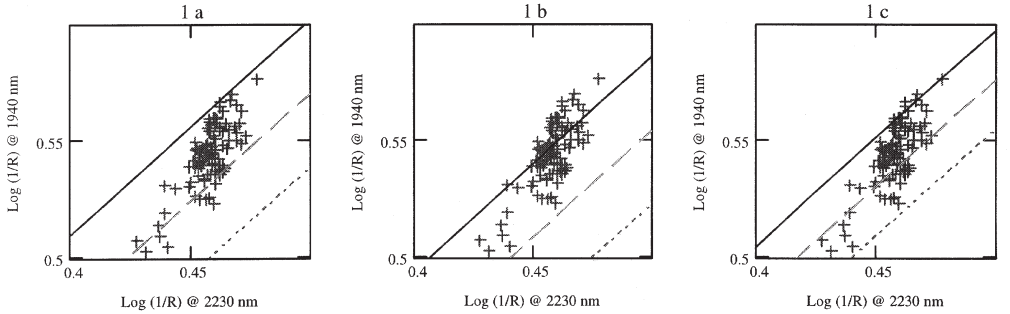

Figure 1. Technicon 400[a] calibration lines of constant predicted moisture (solid, 12%; dash, 10%, dot, 8%) superimposed on Hard Red Spring wheat reflectance data at 1940 nm and 2230 nm. Figure 1(a) shows unadjusted calibration lines. Figure 1(b) shows calibration lines with bias adjustment of +1.0 % moisture. Figure 1(c) shows calibration lines with slope adjustment of 1.5 and bias adjustment of –4.0%.

$\%C = 11.0\%$ lies on another line that is parallel to the first. For any predicted value, there exists a "calibration line" in spectral space parallel to and regularly spaced from the calibration lines associated with all other predicted values. Figure 1(a) illustrates such a set of moisture calibration lines, superimposed on a set of two-wavelength near infrared reflectance data.

Slope and bias adjustments are often used to reduce differences among instruments. It is instructive to observe how these adjustments affect the calibration lines in spectral space. Figure 1(b) shows the calibration lines (of constant predicted values) with a bias adjustment of +1.0%. The calibration lines are offset relative to the stationary data. Figure 1(c) shows calibration lines with a slope adjustment of 1.5 (and a compensating bias adjustment to keep the calibration lines on the same scale). Slope adjustment changes the spacing of the calibration lines, but it does not change the direction of the calibration lines.

It can be seen that any change to the spectral data that causes movement parallel to the calibration lines has no effect on the predicted constituent values. Therefore, the goal of calibration development can be seen as establishing calibration lines that are parallel to the direction of change for all interfering factors and perpendicular to the direction of change resulting from differences in the factor of interest.

At this point, the reader may question the relevance of this simple geometric discussion to "real" NIR calibrations. The relevance emerges when the geometric concept of calibration lines is extended to higher order spectral space. With a three-wavelength calibration, the lines now become parallel *planes* in spectral space. The *hyperplanes* associated with calibrations involving more than three wavelengths cannot be visualised directly, but the conceptual leap from lines (in two dimensions) to planes (in three dimensions) to hyperplanes (in more than three dimensions) is really not that difficult. In each case, all spectra whose components lie on the same n-dimensional surface are assigned the same predicted value and the n-dimensional surfaces associated with increments of the predicted quantity are everywhere parallel and evenly spaced.

[a]The mention of firm names or trade products does not imply that they are endorsed or recommended by the US Department of Agriculture over the firms or similar products not mentioned.

Linear calibrations can only linearly approximate data that contain inherent non-linear relationships between the spectral measurements and the constituent of interest because the planes are constrained to be everywhere flat, parallel and evenly spaced in spectral space. Locally weighted (linear) regression addresses inherent non-linearities in the relationships by creating linear calibrations that are optimised in the vicinity of each unknown sample in spectral space. Therefore, for different samples at different points in spectral space, the calibration planes are not necessarily parallel to, nor spaced the same as, the calibration planes that would be generated for the sample population as a whole.

Transition to non-linear calibrations

Locally weighted regression is a valuable starting point for visualising non-linear calibrations such as artificial neural networks. Imagine many sets of calibration planes associated with locally weighted regression calibrations generated at points throughout spectral space. Now consider bending and connecting planes of equal predicted constituent values to form smooth surfaces in spectral space. The resulting surfaces are approximately parallel and evenly spaced in a limited local region of spectral space, but they are neither parallel nor evenly spaced over the entire "global" spectral space containing all sample spectra. These merged surfaces are conceptually equivalent to the continuous surfaces generated by artificial neural network (ANN) calibrations.

Linear approximations to non-linear calibrations

ANN calibrations, characterised by multiple "hidden" nodes with non-linear transfer functions, seem to defy any attempt to understand their inner workings. However, the concept developed in the previous section (that ANN calibrations are equivalent to global generalisations of locally linear calibrations) suggests a strategy for attacking the problem. At any particular point in spectral space, corresponding to a given sample spectrum, it is possible to derive a linear calibration that is the local linear approximation to the global behaviour of the non-linear calibration. The method comes directly from the definition of the partial derivative of a function with respect to an independent variable.[2]

$$\frac{\partial F(x_1, x_2, \cdots x_n)}{\partial x_i} = \lim_{\delta \to 0}\left[\frac{F(x_1, x_2 \cdots x_i + \delta, \cdots x_n) - F(x_1, x_2, \cdots x_i - \delta, \cdots x_n)}{2\delta}\right] \quad (4)$$

If F is the value of a function of n independent variables making up the input vector ($\mathbf{x}$), the first partial derivative of F with respect to the i^{th} variable (at a given operating point) can be found by holding all variables but the i^{th} one fixed and observing how F changes as that one variable is changed above and below its nominal value by a small amount (δ). Dividing the total change in F by the total change in the i^{th} variable (2δ) gives the derivative—in the limit that the input change is sufficiently small. Repeating the process for each independent variable yields an n-dimensional vector of partial derivatives.

Any NIR calibration can be considered a single function of many independent variables and a single dependent variable. Any pre-processing of the spectral data (such as multiplicative scatter correction or differentiation with respect to wavelength) is part of the overall calibration function. Equation 4, with F as the predicted value from the calibration and $\mathbf{x}$ as the n-dimensional sample spectrum applied to the calibration, yields an n-dimensional linearised calibration vector for any point in spectral space. If the calibration is linear, the *same* linearised calibration vector results from *every* point in spectral space. If the calibration is non-linear, the linearised calibration vectors obtained at different points in spectral space (different samples' spectra) will vary. Figure 2 compares two linearised calibration vectors extracted from an ANN wheat protein calibration (using mean spectra from five low protein and five high protein samples) and a PLS calibration vector (developed using an entirely differ-

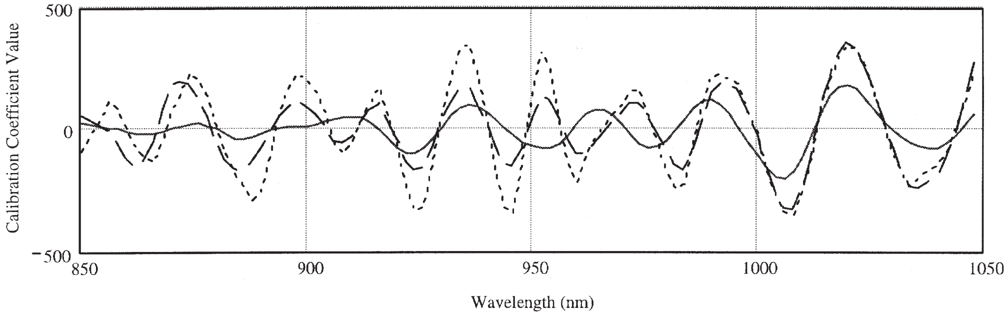

Figure 2. Infratec PLS calibration vector (solid) and extracted calibration vectors for ANN calibration (dot, high protein; dash, low protein) for Hard Red Winter wheat.

ent sample set). Note the similarities between the PLS and extracted ANN calibration vectors and the differences between the extracted calibration vectors for low and high protein wheat.

Assessing and minimising instrumentation effects

The first half of this paper has provided a foundation for visualising linear and non-linear NIR calibrations and extracting linearised calibrations from non-linear calibrations. The remainder will explain some of the ways instrumentation differences affect spectra and how these predictable spectral differences interact with NIR calibrations. Testing for these effects and including transformed spectra during calibration development may significantly improve the robustness of NIR applications.

Noise

If random noise is uniformly distributed over the spectral data, the resultant noise in the predicted results is approximately proportional to the standard deviation of the calibration vector (coefficients). Calibrations with larger coefficients will be more sensitive to noise on the spectra data. Extracted linearised calibration vectors (see Equation 4) can be used to assess noise sensitivity at different points in spectral space. Random noise can be added to spectra to further confirm sensitivity to noise.

Gain

For NIR spectra expressed as absorbance, wavelength-independent gain differences appear as baseline shifts. Referring to Equations 1 and 2, it can be seen that a calibration's sensitivity to baseline shifts is equal to the algebraic sum of the multiplicative coefficients ($K_1 \ldots K_n$). If the sum of coefficients is zero, a constant added to all spectral values causes a shift parallel to the calibration plane in spectral hyperspace, with *zero* resulting change in the predicted value. Extracted calibration vectors can be used to test non-linear calibrations for sensitivity to gain shifts. The sensitivity to a unit baseline shift is the sum of the derivatives obtained using Equation 4. Calibrations that show excessive sensitivity to gain changes can be forced to minimise that sensitivity by adding baseline-shifted spectra to the calibration sample set.

Path length

The multiplicative effects of particle size (scattering) differences are well known from diffuse reflectance research.[3] These differences are assumed to be due to differences in the effective path length. Furthermore, it is obvious from Beer's Law that path length changes induce multiplicative effects for non-scattering samples in transmission mode. The author's work has shown that the effects of physical path length (window spacing) differences for highly scattering materials in transmission mode can also be modelled adequately as multiplicative effects. These multiplicative effects are inherently non-linear. Linear calibrations *cannot* eliminate path length effects.

The concept of calibration planes in hyperspace provides a proof that linear calibrations cannot eliminate multiplicative effects. There is a fundamental difference between offsets and multiplicative effects. Offsets (gain changes) translate data points in spectra space, but do not cause the data to spread out. If the calibration hyperplanes are oriented correctly, an offset doesn't change the predicted result for any data point. On the other hand, multiplicative effects cause the data points to spread out like pellets fired from a shotgun. Since the calibration hyperplanes are parallel and evenly spaced everywhere, it is impossible to orient them so that data points spreading out due to path length differences all move parallel to the hyperplanes. The effects of multiplicative effects on a calibration can be tested by observing changes in predicted results after multiplying a set of spectra by constants slightly less than and greater than one. Such transformed spectra *can* be used to minimise sensitivity to path length in non-linear calibrations.

Stray light

Several different sources of stray light can affect NIR measurements. A dirty, or poor quality, monochromator can "splatter" extensive broadband light over all wavelengths. "Sneak paths" through a granular transmissive sample allow light carrying no sample information to reach the detector. In some cases, light from external sources can reach the detector. All these effects induce positive offset on transmission spectra and compress (flatten) absorbance spectra. The effects of stray light can be simulated by transforming the individual components of an absorbance spectrum (**A**) using Equation 5.

$$A_i' = -\log[10^{-A_i} + (\%StrayLight)/100] \tag{5}$$

The resulting spectra can be used to test different calibrations and calibration techniques for sensitivity to stray light effects.

Monochromator bandwidth

Monochromator bandwidth determines the selectivity of the instrument in isolating sample absorbance characteristics at specific wavelengths. Wider bandwidth "smears" the information from adjacent wavelengths and makes the instrument inherently less sensitive. Therefore, monochromator bandwidth differences are manifested as slope differences between instruments.

The effects of monochromator bandwidth can be simulated by convolving spectra with a window function designed to broaden or sharpen the spectra. A simple window function can be formed by assuming a line shape (i.e. Gaussian or triangular), assigning initial and final bandwidths for the selected line shape, normalising the final line shape to have the same area as the initial line shape, calculating the differences between the two and adding 1.0. The transformed spectra are calculated by convolving the original spectra with this difference function. Applying the original spectra and the transformed spectra to a calibration shows the effects of monochromator bandwidth differences.[b]

Monochromator wavelength axis

Calibrations developed using a "master" instrument are often applied to instruments whose wavelength axes differ appreciably from the master's. When calibration coefficients are applied to data

[b] Unpublished Mathcad documents available on request from the author provide greater detail.

measured at inappropriate wavelengths, calibration performance may be degraded. The author has developed generalised coordinates, (wavelength axis stretch and offset) for describing wavelength axis differences among instruments. Cubic spline interpolation can be used to transform spectra in predictable ways for testing calibrations for sensitivity to wavelength axis differences.[4] Contour plots, showing the effects of various combinations of wavelength axis stretch and offset on instrument bias, slope and standard deviation with respect to nominal conditions, are particularly effective in comparing sensitivities for different calibrations. Wavelength axis transformed spectra can be included in calibration development to "enforce" insensitivity to such differences.

Conclusion

This paper has proposed methods for visualising linear and non-linear NIR calibrations, summarised the effects of instrumentation differences on spectra and discussed interactions of those differences with NIR calibrations. Many of the common types of differences among instruments of a model can be simulated mathematically and used to test and reduce calibration sensitivity to those differences. While these techniques certainly do not eliminate the burden of acquiring representative sample sets to develop and test calibrations, they may minimise the effort required to achieve NIR application robustness with respect to instrumentation differences.

References

1. R. Courant and F. John, *Introduction to Calculus and Analysis*,Volume 2. Wiley Interscience, New York, USA, pp. 13–15 (1974).
2. R. Courant and F. John, *Introduction to Calculus and Analysis*, Volume 1. Wiley-Interscience, New York, USA, p. 27 (1974).
3. G.S. Birth, *Appl. Spectrosc.* **36,** 675 (1983).
4. D.B. Funk, *MS Thesis.* University of Missouri-Kansas City, USA (1997).