Comparison of standardization techniques

John S. Shenk and Mark O. Westerhaus
Agricultural Sciences and Industries Building, The Pennsylvania State University, University Park, PA 16802, USA.

Introduction

The goal of any standardization process is to produce the same predicted results and safety tests on the same sample on all standardized instruments. This is the first step in having a successful network of instruments. To obtain this goal, standardization samples are scanned on a master and host instrument. Mathematical corrections are computed to adjust the spectra of the host instrument to look like spectra from the master instrument. This is a very simple concept but accomplishing this task is still an active area of research. The success of this process depends upon the samples used, the instrument design and sample viewing facilities and the mathematics employed to accomplish the corrections.

Standardization samples

Standardization samples are essential to accomplish the two major goals. First, the wavelengths of the host must be aligned with those of the master instrument (wavelength adjustment). Second, the absorption peaks of the host must be adjusted to match the shape and height of the corresponding peaks of the master instruments (photometric adjustment). Standardization using this method can be considered a calibration problem. Instead of calibrating the spectral data to laboratory values, the host spectral data is calibrated to the spectral data of the master instrument using the scans of the standardization samples. The same principles apply here as in normal constituent calibration. If the standardization is applied on a sample very different from the standardization samples, the agreement may not be acceptable.

Standardization samples must be sealed. Sealing is required to prevent small moisture changes from occurring in the samples between the time the sample is scanned on the host and the time it is scanned on the master instrument. In addition, the sample must be packed tightly so its surface remains constant. These samples must be handled very carefully to prevent minute changes in this surface. If a sample is dropped it must be discarded from the set or rescanned on the master instrument before it is used again. Sample and instrument temperature must also be controlled. All of these precautions must be followed carefully to assure that the spectral differences are due to the instruments and not to the sample.

The Infrasoft International method

The first published standardization method by Shenk and Westerhaus used 30 sealed samples of finely ground agriculture products.[1] Instrument standardization using this method might better be called product standardization. A product is defined as a set of samples with similar composition and physical characteristics. Examples of food and agriculture products would be wheat, flour, cheese, rape seed, forage and feed. A modified form of this standardization procedure is still being

From Near Infrared Spectroscopy: The Future Waves

used today by Infrasoft International Inc. (ISI) for tilting filter instruments and monochromators from different manufacturers.

We chose to use 30 sealed samples to do the standardization. We chose the samples to represent 30 different agriculture products. There was no magic in the fact that 30 samples were used. We tried different numbers of samples and found 30 to be adequate. These samples were scanned on both master and host instruments. Monochromators, with equally spaced wavelengths, were adjusted by calculating a quadratic model to align the wavelengths, then calculating photometric adjustments at each wavelength using linear regression. Since bandpass differences among NIRSystems instruments were negligible from 1100 to 2498 nm, they were not modeled. This procedure could not be applied to filter instruments because the wavelengths were not equally spaced and could not be aligned accurately. The mathematical procedure used on filter instruments was PLS2. Hundreds of instruments are successfully using one of these methods today.

In 1987 the NIRSystems company produced a monochromator with polystyrene mounted inside the instrument for wavelength alignment. Since the design of the instrument was carefully monitored during manufacturing, this single polystyrene sample provided sufficient consistency across instruments to eliminate the wavelength alignment problem. Wavelength alignment of ±0.25 nm was sufficient for most food and agriculture products. Having solved the wavelength alignment with one sample, a single standardization sample was tried to correct the photometric response. Since the spectra of finely ground samples are linear over the 1100–2500 nm range, a simple subtraction or bias correction model was proposed to make the adjustment.

This method is known as single sample standardization. The sample with a spectrum closest to the average spectrum of the original 30 samples was chosen. Although the 30 sample method was sometimes better than the single sample method, because of the simplicity and low cost of the single sample method, it was, and is, the preferred method used today for NIRSystems bench top instruments used in food and agriculture applications.

Problems and alternative solutions

Although these ISI standardization procedures work well under the conditions described above, the method may not perform satisfactorily when applied to a broad range of instruments or products with diverse spectral properties.[2] This was reported in Germany when trying to standardize instruments for rape seed analysis. Agreement between instruments was not acceptable using the normal set of 30 finely ground sealed samples. This was because the optical properties of whole black rape seed were not within the range of the 30 standardization sample set. To correct this problem it was necessary to add a rape seed sample to the 30 sample set.

Others trying to use the ISI method found that it did not work well enough on samples with very sharp spectra. This was known at the time of our earlier development, but it was not important because the majority of food and agriculture samples do not have sharp peaks. The paper by Bouveresse *et al.*[3] describes this problem in detail. They have attempted to solve this problem by using locally weighted regression. Their method appears to hold promise but has not been thoroughly tested. Also, the authors state that they have not found the best set of samples to perform the standardization across a broad spectrum of pure compounds.

A second area where standardization sometimes does not produce satisfactory results is when the optical density of the product goes above 1.0 in log 1/*R*. Even though the manufacturer tries to keep the optical response linear, stray light and sample factors cause non-linear responses at high log values. The amount of non-linearity varies between instruments. An example would be an unground sample of forage containing 60% water. If the same sample is scanned on two instruments that differ slightly in stray light, the difference between the instruments would be non-linear. This, of course, will make the predicted values of one instrument different from the

other. Although it is entirely possible that a non-linear model can be found to make these two instruments more similar, the model may not work well on a third instrument. In addition, making a sealed sample of a coarse unground sample is difficult. This can be done for the NIRSystems transport system for seeds like corn, wheat and barley, but these standards are specific for the NIRSystems instrument and cannot be used by other instrument manufacturers.

A third area of concern with the present approach of using one sample near the center of all agriculture products on the NIRSystems instruments is that equation bias among instruments for a particular product may not be acceptable. This is because a single sample of a generic product with a simple linear offset can not correct for the small differences among products. This too has been known for a long time, but since the problem is usually associated with bias only, customers make the bias adjustment if needed. Our current approach to this problem is to use a single sample standardization for each product. We have found this necessary, especially when predicting the amino acid concentration of corn grain and soymeal. Since the concentrations of the amino acids are so small, it is possible to get a bias between instruments almost as large as the predicted value itself. We found that this was easily corrected by using a single sealed sample of each product with spectral properties near the center of each product as the standardization sample.

The above discussions have been about reflectance, but the same problems and challenges exist for transmission. Some people think that because a blank cuvette or air is used as a reference for transmission, standardization among instruments is not necessary. This, of course, is not true. We have looked at the spectra of sugar syrup in transmission among instruments. One might think that a single sample of water might correct for the differences in predicted values among instruments, but here again the best sample to perform this correction is a single sample of sugar syrup. The same is true for transmission of whole grain analysis. Although the analyses are very repeatable from instrument to instrument, suggesting good wavelength alignment by the manufacturer, the bias error among instruments is unacceptable, indicating that a photometric correction is needed. On-line analysis with fiber optics is another problem. At the moment only bias and slope adjustment can be offered as a solution.

We applied the new absolute absorbance standardization method developed by Karl Norris[4] to our instruments and equations. The test was carried out with three NlRSystems 6500 instruments using the samples and mathematics described by Karl in his previous paper. The results were very good when tested with four product calibration equations (hay, wheat, meat meal and soymeal). The major error was bias. Our single sample standardization method corrects for bias. Using Karl's present method a bias adjustment would still be needed. In addition, this method could be very expensive for a food or agriculture customer since the samples are expensive and must be used to restandardize the instrument even if a light bulb burns out in the instrument.

Prediction equations

Regardless of the method of standardization that is used, an accurate and robust equation for the standardized instruments must be produced. One of the most important attributes of the prediction equation is the ability to minimize the effect of unwanted variation that can occur across instruments during routine operation. We accomplish this task with a repeatability file rather than trying to add this unwanted variation directly to the calibration file.

The repeatability file is developed by using a few samples scanned at different temperatures on different standardized instruments. The repeatability file does not need laboratory reference values since repeatability scans are treated as deviations around the mean spectrum in the regression analysis. The purpose of the repeatability file is to downweight the coefficients in regions of the spectra where the unwanted variation exists. Our experience suggests that using a first derivative with a small gap and smooth breaks up the intercorrelation in the spectral data and

minimizes the effects of the noise in the repeatability file. Combining the variation in the spectra for the constituent with the unwanted variation caused by temperature, and leftover variation in the instrument standardization process, provides a very good combination to produce a robust calibration for instruments on a network.

Testing and summary

The final task is testing the performance of a network, regardless of the standardization and equation generation method. In the food and agriculture areas, we use the following test. If the pooled standard deviation of the predicted values across instruments is less than or equal to the standard deviation of subsamples chosen from a container of the sample, we say the network is performing satisfactorily. Standardized instruments are the starting point for a successful network but the development of a robust prediction equation fine-tuned for this network is essential for successful operation. If all parts of this process are completed successfully, the wavelengths will be within tolerance across instruments, the photometric response will be controlled to prevent bias and all safety tests will function across instruments to make the analysis trustworthy.

References

1. J.S. Shenk, J.J Workman Jr, and M.O. Westerhaus, "Application of NIR Spectroscopy to Agriculture Products" in *Handbook of Near Infrared Analysis*, Ed by D.A. Burns and E.W. Ciurczak. Marcel Dekker Inc., New York (1992).
2. Y. Wang, D.J. Veltkamp and B.R. Kowalski, *Anal. Chem.* **63,** 2750 (1991).
3. E. Bouveresse, D.L. Massart and P. Dardenne, *Anal. Chem.* **67**, 1381 (1995).
4. K. Norris, *personal assistance and communication* (1995).