Anal. Chem. **1995**, *67*, 1381–1389

Modified Algorithm for Standardization of Near-Infrared Spectrometric Instruments

E. Bouveresse and D. L. Massart*

ChemoAC, Pharmaceutical Institute, Vrije Universiteit Brussel, Laarbeeklaan 103, B-1090 Brussels, Belgium

P. Dardenne

Station de Haute Belgique, Rue de Serpont 100, B-6800 Libramont-Chevigny, Belgium

The transfer of calibration models between near-infrared spectrometric instruments with the algorithm proposed by Shenk and Westerhaus as a standardization method is studied. This algorithm does not yield optimal results when standardization samples are used that are not of the same nature as the samples to be analyzed. In an attempt to improve the results obtained with this algorithm when one uses such standardization samples, five different modifications of the spectral intensity correction step of the algorithm are investigated. The use of locally weighted regression, which gives more weight to the standardization samples close to the same spectral intensity range where the samples to be predicted are located and less weight to the farthest samples, seems to be the most suitable one.

Because of the differences between instrumental responses of near-infrared spectrometers, standardization has become a necessary step to correct those differences and to avoid time-consuming recalibration procedures. Among the standardization methods proposed in the literature, the patented method proposed by Shenk and Westerhaus[1,2] is widely applied. In this paper, this algorithm will be referred to as the "patented" algorithm. It is based on the transfer of spectra from the "slave" instrument to the "master" one and involves two steps, namely wavelength index correction and spectral intensity correction. The spectral intensity correction involves a linear regression relating wavelength by wavelength the spectral intensities obtained on a first instrument with those obtained on a second one. In a previous study,[3] calibration transfer involving standardization samples, independent from the samples to be predicted, was tested. Two different approaches were investigated: the first involved using an independent standardization set containing samples similar to those to be predicted, and the second involved using two independent standardization sets containing chemically more stable samples very different from those to be predicted. A mixed approach was studied by Wang and Kowalski:[4] they used two generic standards and a sample selected from the set of samples to be predicted as the standardization set and applied the piecewise direct standardization method.[5] As far as we know, a pure generic standard approach using the patented algorithm has not been proposed yet. Our previous study[3] showed that the use of standardization samples similar to the samples to be predicted can lead to good standardization but is applicable only for those samples. The use of standardization samples of different nature, which cover a larger spectral intensity range, leads to poor results. The use of such standardization samples would be more suitable, because very stable standardization samples could then be measured to transfer spectra of different sets of samples. To explain the results obtained, the first reason proposed is the different nature of the samples involved. Yet, it is also possible that the algorithm presents some limitations when the data sets involved exceed the usual framework of application. For instance, if the spectral variation range covered by the standardization samples is really wide, nonlinearities might occur, and the linear regression used in the patented algorithm might lead to bad results. Changing the regression step to a more suitable one may then improve standardization. Moreover, high spectral intensity values are probably more subject to noise. Since a low number of standardization samples are involved to compute standardization parameters, points located in high spectral intensity ranges and subject to relatively large noise can have a bad influence on the linear regression model. Changing the simple linear regression to a more suitable one may also improve standardization. As explained later, it is not possible to simply eliminate high absorbances from consideration as this would lead to artifacts (discontinuities) in the standardized spectra.

In other words, if the patented algorithm enables good standardization in some cases, the bad results obtained in other cases can perhaps be improved by some modifications of the algorithm.

To cope with the problems previously mentioned, the following techniques are investigated.

Use of Univariate Quadratic Models. In order to better fit nonlinearities due to wide spectral intensity ranges, the spectral intensities of both instruments are related wavelength by wavelength with quadratic models.

Iteratively Reweighted Least Squares.[6–8] This iterative technique is applied in order to decrease the effect of spectral intensities subject to noise on the regression lines obtained.

(1) Shenk, J. S.; Westerhaus, M. O. *Crop Sci.* **1991**, *31*, 1694–1696.
(2) Shenk, J. S.; Westerhaus, M. O. U.S. Patent 4866644, Sept 12, 1989.
(3) Bouveresse, E.; Massart, D. L.; Dardenne, P. *Anal. Chim. Acta* **1994**, *297*, 405–416.
(4) Wang, Y.; Kowalski, B. R. *Appl. Spectrosc.* **1992**, *46*, 764–771.
(5) Wang, Y.; Veltkamp, D. J.; Kowalski, B. R. *Anal. Chem.* **1991**, *63*, 2750–2756.
(6) Mosteller, F.; Tukey, J. W. *Data analysis and regression*; Addison-Wesley: Reading, MA, 1977.
(7) Phillips, G. R.; Eyring, E. R. *Anal. Chem.* **1983**, *55*, 1134–1138.

0003-2700/95/0367-1381$9.00/0