Spectral standardisation methods and modelling techniques applied to a diverse population of forage samples

P. Dardenne,[a] I.A. Cowe[b] and M. Lagerholm[b]

[a] Agricultural Research Center (CRA), 24 Chaussee de Namur, B-5030 Gembloux, Belgium, dardenne@cra.walonie.be
[b] Foss Tecator AB, Box 70, S-26121, Hoganas, Sweden

Introduction

The day is fast approaching when forage, like grain, will become a commodity traded on quality parameters determined by near infrared spectroscopy (NIR). To achieve this goal it is increasingly important to accumulate the large databases needed to cover the variability found in different types of forage worldwide. The utility of such databases depends upon a common standard of laboratory accuracy[1] and our ability to make NIR instruments predict within acceptable limits when using the same calibration equation. While the development of robust prediction models does help, spectral standardisation of instruments[2] is a key issue in the development of forage networks. The standardisation techniques developed by Infrasoft International (ISI)[3,4] have been the most widely used but there are a number of other techniques that have been proposed which may be useful in this respect. In this report we examine the performance of four standardisation techniques, ISI single sample standardisation (ISIS), ISI multi sample standardisation (ISIM), piecewise direct standardisation (PDS)[5] and Fearn standardisation (FS)[6] with models developed using three modelling techniques modified partial least squares (MPLS),[7] LOCAL[8] and artificial neural networks[9] (ANN) when applied to a very diverse set of forage samples.

Material and methods

Samples

The samples used to standardise the instruments and to evaluate the performance of the standardisation methods and prediction models were part of a 74-sample data set collected worldwide in 2001 as part of the activities of the Foss/DeLaval World Forage Board.[1] Ten samples that were common to all instruments were selected for standardisation while the remaining samples were used as a test set. Not all samples were scanned on every instrument and therefore the test sets consisted either of 64 or 42 samples. All samples used for standardisation and testing were subsamples of bulk samples that were ground, mixed, subsampled and distributed to participants in powdered form. Packing error was therefore compounded in the measurements used for standardisation.

The data set used to develop models consisted of more than 16000 sample spectra with reference values for moisture, protein and neutral detergent fibre collected over a period of more than ten years at various laboratories across the world. These three parameters were considered because the methodology was most common worldwide over the period of sample selection. Because of difficulties involved with transport of samples,[1] and because of space, results for moisture are not shown.

Near Infrared Spectroscopy: Proceedings of the 11th International Conference.

Instruments

The instruments were all either NIRSystems 6500 or 5000, and the common data range 1100 nm to 2498 nm was used for this exercise. Instruments in Australia, USA, Canada, Sweden, Germany, and Belgium were included in the trial.

Standardisation techniques

The ISIS standardisation technique creates a standardisation file with a difference spectrum between master and slave instrument that is then subtracted from each spectrum measured on the slave. ISIM standardisation creates a file that has information that first corrects the wavelength axis on the slave to that of the master and then applies a photometric correction to the slave spectrum. PDS involves a moving window within a spectrum with principal components regression being used to calculate a correction for the current wavelength based on the differences between a master and slave measurement of the same sample. Fearn standardisation is a new technique and differs from the other methods in that there is no master instrument as such. Here a representative sample (or set of representative samples) is scanned on a number of instruments and principal components analysis is used on the spectra (or on mean spectra) to identify a subset of components that describe the variation that relates to instrument characteristics. The data set used in model development is then orthogonalised to this variation before modelling and all prediction spectra are similarly treated before values are predicted. WINISI.® was used for ISI standardisations while Matlab.® was used for all other operations.

Modelling techniques

Standard normal variate (*SNV*) plus detrend followed by first derivative (1:4:4:1) was applied to all data. MPLS and LOCAL were carried out using WINISI® while ANN models were developed using software proprietary to Foss Tecator and coded in Matlab.®

Statistics considered

Four statistics were considered as indicators of performance. Root mean square error of prediction (*RMSEP*) is a measure of the average deviation between predicted values for a slave and those for the master instrument. standard error of prediction (*SEP*) is the equivalent measurement once bias has been taken into account, where bias is defined as the difference between means for slave and master instruments. The final indicator is the slope of a regression line of a plot of predicted values for slave versus master. Ideally, *RMSEP* and *SEP* should be as low as possible while bias should be 0.0 and the slope 1.0.

Results and discussion

Figure 1 shows average results over all six data sets with and without standardisation for protein and NDF. Grey bars show *RMSEP*, and within each grey bar the coloured bar shows *SEP*. For protein, ANN gave marginally better unstandardised results for *RMSEP* than either MPLS or LOCAL For each type of model, all four standardisation techniques showed improvements in *RMSEP*. An improvement could either be a reduction in bias or an improvement in fit about a regression line. All the standardisation methods were very efficient at reducing *RMSEP*, but for PDS, Fearn and ISIM there was also a clear reduction in *SEP* indicating an improvement in fit about the regression line. ISIS standardisation had only a minor improvement in *SEP*.

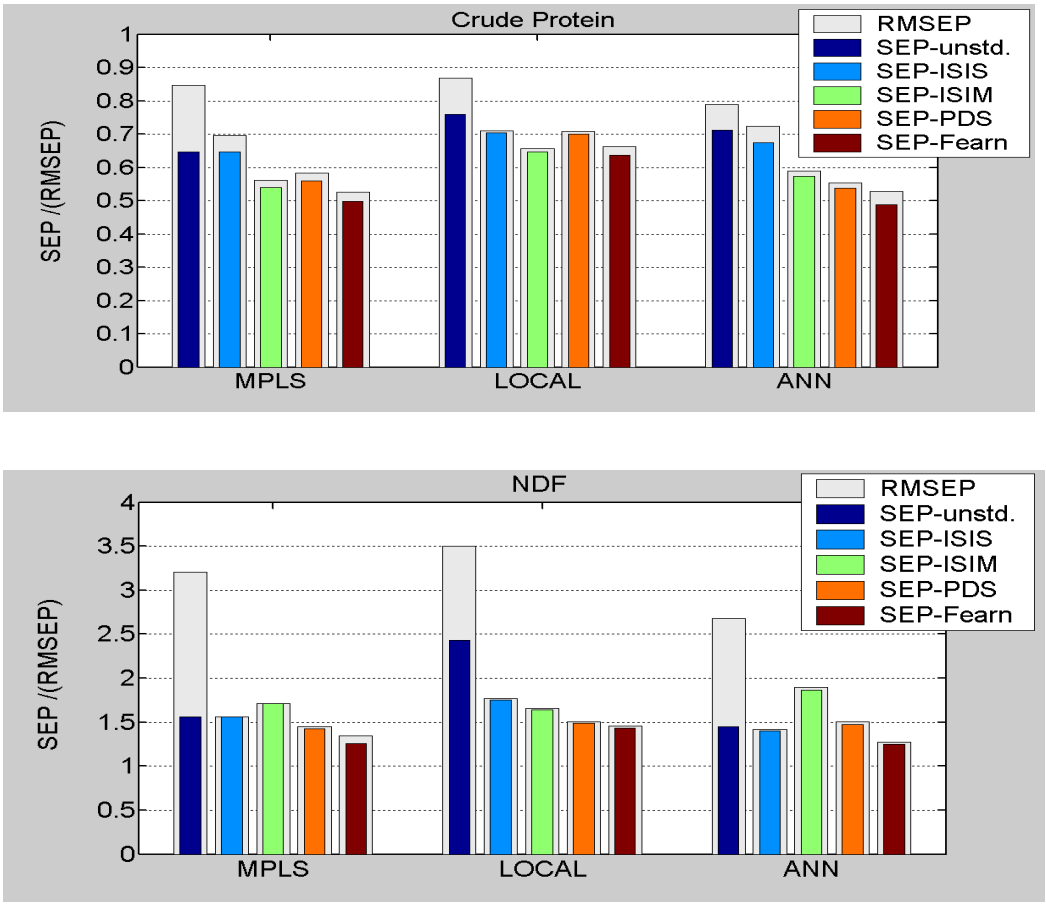

Figure 1. Comparison of errors for unstandardised data and data standardised using four methods and three modelling techniques.

For NDF, reduction in bias producing improvements in *RMSEP* was the main effect. ANN gave the best results for *RMSEP*, but when *SEP* was considered ANN and MPLS were clearly better than LOCAL. However, where LOCAL was used, standardisation improved *SEP*s to levels seen for the other modelling techniques. Of all the standardisation methods, only the Fearn standardisation improved *SEP* for all the modelling techniques.

When average responses are considered we see that all the standardisation techniques improved performance, particularly where bias was considered. Improvements to *SEP* were less consistent and in one case, NDF predicted using ISIM combined with ANN, the average *SEP* was higher after standardisation than before. Even with MPLS, the only technique based on linear modelling, a small improvement in *SEP* was seen both for protein and NDF when ISIS was applied. This is interesting, because the correction applied by ISIS is purely a spectral offset and this would tend to affect bias rather than *SEP*. However we must also consider the effect of the spectral pre-treatment, which may be introducing non-linearities into the data after the spectra have been standardised.

While average response across a population of instruments tells us something about the usefulness of a particular standardisation technique, it is only by examining the response of individual instruments that we see how the interaction between a modelling method and a standardisation technique differs from instrument to instrument.

In Figure 2, and in subsequent plots of this type, predicted values for each of seven instruments are plotted against those for a master instrument. For clarity, an offset has been added to each data set so that multiple plots can be seen together. The line shown for each data set is the 45° line, i.e. the line along which data ideally should lie if the master and slave produce identical responses.

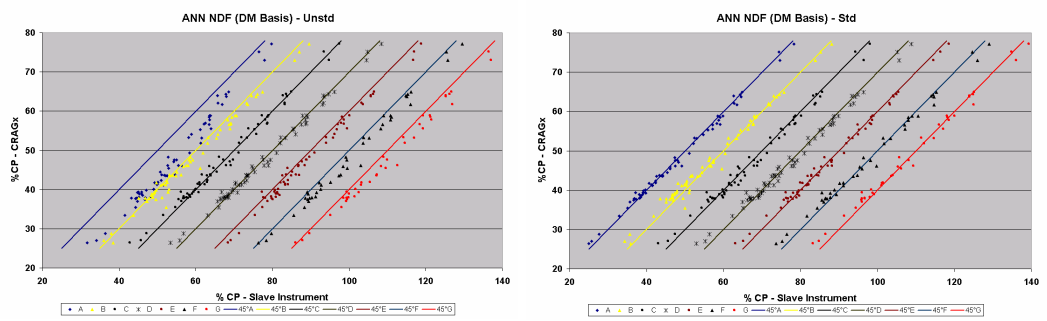

Figure 2. Scatter plots for seven instruments (A…G) plotted against a master instrument (*Y*-axis). ANN model for NDF before (left) and after (right) standardisation by ISIS.

Figure 2 illustrates the typical response of instruments to ISIM standardisation. If we look at the unstandardised plot we see that biases vary considerably between instruments. In the plot on the right we see that ISIM standardisation has removed these biases but has generally left the relative positions of the data points unchanged.

In Figure 3, we see ISIM standardisation coupled with models using LOCAL. In this case not only has the standardisation removed bias but for most of the instruments it has also tightened up the distribution of points along the 45° line.

Figure 4 shows PDS used with MPLS as the modelling technique with crude protein as the example. Here we see that, for protein, bias was much smaller and the effect of standardisation was both to reduce this bias and to improve the fit.

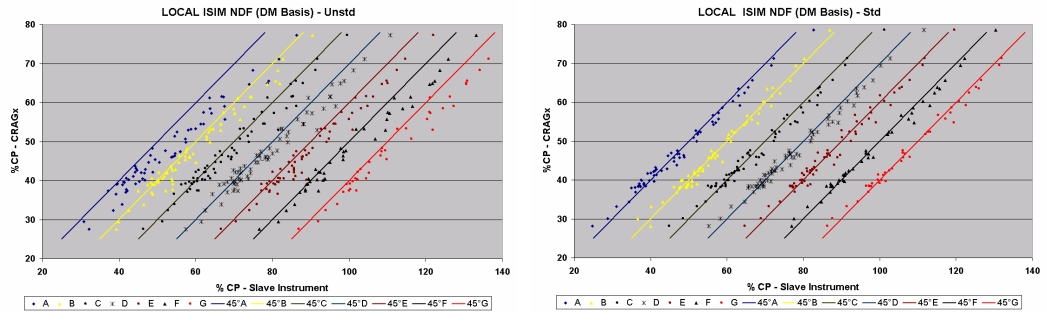

Figure 3. Scatter plots for seven instruments (A…G) plotted against a master instrument (*Y*-axis). LOCAL model for NDF before (left) and after (right) standardisation by ISIM.

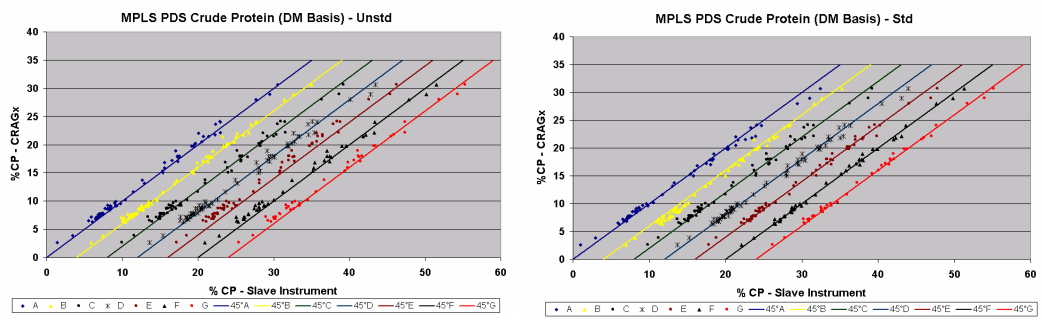

Figure 4. Scatter plots for seven instruments (A…G) plotted against a master instrument (*Y*-axis). MPLS model for protein before (left) and after (right) standardisation by PDS.

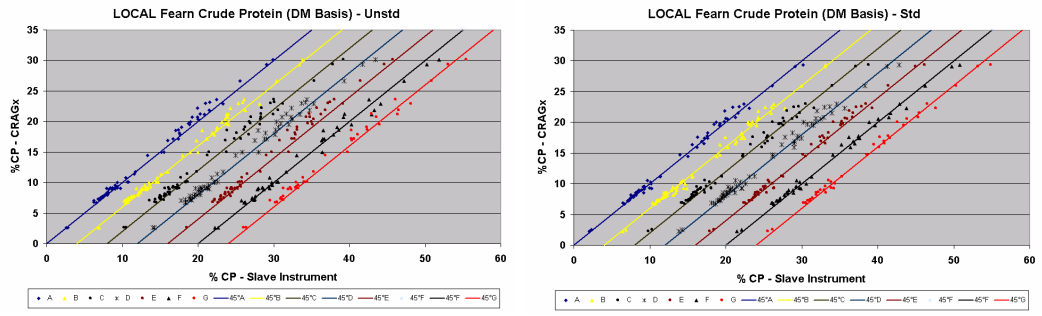

Figure 5. Scatter plots for seven instruments (A…G) plotted against a master instrument (*Y*-axis). LOCAL model for protein before (left) and after (right) standardisation by Fearn standardisation.

In Figure 5, we see Fearn standardisation with LOCAL as the modelling technique for protein. In this case every instrument showed an improvement in *SEP*, with the average down from 0.463 to 0.276% protein.

Conclusions

The effects of standardisation are difficult to quantify because there are considerable interactions between modelling methods, the standardisation technique and instrument characteristics. All of the four techniques tested here worked well with improvements not only to *RMSEP*, which can be improved relatively easily by reduction in bias, but also for *SEP* which is a far more severe test. ISIS, the simplest technique, while correcting primarily for bias, still showed some improvement to *SEP*. The success of Fearn standardisation is interesting in that this was the only method to break the link between a master standardisation instrument and individual slaves. The orthogonalisation of both the calibration and test data sets is based on properties of all the seven instruments that were the test set of instruments. These instruments had contributed to the calibration set but there were many more instruments not characterised in the calibration set. When the data transformation was complete it was necessary to develop completely new models because the data had changed due to the orthogonalisation. A direct comparison may therefore be complicated by small differences in

complexity of the models, but this is to be expected if some variation not relating to composition has been removed from the original data.

Acknowledgements

Particular thanks are due to Professor Tom Fearn, University College, London, for advice and for Matlab code to implement his standardisation technique and to John Shenk and Mark Westerhaus for advice relating to ISI standardisation. Thanks also to Christian Paul, Peter Flinn, Paolo Berzaghi, Dan Undersander, Bo Buchmann and Hua Hsu for the use of data collected as part of the Foss/Delavel World Forage Board activities.

References

1. D.J. Undersander, I.A. Cowe, P. Dardenne, P.C. Flinn, C. Paul, P. Berzaghi, N.P. Martin and N.B. Büchmann, "An International Ring Trial to Assess Differences Between Laboratories in Forage Analysis: A New Slant on an Old Problem", Proceedings of NIR2003, in press.
2. T Fearn, *J. Near Infrared Spectrosc*. **9,** 229 (2001).
3. J.S. Shenk and M.O. Westerhaus, *US Patent No. 4,866,644* (12 September, 1989).
4. J.S. Shenk, M.O. Westerhaus and W.C. Templeton, Jr, *Crop Sci.* **25,** 159 (1985).
5. Y Wang, D.J. Veltkamp and B.R. Kowalski, *Anal. Chem*. **67,** 2379 (1995).
6. Personal communication.
7. *WINISI II Manual*. ISI Windows Near Infrared Software, Infrasoft International, State College, Pennsylvania, USA.
8. G. Sinnaeve, P. Dardenne and R. Agneessens, *J. Near Infrared Spectrosc*. **2,** 163 (1994).
9. Foss Tecator, Hoganas, Sweden, proprietary software.