The effect of ISI standardisation on the performance of Infratec instruments

Ian A. Cowe*

Foss Electric Development (UK) Ltd, Wheldrake, York, YO19 6NA, UK.

John S. Shenk and Mark O. Westerhaus

Infrasoft International, LLC, 109 Sellers Lane, Port Matilda, PA 16870, USA.

Introduction

Standardisation of near infrared (NIR) monochromators is a concept that has been developed extensively by Shenk and Westerhaus[1,2] as part of the ISI Software Package, mainly using Foss NIRSystems instruments. Using a system of sealed sample cups, it is possible to measure differences in spectral response between instruments with high precision, thereby allowing adjustments of spectra from slave instruments to produce spectra similar to those which a master instrument would have produced.

Whole grain analysers, such as the Foss Tecator Infratec series, present different problems with regard to instrument standardisation. First, they work in a region of the spectrum (850 to 1050nm) where absorbance bands are broader and there are fewer peaks that can be used for wavelength standardisation. Second, cereal grains or oilseeds pass through an instrument as freely-flowing samples where periodically the flow is halted to allow measurement of a sub-sample. When enough sub-samples have been measured (typically 10), a prediction of composition is obtained and a mean spectrum is calculated. Under these conditions, variations in packing density produce sub-sample spectra having much greater baseline shifts than with reflectance measurements. The mean spectra are, therefore, less reproducible than is possible with sealed sample cups in reflectance.

While modelling systems such as partial least squares (PLS) and particularly artificial neural networks (ANN) are relatively insensitive to baseline shifts in spectra when predicting composition, there are measurable effects, usually bias, when a calibration developed on one instrument is mounted on another. These effects are usually removed in a slave instrument by adding a bias correction term to predicted values, a process that is time consuming and has to be done for each constituent measured on each commodity.

Spectral standardisation systems such as ISI Standardisation have, therefore, the potential to improve transferability of calibrations by correcting at source the variations which later translate to differences in predicted values. Moreover, this should be possible on a commodity basis regardless of the number of constituents measured on that commodity. Standardisation, by making instruments spectrally alike, should have the effect of making all samples analysed by standardised instruments appear spectrally as if they had been produced by a single instrument. In such a situation, measurements of Mahalanobis distance, such as the Global *H* or Neighbourhood *H* become valid in selecting the right samples for optimising the structure of populations used for calibration.

In this paper we examine some of the factors which have been identified as being important to the successful implementation of ISI standardisation on Infratec whole grain analysers.

Near Infrared Spectroscopy: Proceedings of the 9th International Conference

Materials and methods

ISI standardisation

ISI standardisation[1] is a two-stage process. Initially, the position of peaks and troughs in spectra recorded on a master instrument and those from a slave are compared and then the monochromator *O* and *P* coefficients needed to match the slave spectra with those of the master are calculated. The second stage involves calculation of a scaling factor and a vector of offsets needed to match the photometric response of the slave to that of the master instrument.

Calibrations used

The ANN calibrations used for both wheat and barley in the 1999 Danish Field Trial were WBMO8 for moisture and WPB14 for protein. The PLS calibrations for wheat were DW14-1 (moisture) and DW16-1 (protein) and for barley DB17-1 (moisture) and BG17-1 (protein). Details of these models can be found in Büchmann *et al*.[3]

Results and discussion

Over the past two years, Foss Tecator has conducted two trials of ISI standardisation as applied to Infratec instruments. The first, in the spring of 1998, used Model 1229 instruments, taken straight from the production line. Initially, data from a total of 110 instruments were examined.

In this trial, 15 samples of wheat were scanned through each instrument. Given that the sample transport system of 1229 instruments is designed for freely flowing samples it was not possible to adopt the sealed sample cup concept central to the ISI standardisation process described for reflectance analysers. Instead, five GrainBlocs[4] or five of the 15 free-grain samples were used to standardise the instruments, the remaining ten samples being used as an independent test set. In addition, a subset of three of the five standardisation samples in each set was used to recalculate standardisation factors.

The questions posed in this trial were as follows:

1) Would ISI standardisation work using free grain samples?
2) Would GrainBlocs provide a better standardisation than free grain?
3) Would standardisation be better with five, rather than three, samples?
4) Would standardisation work equally well for moisture and protein?
5) How would models based on artificial neural networks perform compared with PLS models?
6) Which was more important, the wavelength or photometric elements of ISI standardisation?

During the data collection it became apparent that changes in the moisture content of the free-grain samples was becoming a major influence in the results. Over time, the samples were drying out and changes in moisture content were accompanied by changes in spectra, which were being incorporated into the standardisation. The samples were stored in sealed jars but, over a holiday period, it was obvious that they had dried out, presumably due to lids not being correctly sealed.

As, thereafter, moisture content of the samples was relatively constant, a new master instrument was chosen and data from the last 64 instruments were used to assess the standardisation process. The following conclusions were made:

1) Changes in moisture content of samples during standardisation would be incorporated into the standardisation process, producing standardisation errors. Hence the use of sealed sample cups in the standard method.
2) Three samples of free grain were sufficient to standardise instruments. Using five gave no significant improvement in performance.
3) ISI standardisation improved transferability statistics for moisture and protein using both ANN and PLS models. The effects were larger with PLS because there were larger biases.

4) The main improvement was reduction in bias between instruments. However, values for the standard error of prediction (*SEP*) were also improved suggesting that the standardisation process had reduced the scatter about a regression line.
5) Standardisation using GrainBlocs gave poor results, despite the spectra being measured with much higher precision than with free grain. The introduction of large bias effects was the main problem.
6) Examination of the response of individual instruments showed that photometric, rather than wavelength correction, was the more important effect. In addition, there was a tendency for the standardisation process to over- or under-shoot when correcting slave instruments for bias.

Looking at these results, it was not surprising that the wavelength element of ISI standardisation was relatively unimportant. These were new instruments, which had been aligned using production techniques that should have been more accurate than measurements of secondary standards such as whole grain or GrainBlocs.

The poor performance of GrainBlocs was assumed to be due to small absorbance effects from the resin used to produce the block being interpreted by the regression models as large bias effects.

Although the results were encouraging, the trial had not been carried out under field conditions, had not included instruments that were of the older 1221 model or instruments had been subjected to several years of use. A period of assessment of the results and development of the system was, therefore, undertaken with the aim to carry out a full field trial in 1999.

One matter that was examined was the errors inherent in measuring whole grain NIR transmission spectra. For each commodity, the pathlength used is a balance between conflicting requirements. If the pathlength is too short, "pin holes"—air gaps between individual kernels through which light may

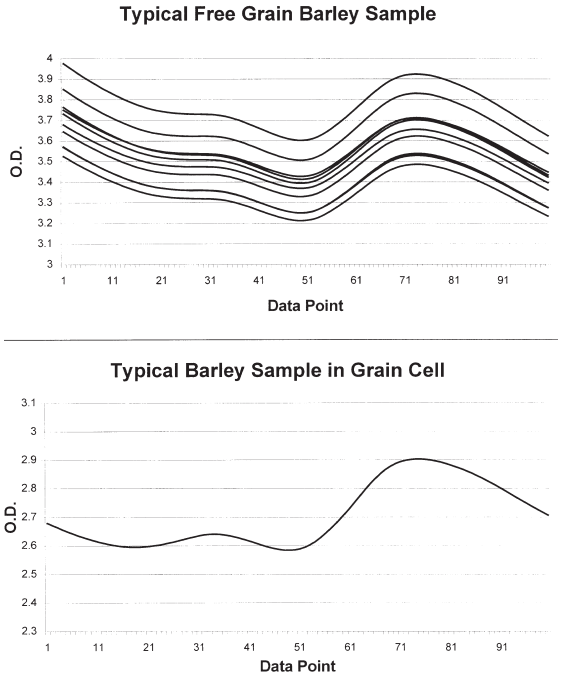

Figure 1.(a) Plot of 10 sub-sample spectra for a sample of barley; (b) plot of 3 sub-sample spectra for a sample of barley in a Grain Cell (sub-sample spectra visually indistinguishable).

pass directly to the detector without interacting with the grain — occur which degrade the signal. If the pathlength is too long, not enough energy reaches the detector to provide a good signal-to-noise ratio needed for reliable prediction of composition.

Samples flow through an instrument and pack the measurement cell chaotically. Even if the same sample is measured many times, it is unlikely that the same grains in the same spatial orientation will ever be seen by the detector. Sub-sample spectra from the same sample exhibit differences in optical density (OD) of as much as 0.5 OD in the general range from 2.0 to 4.0 OD [see Figure 1(a)].

As ISI Standardisation is based on spectral data rather than predicted values, differences in OD of mean spectra for the same sample measured on a master and a slave determine the photometric standardisation coefficients for that slave instrument.

The aim of the GrainBloc system was to present the same sample of grain to an instrument each time. A new system, Grain Cells, was developed as an alternative. Here, samples of grain were packed under compression into sealed metal cells in such a way that the grain would not move over time. When these were presented to different instruments, we obtained the same spectra with sub-sample standard deviations of spectra two orders of magnitude lower than for equivalent free grain samples [see Figure 1(b)].

This work indicated one weakness of previous GrainBloc measurements. It indicated that the tolerances for presenting such samples to an instrument were much tighter than previously allowed and that this might have been part of the reason for the failure of GrainBlocs in the previous trial.

The second field trial was carried out in the Spring of 1999. A total of 23 instruments were visited, 20 in industrial sites in Denmark plus one in Hoganas, Sweden and two in York, England. Six instruments were Infratec Type 1229; one was a Type 1225 and the remainder, Type 1221. Six sample sets of 15 wheat and 15 barley were prepared by subdividing large bulk samples. A new sample set was used each day, with data from up to four slave instruments being collected before the samples were sent to the Danish sub-master at Lingby, where each set was scanned four times through the sub-master. This gave us mean spectra for slaves on the basis of ten sub-samples and mean spectra for the sub-master on the basis of forty sub-samples.

For each of the wheat and barley sets, three samples were used for standardisation, the remaining twelve being used as a test set. In addition, three wheat Grain Cells and three for barley plus an artificial wavelength standard were scanned. During data collection measurements were made of reference and sample gains to monitor the condition of instruments.

The questions asked in this case were:

1) Would ISI Standardisation using, free grain, work under field conditions?
2) Would Grain Cell standardisation outperform standardisation on free grain?
3) Would wavelength standardisation be important with older instruments?

When instruments were visited in the field, a number of problems were identified which were likely to have a bearing on the results from the trial. The most important was cleanliness. Some instruments were very clean, with evidence of careful upkeep, but others were operated under poor conditions and it was obvious that they had not been cleaned for a considerable time. Some instruments were sited, inappropriately, near fluorescent lights and some had sealing gaskets missing.

The quality of sub-sample spectra was also monitored. We identified two forms of aberrant spectra, the first restricted solely to wheat and the second primarily to barley. The first abnormality occurred when a sample settled during a scan. Wheat is very smooth and, in a small percentage of sub-samples, the grain moved slightly during the scan, causing a discontinuity in the spectrum. This condition is almost impossible to identify from visual inspection of log $1/T$ spectra but can have a very large effect on predicted values, depending upon where in the spectrum the discontinuity occurs. In routine prediction mode, an Infratec can be configured to identify such spectra if they produce predicted values, which are outliers. As differences in mean spectra of master and slave instruments are

the basis of ISI standardisation, a sub-sample rejection system was implemented which removed a small number of sub-samples with aberrant spectra.

The second abnormality relates to the "pinhole" problem. Here, we see "flat" spectra where some of the light has passed directly to the detector. This type of abnormality does not produce grossly distorted predicted values but can produce bias effects within the range of normal predicted values.

It became obvious that performance of the standardisation system was influenced by the presence of spectral outliers. Data was, therefore, assessed both with outliers present and with these sub-samples omitted from calculation of mean spectra.

Results using Grain Cells were relatively poor. This was surprising, as the grain in the cells was the same as used for free grain standardisation and there was no resin to add unwanted absorption effects. In addition, the placement of the cells in relation to the detectors on different instruments was of an extremely high standard and sub-sample spectra were extremely repeatable.

The problem was traced, finally, to the cleanliness of the instruments. When using Grain Cells we replaced the instrument's sample cell with an instrument cell, which was clean, was guaranteed to be light tight and which presented Grain Cells exactly to the centre of the detector window. Air reference and then Grain Cell measurements were taken using this attachment and then the original instrument cell was replaced before the free grain samples were measured.

When comparing free-grain standardisation with Grain Cell standardisation, we could see that free-grain standardisation and test samples were measured under identical conditions. With Grain Cells, standardisation was being done with a clean system but the test set was measured with the original sample cell, which could have a dirty cell window. If we cleaned a very dirty window in a cell, the log 1/*T* values for a sample would change by around 0.2 OD. Examination of air reference gain values indicated that the difference was caused by the air reference measurements which are used in calculating optical density. With a clean cell, light passed through the window and fell directly on to the detector. When the window was dirty, the dirt acted as a diffuser, scattering light that would normally pass to the detector. When a sample was measured, the presence of grain in the cell caused scattering and the extra effect of the window dirt was negligible.

This implies that the state of cleanliness of the master instrument, compared with the slave, affects the standardisation process. If either the master or a slave instrument significantly change in terms of cleanliness from the condition in which they were standardised then bias effects will be introduced and the standardisation will no longer be valid. In future, a modification of the software in Infratec instruments is planned to provide a warning when the instrument needs to be cleaned.

Tables 1 to 4 show population statistics over all 23 instruments. In each table, unstandardised results are compared with results for the most successful standardisation option where wavelength standardisation was done using a single plastic standard and three free grain samples were used for calculating photometric adjustments.

From these results we see that ISI standardisation reduced *RMSEP* primarily through the reduction of bias between instruments. Reductions in *SEP* were also found, indicating the spectral standardisation had the effect of tightening the distribution of points about a regression line. On a population basis, results for neural network models were usually better than those for models based on partial least squares.

The results from the Danish trial can be summarised thus:

1) ISI standardisation, using free-grain samples, worked under field conditions.
2) Grain Cell standardisation was not successful, being affected by the cleanliness of the instruments.
3) The implementation of a sub-sample rejection system would be needed before ISI standardisation could routinely be used in a network.
4) Although photometric standardisation was the dominant effect, some instruments were improved by the application of wavelength standardisation.

Table 1. Transferability statistics for barley moisture using ANN and PLS models.

Barley moisture		Original statistics				% of unstandardised results			
		RMSEP	*SEP*	BIAS	*SD* Bias	*RMSEP*	*SEP*	BIAS	*SD* Bias
ANN	UnStd	0.118	0.077	0.065	0.067	100.00	100.00	100.00	100.00
	Std	0.069	0.057	0.006	0.043	58.72	74.23	9.30	64.71
PLS	UnStd	0.257	0.120	–0.175	0.153	100.00	100.00	100.00	100.00
	Std	0.111	0.089	0.009	0.071	43.03	74.22	4.92	46.74

Table 2. Transferability statistics for barley protein using ANN and PLS models.

Barley protein		Original statistics				% of unstandardised results			
		RMSEP	*SEP*	Bias	*SD* Bias	*RMSEP*	*SEP*	Bias	*SD* Bias
ANN	UnStd	0.236	0.156	–0.115	0.145	100.00	100.00	100.00	100.00
	Std	0.166	0.141	0.027	0.094	70.16	90.32	23.15	64.70
PLS	UnStd	0.301	0.172	–0.152	0.206	100.00	100.00	100.00	100.00
	Std	0.188	0.160	0.037	0.105	62.52	93.10	24.57	50.88

Table 3. Transferability statistics for wheat moisture using ANN and PLS models.

Wheat moisture		Original statistics				% of unstandardised results			
		RMSEP	*SEP*	Bias	*SD* Bias	*RMSEP*	*SEP*	Bias	*SD* Bias
ANN	UnStd.	0.118	0.086	0.057	0.064	100.00	100.00	100.00	100.00
	Std.	0.067	0.064	0.006	0.026	56.98	75.20	11.01	40.75
PLS	UnStd.	0.478	0.164	–0.347	0.296	100.00	100.00	100.00	100.00
	Std.	0.119	0.104	0.045	0.049	24.84	63.23	12.91	16.44

Table 4. Transferability statistics for wheat protein using ANN and PLS models.

Wheat protein		Original statistics				% of unstandardised results			
		RMSEP	*SEP*	Bias	*SD* Bias	*RMSEP*	*SEP*	Bias	*SD* Bias
ANN	UnStd	0.269	0.192	–0.122	0.158	100.00	100.00	100.00	100.00
	Std	0.182	0.165	–0.051	0.077	67.48	85.46	41.66	48.79
PLS	UnStd.	0.312	0.183	–0.190	0.178	100.00	100.00	100.00	100.00
	Std.	0.155	0.136	–0.028	0.080	49.61	74.43	14.50	45.04

Conclusions

Standardisation has the potential to minimise differences between instruments in a network to a point where, effectively, all samples analysed by standardised instruments appear spectrally as if they had been produced by a single instrument. The ISI standardisation process has been shown to reduce the differences between free grain near infrared transmission analysers to a point where the residual biases would not be important in a network situation. This work also identified factors such as instrument cleanliness and rejection of corrupted sub-samples that affected the performance of the standardisation process.

Acknowledgements

The authors would like to acknowledge the contributions of the following colleagues and friends. Mikael Personn and production staff at Foss Tecator in Sweden for collection of data in 1998: Tomas Bergkvist from Foss Tecator and Nigel Hewitt from Foss Development in York, who collected data in the field: Nils Bo Büchmann and Torsten Hansen from the Danish Cooperative Farm Supply who provided logistical support for planning and carrying out the 1999 field trial and, finally, Nigel Hewitt, Fiona Bury and Robin Withey from Foss Development and Robin Malm from Foss Tecator who were involved in data processing.

References

1. J.S. Shenk, M.O. Westerhaus and W.C. Templeton, *Crop Sci*. **25,** 159 (1985).
2. J.S. Shenk and M.O. Westerhaus, *Crop Sci*. **31,** 1684 (1991).
3. N.B. Büchmann, H. Josefsson and I.A. Cowe, *J. Near Infrared Spectrosc*., in press.
4. I.A. Cowe and A.M.C. Davies, in *Near Infrared Spectroscopy: The Future Waves*, Ed by A.M.C. Davies and P. Williams. NIR Publications, Chichester, UK, p. 119 (1996).