ELSEVIER

Analytica Chimica Acta 297 (1994) 405–416

ANALYTICA CHIMICA ACTA

Calibration transfer across near-infrared spectrometric instruments using Shenk's algorithm: effects of different standardisation samples

E. Bouveresse [a], D.L. Massart [a,*], P. Dardenne [b]

[a] *ChemoAC, Pharmaceutical institute, Vrije Universiteit Brussel, Laarbeeklaan 103, B-1090 Brussels, Belgium*

[b] *Station de Haute Belgique, Rue de Serpont 100, B-6800 Libramont-Chevigny, Belgium*

Received 19 April 1994; revised manuscript received 24 May 1994

Abstract

In this paper, transfer of calibration models between near-infrared spectrometric instruments using three different standardisation sets is described. The first set contains samples which are very similar to the agricultural samples from three different sets to be analysed, the second set contains generic standards, and the third one contains pure organic and inorganic chemicals. To test the accuracy of each standardisation, root mean square errors and correlation coefficients are computed before and after standardisation. To test the predictive ability, standard error of prediction for the three different prediction sets are also computed before and after standardisation. For standardisation, Shenk's algorithm is used: a description of this algorithm is given.

Keywords: Infrared spectrometry; Calibration models; Shenk's algorithm

1. Introduction

Standardisation

Several problems due to poor instrument performances and a need for all calibrations have recently surfaced in spectroscopy. They concern standardisation, i.e. the possibility to transfer the calibration from one system to another.

The first of the encountered problems is the transfer of a calibration model across two different instruments. If the instrumental response of a first system differs from that which would have been obtained with a second system (this can be due to many reasons such as different sources, different optical systems, different detectors, etc.), the calibration model built on the first system and applied to the second one will give erroneous results. Multivariate instrument standardisation enables to correct the differences between instruments, and avoids to use time-consuming complete recalibration procedures, and to transport recalibration samples from one place to another.

The second problem is the calibration transfer from an instrument to itself over a period of time. The instrumental responses from a single instrument over a period of time can be subject to important fluctuations such as temperature variations [1], wavelength shifts, linear or non-linear drifts, etc. If such fluctuations occur between the calibration procedure and the analysis procedure of unknown samples, or if the instrument needs to be repaired because of technical problems, this could lead to erroneous results. Multivariate instrument standar-

* Corresponding author.

SSDI 0003-2670(94)00237-G

disation enables to correct these fluctuations over time, without repeating the whole recalibration.

Standardisation methods: two different approaches

In order to transfer a calibration model, several standardisation methods have been suggested. The two different approaches used to transfer the calibration across NIR instruments are the following ones:

The first approach consists of trying to transfer the spectra obtained with the secondary system to the primary system, and applying the calibration model built on the primary instrument to these modified spectra. Different methods such as direct standardisation [2], piecewise direct standardisation [2,3], and Shenk's method follow this approach [4,5].

The second approach consists of building a calibration model on a primary system, trying to transfer it to a secondary system, and applying this modified calibration model to the spectra obtained on the secondary instrument. This approach was developed by Forina et al. [6,7].

In this paper, only Shenk's method has been studied. This method is widely applied to transfer calibration between NIR instruments, when results coming from different instruments have to be compared (Quality-control nets)

It should be noted that the method used to correct these instrumental responses, and still applied today in several places, is a very simple method based on bias and slope correction of the model built on the "master" instrument. Some samples were measured on the "slave" instrument, the spectra obtained and the calibration model built on the "master" was used to compute the values to be predicted. Then, the values obtained with the model built on the "master" instrument applied to the spectra measured on the "master" instrument are regressed on those values: a linear model is computed, correcting the values found with the "slave" spectra to the right values obtained with the "master" spectra. The combination of this simple linear model with the calibration model gives a modified calibration model which gives the right values (obtained from wet chemistry) from spectra obtained on the "slave" instrument.

Standardisation samples

Three different sets of samples are involved: the first is the calibration set, the samples of which are used to build a calibration model on the "master" instrument. To transfer calibration models between instruments, two other sets of samples are used:

Samples from the standardisation set are measured on the two instruments to compute the standardisation parameters.

Samples from the prediction set are measured in the "slave" instrument, and the spectra obtained are transferred on the "master" instrument with the standardisation parameters computed with the standardisation set. Then, those transferred spectra and the calibration model built with the calibration set are used to predict the studied variables.

Beside the standardisation method, the selection of the samples used to compute the transfer parameters appears to be very important, and this problem has received very little attention in the literature. Two different approaches are possible, and the choice between these two approaches depends on different practical reasons.

The first possibility is to select the "best" subset samples from the set to be predicted (only measured on the "slave" instrument), to remeasure this set on the "master" instrument, and to compute the transfer parameters with this subset. The main advantage of this method is that the samples used for standardisation are very representative for those used for prediction, and standardisation can be successfully applied [2,3]. This approach has the following practical limitations:

The samples have to be measured on both instruments. If the composition of those samples changes over time, and if the instruments are very far from each other, this can lead to serious problems.

For each type of product to be predicted, one needs a subset of samples. If the number of different products to be analysed is very high, many different subsets have to be remeasured on the "master" instrument.

The second possibility is to use samples not coming from the prediction set, but of a similar nature. These standardisation samples are measured on the two different instruments to compute the standardisation parameters, but all samples of the prediction set are measured only at one place. This method enables to standardise spectra from samples, which are not measurable on both instruments for whatever reason. Nevertheless, a drawback of this method is that it uses samples of the same nature as those used in the prediction set, and these samples can change over time.

In this paper, this second possibility and also a third approach are studied: in this third approach, standardisation is performed with the use of more stable samples (generic standards) which are completely different from the studied samples. This approach should be very helpful to overcome the problems of stability, and should be applicable to samples from different natures. The type and the number of these generic standards should be well chosen, not to lose information during the transfer of spectra. For example, if one wants these standards to be applicable to samples with very different NIR spectra, the spectra of these generic standards should probably cover a larger range of optical density, in order to contain all possible linear and non-linear variations in the optical density (O.D.) field where the measured spectra are located.

In this article therefore, different types of standardisation samples are tested:

30 standardisation cells, made of different agronomic products.

6 standardisation cells, which contain 6 different inert standards.

12 standardisation cells, which contain organic and inorganic pure products.

Interlaboratory experiment

Since 1987, the Station de Haute Belgique (SHB, Libramont, Belgium) has been running a network of NIRS instruments, in order to analyze cereals and forages [8]. The software developed by ISI (InfraSoft International, Port Mathilda, USA), which involves Shenk's algorithm for standardisation, is the key of this network. Recently, the idea of an European Network has surfaced from participants of QUEST (food quality established by spectroscopic techniques), which is a concerted action subsidized by the CEC in the framework of the general FLAIR program (Food-Linked Agro-Industrial Research).

The standardisation and prediction sets were measured by three participants of QUEST (referred to as DA, SP, UK), and by the Station de Haute Belgique (referred to as SHB).

2. Theory: Shenk's method

Shenk's patented [5] method is based on two main steps: wavelength index correction, followed by spectral intensity correction. These two corrections are stored in a standardisation file, which is used to transfer spectra from one "slave" to one "master" instrument. The main advantage is that the software based on this method can be directly linked to NIR spectrometers, and can use raw data obtained on those spectrometers.

The main steps of this method are graphically described in Fig. 1a–e, and the following notations are used: **Xs** = matrix which contains spectra of the standardisation samples measured on the "slave" instrument; **Xm** = matrix which contains spectra of the standardisation samples measured on the "master" instrument; $X_{.i}$ = ith column of the **X** matrix (it corresponds to the responses of all samples at the ith wavelength); $X_{j.}$ = jth row of the **X** matrix (it corresponds to the spectrum of the jth sample); Nw = number of wavelengths; Ns = number of samples; $N = Ns \cdot Nw$ = number of measurements in one **X** matrix.

2.1. Mathematical treatment

Step 1

All spectra from the standardisation set are transformed by a first derivative mathematical treatment. Those spectra will be used for wavelength adjustment (steps 2 to 4).

2.2. Wavelength index correction

Step 2 (see Fig. 1a)

For each "master" instrument wavelength (i), a spectral window ($i-w$,$i+w$) of neighbouring wavelengths on the "slave" instrument is chosen, the correlations between $Xm_{.i}$ and each $Xs_{.k}$ (k from $i-w$ to $i+w$) are computed, and the "slave" instrument wavelength from this window (m), for which the absorptions $Xs_{.m}$ are most highly correlated with those measured on the "master" instrument $Xm_{.i}$ is found.

Step 3 (see Fig. 1b)

To obtain a more precise estimation of the wavelength at which the correlation is maximum, one fits a quadratic model to the wavelength with the highest correlation (m), and its two neighbouring wavelengths ($m-1$ and $m+1$):

$$\text{Correlation} = a + b \cdot i + c \cdot i^2 \qquad (1)$$

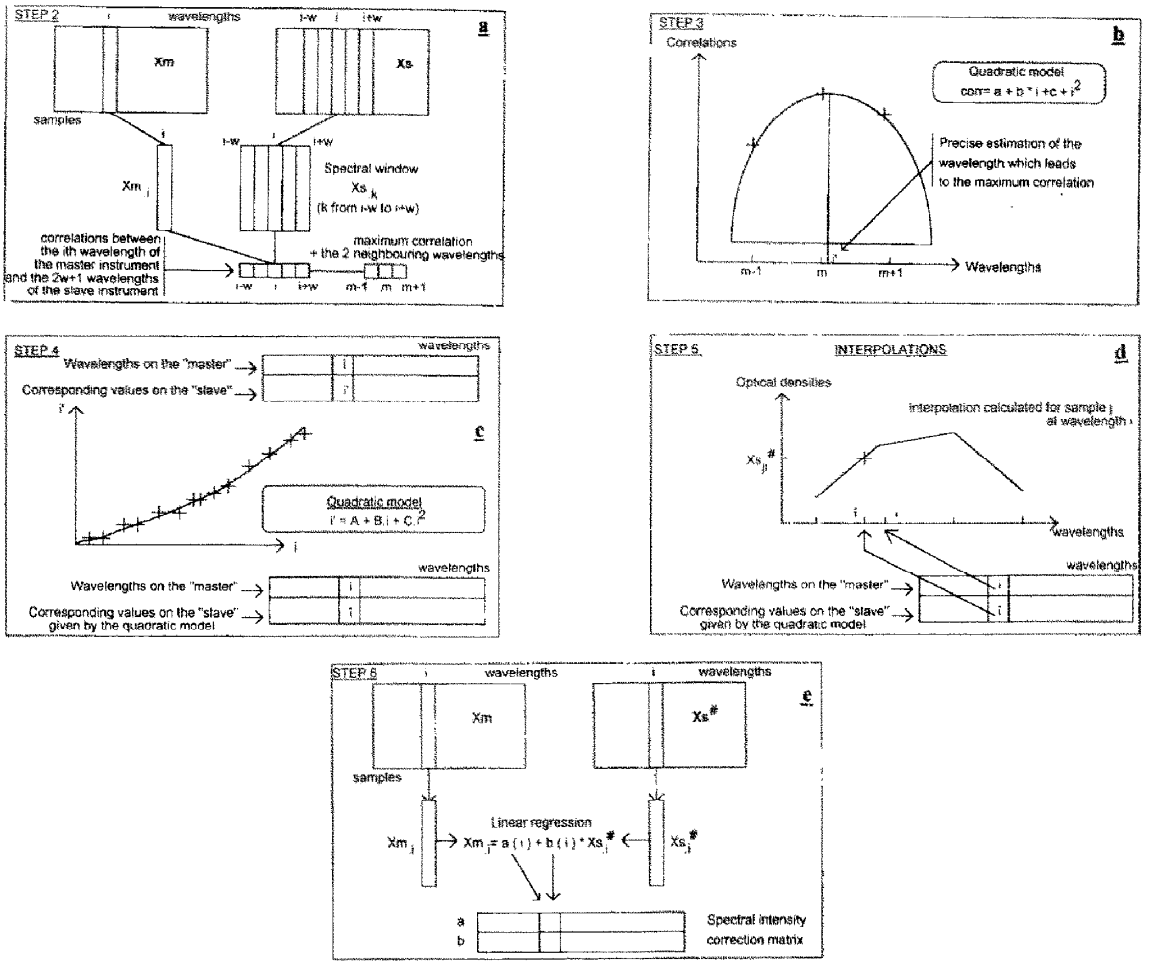

Fig. 1. Graphical description of Shenk's method (steps 2 to 6). a = step 2; b = step 3; c = step 4; d = step 5; e = step 6.

The different locations of the quadratic models obtained for each spectral "master" instrument wavelength are considered as the "slave" wavelengths (i'), that best match the corresponding master wavelengths (i).

Step 4 (see Fig. 1c)

A new quadratic model is fitted relating "master" wavelengths to their matching "slave" wavelengths ("slave" position)

$$i' = A + B \cdot i + C \cdot i^2 \qquad (2)$$

and definitive values for the "slave" wavelength (i') corresponding to a "master" wavelength (i) are obtained.

2.3. Spectral intensity correction

In steps 5 to 8, the spectra without any mathematical treatment (and not the first derivative spectra) are used.

Step 5 (see Fig. 1d)

Interpolations are performed to compute the responses of the spectra measured on the "slave" instrument at the wavelengths suggested by the quadratic model (i'). The **Xs** matrix after interpolation will be referred to as $\mathbf{Xs}^{\#}$.

for $i = 1$ to Nw, (3)

$\mathrm{Xs}^{\#}_{i} = \mathrm{Xs}_{i'}$ (computed by interpolation)

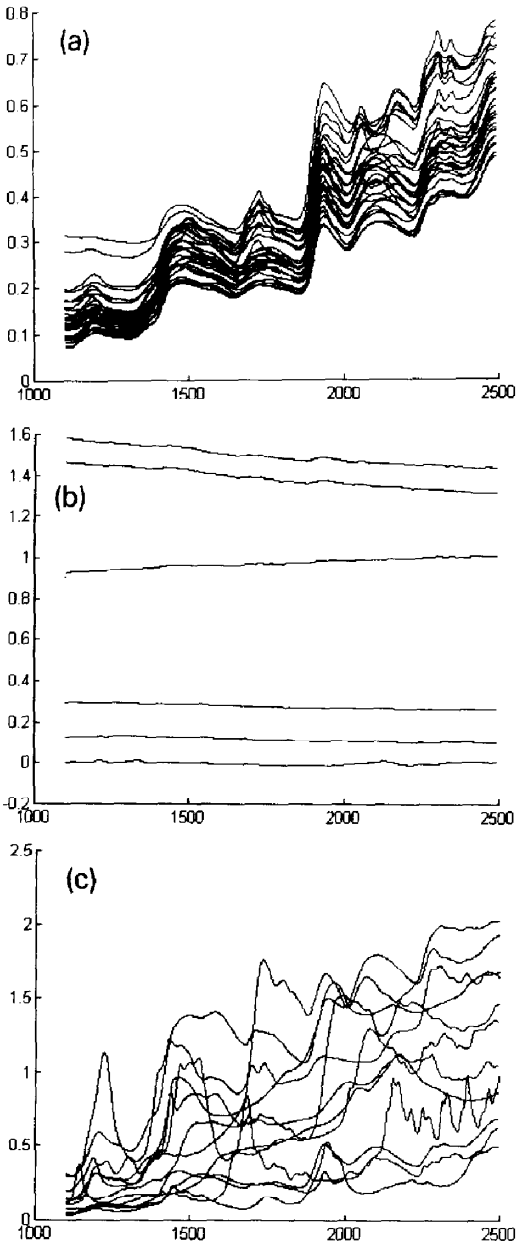

Fig. 2. Spectra of the 3 standardization sets. a = STD1 set: 30 cells with agronomic samples; b = STD2 set: 6 generic standards; c = STD3 set: 12 cells with pure organic and inorganic chemicals.

Step 6 (see Fig. 1e)

Spectral intensity correction is obtained by linear regression of the responses of the "slave" instrument at each shifted wavelength $Xs_{,i}^{\#}$ on the response of the "master" instrument at the corresponding wavelength $Xm_{,i}$.

$$Xm_{,i} = a(i) + b(i) \cdot Xs_{,i}^{\#} \quad (4)$$

where intercept (a) and slope (b) are computed for each wavelength i.

Step 7

Wavelength by wavelength, the response of the "slave" instrument is then adjusted with the corresponding regression coefficients.

$$Xstd_{,i} = a(i) + b(i) \cdot Xs_{,i}^{\#} \quad (5)$$

where **Xstd** is the **Xs** matrix after standardisation.

Step 8

The wavelength index and spectral intensity correction factors are stored in a standardisation file.

2.4. Transfer of spectra

Each spectrum obtained with the "slave" instrument can be standardised by using the standardisation file of step 8. To obtain the standardised spectra, the "slave" instrument wavelengths are shifted with the quadratic model given by Eq. 2, in order to obtain the calculated wavelength values. Interpolations are performed to compute the responses of the spectra measured on the "slave" instrument at the wavelengths suggested by the quadratic model. Finally, these "slave" instrument responses are corrected wavelength by wavelength with the regression coefficients of the spectral intensity correction matrix computed in step 6.

3. Experimental

NIRS instruments

The measurements are made on different spectrometers. The NIRSystem 6500 monochromator of SHB is referred to as the "master" instrument; The NIRSystem 5000 monochromators of UK, DA and SP will be referred to as the "slave" instruments. Before making the measurements, the instruments have been checked up and set up according to the diagnostic procedure of ISI's software: noise level, detector response and wavelength accuracy [9].

Table 1
(a) Name and composition of the standardization and prediction sets

Set name	Description
STD1	30 sealed cups with agronomic products (ISI standardization set)
STD2	6 generic standards (4 from Labsphere + 2 made by SHB)
STD3	12 sealed cups with pure organic and inorganic products
HE	6 grass samples
MA	6 corn samples
CO	17 colza samples

(b) Studied variables for each agronomic samples

Agronomic samples	Studied variables
HE	Proteins, cellulose
MA	Proteins, cellulose
CO	Proteins, fat

Standardisation samples

Three different types of standardisation samples were tested:

30 standardisation cells, made of different agronomic products by InfraSoft International (ISI, Port Mathilda, USA). These cells contain dried powders of different products such as grass, corn, wheat, colza, etc.... This standardisation set will be referred to as STD1.

6 standardisation cells, which contain different standards: 4 standards made by Labsphere, INC. (North Sutton, USA), and 2 made in the Station de Haute Belgique (Libramont, Belgium). These standards have very flat NIR spectra, but they cover a larger range of O.D. than the former 30 samples. This standardisation set will be referred to as STD2.

12 standardisation cells, made by the Station de Haute Belgique (Libramont, Belgium). These cells contain different pure products (urea, glyceride, butyric acid, polyethylene, ammonium carbonate, starch, naphthalene, lactose, saccharose, calcium sulfate, citric acid, oxalic acid). These products have NIR spectra which cover a larger range of O.D. than the former 30 samples. This standardisation set will be referred to as STD3.

Sample measurements

Before being measured by each "slave" system, the samples were measured once in SHB. The measurements consist of 6 different sets of samples (see Table 1a). The STD1, STD2, and STD3 sets contain standardisation samples, and the HE, MA, and CO sets contain agronomic samples (HE = grass, MA = corn, CO = rapeseed) used as the prediction set. The studied variables for all agronomic samples are summarized in Table 1b.

For each sample, two spectra are measured to check the repeatabilities. From these duplicates, the mean spectra are calculated. These are used in the calculations. Trimming of spectra obtained with 6500 NIRSystems is done, to obtain the same wavelength range (1100–2498 nm, step 2 nm) as the spectra obtained with NIRSystems 5000. Spectra of each standardisation set are plotted in Figs. 2a to 2c, and spectra of each prediction set are plotted in Figs. 3a to 3c.

Software

Pretreatment of spectra, determinations of standardisation files, transfers of spectra, and statistical calculations are performed with ISI's software. Graphical displays and further treatments of the results are performed in Matlab 4.0. Conversion of data from ISI's software to MATLAB and vice versa are performed by programs in QBASIC.

Calibration on the master instrument

For each set of agronomic samples, a calibration model is built on the "master" instrument in SHB. The calibration equations obtained are used to estimate the values of the studied variables given in Table 1b.

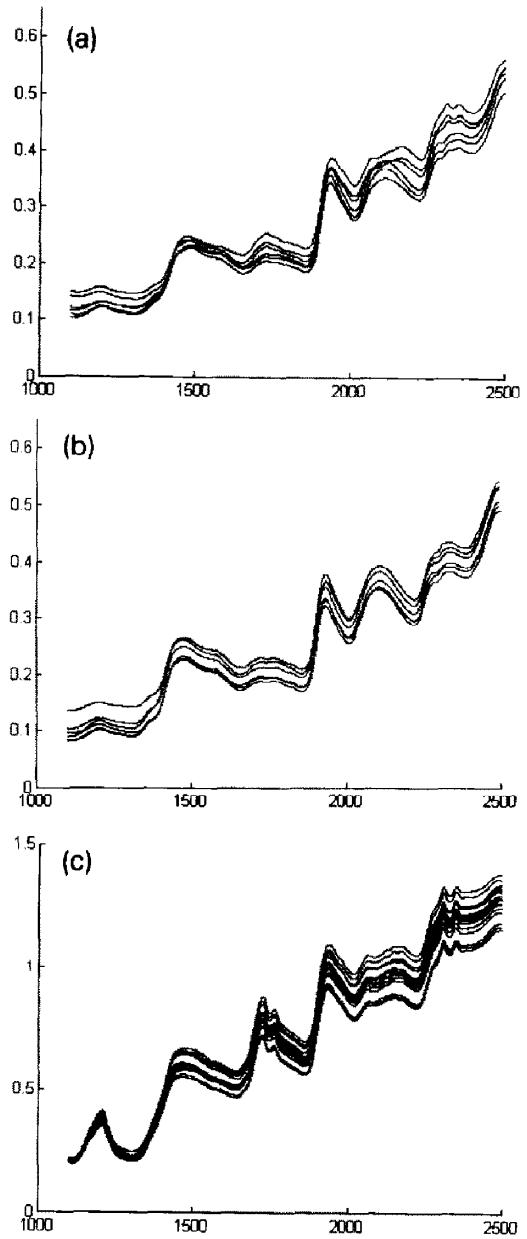

Fig. 3. Spectra of the 3 prediction sets. a = HE set: 6 grass samples; b = MA set: 6 corn samples; c = CO set: 17 colza samples.

Details about those calibration models are given in Table 2.

Standardisation

The root mean square errors RMSE (Eq. 6), the root mean square errors corrected for bias RMSE(c) (Eq. 7) and the determination coefficients R^2 (Eq. 8) between non-standardised spectra obtained on the "slave" instruments and spectra obtained on the "master" instrument are computed.

$$\mathrm{RMSE} = \frac{\sqrt{\sum_{i=1}^{Nw}\sum_{j=1}^{Ns}(Xm_{ji} - Xs_{ji})^2}}{\sqrt{N}} \tag{6}$$

RMSE(c)

$$= \frac{\sqrt{\sum_{i=1}^{Nw}\sum_{j=1}^{Ns}(Xm_{ji} - Xs_{ji})^2 - \left[\sum_{i=1}^{Nw}\sum_{j=1}^{Ns}(Xm_{ji} - Xs_{ji})\right]^2 / N}}{\sqrt{N-1}} \tag{7}$$

$$R^2 = \frac{\left(\sum_{i=1}^{Nw}\sum_{j=1}^{Ns}(Xm_{ji} \cdot Xs_{ji})^2 - \sum_{i=1}^{Nw}\sum_{j=1}^{Ns}(Xm_{ji}) \cdot \sum_{i=1}^{Nw}\sum_{j=1}^{Ns}(Xs_{ji})/N\right)/(N-1)}{(S_m \cdot S_s)} \tag{8}$$

where S_m and S_s are the standard deviations computed with all the values obtained on the "master" and on the "slave" instrument.

With the calibration equations built on the "master" instrument, prediction for the different variables is calculated with the non-standardised spectra, and standard errors of prediction SEP (Eq. 9) are computed. In the literature, SEP gives usually the differences between computed values and reference values. In this paper, the SEP has a quite different meaning: in this case, the reference values used are the values obtained with the spectra measured on the "master" instrument. In other words, SEP gives the differences between values computed with spectra obtained on the "slave" instrument (standardised or not), and values computed with spectra obtained on the "master" instrument.

Table 2
Calibration models: numbers of samples involved, means of the values obtained, Standard error of calibration (SEC), and determination coefficient

Variables	Set	Number of samples	Mean	SEC	R^2
Proteins	HE	564	15.56	0.85	0.98
Cellulose	HE	563	26.21	1.27	0.96
Proteins	MA	482	7.78	0.43	0.90
Cellulose	MA	449	20.70	1.04	0.93
Proteins	CO	382	22.47	0.56	0.94
Fat	CO	354	45.85	0.73	0.91

Table 3
Root mean square errors (microlog) for each prediction set and for each "slave" instrument

		Before	STD1	STD2	STD2′	STD3
UK	HE	5779	6579	7279	6740	9429
	MA	6174	7618	8394	7848	11199
	CO	6513	11126	10599	9332	8187
SP	HE	8737	1460	5395	5875	10484
	MA	8435	1169	5439	5898	11130
	CO	14004	9148	12349	12428	11292
DA	HE	4043	2659	6664	6948	13693
	MA	3703	1957	6164	6119	14786
	CO	10660	9437	15974	16255	17913

Table 4
Root mean square errors corrected for bias (microlog) for each prediction set and for each "slave" instrument. (*) indicates acceptable RMSE(c) values

		Before	STD1	STD2	STD2′	STD3
UK	HE	5306	4457	4966	4619	6072
	MA	4867	4113	4590	4213	6331
	CO	5300	6274	8164	7315	7830
SP	HE	2157	453*	5738	4451	3663
	MA	2115	508*	4231	4503	5873
	CO	7053	5575	4314	7333	6551
DA	HE	3518	1770*	5237	4923	8316
	MA	3583	1748*	5601	5144	9037
	CO	7676	6301	11844	9695	10538

Table 5
Determination coefficients between spectra obtained on the "master" instrument and spectra obtained on the "slave" one, before and after standardization

		Before	STD1	STD2	STD2′	STD3
UK	HE	0.9982	0.9998	0.9984	0.9985	0.9984
	MA	0.9982	0.9998	0.9984	0.9986	0.9984
	CO	0.9998	0.9998	0.9997	0.9997	0.9997
SP	HE	0.9999	1.0000	0.9997	0.9999	0.9994
	MA	0.9999	1.0000	0.9996	0.9998	0.9992
	CO	0.9996	0.9999	0.9995	0.9997	0.9999
DA	HE	0.9996	0.9998	0.9987	0.9992	0.9982
	MA	0.9995	0.9999	0.9984	0.9991	0.9982
	CO	0.9999	0.9999	0.9988	0.9993	0.9998

$$\mathrm{SEP} = \frac{\sqrt{\Sigma (Y_m - Y_{pred})^2}}{\sqrt{Ns}} \qquad (9)$$

where Y_m are the values of the variables to be predicted (e.g., concentrations), obtained with the spectra measured on the "master" instrument, where Y_{pred} are the values of the variables to be predicted, obtained with the "spectra" measured on the "slave" instrument, with or without standardisation.

Four different standardisations are computed: the first three involve the complete method of Shenk and one of the three sets of standardisation samples, and the last one involves the STD2 set of standardisation samples, and Shenk's method without the wavelength adjustment (this standardisation will be referred to as STD2′). This is due to the fact that this wavelength adjustment uses derivative spectra, and in the case of the samples of the STD2 set, which have very flat NIR spectra, derivative spectra will not contain any information.

After all standardisation files are computed, RMSE, RMSE(c), and determination coefficients (R^2) between standardised spectra and spectra obtained on the "master" instrument are computed. With the calibration equations built on the "master" instrument, prediction for the same variables is calculated with the standardised spectra, and standard errors of prediction (SEP) are calculated and compared to those obtained with the non-standardised spectra.

4. Results and discussion

Before we describe the results, it should be noted that all the instruments are 5000 and 6500 NIRSystems and therefore very similar. The results and the conclusions should be interpreted in that context. It is probable that transfer between more different instruments would lead to other conclusions.

The results concerning RMSE, RMSE(c), and correlation coefficients between spectra obtained in SHB and spectra obtained in UK, SP, DA (non-standardised and standardised with the 4 different sets) are summarized in Tables 3, 4 and 5. The values with star in Table 4 are acceptable RMSE(c) values for standardised spectra (values under 2000 μlog).

To test the predictive ability, SEP are given for each test set (HE, MA, CO) measured in UK, SP, and DA (Tables 6–8). SEC obtained on the master instrument are given for comparison. The SEP obtained should be smaller than half the corresponding SEC: if SEP is

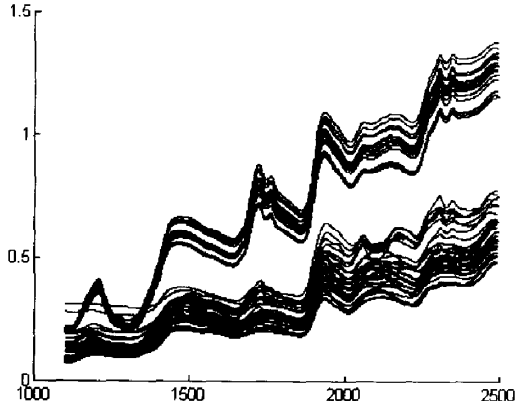

Fig. 4. Comparison between the CO and STD1 sets.

larger or of the same order of magnitude, standardisation is not useful.

Standardisation with the 30 agronomic samples gives the best results: this must be due to the fact that the 30 samples are of the same nature as the samples to be predicted, and that the 30 spectra have a similar optical density (O.D.) range as the HE and MA spectra. Fig. 4 shows that the CO spectra are in a different O.D. range compared to the 30 standardisation samples. The bad results obtained with CO samples can be explained by the different natures of the samples, but can also be due to the fact that one is measuring in different O.D. ranges. This was in fact the reason for trying to use generic standards, which cover a wider O.D. range.

Standardisation with STD2′ gives results, which are not as good as those obtained with STD1; RMSE(c) and R^2 coefficients are not good, SEP are average.

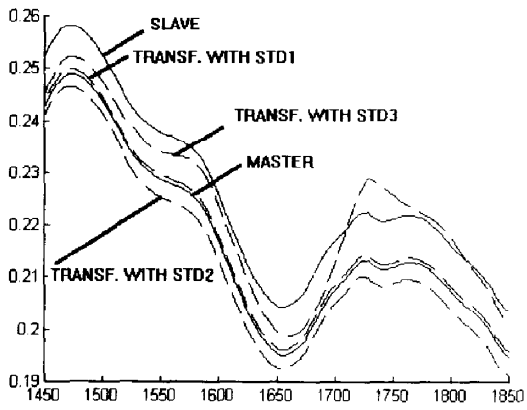

Fig. 5. Comparison of the spectra of one corn sample (MA set) obtained with the "master" instrument, with the "slave" instrument, and transfered with one of the three STD sets.

Standardisation with the STD2 set gives results which show only small differences with the STD2′ set. This proves that wavelength shifts between those NIR-Systems are very small.

Standardisation with the STD3 set gives bad results: spectra after standardisation are often more different than before.

Fig. 5 represents a zoomed part of spectra from one sample of the MA set in order to show how similar the standardised spectra are to the spectra obtained on the "master" instrument.

Figs. 6a and 6b show the mean of the differences and the standard deviation of the differences between spectra obtained in SHB and in SP for the MA, HE, CO and STD1 sets. Those two figures show that the mean and the standard deviation of the differences between SHB and SP for CO and STD1 are very different. This could be due to the different O.D. ranges of the two sets. The comparison of MA and HE with STD1 seems to confirm this assumption: thus, the mean and the standard deviation of the spectra from MA, HE, and STD1 are very close together, and those three sets are exactly in the same O.D. range. This could explain the results obtained in Tables 3, 4 and 5. Differences between spectra from one instrument to another seem to depend on the O.D. range, where the spectra are located.

Figs. 7 and 8 show the mean of the differences and the standard deviation of the differences between spectra obtained in SHB and in SP for the MA, HE, CO and STD2 sets (Fig. 7a–b), and for the MA, HE, CO and STD3 set (Fig. 8a–b). The curves corresponding to the STD2 and STD3 sets are very far from those of the HE, MA, and CO sets, and standardisations with STD2 and

Table 6
(a) Standard errors of prediction (proteins) for the HE set

	SEC	Before	Std1	Std2	Std2′	Std3
UK	0.85	0.66	0.21	0.35	0.45	1.18
SP	0.85	0.37	0.14	0.73	0.69	1.68
DA	0.85	0.53	0.45	1.38	0.47	3.87

(b) Standard errors of prediction (cellulose) for the HE set

	SEC	Before	Std1	Std2	Std2′	Std3
UK	1.27	0.45	0.21	0.93	0.48	7.07
SP	1.27	1.17	0.35	1.02	0.52	0.85
DA	1.27	0.49	0.34	0.82	0.95	1.16

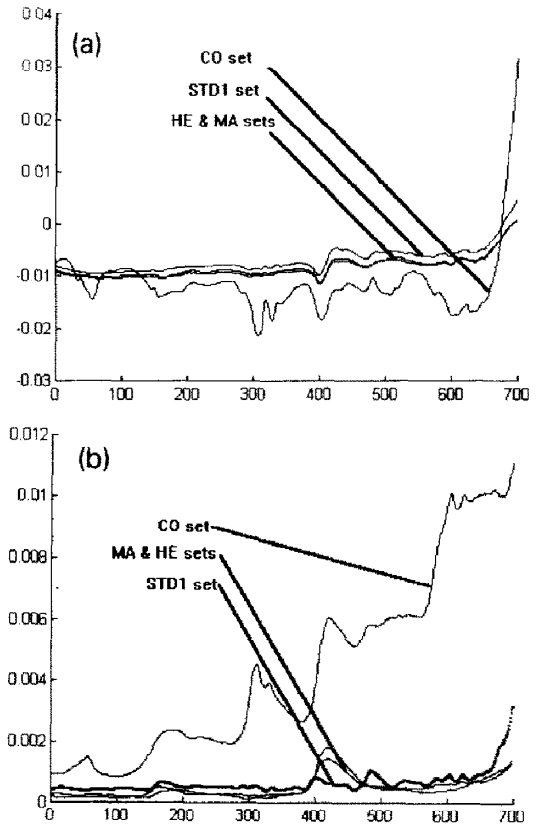

Fig. 6. Mean and standard deviation of the differences between spectra obtained in SP and in SHB for the samples of the MA, HE, CO, and STD1 sets. a = Mean; b = standard deviation.

STD3 do not give good results, as indicated in Tables 3, 4 and 5.

In other words, if one computes differences between instrumental responses of two systems with standardisation samples located in one O.D. range, and if those differences are then used to standardise spectra located in a second O.D. range different from the former, the results obtained will not be acceptable. Differences between spectra obtained on two different instruments depend probably on the O.D. ranges, where the spectra are located.

When standardisation samples are in the same O.D. range as the samples to be predicted, the standardisation parameters computed should give good results. Nevertheless, some unexpected problems might appear: in Fig. 6a, the curve corresponding to the STD1 set is quite the same as those of the MA and HE sets, but it is shifted to the top: this means that spectra of the STD1 set showed smaller differences than samples of the MA and HE sets. It should be noticed that this systematic bias (it does not depend on the wavelength) influences the RMSE, which is significantly larger than the RMSE(c). However, the standard errors of prediction after standardisation with the STD1 set are not influenced by this bias, because the calibration models developed to predict the studied values involve first derivative spectra, which removes constant shifts in absorbency. However in some NIR applications, one uses the raw spectra to build the calibration model. In this case, the bias would give erroneous prediction values.

The presence of this bias proves that attention should be paid to the standardisation samples chosen, and although the standardisation and prediction samples are of the same nature, some unexpected differences might appear.

To explain the presence of this bias, some hypotheses can be advanced:

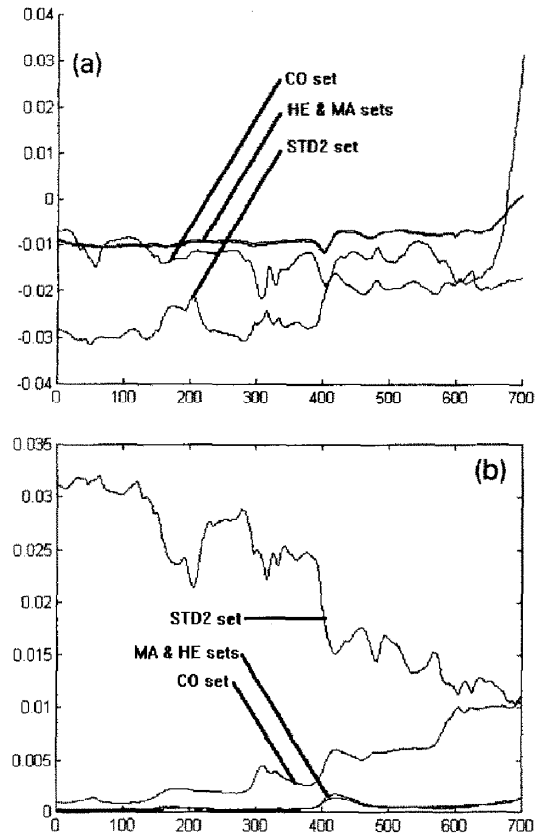

Fig. 7. Mean and standard deviation of the differences between spectra obtained in SP and in SHB for the samples of the MA, HE, CO, and STD2 sets. a = Mean; b = standard deviation.

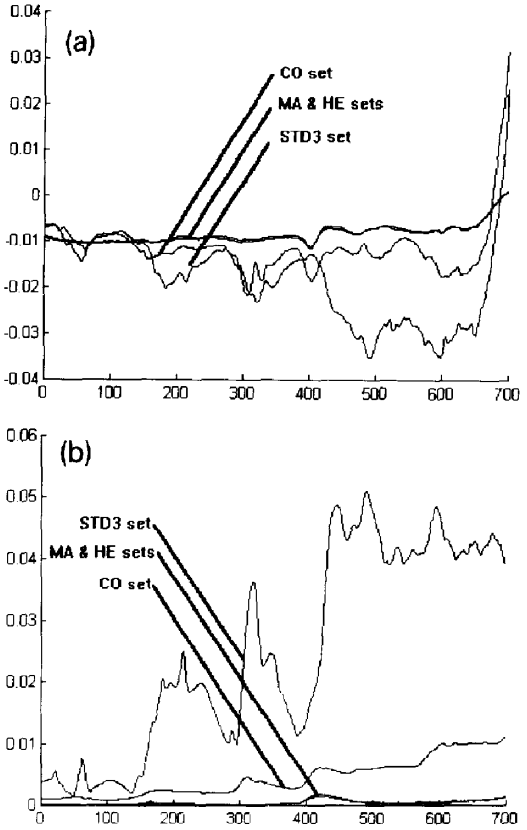

Fig. 8. Mean and standard deviation of the differences between spectra obtained in SP and in SHB for the samples of the MA, HE, CO, and STD3 sets. a = Mean; b = standard deviation.

hypothesis 1: different evolutions of the products over time (different modifications of their compositions).

hypothesis 2: different contributions of the cells used (the cells used from standardisation set and prediction set are not exactly the same).

hypothesis 3: temperature variations between measurements of the standardisation and prediction set.

The last hypothesis is improbable, because the time required for each analysis is very short (a few seconds for each sample), and because all sets are analysed one after the other: in this case, such an important shift is almost impossible.

To explain the presence of the bias, hypothesis 1, hypothesis 2, or a combination of those two seems to be more suitable, than hypothesis 3. There is probably a modification of the amount of water in the samples to be predicted between the measurements made in SHB and in SP. Thus, Fig. 6a shows some fluctuations

Table 7
(a) Standard errors of prediction (proteins) for the MA set

	SEC	Before	Std1	Std2	Std2′	Std3
UK	0.43	0.40	0.12	0.95	0.04	1.24
SP	0.43	0.23	0.10	0.89	0.14	5.71
DA	0.43	0.39	0.22	1.58	0.69	7.65

(b) Standard errors of prediction (cellulose) for the MA set

	SEC	Before	Std1	Std2	Std2′	Std3
UK	1.04	0.47	0.30	0.38	0.60	2.89
SP	1.04	0.74	0.15	0.85	0.38	2.45
DA	1.04	0.92	0.33	1.22	0.61	1.72

Table 8
(a) Standard errors of prediction (proteins) for the CO set

	SEC	Before	Std1	Std2	Std2′	Std3
UK	0.56	0.97	1.33	0.37	0.15	0.99
SP	0.56	0.95	1.08	2.65	1.13	1.20
DA	0.56	0.51	1.23	4.23	2.08	1.93

(b) Standard errors of prediction (fat) for the CO set

	SEC	Before	Std1	Std2	Std2′	Std3
UK	0.73	1.30	0.64	1.14	0.74	6.79
SP	0.73	0.64	1.41	0.52	1.06	4.16
DA	0.73	0.54	0.95	2.15	1.20	5.54

of the bias between standardisation and prediction samples. Moreover, Fig. 6b shows humps at 1900–2000 nm for the spectra to be predicted, but no hump for the standardisation samples is present. In fact, those different origins of the amount of water are probably due to the fact, that standardisation samples are in cells, but prediction samples are in normal cups (they are more sensitive to external variations). As can be seen here, different evolutions of the samples do not lead to a constant bias, but to local variations at particular wavelengths. This remark shows that hypothesis 2 seems to be improbable to explain the origin of the bias.

5. Conclusions

Differences between spectra obtained on two different instruments depend probably on the O.D. ranges,

where the spectra are located. If the standardisation samples used to correct differences between instruments are very different from the samples of the prediction set, and particularly if the differences between "master" and "slave" instruments in the same spectral region are different in the standardisation set compared to the prediction set, this could lead to poor predictions.

It appears that one needs standardisation samples which cover exactly the same O.D. range as the prediction samples. In this case, the best standardisation set used to standardise spectra from one prediction set should be the prediction set itself, or a representative subset of it. Another possibility could also be to find standardisation samples which cover exactly the same range for each spectral region, but to know in which region they are, one needs to remeasure some samples on the "master" instrument to compute the differences. If some samples have to be remeasured, standardisation with subset selection seems to be a better choice.

If we decide that products very similar to the prediction samples (e.g., grass samples to predict other grass samples) can be used for standardisation without any remeasure of prediction samples on the "master" instrument, some problems might appear: in this paper, it is shown that a systematic bias can surface, although the products used in the two sets (standardisation and prediction) are very similar. If possible to remeasure some samples on the "master" instrument, this is to be recommended. In that case, as said higher, standardisation with subset selection has to be applied, since no systematic unexpected difference between standardisation and prediction sets will occur with this method.

In case of standardisation samples different from those used for prediction, a good compromise has to be found, between too similar standardisation samples (which could only be applicable for the prediction of particular samples) and too different standardisation samples (which could lead to bad standardisations).

Acknowledgements

We would like to thank I. Murray (SAC Aberdeen, UK), A. Garrido-Varo (Escuela Tecnica Superior de Ingenieros Agronomos y de Montes, Cordoba, Spain) and L.K. Sørensen (Steins Laboratorium, Albertslund, Denmark), participants of QUEST, for NIRS measurements.

References

[1] Y. Wang and B.R. Kowalski, Anal. Chem., 65 (1993) 1301.
[2] Y. Wang, D.J. Veltkamp and B.R. Kowalski, Anal. Chem., 63 (1991) 2750.
[3] Y. Wang, M.J. Lysaght and B.R. Kowalski, Anal. Chem., 64 (1992) 562.
[4] J.S. Shenk and M.O. Westerhaus, Crop Sci., 31 (1991) 1694.
[5] J.S. Shenk, M.O. Westerhaus, U.S. Pat., 4866644, Sept 12 (1989).
[6] M. Forina, C. Armanino and R. Giangiacomo, in Hildrum, Isaksson, Naes, Tandberg (Eds.), Near Infrared Spectroscopy, Ellis Horwood, Chichester, 1992.
[7] M. Forina, G. Drava, C. Armanino, R. Boggia, S. Lanteri, R. Leardi, Chemom. Intell. Lab. Syst., in press
[8] P. Dardenne, R. Biston, in R. Biston and N. Bartiaux-Thill (Eds.), Proc. Third International Conference on Near Infrared Spectroscopy, Agric. Res. Centre Publ., Belgium, 1991, p. 655.
[9] J.S. Shenk, in R. Biston and N. Bartiaux-Thill (Eds.), Proc. Third International Conference on Near Infrared Spectroscopy, Agric. Res. Centre Publ., Belgium, 1991, p. 649.