PAPER

Check for updates

Calibration model transfer in mid-infrared process analysis with *in situ* attenuated total reflectance immersion probes†

Andrew J. Parrott, [a] Allyson C. McIntyre,[a] Megan Holden,[a] Gary Colquhoun,[b] Zeng-Ping Chen,[c] David Littlejohn[a] and Alison Nordon [*a]

Process applications of mid-infrared (MIR) spectrometry may involve replacement of the spectrometer and/or measurement probe, which generally requires a calibration transfer method to maintain the accuracy of analysis. In this study, direct standardisation (DS), piecewise direct standardisation (PDS) and spectral space transformation (SST) were compared for analysis of ternary mixtures of acetone, ethanol and ethyl acetate. Three calibration transfer examples were considered: changing the spectrometer, multiplexing two probes to a spectrometer, and changing the diameter of the attenuated total reflectance (ATR) probe (as might be required when scaling up from lab to process analysis). In each case, DS, PDS and SST improved the accuracy of prediction for the test samples, analysed on a secondary spectrometer–probe combination, using a calibration model developed on the primary system. When the probe diameter was changed, a scaling step was incorporated into SST to compensate for the change in absorbance caused by the difference in ATR crystal size. SST had some advantages over DS and PDS: DS was sensitive to the choice of standardisation samples, and PDS required optimisation of the window size parameter (which also required an extra standardisation sample). SST only required a single parameter to be chosen: the number of principal components, which can be set equal to the number of standardisation samples when a low number of standards ($n < 7$) are used, which is preferred to minimise the time required to transfer the calibration model.

1. Introduction

Quantitative analysis of multi-component samples using spectroscopic techniques such as near- and mid-infrared absorption and Raman scattering often requires the construction of multivariate calibration models. The development of these models generally requires substantial investment of resources and time.[1–3] Unfortunately if the instrumentation or measurement conditions change after the formation of the model, the calibration performance can be significantly degraded.[1,3–6] Therefore, a number of strategies have been devised to maintain the performance of multivariate calibration models.[1,3–5,7]

[a]WestCHEM, Department of Pure and Applied Chemistry and CPACT, University of Strathclyde, 295 Cathedral Street, Glasgow, G1 1XL, UK. E-mail: alison.nordon@strath.ac.uk

[b]Fibre Photonics Australia Pty Ltd, Forestville, Sydney, 2087, NSW, Australia

[c]State Key Laboratory of Chemo/Biosensing and Chemometrics, College of Chemistry and Chemical Engineering, Hunan University, Changsha, 410082, Hunan, China

One solution is to build a calibration model that is robust to changes in conditions.[5,8] This can involve construction of a global model that contains all of the expected variation for a range of measurement or instrumental conditions. However, it is not always possible to anticipate new sources of variation or define the extent of their variability in the data. In addition, global models tend to require a large number of calibration samples.[1,5] Alternatively, data pre-processing, such as multiplicative signal correction,[9] finite impulse response filtering,[10] orthogonal methods,[8,11,12] generalized least squares,[13] and wavelength selection can be employed to reduce the sensitivity of the resulting model to changes in conditions.[14] A down side of these methods can be that they tend to remove variation not present in all of the different conditions (or different instruments), and so can reduce the sensitivity of the model to the analyte of interest.[6,15]

One of the simplest ways of dealing with changes in measurement or instrumental conditions is to update the original model through addition of spectra acquired under the new set of conditions.[6,7] The drawback of this method is that a large number of samples may need to be added to prevent domination of the model by the original calibration samples. However, the relative importance of a small number of new samples can be increased if they are appropriately weighted.[16] A

number of methods, including Tikhonov regularisation, have been proposed for determination of the weighting factors.[17,18]

An alternative to model updating is calibration transfer, where the aim is to maintain the predictive ability of a model developed under an initial set of conditions (denoted primary) when it is applied to spectra collected under another set of conditions (denoted secondary).[1,5–7,19] Calibration transfer methods achieve this by adjustment of the regression coefficients, the predicted values, or the spectral responses.[1,5]

A widely used, and relatively simple approach for standardisation of the predicted values is the univariate slope and bias correction (SBC) method.[1,5,20] SBC tends to work well when the spectral differences arising from the change in instrument are relatively simple and systematic for all samples. However, where this is not the case other approaches, such as standardisation of the spectral responses, should be considered.

Spectral response standardisation generally requires measurement of carefully selected standardisation samples on both the primary and secondary instruments.[1] A transformation matrix, which relates the spectral responses for the standardisation samples measured on the primary and secondary instruments, is then calculated and used to transform spectra acquired on the secondary instrument into the corresponding spectra as if they were acquired on the primary instrument.[1,5,6,19,21] This means that the calibration model developed on the primary instrument can continue to be used under the new conditions. The two most popular and well established spectral response standardisation methods are direct standardisation (DS) and piecewise direct standardisation (PDS).[19,21] In both cases, a linear model is used to relate the spectral responses on the primary and secondary instruments, but for PDS the transformation matrix is estimated by a moving window procedure. As a consequence, PDS can be used to correct non-linearities in the data and so generally outperforms DS. However, there are disadvantages associated with PDS including the need to optimise the window size (which affects the performance of the algorithm), and issues with determining the rank of each local regression model (poor estimation can lead to spectral artefacts).[1,5,22,23]

To address the limitations of DS and PDS a wide range of spectral response standardisation methods have been reported in the literature,[1–3,5–7,24–36] including several methods which do not require standardisation samples but instead rely on measurement of a selection of new samples in the new condition (or with the new instrument).[1,3,5,37–44] However many of these methods require a very high number of samples to be measured in the new condition, in some cases 50% to 100% of the size of the initial calibration set, and so in effect are approaching full recalibration.[24–28,38,42] This limits the benefits of calibration transfer in terms of savings in time and resources. Many methods also have more than one 'tuning' parameter which must be optimised to get good results,[1–3,6,25,29–33,43] which further increases the time burden of the method.

Furthermore, the vast majority of reported calibration transfer methods have only been demonstrated with near-infrared (NIR) spectra.[1,3,5,7,45] There are very few applications of calibration transfer with mid-infrared (MIR) spectra,[46–49]

especially in applications such as *in situ* measurements for process analysis and control. Advances in fibre optics have allowed access to the MIR region with fibre coupled probes, so that there are many new opportunities for *in situ* MIR measurements in process analysis.[50–52] Therefore there is a need for calibration transfer methods which can easily be applied for MIR applications.

Spectral space transformation (SST)[53] is a calibration transfer method which has been shown to have favourable performance compared to standard methods such as DS and PDS.[6,53] Advantages of SST include the ability to work with a small number of standardisation samples, and easy application because it has only one adjustable parameter. SST has been successfully demonstrated with NIR,[53,54] MIR,[53] and Raman spectroscopy.[55]

The aim of this study was to assess the applicability of SST for use with *in situ* MIR measurements using fibre coupled attenuated total reflectance (ATR) probes under a wider set of scenarios than previously reported.[53] The examples considered arise frequently in process applications of MIR spectrometry with an *in situ* ATR probe, but have not been discussed in the literature to date. These include: changing a spectrometer, multiplexing two equivalent probes to a single spectrometer, and increasing the diameter of the ATR probe to mimic transfer of a process from laboratory to pilot plant scale. Seven different combinations of MIR spectrometer and attenuated total reflectance (ATR) probes were used to analyse ternary mixtures of acetone, ethanol and ethyl acetate as a model system. Calibration models developed on one spectrometer–probe combination were then applied to data acquired on other spectrometer–probe combinations to assess the impact of changes in the spectrometer and/or probe on the predictive performance of the models. The performance of SST was compared to the standard methods DS and PDS in terms of predictive accuracy, sensitivity to the choice of standardisation samples, and ease of use. A simple modification to the SST algorithm is proposed that allows it to be easily and accurately used across a wider set of situations than explored previously.

2. Materials and methods

2.1. Samples

Sixteen samples containing acetone, ethanol and ethyl acetate were prepared according to a ternary mixture design (see Fig. 1) to assess the MIR ATR probes and spectrometers. This model system was selected to create MIR spectra typical of hydrocarbon mixtures where there are overlapping absorption features as well as some discrete bands, and taking account of the miscibility and chemical hazards of the three components. The samples were divided into calibration (samples 1 to 10) and test (samples 11 to 16) sets, with the compositions of the samples given in Table S1 in the ESI.†

2.2. Mid-infrared spectrometry

MIR spectra were acquired using three infrared spectrometer systems: an ABB MB3000 FTIR (Clairet Scientific, Northampton,
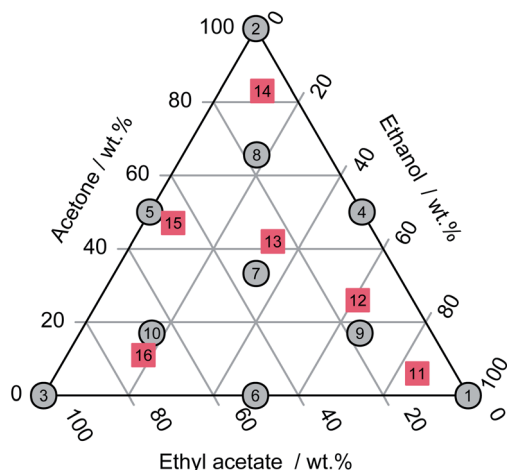
Fig. 1 Ternary mixture design used for preparation of the calibration (samples 1 to 10 denoted by grey circles) and test (samples 11 to 16 denoted by red squares) sets.

UK), an ABB BOMEN MB155 FT-NIR/IR (Clairet Scientific) operated in MIR mode, and an ABB FTLA2000 series FTIR (ABB, Québec, Canada). All spectrometers employed a DTGS detector and a SiC source. An external power supply was used with the MB155 spectrometer to allow short-term operation of the SiC source at a higher voltage to increase the intensity. Each spectrometer was coupled *via* polycrystalline silver halide fibres to Hastelloy bodied immersion probes with diamond ATR crystals of different diameters (Fibre Photonics Ltd, Livingston, UK).[50] Details of the three probes investigated in this study are given in Table 1. The ternary mixtures were analysed using seven different combinations of the three ATR probes and three MIR spectrometers (see Table 2). Spectra were acquired using 51 scans with a resolution of 16 cm^{-1} in the 400 cm^{-1} to 4000 cm^{-1} range using either Horizon MB FTIR software version 2.1.9.0 (ABB) or GRAMS/AI software version 4.04 (Thermo Scientific, UK). Absorbance spectra were calculated using an air background, which was acquired prior to analysis of the samples.

2.3. Data analysis

All data were imported into MATLAB version R2018b (Mathworks Inc., Natick, MA, USA) with PLS_Toolbox version 8.6.2 (Eigenvector Research Inc., WA, USA). The spectral region between 579 cm^{-1} to 1844 cm^{-1} was selected for analysis.

2.3.1. Calibration models. Partial least squares (PLS) calibration models were built using spectra acquired on each of the spectrometer–probe combinations and the corresponding

Table 1 Details of the three probes investigated

Probe	Outer diameter of probe shaft (mm)	Silver halide fibre length (m)	ATR diamond crystal size (mm)
1	12	1.5	3
2	12	1.7	3
3	2.7	1.1	1.2

Table 2 Combinations of spectrometer and ATR probe used to analyse the ternary mixture samples

Combination number	Spectrometer	Probe
1	MB155	1
2	MB155	2
3	MB3000	2
4	MB3000	3
5	MB3000	1
6	FTLA2000	1
7	FTLA2000	2

concentration data for samples 1 to 10. A separate model was built for each analyte (*i.e.* PLS1) resulting in a total of 21 models, for all models mean centring was used as the pre-processing method. The calibration models were then used to predict the concentrations of acetone, ethanol and ethyl acetate in samples 11 to 16 using spectra acquired on the same spectrometer–probe combination.

The models were assessed using the root mean square error of calibration (RMSEC) and root mean square error of prediction (RMSEP). The optimum number of latent variables for each model was determined from a plot of RMSEC against the 2-norm of the regression vector, $\|\mathbf{b}\|_2$, *i.e.* by consideration of both the model bias and variance.[56,57] Each plot exhibited a characteristic L-shaped curve with the optimum (most harmonious) model located in the corner region of the plot, *i.e.* the point where addition of further latent variables resulted in a large increase in variance relative to a small decrease in bias.

The performance of the calibration models when the probe and/or spectrometer was changed was assessed by applying the calibration model built on one spectrometer–probe combination (the primary system) to the test data acquired on the other six spectrometer–probe combinations (secondary system).

2.3.2. Calibration transfer. The study assessed the ability of three calibration transfer algorithms, DS, PDS and SST, to maintain the predictive ability of a model, constructed on one spectrometer–probe combination (the primary system), when applied to spectra acquired on another spectrometer–probe combination (the secondary system).

The DS and PDS algorithms used were from PLS_Toolbox. For PDS, the number of principal components (PCs) used in the calculation of the transformation matrix was determined with a tolerance value of 0.0001 (the default setting); this value gives the minimum relative size of the singular values to include in each model. The optimum window size for PDS was selected from the range 1 to 101 using a step size of 2 on the basis of the RMSEP value obtained for a calibration sample acquired on the secondary system that was not used for standardisation. The SST algorithm used is described in the paper by Du *et al.*[53] and the calculations were performed by an in-house function written in MATLAB. SST eliminates the spectral differences arising from changes in instrumentation through transformation between two spectral spaces spanned by the corresponding spectra of a subset of standardisation samples measured on the two instruments. Determination of only one

model parameter is required, *i.e.* the number of PCs representing information in the spectra of the standardisation samples. This parameter can be set to a value equal to or slightly larger than the number of significant singular values of the combined spectral matrix of the standardisation samples. Alternatively, when only a limited number of standardisation samples are available, the number of principal components can just be set equal to the number of standardisation samples.[53] Therefore, in this case, the number of PCs selected was set equal to the number of standardisation samples, *i.e.* 4, in all cases.

The sensitivity of the three transfer algorithms to the choice of standardisation samples was assessed by selecting different subsets of four calibration samples, spanning different regions in the design space (see Fig. 1). The following combination of samples was used: 1, 2, 3 and 7; 4, 5, 6 and 7; 7, 8, 9 and 10; 2, 4, 5 and 7; 1, 4, 6 and 7; and 3, 5, 6 and 7. For PDS, the window size had to be re-optimised for the different sets of standardisation samples, and was optimised using a calibration sample acquired on the secondary system that was not used for standardisation. Calibration sample 8, 9 or 10 was used in all cases except where the standardisation samples were 7, 8, 9 and 10, in this case the window size was optimised using calibration sample 1, 2 or 3.

3. Results and discussion

3.1. Effect of changing spectrometer–probe combination

Fig. 2 shows an overlay of the absorbance spectra for acetone, ethanol and ethyl acetate acquired using spectrometer–probe combination 5. The spectra of the individual components exhibit extensive overlap and consequently, multivariate regression is required for quantitative analysis of mixtures of the three analytes.

The RMSEP values obtained for calibration models built (using samples 1 to 10) and applied to data (for samples 11 to 16) acquired 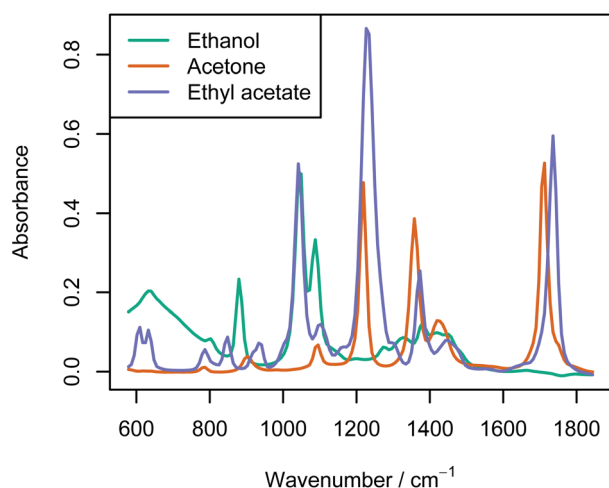on the same spectrometer–probe combination are given on the diagonal of Table 3. Such values provide a measure of the baseline performance of each of the calibrations models. All models exhibit good linearity over the entire calibration range for prediction of each of the three analytes with R^2 of between 0.977 and 1.000, as shown by Fig. S1 (in ESI†). In general, the poorest predictions were obtained for data acquired on the MB155 spectrometer (combination 1 and 2). This was the oldest instrument used in the study, and the gain of the detector had to be set to its maximum setting to obtain a measurable signal.

The off diagonal RMSEP values in Table 3 give the predictive accuracy for models built on one spectrometer–probe combination (the primary system) and used to predict the composition of samples acquired on another spectrometer–probe combination (the secondary system). In general, the predictive accuracy of the calibration models was least affected when the spectrometer was changed (*e.g.*, primary system 6 to secondary system 5). In comparison, the predictive accuracy of the calibration models was degraded most markedly when the probe diameter was changed (*e.g.*, primary system 4 to secondary system 5), as also shown by Fig. S2 in the ESI.† However, there was also some reduction in the performance of the calibration models when they were applied to data acquired using a different 12 mm diameter probe (*e.g.*, primary system 7 to secondary system 6). The reason for this can be observed in Fig. 3. There is minimal difference between the spectra of ethyl acetate acquired using the same probe (probe 1) but different spectrometers (FTLA2000 or MB3000), *i.e.* spectrometer–probe combinations 6 and 5 (Fig. 3a). In comparison, there is a significant difference between the spectra acquired using the same spectrometer (MB3000) but with probes of different diameters (probes 1 or 3), *i.e.* spectrometer–probe combinations 5 and 4 (Fig. 3c). The smaller absorbance obtained with the 2.7 mm diameter probe can be attributed to the smaller diamond cone, which gives rise to a shorter effective pathlength because there are fewer average internal reflections.[58] A smaller spectral change was observed when two probes of the same diameter (probes 1 or 2) were used, *i.e.* spectrometer–probe combinations 6 and 7 (Fig. 3b), these variations can be attributed to small differences in the design and construction of the two probes.[58]

3.2. Comparison of calibration transfer methods

The performance of three calibration transfer algorithms, DS, PDS and SST, was assessed under three different scenarios. The first scenario reflects the case of an upgrade of a spectrometer, and uses spectrometer–probe combination 6 as the primary system and spectrometer–probe combination 5 as the secondary system. The second scenario represents the situation that might occur if two probes are multiplexed to the same spectrometer, but the calibration model is developed with only one of the probes. This scenario uses spectrometer–probe combination 7 as the primary system, and combination 6 as the secondary. The final situation represents the scaling up of the probe diameter which might arise when transferring a method developed in the laboratory to a pilot plant. This example uses

Fig. 2 Absorbance spectra in the selected region, 579 cm^{-1} to 1844 cm^{-1}, for acetone, ethanol and ethyl acetate obtained using spectrometer–probe combination 5.

Table 3 RMSEP values (% w/w) for ternary mixture components obtained using calibration models developed on one spectrometer–probe combination (the primary system) and applied to data acquired on the same and different spectrometer–probe combinations (the secondary system). Figures shown in bold indicate the same calibration and test spectrometer–probe combinations. The values in parentheses indicate the number of latent variables retained on each primary system (and used when predicting concentrations from both primary and secondary test spectra)

Component	Secondary system	Primary system						
		1	2	3	4	5	6	7
Acetone	1	**3.2** (4)	8.6	8.8	65.3	6.1	7.1	6.9
	2	30.6	**5.5** (3)	5.3	91.1	15.8	17.7	10.4
	3	37.0	8.1	**4.3** (3)	92.3	13.7	15.7	8.5
	4	45.8	20.9	20.6	**3.7** (4)	19.1	20.0	22.9
	5	7.8	8.7	7.5	78.4	**2.0** (3)	1.9	6.4
	6	6.8	9.2	8.1	77.7	2.2	**1.7** (3)	7.3
	7	18.9	5.8	4.4	87.2	6.0	6.4	**1.0** (3)
Ethanol	1	**4.5** (3)	7.3	10.1	98.2	6.2	8.1	7.2
	2	13.8	**5.5** (3)	4.3	124.4	11.7	12.9	9.3
	3	13.2	6.1	**3.1** (3)	119.1	10.0	11.8	8.3
	4	27.2	32.4	35.4	**2.9** (5)	24.6	25.1	24.6
	5	7.7	5.8	7.4	108.1	**2.0** (3)	2.8	3.2
	6	8.6	7.2	8.2	110.7	2.3	**1.0** (2)	3.5
	7	12.2	4.9	3.4	129.9	5.3	4.0	**1.2** (2)
Ethyl acetate	1	**1.9** (3)	9.2	10.0	38.7	2.2	3.1	5.4
	2	7.4	**3.4** (2)	3.8	49.6	8.0	7.7	2.7
	3	4.7	3.9	**3.5** (3)	42.8	5.2	4.9	1.9
	4	25.0	32.8	35.7	**0.9** (4)	26.9	29.0	25.1
	5	1.0	8.5	9.6	39.0	**0.9** (3)	2.0	4.6
	6	1.2	8.5	10.0	40.2	0.6	**1.5** (3)	4.4
	7	6.6	2.8	3.3	51.5	6.3	5.4	**0.4** (4)

combinations 4 and 5 as the primary and secondary systems respectively. In all three scenarios, it would be desirable to be able to use the calibration model developed on the primary system with the secondary system without any reduction in predictive performance.

A comparison of the RMSEP values obtained for the different scenarios outlined above when using DS, PDS, or SST are listed in Table 4. Fig. S2 in the ESI† shows the plots of predicted *vs.* actual concentrations for calibration models constructed on a primary system and used to predict the composition of test samples acquired on a secondary system with and without SST standardisation for the three different scenarios. In general, it can be seen from Table 4 and Fig. S2† that use of DS, PDS and SST with four standardisation samples (samples 1, 2, 3 and 7) gives more accurate predictions than when the model developed on the primary system is applied to data acquired on the secondary system without any calibration transfer.

When the spectrometer was changed (scenario 1), use of DS, PDS and SST had only limited impact on the predictive accuracy of a model developed on the primary system when it was applied to test data acquired on the secondary system. This is because there was minimal difference between the spectra acquired on the two spectrometers and hence, the performance of the primary model with test data acquired on the secondary system was only marginally poorer than with that acquired on the primary system. The exception is the model for ethyl acetate where DS, PDS and SST gave slightly improved predictions compared to when the model was applied to test data acquired on the primary system. It may be that any non-linearities in the

data that are not modelled effectively by the original model can be removed by the standardisation procedures; this would give rise to improved predictions. When two different 12 mm probes were used (scenario 2), the predictive accuracy for DS, PDS and SST was comparable to that when the model was applied to test data acquired on the primary system. DS and PDS also retained the predictive accuracy of the model developed on the primary system when the diameter of the probe was changed (scenario 3); however, the performance of SST was poorer than both DS and PDS for the prediction of all analytes and in particular ethyl acetate.

When the magnitude of the primary and secondary absorbance spectra are very different (as is the case for scenario 3, as shown by Fig. 3c), the first few PCs from the singular value decomposition of the combined spectral matrix of the standardisation samples, as used in the SST algorithm,[53] will tend to explain only variations in the minor factors for the spectra with the greater absorbance. A simple method to remove this problem is to scale the spectra acquired on the secondary system. In this work, the spectra acquired on the secondary system were scaled by $\|\mathbf{X_p}\|_F/\|\mathbf{X_s}\|_F$ where $\|\mathbf{X_p}\|_F$ and $\|\mathbf{X_s}\|_F$ are the Frobenius norms[59] of the primary ($\mathbf{X_p}$) and secondary ($\mathbf{X_s}$) standardisation spectra, respectively. It can be seen from the final row in Table 4 that use of scaling prior to application of the SST algorithm reduced the RMSEP values by on average a factor of 2 such that its performance was comparable with that of DS and PDS. It should be noted that other scaling or normalisation methods could potentially be applied, and this method used was chosen as it is very straightforward to apply. The scaling
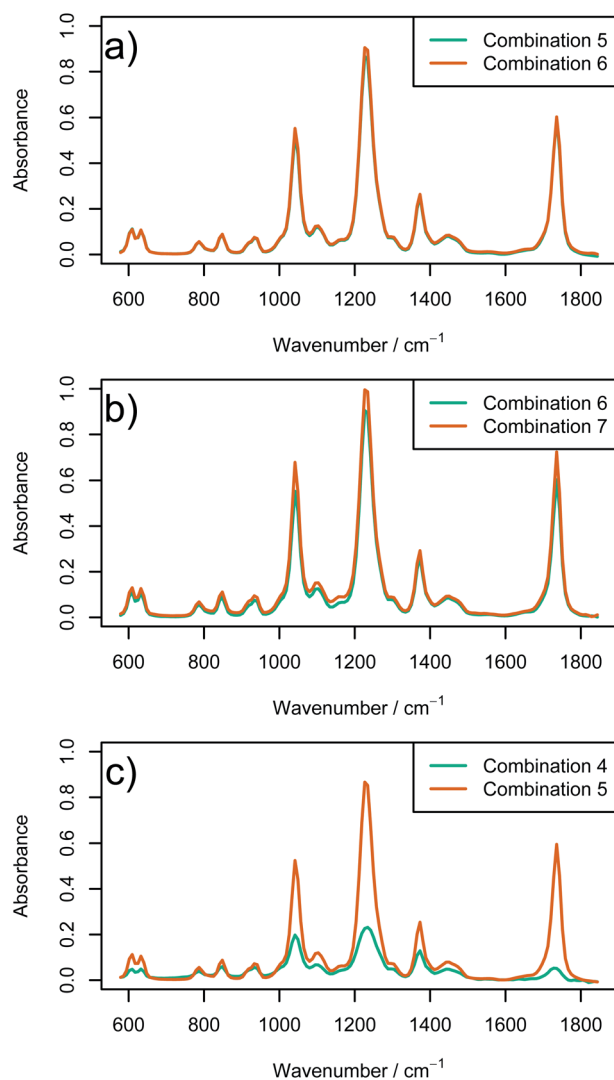
Fig. 3 Absorbance spectra in the selected region, 579 cm^{-1} to 1844 cm^{-1}, for ethyl acetate obtained using spectrometer–probe combinations (a) 5 and 6 the spectrometer upgrade scenario, (b) 6 and 7 the multiplexed probes scenario, and (c) 4 and 5 the scenario when different diameter probes were used.

step has no significant impact on the performance of SST when the magnitude of the primary and secondary absorbance spectra are comparable, as in Fig. 3a and b, as shown by comparing results of SST with and without scaling for scenario 1 and 2 as listed in Tables 4, S2 and S3.†

3.3. Choice of standardisation samples

The variation in the RMSEP values obtained for acetone, ethanol and ethyl acetate as a function of the composition of standardisation samples for DS, PDS and SST is given in Tables S2, S3, and S4 in the ESI† for scenarios 1, 2, and 3 respectively. Regardless of the choice of the standardisation samples, it can be seen from the results listed in Tables S2 to S4† that use of DS, PDS and SST improves the accuracy of the predictions when the model developed on the primary system is applied to data acquired on the secondary system. However, it is only when the

standardisation samples were 1, 2, 3 and 7 (*i.e.* spanning the entire design space) that the predictive accuracy was comparable to that where the calibration and test samples were acquired on the same system (*i.e.* primary), with the three standardisation methods performing equally well.

However, it is noticeable that the results are affected by the degree of difference between the two datasets. When the primary and secondary systems were spectrometer–probe combinations 7 and 6, respectively, the difference between the two datasets is relatively small. In this case, the predictive accuracy obtained when the standardisation samples were 7, 8, 9 and 10 was comparable to that obtained with samples 1, 2, 3 and 7. In addition, the performance of DS was more sensitive to the composition of the standardisation samples than PDS and SST. When the difference between the two datasets was relatively large, *i.e.* when the primary and secondary systems were spectrometer–probe combinations 4 and 5, respectively, the performance of DS and SST was degraded when the composition of the standardisation samples spanned only selected regions in the entire design space. In comparison, PDS was less affected by the composition of the standardisation samples. However, the performance of PDS was affected by the choice of sample used for optimisation of the window size when the standardisation samples were 2, 4, 5 and 7, 1, 4, 6 and 7, and 3, 5, 6 and 7. For these combinations of standardisation samples, which only span a small area of the design space, in general, poorer predictions were obtained for the test samples when the sample for window size optimisation was located in the centre of the design space covered by the standardisation samples. This is because in these examples, calibration transfer with PDS has been optimised to work within a very small compositional range and therefore, when attempts were made to transfer samples outside this region, poorer performance was observed. The optimum window size can vary quite widely with the choice of standardisation samples and the sample used for optimisation of the window size.

3.4. Ease of use

While DS is perhaps the easiest procedure to use in that it requires no user input or parameter optimisation, it is more sensitive to the composition of the standardisation samples than PDS and SST, irrespective of the spectral difference between the primary and secondary systems.

When the difference between the primary and secondary systems is large, PDS is the least sensitive to the choice of standardisation samples. However, PDS requires optimisation of two parameters: window size and number of PCs, both of which have a significant impact on predictive accuracy. The optimisation of the window size requires at least one extra sample in addition to the standardisation samples. In comparison, SST requires optimisation of only one parameter: the number of PCs. This can simply be set equal to the number of standardisation samples when only a limited number of samples are available ($n < 7$).[53] Accordingly, in this work the number of PCs was always set to 4, and SST was essentially performed without an adjustable parameter. For situations

Table 4 RMSEP values (% w/w) for acetone, ethanol and ethyl acetate when the calibration and test spectra were acquired on different spectrometer–probe combinations and using DS, PDS, or SST as the calibration transfer method. Samples 1, 2, 3 and 7 were used for standardisation. Values without calibration transfer are repeated from Table 3 for convenience of comparison

Scenario	Calibration (primary) system	Test (secondary) system	Standardisation method	Acetone	Ethanol	Ethyl acetate
1: Spectrometer upgrade	6	6	None	1.7	1.0	1.5
	6	5	None	1.9	2.8	2.0
	6	5	DS	2.2	2.4	0.9
	6	5	PDS[a]	1.7–2.3	2.0–2.6	0.6–0.9
	6	5	SST	2.3	2.5	0.8
	6	5	SST with scaling	2.3	2.5	0.8
2: Multiplexed probes	7	7	None	1.0	1.2	0.4
	7	6	None	7.3	3.5	4.4
	7	6	DS	1.4	1.0	1.6
	7	6	PDS[a]	1.3–1.5	0.8–0.9	1.0–1.1
	7	6	SST	1.4	1.0	1.0
	7	6	SST with scaling	1.4	1.0	1.0
3: Different probe diameter	4	4	None	3.7	2.9	0.9
	4	5	None	78.4	108.1	39.0
	4	5	DS	3.0	3.0	1.3
	4	5	PDS[a]	3.0–3.1	3.0–3.5	1.1–1.6
	4	5	SST	5.2	4.0	3.6
	4	5	SST with scaling	3.0	2.6	1.0

[a] The window size for PDS was optimised using sample 8, 9 or 10. Therefore, the ranges of RMSEP values obtained for the secondary test samples transformed using these window sizes are given. Full details are listed in Tables S2, S3, and S4 in the ESI for scenarios 1, 2, and 3 respectively.

where a larger number of standardisation samples are available, the performance of SST is relatively insensitive to the choice of number of PCs as long as this number is set higher than the number of significant singular values in the combined spectral matrix of the standardisation samples.[53]

Where spectra acquired on the primary and secondary systems differ in magnitude, then the secondary spectra should be scaled prior to application of SST. It was found that simply scaling by the ratio of the Frobenius norms yielded good results. One advantageous feature of the proposed scaling method is that it has no impact on performance for datasets where the magnitudes are comparable, therefore the SST version with the extra scaling step can be used in all situations without any extra user intervention or decisions.

4. Conclusions

The study has demonstrated that DS, PDS and SST can be used to maintain the performance of calibration models built with MIR spectra acquired using fibre coupled ATR probes, an application largely neglected in the current literature on calibration transfer. It has been demonstrated that all three methods can be effective, irrespective of the different types of spectral change caused by altering the spectrometer–probe combinations. SST, with the addition of a simple scaling step, was found to be a more straightforward method compared to DS or PDS. This is because DS is highly sensitive to the composition of the standardisation samples used, and PDS requires the optimisation of two parameters. The study has demonstrated particular advantages of SST for *in situ* MIR analysis, especially when faced with the need to replace spectrometers, ATR probes,

or to change the size of the ATR crystal during scale-up of manufacturing.

Author contributions

DL and AN conceived and supervised the study, and wrote the original draft of the manuscript. ACM and MH performed the experiments and did preliminary data analysis. GC provided resources including ATR probes. ZPC suggested the use of Frobenius norms within the scaled SST approach and wrote the initial software code. AJP validated the results, curated the data for publication, and wrote the final draft of the manuscript. DL and AN provided critical review and commentary needed for manuscript completion.

Conflicts of interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Acknowledgements

Notes and references

1 R. N. Feudale, N. A. Woody, H. Tan, A. J. Myles, S. D. Brown and J. Ferré, *Chemom. Intell. Lab. Syst.*, 2002, **64**, 181–192.

2 Y. Mou, L. Zhou, S. Yu, W. Chen, X. Zhao and X. You, *Chemom. Intell. Lab. Syst.*, 2016, **156**, 62–71.

3 B. Malli, A. Birlutiu and T. Natschläger, *Chemom. Intell. Lab. Syst.*, 2017, **161**, 49–60.

4 O. E. de Noord, *Chemom. Intell. Lab. Syst.*, 1994, **25**, 85–97.

5 T. Fearn, *J. Near Infrared Spectrosc.*, 2001, **9**, 229–244.

6 B. M. Wise and R. T. Roginski, *IFAC-PapersOnLine*, 2015, **48**, 260–265.

7 J. J. Workman, *Appl. Spectrosc.*, 2017, **72**, 340–365.

8 A. Andrew and T. Fearn, *Chemom. Intell. Lab. Syst.*, 2004, **72**, 51–56.

9 K. E. Kramer, R. E. Morris and S. L. Rose-Pehrsson, *Chemom. Intell. Lab. Syst.*, 2008, **92**, 33–43.

10 H. Tan, S. T. Sum and S. D. Brown, *Appl. Spectrosc.*, 2002, **56**, 1098–1106.

11 B. Igne, J.-M. Roger, S. Roussel, V. Bellon-Maurel and C. R. Hurburgh, *Chemom. Intell. Lab. Syst.*, 2009, **99**, 57–65.

12 Z. Lin, B. Xu, Y. Li, X. Shi and Y. Qiao, *J. Chemom.*, 2013, **27**, 406–413.

13 H. Martens, M. Høy, B. M. Wise, R. Bro and P. B. Brockhoff, *J. Chemom.*, 2003, **17**, 153–165.

14 B. Igne and C. R. Hurburgh, *J. Chemom.*, 2010, **24**, 75–86.

15 E. Andries and J. H. Kalivas, *J. Chemom.*, 2013, **27**, 126–140.

16 X. Capron, B. Walczak, O. de Noord and D. L. Massart, *Chemom. Intell. Lab. Syst.*, 2005, **76**, 205–214.

17 J. H. Kalivas, G. G. Siano, E. Andries and H. C. Goicoechea, *Appl. Spectrosc.*, 2009, **63**, 800–809.

18 P. Shahbazikhah and J. H. Kalivas, *Chemom. Intell. Lab. Syst.*, 2013, **120**, 142–153.

19 Y. Wang, D. J. Veltkamp and B. R. Kowalski, *Anal. Chem.*, 1991, **63**, 2750–2756.

20 E. Bouveresse, C. Hartmann, D. L. Massart, I. R. Last and K. A. Prebble, *Anal. Chem.*, 1996, **68**, 982–990.

21 Y. Wang, M. J. Lysaght and B. R. Kowalski, *Anal. Chem.*, 1992, **64**, 562–564.

22 E. Bouveresse and D. Massart, *Chemom. Intell. Lab. Syst.*, 1996, **32**, 201–213.

23 P. J. Gemperline, J. Cho, P. K. Aldridge and S. S. Sekulic, *Anal. Chem.*, 1996, **68**, 2913–2915.

24 X. Fan, H. Lu and Z. Zhang, *Chemom. Intell. Lab. Syst.*, 2018, **181**, 21–28.

25 J. Zhang, C. Guo, X. Cui, W. Cai and X. Shao, *Anal. Chim. Acta*, 2019, **1050**, 25–31.

26 Y. Liu, W. Cai and X. Shao, *Anal. Chim. Acta*, 2014, **836**, 18–23.

27 W. Fan, Y. Liang, D. Yuan and J. Wang, *Anal. Chim. Acta*, 2008, **623**, 22–29.

28 Y. Liu, H. Xu, Z. Xia and Z. Gong, *Analyst*, 2018, **143**, 1274–1280.

29 T. Skotare, D. Nilsson, S. Xiong, P. Geladi and J. Trygg, *Anal. Chem.*, 2019, **91**, 3516–3524.

30 T. Boucher, M. D. Dyar and S. Mahadevan, *J. Chemom.*, 2017, **31**, e2877.

31 P. Shan, Y. Zhao, Q. Wang, Y. Ying and S. Peng, *Spectrochim. Acta, Part A*, 2020, **227**, 117653.

32 Y.-Y. Shi, J.-Y. Li and X.-L. Chu, *Chin. J. Anal. Chem.*, 2019, **47**, 479–487.

33 B. M. Wise, N. B. Gallagher, R. Bro, J. M. Shaver, W. Windig and R. S. Koch, *Chemometrics Tutorial for PLS_Toolbox and Solo*, Eigenvector Research, Inc., 2006.

34 H.-W. Tan and S. D. Brown, *J. Chemom.*, 2001, **15**, 647–663.

35 M. Kompany-Zareh and F. van den Berg, *Analyst*, 2010, **135**, 1382.

36 J. Yang, X. Lou, H. Yang, H. Yang, C. Liu, J. Wu and J. Bin, *Anal. Lett.*, 2019, **52**, 2188–2202.

37 Z.-P. Chen, L.-M. Li, R.-Q. Yu, D. Littlejohn, A. Nordon, J. Morris, A. S. Dann, P. A. Jeffkins, M. D. Richardson and S. L. Stimpson, *Analyst*, 2011, **136**, 98–106.

38 X. Li, W. Cai and X. Shao, *J. Near Infrared Spectrosc.*, 2015, **23**, 285–291.

39 Y. Liu, W. Cai and X. Shao, *Spectrochim. Acta, Part A*, 2016, **169**, 197–201.

40 F. Zhang, R. Zhang, J. Ge, W. Chen, W. Yang and Y. Du, *Anal. Methods*, 2018, **10**, 2169–2179.

41 C. Zou, H. Zhu, J. Shen, Y. He, J. Su, X. Fan, H. Lu, Z. Zhang and Y. Chen, *Anal. Methods*, 2019, **11**, 4481–4493.

42 Y. Zhao, Z. Zhao, P. Shan, S. Peng, J. Yu and S. Gao, *Molecules*, 2019, **24**, 1802.

43 Q. Li, X. Sun, X. Ma, B. Li, H. Wang, H. Lv, Q. Wang, K. Xu and D. Chen, *Chemom. Intell. Lab. Syst.*, 2019, **191**, 143–147.

44 P. Mishra, R. Nikzad-Langerodi, F. Marini, J. M. Roger, A. Biancolillo, D. N. Rutledge and S. Lohumi, *TrAC, Trends Anal. Chem.*, 2021, **143**, 116331.

45 D. Brouckaert, J.-S. Uyttersprot, W. Broeckx and T. D. Beer, *Anal. Chim. Acta*, 2017, **971**, 14–25.

46 J. Eliaerts, N. Meert, P. Dardenne, F. V. Durme, V. Baeten, N. Samyn and K. D. Wael, *Talanta*, 2020, **209**, 120481.

47 S. R. S. Dangal and J. Sanderman, *Sensors*, 2020, **20**, 6729.

48 R. R. Rodrigues, J. T. Rocha, L. M. S. Oliveira, J. C. M. Dias, E. I. Müller, E. V. Castro and P. R. Filgueiras, *Chemom. Intell. Lab. Syst.*, 2017, **166**, 7–13.

49 C. Grelet, J. A. F. Pierna, P. Dardenne, H. Soyeurt, A. Vanlierde, F. Colinet, C. Bastin, N. Gengler, V. Baeten and F. Dehareng, *J. Dairy Sci.*, 2017, **100**, 7910–7921.

50 V. Artyushenko, A. Bocharnikov, G. Colquhoun, C. Leach, V. Lobachev, T. Sakharova and D. Savitsky, *Vib. Spectrosc.*, 2008, **48**, 168–171.

51 A. W. Owen, E. A. McAulay, A. Nordon, D. Littlejohn, T. P. Lynch, J. S. Lancaster and R. G. Wright, *Anal. Chim. Acta*, 2014, **849**, 12–18.

52 C. A. Damin and A. J. Sommer, *Appl. Spectrosc.*, 2013, **67**, 1252–1263.

53 W. Du, Z.-P. Chen, L.-J. Zhong, S.-X. Wang, R.-Q. Yu, A. Nordon, D. Littlejohn and M. Holden, *Anal. Chim. Acta*, 2011, **690**, 64–70.

54 T. U. Rehman, L. Zhang, D. Ma, L. Wang and J. Jin, *Comput. Electron. Agric.*, 2020, **176**, 105685.

55 S. Zhou, S. Zhu and X. Wei, *Spectrosc. Lett.*, 2020, **53**, 448–457.

56 R. L. Green and J. H. Kalivas, *Chemom. Intell. Lab. Syst.*, 2002, **60**, 173–188.

57 J. H. Kalivas, *Anal. Lett.*, 2005, **38**, 2259–2279.

58 A. C. McIntyre, PhD thesis, University of Strathclyde, 2011.

59 T. S. Shores, *Applied Linear Algebra and Matrix Analysis*, Springer International Publishing, Cham, 2nd edn, 2018.