Contents lists available at ScienceDirect

Analytica Chimica Acta

A two-level strategy for standardization of near infrared spectra by multi-level simultaneous component analysis

Jin Zhang [a], Cheng Guo [a], Xiaoyu Cui [a], Wensheng Cai [a], Xueguang Shao [a, b, c, d, e, *]

[a] Research Center for Analytical Sciences, College of Chemistry, Nankai University, Tianjin, 300071, China
[b] Tianjin Key Laboratory of Biosensing and Molecular Recognition, Tianjin, 300071, China
[c] State Key Laboratory of Medicinal Chemical Biology, Tianjin, 300071, China
[d] Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), Tianjin, 300071, China
[e] Xinjiang Laboratory of Native Medicinal and Edible Plant Resources Chemistry, College of Chemistry and Environmental Science, Kashgar University, Kashgar, 844006, China

HIGHLIGHTS

- A two-level model was used to fit the spectra measured on different instruments.
- Differences between instruments and measurements were described, respectively.
- Correction was obtained by adjusting the coefficients in the models.
- A method for standardizing NIR spectra on different instruments was developed.

GRAPHICAL ABSTRACT
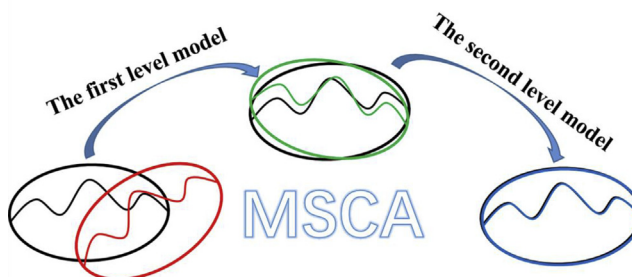
ARTICLE INFO

ABSTRACT

Standardization of near infrared (NIR) spectra is indispensable in practical applications because the spectra measured on different instruments are commonly used and the difference between the instruments must be corrected. A two-level standardization method is proposed in this study based on multi-level simultaneous component analysis (MSCA) algorithm for correcting the spectral difference between instruments. A two-level MSCA model is used to model the difference between instruments (the first level) and samples (the second level). With the two models, the spectral difference due to instruments and measurement operation can be corrected, respectively. Three NIR spectral datasets of pharmaceutical tablet, corn and plant leaf are used to evaluate the efficiency of the proposed method. The results show that the score of the first level model describes the overall spectral difference between instruments, and the score of the second level model depictures the spectral difference of the same sample between the measurements. The latter difference may include the spectral variations caused by instrument, operation and the measurement conditions. Therefore, both the spectral difference due to the instrument and measurement can be corrected by adjusting the coefficients in the scores of the two level models, respectively. The proposed method provides a good way for standardizing the spectra measured on different instruments when the measurement is not reproducible.

© 2018 Elsevier B.V. All rights reserved.

---

* Corresponding author. Research Center for Analytical Sciences, College of Chemistry, Nankai University, Tianjin, 300071, China.
   E-mail address: xshao@nankai.edu.cn (X. Shao).

1. Introduction

Standardization is essential for instrumentation and analytical measurement in, e.g., chemical sensor arrays [1] and Raman spectroscopy [2]. It is more important for near infrared (NIR) spectroscopy when multivariate calibration is needed [3–5]. Standardization or calibration transfer must be used to correct the spectral variation caused by the instrument or the measurement when the prediction samples are measured on a different instrument or under different condition from the calibration samples.

A large number of methods have been proposed for standardizing the NIR spectra measured on different instruments [6]. Standardization (or calibration transfer) can be achieved by three strategies, i.e., the correction of the prediction values, standardization of the spectra or modification of the model coefficients. The univariate slope and bias correction (SBC) [7] is a representative method for the first kind of strategy, i.e., prediction value correction. The method works well only when the prediction biases change linearly with the spectral difference. Standardization of the spectra is the most commonly used strategy. Piecewise direct standardization (PDS) [8] very efficient. In the calculation, the linear relationship is built between the spectra measured on different instruments in a small window region. Instead of correcting the spectra, works were also done by finding the relationship between the principal components extracted from the spectra, such as canonical correlation analysis (CCA) [9] and spectral space transformation (SST) [10]. Alternatively, the methods by correction of model coefficients includes Tikhonov regularization [5,11,12], domain-invariant partial least squares regression (di-PLS) [13], etc. These methods can be viewed as a re-calibration, which handles new spectra measured under different conditions based on the prior knowledge of the original model coefficients. In our recent works, alternating trilinear decomposition (ATLD) [14], PLS-corrected [15], dual model strategy [16] and linear model correction (LMC) [17] have been proposed based on the assumption that there is a linear relationship between the spectra under different conditions. All these methods are straightforward to be understood and remarkably reliable when the constraint is reasonable for all instruments. However, most of the methods correct the spectral difference following a simple relationship between the spectra measured on different instruments. When the spectral difference is affected by more factors, new strategies are still needed.

Multi-level simultaneous component analysis (MSCA) [18] has been proposed as a method for exploring the underlying relationship of the multivariate data involving different factors. The method was designed based on the simultaneous component analysis (SCA) [19], which can simultaneously analyze the relationship of the samples based on the analytical signal of different instruments. MSCA is, however, developed for summarizing different variances with a multi-level model, e.g., between instrument model (the first level) or within instrument model (the second level). The method has been successfully used in metabolomics [20,21], psychology [22] and temperature dependent NIR spectra [23], etc. for the data containing different variances.

In practical application, the NIR spectra measured on different instruments contain not only the absorption following the Lambert-Beer law, but also the signal relating instrumental state, measurement condition, particle size, even temperature and humidity [5]. The spectral profiles of pure components may even be varied over instruments [8,10]. Therefore, it is very hard to describe the variance of the NIR spectra measured on different instruments using only a simple model [18,19,24,25]. In this study, a multi-level calibration transfer method based on MSCA was proposed, and the performance is validated by three NIR spectral datasets.

2. Theory and algorithm

For calibration transfer, the spectra of standard samples measured on different instruments are generally used to build the transfer model. In this study, the spectra of standard samples measured on two instruments were used. If $\mathbf{X}_{raw,1}$ and $\mathbf{X}_{raw,2}$ were used to represent the spectra measured on the two instruments, the two-level MSCA model can be described as,

$$\mathbf{X}_{raw,i} = \mathbf{1}_K\mathbf{m}^T + \mathbf{1}_K\mathbf{t}_{1,i}\mathbf{P}_1^T + \mathbf{T}_{2,i}\mathbf{P}_2^T + \mathbf{E}_{MSCA,i} \tag{1}$$

subjecting to the following constraints:

$$\sum_{i=1}^{2} K\mathbf{t}_{1,i} = 0 \text{ and } \mathbf{1}_K^T\mathbf{T}_{2,i} = 0 \tag{2}$$

where i equals 1 or 2 to represent the two instruments, K is the number of the standard samples, the superscript 'T' represents a transpose operator, $\mathbf{1}_K$ is the column vector of ones with K elements, $\mathbf{m}^T$ is the overall mean of $\mathbf{X}_{raw}$, $\mathbf{t}_{1,i}$ and $\mathbf{T}_{2,i}$ represent the score of the first and second level model, $\mathbf{P}_1$ and $\mathbf{P}_2$ denote the loading of the first and second level model, respectively, and $\mathbf{E}_{MSCA,i}$ is the residual matrix that is not fitted by the model. By imposing the constraints on the scores, the three parts of the MSCA model are forced to be orthogonal to each other [18], i.e., the model at different level can separately describe the different types of variation in the data.

In the calculation of MSCA, X_{raw} is mean centered at first, and then the first level model can be obtained by performing principal component analysis (PCA) on the mean spectra of each instrument. Therefore, the score of the model explains only the variance between the instruments. The second level model can be obtained by performing PCA on the mean-centered spectra after subtracting the variances explained by the first level model. Therefore, the score of the model includes only the variance induced by the measurement of the sample. Clearly, the transfer of the spectra measured on one instrument into another one can be achieved by adjusting the coefficients in the scores of both models.

The details of the calculation can be summarized in two steps as follows:

(1) Build MSCA models: With the standard spectra measured on two instruments, $\mathbf{X}_{raw}$, the two level models, including $\mathbf{m}^T$, $\mathbf{t}_1$, $\mathbf{T}_2$, $\mathbf{P}_1$ and $\mathbf{P}_2$, are obtained by using the algorithm of MSCA. In this step, the number of principal components for the two level models should be determined.

(2) Transfer the spectrum: Taking the transfer of the spectrum of instrument 2 to instrument 1 as an example, the transfer can be achieved by the following calculations:

$$\mathbf{t}_{1,2s} = \left(\mathbf{x}_{2s} - \mathbf{m}^T\right)\left(\mathbf{P}_1^T\right)^+ \tag{3}$$

$$\mathbf{T}_{2,2s} = \left(\mathbf{x}_{2s} - \mathbf{m}^T - \mathbf{t}_{1,2s}\mathbf{P}_1^T\right)\left(\mathbf{P}_2^T\right)^+ \tag{4}$$

$$\mathbf{t}_{1,T2s} = \mathbf{t}_{1,2s}\left(\mathbf{t}_{1,2}\right)^+\mathbf{t}_{1,1} \tag{5}$$

$$\mathbf{T}_{2,T2s} = \mathbf{T}_{2,2s}\left(\mathbf{T}_{2,2}\right)^+\mathbf{T}_{2,1} \tag{6}$$

$$\mathbf{x}_{T2s} = \mathbf{m}^T + \mathbf{t}_{1,T2s}\mathbf{P}_1^T + \mathbf{T}_{2,T2s}\mathbf{P}_2^T \tag{7}$$

where the superscript '+' represents a pseudo inverse operator. By using Eqs. (3) and (4), the scores of the spectrum to be transferred,

x$_{2s}$, for the two level models are calculated, then the scores are transferred from instrument 1 to instrument 2 using the Eqs. (5) and (6), and finally, the transferred spectrum can be obtained by Eq. (7) using the transferred scores and the loadings **P**$_1$ and **P**$_2$.

3. Data description

Three NIR spectra datasets are used in this study for evaluating the proposed method. The pharmaceutical tablet dataset, denoted as dataset 1, was downloaded from the website of the international diffuse reflectance conference, http://www.idrc-chambersburg.org/shootout2002.html. The dataset contains the NIR spectra, weight, hardness, and assay values of the active pharmaceutical ingredient (API) of 655 pharmaceutical tablet samples. The spectra are obtained with two spectrometers (Foss, USA) in the range of 600-1898 nm with the digitization interval of 2 nm. The suggested calibration, validation, and prediction set contain 155, 40 and 460 samples, respectively. In this study, the first instrument is assigned as the master, and the other is taken as the slave. The API is adopted for validating the transfer effect. In the calculations, the suggested 597 variables in the spectral region from 600 to 1792 nm were used because the spectra in the region from 1794 to 1898 nm are too noisy to obtain reasonable calibration results [26]. The typical spectrum measured on the master and slave for the dataset is plotted in Fig. 1 (a).

The corn dataset, denoted as dataset 2, contains the NIR spectra measured on three instruments (m5, mp5, and mp6), and the moisture, oil, protein and starch contents of 80 corn samples. The dataset was downloaded from th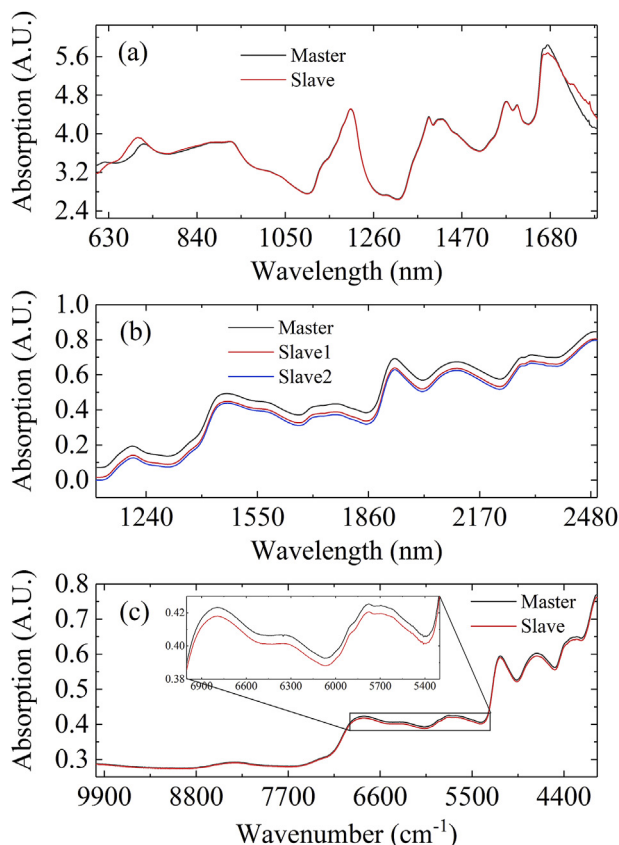e website of eigenvector research incorporated, http://software.eigenvector.com/Data/Corn/index.html. Each spectrum was measured in the range of 1100 - 2498 nm with the digitization interval of 2 nm. In this study, m5 is assigned as the master, mp5 and mp6 are taken as the slaves, respectively. The oil content is used for evaluating the performance of the method. The typical spectra measured on the three instruments are plotted in Fig. 1 (b).

The plant leaf dataset, denoted as dataset 3, contains the NIR spectra and the contents of reducing sugar of 117 plant leaf samples. The samples were ground into 40 mesh and 60 mesh, respectively, and measured on a NIR spectrometer (Thermo Electron Corporation, USA) in the range of 4000−10,000 cm^{-1} with the digitization interval of ca. 3.9 cm^{-1}. In this study, when the 40 and 60 mesh samples were measured, the instrument is assigned as the master and the slave, respectively. The typical spectra are plotted in Fig. 1 (c).

Before calculation, each dataset was divided into a calibration sets, a transfer dataset and a prediction set using Kennard-Stone (KS) algorithm [27]. The calibration set (the spectra of the master) is used for building the multivariate calibration model of the master, the transfer set, including the spectra of all the instruments, is adopted for building the transfer model, and the prediction set (the spectra of the salve) is employed for validating the effect of the transfer model. For dataset 1, the possible outliers (Nos. 19, 122, 126, and 127 in the calibration set and Nos. 11, 145, 267, 295, 294, 342, 313, 341, and 343 in the prediction set) were removed, the rest 642 samples were divided into a calibration set of 400 samples, a transfer set of 30 samples and a prediction set of 212 samples. For dataset 2 and 3, 30, 30, 20 and 50, 30, 37 samples were used for the calibration, transfer and prediction set, respectively.
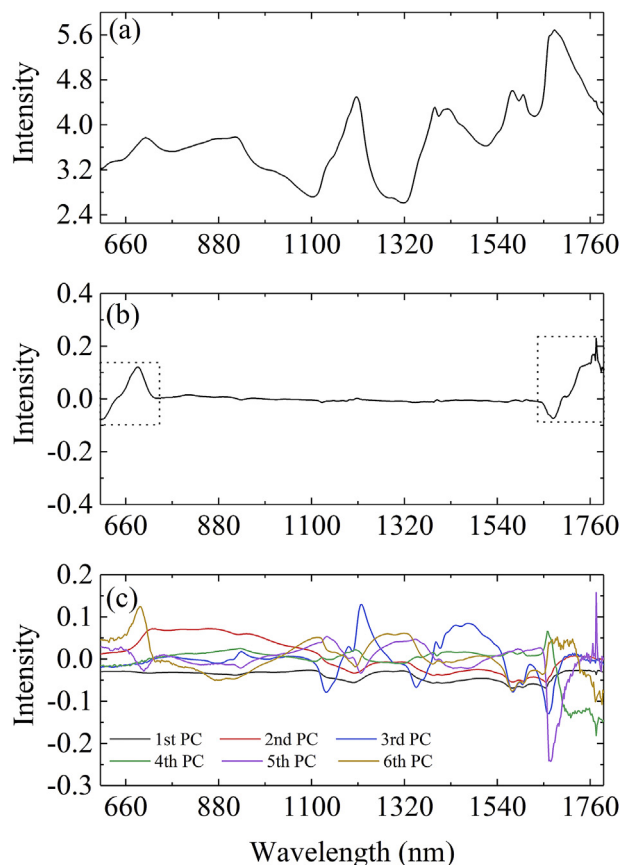

Fig. 1. Typical spectra of dataset 1 (a), dataset 2 (b) and dataset 3 (c).

Fig. 2. The mean spectrum (a) and the loadings for the first (b) and second (c) level model.

4. Results and discussion

4.1. Building the two-level MSCA model

For building the MSCA model, the number of principal components (nPC) for the two level models is needed. In this study, the number was determined by the commonly-used criterion of the 'explained variance'. The nPC that explains 99.9% of the variance was used. For dataset 1, one and six were used as the parameter for the first and second level models, respectively. When the spectra of more instruments are involved in the calculation, more components may be needed to describe the variance between the spectra.

Fig. 2 shows the mean spectrum, **m**, and the loadings of the first and second level model, **P**$_1$ and **P**$_2$, respectively, for dataset 1. Fig. 2 (a) shows the mean spectrum, which is the common variance of the spectra measured on the two instruments. Fig. 2 (b) displays the loading of the first level model, which shows the difference between the spectra measured on the different instruments. Clearly, the difference can be seen in the spectral region of 600−700 nm and 1700 - 1792 nm. The loadings of the second level model were plotted in Fig. 2 (c), from which the spectral differences can be found in the full spectral region. A large difference can be seen in the region of 600−700 nm, 1100−1400 nm, and 1700−1792 nm,
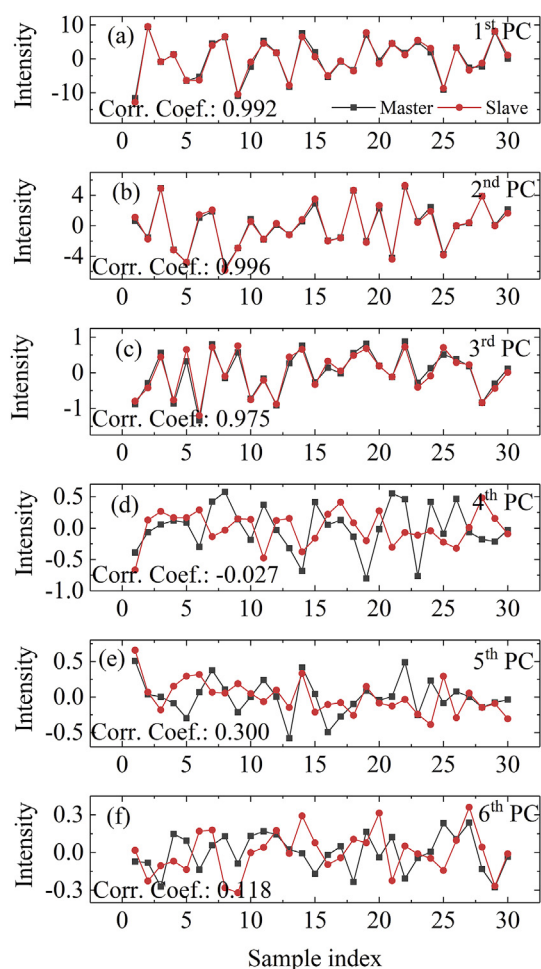
Fig. 3. Comparison of the scores in the 1st - 6th (a−f) principal component in the second level model for the master and the slave spectra. The black and red points represent the spectra of the master and the slave for the second level model, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
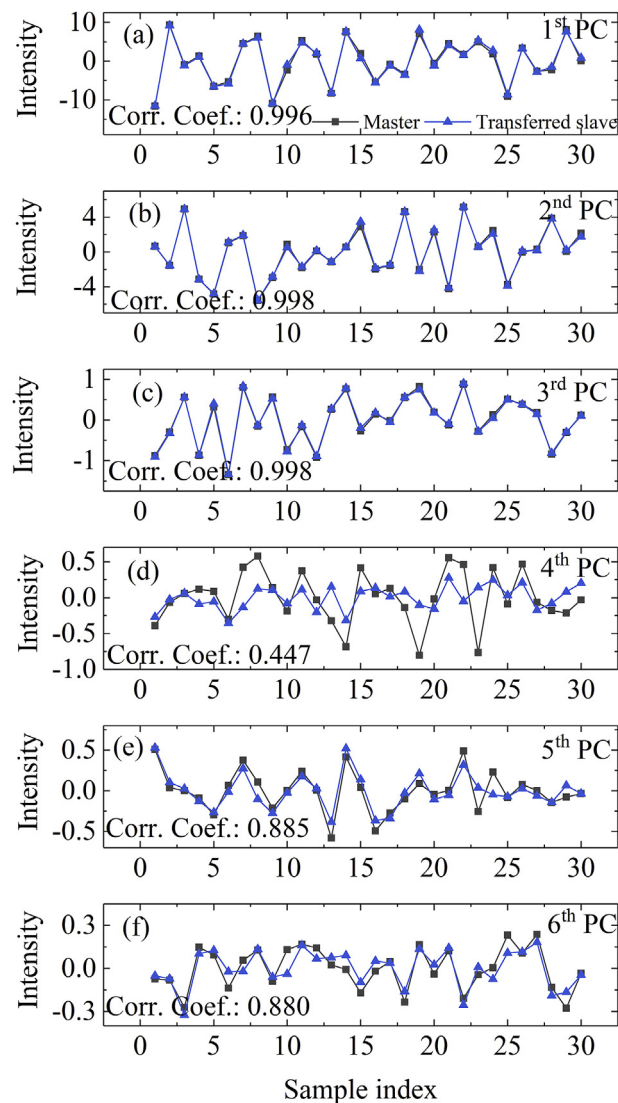
Fig. 4. Comparison of the scores in the 1st - 6th (a−f) principal component in the second level model for the master and the transferred slave spectra. The black and blue points represent the spectra of the master and the transferred slave for the second level model, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

which correspond to the third, second and first overtone of C-H, respectively. These difference may be resulted in by the difference between the samples as well as the measurement of the spectra.

The scores in the models are more important in this study because the transfer is achieved by adjusting the coefficient values in the scores. In the first level model, only one principal component is used, there is only one score for each instrument. The values for the master and slave are 1.26 and −1.26, respectively. This is reasonable, each score representing the half of the difference. Fig. 3 shows the 1st - 6th PCs of the scores for the second level model. The black and red points represent the score of master and slave, respectively. Clearly, there is only a small difference in the 1st − 3rd PCs, but obvious difference can be found in the 4th -6th scores. The correlation coefficients (Corr. Coef.) are labeled in the figure, which can be used for evaluation of the correlation of these values. The similarity of the first three scores indicates that the main profiles of the spectra measured on different instruments are similar to each other, and the difference in the later three scores suggests that the variation occurs when the spectra were measured on the different
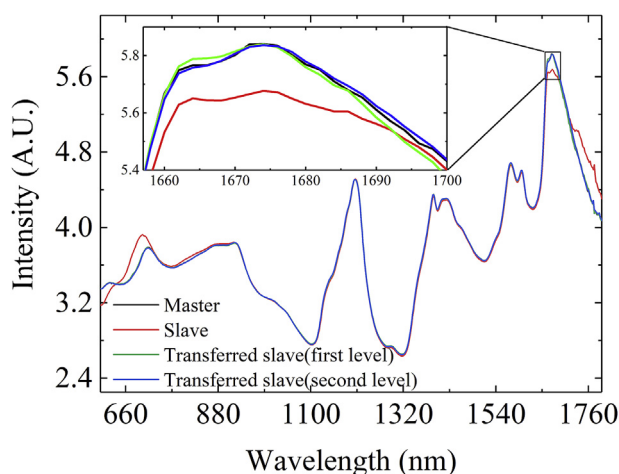
instruments.

4.2. Standardization of the spectra

According to the principle of the proposed method, spectral standardization, or the transfer of the spectrum measured on the slave to the master, can be achieved by adjusting the coefficient values in the scores of the models. Clearly, in the case of dataset 1, the transfer can be completed by simply replacing the score of the slave to the master, because there is only one value in each score. Generally, the transfer can be achieved according to Eq. (5). Similarly, the transfer of the second level model can be obtained with Eq. (6).

Fig. 4 shows the scores of the six PCs before and after the transfer. The black points are the scores of the master, and the red points are the 'transferred scores' of the slave. Similarly, as in Fig. 3, correlation coefficients are labeled in the figure. Compared with that in Fig. 3, the similarity of the two scores is greatly improved, although the effect for the first three scores is not so obvious. It can be seen, however, that there is still a difference between the two scores, particularly in the later three ones. This indicates that not all the difference can be fitted by the models, implying that more factors in the measurement may be included to affect the measured spectra. Multi-level strategy may be a good choice for correcting the complicated effects.

Fig. 5 shows the spectra measured on the master and the slave for a sample arbitrarily selected from the prediction set of dataset 1. To show the effect of the standardization by the proposed method, the corrected spectra by the first level and the second level models were also displayed in the figure. Clearly, there is a clear difference between the spectra of the two instruments, particularly in the spectral region 600−700 nm and 1700 - 1792 nm. After the correction by the first level model, the slave spectrum moves closer to the master spectrum, but there is still a slight difference in the spectral region 1660−1700 nm. After the correction of the second level model, the spectrum of the master and the slave is almost superposed. The results clearly show that each model takes its effect in the transfer of the spectrum.

To further investigate the transfer effect of the proposed method, PCA was performed on the spectra measured on the two instrument and the transferred spectra. The distribution of the

spectra in the 1st - 2nd PC space and the 3rd - 4th PC space is shown in Fig. 6 (a) and (b), respectively. The black, red and blue points represent the spectra of the master, slave and the transferred spectra. The ellipses are determined with the confidence 95%. Similar to that observed in Fig. 5, only a small difference can be seen in the 1st - 2nd PC space, but there is an obvious difference in the 3rd - 4th PC space. After the transfer, however, the distribution of the transferred spectra is well overlapped with that of the master,
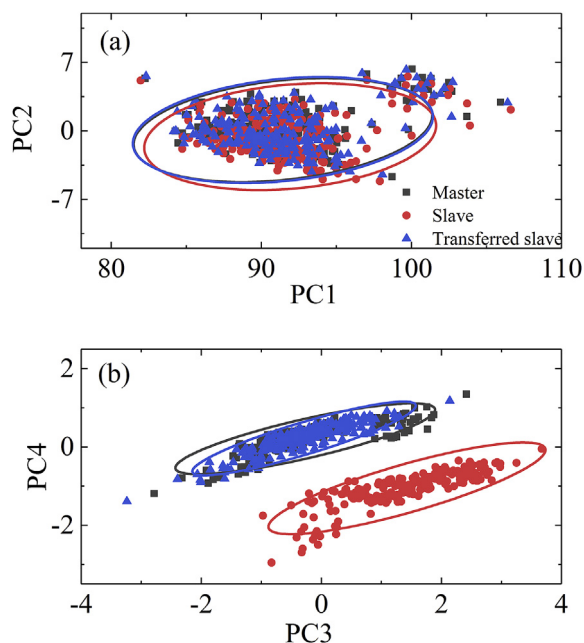
Fig. 6. The transfer of the spectra from the master to the slave in PC1-PC2 (a) and PC3-PC4 (b) space.
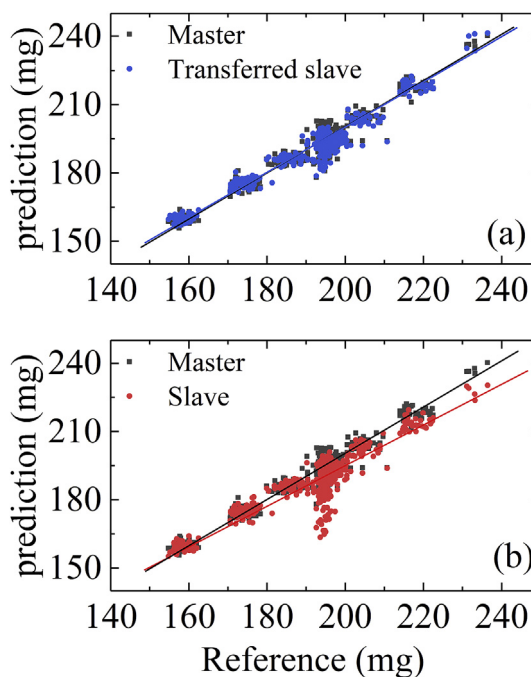

Fig. 5. The transfer of an arbitrarily selected spectrum from the master to the slave. The black, red, green and blue curve represents the spectrum of the master, the slave, and the spectrum transferred by the first level and second level model, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Fig. 7. Relationship of the reference API contents and the prediction values obtained from the master and transferred spectra (a), and that for the master and the slave spectra (b), respectively.

Table 1
Comparison of the results obtained by PDS, SST and MSCA for dataset 1.

Calibration spectra	Validation spectra	Parameters	RMSEP
Master	Master	nLV = 3	3.4
	Slave	nLV = 3	5.7
	Transferred Slave (PDS)	w^a = 17, nLV = 4	3.7
	Transferred Slave (SST)	nPC^b = 4	3.4
	Transferred Slave (MSCA)	nPC$_1^c$ = 1, nPC$_2$ = 6	**3.4**
Slave	Slave	nLV = 3	3.2
	Master	nLV = 3	12.6
	Transferred Master (PDS)	w^a = 17, nLV = 4	3.6
	Transferred Master (SST)	nPC^b = 4	3.4
	Transferred Master (MSCA)	nPC$_1$ ^{c} = 1, nPC$_2$ = 6	**3.3**

The bold numbers represent the results obtained by the proposed method.
  [a] The window size used in PDS.
  [b] The number of principal components used in SST.
  [c] nPC$_1$ and nPC$_2$ represent the number of principal components for the first and second level model in MSCA, respectively.

indicating clearly that the difference is corrected.

4.3. Validation of the standardized spectra

For final evaluation of the proposed method, prediction of the transferred spectra in the prediction set using the PLS model built with the calibration set of the master was investigated. The optimal number of latent variables (nLV) in building the PLS model is determined with a 7-fold cross-validation. Three LVs were used for the calibration set of dataset 1. The blue points in Fig. 7 (a) shows the relationship between the predicted and reference values. As a comparison, the predicted values of the spectra measured on the master were also plotted in the figure as the black points. Clearly, there is no obvious difference between the blue and black points, indicating that the transferred spectra can be well predicted by the PLS model of the master. For further comparison, the relationship

between the predicted results of the spectra measured on the slave (before the standardization) and the reference values are plotted in Fig. 7 (b) by red points. There is a clear deviation between the red and black points, particularly for the points with larger reference values.

To further investigate the efficiency of the proposed method, the results obtained by PDS and SST for dataset 1 was compared in Table 1. In the table, the results by interchanging the master and the slave were also listed. Clearly, the prediction of the slave spectra directly by the model of the master is much worse than that of the master spectra. When the transferred spectra are predicted, the prediction error, root mean squared error of the prediction set (RMSEP), can be significantly reduced. Comparing the three methods, the proposed method produces a comparable result with SST, and slightly better than the commonly-used PDS method. Furthermore, similar results are obtained when the master and the slave are interchanged. The results may give a further validation of the practicability of the method.

Tables 2 and 3 shows the results for dataset 2 and 3, respectively. The nLV used for the PLS model of the master is 4 and 7. From both the tables, it can be seen that calibration transfer is needed when the spectra of the slave are predicted with the model of the master. The proposed method can obtain the similar or slightly better results compared with PDS and SST method.

5. Conclusions

A multi-level standardization method is proposed for the transfer of the NIR spectra measured on different instruments. A two-level MSCA was used for modeling the NIR spectra. The first level model describes the overall difference between the instruments and the second level model depictures the difference caused in the measurement of the spectra. By adjusting the coefficients in the scores of the models, the spectra measured on one

Table 2
Comparison of the results obtained by PDS, SST and MSCA for dataset 2.

Calibration spectra	Validation spectra	Parameters	RMSEP
Master	Master	nLV = 4	0.09
	Slave1	nLV = 4	0.15
	Slave2	nLV = 4	0.22
	Transferred Slave1 (PDS)	w^a = 17, nLV = 2	0.14
	Transferred Slave1 (SST)	nPC^b = 2	0.13
	Transferred Slave1 (MSCA)	nPC$_1$ ^{c} = 2, nPC$_2$ = 10	**0.10**
	Transferred Slave2 (PDS)	w^a = 17, nLV = 2	0.12
	Transferred Slave2 (SST)	nPC^b = 2	0.12
	Transferred Slave2 (MSCA)	nPC$_1$ ^{c} = 2, nPC$_2$ = 10	**0.10**
Slave1	Slave1	nLV = 4	0.12
	Master	nLV = 4	1.57
	Slave2	nLV = 4	0.18
	Transferred Master (PDS)	w^a = 17, nLV = 2	0.14
	Transferred Master (SST)	nPC^b = 2	0.20
	Transferred Master (MSCA)	nPC$_1$ ^{c} = 2, nPC$_2$ = 10	**0.13**
	Transferred Slave2 (PDS)	w^a = 17, nLV = 2	0.10
	Transferred Slave2 (SST)	nPC^b = 2	0.11
	Transferred Slave2 (MSCA)	nPC$_1$ ^{c} = 2, nPC$_2$ ^{d} = 10	**0.12**
Slave2	Slave2	nLV = 4	0.13
	Master	nLV = 4	0.62
	Slave1	nLV = 4	0.17
	Transferred Master (PDS)	w^a = 17, nLV = 2	0.14
	Transferred Master (SST)	nPC^b = 2	0.15
	Transferred Master (MSCA)	nPC$_1$ ^{c} = 2, nPC$_2$ = 10	**0.14**
	Transferred Slave1 (PDS)	w^a = 17, nLV = 2	0.17
	Transferred Slave1 (SST)	nPC^b = 2	0.17
	Transferred Slave1 (MSCA)	nPC$_1$ ^{c} = 2, nPC$_2$ = 10	**0.13**

The bold numbers represent the results obtained by the proposed method.
  [a] The window size used in PDS.
  [b] The number of principal components used in SST.
  [c] nPC$_1$ and nPC$_2$ represent the number of principal components for the first and second level model in MSCA, respectively.

Table 3
Comparison of the results obtained by PDS, SST and MSCA for dataset 3.

Calibration spectra	Validation spectra	Parameters	RMSEP
Master	Master	$nLV = 7$	0.52
	Slave	$nLV = 7$	0.87
	Transferred Slave (PDS)	$w^a = 13$, $nLV = 2$	0.54
	Transferred Slave (SST)	$nPC^b = 2$	0.57
	Transferred Slave (MSCA)	$nPC_1{}^c = 1$, $nPC_2 = 13$	**0.55**
Slave	Slave	$nLV = 7$	0.57
	Master	$nLV = 7$	0.95
	Transferred Master(PDS)	$w^a = 13$, $nLV = 2$	0.70
	Transferred Master (SST)	$nPC^b = 2$	0.64
	Transferred Master (MSCA)	$nPC_1{}^c = 1$, $nPC_2 = 13$	**0.56**

The bold numbers represent the results obtained by the proposed method.
[a] The window size used in PDS.
[b] The number of principal components used in SST.
[c] nPC_1 and nPC_2 represent the number of principal components for the first and second level model in MSCA, respectively.

instrument can be transferred to that of the other. With a NIR dataset of pharmaceutical tablets, it was shown that the two level models both take effect for the transfer of the spectra, and multi-level strategy may be superior to the one-model-based methods when the factors affecting the measured spectra are complicated. Through the comparison of the proposed method with PDS and SST using the three datasets, the former produces slightly better results for most of the cases, although for some cases similar results are obtained. The multi-level methods may be a better choice for standardization of NIR spectra measured on different instruments when affecting factors are complicated.

Acknowledgements

References

[1] K. Yan, L. Kou, D. Zhang, Learning domain-invariant subspace using domain features and independence maximization, IEEE Trans. Cybern. 48 (2018) 288–299.

[2] S.X. Guo, A. Kohler, B. Zimmermann, R. Heinke, S. Stöckel, P. Rösch, J. Popp, T. Bocklitz, Extended multiplicative signal correction based model transfer for Raman spectroscopy in biological applications, Anal. Chem. 90 (2018) 9787–9795.

[3] T. Fearn, Standardisation and calibration transfer for near infrared instruments: a review, J. Near Infrared Spectrosc. 9 (2001) 229–244.

[4] R.N. Feudale, N.A. Woody, H.W. Tan, A.J. Myles, S.D. Brown, J. Ferré, Transfer of multivariate calibration models: a review, Chemometr. Intell. Lab. Syst. 64 (2002) 181–192.

[5] J.H. Kalivas, Overview of two-norm (L2) and one-norm (L1) Tikhonov regularization variants for full wavelength or sparse spectral multivariate calibration models or maintenance, J. Chemom. 26 (2012) 218–230.

[6] F.Y. Zhang, R.Q. Zhang, J. Ge, W.C. Chen, W.Y. Yang, Y.P. Du, Calibration transfer based on the weight matrix (CTWM) of PLS for near infrared (NIR) spectral analysis, Anal. Methods 10 (2018) 2169–2179.

[7] E. Bouveresse, C. Hartmann, D.L. Massart, I.R. Last, K.A. Prebble, Standardization of near-infrared spectrometric instruments, Anal. Chem. 68 (1996) 982–990.

[8] Y.D. Wang, D.J. Veltkamp, B.R. Kowalski, Multivariate instrument standardization, Anal. Chem. 63 (1991) 2750–2756.

[9] W. Fan, Y.Z. Liang, D.L. Yuan, J.J. Wang, Calibration model transfer for near-infrared spectra based on canonical correlation analysis, Anal. Chim. Acta 623 (2008) 22–29.

[10] W. Du, Z.P. Chen, L.J. Zhong, S.X. Wang, R.Q. Yu, A. Nordon, D. Littlejohn, M. Holden, Maintaining the predictive abilities of multivariate calibration models by spectral space transformation, Anal. Chim. Acta 690 (2011) 64–70.

[11] J.H. Kalivas, G.G. Siano, E. Andries, H.C. Goicoechea, Calibration maintenance and transfer using Tikhonov regularization approaches, Appl. Spectrosc. 63 (2009) 800–809.

[12] M.R. Kunz, J.H. Kalivas, E. Andries, Model updating for spectral calibration maintenance and transfer using 1-norm variants of Tikhonov regularization, Anal. Chem. 82 (2010) 3642–3649.

[13] R. Nikzad-Langerodi, W. Zellinger, E. Lughofer, S. Saminger-Platz, Domain-invariant partial least squares regression, Anal. Chem. 90 (2018) 6693–6701.

[14] Y. Liu, W.S. Cai, X.G. Shao, Standardization of near infrared spectra measured on multi-instrument, Anal. Chim. Acta 836 (2014) 18–23.

[15] X.Y. Li, W.S. Cai, X.G. Shao, Correcting multivariate calibration model for near infrared spectral analysis without using standard samples, J. Near Infrared Spectrosc. 23 (2015) 285–291.

[16] J.J. Wang, Z.F. Li, Y. Wang, Y. Liu, W.S. Cai, X.G. Shao, A dual model strategy to transfer multivariate calibration models for near-infrared spectral analysis, Spectrosc. Lett. 48 (2016) 348–354.

[17] Y. Liu, W.S. Cai, X.G. Shao, Linear model correction: a method for transferring a near-infrared multivariate calibration model without standard samples, Spectrochim. Acta 169 (2016) 197–201.

[18] M.E. Timmerman, Multilevel component analysis, Br. J. Math. Stat. Psychol. 59 (2006) 301–320.

[19] M.E. Timmerman, H.A.L. Kiers, Four simultaneous component models for the analysis of multivariate time series from more than one subject to model intraindividual and interindividual differences, Psychometrika 68 (2003) 105–121.

[20] I. Montoliu, O. Cominetti, C.L. Boulange, B. Berger, J. Siddharth, J. Nicholson, F.P.J. Martin, Modeling longitudinal metabonomics and microbiota interactions in C57BL/6 mice fed a high fat diet, Anal. Chem. 88 (2016) 7617–7626.

[21] A. Lemanska, M. Grootveld, C.J.L. Silwood, R.G. Brereton, Chemometric variance analysis of H-1 NMR metabolomics data on the effects of oral rinse on saliva, Metabolomics 8 (2012) S64–S80.

[22] S. Jeroen, C. Eva, T.M. E, V.H. Alain, Tolerance of justice violations: the effects of need on emotional reactions after violating equality in social dilemmas1, J. Appl. Soc. Psychol. 41 (2011) 357–380.

[23] R.F. Shan, Y. Zhao, M.L. Fan, X.W. Liu, W.S. Cai, X.G. Shao, Multilevel analysis of temperature dependent near-infrared spectra, Talanta 131 (2015) 170–174.

[24] J.J. Jansen, H.C.J. Hoefsloot, J. van der Greef, M.E. Timmerman, A.K. Smilde, Multilevel component analysis of time-resolved metabolic fingerprinting data, Anal. Chim. Acta 530 (2005) 173–183.

[25] A.K. Smilde, J.J. Jansen, H.C. Hoefsloot, R.J. Lamers, J. van der Greef, M.E. Timmerman, ANOVA-simultaneous component analysis (ASCA): a new tool for analyzing designed metabolomics data, Bioinformatics 21 (2005) 3043–3048.

[26] J. Zhang, X.Y. Cui, W.S. Cai, X.G. Shao, Combination of heuristic optimal partner bands for variable selection in near-infrared spectral analysis, J. Chemom. (2017), e2971. https://doi.org/10.1002/cem.2971.

[27] R.W. Kennard, L.A. Stone, Computer aided design of experiments, Technometrics 11 (1969) 137–148.