

Document downloaded from:

<http://hdl.handle.net/10251/108346>

This paper must be cited as:

Folch-Fortuny, A.; Vitale, R.; De Noord, OE.; Ferrer, A. (2017). Calibration transfer between NIR spectrometers: new proposals and a comparative study. *Journal of Chemometrics*. 31(3):1-11. doi:10.1002/cem.2874



The final publication is available at

<https://doi.org/10.1002/cem.2874>

Copyright John Wiley & Sons

Additional Information

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/314095847>

Calibration transfer between NIR spectrometers: New proposals and a comparative study

Article in *Journal of Chemometrics* · February 2017

DOI: 10.1002/cem.2874

CITATION

1

READS

162

4 authors:



Abel Folch-Fortuny

Royal DSM

14 PUBLICATIONS 72 CITATIONS

SEE PROFILE



Raffaele Vitale

Katholieke Universiteit Leuven - Université d...

23 PUBLICATIONS 54 CITATIONS

SEE PROFILE



Onno De Noord

Consultant

55 PUBLICATIONS 2,228 CITATIONS

SEE PROFILE



Alberto Ferrer

Universitat Politècnica de València

109 PUBLICATIONS 1,883 CITATIONS

SEE PROFILE

Calibration transfer between NIR spectrometers: new proposals and a comparative study

Abel Folch-Fortuny^{a,b,*}, Raffaele Vitale^{a,*}, Onno E. de Noord^c, Alberto Ferrer^a

^a*Grupo de Ingeniería Estadística Multivariante, Departamento de Estadística e Investigación Operativa Aplicadas y Calidad, Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain*

^b*Genetics BioIT DBC Department, DSM Food Specialties, Alexander Fleminglaan 1, 2613 AX Delft, The Netherlands*

^c*Shell Global Solutions International B.V., Shell Technology Centre Amsterdam, PO Box 38000, 1030 BN Amsterdam, The Netherlands*

Abstract

Calibration transfer between near-infrared (NIR) spectrometers is a subtle issue in chemometrics and process industry. In fact, as even very similar instruments may generate strongly different spectral responses, regression models developed on a first NIR system can rarely be used with spectra collected by a second apparatus. In this work, two novel methods to perform calibration transfer between NIR spectrometers are proposed. Both of them permit to exploit the specific relationships between instruments for imputing new unmeasured spectra, which will be then resorted to for building an improved predictive model, suitable for the analysis of future incoming data. Specifically, the two approaches are based on Trimmed Scores Regression (TSR) and Joint-Y Partial Least Squares regression (JYPLS), respectively. The performance of these novel strategies is assessed and compared to that of well-established techniques such as Maximum Likelihood Principal Component Analysis (MLPCA) and Piecewise Direct Standardisation (PDS) in two real case-studies.

Keywords: multivariate calibration, calibration transfer, missing data imputation, Trimmed Scores Regression (TSR), Joint-Y Partial Least Squares regression (JYPLS)

*Corresponding authors:

Telephone number: +34963877007 ext. 74935

Email address: abfolfor@upv.es (Abel Folch-Fortuny); rvitale86@gmail.com (Raffaele Vitale)

These authors have equal contributions

1. Introduction

Multivariate calibration is of crucial importance when useful quantitative information needs to be extracted from complex spectroscopic signals. Numerous applications of e.g. Principal Component Regression (PCR) or Partial Least Squares regression (PLS) aimed at this end have been reported over the past decades [1, 2]. However, practical limitations to the deployment of such techniques appear when existing calibration models are to be applied to measurements recorded by a new instrument and/or in different environmental conditions. In fact, even very similar spectrometers generally exhibit variations in their responses, which may seriously jeopardise this so-called *calibration transfer*.

Several methods have been proposed to overcome this issue and avoid at the same time an expensive and time-consuming full recalibration, using the newly acquired spectra. One of these approaches consists in updating the calibration model by merging measurements collected by both the first and the second spectrometer. However, that is commonly effective only when the two sets of spectral profiles are rather similar [3].

For the sake of simplicity, suppose now that a certain number of samples has been analysed by the primary instrument and a subgroup of these samples characterised also by the secondary oneⁱ. Among all the other strategies proposed in the scientific literature for dealing with such a scenario, Piecewise Direct Standardisation (PDS) [4] has been unanimously pointed out as *a reference for novel techniques* due to its local and multivariate nature [3, 5–7]. PDS basically transforms the spectra recorded by the secondary instrument so that its spectral response matches the one of the primary instrument. This allows any calibration model, built on the data resulting from the primary spectrometer, to be used for the analysis of those acquired by the secondary apparatus.

From a slightly different perspective, the transfer of a calibration model from a NIR spectrometer to another can be looked at as a missing data imputation problem. In this circumstance, all the information contained in the available primary and secondary spectra can be exploited to entirely reconstruct the profiles associated to those samples that were not analysed by the secondary instru-

ⁱAlternatively, two distinct or partially distinct sets of samples may be analysed by the two spectrometers. However, this contingency will not be contemplated here.

ment. These profiles can be then utilised for fitting an improved predictive model, suitable for the assessment of future incoming recordings. Maximum Likelihood Principal Component Analysis (MLPCA) [8] has been the first computational methodology to be applied for solving the calibration transfer issue in this peculiar fashion.

In this article, two innovative strategies to perform calibration transfer based on Trimmed Scores Regression (TSR) [9] and Joint-Y Partial Least Squares regression (JYPLS) [10] are proposed. Specifically, their performance will be assessed and compared to that of MLPCA and PDS in two real case-studies, in which the same set of samples was characterised by two different Near-Infrared (NIR) spectrometers.

2. Materials

The first dataset analysed here contains 60 spectra measured on 30 pseudo-gasoline samples within 800 and 1600 nm (401 scanned wavelengths, 30 spectra per instrument). Heptane, iso-octane, toluene, xylene and decane concentration are the properties of interest to be predicted.

The second dataset relates to 80 corn samples, whose spectral profiles were registered within 1100 and 2498 nm (700 scanned wavelengths for a total number of 160 spectra, 80 per each spectrometer). The response variables are moisture, oil, protein and starch content.

Both of them have been widely used to compare calibration transfer methods [11–13]. The gasoline dataset is included in the PLS_Toolbox for MATLAB [14], the corn dataset can be downloaded from <http://www.eigenvector.com/data>.

Prior to any analysis the gasoline and the corn spectra were simply mean-centred, as also proposed in [13], while the respective response variables were auto-scaled.

3. Methods

Let \mathbf{X}_a ($N_a \times J_a$) and \mathbf{X}_b ($N_b \times J_b$) be the matrices containing the spectral profiles collected by the primary and the secondary spectrometer, respectively. Mind that here the N_b samples characterised by the secondary instrument were also analysed by the primary one.

3.1. Piecewise Direct Standardisation (PDS)

PDS executes a series of local linear transformations of the spectra collected by the secondary instrument to subsequently allow the calibration model built for the primary spectrometer to be exploited for prediction purposes. Specifically, at each j -th wavelength, the absorbance values registered by the primary instrument ($\mathbf{x}_{a,j}$) are related by Principal Component Regression (PCR) to a specific spectral window of the profiles of the same samples collected by the secondary spectrometer ($\mathbf{X}_{b,j}$):

$$\mathbf{x}_{a,j} = \mathbf{1}b_j + \mathbf{X}_{b,j}\mathbf{f}_j \quad (1)$$

where $\mathbf{1}$ represents a vector of ones of appropriate dimensions. Incoming secondary instrument data are then adjusted through the estimated standardisation parameters, \mathbf{f}_j and b_j . Here, PDS was applied as coded in the PLS.Toolbox package for MATLAB [14]: all the principal components, whose eigenvalue (divided by the first one) was found to be larger than 0.0001, were included in each local regression model. On the other hand, the spectral window width was automatically optimised within the modelling procedure (see Section 4).

3.2. Maximum Likelihood Principal Component Analysis (MLPCA)

The adaptation of the MLPCA algorithm to model building with missing data is an iterative procedure based on an imputation alternatively performed by rows and columns. It has been proven [15–17] that the MLPCA object-wise imputation step is equivalent to performing a Projection to the Model Plane (PMP) for both PCA model building [9] (i.e. when a PCA model has to be fitted on incomplete data) and model exploitation [18] (i.e. when a PCA model is fitted on complete data and exploited to predict the scores of new incomplete observations).

Let \mathbf{X} be a matrix of dimensions $N \times J$. When data are missing in its n -th row, \mathbf{x}_n^T , \mathbf{X} can be rearranged so that such missing values are located in its last, say R , columns. Thus,

$$\mathbf{x}_n^T = [\mathbf{x}_n^{*T} \ \mathbf{x}_n^{\#T}] \quad (2)$$

and

$$\mathbf{X} = [\mathbf{X}^* \ \mathbf{X}^{\#}] \quad (3)$$

with $*/\#$ connoting the available/missing entries in \mathbf{x}_n^T , respectively.

Based on the following partition of the Singular Value Decomposition (SVD) of \mathbf{X} :

$$[\mathbf{X}^* \ \mathbf{X}^\#] = \mathbf{U}\mathbf{D}[\mathbf{P}^{*\top} \ \mathbf{P}^{\#\top}] + \mathbf{E} \quad (4)$$

the MLPCA algorithm imputes $\mathbf{x}_n^{\#\top}$ as:

$$\hat{\mathbf{x}}_n^\# = \mathbf{P}^\#(\mathbf{P}^{*\top}\mathbf{P}^*)^{-1}\mathbf{P}^{*\top}\mathbf{x}_n^* \quad (5)$$

Concerning the missing data in the j -th column of \mathbf{X} , \mathbf{x}_j , the data partition is performed according to the available and missing observations of the corresponding variable. This way:

$$\begin{bmatrix} \mathbf{X}^* \\ \mathbf{X}^\# \end{bmatrix} = \begin{bmatrix} \mathbf{U}^* \\ \mathbf{U}^\# \end{bmatrix} \mathbf{D}\mathbf{P}^\top + \mathbf{E} \quad (6)$$

and:

$$\hat{\mathbf{x}}_j^\# = \mathbf{U}^\#(\mathbf{U}^{*\top}\mathbf{U}^*)^{-1}\mathbf{U}^{*\top}\mathbf{x}_j^* \quad (7)$$

The imputation is iteratively executed until the reconstruction of the available values stabilises.

The MLPCA algorithm is comprehensively detailed in [8]. Besides, a thorough assessment of the use of MLPCA for missing data imputation is provided in [17].

To transfer a calibration model using this methodology, the complete set of N_a primary instrument spectra, \mathbf{X}_a , has to be concatenated with the N_b spectra collected by the secondary spectrometer, \mathbf{X}_b . An augmented data matrix \mathbf{X}_{ab} ($N_a \times (J_a + J_b)$) is then constructed, where the unrecorded secondary instrument profiles are missing (see Figure 1, *Imputation* box). In other words, if the sample associated to the n -th row of \mathbf{X}_{ab} has not been analysed by the secondary spectrometer, the partition in Eq. 2 applies: $\mathbf{x}_n^{*\top}$ and $\mathbf{x}_n^{\#\top}$ denote its available primary and its missing secondary instrument spectrum, respectively.

\mathbf{X}_{ab} is finally subjected to MLPCA.

3.3. Trimmed Scores Regression (TSR)

TSR is an iterative missing data imputation method, originally proposed for PCA model exploitation [19, 20]. Afterwards, it was adapted to the more general framework of PCA model

building in the presence of missing data [9].

TSR imputes the missing values in a dataset by carrying out a regression using the scores of its available entries. Considering the partition of \mathbf{X} in Eq. 3 and its decomposition in Eq. 4, the missing elements in \mathbf{x}_n^T are estimated as:

$$\hat{\mathbf{x}}_n^\# = \mathbf{S}^{\#*} \mathbf{P}^* (\mathbf{P}^{*\top} \mathbf{S}^{**} \mathbf{P}^*)^{-1} \mathbf{P}^{*\top} \mathbf{x}_n^* \quad (8)$$

where $\mathbf{S}^{**} = \frac{\mathbf{x}^{*\top} \mathbf{x}^*}{N-1}$ and $\mathbf{S}^{\#*} = \frac{\mathbf{x}^{\# \top} \mathbf{x}^*}{N-1}$.

The imputation is iteratively executed until the reconstruction of the missing values stabilises.

A complete survey on TSR can be found in [9]. A Graphical User Interface for TSR-based missing data imputation, the Missing Data Imputation Toolbox for MATLAB [21], is also available at http://mseg.webs.upv.es/Software_e.html.

Calibration transfer by TSR is achieved in the same way as for MLPCA, that is building the augmented array \mathbf{X}_{ab} and submitting it to the computational procedure described before (see Figure 1, *Imputation* box).

3.4. Joint-Y Partial Least Squares regression (JYPLS)

JYPLS [10] is a Non-linear Iterative PARTial Least Squares (NIPALS) algorithm variant, initially developed for modelling the latent variable structure shared by two or more sets of data (say \mathbf{X} s) via a PLS-based regression against their corresponding responses (say \mathbf{Y} s). When only two different couples of data blocks are dealt with, namely \mathbf{X}_a - \mathbf{Y}_a and \mathbf{X}_b - \mathbf{Y}_b , the mathematical formulation of the JYPLS model is given by:

$$\mathbf{Y}_J = \begin{bmatrix} \mathbf{Y}_a \\ \mathbf{Y}_b \end{bmatrix} = \begin{bmatrix} \mathbf{T}_a \\ \mathbf{T}_b \end{bmatrix} \mathbf{Q}_J^\top + \mathbf{E}_{\mathbf{Y}_J} \quad (9)$$

$$\mathbf{X}_a = \mathbf{T}_a \mathbf{P}_a^\top + \mathbf{E}_{\mathbf{X}_a} \quad (10)$$

$$\mathbf{X}_b = \mathbf{T}_b \mathbf{P}_b^\top + \mathbf{E}_{\mathbf{X}_b} \quad (11)$$

$$\mathbf{T}_a = \mathbf{X}_a \mathbf{W}_a^* \quad (12)$$

$$\mathbf{T}_b = \mathbf{X}_b \mathbf{W}_b^* \quad (13)$$

where $\mathbf{T}_a/\mathbf{T}_b$, $\mathbf{P}_a/\mathbf{P}_b$ and $\mathbf{W}_a^*/\mathbf{W}_b^*$ are the JYPLS scores, loadings and weighing matrices related to $\mathbf{X}_a/\mathbf{X}_b$, respectively. The originality of this approach concerns the fact that only one single set of loadings, \mathbf{Q}_J , is derived for both \mathbf{Y}_a and \mathbf{Y}_b , which defines a combined plane mapped by the \mathbf{Y}_J joint array (see Equation 9).

Until now, JYPLS has been mainly resorted to for product transfer between different manufacturing sites, but here its application is extended to cases in which the common sources of variation underlying measurements resulting from multiple instruments and mostly related to specific properties of interest need to be modelled (i.e. calibration transfer). To this end, two possible JYPLS-based computing strategies were implemented, namely JYPLS-noinv and JYPLS-inv.

- JYPLS-noinv - Let \mathbf{Y}_a and \mathbf{Y}_b be the matrices including the measured response variables associated to the samples analysed by the primary and the secondary spectrometer, respectivelyⁱⁱ. Once a JYPLS model is built as in Equations 9-13 (see Figure 1, *Model transfer* box), the same responses for new samples characterised by the secondary instrument, $\mathbf{Y}_{b,new}$, can be predicted from their spectral profiles, $\mathbf{X}_{b,new}$, as (see Figure 1, *External validation II* box):

$$\mathbf{Y}_{b,new} = \mathbf{X}_{b,new} \mathbf{W}_b^* \mathbf{Q}_J^T \quad (14)$$

- JYPLS-inv - On the other hand, as for TSR, spectra unrecorded by the secondary instrument can be reconstructed, provided they are associated to samples analysed by the primary one and whose response values ($\mathbf{Y}_{b,unrecorded}$) are then present in \mathbf{Y}_a , by the following inversion (see Figure 1, *Model inversion* box):

$$\mathbf{X}_{b,unrecorded} = \mathbf{Y}_{b,unrecorded} (\mathbf{Q}_J \mathbf{Q}_J^T)^\dagger \mathbf{Q}_J \mathbf{P}_b^T \quad (15)$$

where † denotes the Moore-Penrose pseudo-inverse [22]. Such *imputed* spectra, fused to \mathbf{X}_b , are then exploited for fitting an improved PLS predictive model (see Figure 1, *Model calibration* box), suitable for the assessment of future incoming data (see Figure 1, *External validation I* box).

ⁱⁱHere the rows of \mathbf{Y}_b are also contained in \mathbf{Y}_a , as they relate to samples analysed by both the secondary and the primary instrument. This is, however, not a necessary requirement to apply JYPLS.

4. Modelling procedure

The comparative study among PDS, MLPCA, TSR and JYPLS was carried out according to a 5-step procedure (see Figure 1):

1. both the primary and secondary instrument data blocks were randomly split into calibration (2 thirds of the original spectra) and validation (1 third of the original spectra) sets (see Figure 1, *Data split* box). 20 split rounds were conducted to prevent spurious results from being yielded;
2. secondary instrument calibration subsets of increasing size were generated to determine the minimum number of measurements needed to be collected for accomplishing an accurate calibration transfer. The samples belonging to each one of these subsets were selected by the Kennard-Stone (KS) algorithm [23], probably the most popular computational procedure for data-representative object identification [24, 25] (see Figure 1, *Sample selection* box)ⁱⁱⁱ;
3. the four methods under study were then applied in the following fashion:
 - when TSR, MLPCA and JYPLS-inv, which are missing data imputation-based approaches (see Section 1), were handled, the secondary instrument calibration spectra left out of each subset were consecutively reconstructed as described before (see Figure 1, *Imputation, Model transfer* and *Model inversion* boxes). They were then merged with those belonging to the respective calibration subset to fit a new PLS regression model (see Figure 1, *Model calibration* box);
 - by JYPLS-noinv, predictive JYPLS models were constructed fusing both the primary spectrometer calibration set and the different secondary spectrometer calibration subsets as detailed in Section 3.4 (see Figure 1, *Model transfer* box);
 - the PDS standardisation was performed relating the secondary instrument calibration subsets of spectra to their corresponding profiles registered by the primary spectrometer (see Figure 1, *Parameter fitting* and *Standardisation* boxes). Notice that the properties of interest of the corrected spectra are thereafter predicted by a PLS regression

ⁱⁱⁱHere, KS is run on the scores of a PLS model resulting from the primary spectrometer calibration data.

model built on the whole primary instrument calibration set (see Figure 1, *External validation III* box).

For the various strategies, the parameters to be optimised (number of components of the imputation model, number of components of the regression model, PDS spectral window width) were adjusted in order to minimise the average Root Mean Square Error in Cross-Validation (RMSECV), defined as:

$$RMSECV = \frac{\sum_{k=1}^K \sqrt{\frac{\sum_{n=1}^N (y_{n,k} - \hat{y}_{n,k})^2}{N}}}{K} \quad (16)$$

where $y_{n,k}$ represents the actual value of the k -th response variable associated to the n -th calibration sample and $\hat{y}_{n,k}$ is its final prediction^{iv}.

4. The performance of PDS, MLPCA, TSR and JYPLS were finally assessed according to the average Root Mean Square Error in Prediction (RMSEP):

$$RMSEP = \frac{\sum_{k=1}^K \sqrt{\frac{\sum_{n'=1}^{N'} (y_{n',k} - \hat{y}_{n',k})^2}{N'}}}{K} \quad (17)$$

where $y_{n',k}$ represents the actual value of the k -th response variable associated to the n' -th validation sample and $\hat{y}_{n',k}$ is its final prediction, while N' equals the total number of spectra included in the validation set^{iv} (see Figure 1, *External validation I*, *External validation II* and *External validation III* boxes);

5. Statistically significant differences among the considered approaches were finally evaluated via a mixed-effect ANalysis Of VAriance (ANOVA), taking into account four factors: calibration transfer technique, size of the secondary instrument calibration subset and their interaction (fixed-effect factors) as well as split round (random-effect factor, nested to the size of the secondary instrument calibration subset). If any effect or interaction was found to be statistically significant, the 95% Least Significance Difference (LSD) intervals were calculated to assess which methods were different from the others.

^{iv}The reported RMSECV and RMSEP values concern autoscaled response variables owing to the differences in their original units of measurements.

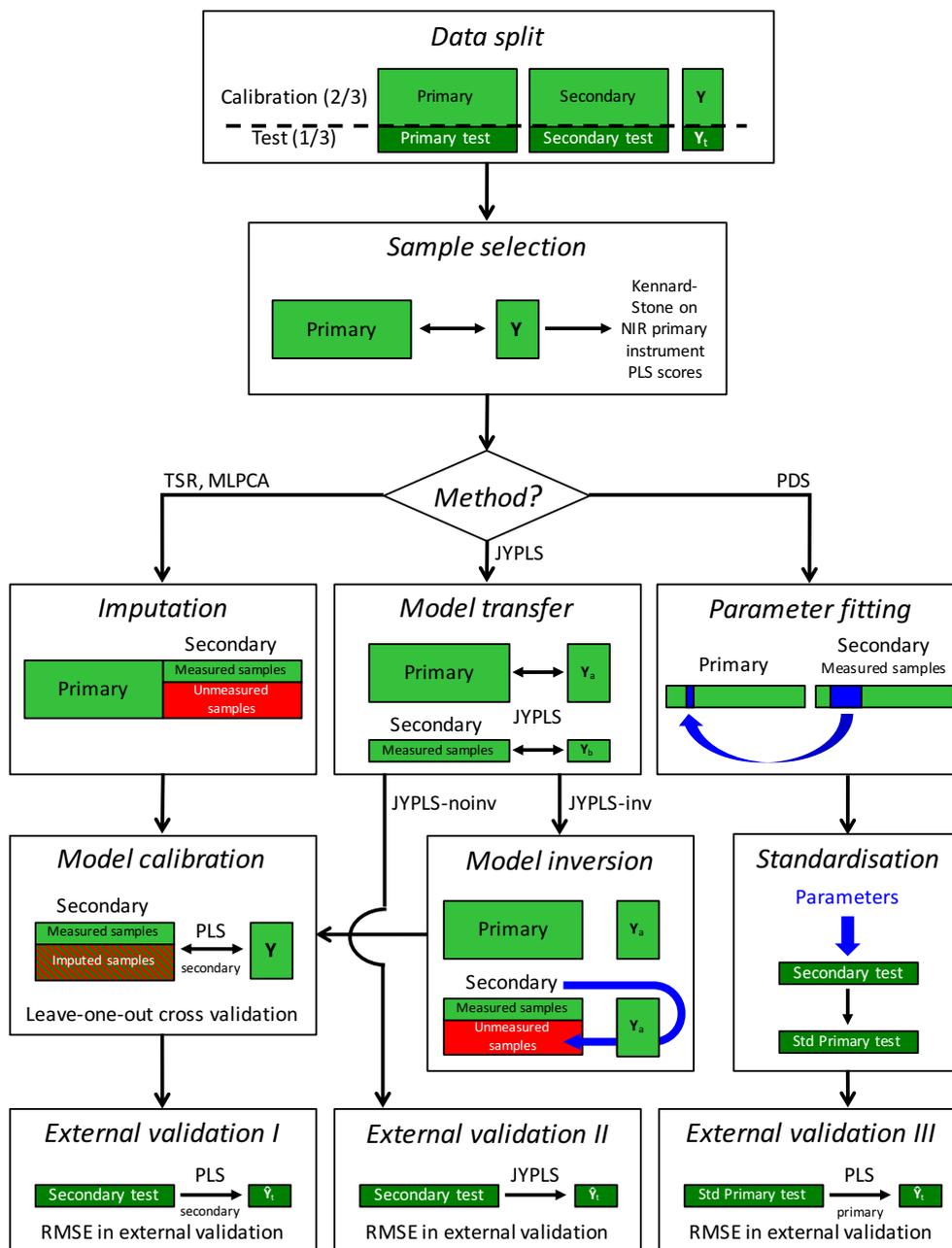


Figure 1: Flow-chart of the comparative study. Std stands for *standardised*. $\hat{\cdot}$ refers to predicted values. Notice that part of the whole secondary instrument calibration set is assumed to be unmeasured when addressing the calibration transfer

It is worth noting that in all the cases the values of the various response variables were predicted simultaneously by global PLS2-based models. Although performing a single PLS1 regression for each one of them could have constituted a better solution from a predictive perspective, the adopted procedure enabled an easier and more immediate global comparison.

5. Results

5.1. Gasoline dataset

For each spectrometer, 20 pseudo-gasoline samples were assigned to the calibration set and the remaining 10 to the validation set. 15 secondary instrument calibration subsets, containing from 5 to 19 spectral profiles, were generated.

5.1.1. Missing data imputation

As TSR, JYPLS-inv and MLPCA rely on a preliminary missing data imputation step, it is worth assessing the accuracy of the reconstruction of the unmeasured spectra, since they will be then resorted to for building the final predictive PLS model.

Figure 2 permits to compare original and imputed profiles for one of the 20 split rounds. Their correlation and χ^2 distance are represented in Figures 2A, 2D, 2G and 2B, 2E, 2H, respectively. Each line refers to the best model selected for one specific secondary instrument calibration subset. High correlations (larger than 0.9999) and low χ^2 distance values (smaller than 0.001) were yielded by TSR and JYPLS-inv, while several issues appeared when dealing with MLPCA. First, it often suffered from convergence problems (as already pointed out by Feudale *et al.* [3]), which dramatically slowed the computational procedure down. Consequently, the reconstructed spectra were found to be considerably different from their actual profiles (see Figure 2I). For these reasons, MLPCA was not taken into account in the final study.

5.1.2. Comparative study

Figure 3A allows the performance of the different calibration transfer techniques under study to be examined. Each point in the plot represents the average *RMS EP* value, estimated from the 10-sample external validation set, over the 20 split rounds (for 5- to 19-sample secondary instrument calibration subsets). As expected, for all the approaches, the higher the size of the secondary

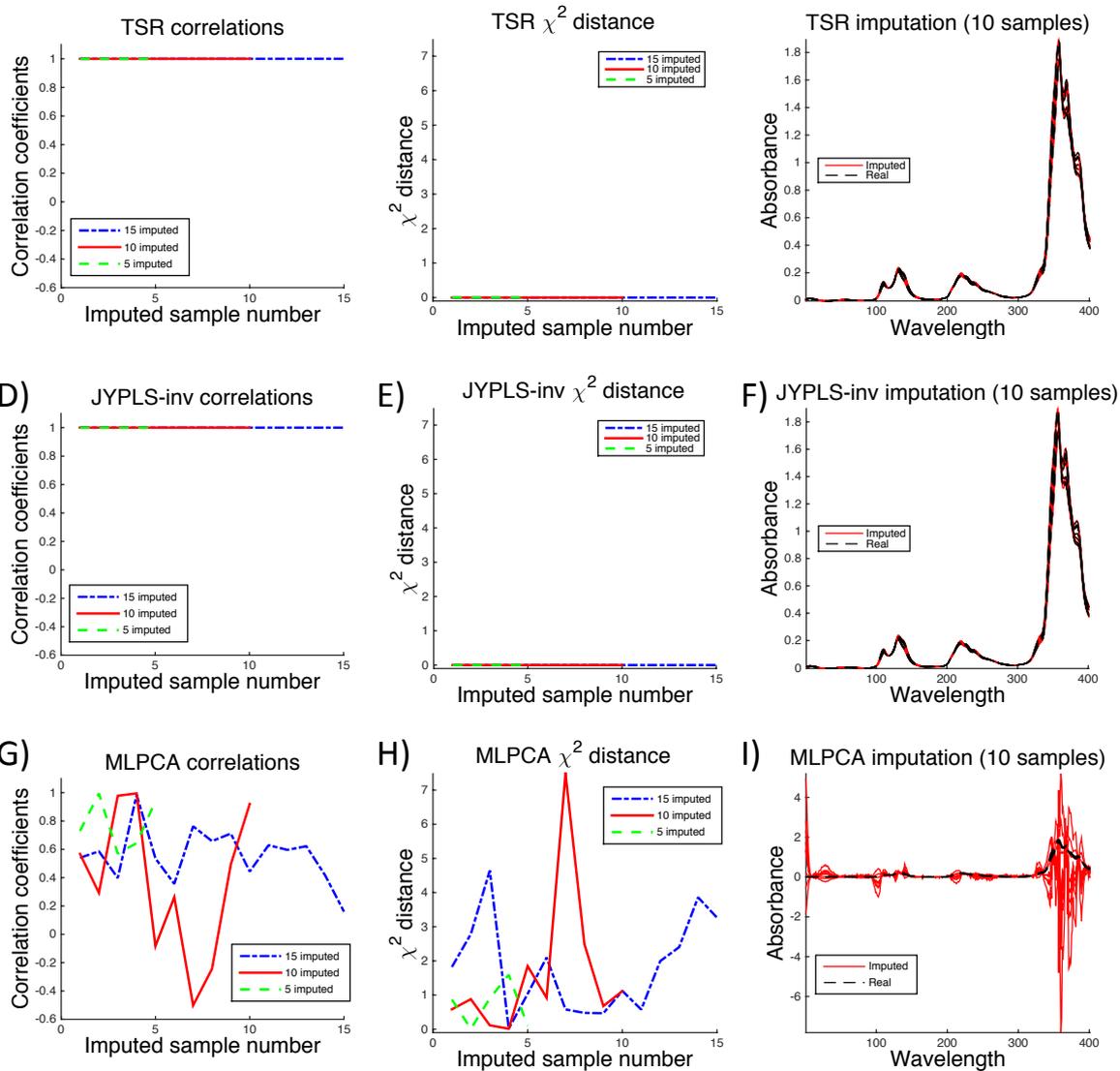


Figure 2: Gasoline dataset - A), D) and G) show the correlation coefficients between the original spectra and those imputed by TSR, JYPLS-inv and MLPCA, respectively. B), E) and H) represent their corresponding χ^2 distance values. The dotted-dashed blue lines refer to the case in which the secondary instrument calibration subset was constituted by 5 samples and 15 spectra were imputed. The solid red lines refer to the case in which the secondary instrument calibration subset was constituted by 10 samples and 10 spectra were imputed. The dashed green lines refer to the case in which the secondary instrument calibration subset was constituted by 15 samples and 5 spectra were imputed. C), F) and I) display the original and reconstructed profiles in the second of these three cases

instrument calibration subset, the lower the *RMS EP*.

As the effect of all the factors included in the ANOVA model was found to be statistically significant (p -value < 0.05), the 95% LSD intervals were calculated to point out existing differences among methods. For the sake of an easy visualisation, dashed-line ellipses are used in Figure 3A to highlight them. Specifically, methods embraced by the same ellipse show no statistical difference. On the other hand, methods embraced by different ellipses are statistically different.

Clearly, PDS guaranteed the lowest *RMS EP* when the secondary instrument calibration subset consisted of 5/6 samples. No statistically significant differences were detected between PDS and TSR for a 7-sample secondary instrument calibration subset and between PDS and JYPLS-inv when a 10-sample secondary instrument calibration subset was concerned. From 10 samples onwards, the *RMS EP* stabilised around 0.09 for PDS, but it continuously decreased for TSR and JYPLS-inv, reaching values around 0.05-0.06 (for 12-13 to 19 samples). The straight line in Figure 3A indicates the *RMS EP* value obtained when a full recalibration was performed, i.e. when the whole set of 20 secondary instrument calibration samples was used to build a new predictive model. Although it cannot be directly compared to the outcomes resulting from PDS, TSR, JYPLS-inv and JYPLS-noinv, it eases the determination of the number of spectra needed to be collected by the secondary spectrometer for generating no statistically significant differences with respect to full recalibration. TSR required 12, while JYPLS-inv 13. On the other hand, PDS and JYPLS-noinv always showed a statistically worse performance than full recalibration.

5.1.3. Instruments with different resolutions

A common situation faced by practitioners in industrial environments is transferring calibration models between instruments with diverse spectral resolution. This problem has already been addressed in [26], where the authors propose a novel PLS-based approach resulting in similar outcomes as PDS.

Figures 3B-3D show the results of the whole analysis, conducted gradually reducing the spectral resolution of the secondary instrument. The performance of all the methods was basically the same as in the full resolution case described in Section 5.1.2. However, for PDS, a slight gradual decrease in the quality of the calibration transfer can be noticed.

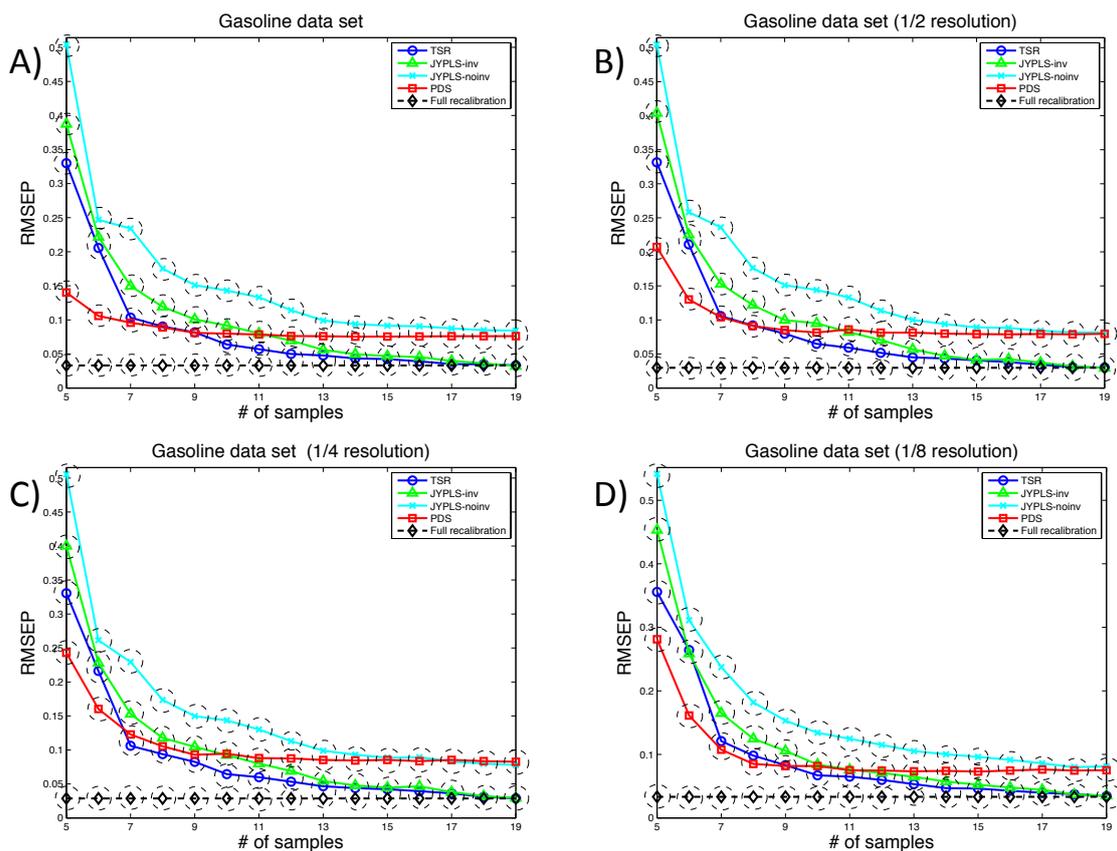


Figure 3: Gasoline dataset - RMSEP values obtained for the different sizes of the secondary spectrometer calibration subset with A) the same spectral resolution for both instruments, B) $\frac{1}{2}$, C) $\frac{1}{4}$ and D) $\frac{1}{8}$ of the primary instrument spectral resolution for the secondary spectrometer. Dashed ellipses mark the statistically significant differences among groups of methods (p -value < 0.05)

5.1.4. Sample selection effect

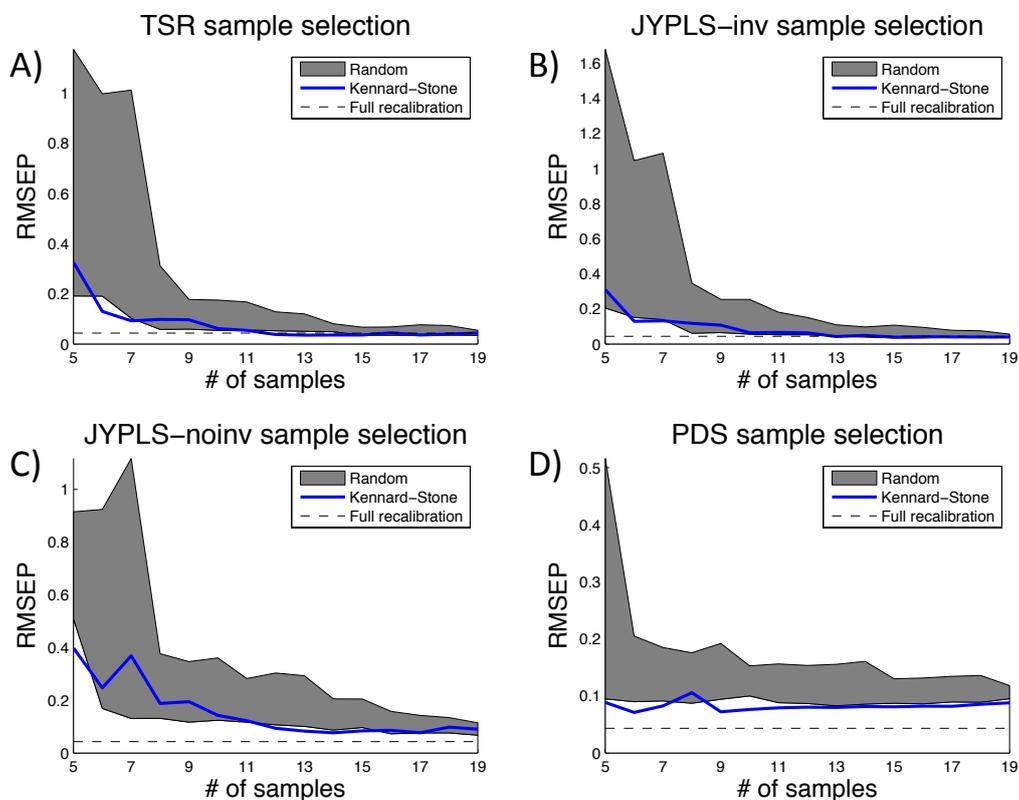


Figure 4: Gasoline dataset - Effect of the Kennard-Stone algorithm-based sample selection on the performance of the calibration transfer methods under study

The effect of the secondary spectrometer calibration subset sample selection is here assessed. 10 random selections were performed for one particular split round and the final *RMSEP* values were then compared to those obtained by preliminarily running KS. It is clear from Figure 4 that KS generally returned a lower RMSEP, very close to that achievable through a full recalibration. It then enabled a better calibration transfer plausibly due to the fact that it permits to choose a subset of samples, which is statistically representative of the experimental domain related to the spectral data collected by the primary instrument. This is not necessarily the case when such a selection is carried out at random.

5.2. Corn dataset

For each spectrometer, 54 corn samples were assigned to the calibration set and the remaining 26 to the validation set. 10 secondary instrument calibration subsets, containing from 5 to 50 spectral profiles (5-spectra intervals), were generated.

5.2.1. Missing data imputation

Figure 5 permits to compare original and imputed corn sample spectral profiles for one of the 20 split rounds. TSR preserved its reconstruction ability and MLPCA suffered from the same problems observed for the gasoline dataset. Regarding JYPLS-inv, the correlation coefficients/ χ^2 distance values were rather high/low, but the imputed spectra showed less variability than the real ones (see e.g. Figure 5F). This happened because the large difference in the offset of these latter is scarcely related to the properties to be predicted. As the imputation here involves the joint- \mathbf{Y} loadings matrix, \mathbf{Q}_J , such a difference is not transferred to the reconstructed spectra (see Equation 9). Thus, one can think of JYPLS-inv as filtering spectral variations, which is uninteresting from a predictive point of view.

5.2.2. Comparative study

Existing differences among methods were investigated as in the previous case-study (also here the effect of all the ANOVA factors was found to be statistically significant). Figure 6A displays the results of the comparative study conducted on the corn dataset. Again, PDS showed a better performance for small secondary instrument calibration subsets (5-10 samples). For 20-25 samples, there were no statistical differences between PDS, TSR, JYPLS-inv and JYPLS-noinv. Finally, from 30 samples onwards, the proposed approaches outperformed PDS, similarly as for the gasoline dataset. From 40 samples onwards, TSR, JYPLS-inv and JYPLS-noinv exhibited no significant differences with respect to full recalibration.

5.2.3. Instruments with different resolutions

In this case, the reduction of the spectral resolution of the secondary instrument strongly affected the quality of the PDS-based calibration transfer. In fact, when the resolution of the secondary spectrometer was decreased to $\frac{1}{8}$, even for small secondary calibration subsets, the per-

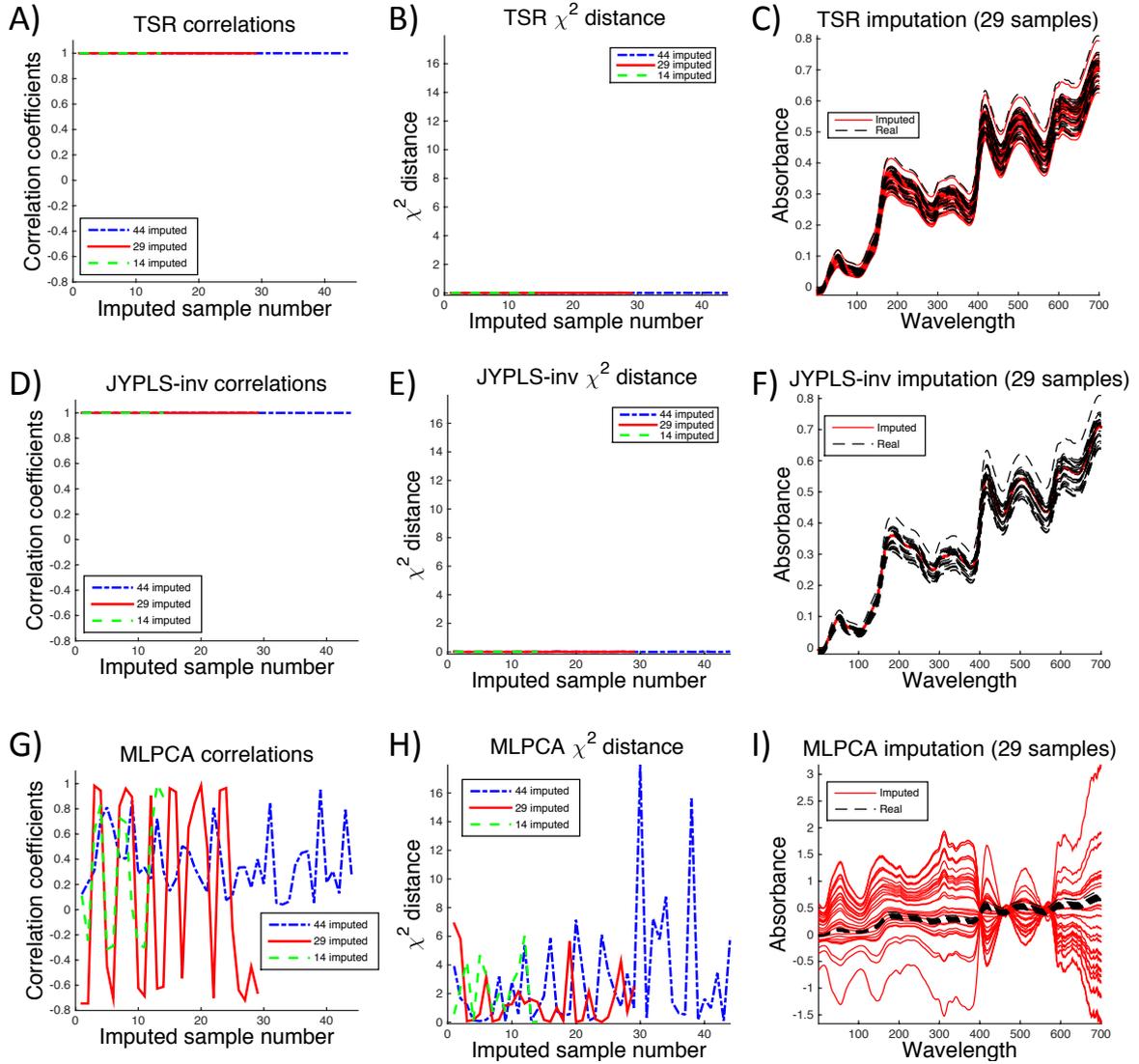


Figure 5: Corn dataset - A), D) and G) show the correlation coefficients between the original spectra and those imputed by TSR, JYPLS-inv and MLPCA, respectively. B), E) and H) represent their corresponding χ^2 distance values. The dotted-dashed blue lines refer to the case in which the secondary instrument calibration subset was constituted by 10 samples and 44 spectra were imputed. The solid red lines refer to the case in which the secondary instrument calibration subset was constituted by 25 samples and 29 spectra were imputed. The dashed green lines refer to the case in which the secondary instrument calibration subset was constituted by 40 samples and 14 spectra were imputed. C), F) and I) display the original and reconstructed profiles in the second of these three cases

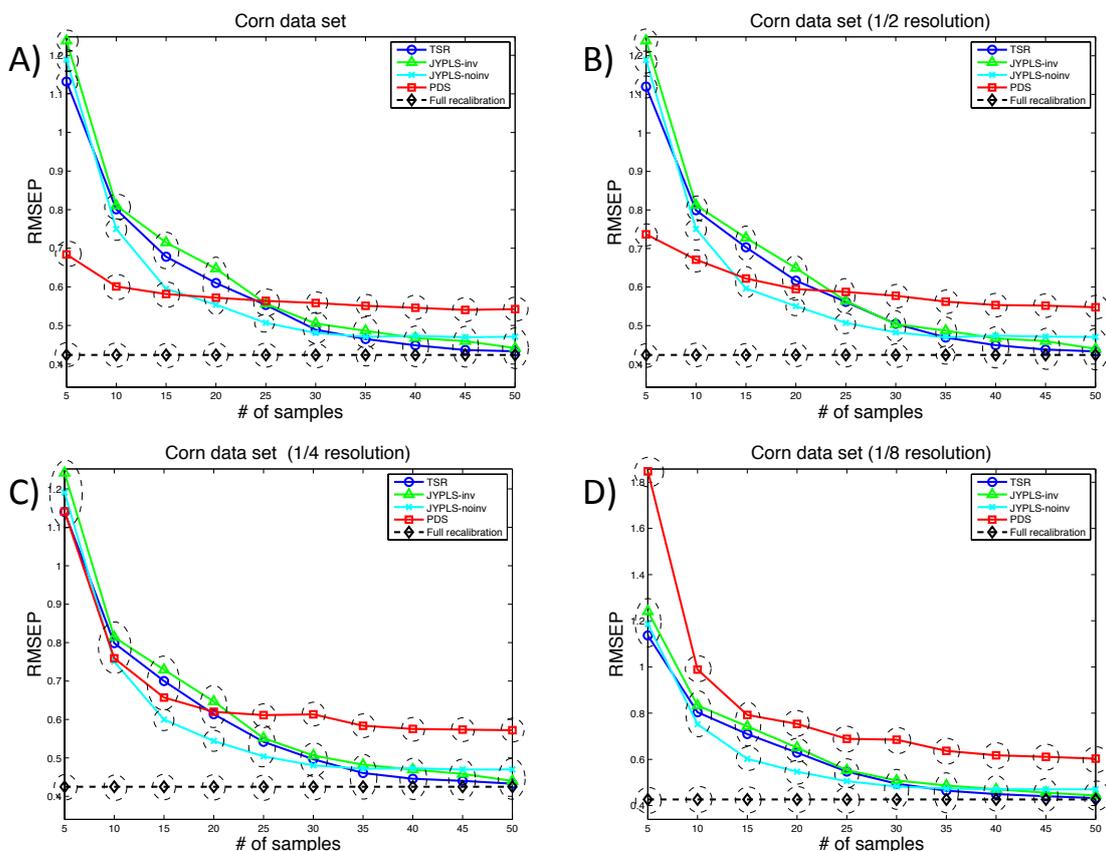


Figure 6: Corn dataset - RMSEP values obtained for the different sizes of the secondary spectrometer calibration subset with A) the same spectral resolution for both instruments, B) $\frac{1}{2}$, C) $\frac{1}{4}$ and D) $\frac{1}{8}$ of the primary instrument spectral resolution for the secondary spectrometer. Dashed ellipses mark the statistically significant differences among groups of methods (p -value < 0.05)

formance of PDS was found to be statistically worse than the other compared approaches. On the other hand, TSR, JYPLS-inv and JYPLS-noinv were found to be quite robust towards such a change (see Figures 6B-6D).

5.2.4. Sample selection effect

The effect of the secondary spectrometer calibration subset sample selection can be evaluated by looking at Figure 7. Here, especially when the number of spectra included in this subset was not particularly high, some random selection runs permitted to obtain better results in terms of *RMSEP*. However, for larger subsets, KS-based selection enabled better prediction than random

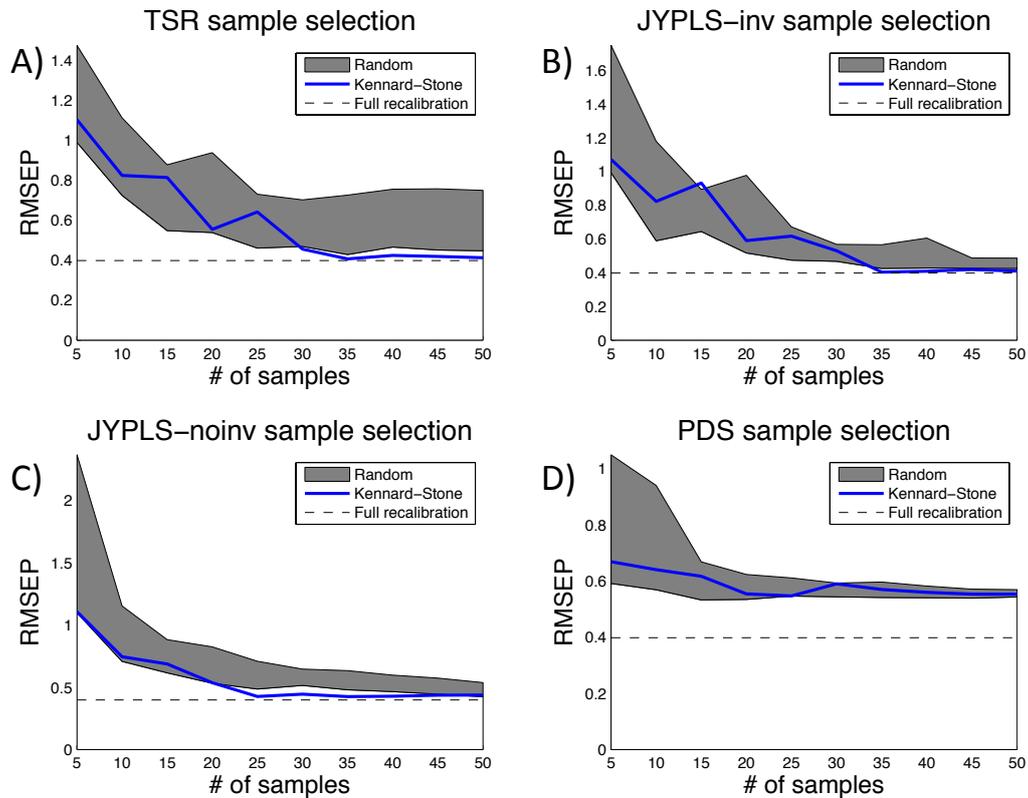


Figure 7: Corn dataset - Effect of the Kennard-Stone algorithm-based sample selection on the performance of the calibration transfer methods under study

ordering.

6. Discussion

When carrying out a calibration transfer with a very small secondary instrument calibration subset (around 5-10 samples), PDS showed better (or equal) results than TSR and JYPLS-inv, but its performance was worse than full recalibration. Nevertheless, when the size of the secondary instrument calibration subset was enlarged, TSR and JYPLS-inv clearly outperformed PDS, achieving a similar error rate as for full recalibration. No evident conclusions can be drawn regarding the performance of JYPLS-noinv, as the quality of its outcomes changed depending on the analysed dataset. On the other hand, it can be said that it was found to be, in general, as reliable as TSR and

JYPLS-inv when the corn dataset was dealt with, but statistically worse in the gasoline case-study. The number of spectra to be collected by the secondary spectrometer for a precise calibration transfer was also assessed. TSR and JYPLS-inv yielded very similar results to full recalibration even if only about the 60% of the available spectra were included in the corresponding calibration subset. On the other hand, PDS never reached such a degree of accuracy.

PDS was strongly affected by the reduction of the spectral resolution of the secondary instrument when the corn dataset was concerned, while TSR, JYPLS-inv and JYPLS-noinv seemed not to suffer from the same issue.

In terms of unmeasured spectra reconstruction, TSR resulted in the best performance. In contrast, JYPLS-inv acted as a sort of filter removing the variations in the spectra not related to the properties to be predicted, and consequently producing deviations from their original shape.

Moreover, it is worth saying that both JYPLS-inv and JYPLS-noinv are the unique strategies, which could be resorted to when distinct or partially distinct sets of samples^v are analysed by the two spectrometers.

Finally, it was shown that selecting the samples using KS generally permitted to achieve better results, regardless the calibration transfer technique resorted to.

7. Conclusions

Two novel methods to perform calibration transfer between NIR spectrometers, based on TSR and JYPLS, respectively, were proposed. They outperformed PDS and guaranteed a very similar performance to that resulting from a full recalibration when only about the 60% of the spectra collected by the secondary instrument was available. Both the approaches also showed a sufficient robustness towards the reduction of its spectral resolution. In addition, TSR allowed unmeasured spectra to be accurately imputed, while the inversion of the JYPLS models yielded reconstructed spectral profiles filtered of all the variation not of interest from a predictive point of view.

^vProvided that the values of their properties of interest are known.

8. Acknowledgments

This research work was partially supported by the Spanish Ministry of Economy and Competitiveness under the project DPI2014-55276-C5-1R and Shell Global Solutions International B.V. (Amsterdam, The Netherlands).

9. References

- [1] H. Martens, T. Næs, *Multivariate Calibration*, 1st Edition, John Wiley & Sons Ltd., 1989.
- [2] K. Esbensen, *Multivariate Data Analysis - in practice*, 5th Edition, CAMO Process AS, 2002.
- [3] R. Feudale, N. Woody, H. Tan, A. Myles, S. Brown, J. Ferré, Transfer of multivariate calibration models: a review, *Chemometr. Intell. Lab.* 64 (2002) 181–192.
- [4] Y. Wang, D. Veltkamp, B. Kowalski, Multivariate instrument standardization, *Anal. Chem.* 63 (1991) 2750–2756.
- [5] E. Bouveresse, D. Massart, Standardisation of near-infrared spectrometric instruments: a review, *Vib. Spectrosc.* 11 (1996) 3–15.
- [6] T. Fearn, Standardisation and calibration transfer for near infrared instruments: a review, *J. Near Infrared Spec.* 9 (2001) 229–244.
- [7] O. de Noord, Multivariate calibration standardisation, *Chemometr. Intell. Lab.* 25 (1994) 85–97.
- [8] D. Andrews, P. Wentzell, Applications of Maximum Likelihood Principal Component Analysis: incomplete data sets and calibration transfer, *Anal. Chim. Acta* 350 (1997) 341–352.
- [9] A. Folch-Fortuny, F. Arteaga, A. Ferrer, PCA model building with missing data: new proposals and a comparative study, *Chemometr. Intell. Lab.* 146 (2015) 77–88.
- [10] S. García Muñoz, J. MacGregor, T. Kourti, Product transfer between sites using Joint-Y PLS, *Chemometr. Intell. Lab.* 79 (2005) 101–114.
- [11] B. Wise, H. Martens, M. Høy, R. Bro, P. Brockhoff, Calibration transfer by Generalized Least Squares, Tech. rep., Eigenvector Research Inc. (www.eigenvector.com/Docs/GLS_Standardization.pdf).
- [12] E. Bouveresse, D. Massart, Improvement of the piecewise direct standardisation procedure for the transfer of NIR spectra for multivariate calibration, *Chemometr. Intell. Lab.* 2 (1996) 201–213.
- [13] Y. Wang, B. Kowalski, Calibration transfer and measurement stability of near-infrared spectrometers, *Appl. Spectrosc.* 46 (1992) 764–771.
- [14] *PLS_Toolbox 7.0.2*, Eigenvector Research Inc., Manson, Washington, USA (2012).
- [15] P. R. Nelson, The treatment of missing measurements in PCA and PLS models, Ph.D. thesis, MacMaster University (2002).

- [16] F. Arteaga, Control estadístico multivariante de procesos con datos faltantes mediante Análisis de Componentes Principales, Ph.D. thesis, Universitat Politècnica de València (2003).
- [17] A. Folch-Fortuny, F. Arteaga, A. Ferrer, Assessment of Maximum Likelihood PCA missing data imputation, *J. Chemometr.* 30 (2016) 386–393.
- [18] P. R. Nelson, P. A. Taylor, J. F. MacGregor, Missing data methods in PCA and PLS: score calculations with incomplete observations, *Chemometr. Intell. Lab.* 35 (1996) 45–65.
- [19] F. Arteaga, A. Ferrer, Dealing with missing data in MSPC: several methods, different interpretations, some examples, *J. Chemometr.* 16 (2002) 408–418.
- [20] F. Arteaga, A. Ferrer, Framework for regression-based missing data imputation methods in on-line MSPC, *J. Chemometr.* 19 (2005) 439–447.
- [21] A. Folch-Fortuny, F. Arteaga, A. Ferrer, Missing data imputation toolbox for MATLAB, *Chemometr. Intell. Lab.* 154 (2016) 93–100.
- [22] J. Alves Barata, M. Hussein, The Moore-Penrose pseudoinverse: a tutorial review of the theory, *Braz. J. Phys.* 42 (2012) 146–165.
- [23] R. Kennard, L. Stone, Computer aided design of experiments, *Technometrics* 11 (1969) 137–148.
- [24] M. Daszykowski, B. Walczak, D. Massart, Representative subset selection, *Anal. Chim. Acta* 468 (2002) 91–103.
- [25] R. Vitale, M. Bevilacqua, R. Bucci, A. Magrì, A. Magrì, F. Marini, A rapid and non-invasive method for authenticating the origin of pistachio samples by NIR spectroscopy and chemometrics, *Chemometr. Intell. Lab.* 121 (2013) 90–99.
- [26] J. Fernández-Pierna, P. Vermeulen, B. Lecler, V. Baeten, P. Dardenne, Calibration transfer from dispersive instruments to handheld spectrometers, *Appl. Spectrosc.* 64 (2010) 644–647.