Talanta 179 (2018) 386–392

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Calibration transfer of a Raman spectroscopic quantification method for the assessment of liquid detergent compositions from at-line laboratory to in-line industrial scale

D. Brouckaert[a], J.-S. Uyttersprot[b], W. Broeckx[b], T. De Beer[a,*]

[a] Laboratory of Pharmaceutical Process Analytical Technology, Ghent University, Ottergemsesteenweg 460, 9000 Ghent, Belgium
[b] Procter & Gamble, Brussels Innovation Centre, Temselaan 100, 1853 Strombeek-Bever, Belgium

ARTICLE INFO

Keywords:
Process analytical technology
Raman spectroscopy
Calibration transfer
In-line calibration
Real-time quantification

ABSTRACT

Calibration transfer or standardisation aims at creating a uniform spectral response on different spectroscopic instruments or under varying conditions, without requiring a full recalibration for each situation. In the current study, this strategy is applied to construct at-line multivariate calibration models and consequently employ them in-line in a continuous industrial production line, using the same spectrometer.

Firstly, quantitative multivariate models are constructed at-line at laboratory scale for predicting the concentration of two main ingredients in hard surface cleaners. By regressing the Raman spectra of a set of small-scale calibration samples against their reference concentration values, partial least squares (PLS) models are developed to quantify the surfactant levels in the liquid detergent compositions under investigation. After evaluating the models performance with a set of independent validation samples, a univariate slope/bias correction is applied in view of transporting these at-line calibration models to an in-line manufacturing set-up. This standardisation technique allows a fast and easy transfer of the PLS regression models, by simply correcting the model predictions on the in-line set-up, without adjusting anything to the original multivariate calibration models.

An extensive statistical analysis is performed in order to assess the predictive quality of the transferred regression models. Before and after transfer, the R^2 and RMSEP of both models is compared for evaluating if their magnitude is similar. T-tests are then performed to investigate whether the slope and intercept of the transferred regression line are not statistically different from 1 and 0, respectively. Furthermore, it is inspected whether no significant bias can be noted. F-tests are executed as well, for assessing the linearity of the transfer regression line and for investigating the statistical coincidence of the transfer and validation regression line. Finally, a paired *t*-test is performed to compare the original at-line model to the slope/bias corrected in-line model, using interval hypotheses.

It is shown that the calibration models of Surfactant 1 and Surfactant 2 yield satisfactory in-line predictions after slope/bias correction. While Surfactant 1 passes seven out of eight statistical tests, the recommended validation parameters are 100% successful for Surfactant 2. It is hence concluded that the proposed strategy for transferring at-line calibration models to an in-line industrial environment via a univariate slope/bias correction of the predicted values offers a successful standardisation approach.

1. Introduction

Calibration transfer or standardisation can be defined as the uniformisation of the spectral response on different spectroscopic instruments or under altered environmental conditions in view of eliminating the need for time-consuming recalibration procedures [1–4]. Such standardisation strategies are desirable on a number of occasions that result in the failure of the originally developed multivariate calibration methods.

Firstly, drifts or nonlinearities in the instrumental response function of a spectrometer may be witnessed over time due to ageing of the equipment. Analogously, within one spectrometer, the instrumental response can alter due to a repair, causing a shift in the wavelength axis. Secondly, multivariate calibration models may become invalid when it is pursued to transport an existing chemometric model from one instrument to a second instrument. A variety of reasons may render

* Corresponding author.
E-mail addresses: davinia.brouckaert@ugent.be (D. Brouckaert), uyttersprot.js@pg.com (J.-S. Uyttersprot), broeckx.wa@pg.com (W. Broeckx), thomas.debeer@ugent.be (T. De Beer).

https://doi.org/10.1016/j.talanta.2017.11.025
Received 9 September 2017; Received in revised form 14 November 2017; Accepted 16 November 2017
Available online 20 November 2017
0039-9140/

erroneous results on this secondary or slave apparatus, as the measured intensity values from two spectrometers will generally be different, as may the wavelength axes and peak locations be. Thirdly, changes in the physical or chemical constitution of the samples between the calibration and prediction step may cause issues. When modifications to the viscosity, particle size or surface texture arise, the model predictions may become inaccurate. Likewise, when batch-to-batch variations occur due to modifications in raw materials or sample preparation, this may lead to incorrect prediction results. Furthermore, changes in the instruments environment exhibit another situation where a multivariate calibration model may become invalid. Fluctuations in temperature or humidity, for instance, may strongly influence the spectral outcome of the instrument [1–4].

As the construction of qualitative or quantitative multivariate models is often an elaborate process that requires quite a lot of effort and resources, it would negatively impact the business if full recalibrations would be required under all these circumstances, since they involve a lot of development costs and time delays. Therefore, a number of standardisation strategies has been developed with the objective of avoiding the need for a full recalibration. Generally, they can be divided into two main classes: strategies that can be applied prior to model implementation and instrument standardisation methods that can be of use after the model is already in operation.

Methodologies that can be performed before model implementation are instrument matching, global modelling, model updating and sensor selecting. These techniques aim at avoiding the need for data transformation by carefully controlling the experimental and environmental conditions and by selecting appropriate instrumental parameters [1–3].

However, when it is impossible to foresee all future sources of variation or these instrumental and experimental design techniques offer insufficiently accurate results, standardisation techniques are required to deal with non-calibrated variations that are observed after the model is already in use. Herein, a distinction can be made between transfer strategies that standardise either the model coefficients, the spectral responses or the predicted values. Modification of the spectral responses is the most popularly applied approach for diminishing the observed dissimilarities in model outcome. Univariate procedures such as the Shenk and Westerhaus standardisation or single wavelength standardisation (SWS), for example, allow to correct simple wavelength shifts or linear intensity differences in spectra. In case the transfer issue is more complex and involves the correction of peak broadening or effects in the data where the covariance between separate channels cannot be neglected, multivariate standardisation methods are employed. Among the most commonly suggested standardisation strategies in this category are direct standardisation (DS) and piece-wise direct standardisation (PDS) [1–4]. In view of correcting the predicted values rather than the spectral responses, only one standardisation approach is widely used, namely the slope/bias correction (SBC). In this univariate method, a linear relationship is assumed between the predictions from the secondary instrument or conditions and the corresponding predictions that would have been obtained on the primary instrument or conditions. This standardisation approach allows to rapidly compensate for simple and systematic differences and has proven valuable on a number of occasions [1,2,5–10].

In this study, it is aimed to transfer PLS regression quantification models from an at-line to an in-line set-up within one spectrometer. Firstly, quantitative calibration models are constructed for predicting the concentration levels of two surfactants in a liquid detergent composition based on Raman spectra. In order to allow the preparation of calibration samples with deviating compositions at a small scale, these multivariate regression models need to be developed at-line. Secondly, once it is validated that the model can adequately predict the concentration of unknown samples, a transfer of these at-line calibration models to an in-line manufacturing environment at industrial scale is pursued. Running the constructed quantitation models in a production line would allow to evaluate the quality of the fabricated liquid detergent compositions in real-time. This would offer the opportunity of quickly observing trends and regulating the process when products tend to go out of specification. Such fast interventions could result in decreased manufacturing costs and increased product quality, while requiring less manpower [11].

The calibration transfer challenge in this specific case lies in overcoming the differences in spectral response that are witnessed between the at-line and the in-line set-up. As only one Raman instrument is employed, it is a change in environmental factors and sample presentation that defines the observed variation between the primary (at-line) and secondary (in-line) conditions. It is expected that this alteration in sampling strategy affects the response of the Raman spectrometer. Previously, Karande et al. have illustrated that shifting from static to dynamic sampling influences the acquired spectral data and hence has an impact on the quality of the resulting multivariate models [12]. This study on near-infrared spectroscopy stresses the importance of appropriate sampling and shows the negative consequences of dissimilar sampling strategies between the calibration and the prediction step.

Furthermore, previous work by the authors has revealed that such dissimilarities in spectral outcome between static and dynamic sampling are to be expected in Raman spectroscopy as well [10]. As this experimental work, which served as a proof of concept for this standardisation strategy, has illustrated that a univariate slope/bias correction allows to transfer calibration models from an at-line to an in-line set-up at laboratory scale, the same approach is addressed here [10]. Slope/bias corrections are hence employed to achieve a transfer of the two PLS calibration models between the stationary at-line and dynamic in-line conditions. This univariate standardisation of the predicted values offers the benefit of being simple, practical and straightforward, without requiring sophisticate software packages or complex calculations [1–3,13,14].

A handful of studies are described in literature, which value the use of univariate slope/bias correction in achieving a successful standardisation [5–9]. A first series of investigations aimed to transfer quantitative multivariate calibrations between different near-infrared spectrometers [5–7]. Comparing several standardisation strategies to achieve their goal, the authors have witnessed that a simple slope/bias correction could successfully be applied and yielded small prediction errors. Analogously, Sales et al. have investigated both univariate and multivariate standardisation techniques for extending the lifetime of a PLS calibration model based on UV–VIS spectra [8]. Herein, the authors highlight the preference of slope/bias correction over multivariate PDS due to the simplicity of the univariate method. More recently, Brito et al. evaluated the feasibility of instrument standardisation to transfer calibration models between a bench scanning and a submersible diode array spectrophotometer [9]. It was concluded that slope/bias correction yielded the most adequate results in comparison to other univariate transfer approaches.

To the best of our knowledge, only one study has previously been performed on the calibration transfer of multivariate calibration models from an at-line to an in-line measurement set-up via Raman spectroscopy [10]. Since the results of this small-scale investigation showed promising perspectives for in-line quantification based on at-line calibration, the proposed strategy is expanded in this study to a continuous manufacturing process at industrial scale.

2. Materials and methods

2.1. Liquid detergent composition

The liquid detergent compositions under investigation are hard surface cleaners (HSC), consisting of fourteen ingredients in a largely water-based carrier. These constituents are a combination of surfactants, builders, polymers, solvents, dyes, perfumes, preservatives, viscosity modifiers and pH-adjustment agents. As the proprietary

confidential formulas cannot be concealed, the components of interest are named Surfactant 1 and Surfactant 2 in this manuscript. Their target concentration is 2.640% and 6.000%, respectively. Except for water, all other components are present in quantities smaller than 1%. Four HSC formulas are inspected. These cleaning solutions are identical from a chemical point of view, but differ in the type of dye and perfume added to finalise the formulations. Depending on the raw material availability, two different types of Surfactant 2 are included as well. Both are non-ionic alkyl polyethylene glycol ethers, which slightly differ in terms of chain length and presence or absence of side chains.

2.2. Raman spectroscopy

A Viserion Raman instrument, compiled by Indatech (Montpellier, France), is employed for both at-line and in-line spectral data acquisition. The spectrometer is built up of a CCD detector (Horiba, Kyoto, Japan), a 785 nm diode laser (PD-LD, Pennington, USA) and a fibre-optic probe (Indatech, Montpellier, France). This fibre-optic probe allows to acquire spectral information directly in-line in a production environment. The liquid detergent compositions, produced at a speed of multiple tonnes per hour, are hence passing by the Raman probe, which is implemented in the continuous manufacturing line. The accompanying Viserion 1.06 software (Indatech, Montpellier, France) is engaged for data acquisition. An integration time of six seconds is selected in view of obtaining in-line spectral data with a satisfying signal-to-noise ratio. The same acquisition parameters are respected for the at-line data collection. There, liquid detergent compositions are prepared at laboratory scale in 100 mL cups. An at-line Raman set-up is created by fixating the fibre-optic probe at the edge of a box, facing downwards inside the sample cups and closing the box for preventing the interference of ambient light.

2.3. Calibration model development

Two individual calibration models are constructed to predict the concentration of Surfactant 1 and Surfactant 2, respectively, in all liquid detergent compositions, whatever their dye or perfume. PLS regression is selected for creating the quantitative models by linking the multivariate dataset acquired via Raman spectroscopy to the reference concentration values of the corresponding samples. Therefore, a calibration set is created at laboratory scale, consisting of thirty-six samples with the concentration levels of Surfactant 1 and Surfactant 2 varying at five levels: target – 15%, target – 5%, target, target + 5% and target + 15%. Furthermore, the types of dye, perfume and Surfactant 2 are altered in order to include this source of variability in the calibration model. Three repetitive measurements of these thirty-six calibration samples are collected at-line and their Raman spectra are regressed against the concentration levels of Surfactant 1 and Surfactant 2, respectively. These reference concentration data are derived from the amounts of raw materials weighed during sample preparation, taking the impurity of the batches (such as divergent water levels) into account.

As a validation set, nineteen samples are collected from in-line runs, during which the concentration of the components of interest is varied between 2.244% and 3.036% for Surfactant 1 and between 5.100% and 6.900% for Surfactant 2. Raman spectra are continuously collected in-line during manufacturing, as well as at-line after gathering of the produced formulas. As only one type of dye and perfume is present in these nineteen samples with varying composition, three other production samples, in which all components are present in their target concentration, are added to the validation set, in view of representing all the dye types.

The acquired Raman spectra are loaded into MATLAB R2016a (The MathWorks, Inc., USA) for further processing. PLS calibration models are constructed by regressing the spectral information from the at-line calibration samples against their gravimetrically acquired reference values. Using the model optimiser of PLS_Toolbox (Eigenvector Research, Inc., Washington, USA), an assortment of spectral ranges, pre-processing parameters and numbers of latent variables is compared, as described in earlier work by the authors [13]. Spectral pre-processing is always followed by mean centring and all listed filters are tested either individually or successively.

The at-line acquired spectra of the validation set are employed for selecting the optimal quantitative model. Based on the RMSEP (root mean square error of prediction) of the validation set, the combination of model parameters exhibiting the best predictive results is chosen. The RMSEP is defined as

$$RMSEP = \sqrt{\frac{(\sum_{i=1}^{n} (y_{pred,i} - y_{obs,i})^2)}{n}} \quad (1)$$

where $y_{pred,i}$ is the y-value of object i as predicted by the model under investigation, while $y_{obs,i}$ is the reference value of y for object i and n is the number of objects for which $y_{pred,i}$ is obtained by prediction.

Furthermore, the average recovery and relative standard deviation are calculated for all samples. For each object i, the recovery is computed as:

$$Recovery = \left(\frac{y_{pred,i}}{y_{obs,i}}\right) \times 100\% \quad (2)$$

and expressed in percentages. The average recovery is hence determined for the calibration and validation set, next to the standard deviation on these values.

2.4. Calibration transfer

The objective of calibration transfer or standardisation is to avoid a full recalibration in case the instrumental response in the calibration step is different from the one in the prediction step, by applying chemometric techniques that allow to correct for instrumental or environmental differences [1–3]. Consequently, the constructed multivariate models are transferable between instruments or conditions.

Here, the challenge lies in making at-line developed calibration models for ingredient quantification applicable in a real-time in-line environment. The change in conditions between the calibration and prediction step is hence expressed by a change in sample presentation and surrounding circumstances, while the spectroscopic instrument itself is unaltered. While in the at-line set-up, the samples are stationary during spectral acquisition, liquid detergent compositions are passing by the Raman probe at a speed of tens of tonnes per hour in the in-line production environment. It is expected that this difference in sampling influences the spectral response. Earlier work by the authors and by Karande et al. has demonstrated that transformations in the spectral outcome are to be expected when comparing dynamic versus static sampling strategies [10,12].

2.4.1. Standardisation samples

A standardisation procedure generally consists of two steps. In first order, a set of well-chosen samples is measured under both conditions in order to estimate the differences in response between the calibration and the prediction step. These samples are often referred to as standardisation samples. Subsequently, the data of these standardisation samples are then employed to compute the standardisation parameters, via the transfer method of choice [1,15].

For selecting the standardisation samples, two factors have to be carefully considered, namely the representativity and the stability of the sample set [1,15]. In view of accurately estimating the differences between both measurement conditions and obtaining an efficacious calibration transfer, it is of utmost importance that the employed samples are representative of the future samples to which the computed standardisation parameters will be applied. Therefore, an often applied procedure is to pick a subset of the calibration set and to remeasure

those samples under the secondary conditions [1,10,14]. As it is physically infeasible to measure 100 mL laboratory samples in a tens of tonnes per hour production line, this selection strategy cannot be realised in this study. Therefore, the standardisation samples are selected from the prediction set, rather than the calibration set and remeasured under the primary conditions of the calibration step [1,15]. In our study, this strategy, proposed by Bouveresse and Massart, involves gathering validation samples from the continuous manufacturing line at production scale and collecting at-line Raman spectra in 100 mL laboratory containers from these samples. Comparing these at-line acquired spectra to their corresponding in-line data then allows to estimate the differences in spectral response function between both measurement conditions. In terms of stability, it is aimed to limit the time delay between in-line and at-line acquisition in order to assure that the samples themselves are physically and chemically identical under both conditions.

2.4.2. Slope/bias correction

Slope/bias correction is a straightforward, univariate standardisation strategy that does not correct the spectral data themselves, but only corrects the predictions from the untransformed spectra on the second instrument or under the secondary conditions [1,2]. In this study, this is achieved by plotting the uncorrected predictions of the in-line standardisation samples in function of the predictions of the at-line acquired spectra. Fig. 1 visualizes this for both PLS models, respectively predicting Surfactant 1 and Surfactant 2.

By assuming a linear relationship between both prediction sets, a linear equation can be fitted using ordinary least squares, as displayed in blue in Fig. 1. The slope and intercept of these curves can hence be calculated and the in-line predictions of each surfactant are then corrected via Eq. (3).

$$y_{In,T} = bias + slope.\, y_{In} \quad (3)$$

In this equation, y_{In} represents a concentration prediction from an in-line acquired Raman spectrum, while $y_{In,T}$ is the slope/bias corrected value of this prediction [1,2].

2.5. Method validation

Once the calibration models are constructed and transfer via slope/bias correction is pursued, the RMSEP, average recovery and standard deviation of the standardisation set are primarily compared before and after transfer to get a first impression of the predictive quality of the models. Herein, it is evaluated whether the RMSEP after correction is in the same order of magnitude as the one prior to transfer, if the average recovery values lies within 98% and 102% and if the relative standard deviation is below 5%.

Table 1
Statistical parameters for evaluating the transferred calibration models as proposed by Smith et al. [16,17].

Test/parameter	Condition
R^2	Similar magnitude to original calibration model
RMSEP	$RMSEP_{B,T} \leq 1.7 \times RMSEP_A$
Paired *t*-test versus reference data	No statistical difference ($\mu_{B,T} = \mu_R$)
Slope of transfer regression line	No statistical difference from 1 ($b_1 = 1$)
Intercept of transfer regression line	No statistical difference from 0 ($b_0 = 0$)
Bias	Not significant (bias < $bias_{max}$)
Linearity	Use of quadratic term does not improve fit (F-test)
Paired *t*-test versus at-line data	No statistical difference ($\mu_{B,T} = \mu_A$)
Coincidence test	Transfer and validation regression line statistically coincident

Secondly, a statistical evaluation of the transferred models is performed, based on the guidelines described by Smith et al. [16,17]. They have launched some statistical recommendations for proving successful transfer and for justifying the utilisation of a certain correction method. Table 1 lists the nine suggested statistical examinations for appraising whether the performed correction offers acceptably accurate predictive results. For a detailed description of the mathematics behind these statistical analyses, the authors would like to refer to the original work by Smith et al. [16]. Where applicable, $P = 0.05$ is used.

In their work, Smith et al. have performed these examinations on the calibration transfer of a near-infrared assay for tablets. Although both standardisation challenges exhibit many analogies, some adaptations to the guidelines are required in order to make them applicable to the Raman assay discussed in this manuscript. For instance, the paired *t*-test versus the reference data is discarded since no analytical data from a reference method are available for these samples. Furthermore, the second paired *t*-test, performed for assessing the statistical equivalence of the slope/bias corrected in-line data and the original at-line predictions, is altered based on the findings of Hartmann et al. [18].

Hartmann et al. have illustrated that for method validation purposes, the classically performed point hypothesis tests (as suggested in

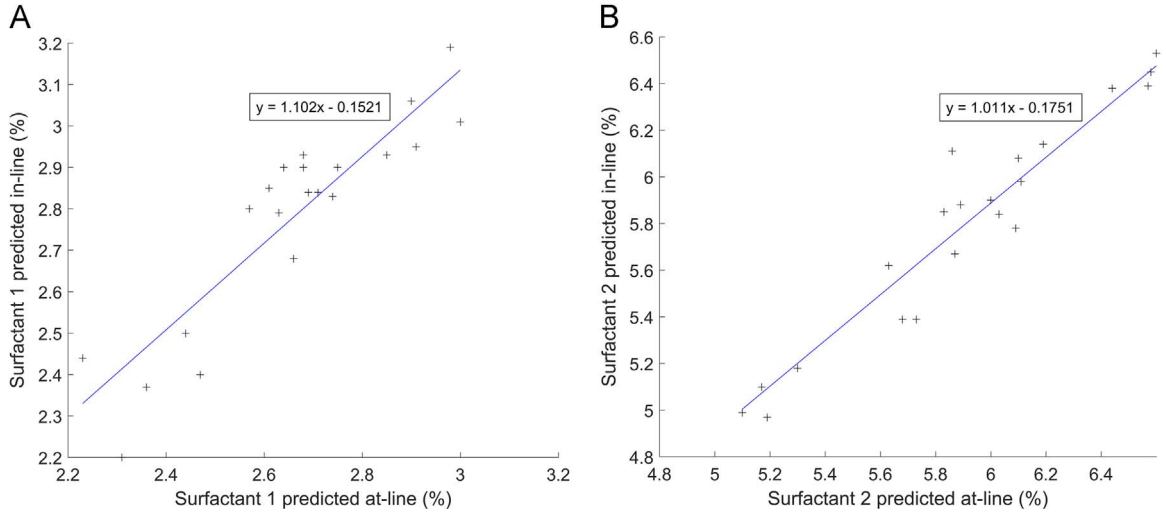

Fig. 1. Uncorrected predictions of the standardisation set under the secondary conditions (in-line) plotted in function of the predictions of the corresponding spectra under the primary conditions (at-line) by the calibration model of Surfactant 1 (plot A, left) and Surfactant 2 (plot B, right). The slope and intercept of the linear regression lines are determined to extract the parameters for the slope/bias correction of the in-line data.

Table 1) often provide insufficient information and therefore the definition of an interval of acceptable values is to be preferred. In view of testing the equivalence of two measurement procedures, one should fix the β-error and reverse the usual null and alternative hypotheses. This suggestion is made based on two disadvantages inherent to the conventionally applied point hypotheses. Firstly, it is extremely unlikely that the two methods under investigation should have no bias whatsoever. Therefore, the research question should not be whether the means are the same, but rather whether they do not differ more than a certain amount. Secondly, in method validation, the α-error is of inferior importance to the β-error. As concluding that there is bias, when in fact there is none (α-error) has less severe consequences than concluding that there is no bias when in fact there is (β-error), the β-error should be fixed rather than the α-error. Hartmann et al. have shown that statistical examination against acceptance limits benefits from limiting the β-error, allowing to focus on not incorrectly accepting a method that is largely biased. Furthermore, they have demonstrated that this significance testing at a fixed β-error imposes that the usual null and alternative hypotheses should be reversed.

Therefore, in the paired *t*-test suggested by Smith et al. for evaluating whether or not the corrected in-line data or statistically equivalent to the original at-line data, the null and alternative hypotheses are reversed. The risk of wrongly accepting that there is no bias is set at 0.05 and a bias smaller than or equal to 5% is considered acceptable [16,18].

3. Results and discussion

3.1. Calibration model development

Two PLS regression models are constructed based on the spectral data of the at-line measured calibration samples, predicting the concentration of Surfactant 1 and Surfactant 2 in the hard surface cleaners, respectively. For Surfactant 1, the spectral range from 313 until 3219 cm^{-1} is selected and EMSC (extended multiplicative scatter correction) pre-processing is selected, followed by first derivative (Savitzky-Golay, second order polynomial, window width = 15) calculation and mean centring. Eight latent variables are retained. Fig. 2 displays the observed versus predicted plot of this model. Herein, the concentration levels of Surfactant 1 as predicted by the Raman assay

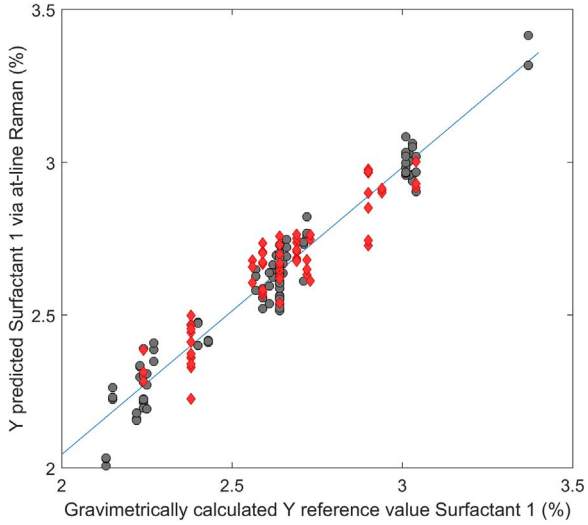

Fig. 2. Observed versus predicted plot showing the at-line PLS model predictions of Surfactant 1 in function of their gravimetrically acquired reference values for all samples in the calibration (grey circles) and validation (red diamonds) set. The blue line represents the fit to the calibration and validation samples. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

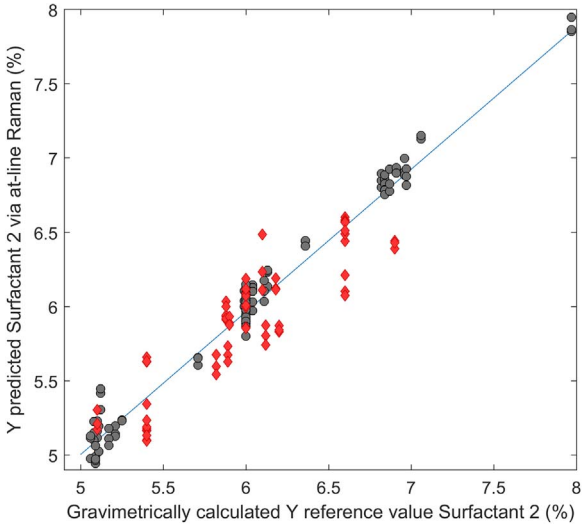

Fig. 3. Observed versus predicted plot showing the at-line PLS model predictions of Surfactant 2 in function of their gravimetrically acquired reference values for all samples in the calibration (grey circles) and validation (red diamonds) set. The blue line represents the fit to the calibration and validation samples. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

are visualised in function of their gravimetrically calculated reference values.

Analogously, Fig. 3 shows the observed versus predicted plot of the Surfactant 2 regression model. For this component, spectra are cropped in the range from 605 until 2962 cm^{-1}. The spectral data are again EMSC pre-processed, now followed by a baseline correction via automatic weighted least squares and mean centring. Ten latent variables are required in order to adequately capture the variance, which can be justified by the complexness of the liquid detergent composition and the high number of variation sources present within the dataset (cfr. different types of dye, perfume and surfactant).

Table 2 lists the calculated model statistics of these at-line calibration models. It is witnessed that both models meet the predefined acceptance values of an average recovery between 98% and 102% and a relative standard deviation below 5%.

3.2. Calibration transfer

3.2.1. Standardisation samples

The raw data of the selected standardisation samples are visualised in Fig. 4. While all spectra display well-aligned peak locations, dissimilarities in intensity are perceived between the in-line acquired Raman data and the at-line spectral data of the same samples. This is due to the changes in presentation (static versus dynamic) between both set-ups. Furthermore, intensity variations are observed within each dataset, which can be ascribed to the different dyes causing different degrees of fluorescence.

3.2.2. Slope/bias correction

The spectra of the standardisation samples, collected both in-line and at-line, are introduced in the PLS regression models described higher and their Surfactant 1 and Surfactant 2 concentration levels are predicted. For both constituents, the estimates from the in-line manufacturing line are plotted in function of their corresponding at-line predictions, as visualised in Fig. 1. By determining the slope and intercept of their linear regression lines, the in-line predictions are then corrected via Eq. (3).

Table 3 shows the RMSEP, average recovery and standard deviation of the standardisation samples before and after this univariate calibration transfer. From this table, one learns that the original at-line

Table 2
Model statistics of the at-line developed PLS regression models. The root mean square error of prediction (RMSEP) values of the validation set are shown for the estimates of Surfactant 1 and Surfactant 2, next to the average recovery and standard deviation (stdev) of the calibration and validation set predictions.

	Surfactant 1			Surfactant 2		
At-line models	**Stdev (%)**	**RMSEP (%)**	**Recovery (%)**	**Stdev (%)**	**RMSEP (%)**	**Recovery (%)**
Calibration set	/	100.1	2.549	/	100.0	1.582
Validation set	0.07657	100.7	2.868	0.2278	98.49	3.423

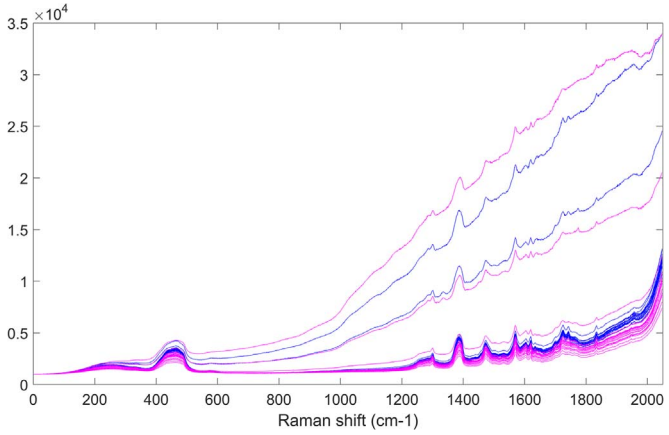

Fig. 4. Raw spectral data of the standardisation samples measured on both the at-line (blue) and in-line (magenta) set-up.

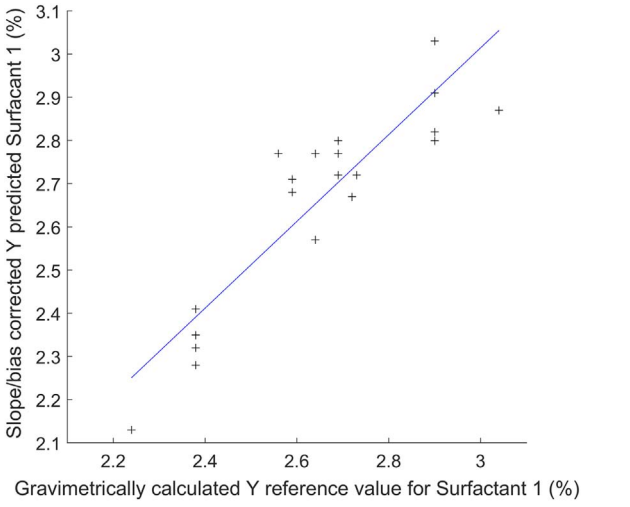

Fig. 5. Observed versus predicted plot showing the slope/bias corrected in-line PLS model predictions of Surfactant 1 in function of their gravimetrically calculated reference values for all standardisation samples.

calibration models yield satisfying prediction results for this sample set, meeting the pre-set acceptance criteria. Without any correction, the concentration predictions of the in-line acquired spectral data do not meet these specifications. For both quantification models, the average recovery falls outside the range from 98% to 102%. This indicates that a standardisation strategy is required to overcome the dissimilarities in spectral response observed between both measurement conditions. It is furthermore noticed that this deviation from the target can be corrected by applying a univariate slope/bias correction, resulting in average recovery values of 100.5% and 98.50% for Surfactant 1 and Surfactant 2, respectively. All models exhibit an acceptable relative standard deviation.

Figs. 5 and 6 show the corresponding observed versus predicted plots of these slope/bias corrected in-line predictions of the Surfactant 1 and Surfactant 2 concentrations, respectively. Their concentration levels after univariate standardisation are plotted in function of the gravimetrically acquired reference values for these surfactant quantities.

3.3. Method validation

While Table 3 already hints at a successful transfer from at-line to in-line quantification, an adequate statistical evaluation is required to justify the applicability of the performed transfer technique. The suggested statistical tests for assessing whether calibration transfer has been successful (Table 1) are applied to the slope/bias corrected

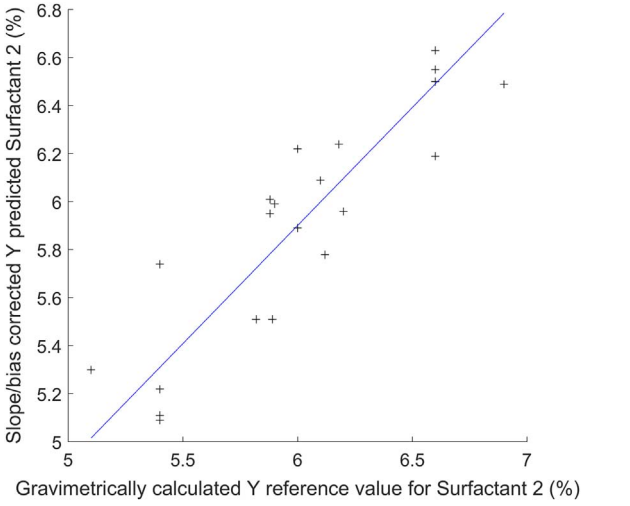

Fig. 6. Observed versus predicted plot showing the slope/bias corrected in-line PLS model predictions of Surfactant 2 in function of their gravimetrically calculated reference values for all standardisation samples.

Table 3
Model statistics of the PLS regression models before and after transfer. The root mean square error of prediction (RMSEP) values of the standardisation set are shown for the estimates of Surfactant 1 and Surfactant 2, next to the average recovery and standard deviation (stdev) of the these predictions.

	Surfactant 1			Surfactant 2		
Calibration model	**RMSEP (%)**	**Recovery (%)**	**Stdev (%)**	**RMSEP (%)**	**Recovery (%)**	**Stdev (%)**
Original at-line	0.07691	100.6	3.038	0.2232	98.51	3.436
Uncorrected in-line	0.1732	105.0	4.293	0.3036	96.61	3.845
In-line after SBC	0.1005	100.5	3.849	0.2407	98.50	3.855

Table 4
Statistical evaluation of the slope/bias corrected Surfactant 1 and Surfactant 2 calibration models.

	Transferred model	At-line model	Conclusion
Surfactant 1			
R^2	0.8024	0.8600	OK
RMSEP	0.1005	0.07691	OK
Slope	$b_1 = 1$	/	OK
Intercept	$b_0 = 0$	/	OK
Bias	bias < bias_{max}	/	OK
Linearity	$\sigma_A^2 \neq \sigma_B^2$	/	Not OK
Interval hypothesis	/	$\theta_1 < \mu_{B,T}/\mu_A < \theta_2$	OK
Coincidence	/	$\sigma_A^2 = \sigma_B^2$	OK
Surfactant 2			
R^2	0.7870	0.8174	OK
RMSEP	0.2407	0.2232	OK
Slope	$b_1 = 1$	/	OK
Intercept	$b_0 = 0$	/	OK
Bias	bias < bias_{max}	/	OK
Linearity	$\sigma_A^2 = \sigma_B^2$	/	OK
Interval hypothesis	/	$\theta_1 < \mu_{B,T}/\mu_A < \theta_2$	OK
Coincidence	/	$\sigma_A^2 = \sigma_B^2$	OK

standardisation set data in view of confirming the effectiveness of this transfer strategy in overcoming the dissimilarities in spectral response witnessed between the at-line and in-line acquired Raman data. Table 4 lists the results from this statistical validation for both the Surfactant 1 and Surfactant 2 models.

It is shown that all relevant criteria are met for the transfer of Surfactant 2, while Surfactant 1 successfully passes seven out of eight tests. It is hence confirmed that the calibration assay of Surfactant 2 is efficaciously transferred from the at-line to the in-line manufacturing set-up, allowing to accurately estimate the concentration levels of this component in the liquid detergent compositions in-line and in real-time. As the Surfactant 1 assay also yields positive results in the most discriminating tests, it is concluded that slope/bias correction can likewise be considered as an efficient correction method for this calibration model. As the linearity test implies that the use of a quadratic term is not found to significantly improve the fit, a failure in this guideline means that the slope/bias corrected model predictions would be better fit by a quadratic term. This is most plausibly due to the presence of some moderate outliers and the slight variation in predicted values at certain concentration levels. However, as the obtained recovery and standard deviation meet the pre-set specifications, this is not found to be alarming. Moreover, in their recommendations, Smith et al. have highlighted that the most discriminating tests are the paired *t*-tests and the other statistical tests suggested are not critical. As the surfactant 1 assay passes the paired *t*-test with the fixed β-error, as well as six of the other recommended evaluations, the failure of one test is considered insufficient to write the slope/bias correction off as being invaluable for standardisation of this specific model.

4. Conclusion

A strategy is suggested for transferring two multivariate calibration models from an at-line to in-line set-up. Firstly, two PLS regression models are developed by regressing the at-line acquired Raman spectra of a set of calibration samples against their reference concentration values. These models allow to quantify the concentration levels of two surfactants in a series of hard surface cleaners. Subsequently, the models are transferred to a continuous manufacturing line at production scale, where immersion of the Raman probe in the production pipe allows the in-line collection of spectral information. Predicting the concentration of both surfactants using the at-line calibration models is pursued. Therefore, a univariate slope/bias correction of the predicted values is required to overcome the dissimilarities in spectral response witnessed between both measurement conditions. An extensive statistical validation of the proposed quantification method is performed in view of assessing the truthfulness of the slope/bias corrected in-line predictions. This evaluation shows that both calibration models produce satisfactory accurate values. Since the transferred regression model of Surfactant 1 passes seven out eight statistical tests, it is concluded that the slope/bias correction of the in-line acquired results offers acceptably accurate predictions. Similarly, the concentration of Surfactant 2 can be predicted adequately in-line, since the statistical validation of this transferred model yields positive results for all eight performed evaluations. It is hence demonstrated that real-time information on critical quality attributes such as surfactant concentrations can be obtained in-line without the need for an extensive recalibration. Applying a straightforward univariate slope/bias correction to the at-line developed multivariate calibration models is proven to be sufficient to resolve the variations in spectral response observed between the altered measurement conditions.

References

[1] E. Bouveresse, D.L. Massart, Standardisation of near-infrared spectrometric instruments: a review, Vib. Spectrosc. 11 (1996) 3–15, http://dx.doi.org/10.1016/0924-2031(95)00055-0.

[2] R.N. Feudale, N.A. Woody, H. Tan, A.J. Myles, S.D. Brown, J. Ferré, Transfer of multivariate calibration models: a review, Chemom. Intell. Lab. Syst. 64 (2002) 181–192, http://dx.doi.org/10.1016/S0169-7439(02)00085-0.

[3] O.E. de Noord, Multivariate calibration standardization, Chemom. Intell. Lab. Syst. 25 (1994) 85–97, http://dx.doi.org/10.1016/0169-7439(94)85037-2.

[4] Y. Wang, D.J. Veltkamp, B.R. Kowalski, Multivariate instrument standardization, Anal. Chem. 63 (1991) 2750–2756, http://dx.doi.org/10.1021/ac00023a016.

[5] E.L. Bergman, H. Brage, M. Josefson, O. Svensson, A. Sparén, Transfer of NIR calibrations for pharmaceutical formulations between different instruments, J. Pharm. Biomed. Anal. 41 (2006) 89–98, http://dx.doi.org/10.1016/j.jpba.2005.10.042.

[6] H. Leion, S. Folestad, M. Josefson, A. Sparén, Evaluation of basic algorithms for transferring quantitative multivariate calibrations between scanning grating and FT NIR spectrometers, J. Pharm. Biomed. Anal. 37 (2005) 47–55, http://dx.doi.org/10.1016/j.jpba.2004.09.046.

[7] L. Salguero-Chaparro, B. Palagos, F. Peña-Rodríguez, J.M. Roger, Calibration transfer of intact olive NIR spectra between a pre-dispersive instrument and a portable spectrometer, Comput. Electron. Agric. 96 (2013) 202–208, http://dx.doi.org/10.1016/j.compag.2013.05.007.

[8] F. Sales, A. Rius, M.P. Callao, F.X. Rius, Standardization of a multivariate calibration model applied to the determination of chromium in tanning sewage, Talanta 52 (2000) 329–336, http://dx.doi.org/10.1016/S0039-9140(00)00366-0.

[9] R.S. Brito, H.M. Pinheiro, F. Ferreira, J.S. Matos, A. Pinheiro, N.D. Lourenço, Calibration transfer between a bench scanning and a submersible diode array spectrophotometer for in situ wastewater quality monitoring in sewer systems, Appl. Spectrosc. 70 (2016) 443–454, http://dx.doi.org/10.1177/0003702815626668.

[10] D. Brouckaert, J. Uyttersprot, W. Broeckx, T. De Beer, Calibration transfer of a Raman spectroscopic quantification method from at-line to in-line assessment of liquid detergent compositions, Anal. Chim. Acta 971 (2017) 14–25, http://dx.doi.org/10.1016/j.aca.2017.03.049.

[11] J.-S. Uyttersprot, W. Broeckx, D. Brouckaert, T. De Beer, Process for Making A Liquid Laundry Detergent Composition, European Patent no. 16162424.2-1375, 2016.

[12] A.D. Karande, C.V. Liew, P.W.S. Heng, Calibration sampling paradox in near infrared spectroscopy: a case study of multi-component powder blend, Int. J. Pharm. 395 (2010) 91–97, http://dx.doi.org/10.1016/j.ijpharm.2010.05.011.

[13] M.S. Larrechi, M.P. Callao, Strategy for introducing NIR spectroscopy and multivariate calibration techniques in industry, TrAC - Trends Anal. Chem. 22 (2003) 634–640, http://dx.doi.org/10.1016/S0165-9936(03)01005-7.

[14] D. Brouckaert, J. Uyttersprot, W. Broeckx, T. De Beer, Calibration transfer of a Raman spectroscopic quantification method for the assessment of liquid detergent compositions between two at- line instruments installed at two liquid detergent production plants, Anal. Chim. Acta (2017) 1–18, http://dx.doi.org/10.1016/j.aca.2017.07.044.

[15] E. Bouveresse, D.L. Massart, P. Dardenne, Calibration transfer across near-infrared spectrometric instruments using Shenk's algorithm: effects of different standardisation samples, Anal. Chim. Acta 297 (1994) 405–416, http://dx.doi.org/10.1016/0003-2670(94)00237-1.

[16] M.R. Smith, R.D. Jee, A.C. Moffat, The transfer between instruments of a reflectance near-infrared assay for paracetamol in intact tablets, Analyst 127 (2002) 1682–1692, http://dx.doi.org/10.1039/b209269g.

[17] M.R. Smith, R.D. Jee, A.C. Moffat, D.R. Rees, N.W. Broad, A procedure for calibration transfer between near-infrared instruments – a worked example using a transmittance single tablet assay for piroxicam in intact tablets, Analyst 129 (2004) 806–816, http://dx.doi.org/10.1039/b401267d.

[18] C. Hartmann, J. Smeyers-Verbeke, W. Penninckx, Y. Vander Heyden, P. Vankeerberghen, D.L. Massart, Reappraisal of hypothesis testing for method validation: detection of systematic error by comparing the means of two methods or of two laboratories, Anal. Chem. 67 (1995) 4491–4499, http://dx.doi.org/10.1021/ac00120a011.