Title: **Calibration transfer for bioprocess Raman monitoring using Kennard Stone Piecewise Direct Standardization and multivariate algorithms**.

Authors: Laure Pétillot[1], Fiona Pewny[2], Martin Wolf[2], Célia Sanchez[1], Fabrice Thomas[1], Johan Sarrazin[1], Katharina Fauland[2], Herman Katinger[2], Charlotte Javalet[1] and Christophe Bonneville[1*].

[1] RESOLUTION Spectra Systems, Meylan, France. [2] Polymun Scientific, Klosterneuburg, Austria.

*Corresponding author.

Corresponding author information:

Christophe Bonneville

13, chemin du Vieux Chêne

38240 Meylan, France

+33 (0)4 58 00 12 49

Christophe.Bonneville@resolutionspectra.com

Abstract: In the biopharmaceutical industry, Raman spectroscopy is now a proven PAT tool that enables in-line simultaneous monitoring of several CPPs and CQAs in real-time. However, as Raman monitoring requires multivariate modeling, variabilities unknown by models can impact the monitoring prediction accuracy. With the widespread use of Raman PAT tools, it is necessary to fix instrumental variability impacts, encountered for instance during a device replacement. In this work, we investigated the impact of instrumental variability between probes inside a multi-channel analyzer and between two analyzers, and explored solutions to correct them on model prediction errors in cell cultures. We found that the Kennard Stone Piecewise Direct Standardization (KS PDS) method enables to lower model prediction errors and that only one batch with the unknown device in the calibration dataset was sufficient to correct the prediction gap induced by instrumental variability. As a matter of fact, during device replacement a first cell culture monitoring can be performed with the KS PDS method. Then, the new data obtained can be inserted in the calibration dataset to integrate instrumental variability in the chemometric model. This methodology provides good multivariate calibration model prediction errors throughout the instrumental changes.

Keywords: Raman Spectroscopy, Cell Culture Monitoring, Model Transferability, Multivariate Model Maintenance, Process Analytical Technology.

Main text:

1 INTRODUCTION

The monitoring of mammalian cell cultures with Raman spectroscopy and chemometric tools has been well demonstrated and documented (Abu-Absi et al., 2011; Berry et al., 2016). The main metabolites and nutrients of a cell culture can be predicted by Partial Least Square (PLS) models with Root Mean Square Error of Prediction (RMSEP) quite close to off-line reference measurement accuracy (Mehdizadeh et al., 2015). This real-time monitoring of Critical Process Parameters (CPPs) has been used to implement a glucose feeding control loop, leading to improved productivity (Craven, Whelan, & Glennon, 2014). Several Critical Quality Attributes (CQAs) of a cell culture have also been successfully predicted with Raman spectroscopy: protein titer (André et al., 2015), glycosylation (Li et al., 2018) and aggregation (Ettah & Ashton, 2018). These results demonstrate that Raman spectroscopy can be efficiently used to monitor cell cultures in real-time and *in situ*, automate processes and even open the door to the use of Raman spectroscopy for Real-Time Release (RTR) of batches (Swann, Brophy, Strachan, Lilly, & Jeffers, 2017). Another approach in the field of Raman monitoring of bioprocesses has consisted in the attempt of developing generic models (Webster, Hadley, Hilliard, Jaques, & Mason, 2018). This experiment shows that models based on wide process variability provide poor accuracy with regards to models build with and for a limited design space, as defined by the Quality-by-Design (QbD) rules. An interesting way to proceed may be to find a methodology to select a dataset consistent with a given process inside a large dataset, as presented by Rowland-Jones et al. (2017). Recently, Tulsyan et al. (2019) have proposed a novel machine-learning procedure based on Just-In-Time Learning (JITL) to calibrate Raman models. However, being able to use an existing multivariate model on different hardware configurations is probably a first priority, because generic models and generic datasets may have no use if they cannot be exploited on a variety of hardware units with at least the same design.

Most of these mentioned works are based on model building and prediction based on the same Raman analyzer hardware. Then, they leave aside a key issue when using Raman spectroscopy in biopharmaceutical environments: instrumental variability, including replacements or changes of hardware. Instrumental variability over time on the same Raman analyzer is one issue which can be overcome through regular instrumental re-calibration or multivariate model updates. But the case of instrumental changes needs further attention, since this is related to more radical variabilities, with sometimes operational unforeseen constraints, and given that it has not been much

documented in the engineering literature on Raman monitoring. The same question applies to other transfers: the transfer of models from one device to another and the transfer of model from one hardware configuration to another. This question is crucial since generating a dataset suitable to multivariate models represents a significant effort corresponding to typically three to more than ten cell culture runs. For example, the calibration dataset may be generated on several channels in parallel and the models built to be used on each of the channels, or models may be built on an R&D analyzer and then transferred to a pilot plant analyzer, or the tube of a Raman probe may be changed because its lifetime has expired, the same for the excitation laser, etc.

This topic is actually fully part of the multivariate model maintenance issue. From a chemometric perspective, this question is equivalent to taking into account the principal components related to instrumental variability in the way of managing the model space, which can be enlarged or translated for this purpose. Wise and Roginsky (2015) have proposed a quite comprehensive roadmap to maintain multivariate models. In this paper, we propose similar tools in the perspective of instrumental variability which are not only unexpected problems but anticipated steps in the use of Raman monitoring in a biopharmaceutical environment. The ambition of the present study is limited to calibration transfers between different hardware units with the same design.

It is worth noticing that we have made the choice to limit the exploration to PLS models, as a basic multivariate tool. Indeed, PLS is the most common one and it generally provides very good results in most of the bioprocessing applications (Buckley & Ryder, 2017). On top of that, well documented mathematical tools, such as Piecewise Direct Standardization (PDS) (Bouveresse & Massart, 1996), have been used to perform calibration transfers. The comparison has been made with another conventional chemometric method consisting in the enrichment of the calibration dataset with data from the new Raman system configuration performed on the same process.

This work is specific to the monitoring of cell cultures. Indeed, the goal of this study is to assess a methodology to lower the impact of instrumental variability on monitoring accuracy and robustness, so that it is negligible with regards to process variability itself. In the case of cell cultures, the biological and setup variability is for example much more significant than in a pure chemical process.

Finally, this paper explores two cases which have been considered crucial: the first one is related to the use of a four-channel Raman system to efficiently establish first multivariate models to be used with these same

channels (inter-probe transferability study). The studied methodology aims to accelerate the implementation of Raman monitoring in a process development context, or the transfer of a process in new bioreactors. The second case deals with multivariate calibration transfer from one analyzer, where models show acceptable performances, to a second analyzer (inter-analyzer transferability study). This is typically the case when a process is transferred from process development to pilot plant to production or when an analyzer needs to be replaced for maintenance.

2 MATERIALS AND METHODS

2.1. Cell culture methods, sample collection and offline analysis

For the inter-probe transferability study: the CHO-K1 cells expressing monoclonal antibody IgG1 (ECACC 85051005) were inoculated in four DASGIP® parallel bioreactor systems (Eppendorf) in 1 L glass vessels at 0.5 x 10^6 cells/mL in ActiCHO™ P medium (GE Healthcare) supplemented with 8 mM of glutamine, 1.8 g/L $NaHCO_3$ and approximatively 30 mM of NaOH. The cultivation conditions were set at 37 °C, pH 7.0 and pO_2 30% regulated by sparging CO_2 for pH and a mix of air, O_2, CO_2 and N_2 for pO_2. The foam level was regulated by 0.3% of Dow Corning® Antifoam C (DuPont™) based on visual inspection. The cells were agitated with a pitched blade impeller at 90 rpm. The aeration of the culture was performed with a sparger with a flow rate up to 0.4 vvm. Starting on day 3 or 4, cultures were fed daily with ActiCHO Feed A and Feed B medium (GE Healthcare). Feeds were added to a calculated glucose concentration of 4.5 to 10 g/L for Feed A and 0.28% v/v for Feed B. Alternatively spikes of glucose, glutamine and glutamic acid were performed on one batch. In addition, for some batches, temperature shifts down to 33 °C were performed. Samples were collected once a day when no feed was performed. For each feed, a sample was taken before and after the feed instead of the daily samples. Samples were analyzed with High-Performance Liquid Chromatography (HPLC, Agilent 1200 Series) for quantitation of glucose and lactate, with the BioProfile® 100 Plus (Nova Biomedical) for quantitation of ammonium, with the BioProfile® CDV (Nova Biomedical) for Viable Cell Density (VCD) and with the Multisizer™ 4 Coulter Counter® (Beckman Coulter) for quantitation of Total Cell Density (TCD).

For the inter-analyzer transferability study: the FreeStyle™ CHO-S (Gibco™) cells were inoculated in a 3 L glass vessel with the BioFLO®320 (Eppendorf) at 0.4 x 10^6 cells/mL in CD-CHO medium (Gibco™) supplemented with 8 mM of glutamine, 1‰ of Anti-Clumping Agent (Gibco™) and 0.5% of Penicillin/Streptomycin. The cultivation conditions were set at 37 °C, pH 7.0 and pO_2 40% regulated by sparging CO_2 and 0.5N NaOH for pH and a mix of air

and O_2 for pO_2. The cells were agitated with a pitched blade impeller at 80 rpm. The aeration of the culture was performed with a ring sparger with a flow rate up to 0.1 vvm. The culture was fed with 15% v/v on day 0 and 10% v/v every three days with EfficientFeed™ B (Gibco™). Glutamine was added when under 4 mM and in addition, a constant glutamine feed starts on day three. In addition to the EfficientFeed™ B, glucose was added when under 4 g/L. Samples were collected twice a day when no feed was performed. For each feed, a sample was taken before and after the feed in addition to the daily sample. All samples were taken in triplicate and each triplicate was analyzed with FLEX2 (Nova Biomedical) for quantitation of: VCD, TCD, glucose, glutamic acid, ammonium and lactate.

2.2. Raman spectral data collection

To perform these experiments, all spectra acquired in cell cultures have been performed *in situ* by ProCellics™ in-line analyzer. This solution is a Raman analyzer, compliant with Good Manufacturing Practices (GMP), using a 785 nm excitation laser source with an approximate power of 350 mW at the probe tip; it is associated with a high-sensitivity spectrometer with about 14 cm^{-1} average spectral resolution and 3 cm^{-1} sampling step in the +150 – +4,000 cm^{-1} Raman shifts bandwidth (Stokes signal), using a back-thinned charge coupled device (CCD) detector operating at -10 °C. This large spectral bandwidth up to 4,000 cm^{-1} Raman shifts gives a strategic full access to major O-H contributions from the aqueous media. The Raman spectra were acquired *in-situ* using 316 L stainless-steel optical ProCellics™ probes with sapphire windows. The probes were directly immersed into the bioreactors using PG13.5 cable glands adaptors before autoclave sterilization. To avoid disturbing the optical measurement, an appropriate solution has been found in order to isolate the medium from external straylight (daylight and artificial light), using either an opaque double-layer of thick aluminum foil around the vessel or a light proof fabric (Thorlabs). ProCellics™ analyzer can also be coupled with a Multi-Channel Unit (MCU) which is an add-on with the same compacity as the ProCellics™ main base unit that allows monitoring up to four probes with the same ProCellics™ analyzer. ProCellics™ analyzer and the spectral data collection were controlled by ProCellics™ Software using an Ethernet connection between the instrument and the computer that allows remote and network control. For the study of the inter-probe transferability using ProCellics™ Multi-Channel Unit, an exposure time of 47 seconds and averaging of 25 spectra were used for each 20-minute acquisition. For the study of the inter-analyzer transferability based on single-channel configuration, an exposure time of 45 seconds and averaging of 20 spectra were used for

each 15-minute acquisition. The sampling interval for the Raman spectra was set every 30 minutes in a single-channel configuration and continuously for the 4-probe Multi-Channel Unit sequential acquisitions. All the data and metadata (user actions, timestamps, instrument information, etc.) were recorded by ProCellics™ Software in a secure SQL-database on the computer. Cosmic ray removal and automatic dark subtraction were directly operated in ProCellics™ Software during the acquisition. Deviant measurements were automatically sorted before the spectra coaddition using a 3-sigma rejection algorithm. Furthermore, ProCellics™ Software integrates configurable preprocessing (normalization, Savitzky-Golay filter, spectral selection) and chemometric pre-analysis (based on Principal Components Analysis (PCA)) solution tools in a model building module, and is running SIMCA-Q engine (Sartorius Stedim Biotech) for real-time quantitation in a monitoring module.

Instrumental variability may come from different hardware elements of the analyzers, from their lasers through their probes to their spectrometers, typically in function of each laser power, each probe spectral transmission (with filters, collimating optics, and tube lengths) or each spectrometer sensitivity and resolution. Most of these variabilities must be taken into account by factory calibration. For each ProCellics™ analyzer, all raw data are thus calibrated in Raman intensities and Raman shift scales, using the same references and protocols compliant with ASTM standards (ASTM, 2014). This factory calibration procedure must enable the inter-comparison of Raman spectra acquired from several instruments, and from the same instrument associated to a Multi-Channel Unit and with different probes.

2.3. Instrument standardization calibration transfer

To enhance the calibration and inter-comparison of Raman spectra acquired from different probes or analyzers, the Piecewise Direct Standardization (PDS) method was used (Bouveresse & Massart, 1996). The PDS method allows transferring a set of data from an instrument (called slave instrument) to another (called master instrument) by correcting the differences induced by the sensors. First of all, a collection of spectra from the same experiment with the master instrument (${}^{m}T$) and with the slave instrument (${}^{s}T$) are acquired. From these ${}^{m}T$ data, a consistent subset of spectra (${}_{sub}^{m}T$) can be selected from the data collection. The selection is performed by applying a Principal Component Analysis (PCA) over the data collection followed by the Kennard Stone algorithm in the PCA space to uniformly cover the whole dataset. The same data subset acquired with the slave instrument (${}_{sub}^{s}T$) is then retrieved from ${}^{s}T$. The next step is to apply the PDS algorithm which consists in a multivariate

Principal Component Regression (PCR) on the subsets ${}_{sub}^{m}T$ and ${}_{sub}^{s}T$ for a given range of wavelengths. From each wavelength index ($i = 0, .., n$) of ${}_{sub}^{m}T$, a corresponding window (w^i) in ${}_{sub}^{m}T$ is used to compute the PCR, $w^i = {}_{sub}^{m}T\ [i - k; i + k]$ where w^i has a size of $2k + 1$ wavelengths and the PCR is solved by: ${}_{sub}^{m}T(i) = w^i b(i) + err(i)$. The resulting regression coefficients $b(i)$ are inserted in a banded diagonal transfer matrix (${}^{f}T$). Finally, ${}^{f}T$ is used to transfer all the spectra from the slave instrument in the master instrument: ${}^{s}\hat{T} = {}^{s}T{}^{f}T$, where ${}^{s}\hat{T}$ is the spectra data collection of the slave instrument transferred in the master instrument. In this study, KS PDS standardization was applied to the 300 – 3,900 cm^{-1} Raman shifts range, with a window size equal to 1 (k = 0) since the instrumental factory calibration allowed to achieve an accuracy better than ± 2 cm^{-1} in Raman shifts which was lower than the spectral resolution sampling.

2.4. Spectral preprocessing and multivariate modeling techniques

Preprocessing steps were performed with ProCellics™ Software (RESOLUTION Spectra Systems). Bands appearing on the spectra can be easily linked to chemical structures ("fingerprints"). The spectral regions can thus be selected according to their interest in the creation of the model. In the present typical cell culture cases, the spectral regions of interest were between 350 cm^{-1} and 1,775 cm^{-1} and between 2,800 cm^{-1} and 3,000 cm^{-1}. In addition to the selection of the spectral values, two treatments were chosen: a customized Standard Normal Variate (SNV) on the Region Of Interest (ROI) of water and the first derivative according to the Savitzky-Golay (SG) algorithm (Savitzky, A., & Golay, 1964). Performing customized SNV on water signal (between 3,100 cm^{-1} and 3,600 cm^{-1}), which is the only invariable element between different batches and conditions, enabled to normalize all spectra according to an invariant element. The goal of the derivative (derivative order 1, step 15 cm^{-1}, polynomial order 2) was to numerically increase the apparent resolution of the spectra. The data processes carried out enabled a decrease of the fluorescence impact, and an increase of spectral differences was noticed. PCA is an unsupervised data transformation procedure of complex datasets (Jackson, 1991). PCA has been used as first analysis of the data in order to see any trends, dispersion of the data or bundles. In addition to the spectrometric calibration, a PLS regression was applied to monitor and control industrial processes (Wold, Sjöström, & Eriksson, 2001). Multivariate data analyses were performed on SIMCA 15 Users Software (Sartorius Stedim Biotech) and PLS modeling was used to build linear models that specify the relationship between the observed dependent variables (Y) and some predictable variables (X). The (X) matrix included the spectral range under study and the (Y) matrix was composed by

the molecule concentrations measured off-line (glucose, lactate, etc.). After developing the reduced model, the omitted data were used as a test set, and the differences between actual and predicted Y-values were calculated for these data points. A popular parameter to interpret the performances of the models is the RMSEP. It is computed as $RMSEP = \sqrt{\frac{\sum(Yobs-Ypred)^2}{(N)}}$, where *Yobs-Ypred* refers to the predicted residuals for the observations in the predictionset. RMSEP was used to measure the predictive power of the model.

3 RESULTS AND DISCUSSION

3.1. Inter-probe transferability inside a Raman multi-channel analyzer

Instrumental variability between probes inside a multi-channel analyzer (ProCellics™ MCU) has been studied. The study consisted in comparing two calibration model strategies, given that data were collected by all probes: a global chemometric calibration model integrating the inter-probe variability in the model and a per channel chemometric calibration model. Then, model prediction errors obtained were compared to determine the best calibration model strategy. For global chemometric calibration models, off-line data from eight cell culture runs, performed during two process sessions of four batches over the four channels, were combined with their respective Raman spectra. These latter were used to produce calibration models for glucose, TCD, VCD, lactate and ammonium.

TABLE 1: Chemometric calibration and results for channel transfer: global models and channel models. PLS, Partial Least Square; Cum, cumulated; TCD, Total Cell Density; VCD, Viable Cell Density; RMSEP, Root Mean Square Error of Prediction; A, Latent variables; N, Number of points.

PLS model calibration summary				
	A	N	R^2Y (cum)	Q^2 (cum)
Global models				
Glucose	7	251	0.961	0.935
TCD	9	137	0.967	0.915
VCD	9	128	0.970	0.930
Lactate	10	250	0.972	0.943
Ammonium	8	256	0.753	0.617
Models per channel				
Channel 1				
Glucose	6	79	0.980	0.961
TCD	4	38	0.979	0.966
VCD	4	38	0.963	0.935
Lactate	8	77	0.971	0.859
Ammonium	8	79	0.966	0.885
Channel 2				
Glucose	7	54	0.980	0.899
TCD	4	33	0.953	0.930
VCD	5	33	0.925	0.754
Lactate	7	56	0.981	0.952
Ammonium	10	59	0.966	0.721
Channel 3				
Glucose	6	64	0.976	0.900
TCD	3	35	0.957	0.894
VCD	3	35	0.946	0.834
Lactate	7	62	0.991	0.959
Ammonium	6	65	0.932	0.797

PLS model validation results					
		Global models		Models per channel	
	Range	RMSEP	Max value % error	Global RMSEP	Max value % error
Glucose (g/L)	0 - 6.30	0.48	7.5	0.49	7.7
TCD (x10^6 cells/mL)	0.5 - 31.6	2.20	6.9	3.73	11.8
VCD (x10^6 cells/mL)	0.5 - 31.0	3.30	10	4.82	15.5
Lactate (g/L)	0 - 5.22	0.46	8.8	0.47	9
Ammonium (mM)	0.7 - 16.7	2.03	11.9	1.92	11.5

As shown in Table 1, the models have high explained ($R^2Y > 0.90$) and predicted ($Q^2 > 0.90$) performances for all parameters, except for ammonium, probably because the principal molecular bounds of NH_4 are found in several elements and consequently difficult to decorrelate. Then, models were validated using new data, not included in the

calibration models from a third session of three batches over three channels, and performed under similar process conditions than the ones used for the calibration batches. For chemometric calibration models per channel, for each of three channels, off-line data from two cell culture runs performed during two process sessions, with one batch from both sessions per channel, were combined with their respective Raman spectra. These latter were used to produce calibration models for glucose, VCD, TCD, lactate and ammonium. As shown in Table 1, the models have high explained ($R^2Y > 0.90$) and predicted ($Q^2 > 0.80$) performances for all parameters, except for ammonium, probably for the same reason as in global models. Then, models were validated using new data, not included in the calibration models, from a third session of three batches on three channels, performed under similar process conditions as the ones used for the calibration batches. To compare the effect of both strategies on model prediction errors, the maximum value percentage errors of RMSEPs were calculated. Validation results showed that models for glucose, TCD, VCD and lactate performed with a global dataset presented results with higher prediction accuracy than the results obtained with the channel datasets (Table 1). These results may be explained by the huge difference in the dataset size (x4 for global model datasets). In addition, global calibration models presented low model prediction errors, less than or equal to 10% for glucose, TCD, VCD, lactate and less than 12% for ammonium (Table 1). As shown in Figure 1, predicted kinetics were highly reliable for each component among the three batches.

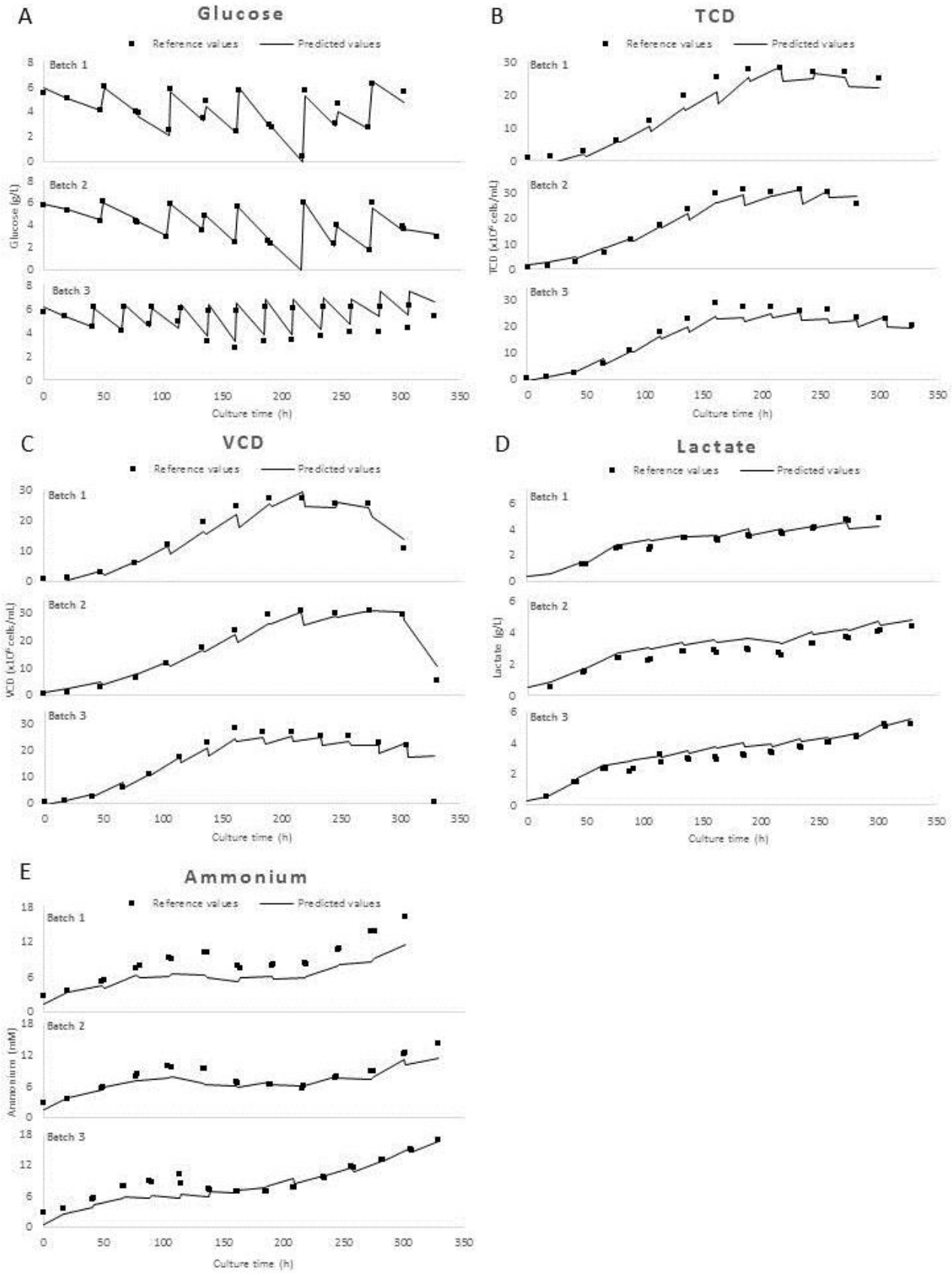

FIGURE 1: PLS model prediction results from a calibration dataset performed with the channel 1, 2, 3, 4 samples and from four calibration datasets performed with each channel sample, predicting three validation batches for (A) glucose, (B) TCD, (C) VCD, (D) lactate and (E) ammonium. TCD, Total Cell Density; VCD, Viable Cell Density.

After studying instrumental variability between probes with data collected by all probes, we investigated instrumental transferability by predicting data collected with one probe unknown by the models. In other words, data used in the model building step were collected with three of the four probes and predicted data were collected with the fourth probe. Consequently, the instrumental variability between probes, highlighted by the PCA (Figure 2A), could not be included in the model for the fourth probe.

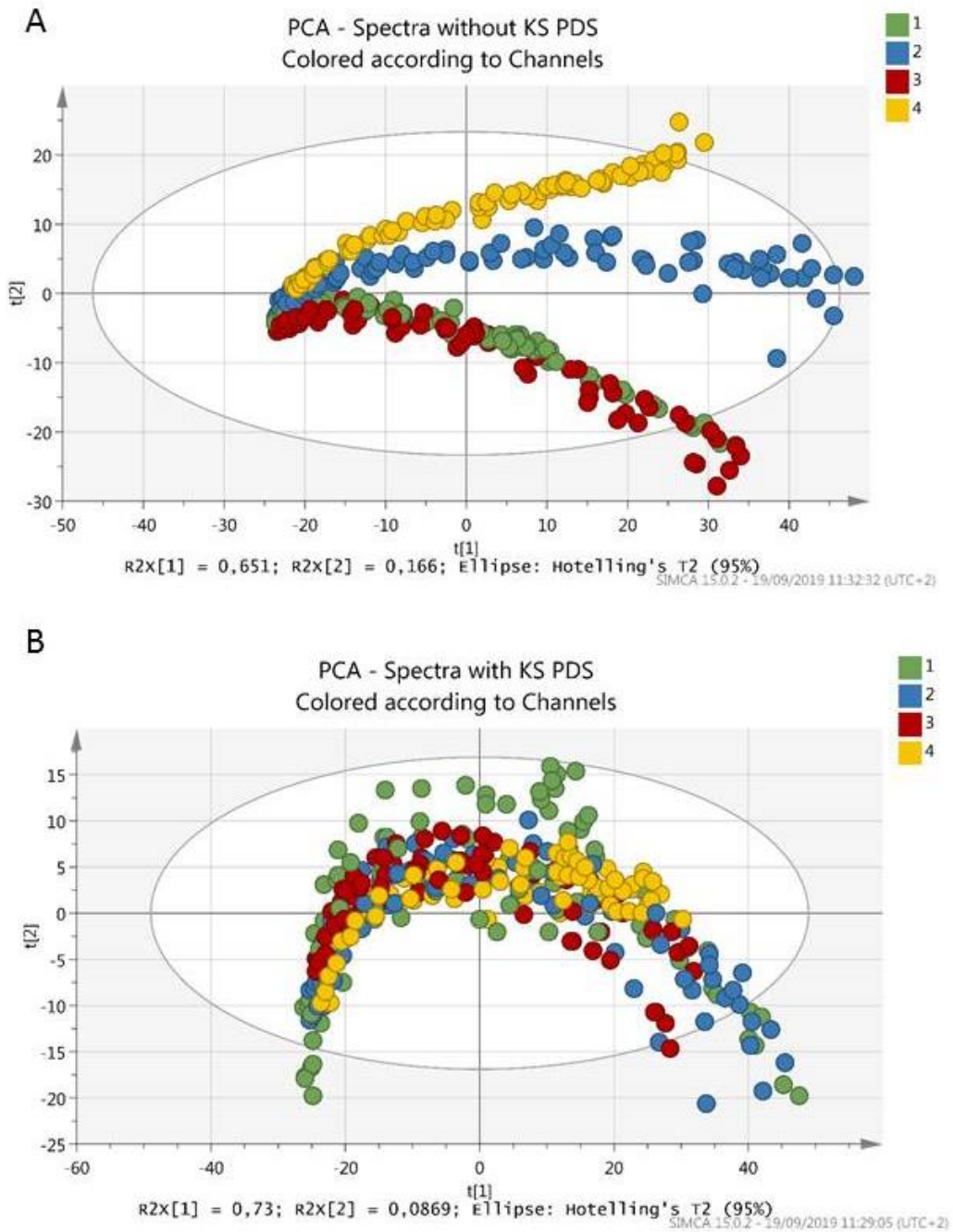

FIGURE 2: PCA demonstrating instrumental variability between the four channels and its decrease due to the Kennard-Stone PDS instrumental transfer calibration. PCA, Principal Component Analysis; KS PDS, Kennard-Stone Piecewise Direct Standardization.

To correct it and maintain the model prediction accuracy, the Kennard-Stone Piecewise Direct Standardization (KS PDS) method was tested. Off-line data from nine cell culture runs were combined with their respective Raman spectra obtained with the channels 1, 2 and 3 of the ProCellics™ MCU and used to produce global calibration models for glucose, TCD, VCD, lactate and ammonium (Table 2).

TABLE 2: Instrumental calibration, chemometric models and results for channel transfer: master batches analyzed with channel 1 and slave batches analyzed with channels 2, 3, 4 without PDS or processed with a 10-point Kennard-Stone PDS transfer matrix, predicting validation batches analyzed with channel 4. PLS, Partial Least Square; Cum, cumulated; PDS, Piecewise Direct Standardization; TCD, Total Cell Density; VCD, Viable Cell Density; RMSEP, Root Mean Square Error of Prediction; A, Latent variables; N, Number of points.

PLS model calibration summary				
	A	N	R^2Y (cum)	Q^2 (cum)
Without PDS				
Glucose	8	271	0.969	0.950
TCD	6	150	0.952	0.933
VCD	7	141	0.966	0.939
Lactate	10	267	0.970	0.945
Ammonium	8	269	0.838	0.775
With PDS				
Glucose	7	237	0.937	0.881
TCD	7	137	0.971	0.941
VCD	7	132	0.976	0.932
Lactate	9	240	0.952	0.881
Ammonium	10	242	0.821	0.682

PLS model validation results					
		Without PDS		With PDS	
	Range	Global RMSEP	Max value % error	Global RMSEP	Max value % error
Glucose (g/L)	0.04 - 6.70	0.40	5.9	0.64	9.5
TCD (x10^6 cells/mL)	0.5 - 34.1	6.38	18.7	3.51	10.2
VCD (x10^6 cells/mL)	0.5 - 33.5	6.51	19.4	3.48	10.3
Lactate (g/L)	0 - 4.33	0.50	12.7	0.45	10.3
Ammonium (mM)	0.7 - 15.06	2.71	17.9	1.61	10.6

Two different instrumental calibrations were applied: the classical instrumental calibration and the particular instrumental calibration with a 10-point KS PDS where channel 1 was considered the master and channel 2, 3 and 4 were considered the slaves.

As shown on the PCA (Figure 2B), the 10-point KS PDS instrumental calibration allowed to drastically reduce the instrumental variability present between the different channels (Figure 2A). Then, for both instrumental calibrations, the models have high explained ($R^2Y > 0.90$) and predicted ($Q^2 > 0.90$) performances for all parameters (except for ammonium). Models were validated using new data from three batches with Raman spectra acquired with the channel 4 of ProCellics™ MCU and performed under similar process conditions as the ones used for the calibration batches. To test the effect of the KS PDS method on model prediction errors, the maximum value percentage errors of RMSEPs were calculated for both calibration methods. Validation results showed that the model for glucose was very few impacted by instrumental variability, probably because glucose is the most defined and characterized compound in the medium and then, the KS PDS method did not improve the prediction accuracy. In both cases, with and without the KS PDS, glucose models presented less than 10% of prediction errors (Table 2 and Figure 3A). However, we found that models for TCD, VCD, lactate and ammonium performed with the PDS calibrated dataset presented results with a higher prediction accuracy (around 10% of prediction errors) than the results obtained with the classical calibrated dataset (Table 2 and Figure 3B-E).

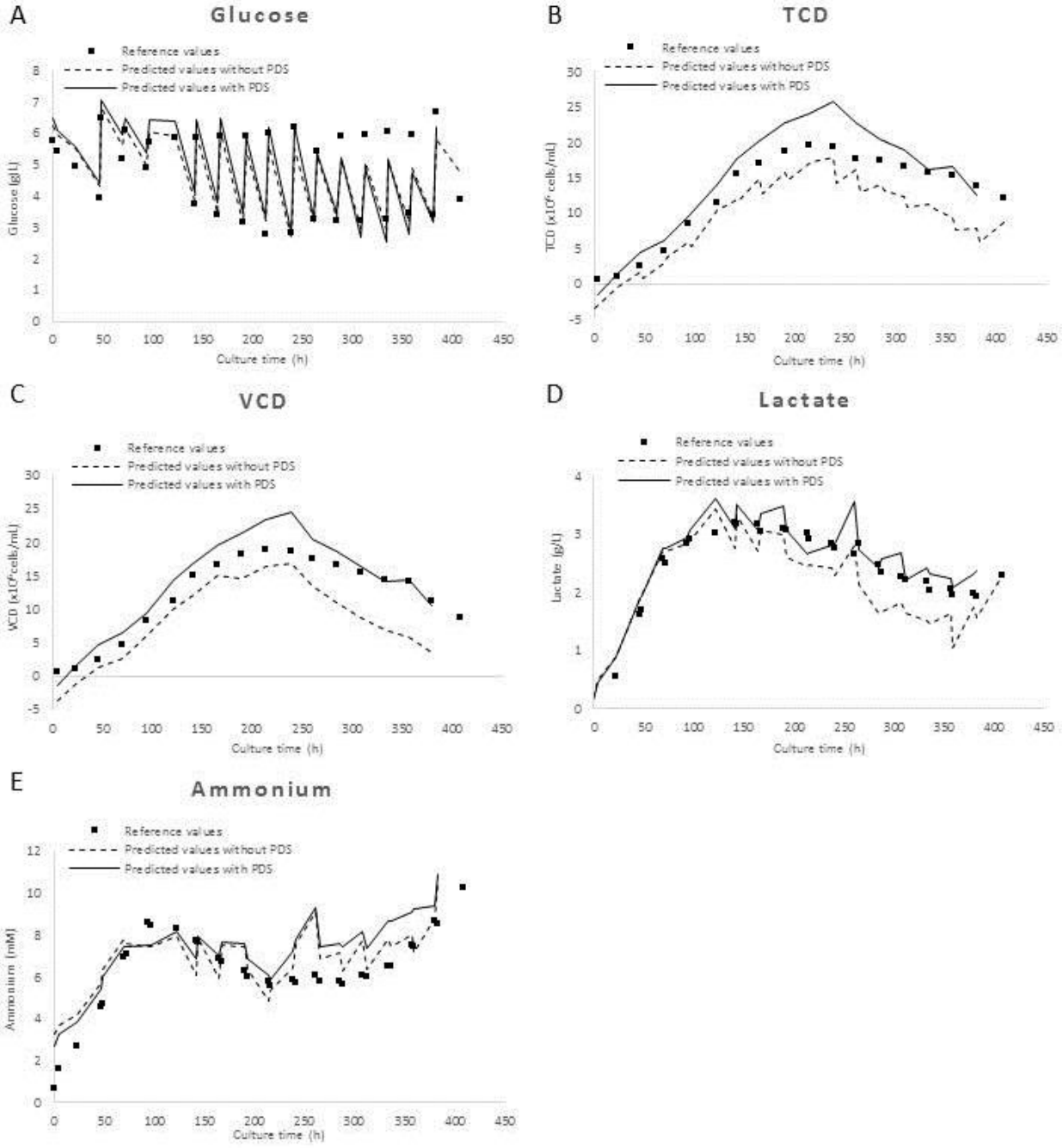

FIGURE 3: PLS model prediction results from Kennard-Stone PDS calibration datasets (without PDS or processed with PDS transfer matrix) predicting validation batches analyzed with channel 4 for (A) glucose, (B) TCD, (C) VCD, (D) lactate and (E) ammonium. TCD, Total Cell Density; VCD, Viable Cell Density; PDS, Piecewise Direct Standardization.

3.2. Inter-instrument transferability between two analyzers (device to device)

In the second part, we investigated the impact of instrumental variability between two analyzers. First, two CHO cell culture runs were performed with two ProCellics™ analyzers at the same time to test this variability: for

each cell culture, the probe from each analyzer was simultaneously placed in the same bioreactor. Due to the absence of biological variability and position of both probes inside the bioreactor, variability shown on the PCA (Figure 4) reflected the instrumental variability present between both analyzers (ProCellics™ n°1 and n°2) only.

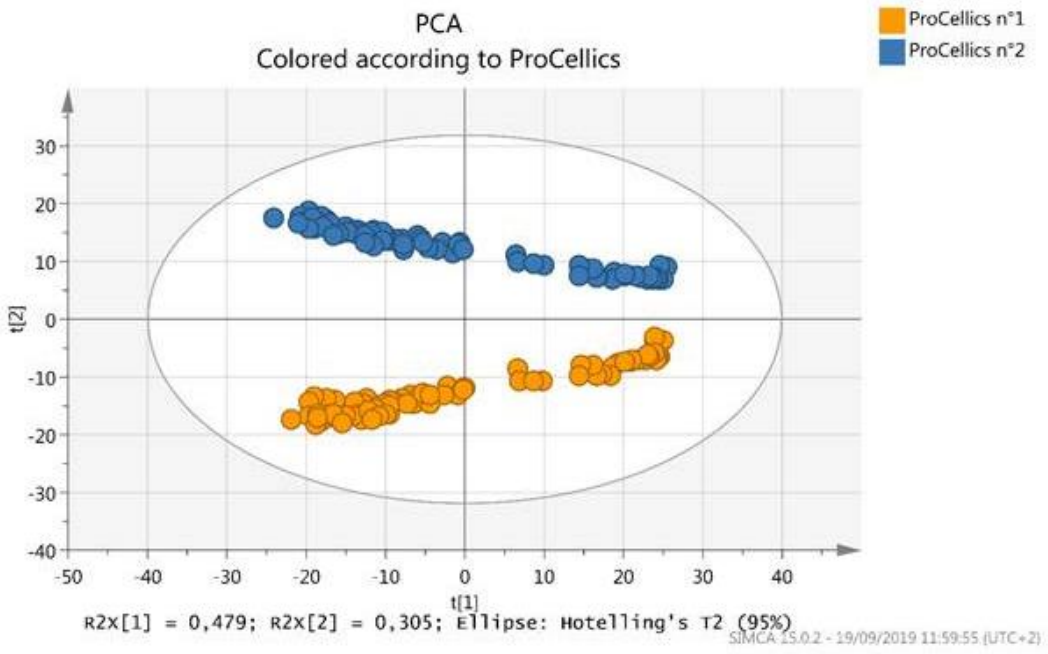

FIGURE 4: PCA demonstrating instrumental variability between ProCellics™ n°1 and ProCellics™ n°2 and its impact on the two CHO cell culture runs. PCA, Principal Component Analysis.

Based on user cases, we tested two solutions to reduce the impact of calibration transfer on model prediction errors: a calibration transfer solution and a chemometric solution.

The calibration transfer solution was tested in order to meet the need of a direct monitoring with a new analyzer unknown by the models. It aims to respond to a classical situation of a direct instrument substitution without the possibility to perform a new calibration batch. As we decided to consider the new analyzer (ProCellics n°2) as the reference, the strategy was to calibrate data obtained with the old analyzer in order to do as if they were performed with the new one. Resulting data were used to produce calibration models as if they were obtained with the new analyzer. The calibration transfer solution was based on the KS PDS method where the new analyzer was considered as the master and the old one as the slave. Calibration data from two batches were used to produce eight datasets to build calibration models for glucose, TCD, VCD, glutamic acid, lactate and ammonium (Table 3).

TABLE 3: Instrumental calibration, chemometric models and results for instrumental transfer: master batches and slave batches without PDS or processed with PDS transfer matrices of 5 points (6%), 10 points (12%), 20 points (25%), 30 points (38%), 40 points (50%), 60 points (75%), 80 points (100%) over 80 references, predicting a validation batch analyzed with the master analyzer. PLS, Partial Least Square; Cum, cumulated; PDS, Piecewise Direct Standardization; TCD, Total Cell Density; VCD, Viable Cell Density; RMSEP, Root Mean Square Error of Prediction; A, Latent variables; N, Number of points.

PLS model calibration summary									
	A	N	R^2Y (cum)	Q^2 (cum)		A	N	R^2Y (cum)	Q^2 (cum)
	Without PDS					PDS 30 pts			
Glucose	3	55	0.997	0.995	Glucose	3	55	0.998	0.992
TCD	3	55	0.987	0.981	TCD	3	55	0.992	0.989
VCD	3	55	0.983	0.971	VCD	3	55	0.990	0.984
Ammonium	3	55	0.979	0.964	Ammonium	3	55	0.972	0.958
Glutamic acid	3	55	0.916	0.850	Glutamic acid	3	55	0.900	0.851
Lactate	3	55	0.974	0.949	Lactate	3	55	0.973	0.957
	PDS 5 pts					PDS 40 pts			
Glucose	3	55	0.998	0.993	Glucose	3	55	0.998	0.992
TCD	3	55	0.991	0.987	TCD	3	55	0.992	0.989
VCD	3	55	0.989	0.982	VCD	3	55	0.990	0.984
Ammonium	3	55	0.974	0.960	Ammonium	3	55	0.972	0.959
Glutamic acid	3	55	0.906	0.849	Glutamic acid	3	55	0.899	0.850
Lactate	3	55	0.978	0.960	Lactate	3	55	0.973	0.958
	PDS 10 pts					PDS 60 pts			
Glucose	3	55	0.997	0.992	Glucose	3	55	0.998	0.992
TCD	2	55	0.986	0.984	TCD	3	55	0.992	0.989
VCD	3	55	0.989	0.983	VCD	3	55	0.990	0.984
Ammonium	3	55	0.972	0.959	Ammonium	3	55	0.972	0.958
Glutamic acid	3	55	0.901	0.850	Glutamic acid	3	55	0.900	0.851
Lactate	3	55	0.975	0.958	Lactate	3	55	0.974	0.958
	PDS 20 pts					PDS 80 pts			
Glucose	3	55	0.998	0.993	Glucose	3	55	0.998	0.993
TCD	3	55	0.993	0.989	TCD	3	55	0.993	0.989
VCD	3	55	0.991	0.984	VCD	3	55	0.990	0.984
Ammonium	3	55	0.972	0.957	Ammonium	3	55	0.972	0.958
Glutamic acid	3	55	0.902	0.853	Glutamic acid	3	55	0.902	0.852
Lactate	3	55	0.973	0.957	Lactate	3	55	0.973	0.958

PLS model validation results								
	Without PDS		PDS 5 pts		PDS 10 pts		PDS 20 pts	
	RMSEP	Max value % error	RMSEP	Max value % error	RMSEP	Max value % error	RMSEP	Max value % error
Glucose (g/L)	2.88	31	0.69	7	0.70	8	0.73	8
TCD (x10^6 cells/mL)	2.13	18	1.07	9	1.02	9	1.15	10
VCD (x10^6 cells/mL)	4.37	38	1.09	9	1.10	9	1.24	11
Ammonium (mM)	6.00	31	3.56	19	3.29	17	3.28	17
Glutamic acid (mM)	0.87	15	0.94	16	0.89	15	0.90	15
Lactate (g/L)	0.22	8	0.19	7	0.19	7	0.20	7
	PDS 30 pts		PDS 40 pts		PDS 60 pts		PDS 80 pts	
	RMSEP	Max value % error	RMSEP	Max value % error	RMSEP	Max value % error	RMSEP	Max value % error
Glucose (g/L)	0.68	7	0.68	7	0.68	7	0.70	7
TCD (x10^6 cells/mL)	1.10	9	1.07	9	1.09	9	1.10	9
VCD (x10^6 cells/mL)	1.22	10	1.22	11	1.23	11	1.22	11
Ammonium (mM)	3.27	17	3.25	17	3.31	17	3.27	17
Glutamic acid (mM)	0.90	15	0.90	15	0.90	15	0.90	15
Lactate (g/L)	0.20	7	0.20	7	0.20	7	0.20	7

Thus, off-line data from two cell culture runs were combined with their respective Raman spectra obtained with ProCellics™ n°1 that we considered as the old one and consequently as the slave. To do as if these data were obtained with a new analyzer (named ProCellics™ n°2), seven different instrumental calibrations were set up for ProCellics™ n°1 with PDS transfer matrices of 5 points (6%), 10 points (12%), 20 points (25%), 30 points (38%), 40 points (50%), 60 points (75%), 80 points (100%) over 80 references. At this stage of the study, the KS PDS matrices were built from three culture runs analyzed with both analyzers simultaneously. Models were performed for each seven created datasets as well as for the original dataset. As shown in Table 3, the models have high explained ($R^2Y > 0.90$) and predicted ($Q^2 > 0.80$) performances for all parameters. To test the effect of the KS PDS method on model prediction errors, a third batch analyzed by ProCellics™ n°2 (the new one) was considered as a validation batch and performed under similar process conditions as calibration batches (analyzed by ProCellics™ n°1). The maximum value percentage errors of RMSEPs were calculated for each calibration method. Validation results showed that the models for glucose, TCD, VCD, ammonium and lactate performed with the KS PDS calibrated datasets presented results with a higher prediction accuracy than the results obtained with the classical calibrated dataset without PDS (Table 3 and Figure 5).

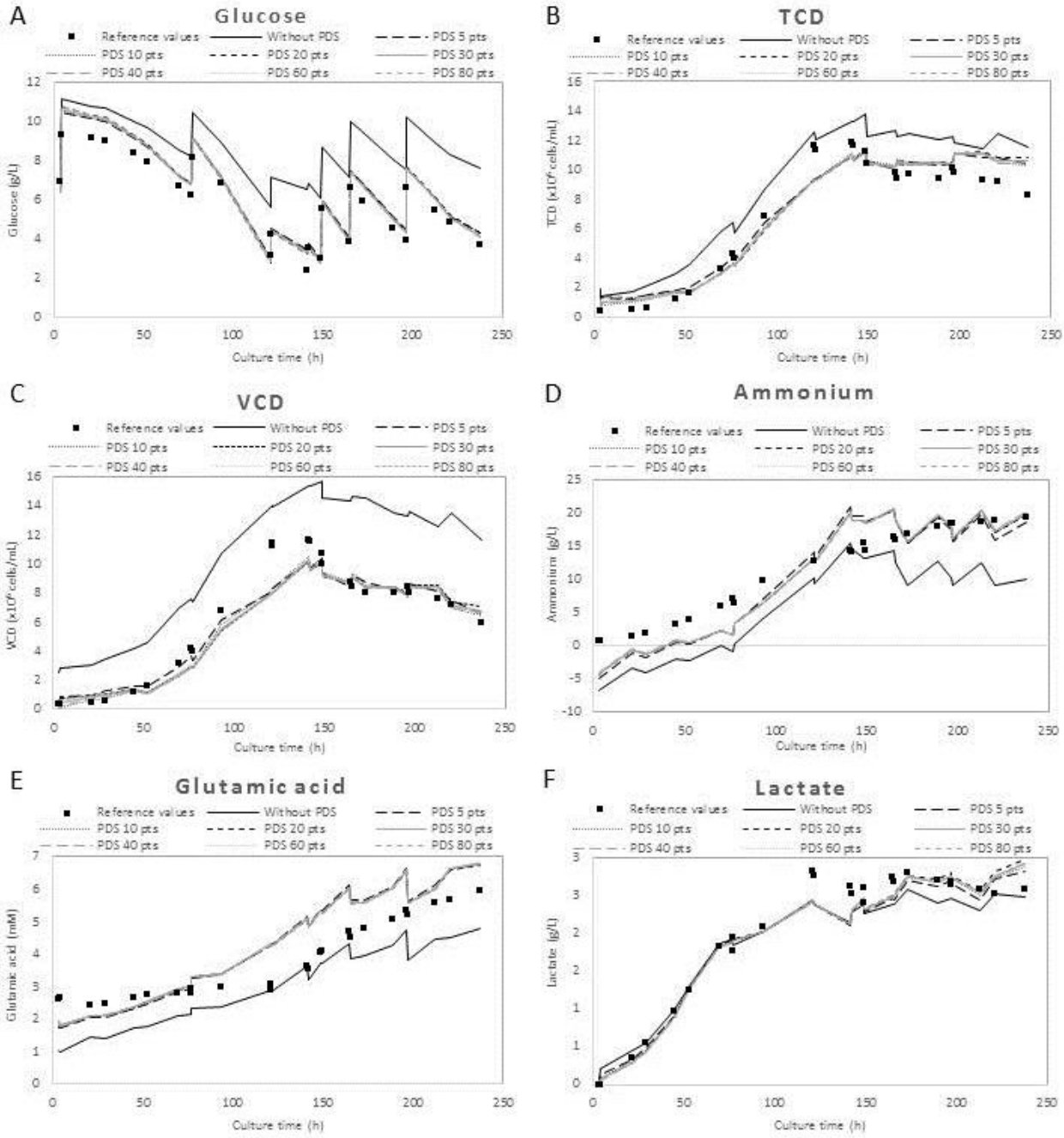

FIGURE 5: PLS model prediction results from Kennard-Stone PDS calibration datasets (without PDS or processed with PDS transfer matrices of 5 points (6%), 10 points (12%), 20 points (25%), 30 points (38%), 40 points (50%), 60 points (75%), 80 points (100%) over 80 references predicting a validation batch analyzed with the master analyzer for (A) glucose, (B) TCD, (C) VCD, (D) ammonium, (E) glutamic acid and (F) lactate. TCD, Total Cell Density; VCD, Viable Cell Density; PDS, Piecewise Direct Standardization.

For instance, the KS PDS calibration allows dividing the model prediction errors by four for glucose and by two for TCD. As shown in Figure 5, predicted kinetics were highly reliable for each component. Moreover, the comparison

between the results of the different KS PDS transfer matrix sizes showed higher prediction accuracy for 10 points (Table 3 and Figure 5). Consequently, we can assume that the selection of only 5 to 10 points (determined with optimization by the KS algorithm) to build the PDS matrix is sufficient and that this kind of transfer method could be used in a regular instrument calibration process for a new analyzer with reference samples measured before the first use of the analyzer. Using such instrument calibration method, the old analyzer could then be considered as the master, and the new analyzer as the slave depending on the model maintenance strategy.

The second solution tested to reduce the impact of instrumental variability between two analyzers on model prediction errors was a chemometric solution, based on the integration of new calibration data collected with the unknown analyzer. This solution responds to a situation in which the analyzer substitution could be anticipated and with the possibility to perform new calibration batches. To correct the instrumental variability and maintain the model prediction accuracy, a comparison study between models that integrate, or not, new calibration batches (one or two) was performed. Calibration data from three batches and two analyzers were used to produce three different datasets to build calibration models for glucose, TCD, VCD, ammonium, glutamic acid and lactate. Off-line data from the three cell culture runs were combined with their respective Raman spectra obtained by two different analyzers to build the three datasets: either with the previous analyzer (ProCellics™ n°1) only, or with two batches from the previous analyzer (ProCellics™ n°1) and the third batch with the new one (ProCellics™ n°2), or with one batch from the previous analyzer (ProCellics™ n°1) and with two batches from the new one (ProCellics™ n°2).

TABLE 4: Chemometric calibration and results for instrumental transfer: 0/3, 1/3 or 2/3 batches analyzed with ProCellics™ n°2 predicting a validation batch analyzed with ProCellics™ n°2. PLS, Partial Least Square; Cum, cumulated; PC2, ProCellics™ n°2; TCD, Total Cell Density; VCD, Viable Cell Density; RMSEP, Root Mean Square Error of Prediction; A, Latent variables; N, Number of points.

PLS model calibration summary				
	A	N	R^2Y (cum)	Q^2 (cum)
0 batch/3 ProCellics™ n°2				
Glucose	3	73	0.986	0.978
TCD	5	76	0.990	0.978
VCD	5	76	0.981	0.920
Ammonium	3	76	0.983	0.979
Glutamic acid	7	76	0.985	0.927
Lactate	5	76	0.990	0.981
1 batch/3 ProCellics™ n°2				
Glucose	3	69	0.990	0.987
TCD	7	76	0.995	0.988
VCD	7	76	0.990	0.975
Ammonium	4	76	0.986	0.982
Glutamic acid	5	76	0.943	0.914
Lactate	6	76	0.990	0.982
2 batches/3 ProCellics™ n°2				
Glucose	3	69	0.991	0.985
TCD	6	76	0.990	0.984
VCD	8	76	0.995	0.984
Ammonium	5	76	0.979	0.967
Glutamic acid	4	71	0.951	0.936
Lactate	6	76	0.990	0.977

PLS model validation results							
		0 batch/3 ProCellics™ n°2		1 batch/3 ProCellics™ n°2		2 batches/3 ProCellics™ n°2	
	Range	RMSEP	Max value % error	RMSEP	Max value % error	RMSEP	Max value % error
Glucose (g/L)	2.4 - 9.3	2.92	31	0.68	7	0.55	6
TCD (x10^6 cells/mL)	0.3 - 11.8	1.60	13	0.80	7	0.56	5
VCD (x10^6 cells/mL)	0.3 - 11.6	4.46	38	1.13	9	0.71	6
Ammonium (mM)	0.5 - 19.2	9.62	50	5.52	18	3.42	16
Glutamic acid (mM)	2.4 - 6.0	0.72	12	0.32	5	0.46	7
Lactate (g/L)	0 - 2.8	0.86	30	0.21	7	0.28	10

As shown in Table 4, all models have high explained ($R^2Y > 0.90$) and predicted ($Q^2 > 0.90$) performances for all parameters. Then, models were validated using one batch measured by the new analyzer (ProCellics™ n°2). This fourth validation batch was not included in the calibration models and was performed under similar process

conditions as those used for the calibration batches. To evaluate the chemometric capability to integrate instrument variability, the maximum value percentage errors of RMSEPs were calculated. As soon as the new instrument variability was integrated in the models, validation results showed much lower model prediction errors, less than 10% for glucose, VCD, TCD, glutamic acid and lactate and less than 20% for ammonium (Table 4).

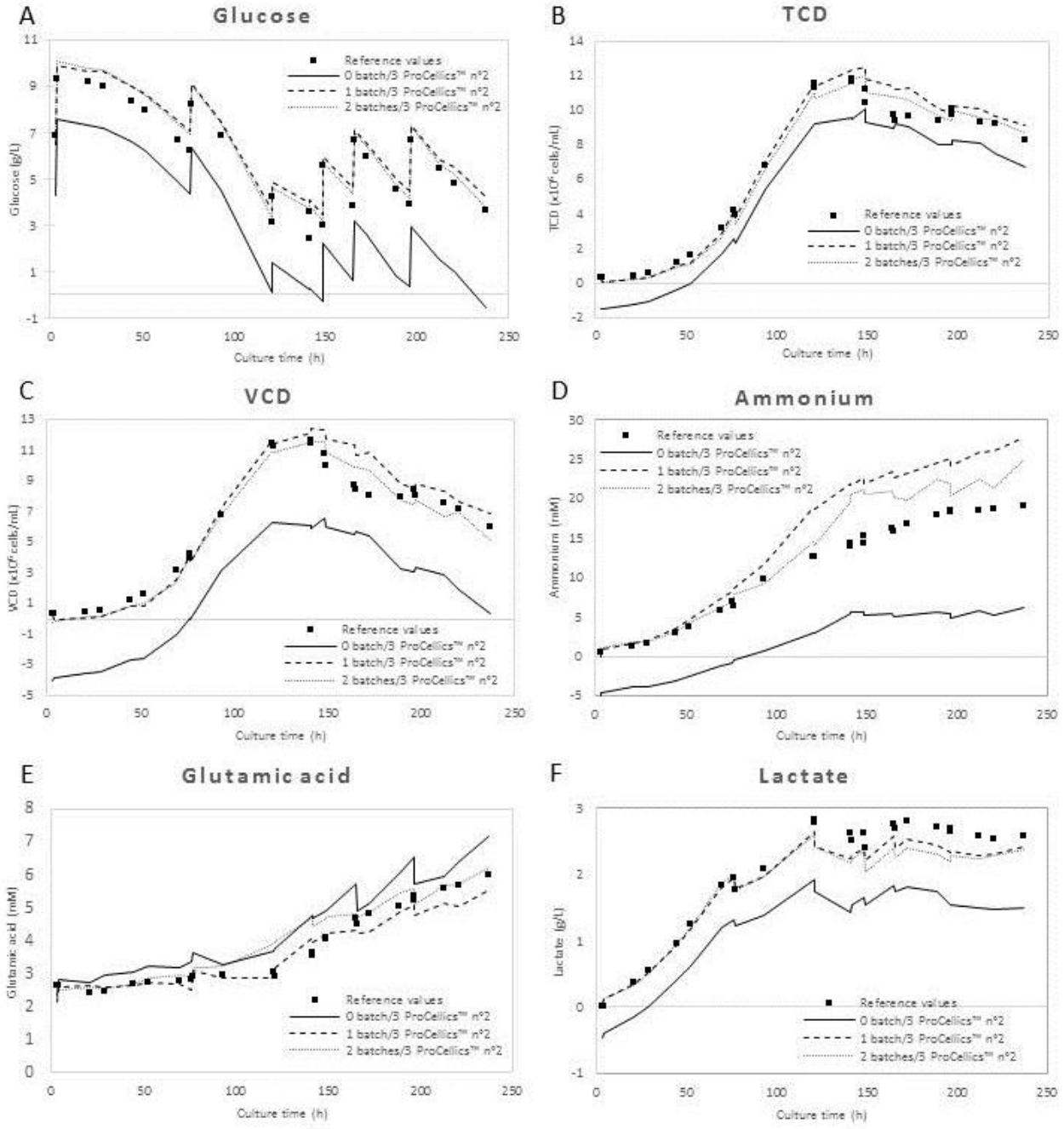

FIGURE 6: PLS model prediction results from calibration datasets containing 0 batch/3, 1 batch/3 or 2 batches/3 analyzed with ProCellics™ n°2 predicting a validation batch analyzed with ProCellics™ n°2 for (A) glucose, (B) TCD, (C) VCD, (D) ammonium, (E) glutamic acid and (F) lactate. TCD, Total Cell Density; VCD, Viable Cell Density.

As shown in Figure 6, predicted kinetics were highly reliable for each component. These results demonstrated that inserting only one batch analyzed with the unknown analyzer in the calibration dataset to predict a batch analyzed with this same analyzer was sufficient to get rid of the prediction gap induced by instrumental variability.

3.3. Discussion

These transferability studies validate some methods to manage the calibration transfer between different devices and different hardware configurations in the context of an extended use of Raman monitoring inside bioprocessing sites. It is worth noticing that the presented study is based on real Raman and off-line data as well as on predictions performed with multivariate models. This *posteriori* prediction was necessary in order to compare different methods on the same datasets. However, the objective was to develop solutions which can be integrated in a real-time monitoring data processing. The good results presented in these studies are nonetheless drawn from a limited number of hardware devices (four probes for the inter-probe transferability study and two analyzers for the transferability between instruments study). The extrapolation of the performances of these methods may be studied on a larger scale with a greater number of instruments and multiplexed probes, typically corresponding to production lines in the biopharmaceutical industry.

Mathematical instrumental standardization methods such as the PDS can be added to the classical calibration step of a single instrument, in order to facilitate the transfer of PAT models between devices before or/and in addition to the chemometric transfer. Strategies for creating and optimizing the parameters of these methods must be carefully considered. The PDS method works well when the transferred associated instruments are similar, sharing the same wavelength range and sampling frequency. However, PDS cannot be used directly to transfer spectra of different wavelengths or recorded by instruments with different spectrometer configurations in resolution and sampling. The combined use of other calibration transfer algorithms based on dataset matrix orthogonalization - such as Dynamic Orthogonal Projection (Zeaiter, Roger, & Bellon-Maurel, 2006), External Parameter Orthogonalization (Roger, Chauchard, & Bellon-Maurel, 2003), and Orthogonal Signal Correction (Sjöblom, Svensson, Josefson, Kullberg, & Wold, 1998; Wold, Antti, Lindgren, & Öhman, 1998) - or directly on Raman spectra processing - such as Shenk-Westerhaus algorithm (Bouveresse, Massart, & Dardenne, 1994; Shenk, Westerhaus, & Templeton, 1985) or Spectral Space Transformation (Du et al., 2011) - are attractive openings to further maintaining the predictive abilities of multivariate calibration models with large scale instrument

installations. A challenging extension of this work could be to develop a similar methodology for calibration transfer between different Raman analyzers from different suppliers. The construction of chemometric models on a large number of instruments will require the optimized selection of reference samples for different processes, potentially via the building of local models based on Locally Weighted Partial Least Square (LW-PLS) used as a Just-In-Time (JIT) learning method (Hazama & Kano, 2015). Finally, the long-term maintenance of models is a major question to be taken into account from the beginning of the instrumentation implementation to ensure the continuity of measurements without increasing the chemometric prediction errors, and even improve models over time and changing instruments with the least amount of time and cost (Wise & Roginski, 2015). The support of real-time machine learning methods is thus interesting in this perspective, even if the authorization of automatic updating methods within a regulated environment seems still complex.

4. CONCLUSION

In this paper, based on cell culture user cases, we studied two types of Raman instrumental variability and their impacts on PLS model prediction errors. Firstly, we explored instrumental variability between probes inside a Raman multi-channel analyzer. It is worth noticing that in our study calibration data were collected with one probe unknown by the model. Our results showed that the KS PDS calibration method used allowed to correct the instrumental variability and to obtain good model prediction accuracy. Secondly, we investigated instrumental variability between two analyzers, in the case of an analyzer replacement. To perform monitoring directly with the new analyzer, we found that the KS PDS method enabled to lower the impact on model prediction errors. However, if it was possible to perform new calibration batches, we demonstrated that only one batch with the unknown analyzer in the calibration dataset was sufficient to get rid of the prediction gap induced by instrumental variability. Consequently, to meet the need of an analyzer substitution, a first monitoring can be directly performed with the KS PDS strategy. Then, data obtained should be inserted in the calibration dataset to integrate the instrumental variability in the model and thereby to correct it. Another possibility is to anticipate this transfer with a regular instrumentation calibration process including always the same reference samples. To conclude, we demonstrated that impacts of Raman instrumental variability over probes and analyzers on multivariate calibration model prediction errors can be corrected to maintain an accurate cell culture monitoring.

Acknowledgments:

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 779218. The authors declare no conflict of interest. We thank D. Simiand for the critical reading of the manuscript and K. Grollier for applicative support throughout this work. We acknowledge the LBFA Laboratory including U. Schlattner and particularly F. Lamarche for his technical advice and help.

Authors have no conflict of interest relevant to this article.

Grant numbers: *European Union's Horizon 2020 research and innovation programme* agreement No 779218

References:

Abu-Absi, N. R., Kenty, B. M., Cuellar, M. E., Borys, M. C., Sakhamuri, S., Strachan, D. J., … Li, Z. J. (2011). Real time monitoring of multiple parameters in mammalian cell culture bioreactors using an in-line Raman spectroscopy probe. *Biotechnology and Bioengineering*, *108*(5), 1215–1221. https://doi.org/10.1002/bit.23023

André, S., Cristau, L. Saint, Gaillard, S., Devos, O., Calvosa, É., & Duponchel, L. (2015). In-line and real-time prediction of recombinant antibody titer by in situ Raman spectroscopy. *Analytica Chimica Acta*, *892*, 148–152. https://doi.org/10.1016/j.aca.2015.08.050

ASTM. (2014). Standard Guide for Raman Shift Standards for Spectrometer Calibration. *ASTM International*, *96*(2002 (Reapproved 2014)), 1–11. https://doi.org/10.1520/E1840-96R07.2

Berry, B. N., Dobrowsky, T. M., Timson, R. C., Kshirsagar, R., Ryll, T., & Wiltberger, K. (2016). Quick generation of Raman spectroscopy based in-process glucose control to influence biopharmaceutical protein product quality during mammalian cell culture. *Biotechnology Progress*, *32*(1), 224–234. https://doi.org/10.1002/btpr.2205

Bouveresse, E., & Massart, D. L. (1996). Improvement of the piecewise direct standardisation procedure for the transfer of NIR spectra for multivariate calibration. *Chemometrics and Intelligent Laboratory Systems*, *32*(2), 201–213. https://doi.org/10.1016/0169-7439(95)00074-7

Bouveresse, E., Massart, D. L., & Dardenne, P. (1994). Calibration transfer across near-infrared spectrometric instruments using Shenk's algorithm: effects of different standardisation samples. *Analytica Chimica Acta*,

297(3), 405–416. https://doi.org/10.1016/0003-2670(94)00237-1

Buckley, K., & Ryder, A. G. (2017). Applications of Raman Spectroscopy in Biopharmaceutical Manufacturing: A Short Review. *Applied Spectroscopy*, *71*(6), 1085–1116. https://doi.org/10.1177/0003702817703270

Craven, S., Whelan, J., & Glennon, B. (2014). Glucose concentration control of a fed-batch mammalian cell bioprocess using a nonlinear model predictive controller. *Journal of Process Control*, *24*(4), 344–357. https://doi.org/10.1016/j.jprocont.2014.02.007

Du, W., Chen, Z. P., Zhong, L. J., Wang, S. X., Yu, R. Q., Nordon, A., … Holden, M. (2011). Maintaining the predictive abilities of multivariate calibration models by spectral space transformation. *Analytica Chimica Acta*, *690*(1), 64–70. https://doi.org/10.1016/j.aca.2011.02.014

Ettah, I., & Ashton, L. (2018). Engaging with Raman Spectroscopy to Investigate Antibody Aggregation. *Antibodies*, *7*(3), 24. https://doi.org/10.3390/antib7030024

Hazama, K., & Kano, M. (2015). Covariance-based locally weighted partial least squares for high-performance adaptive modeling. *Chemometrics and Intelligent Laboratory Systems*, *146*, 55–62. https://doi.org/10.1016/j.chemolab.2015.05.007

Jackson, J. E. (1991). A User's Guide to Principal Components John Wiley. *New York*.

Li, M. Y., Ebel, B., Paris, C., Chauchard, F., Guedon, E., & Marc, A. (2018). Real-time monitoring of antibody glycosylation site occupancy by in situ Raman spectroscopy during bioreactor CHO cell cultures. *Biotechnology Progress*, *34*(2), 486–493. https://doi.org/10.1002/btpr.2604

Mehdizadeh, H., Lauri, D., Karry, K. M., Moshgbar, M., Procopio-Melino, R., & Drapeau, D. (2015). Generic Raman-based calibration models enabling real-time monitoring of cell culture bioreactors. *Biotechnology Progress*, *31*(4), 1004–1013. https://doi.org/10.1002/btpr.2079

Roger, J. M., Chauchard, F., & Bellon-Maurel, V. (2003). EPO-PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits. *Chemometrics and Intelligent Laboratory Systems*, *66*(2), 191–204. https://doi.org/10.1016/S0169-7439(03)00051-0

Rowland-Jones, R. C., van den Berg, F., Racher, A. J., Martin, E. B., & Jaques, C. (2017). Comparison of spectroscopy technologies for improved monitoring of cell culture processes in miniature bioreactors. *Biotechnology Progress*, *33*(2), 337–346. https://doi.org/10.1002/btpr.2459

Savitzky, A., & Golay, M. J. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, *36*(8), 1627–1639.

Shenk, J. S., Westerhaus, M. O., & Templeton, W. C. (1985). Calibration Transfer Between near Infrared Reflectance Spectrophotometers1. *Crop Science*, *25*(1), 159. https://doi.org/10.2135/cropsci1985.0011183x002500010038x

Sjöblom, J., Svensson, O., Josefson, M., Kullberg, H., & Wold, S. (1998). An evaluation of orthogonal signal correction applied to calibration transfer of near infrared spectra. *Chemometrics and Intelligent Laboratory Systems*, *44*(1–2), 229–244. https://doi.org/10.1016/S0169-7439(98)00112-9

Swann, P., Brophy, L., Strachan, D., Lilly, E., & Jeffers, P. (2017). *Biomanufacturing Technology Roadmap - In-line monitoring and real-time release*. 1–40.

Tulsyan, A., Schorner, G., Khodabandehlou, H., Wang, T., Coufal, M., & Undey, C. (2019). A machine-learning approach to calibrate generic Raman models for real-time monitoring of cell culture processes. *Biotechnology and Bioengineering*, *116*(10), 2575–2586. https://doi.org/10.1002/bit.27100

Webster, T. A., Hadley, B. C., Hilliard, W., Jaques, C., & Mason, C. (2018). Development of generic raman models for a GS-KOTM CHO platform process. *Biotechnology Progress*, *34*(3), 730–737. https://doi.org/10.1002/btpr.2633

Wise, B. M., & Roginski, R. T. (2015). A calibration model maintenance roadmap. *IFAC-PapersOnLine*, *28*(8), 260–265. https://doi.org/10.1016/j.ifacol.2015.08.191

Wold, S., Antti, H., Lindgren, F., & Öhman, J. (1998). Orthogonal signal correction of near-infrared spectra. *Chemometrics and Intelligent Laboratory Systems*, *44*(1–2), 175–185. https://doi.org/10.1016/S0169-7439(98)00109-9

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and*

Intelligent Laboratory Systems, *58*(2), 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1

Zeaiter, M., Roger, J. M., & Bellon-Maurel, V. (2006). Dynamic orthogonal projection. A new method to maintain the on-line robustness of multivariate calibrations. Application to NIR-based monitoring of wine fermentations. *Chemometrics and Intelligent Laboratory Systems*, *80*(2), 227–235. https://doi.org/10.1016/j.chemolab.2005.06.011