

## RESEARCH ARTICLE

# Feasibility and performance of cross-clone Raman calibration models in CHO cultivation

Rafael Machleid<sup>1,3</sup> | Marek Hoehse<sup>1</sup> | Steffi Scholze<sup>1</sup> | Kleanthis Mazarakis<sup>2</sup> |  
David Nilsson<sup>3</sup> | Erik Johansson<sup>4</sup> | Christoph Zehe<sup>5</sup> | Johan Trygg<sup>3,6</sup> |  
Christian Grimm<sup>1</sup> | Izabella Surowiec<sup>4</sup> 

<sup>1</sup>Sartorius Stedim Biotech GmbH, Göttingen, Germany

<sup>2</sup>Sartorius Stedim UK Ltd., Epsom, UK

<sup>3</sup>Computational Life Science Cluster (CLiC), Umeå University, Umeå, Sweden

<sup>4</sup>Sartorius Stedim Data Analytics AB, Umeå, Sweden

<sup>5</sup>Sartorius Corporate Research, Ulm, Germany

<sup>6</sup>Sartorius Corporate Research, Umeå, Sweden

## Correspondence

Izabella Surowiec, Sartorius Stedim Data Analytics AB, Östra Strandgatan 24, 903 33 Umeå, Sweden.

Email: [izabella.surowiec@sartorius.com](mailto:izabella.surowiec@sartorius.com)

## Abstract

Raman spectroscopy is widely used in monitoring and controlling cell cultivations for biopharmaceutical drug manufacturing. However, its implementation for culture monitoring in the cell line development stage has received little attention. Therefore, the impact of clonal differences, such as productivity and growth, on the prediction accuracy and transferability of Raman calibration models is not yet well described. Raman OPLS models were developed for predicting titer, glucose and lactate using eleven CHO clones from a single cell line. These clones exhibited diverse productivity and growth rates. The calibration models were evaluated for clone-related biases using clone-wise linear regression analysis on cross validated predictions. The results revealed that clonal differences did not affect the prediction of glucose and lactate, but titer models showed a significant clone-related bias, which remained even after applying variable selection methods. The bias was associated with clonal productivity and lead to increased prediction errors when titer models were transferred to cultivations with productivity levels outside the range of their training data. The findings demonstrate the feasibility of Raman-based monitoring of glucose and lactate in cell line development with high accuracy. However, accurate titer prediction requires careful consideration of clonal characteristics during model development.

## KEYWORDS

bioprocess development, bioprocess engineering, bioprocess monitoring, CHO cells

**Abbreviations:** AsLS, asymmetric least squares smoothing; CHO, Chinese hamster ovary; CPP, critical process parameter; CQA, critical quality attribute; CV, cross validation; DoE, design of experiments; FDA, food and drug administration; HEK, Human embryonic kidney; HeLa, Henrietta Lacks cell line; IgG, Immunoglobulin G; IR, importance ratio; LV, latent variables; mAb, monoclonal antibody; NIR, Near-infrared; OPLS, orthogonal projection to latent structures; PAT, process analytical technology; PLS, projection to latent structures; PS, prediction set; RMSE<sub>CV</sub>, root mean square error of cross validation; RMSE<sub>E</sub>, root mean square error of estimation; RMSE<sub>P</sub>, root mean square error of prediction; Sf9, Spodoptera frugiperda cell line; SR, selectivity ratio; TCD, total cell density; VCD, viable cell density; VIP, variable importance on projections; VIP<sub>total</sub>, total sum of the predictive and orthogonal VIP parts in OPLS.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Biotechnology Journal* published by Wiley-VCH GmbH.

## 1 | INTRODUCTION

The global market size of monoclonal antibodies (mAbs) is projected to attain a value of USD 500 billion by the year 2030, exhibiting an estimated compound annual growth rate of 11.3% from 2022 to 2030.<sup>[1]</sup> The most popular representative mAbs are Immunoglobulin G (IgG) antibodies, which are highly specific glycoproteins employed as therapeutics for a variety of conditions. Therapeutic protein manufacturing involves a set of complex biological processes, typically involving

cultivating genetically modified mammalian cell lines, such as Chinese hamster ovary (CHO) cells.

Traditionally, the bioprocess has been monitored and controlled through regular manual sampling and off-line measurements of various critical process parameters (CPPs) and critical quality attributes (CQAs). Driven originally by the FDA's process analytical technologies (PAT) initiative in 2004,<sup>[2]</sup> the biopharmaceutical industry has been continuously striving for the incorporation of PAT tools, such as Raman and NIR spectroscopy, for inline or online process control.

The implementation of Raman spectroscopy for real-time, non-invasive monitoring of various cell culture parameters has been demonstrated in several studies.<sup>[3-5]</sup> More recently, Raman measurements with feedback control loops were utilized to maintain glucose and lactate concentrations in CHO cell cultivations at specified levels.<sup>[6,7]</sup> Typically, CPPs like glucose, lactate, glutamine, glutamate, viable cell density (VCD) and total cell density (TCD), CQAs like glycation and glycosylation and other key performance indicators like titer are measured.<sup>[8,9]</sup>

However, Raman-based monitoring of such parameters requires reliable and accurate calibration models. These are necessary to convert the Raman spectra into quantitative measurements of the analyte of interest and are typically developed using multivariate statistical methods. To increase model specificity and thereby the prediction performance, variable selection methods, such as variable importance on projections (VIP) or selectivity ratio (SR), are frequently employed.<sup>[3]</sup>

The development of accurate and sustainable calibration models over time is challenging. The procedure requires multiple cell cultivations to be executed under varying conditions to cover all relevant process variabilities. This renders calibration model development being costly and time-consuming. Therefore, the biopharmaceutical industry is shifting to small-scale, multi-parallel automated bioreactor platforms, such as the Ambr<sup>®</sup> system, for process development. This approach directly reduces the associated costs, accelerates calibration model development and increases the data availability by facilitating experiment execution within the Design of Experiments (DoE) methodology.<sup>[10,11]</sup> It has been shown that calibration models built on small scales can retain high prediction accuracies when applied for monitoring in larger scale bioreactors.<sup>[12-14]</sup> For some analytes however, incorporating data from large-scale batches into the small-scale model was necessary to achieve accurate predictions, resulting in so called *generic* calibration models.<sup>[12]</sup>

The generic approach has become prevalent in addressing the challenges of transferring models from one process modality to another. These models typically incorporate cultivation data from various process modalities like different process scales, cell lines and/or cultivation modes, and often leverage historical data along with new data specific to the given target process. Thereby, the generic approach effectively reduces the number of required new experiments for calibration model development and lowers the associated costs and time efforts.

Apart from scale transfer, transferability of generic calibration models between different cell line platforms has been demonstrated, for example by André et al.<sup>[15]</sup> Initially, the calibration model built on CHO, HeLa, and Sf9 cultivation data showed large errors for predictions of

glucose and lactate concentration in HEK-293 cultivation. However, the prediction errors were significantly reduced by incorporating even a single batch of HEK-293 cultivation data into the calibration model (representing 2.75% of the total sample size). This result demonstrated that Raman calibration models were transferable between diverse cell line platforms with low additional experimental effort as compared to creating a new calibration model for each specific cell line platform.

For the biopharmaceutical industry, the seamless transfer of calibration models to new cell lines of an established platform process would be highly beneficial as it could significantly accelerate cell line, process, and calibration model development. There exist several studies in the literature that evaluate calibration model transfer based on specific platform processes, regardless of the cell line.<sup>[13,16,17]</sup> For instance, Webster et al. demonstrated successful calibration transfer between cell lines of the CHOK1SV GS-KO platform for generic models predicting glucose, lactate, ammonium, VCD, and TCD.<sup>[16]</sup>

The presented studies demonstrate a comprehensive understanding of calibration model transfer between different cell line platforms and different cell lines of a specific platform process, while the transferability to CHO cultivations of a specific cell line with varying clonal characteristics is not yet well understood. This knowledge, however, is crucial to assess the applicability of Raman-based monitoring in early drug development stages like cell line development and final clone selection. These are typically characterized by large clonal variations, that is, differences in clonal characteristics such as productivity, growth rates, and metabolism.

The implementation of Raman-based monitoring in the cell line development presents the opportunity to acquire initial Raman data that can later be utilized to develop calibration models for the manufacturing stage. Thereby calibration model development could be streamlined, and future experimental efforts reduced. It also provides the increased real-time data for multiple cell culture parameters that can support researchers in gaining better understanding of how cells are behaving and responding to changes in the environment and how these changes might impact growth and productivity. This can lead to more informed choices in terms of media formulation, feeding strategies, and culture conditions. Further, Raman spectroscopy enables early detection of issues or deviations in the cell culture, which helps in identifying problems before they become critical. Quick identification and resolution of issues can significantly reduce the time required for cell line development and therefore accelerate the development of new cell lines and get products to market more rapidly.

A critical component of the required assessment is to understand the potential impact of clonal variations on the development of Raman calibration models and their transferability. Despite the availability of literature, it remains challenging to accurately evaluate these potential influences, as the described models are typically constructed using data from diverse cell lines, multiple products, and process conditions,<sup>[18]</sup> or use only a limited number of clones of a specific cell line.<sup>[13,18]</sup>

To bridge this knowledge gap, we cultivated twelve CHO clones from a single cell line that produced the same IgG antibody but varied in their productivity and growth characteristics. Using this data, we

developed generic Raman calibration models with a focus on titer, glucose, and lactate. We then evaluated these calibration models regarding their dependence on clonal characteristics and transferability to bioprocesses from different CHO clones.

## 2 | EXPERIMENTAL SECTION

### 2.1 | Experimental setup

We performed the experiment in an automated fashion using the Ambr<sup>®</sup> 250 High Throughput system (Sartorius, Germany) equipped with the Ambr<sup>®</sup> Analysis Module (Sartorius, Germany), the BioPAT<sup>®</sup> Spectro (Sartorius, Germany) and the HyperFlux<sup>™</sup> PRO spectrometer (Tornado Spectral Systems, Canada) for Raman measurements (Figure S1). The BioProfile<sup>®</sup> FLEX2 device (Nova Biomedical Corporation, USA) was directly connected to the Ambr<sup>®</sup> Analysis Module for automated measurement of cell culture parameters. We used the Octet<sup>®</sup> R8 system (Sartorius, Germany) for off-line titer measurements.

### 2.2 | Cell cultivations

To study the influence of clonal characteristics on Raman calibration models, we performed fed-batch cultivations with a mAb expressing CHO-DG44 cell line (Sartorius, Germany). Twelve cellular clones (one triplicate: clones 1, 6, and 12), which produced the same IgG antibody but differed in their specific productivity and growth rates, were used. The cultivation of clone 2 was contaminated and terminated early in the process, leading to exclusion of the data from all further analysis. We expanded and cultured the cells in a proprietary, chemically defined medium (Sartorius, Germany). We inoculated the Ambr<sup>®</sup> bioreactors to a target viable cell density of  $0.3 \times 10^6$  cells mL<sup>-1</sup>, operating each cultivation at 36.8°C with 40% dissolved oxygen saturation for 14 days. The pH was maintained at a setpoint of 7.1 ( $\pm 0.1$ ) controlled by CO<sub>2</sub> sparging and NaCO<sub>3</sub> base addition. We performed volume adjusted bolus feed starting on cultivation day 3 following the standard applications implemented at Sartorius. Additionally, if the measured glucose concentration was less than 4 g L<sup>-1</sup>, a daily glucose bolus (400 g L<sup>-1</sup> glucose stock solution) was performed starting on cultivation day 5. To prevent foam formation, we manually added 30  $\mu$ L antifoam solution (Sigma antifoam C 2%) daily.

### 2.3 | Raman and reference data acquisition

We recorded the Raman spectra with the HyperFlux<sup>™</sup> PRO Raman spectrometer (785 nm laser, 495 mW) (Tornado Spectral Systems, Canada) within the BioPAT<sup>®</sup> Spectro measurement chamber (Sartorius, Germany). For each sample, we measured five spectra, each consisting of 80 acquisitions with 750 ms acquisition time. We performed a dark scan before each sample measurement to mitigate the impact of potential variations of ambient light straying into the

shielded measurement chamber. For each bioreactor, we measured the respective Raman spectra once per day. The samples were automatically taken by the liquid handler and transferred from the bioreactor to the Ambr<sup>®</sup> Analysis Module sampling cup and a separate vial for off-line titer measurements. From the sampling cup, the samples were pumped through the flow cell of the BioPAT<sup>®</sup> Spectro and BioProfile<sup>®</sup> FLEX2 device for Raman and cell cultivation parameter measurements, respectively.

### 2.4 | Calibration model development

We averaged the five Raman spectra taken for each sample to increase the signal-to-noise ratio. Before analysis, we discarded ~8% of all Raman spectra due to low signal intensities (empty measuring cell) and further ~20% of the spectra (primarily from the death phase of the cultivations) because of detector saturation caused by strong fluorescence.

For the model building and analysis, we imported the time-matched reference values and spectral data into SIMCA<sup>®</sup> version 17.0 (Sartorius, Germany). We corrected the background of each spectrum through asymmetric least squares smoothing<sup>[19]</sup> (AsLS) baseline correction ( $p = 0.05$ ,  $\lambda = 10^7$ ). Following background correction, we normalized each spectrum to the integral of the water band around 1650 cm<sup>-1</sup>, which helps to account for potential variability in optical sampling volume caused by variations in cell density across the samples. The raw and preprocessed Raman spectra are provided in Figures S2 and S3.

Using principal component analysis (PCA), we studied the variability within the spectral data. We used orthogonal projection to latent structures<sup>[20]</sup> (OPLS) regression for calibration modeling. OPLS removes systematic variance from the X-block that is not correlated to the Y-response leading to separation into orthogonal and predictive components. The predictive loading is used to describe the correlation between the (preprocessed) X-block and the Y-variable. OPLS models exhibit increased interpretability as compared to classical projection to latent structures<sup>[21]</sup> (PLS) models.

We built separate (single-Y) OPLS model for each analyte using the scaled reference values (unit variance) as Y-response. For each and every one of the models, we used the mean-centered variables of the Raman spectra as the X-block. In this way, we built initial calibration models over the Raman shift range of 450–1800 cm<sup>-1</sup> for titer, glucose, and lactate. We modified the initial titer model later based on the outcome of variable selection methods, as described in section 2.7 Variable selection methods.

We validated all calibration models internally using leave-one-group-out cross validation, with each clone assigned to a unique cross validation group (e.g., 12 clones result in 12 groups, with one clone per group). To quantify the average error of the models, we employed the root mean square error of estimation (RMSE<sub>E</sub>) and root mean square error of cross validation (RMSE<sub>CV</sub>) metrics given as

$$\text{RMSE}_E = \sqrt{\frac{\sum (y_{\text{actual}} - y_{\text{cal}})^2}{N}}, \quad (1)$$

$$\text{RMSE}_{\text{CV}} = \sqrt{\sum \frac{(y_{\text{actual}} - y_{\text{predCV}})^2}{N}} \quad (2)$$

with the measured values  $y_{\text{actual}}$ , the predicted values of the model  $y_{\text{cal}}$  or the predicted values from the internal cross validation  $y_{\text{predCV}}$  and the total number of predictions  $N$ . Accurate models are characterized by low  $\text{RMSE}_E$  and  $\text{RMSE}_{\text{CV}}$  values, while a  $\text{RMSE}_E/\text{RMSE}_{\text{CV}}$  ratio of one is considered ideal. Ratios below one indicate overfitting and, specific to our experiment, hint toward a clone-related bias within the calibration model. Further, we have calculated the relative RMSEs in % by dividing the RMSE values by the respective mean measured concentrations (and multiplying with factor 100).

## 2.5 | Clone-related bias analysis

We analyzed the initial calibration models for clone-related bias based on a cross validation procedure. We assigned unique cross validation groups for each clone within a model. During each cross-validation iteration, data from a single clone was excluded from the model building. The reference values of this specific clone were subsequently predicted by the model, which was built only using the remaining clones.

For each clone, we conducted a linear regression analysis on the obtained observed versus predicted plots. We compared the resulting linear regression slopes across all clones. In an ideal scenario, for each clone a slope of one (and intercept of zero) would be obtained from the analysis, implying accurate predictions. Differences in slopes and intercepts between the clones, however, hint towards clone-related bias within the calibration models. The initial titer model showed indications of clone-related bias. Therefore, we applied various variable selection methods to select spectral regions within the Raman spectra for model building. Further, we analyzed the resulting models for clone-related bias employing the same methodology as for the initial calibration models.

## 2.6 | Productivity dependency of titer calibration models

The clone-related bias analysis indicated influence of clonal productivity on the titer calibration models. To assess this effect and its severity, we grouped the clones according to the final titer values reached on the last day of the experiment (Table S1). We considered clones with titer values below  $2.5 \text{ g L}^{-1}$  as low-producing clones (clones 3, 5, 7, 10, 11, and 12) and clones above this threshold as high-producing ones (clones 4, 8, and 9). No classification was made for clones with missing titer values on the final day of the experiment (clones 1, 2, and 6).

According to this grouping, we constructed high-producer and low-producer titer models. We applied the variable selections to the initial titer model and used the selected X-block variables for the high-producer as well as the low-producer models. In this way, we facilitated the comparison of the models obtained after variable selection to the respective high- and low-producer models based on the Raman shift range of  $450\text{--}1800 \text{ cm}^{-1}$ . We then used the high-producer models to

predict the titer values of the low-producing clones and vice versa. This required the models to predict titer values of clones with productivity levels outside the range of their respective training data, which represents a worst-case scenario (e.g., model extrapolation). We evaluated the prediction accuracies based on the Root Mean Square Error of Prediction ( $\text{RMSE}_p$ ) given as

$$\text{RMSE}_p = \sqrt{\sum \frac{(y_{\text{actual}} - y_{\text{predPS}})^2}{N}} \quad (3)$$

with the measured values  $y_{\text{actual}}$ , the predicted values of the prediction set (PS)  $y_{\text{predPS}}$  and the total number of predictions  $N$ .

## 2.7 | Variable selection methods

Both PLS and OPLS methods cannot set the loadings of uncorrelated, non-informative X-block variables to zero (e.g., excluding specific variables completely). Such X-block variables may negatively impact the model performance by inflating the prediction error. To circumvent this issue, multivariate statistical models are frequently used in combination with variable selection methods to determine the most relevant X-block variables. This approach has been shown to increase model specificity and performance by removing spectral regions not relevant to the analyte of interest.<sup>3</sup>

In this study, we employed the established VIP and SR methods and the novel Importance ratio (IR)<sup>[22]</sup> variable selection method. We evaluated whether these methods could reduce the clone-related bias of the initial titer model. We refer to the resulting models as VIP titer model, SR titer model, and IR titer model, indicating which specific method was used for variable selection.

### 2.7.1 | Variable importance on projections variable selection (VIP)

The VIP method quantifies the impact of individual X-block variables on PLS<sup>[23]</sup> and OPLS<sup>[24]</sup> models. For OPLS, VIP vectors for the predictive components, the orthogonal components, and the total sum of the predictive and orthogonal parts ( $\text{VIP}_{\text{total}}$ ) are available. In this study, the  $\text{VIP}_{\text{total}}$  scores were used, which for each X-block variable  $i$  are given as

$$\text{VIP}_{\text{total},i} = \sqrt{\frac{K}{2} \left( \frac{\left[ \sum_{a_o=1}^{A_o} (p_{a_o}^2 \cdot \text{SSX}_{\text{comp},a_o}) \right]}{\text{SSX}_{\text{cum}}} + \frac{\left[ \sum_{a_p=1}^{A_p} (p_{a_p}^2 \cdot \text{SSX}_{\text{comp},a_p}) \right]}{\text{SSX}_{\text{cum}}} \right) + \left( \frac{\left[ \sum_{a_o=1}^{A_o} (p_{a_o}^2 \cdot \text{SSY}_{\text{comp},a_o}) \right]}{\text{SSY}_{\text{cum}}} + \frac{\left[ \sum_{a_p=1}^{A_p} (p_{a_p}^2 \cdot \text{SSY}_{\text{comp},a_p}) \right]}{\text{SSY}_{\text{cum}}} \right)} \quad (4)$$

with the indices  $a_o$  and  $a_p$  denoting orthogonal or predictive components respectively. Analogously,  $A_o$  and  $A_p$  describe the total number of orthogonal and predictive components respectively.  $p_{a_o}$  and  $p_{a_p}$  denote the normalized orthogonal and predictive loadings of the indicated component respectively. Likewise,  $\text{SSX}_{\text{comp}}$  and  $\text{SSY}_{\text{comp}}$  denote the explained sum of squares of the indicated orthogonal or predictive component in the X- and Y-space respectively.  $K$  indicates the total number of variables and  $\text{SSX}_{\text{cum}}$  and  $\text{SSY}_{\text{cum}}$  the cumulative explained

sum of squares by all model components in the X- and Y-space respectively.

Typically, variables with VIP scores greater than one are selected for retention, while all other variables are discarded. In this study, we adopted this threshold for the sake of comparability to typical VIP method implementations.<sup>[3,25]</sup>

## 2.7.2 | Selectivity ratio variable selection (SR)

The SR method<sup>[26]</sup> evaluates the relative contribution of each X-block variable to the explained and residual variance in the target projection (TP) component. The selectivity ratio  $SR_i$  of the  $i^{\text{th}}$  variable is calculated as the ratio of the explained variance  $SS_{i, \text{explained}}$  to the residual variance  $SS_{i, \text{residual}}$ , given as

$$SR_i = \frac{SS_{i, \text{explained}}}{SS_{i, \text{residual}}} \quad (5)$$

The explained and residual variances of the  $i^{\text{th}}$  variable are calculated as

$$SS_{i, \text{explained}} = \|t_{\text{TP}} p_{\text{TP}, i}^T\|^2 \quad (6)$$

$$SS_{i, \text{residual}} = \|e_{\text{TP}, i}\|^2 \quad (7)$$

with the scores  $t_{\text{TP}}$ , loadings  $p_{\text{TP}, i}$  and residuals  $e_{\text{TP}, i}$  of the target projection.

A commonly used approach to determine the threshold between high and low discriminating variables is to conduct an F-test. The calculated F value,  $F_{\text{calc}}$ , must exceed the critical value  $F_{\text{crit}}$  for the F distribution to reject the null hypothesis of equal explained and residual variances.  $F_{\text{calc}}$  is given as

$$F_{\text{calc}} = SR_i > F_{\text{crit}} = F(\alpha, N-2, N-3) \quad (8)$$

with the significance level  $\alpha$  and sample size  $N$ . The degrees of freedom for the explained variance is equal to  $N-2$ , it is equal to  $N-3$  for the residual variance. In this study, we employed a significance level of 95% ( $\alpha = 0.95$ ) to enhance comparability to typical SR method implementations.<sup>[3,25]</sup>

## 2.7.3 | Importance ratio variable selection (IR)

We used the IR method<sup>[22]</sup> to identify the most relevant Raman spectral regions for the calibration models. The importance ratio  $R_i$  for each variable  $i$  is given as

$$R_i = \frac{|p_i|}{pSE_{CV, i}} \quad (9)$$

with the predictive loading  $p_i$  and the jackknife standard error of cross validation  $pSE_{CV, i}$ . In general, the higher the computed  $R_i$  value of a variable, the more relevant the variable is deemed for the prediction of the

Y-response. X-block variables, which are strongly correlated to the Y-response, are characterized by large absolute predictive loadings. The jackknife standard error of cross validation is a measure of confidence in the computed predictive loadings value of each variable. Therefore, a high value of  $R_i$  implies that the variable has a strong correlation to the Y-response (large absolute predictive loadings) or high confidence in the predictive loading (small jackknife standard errors). On the contrary, a low value of  $R_i$  suggests a weak correlation to the Y-response (small absolute predictive loading) or low confidence in the predictive loading (large jackknife standard error). As cross validation is utilized for the confidence estimation, the IR method can be tuned according to the general cross validation setup. Variables with varying correlation structures across cross validation groups may result in high jackknife standard errors, allowing for variable selection (and exclusion) based on group-to-group variations.

We implemented the IR method as an iterative procedure. In each iteration, the  $R$  value of each variable in the current OPLS model is calculated and  $N$  variables with the lowest IR value are excluded from the X-block (Figure S4). A new OPLS model with the same number of latent variables as the starting model is then fitted and the process is repeated until no variables remain in the X-block. The final model is selected based on the highest  $Q^2$  value.

In this study, we applied the IR method to the initial titer model, which used the Raman shift range from 450–1800  $\text{cm}^{-1}$  as X-block variables (one variable per wavenumber). We set the method to exclude 25 variables per iteration ( $N = 25$ ) and assigning unique cross validation groups to each clone (leave-one-group-out) throughout the procedure.

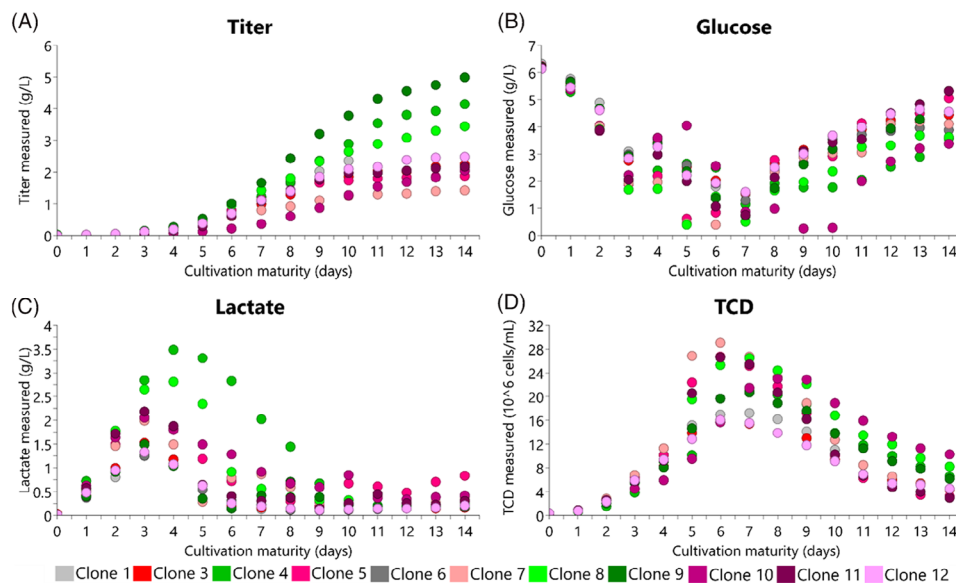
## 3 | RESULTS AND DISCUSSION

### 3.1 | Spectral data analysis

We developed a PCA model to analyze the variability within the Raman spectra acquired during the 14-day lasting cultivation processes. The first, second, third, and fourth principal components captured 72%, 13%, 7%, and 4% of the spectral variation, respectively. We found that the variability within the trajectories of triplicate clones 1, 6, and 12 was considerably smaller than the spread among all the other clones (Figure S5).

The scores plot identified clusters within the spectral dataset that are primarily associated with the process maturity (Figure S6 and loadings plot in Figure S7). Tight clustering was observed in the initial stage of the cultivations (days 0–3), which indicates high spectral similarity between the clones. This implied minor impact of clonal variations on the Raman spectra in the early stage of the cultivations.

The clusters expanded considerably with the maturity of the process. The increasing spectral variation must be associated with the clonal differences and their effects, as the process conditions, such as pH, dissolved oxygen, feeding, and seeding strategy, were equal and strictly controlled for all cultivations. We expected an influence of productivity and growth rates on the Raman spectra due to their direct



**FIGURE 1** Daily profiles for (A) titer, (B) glucose, (C) lactate, and (D) TCD colored according to clones; for all plots high-producing clones are colored in green shades and low-producing clones in red/violet shades.

effects on the chemical composition within the cultivation. The impact of clonal productivity on Raman spectra of cell cultivations has been noted in the literature before.<sup>[13]</sup>

Despite increasing spectral variance during the cultivation processes, the scores plot showed similar process trajectories for all clones. This result gave a first indication for the feasibility of transferring calibration models between different clones of a specific cell line.

### 3.2 | Reference data analysis

The daily profiles, as depicted in Figure 1, visualize the impact of differences between clones on the performed cultivations. Generally, all clones showed similar trends for titer, glucose, lactate, and TCD. However, some clones produced titer in higher amounts than others, as expected due to the design of the experiment. Likewise, the TCD profiles showed differences between the clones due to varying growth rates and peak cell densities. All cultivations reached their peak cell densities on cultivation days 6 or 7, except clone 10 (day 8). The varying cell densities further caused slightly different glucose and lactate profiles, which could also be attributed to metabolic differences between the clones. It is noteworthy that clone 4 produced considerably higher lactate levels during cultivation days 3 to 8 than all other clones.

The correlation matrix was computed to assess the correlation between different analyte trends (Table S2). Correlation coefficients between absolute values of 0.7 and 1.0 define strong correlations, while absolute values between 0.3 and 0.7 indicate moderate correlations, whereas positive values indicate a positive correlation and negative values an inverse correlation.<sup>[27]</sup> For the analytes of interest (titer, glucose, and lactate) strong correlations between glucose and glutamine and glucose and TCD were found (0.79 and  $-0.84$  respec-

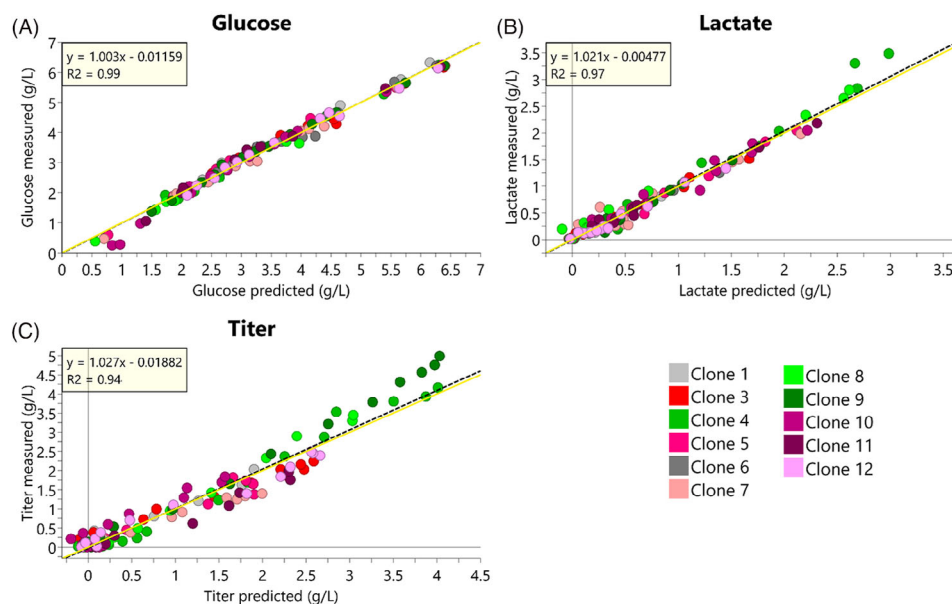
tively). The respective Raman calibration models might lack specificity to the analyte of interest and therefore be confounded and influenced by correlated analytes. Moderate correlations, such as those found for titer and lactate, titer and glutamine, titer and glutamate, lactate and glucose, and glucose and glutamate, may as well affect the respective calibration models. One has to note, that the correlation coefficients alone can merely provide an indication of potential confounding effects. Their severity is further influenced by other factors such as the intensity of the respective Raman scattering signals.

### 3.3 | Initial calibration model building

The initial calibration models for titer, glucose, and lactate, built on the Raman shift range from  $450\text{--}1800\text{ cm}^{-1}$ , showed high prediction accuracy despite the clonal variations encountered within the dataset (Figure 2 and Table 1, loadings plots in Figure S8). Specifically, the initial glucose and lactate models reached  $\text{RMSE}_{\text{CV}}$  values of  $0.17\text{ g L}^{-1}$  (4.9%) and  $0.13\text{ g L}^{-1}$  (18.5%), respectively. These results are comparable to a previous study with a similar experimental set-up (Ambr<sup>®</sup> 250 HT with BioPAT<sup>®</sup> Spectro and HyperFlux<sup>™</sup> PRO spectrometer), achieving comparable  $\text{RMSE}_{\text{CV}}$  for modeling of glucose and lactate ( $\text{RMSE}_{\text{CV}}$  values of  $0.20\text{ g L}^{-1}$  and  $0.23\text{ g L}^{-1}$ , respectively).<sup>11</sup>

Strong correlations between observed and predicted values were found for both models ( $R^2_{\text{CV}} > 0.97$ ). The initial glucose model demonstrated a high  $\text{RMSE}_E/\text{RMSE}_{\text{CV}}$  ratio of 0.95, implying good model generalization.

On the contrary, a low ratio of 0.71, indicating overfitting, was observed for the initial lactate model. Further analysis revealed that the ratio significantly increased to 0.92 upon exclusion of the two observations with the highest lactate concentrations, which exhibited increased prediction errors. Hence, the initial low  $\text{RMSE}_E/\text{RMSE}_{\text{CV}}$



**FIGURE 2** Observed vs predicted (CV) plots of the initial (A) glucose (126 observations), (B) lactate (137 observations), and (C) titer (126 observations) calibration models built on the 450–1800  $\text{cm}^{-1}$  Raman shift range. For all plots, high-producing clones are colored in green shades and low-producing clones in red/violet shades, the identity line is colored in yellow, and the regression line is in dashed black. The corresponding predictive loadings plots are provided in Figure S7.

**TABLE 1** Summary of key parameters for initial calibration models.

Analyte	Range ( $\text{g L}^{-1}$ )	LV	$R^2\text{CV}$	$\text{RMSE}_E$ ( $\text{g L}^{-1}$ )	$\text{RMSE}_{\text{CV}}$ ( $\text{g L}^{-1}$ )	$\text{RMSE}_{\text{CV}}$ (%)	$\text{RMSE}_E/\text{RMSE}_{\text{CV}}$
Glucose	0–6.3	1+3	0.99	0.16	0.17	4.9	0.95
Lactate	0–3.5	1+4	0.97	0.09	0.13	18.5	0.71
Titer	0–5.0	1+4	0.94	0.21	0.29	25.5	0.73

Note: Initial glucose, lactate, and titer models were generated from 126, 137, and 126 observations, respectively. Range states the concentration range covered by the model. LV indicates the number of predictive and orthogonal components;  $R^2\text{CV}$  denotes the cross validated coefficient of determination of the measured and predicted values;  $\text{RMSE}_E$  equals the root mean square error of estimation;  $\text{RMSE}_{\text{CV}}$  denotes the root mean square error of cross validation.  $\text{RMSE}_E/\text{RMSE}_{\text{CV}}$  states the ratio between the  $\text{RMSE}_E$  and the  $\text{RMSE}_{\text{CV}}$ .

ratio was rather caused by large prediction errors due to model extrapolation than overfitting. To avoid extensive extrapolation errors, additional observations with high lactate concentration, preferably spread across multiple cultivations, should be incorporated in the model.

The initial titer model demonstrated slightly lower prediction accuracy.  $\text{RMSE}_{\text{CV}}$  and  $R^2\text{CV}$  values of  $0.29 \text{ g L}^{-1}$  (25.5%) and 0.94 were reached, respectively. Moreover, a low  $\text{RMSE}_E/\text{RMSE}_{\text{CV}}$  ratio of 0.73 was found. The low ratio gives a first indication of model overfit and hints toward clone-related bias of the initial titer model.

### 3.4 | Clone-related bias analysis of initial calibration models

We analyzed the initial calibration models for clone-related bias based on a linear regression analysis, as described in the methods section.

The initial glucose model was not affected by the clonal variations present in the dataset. All clones showed linear regression slopes between 0.91 and 1.09 (Table S3), which are close to the ideal value of one. These were accompanied by a high linearity of prediction ( $R^2\text{CV} > 0.98$ , Table S3). We expected minor deviations in the slopes as found for clones 3 and 10, which can be attributed to the models' prediction errors stemming from noisy Raman spectra and the uncertainties of the off-line reference measurements.

The analytes previously found to be strongly correlated with glucose, namely glutamine and TCD (see 3.2 Reference data analysis), did not notably impair the accuracy of the initial glucose model. On one hand, this could mean that these analytes do not exhibit sufficiently strong Raman scattering signals as compared to glucose resulting in negligible confounding effects. On the other hand, the glucose models might even improve by the correlations, providing that all clones show highly similar underlying correlation structures in the correlated analytes.

The initial lactate model showed no indication of clone-related bias. The analysis resulted in linear regression slopes between 0.89 and 1.01, with exception of clone 4 ( $m = 1.18$ , Table S4), and high linearity of prediction for all clones ( $R^2_{CV} > 0.93$ , Table S4). The slope deviation of clone 4 was attributed to the large prediction errors found for the two observations with the highest lactate levels across all cultivations and are explained by model extrapolation.

The clone-related bias analysis performed on the initial titer model revealed strong deviations from the ideal slope with values ranging from 0.71 to 1.20 (Table S5) with overall high linearity of prediction for all clones ( $R^2_{CV} > 0.92$ ). This implies systematic prediction errors, which indicate a strong clone-related bias of the model. Further analysis revealed a strong trend between the linear regression slopes and clonal productivity. Specifically, high-producing clones exhibited slopes greater than 1.12, while low-producing clones demonstrated slopes below 0.87. The only exception from this trend was low-producing clone 10 with a slope closer to high-producing than to low-producing clones ( $m = 1.06$ ).

It is noteworthy that the TCD profiles (Figure 1) showed a distinct separation of the clones into high and low cell densities after day 9. These groups correspond well with the established productivity groups except for clone 10, whose TCD during that timeframe was closer to high-producing than to low-producing clones. This agrees with the results of the clone-related bias analysis, which showed the slope of clone 10 being closer to high-producing than to low-producing clones and suggests a relation between suspended cells and the found bias. The most influential data points for the titer model development were collected toward the end of the cultivations (highest titer values due to accumulation of the protein in the bioreactor). This coincides with the TCD profile separation into groups of high and low cell densities and might contribute to the observed effect. The initial glucose and lactate models were not affected by these variations in cell densities to such an extent, which might be due to the most relevant samples (highest concentrations) of these models being collected in the beginning of the cultivations.

The clone-related bias analysis revealed that the clonal variations present in the dataset did not affect the initial glucose and lactate models. Highly accurate predictions were demonstrated regardless of the specific clone. Therefore, we state that these models are transferable between clones of a specific cell line, even when these exhibit different clonal characteristics.

For the initial titer model, however, the analysis revealed a strong clone-related bias. This highlights the need for taking clonal variations into consideration when developing calibration models for prediction of titer. Failure to account for clonal characteristics that fall outside the range captured by the calibration model will result in increased prediction errors (as discussed later). This presents a significant challenge for the use of titer Raman calibration models in environments of high clonal variations like cell line development. Furthermore, higher-scale cultivations usually exhibit larger cell counts and productivity than cultivations in miniature bioreactors, which could be limiting the scale-transferability of titer calibration models.

**TABLE 2** Summary of key parameters for initial, IR, VIP, and SR titer models with 1+4 latent variables and 126 observations for each model.

	Initial	VIP	SR	IR
$R^2_{CV}$	0.94	0.95	0.95	<b>0.96</b>
$RMSE_E$ (g L <sup>-1</sup> )	0.21	0.23	0.21	<b>0.19</b>
$RMSE_{CV}$ (g L <sup>-1</sup> )	0.29	0.28	0.29	<b>0.26</b>
$RMSE_{CV}$ (%)	25.5	24.6	24.9	<b>22.2</b>
$RMSE_E/RMSE_{CV}$	0.73	<b>0.80</b>	0.73	0.76

Note: For each metric, the values of the respective best performing model are highlighted.  $R^2_{CV}$  denotes the cross validated coefficient of determination of the measured and predicted values;  $RMSE_E$  equals the root mean square error of estimation;  $RMSE_{CV}$  denotes the root mean square error of cross validation.  $RMSE_E/RMSE_{CV}$  states the ratio between the  $RMSE_E$  and the  $RMSE_{CV}$ .

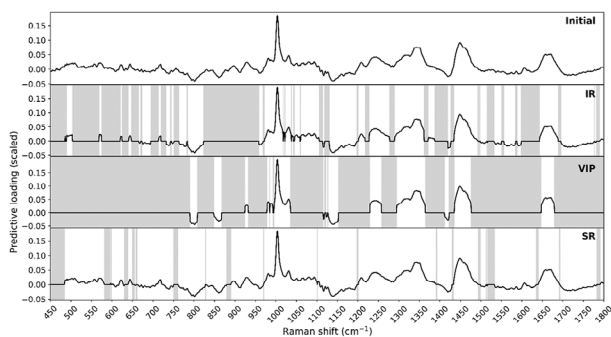
### 3.5 | Titer calibration models after variable selection

The clone-related bias found for the initial titer model could indicate a lack of specificity towards the respective IgG antibody. Low specificity can result in decreased prediction accuracy caused by influences of correlated analytes. We found moderate correlations between titer and lactate, titer and glutamine, as well as titer and glutamate, which could be responsible for confounding effects (see 3.2 Reference data analysis). With the aim to increase the titer models' specificity and reduce potential influences of correlated analytes on the models' predictions, we applied IR, VIP, and SR variable selection methods to the initial titer model.

The resulting models, as summarized in Table 2 and shown in Figure S9, demonstrated similar performance as the initial model. The IR titer model slightly outperformed the other models in terms of  $RMSE_E$ ,  $RMSE_{CV}$ , and  $R^2_{CV}$  metrics. The VIP titer model slightly improved in  $RMSE_{CV}$  and  $R^2_{CV}$  but had worse  $RMSE_E$  as compared to the initial model, resulting in the highest  $RMSE_E/RMSE_{CV}$  ratio overall. The SR titer model showed no significant changes as compared to the initial model.

To compare the titer models before and after variable selection, we analyzed the predictive loadings of the models as shown in Figure 3. The predictive loadings of all titer models included major characteristic bands of IgG Raman spectra<sup>[28,29]</sup> presented in the literature. This includes bands originating from phenylalanine residues (1004 cm<sup>-1</sup>), amide III vibrations (1200–1270 cm<sup>-1</sup>), tryptophan residues (1300–1370 cm<sup>-1</sup>), CH<sub>2</sub> bending (1435–1480 cm<sup>-1</sup>) and amide I vibrations (1645–1685 cm<sup>-1</sup>). Further, the predictive loadings corresponding to these band regions showed high values, indicating their high impact on the titer models. An exemplary Raman spectrum of an IgG molecule is provided in Figure S10.<sup>[28]</sup>

Still, we found significant differences in the selected variables of the tested methods. The IR method effectively selected the characteristic Raman band regions of IgG molecules. The VIP method gave a similar result, though fewer variables with low absolute predictive loadings



**FIGURE 3** Predictive loadings of titer calibration models before and after applying variable selection methods with 126 observations for each model. Variables excluded through variable selection are colored in grey. The corresponding observed versus predicted (CV) plots are depicted in Figures 2C and S9.

were selected as compared to the IR method. The SR method excluded only a few variables, resulting in a model that was comparable to the initial model.

It is noteworthy that only the IR method excluded the Raman band region around  $855\text{ cm}^{-1}$ , which is highly specific for carboxyl acid stretching modes found in lactate, glutamine, and glutamate molecules. The predictive loadings of this Raman band exhibited negative values, indicating negative correlation to titer. This finding is consistent with previous results, which showed moderate negative correlations between titer and lactate, titer and glutamine, and titer and glutamate levels (see 3.2 Reference Data Analysis). However, the band region around  $1420\text{ cm}^{-1}$ , which is as well associated with carboxyl acid stretching modes, was not excluded by any of the variable selection methods. This could be attributed to low signal intensity of the bands or overlapping with IgG antibody associated bands. Nevertheless, we expected a diminished influence of lactate, glutamine, and glutamate on the IR titer model as compared to the initial, VIP, and SR titer models due to the exclusion of the Raman band region around  $855\text{ cm}^{-1}$ .

Indeed, we confirmed these expectations by a correlation analysis of the measured lactate levels and titer predictions errors performed on each titer models. The correlation analysis was restricted to all samples from cultivation days 3 and 4, where most clones reached maximum lactate concentrations. The results, as shown in Figure 4, revealed strong correlation between measured lactate levels and titer predictions errors for the initial and SR titer models ( $R^2 > 0.79$ ). The VIP titer model demonstrated slightly reduced correlation ( $R^2 = 0.75$ ), while the IR titer models exhibited considerably reduced correlation ( $R^2 = 0.33$ ).

Our analysis revealed that the prediction errors for most clones were only slightly affected after IR variable selection, with the exception for clone 4 (Figure 4). This clone considerably differed from the other ones in its daily lactate profile (Figure 1) and thus in the underlying correlation structure between lactate and titer. This explains why excluding lactate associated Raman band regions exhibits a strong impact on the titer prediction errors, particularly for clone 4, and significantly reduces the correlation between measured lactate levels and titer prediction errors.

Our results display that confounded analytes can indeed negatively influence calibration models and their prediction accuracy. The observed effect can be considered as a form of clone-related bias, as it strongly connects to the metabolic characteristics of the clones.

Identifying and excluding variables of confounded analytes is a non-trivial task and a common challenge encountered within multivariate calibration modeling. The challenge was successfully addressed for the titer calibration models by applying the novel IR variable selection method. We attribute the superiority of the method, compared to the other tested ones, to its calculations being based on clone-specific cross validation. In contrast to VIP and SR methods, this allows for the consideration of batch-to-batch (in our case clone-to-clone) variations in the variable selection procedure.

### 3.6 | Clone-related bias analysis of titer calibration models after variable selection

The titer calibration models after variable selection were analyzed for clone-related bias based on a linear regression analysis, as described in the methods section.

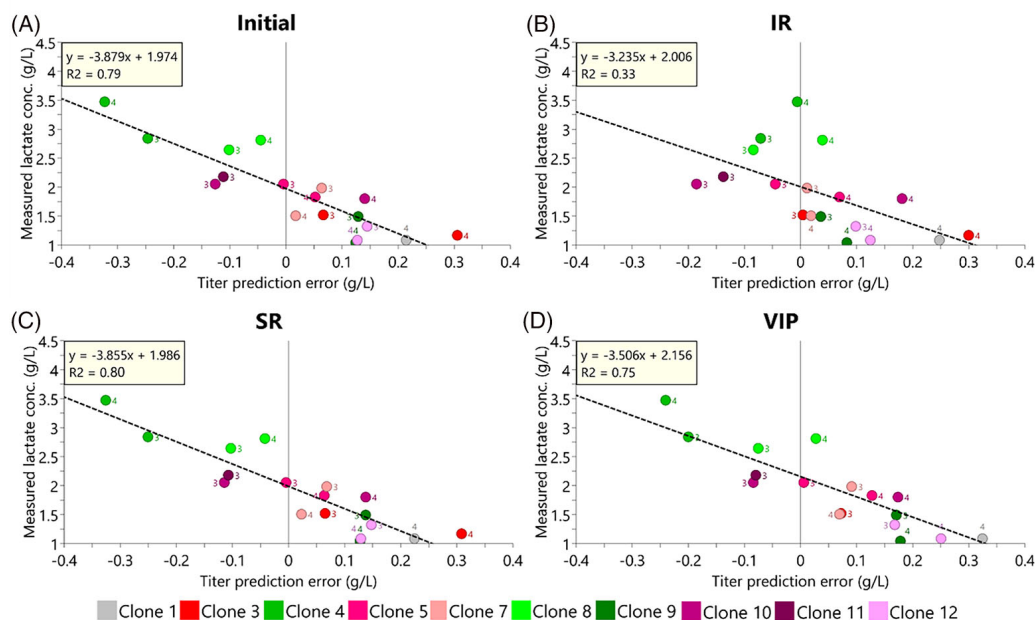
Our results show that the linear regression slopes after variable selections were comparable to the ones obtained for the initial titer model (Tables S5 and S6). As before, low-producing clones exhibited slopes below one, while high-producing clones demonstrated slopes greater than one.

In general, the slopes of low-producing clones increased after IR and VIP variable selection. The only low-producing clone that showed decreased slope after applying these methods was clone 10. We discussed previously that this clone demonstrated a slope closer to high-producing than to low-producing clones for the initial titer model (see 3.4 Clone-related bias analysis of initial calibration models). This exception was removed after VIP and IR variable selection, decreasing the slope of clone 10 from 1.06 for the initial model to 0.88 and 0.97 respectively, which was more consistent to other low-producing clones.

For the high-producing clones, only the slope of clone 4 was significantly affected by VIP and IR variable selections. It was reduced from 1.12 for the initial model to 1.00 and 0.94 for the VIP and IR titer models, respectively. As the cultivation of clone 4 was characterized by increased lactate levels, we expected a strong slope change after applying the IR method due to the exclusion of lactate associated Raman bands.

Generally, the linear regression slopes were shifted towards the ideal value of one after applying IR and VIP methods. This explains the slightly improved performances of the IR and VIP titer models and indicates reduced clone-related bias. This was further supported by the increase of the  $RMSE_E/RMSE_{CV}$  ratios (from 0.72 for the initial model to 0.80 and 0.78 for VIP and IR models respectively) for these models.

The SR method did not considerably affect any of the clones' slopes, which was explained by only few variables being excluded by that method.



**FIGURE 4** Measured lactate versus titer prediction errors (CV) for on cultivation days 3 and 4 for (A) initial, (B) IR, (C) SR, and (D) VIP titer models with cultivation day annotated and regression line marked in dashed black.

Though reduced, the titer models' clone-related bias was still visible after applying variable selection methods. This indicates that either the titer model' specificity was still not sufficient, even after variable selections, or that the clone-related bias had other cause.

### 3.7 | The influence of clonal productivity on titer calibration models

To assess the contribution of clonal productivity to the titer models' clone-related bias, we separated the models into high- and low-producer titer models, as described in the methods section. The models' specifications and performances are given in Tables S7 and S8. We used the high-producer models to predict the titer values of the low-producing clones and vice versa. In this way, the models were confronted with productivity levels outside the range of their respective training data, which represents a worst case scenario.

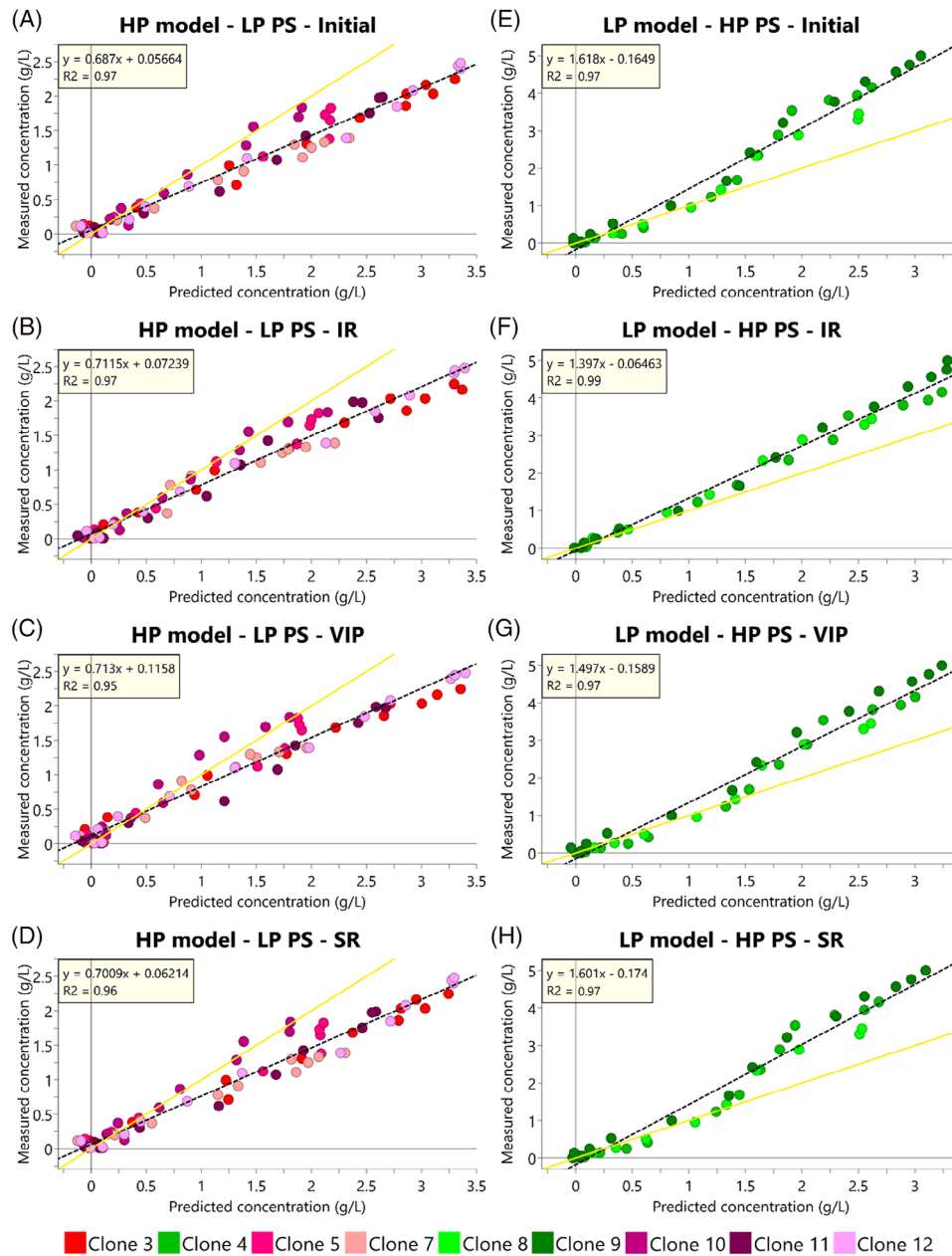
All low-producer models systematically underestimated the titer values of the high-producing clones, while all high-producer models systematically overestimated the titer values of the low-producing clones, as displayed in Figure 5. This resulted in increased RMSE<sub>P</sub> values for the high-producer (0.69 g L<sup>-1</sup>) and the low-producer (0.43 g L<sup>-1</sup>) prediction sets (PS). The large prediction errors for titer of high-producing clones could partially be caused by model extrapolation, as the low-producer models only range to a maximal titer of ~2.5 g L<sup>-1</sup>. However, model extrapolation cannot be the explanation for the strong systematic under-estimation of low-producing clones titer values.

The only exception from the observed trend was the clone 10, whose titer values were predicted with higher accuracy than all other low-

producing clones (RMSE<sub>P</sub> = 0.11 g L<sup>-1</sup> vs. 0.54 g L<sup>-1</sup>, respectively). That clone was similar to the high-producing clones in the daily TCD profile, especially during the last days of the process (Figure 1). This supports the hypothesis that cell density influences the titer models' predictions, as previously discussed in 3.4 Clone-related bias analysis of initial calibration models. This phenomenon has been studied in literature before, showing that suspended cells have significant impact on Raman spectra in general and further on the resulting calibration models.<sup>[30,31]</sup>

Despite the systematic prediction errors, all models demonstrated high linearity of predictions (R<sup>2</sup>PS > 0.95). The highest linearities were found after IR variable selection, resulting in R<sup>2</sup>PS values of 0.97 and 0.99 for prediction of low-producing and high-producing clones, respectively. Particularly for the prediction of high-producing clones, a curvature (S-shape) was found in the observed versus predicted plots of all models, except after IR variable selection (Figure 5). We attributed this curvature mainly to the influence of the confounded lactate analyte on the titer prediction (see 3.5 Titer calibration models after variable selection). This explains the high linearity and superiority of the IR method, which reduced this effect through exclusion of lactate associated Raman bands.

In general, the VIP and IR methods improved the prediction accuracy as compared to high- and low-producer models before variable selection. This is reflected by the decrease in RMSE<sub>P</sub> for the prediction of high-producing clones (from 0.90 g L<sup>-1</sup> to 0.69 g L<sup>-1</sup> for IR and 0.79 g L<sup>-1</sup> for VIP) as well as for low-producing clones (0.49 g L<sup>-1</sup> to 0.43 g L<sup>-1</sup> for IR and 0.40 g L<sup>-1</sup> for VIP). Despite the accuracy improvements, the models after IR and VIP variable selection still exhibited large prediction errors, especially in the higher protein concentration range. The SR method did not yield any significant improvements.



**FIGURE 5** Titer predictions for low-producing (LP) clones using (A) initial, (B) IR, (C) VIP, and (D) SR high-producing (HP) models and predictions for HP clones using (E) initial, (F) IR, (G) VIP, and (H) SR LP titer models. All HP models were generated using 40 observations, while all LP models were based on 77 observations. For all plots, high-producing clones are colored in green shades and low-producing clones in red/violet shades, the identity line is colored in yellow, and the regression line is in dashed black. The corresponding predictive loadings plots of each model are provided in Figure S11.

Compared to the titer models that contained all clones, specifically the  $RMSE_{CV}$  values (Table 2), the prediction errors of the high- and low-producer models increased by a factor of  $\sim 2$  and  $\sim 3$  respectively, regardless of the variable selection method used. This means that the titer models demonstrated significantly reduced prediction accuracies when confronted with productivity levels outside the range of their respective training data. The results of the presented analysis show that clonal productivity contributes most to the titer models' clone-related bias, and that it can't be fully removed with the vari-

able selection approach. Further, variations in the cell density might contribute to the observed effects.

The findings highlight the need for caution when deploying titer calibration models to productivity levels outside the range of their respective training data. The predictions are more accurate when applied to clones with productivity levels captured by the range of the calibration model. These factors need to be considered when developing Raman calibration models for dynamic environments with large clonal variations.

## 4 | CONCLUSIONS

In this work, we developed OPLS Raman calibration models for predicting titer, glucose, and lactate for multiple CHO cultivations of various clones of a single cell line with differences in clonal characteristics. These models were evaluated for any clone-related biases affecting their predictive performances. The assessment is crucial to evaluate the applicability of Raman-based monitoring and control in dynamic environments with large clonal variations, as for example in cell line development.

It was demonstrated that Raman calibration models developed for predicting glucose and lactate were not notably affected by the differences in clonal productivity and growth rates. The generic models were transferable between bioprocess batches of different CHO clones, while consistently demonstrating high prediction accuracies. These results imply that Raman-based monitoring of glucose and lactate in dynamic environments with large clonal variations is feasible. Further, Raman data acquired from different clone cultivations in the cell line development stage should be usable later in calibration model development for application in manufacturing stages.

In sharp contrast to the results for glucose and lactate, it was found that titer calibration models exhibit pronounced clone-related bias associated with clonal productivity. When titer models were applied to cultivations with productivity levels outside the range of their training data, increased prediction errors are found. These errors could not be reduced to satisfactory levels by utilizing variable selection methods. The findings highlight the need for caution when using Raman-based monitoring for titer in dynamic environments with large clonal variations.

To overcome the limitations, titer calibration models should be developed according to the generic calibration modeling approach. Such models would incorporate data from multiple clones of the specific cell line to cover a large range of clonal variations within the model and thereby increasing its robustness. Potentially, historical data from other cell lines (and product molecules) of the same platform process, which should be readily available at biopharmaceutical companies, could be incorporated. Additionally, we propose a hierarchical modeling approach combining discriminant analysis and regressive modeling for initial real-time clustering of clones into groups of high- and low-productivity based on their respective Raman spectra. Following that, concentration predictions could be obtained either from a low- or a high-producer regression model, depending on the outcome of the initial classification. Likely, the separation quality would be lower at the batch start but should significantly gain in accuracy when the cultivation progresses, and the group differences get larger.

### AUTHOR CONTRIBUTIONS

**Rafael Machleid:** Formal analysis; conceptualization; investigation; methodology; validation; visualization; writing—original draft; writing—review & editing. **Marek Hoehse:** Conceptualization; investigation; methodology; supervision; writing—review & editing. **Steffi Scholze:** Data curation; conceptualization; writing—review & editing. **Kleanthis Mazarakis:** Methodology; writing—review & editing. **David**

**Nilsson:** Investigation; writing—review & editing. **Erik Johansson:** Methodology; writing—review & editing. **Christoph Zehe:** Conceptualization; methodology; writing—review & editing. **Johan Trygg:** Project administration; supervision; writing—review & editing. **Christian Grimm:** Project administration; writing—review & editing. **Izabella Surowiec:** Conceptualization; investigation; methodology; validation; project administration; supervision; writing—review & editing.

### ACKNOWLEDGMENTS

The authors thank Pär Jonsson and Lennart Eriksson for providing valuable comments to improve the scientific presentation of the work.

### CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

### DATA AVAILABILITY STATEMENT

The datasets generated during the current study are not publicly available but are available from the corresponding author on reasonable request.

### ORCID

Izabella Surowiec  <https://orcid.org/0009-0002-3709-2458>

### REFERENCES

- Grand View Research. Monoclonal Antibodies Market Size & Share Report, 2030. (2021).
- U.S. Department of Health and Human Services, Guidance for Industry PAT—A Framework for Innovative Pharmaceutical Development, manufacturing, and Quality Assurance (2004).
- Santos, R. M., Kessler, J. M., Salou, P., Menezes, J. C., & Peinado, A. (2018). Monitoring mAb cultivations with in-situ raman spectroscopy: The influence of spectral selectivity on calibration models and industrial use as reliable PAT tool. *Biotechnology Progress*, 34, 659–670.
- Whelan, J., Craven, S., & Glennon, B. (2012). In situ Raman spectroscopy for simultaneous monitoring of multiple process parameters in mammalian cell culture bioreactors. *Biotechnology Progress*, 28, 1355–1362.
- Abu-Absi, N. R., Kenty, B. M., Cuellar, M. E., Borys, M. C., Sakhamuri, S., Strachan, D. J., Hausladen, M. C., & 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, 1215–1221.
- Matthews, T. E., Berry, B. N., Smelko, J., Moretto, J., Moore, B., & Wiltberger, K. (2016). Closed loop control of lactate concentration in mammalian cell culture by Raman spectroscopy leads to improved cell density, viability, and biopharmaceutical protein production. *Biotechnology and Bioengineering*, 113, 2416–2424.
- Graf, A., Lemke, J., Schulze, M., Soeldner, R., Rebner, K., Hoehse, M., & Matuszczyk, J. (2022). A novel approach for non-invasive continuous in-line control of perfusion cell cultivations by raman spectroscopy. *Frontiers in bioengineering and biotechnology*, 10, 719614.
- 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, 486–493.
- Schwarz, H., Mäkinen, M. E., Castan, A., & Chotteau, V. (2022). Monitoring of amino acids and antibody N-glycosylation in high cell density perfusion culture based on Raman spectroscopy. *Biochemical Engineering Journal*, 182, 108426.

10. Sandner, V., Pybus, L. P., McCreath, G., & Glassey, J. (2019). Scale-down model development in Ambr systems: an industrial perspective. *Biotechnology Journal*, 14, e1700766.
11. Graf, A., Woodhams, A., Nelson, M., Richardson, D. D., Short, S. M., Brower, M., & Hoehse, M. (2022). Automated data generation for raman spectroscopy calibrations in multi-parallel mini bioreactors. *Sensors*, 22, 3397.
12. Berry, B., Moretto, J., Matthews, T., Smelko, J., & Wiltberger, K. (2015). Cross-scale predictive modeling of CHO cell culture growth and metabolites using Raman spectroscopy and multivariate analysis. *Biotechnology Progress*, 31, 566–577.
13. Rowland-Jones, R. C., Graf, A., Woodhams, A., Diaz-Fernandez, P., Warr, S., Soeldner, R., Finka, G., & Hoehse, M. (2021). Spectroscopy integration to miniature bioreactors and large scale production bioreactors—Increasing current capabilities and model transfer. *Biotechnology Progress*, 37, e3074.
14. Classen, J., Langer, M., Jockwer, A., & Traenkle, J. (2022). Raman spectrometric PAT models: Successful transfer from minibioreactors to larger-scale, stirred-tank bioreactors. *Bioprocess International*, 20, 34–38.
15. André, S., Lagresle, S., Da Silva, A., Heimendinger, P., Hannas, Z., Calvosa, É., & Duponchel, L. (2017). Developing global regression models for metabolite concentration prediction regardless of cell line. *Biotechnology and Bioengineering*, 114, 2550–2559.
16. Webster, T. A., Hadley, B. C., Hilliard, W., Jaques, C., & Mason, C. (2018). Development of generic Raman models for a GS-KO™ CHO platform process. *Biotechnology Progress*, 34, 730–737.
17. 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, 1004–1013.
18. Santos, R. M., Kaiser, P., Menezes, J. C., & Peinado, A. (2019). Improving reliability of Raman spectroscopy for mAb production by upstream processes during bioprocess development stages. *Talanta*, 199, 396–406.
19. Eilers, P. H. C., & Boelens, H. F. M., *Baseline correction with asymmetric least squares smoothing*. Leiden University Medical Centre Report. (2005). 1.
20. Trygg, J., & Wold, S. (2002). Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics*, 16, 119–128.
21. Wold, S., Martens, H., & Wold, H. *Lecture Notes in Mathematics*. (1982).
22. Johannson, E. A., & Mazarakis, K. (Sartorius Stedim Data Analytics AB), Sweden, EP4036919. (2021).
23. Wold, S., Johannson, E., & Cocchi, M. 3D QSAR in Drug Design; Theory, Methods and Applications, ESCOM. (1993). 523–550.
24. Galindo-Prieto, B., Eriksson, L., & Trygg, J. (2014). Variable influence on projection (VIP) for orthogonal projections to latent structures (OPLS). *Journal of Chemometrics*, 28, 623–632.
25. Farrés, M., Platikanov, S., Tsakovski, S., & Tauler, R. (2015). Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation. *Journal of Chemometrics*, 29, 528–536.
26. Rajalahti, T., Arneberg, R., Berven, F. S., Myhr, K.-M., Ulvik, R. J., & Kvalheim, O. M. (2009). Biomarker discovery in mass spectral profiles by means of selectivity ratio plot. *Chemometrics and Intelligent Laboratory Systems*, 95, 35–48.
27. Ratner, B. (2009). The correlation coefficient: Its values range between +1/-1, or do they? *Journal of Targeting, Measurement and Analysis for Marketing*, 17, 139–142.
28. Kengne-Momo, R. P., Daniel, P., Lagarde, F., Jeyachandran, Y. L., Pilard, J. F., Durand-Thouand, M. J., & Thouand, G. (2012). Protein interactions investigated by the raman spectroscopy for biosensor applications. *International Journal of Spectroscopy*, 2012, 1–7.
29. Ota, C., Noguchi, S., Nagatoishi, S., & Tsumoto, K. (2016). Assessment of the protein–protein interactions in a highly concentrated antibody solution by using raman spectroscopy. *Pharmaceutical Research*, 33, 956–969.
30. Iversen, J. A., Berg, R. W., & Ahring, B. K. (2014). Quantitative monitoring of yeast fermentation using Raman spectroscopy. *Analytical and Bioanalytical Chemistry*. *Anal. Bioanal.*, 406, 4911–4919.
31. Voss, J. P., Mittelheuser, N. E., Lemke, R., & Luttmann, R. (2017). Advanced monitoring and control of pharmaceutical production processes with *Pichia pastoris* by using Raman spectroscopy and multivariate calibration methods. *Engineering in Life Sciences*, 17, 1281–1294.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Machleid, R., Hoehse, M., Scholze, S., Mazarakis, K., Nilsson, D., Johansson, E., Zehe, C., Trygg, J., Grimm, C., & Surowiec, I. (2024). Feasibility and performance of cross-clone Raman calibration models in CHO cultivation. *Biotechnology Journal*, 19, e2300289. <https://doi.org/10.1002/biot.202300289>