A digital tool for the automatic identification of anomalous cell cultures in biopharmaceutical process development

G. Barberi*, P. Diaz-Fernandez, D. Lega**, P. Kotidis**, G. Finka**, P. Facco ***

**CAPE-Lab – Computer-Aided Process Engineering Laboratory, Department of Industrial Engineering, University of Padova, Italy (e-mail: gianmarco.barberi@unipd.it; pierantonio.facco@unipd.it).*
***Biopharm Process Research, Biopharm Product Development and Supply, GlaxoSmithKline R&D, Stevenage, UK*

Abstract: The development of new monoclonal antibodies (mAb) is a long-lasting and expensive procedure. Digital models can be adopted to reduce research costs and accelerate timelines. During mAb development, Ambr®15 is a small-scale, multi-parallel bioreactor platform used to assess performance of different cell lines and find the most productive and stable ones. Many factors affect the culture performance variability and often determine anomalies in the experimental batches. Those anomalies are neither easy, nor fast to be identified even by expert scientists. In this work, a tool for the automatic identification of cell culture anomalies and outlier experimental batches in Ambr®15 scale is presented. The software, calibrated on historical data of the experimental batches, effectively identifies through assumption-free modeling anomalies and diagnoses the root cause of cell lines non-standard behavior, representing also the first application of these methodologies for the development of mAbs. Accordingly, it represents a tool of invaluable importance to speed-up analysis of experimental data and reduce the effort of operators, thus reducing development timeline and costs.

Keywords: monoclonal antibodies, anomaly detection, digital models, cell culture, assumption-free modelling.

1. INTRODUCTION

Biopharmaceuticals are gaining a great interest as innovative highly specialized drugs. Among those, monoclonal antibodies (mAbs) represent the biggest selling class, whose market reached 271 billion dollars in 2021 (Walsh and Walsh, 2022). Monoclonal antibodies are produced in cultures of mammalian cells, commonly Chinese Hamster Ovary (CHO) cells, and are used to treat oncological and immunological diseases.

The development of new mAbs is a long and expensive process, which lasts up to 10 years and requires more than 2 billion dollars of investment. For this reason, biopharmaceutical companies look for innovative solutions to exploit the information available in historical data to accelerate the development of new assets and reduce research costs and timelines (Barberi et al., 2022).

One key step of mAb development is cell line selection, in which cell lines reaching the desired performance in terms of productivity, stability, and product quality are selected. At this stage, a specific cell line, which will also be used at the manufacturing stage, is selected among a pool of originally transfected (i.e., genetically modified cell lines to produce the desired product) clonal cell lines (Li et al., 2010). This procedure is typically performed at different laboratory scales, from nano-well pens, where thousands of cell lines are screened in nanoliter volumes, to shake flasks and high-throughput automated bioreactors (i.e., cL-dL volumes), where only the best performing cell lines are cultured to properly assess their performance at scales and processes that better resemble the production one. One automated bioreactor typically used during cell line screening and selection is Ambr®15 (Rameez et al., 2014). It is a high-throughput bioreactor system with 48 single-use bioreactors, that mimics the performance of larger scale stirred bioreactors. The bioreactors (whose working volume is 15 mL) are equipped with an internal impeller, are individually supplied with gasses and support automatic feeding, sampling, and control of pH and dissolved oxygen. Typically, time profiles of the main process variables (such as nutrients concentration, pH and dissolved oxygen) are available.

During experimental runs in Ambr®15 several events can affect the correct growth and productivity of cells, leading the experimental batches (each one held in one of the 48 single-use bioreactors) to anomalies. Since process variables are extremely correlated, a change in the culture state corresponds to variations in several process variables. The identification of anomalies is typically performed by expert scientists, who inspect the variable time trajectories of each experimental batch to identify the outlier ones. This identification is neither fast nor trivial, if not supported by science-based modelling methodologies. In fact, the main challenge is the fast and reliable inspection of time trajectories of many variables for a large number of batches to discriminate the ones that behave differently from the standard historical cultures.

Multivariate latent variable models represent well-known effective methodologies to monitor the time evolution of batch processes, in order to identify anomalies and

outliers. Typically, multiway principal component analysis (MPCA) and its diagnostics, Hotelling's T^2 and squared prediction error (SPE) are used to uncover process anomalies and the nature or root cause of these anomalies (Joe Qin, 2003). For example, MPCA is used for anomaly detection and diagnosis of a biopharmaceutical drug production process (Zeberli et al., 2021). However, this methodology handles the batch time dimension as additional variables (i.e., batch-wise unfolding of the data) and is effective if all the batches are at each sampling time in the same operating stage and biological state, namely they are characterized by the same chemical, physical and biological phenomena. This does not happen in cell cultures, because culture might grow and progress differently from each other, due to their biological variability. Furthermore, to obtain dynamic diagnostics, online monitoring methods require the artificial completion of the batch, which may introduce artifacts that are not realistic.

To overcome these issues, an assumption-free method for batch monitoring and anomaly detection has previously been developed (Westad et al., 2015). This method considers batches that may be started and concluded in different states, without requiring synchronization. Furthermore, it handles each time point independently from the correlation with the others, thus providing dynamic diagnostics which do not depend on the specific operating stage or biological state. However, application of the assumption-free method in the biopharmaceutical sector and especially in the development of new mAbs is missing. Furthermore, the method has not been tested in applications where only a reduced number of time measurements is available.

This work aims at developing a tool for the identification of anomalies and batch with non-standard behavior (i.e., outlier) in the development of new mAbs at Ambr®15 scale. This method is able to define the standard behavior of the cell cultures and effectively performs the analysis of batch time trajectories, identifying possible outliers and diagnosing the causes of the anomalies in a reliable, fast and automatic manner. The developed tool provides precious support in inspecting the experimental results for cell line selection during mAb development. First of all, it is a objective and science-based method for cell lines selection and informed decision making. Furthermore, being a fast and automated tool to highlight non-standard batches, it relieves the scientists from the manual inspection of many experimental results, allowing to review only the non-standard batches. Hence, it saves scientists time in the analysis of experimental data for product development; thus, supporting the reduction of development timelines.

2. MATERIALS AND METHODS

2.1 Cell culture data

Experimental batches culturing CHO cell lines (GSK proprietary) were run in the Ambr®15 miniature bioreactor system (Sartorius Stedim Biotech, Sartorius AG, Goettingen, Germany) to produce several therapeutic mAbs. All runs were performed for 15 days in fed-batch mode, with glucose as the main carbon source. All process conditions and their changes along the culture (i.e., feeding strategy, pH and temperature set points) are the same across all experimental runs. Variability in the profiles of controlled variables is due to the specificity of each clone and to the timing of controlled actions. $N = 1160$ experimental batches are available for the analysis.

Process variables were measured along the experimental batch at $K = 7$ time points (namely days 0, 3, 6, 8, 10, 13, 15). $V = 10$ process variables are available for the analysis, namely *antibody titer*, *viable cell concentration*, *cell culture viability*, concentration of *ammonium*, *glutamate*, *glutamine*, *lactate*, *LDH*, *pH*, and *dissolved oxygen (DO)*. Batch data are organized in a three-dimensional array $\underline{\mathbf{X}}$ $[N \times V \times K]$.

To test the proposed method for the identification of anomalies and outliers, a single Ambr®15 run (i.e., 48 test batches) is randomly selected and used as external validation set, while the remaining 1112 batches are used to calibrate the model. Testing batches are left aside before any processing and analysis.

2.2 Data pre-processing

Missing values along the process variable time profiles are imputed by linear interpolation. The remaining missing values (e.g., at the beginning of the culture) are imputed using the column-wise mean value across all batches. This is done to reduce the impact of the imputed values on the multivariate model.

Since experimental batches comprise data from cell lines expressing different mAbs, data were scaled within each experiment to allow multivariate comparison among process variable time profiles. In this work, data were scaled to zero mean and unit variance.

Test batches are preprocessed similarly to calibration ones. Missing value imputation is performed using the column-wise mean value calculated over calibration batches.

2.3 Multiway PCA

Multiway Principal Component Analysis (MPCA; Nomikos and MacGregor, 1994) is used as a dimensionality reduction method to identify anomalies and outlier batches (i.e., batches with non-standard behavior) while properly handling the time-dimension of the batch data.

In MPCA data is unfolded prior PCA modelling. Variable-wise unfolding is used in this work. It retains the time behavior of batches along the observation direction, allowing to study time-averaged correlations across variables. This specific handling of the time-dimension of data is required by the assumption-free method used for the identification of outliers. In variable-wise unfolding, data at each time points k, $\mathbf{X}_k$ $[N \times V]$, are concatenated vertically (i.e., along the variable dimension) to generate the unfolded bi-dimensional matrix $\mathbf{X}$ $[N \cdot K \times V]$.

PCA (Jolliffe, 2022) is then applied. It is a multivariate technique that captures the direction of maximum variability of and the correlation among variables of $\mathbf{X}$ scaled data. It projects the data into a reduced space of A principal components (PCs; where $A \ll V$) as:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathbf{T}} + \mathbf{E} \quad , \tag{1}$$

where $\mathbf{T}$ $[N \cdot K \times A]$ is the score matrix, $\mathbf{P}$ $[V \times A]$ is the loading matrix and $\mathbf{E}$ $[N \cdot K \times V]$ is the residual matrix which is minimized in the least-square sense. In MPCA, scores capture the relationship between batches along their time evolution, while loadings capture the time-averaged correlation among process variables.

Squared Prediction Error (SPE; Nomikos and MacGregor, 1995) is used to assess if observations (i.e., any batch at time point) follow the correlation structure captured by the MPCA model. SPE is defined as:

$$SPE_n = \mathbf{e}_n^{\mathrm{T}}\mathbf{e}_n \quad , \tag{2}$$

where $\mathbf{e}_n$ is the residual vector for an observation n from the MPCA model (i.e., a row of matrix $\mathbf{E}$). One-sided confidence limit can be calculated for the SPE diagnostic as:

$$SPE_{lim} = \frac{\sigma}{2\mu}\chi^2_{2\mu/\sigma,\alpha} \quad , \tag{3}$$

where μ and σ are the mean and variance of the SPE, respectively, χ^2 is a chi-square distribution with $2\mu/\sigma$ degrees of freedom at a confidence level α.

2.4 Assumption-free modeling

Assumption-free modelling (Sartori, 2024; Westad et al., 2015) is used in this work to identify anomalies in batch time progression because it can handle batches which: *i*) do not start in the same state (i.e., seeding and feeding conditions might vary), *ii*) have a different state of the final sampling point (i.e., due to a different degree of batch evolution), and *iii*) follow a different time progression.

The assumption free algorithm considers the score matrix $\mathbf{T}$ to identify the common batch trajectory and dynamic confidence limits that identify outlier batches. The algorithm goes through the following steps:

1. Use a grid-search algorithm to build a grid of cells in the score space of the MPCA model (Sartori, 2024). The algorithm tests different space subdivisions and identifies the best one, namely the one which maximizes the number of valid grid cells. A grid cell is valid if it contains scores from at least a fraction β of all batches. Grid cells must be valid because they are used to determine the common batch trajectory; hence, they should be representative of a sufficiently large portion of batches. The algorithm also ensures that at least a fraction γ of all MPCA scores are contained in valid cells to guarantee that the identified grid of cells appropriately capture the MPCA scores. In this work, $\beta = 0.5$ and $\gamma = 0.88$ were heuristically set during preliminary studies. The parameters should be tuned to ensure that the resulting common batch trajectory properly follows the time evolution of batches. The automatic identification of these parameters is object of further studies.
2. Calculate the mean of all scores and the means of individual batches (irrespectively of their time) within a valid cell. Multiple time points of a batch can be contained in a cell, as it might have a time evolution which is different from the common one. The mean of all scores is the average state of the batches in a cell, and it is thereafter named cell center.
3. Determine the common batch trajectory by linear interpolation of all the centers of valid cells.
4. Project the mean of the individual batches onto the average batch trajectory and estimate the distance statistic. The distance statistic for an individual batch mean b is calculated as:

$$d_b = \sqrt{\sum_{a=1}^{A}(t_b - t_b\ )^2} \quad , \tag{4}$$

   where t_b is the individual batch mean and t_b is the projection of the individual mean onto the average batch trajectory.
5. Estimate the confidence limit for the population of individual batch distance statistics within each valid cell. This confidence limit will be used to identify outlier observations (i.e., a time point of a batch) within each cell.
6. Calculate the SPE for all batch observations within each valid cell and determine its confidence limit (2-3). This confidence limit will be used to identify in each cell batch observations that do not follow the correlation structure captured by the model.

When a new batch is available the following steps are performed:

1. Preprocess the new batch;
2. Project the new batch into the MPCA model $\mathbf{t}_{new} = \mathbf{x}_{new}\mathbf{P}$;
3. Estimate the distance statistic and SPE to identify possible outliers.

2.5 Outlier detection and diagnosis

The anomalies detection is performed using the dynamic control limits of SPE and distance from the average batch trajectory (i.e., distance statistic) calculated through the assumption-free modelling (Section 2.4). Outlier detection identifies batch that deviates from the average batch trajectory (i.e., large distance) and/or does not conform with the correlation structure captured by the MPCA model (i.e., large SPE). Accordingly, a batch is considered faulty if a defined number of observations violates either the SPE or the distance confidence limit in the score space. The number of consecutive observations outside the

confidence limits can be tuned to minimize false positives (Rato et al., 2016). However, in this work, due to the small number of time points, a batch is identified as an outlier if a single time observation is outside the confidence limit.
Outlier diagnosis (i.e., identification of the root causes of the anomalies) is performed using the contributions to the distance from the average batch trajectory and the SPE. Contributions to the SPE for an observation n are provided by the residual vector $\mathbf{e}_n$. Distance contributions, instead, are calculated as:

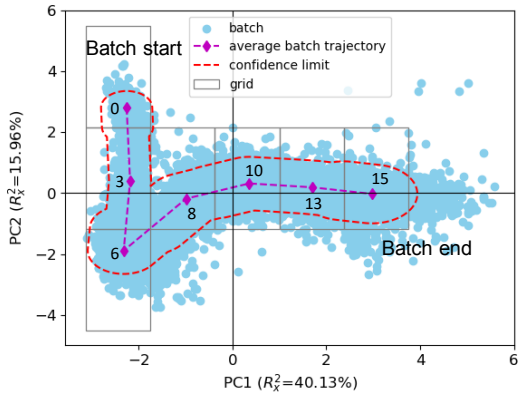
$$\mathbf{d}_{cont,n} = \mathbf{t}_n \mathbf{P}^{\mathbf{T}} \quad . \tag{5}$$

To properly assess the deviations of a batch with respect to the average batch trajectory, the relative contributions of the distance statistic are calculated. Two-sided confidence limit for contributions can be calculated from the population of contributions resulting during the calibration phase.

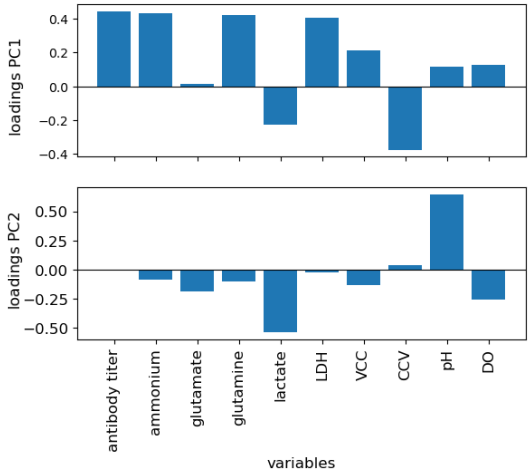
3. RESULTS AND DISCUSSION

3.1 Model calibration for the detection of outlier cell-lines

To perform the detection of outlier batches, both the MPCA and the assumption-free models have to be calibrated. However, not all the N available batches are representative of standard batch conditions, due to the intrinsic difficulty in defining standard conditions for mAb cell cultures, and the absence of batch labeling (standard/outliers) and normal operating conditions for cell cultures. Accordingly, to identify the standard batches among the available ones, a first global version of both MPCA and assumption-free model are built on all the available data, and the calculated distance statistic and SPE limits are used to exclude calibration batches with non-standard conditions. In this way, the model can be calibrated on well-behaving batches, excluding unknown outliers, which cannot be otherwise determined due to the absence of ground truth labels. Batches with at least one observation outside 2 times the distance confidence limit and 2.5 times the SPE confidence limit are considered as non-standard and are accordingly excluded. The remaining $N' = 994$ batches are considered to operate in standard conditions and are used to calibrate the model. These thresholds have been heuristically selected (during preliminary studies not showed) as they exclude batches with highly anomalous behavior, without limiting excessively the variability of the historical data. However, the presence of ground truth labels for batches would have strengthened the capability of the model.
MPCA captures with 2 PCs the 56.1% of $\mathbf{X}$ variability. Figure 1a reports the score space of the calibration data with the average batch trajectory (purple line), the confidence limit of the distance within the model (red line), and the optimal grid. The time behavior of batches is captured in the score space. In fact, batches start in the top-left quadrant and move in time toward the bottom-left quadrant and then the right part of the score space. The confidence limit identifies the region where standard batches are located. Note that the presence of some observations outside the confidence limit is physiological, because the 95% confidence limit leaves 5% of observations outside by definition.



(a)



(b)

Fig. 1. Results of MPCA and assumption free section calibration: (a) score plot with average batch trajectory, distance confidence limit, and grid; (b) loading plot of the MPCA model.

The loading plot reports the correlation structure among variables, averaged throughout the entire batch duration, captured by the model (Figure 1b) used for the identification of anomalies and outlier batches. The first PC captures the positive correlation between variables *antibody titer*, *ammonium*, *glutamine*, *LDH*, and *viable cell concentration*, which are anticorrelated to *lactate* and *cell culture viability*. Accordingly, observations (i.e., batch at each time point) located at positive values of PC1 are characterized by high values of *antibody titer*, *ammonium*, *glutamine*, *LDH* and *viable cell concentration*, and by low values of *lactate* and *cell culture viability*.
The second PC, instead, mainly captures the anticorrelation between *lactate* and *pH*. This anticorrelation is physically reasonable, because it captures the physical effect of *pH* reduction when the acid in the culture (i.e., *lactate*) increases. According to this

anticorrelation, batches with positive value of PC2 are characterized by low values of *lactate* and high value of *pH*. Other minor effects indicate that these batches are also characterized by low value of *viable cell concentration* and *DO*.

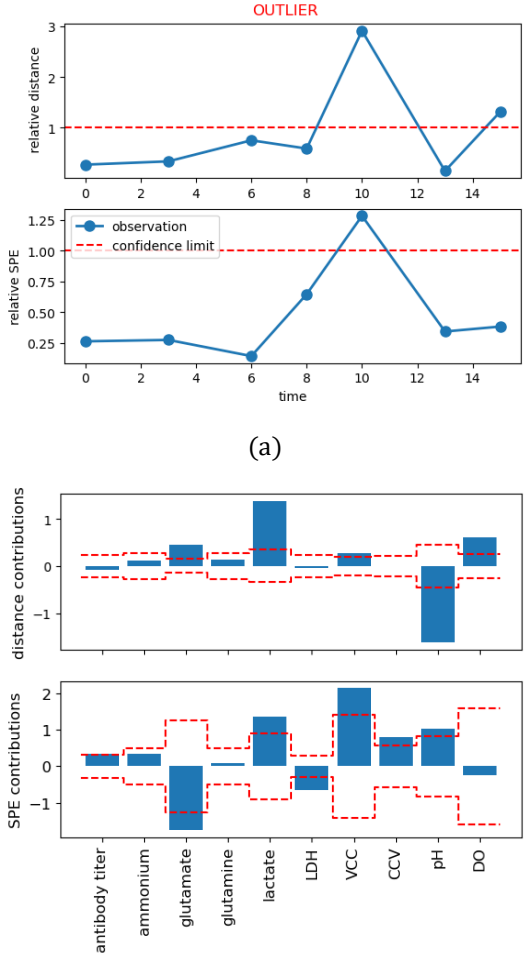


Fig. 2. Results of outlier detection and diagnosis for test batch #32: (a) time trajectory of relative distance statistic and relative SPE, with confidence limit; (b) distance and SPE contributions with confidence limits at day 10 for root cause diagnosis.

The joint analysis of scores and loadings indicates that at the beginning of the culture batches are characterized by low value of *antibody titer*, *ammonium*, *glutamine*, *LDH*, *viable cell concentration*, *lactate*, and by high values of *cell culture viability* and *pH*, because they are located in the top-left quadrant (negative PC1 and positive PC2 values). With the growth phase, the state of batches moves towards the bottom-left quadrant (negative PC2 values), which indicates an increase in *lactate* and *viable cell concentration* and a corresponding decrease in *pH*, while all other culture variables are not subject to large variations. Then, during stationary and decline phases, batches move towards the right side of the score space with PC2 ~0. This indicates that the progression through stationary and decline phases induces an increase in *antibody titer*, *ammonium*, *glutamate*, *LDH*, *viable cell concentration*, and a decrease in *cell culture viability* and *lactate*, while *pH* remains on intermediate values.

The batch behavior captured by the MPCA model is coherent with the expected time evolution of mAb cell cultures, which are subject to the accumulation of some compounds, such as *antibody titer*, *ammonium*, *glutamate*, together with an increase in *viable cell concentration* and a decrease in *cell culture viability*. *Lactate*, instead, is initially produced by cells, and subsequently consumed after the lactate shift. According to these results, the MPCA model can be used to identify outlier batches, as the ones that deviate from this standard behavior.

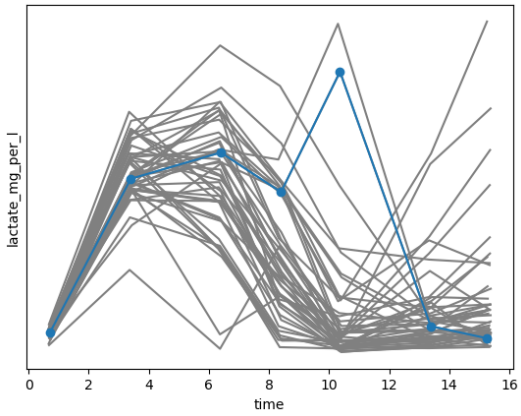
3.2 Detection of outlier cell lines and root causes analysis

The developed tool is tested on 48 validation batches not included in calibration to identify outliers and anomalies in the testing set. Test batches follow the following analysis procedure (Section 2.4): *i*) preprocessing; *ii*) projection onto the MPCA model; *iii*) calculation of distance statistic and SPE; *iv*) outlier detection and diagnosis.

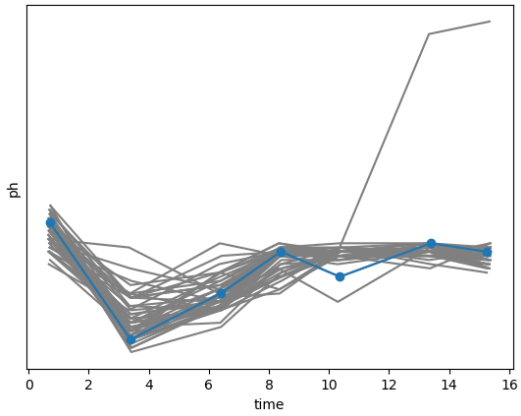
The tool identifies 25 test batches over 48 as outliers. These batches are highlighted as having at least one time observation outside the distance statistic or the SPE confidence limit. The large number of outliers, identified in this case, is expected since the testing Ambr®15 run presents large variability in the main process trajectories (Figure 3). Furthermore, due to the absence of ground truth labels for outliers, the accuracy and false discovery rate of the method cannot be determined. The identified batches, together with the root cause of the non-standard behavior are provided to expert developers for further analysis. As an example, the analysis of the batch #32, which is identified as an outlier, is presented. The distance statistic and SPE over the batch time evolution is reported in Figure 2. Note that the Figure is represented in terms of relative distance (i.e., the ratio between the distance *d* and the respective limit) and relative SPE (i.e., the ratio between SPE of the observation and the respective limit), so the confidence limit is equal to 1.

This representation is used because the confidence limits change according to the position of the observation in the score space. Test batch #32 is identified as an outlier because it exceeds the confidence limit of distance statistic at days 10 and 15 and the confidence limit of SPE at day 10. Accordingly, at day 10 the batch is strongly different from the standard conditions of the Ambr®15 batches (given by the average batch trajectory), but at the same time it does not follow the correlation structure identified by the model, being outside the SPE confidence limit.

The final state of the batch (at day 15) is statistically different from the standard state of batches. In particular, at this time point, the batch has progressed further than standard conditions, showing slightly higher values of *antibody titer*, *ammonium*, *glutamine*, *LDH*, and lower values of *lactate* and *cell culture viability*.



(a)



(b)

Fig. 3. Relevant process variables time profile for test batch #32 (blue line – batch #32; gray lines - time profiles of the other test batches): (a) *lactate* concentration; (b) *pH*. Axis values are anonymized for confidentiality reasons.

The root cause diagnosis of the non-standard behavior is performed by analyzing the diagnostic contribution plot at the relevant time points (i.e., where the batch exceeds a confidence limit). The diagnostics contributions plot at day 10 for the test batch #32 is reported in Figure 2b, together with the respective confidence limits. The batch shows values *glutamate*, *lactate*, *viable cell concentration*, and *DO* higher than standard batches, while lower values of *pH* (from contribution of the distance statistics). Furthermore, the same variables with the addition of LDH make the batch not following the correlation structure of the MPCA (i.e., is outside SPE limit).

The behavior highlighted by the software can be observed in the process variables time profiles (Figure 3). Batch #32 shows a spike in *lactate* concentration at day 10 (Figure 3a), which is not a standard behavior, as can be observed also in the other test batches (gray lines). At the same time, the *lactate* spike produces an increased culture acidity, resulting in a corresponding decrease of *pH* (Figure 3b).

4. CONCLUSIONS

In this work, we presented a tool for the detection of anomalies and outliers to support the development of new monoclonal antibodies. The tool uses multivariate and assumption-free methods to identify batches showing anomalies in their process trajectories at Ambr®15 scale. Accordingly, the detection of anomalies is accomplished on all process variables together, not just considering the single variables in a univariate fashion. Being calibrated on historical Ambr®15 data, even without knowing the standard conditions for cell cultures or ground truth labels for outliers, the tool is able to highlight outlier batches to expert scientists and provide an indication of the root cause of the anomaly.

The developed tool supports the digitalization of the development of mAbs, allowing a fast and automatic evaluation of Ambr®15 batch time trajectories. This simplifies and accelerates the analysis conducted by scientists, reducing the experimental effort and providing a science-based tool to support decision making during the development and approval of new mAbs.

REFERENCES

Barberi, G., Benedetti, A., Diaz-Fernandez, P., Sévin, D.C., Vappiani, J., Finka, G., Bezzo, F., Barolo, M., Facco, P., 2022. Integrating metabolome dynamics and process data to guide cell line selection in biopharmaceutical process development. Metabolic Engineering 72, 353–364. https://doi.org/10.1016/j.ymben.2022.03.015

Joe Qin, S., 2003. Statistical process monitoring: basics and beyond. Journal of Chemometrics 17, 480–502. https://doi.org/10.1002/cem.800

Jolliffe, I.T., 2022. Principal Component Analysis. Springer, New York.

Li, F., Vijayasankaran, N., Shen, A. (Yijuan), Kiss, R., Amanullah, A., 2010. Cell culture processes for monoclonal antibody production. mAbs 2, 466–479. https://doi.org/10.4161/mabs.2.5.12720

Nomikos, P., MacGregor, J.F., 1995. Multivariate SPC Charts for Monitoring Batch Processes. Technometrics 37, 41–59. https://doi.org/10.1080/00401706.1995.10485888

Nomikos, P., MacGregor, J.F., 1994. Monitoring batch processes using multiway principal component analysis. AIChE Journal 40, 1361–1375. https://doi.org/10.1002/aic.690400809

Rameez, S., Mostafa, S.S., Miller, C., Shukla, A.A., 2014. High-throughput miniaturized bioreactors for cell culture process development: Reproducibility, scalability, and control. Biotechnology Progress 30, 718–727. https://doi.org/10.1002/btpr.1874

Rato, T.J., Rendall, R., Gomes, V., Chin, S.-T., Chiang, L.H., Saraiva, P.M., Reis, M.S., 2016. A Systematic Methodology for Comparing Batch Process Monitoring Methods: Part I—Assessing Detection Strength. Ind. Eng. Chem. Res. 55, 5342–5358. https://doi.org/10.1021/acs.iecr.5b04851

Sartori, F., 2024. EFFECTIVE IMPLEMENTATION OF INDUSTRY 4.0 APPROACHES FOR PROCESS MONITORING IN THE BATCH MANUFACTURING OF SPECIALTY CHEMICALS (Ph.D. Dissertation). University of Padova.

Walsh, G., Walsh, E., 2022. Biopharmaceutical benchmarks 2022. Nat Biotechnol 40, 1722–1760. https://doi.org/10.1038/s41587-022-01582-x

Westad, F., Gidskehaug, L., Swarbrick, B., Flåten, G.R., 2015. Assumption free modeling and monitoring of batch processes. Chemometrics and Intelligent Laboratory Systems 149, 66–72. https://doi.org/10.1016/j.chemolab.2015.08.022

Zeberli, A., Badr, S., Siegmund, C., Mattern, M., Sugiyama, H., 2021. Data-driven anomaly detection and diagnostics for changeover processes in biopharmaceutical drug product manufacturing. Chemical Engineering Research and Design 167, 53–62. https://doi.org/10.1016/j.cherd.2020.12.018