Received: 30 April 2023 | Revised: 12 September 2023 | Accepted: 21 September 2023

DOI: 10.1002/bit.28571

ARTICLE

BIOTECHNOLOGY AND BIOENGINEERING WILEY

Enhanced harvest performance predictability through advanced multivariate data analysis of mammalian cell culture particle size distribution

Martina Sebastian[1] | Stephen Goldrick[1] | Matthew Cheeks[2] | Richard Turner[3] | Suzanne S. Farid[1]

[1]Department of Biochemical Engineering, University College London, London, UK

[2]Cell Culture & Fermentation Sciences, Biopharmaceuticals Development, R&D, AstraZeneca, Cambridge, UK

[3]Purification Process Sciences, Biopharmaceuticals Development, R&D, AstraZeneca, Cambridge, UK

Correspondence
Suzanne S. Farid, Department of Biochemical Engineering, University College London, Gower St, London WC1E 6BT, UK.
Email: s.farid@ucl.ac.uk

Funding information
AstraZeneca, Grant/Award Number: EP/L01520X/1; Engineering and Physical Sciences Research Council

Abstract

The industry's pursuit for higher antibody production has led to increased cell density cultures that impact the performance of subsequent product recovery steps. This increase in cell concentration has highlighted the critical role of solids concentration in centrifugation yield, while recent product degradation cases have shed light on the impact of cell lysis on product quality. Current methods for measuring solids concentration and cell lysis are not suited for early-stage high-throughput experimentation, which means that these cell culture outputs are not well characterized in early process development. This article describes a novel approach that leveraged the data from a widely-used automated cell counter (Vi-CELL™ XR) to accurately predict solids concentration and a common cell lysis indicator represented as lactate dehydrogenase (LDH) release. For this purpose, partial least squares (PLS) models were derived with k-fold cross-validation from the particle size distribution data generated by the cell counter. The PLS models showed good predictive potential for both LDH release and solids concentration. This novel approach reduced the time required for evaluating the solids concentration and LDH for a typical high-throughput cell culture system (with 48 bioreactors in parallel) from around 7 h down to a few minutes.

KEYWORDS
automated cell counter, centrifugation, monoclonal antibodies, multivariate data analysis (MVDA), particle size distribution, shear

1 | INTRODUCTION

Therapeutic monoclonal antibodies (mAbs) are one of the leading classes of therapeutics accounting for one in five FDA approvals with the 100th mAb approved by the FDA in 2021 and expected revenue of $300 billion by 2025 (Lu et al., 2020; Mullard, 2021). Mammalian cell lines are used to produce mAbs, with Chinese hamster ovary (CHO) cells established as the most commonly used expression system (Khan, 2013). During cell culture and primary recovery stages of a mammalian cell-based mAb manufacturing process, it is

Abbreviations: CSD, capillary shear device; DOE, design of experiment; LDH, lactate dehydrogenase; PCV, packed cell volume; PLS, partial least squares; PRESS, predicted residual error sum of squares; PSD, particle size distribution; TCD, total cell density; VCD, viable cell density; VIP, variable importance on projection.

important to monitor solids concentration and the degree of cell membrane lysis since these can impact yield and product quality. However, current methods for measuring solids concentration and cell lysis are often time- and resource-consuming and hence rarely performed in early-stage high-throughput experimentation when screening multiple conditions. The aim of this article is to develop reliable and rapid tools for evaluating cell lysis and solids concentration that will be compatible with high-throughput small-scale experimentation platforms deployed for early process development.

Primary clarification of mammalian cell culture is usually achieved via centrifugation. Increases in cell culture solids concentration result in higher product losses from the disc-stack centrifugation operation due to the need for more frequent discharges. Hence, it is important to be able to measure solid concentration load in a timely manner to be able to predict yield more accurately and help design more efficient processes.

Another important cell culture parameter is the loss of membrane integrity or the degree of cell lysis, since it can lead to the release of intracellular species with detrimental impact on product quality such as product degradation via antibody reduction at the disulfide bridge (Aucamp et al., 2014; Handlogten et al., 2017, 2020; Hutterer et al., 2013; Kao et al., 2010; O'Mara et al., 2019; Tebbe et al., 1996). The cells can lose membrane integrity towards the end of cell culture as the cell culture viability drops. Cells can also lyse due to mechanical forces such as the shear experienced at the centrifuge feedzone, where the feed is accelerated to the bowl speed (Joseph et al., 2017). Rapid detection of cell lysis at the end of cell culture prior and post centrifugation would help predict and understand the conditions that lead to product degradation and devise informed strategies to minimize quality deviations.

The impact of solids concentration on centrifugation yield and cell lysis on product quality highlights the need to evaluate these measurements early on in process development as it allows the optimum process conditions to be identified aiding in tech-transfer and scale-up activities. However, at this stage of process development, studies are usually carried out using high-throughput cell culture systems and due to sample volume and time constraints the analysis is limited to cell counts via automated cell counters, high throughput metabolites analytics and in-process samples are frozen for later analysis.

Current methods for evaluating cell lysis and solids concentration are usually not compatible with small-scale, high-throughput experimentation. The cell culture solids concentration is typically defined empirically using packed cell volume measurements (Li et al., 2010). It is a manual measurement requiring triplicates and at least 40 µL per sample, which makes it cumbersome for routine use with high-throughput cell culture. A common approach for measuring cell lysis is by measuring the levels of lactate dehydrogenase (LDH) release. LDH is an ubiquitous intracellular enzyme found in nearly all mammalian cells. Its release in the cell culture media indicates that the integrity of the cell membrane has been compromised (Riss et al., 2004). Although absolute LDH can be measured with high throughput methods, it suffers from being sensitive to media composition and culture operating conditions (Hiebl et al., 2017; Méry et al., 2017). Therefore, it is common to measure LDH release based on the ratio of the sample's LDH relative to the total LDH released upon chemical lysis (Cummings & Schnellmann, 2004). The need for triplicates, additional cell lysing step and time to read the samples increases the experimental workload for the measurement.

Another approach for evaluating cell lysis is through viability measurements. The levels of cell culture viability can be defined as the ratio of dead-to-live cells, where dead cells are considered cells that have lost plasma membrane integrity or have broken into discrete fragments (Galluzzi et al., 2015). Typically, cell culture viability can be determined through using a live/dead dye for counting live and dead cells. The trypan blue staining method is a gold standard for routine viability measurements; it uses the permeability of the dead cells to the dye to differentiate between viable and nonviable cells. Nowadays, cell counts based on this method are often performed using automated cell counters (Cedex; Innovatis and Vi-Cell™ XR, Beckamn Coulter), which provide cell counts, viability and particle size distribution data (Li et al., 2010). This method relies on the ability of the instrument to accurately count the number of dead and live cells based on the images. However, dead cells may have completely disintegrated or shrunk and not be accounted for.

The aim of this article is to develop soft sensors that leverage existing data available from cell counters to accurately and rapidly predict solids concentration and cell lysis without increasing the experimental burden and thus facilitating the analysis for high throughput systems. This will equip manufacturers with tools to gain insights about the impact of upstream conditions and the harvest operation on process performance with the goal of designing upstream and recovery processes, which currently is only possible for process development efforts at pilot scale. To reduce the experimental burden, the methodology described in this article leverages the routinely collected particle size data (PSD) from the automated cell counter (Vi-CELL™ XR) and advanced statistical techniques to predict the solids concentration and cell lysis.

In this study, first the relationship between LDH release and viability was investigated to see if the viability can be correlated to LDH release for fresh cell culture samples as well as for samples that are lysed due to stress similar to the levels experienced in the centrifuge feedzone. A thorough analysis of the automated cell counter PSD data was performed for different cell culture conditions and day of harvest. A novel method was applied that successfully leveraged the PSD using partial least squares (PLS) models to predict the levels of LDH release and solids concentration leading to significant time and resource savings compared to traditional methods. The PLS algorithm was selected here based on its proven ability to extract useful relationships between complex biological data sets (Banner et al., 2021). This method was integrated in an overall harvest toolbox that is compatible with upstream small scale experimentation and uses a centrifugation scale-down mimic to characterize different harvests in terms of LDH release and product quality.

2 | MATERIALS AND METHODS

2.1 | Cell culture preparation

2.1.1 | Cell culture process

CHO cell lines were utilized in this study, that were either nonproducing (Null) one or expressing a number of different mAbs defined as projects A, B, C and D. Table 1 summarizes the different conditions of the experiments, including their scale, level of shear, harvest day and seeding density. Chemically defined media was used in all bioreactors, with a two-part proprietary feed addition every other day, between Day 2 and 12. The temperature was maintained at 35.50°C and pH was set to 6.95 (lower bound = 6.9 and upper band = 7.10) for the standard process. In some cases pH shift, temperature shift and different pH set points were applied (Table 1). Cell culture material from various scales was used in this study, with focus on ambr^{TM}15 and bench-scale 7 L bioreactors, and some additional samples taken from pilot scale 50 and 200 L single use bioreactors. The ambr^{TM}15 systems in this study used 24 and 48 single micro-bioreactors. Each vessel has its individual pH and DO control and gas supply, while temperature and impeller rotational speed are controlled per culture station. Detailed information about the system can be found in Hsu et al. (2012). The bioreactors' working volume was maintained between 11 and 16 mL and vessels were seeded at different target densities through varying the inoculum concentrations and harvested on different days. The bench-scale experiments were performed in 7-L glass bioreactors (Chemglass) with approximately 5-L working volume, controlled by DASGIP® (Eppendorf, DASGIP, Germany) or ez-Control® (Applikon Biotechnology B.V.) controllers. Feeds for all scales were through bolus additions.

2.1.2 | Analytical methods for cell culture daily monitoring

Viable cell density and total cell density were measured in the Vi-CELL™ XR (Beckman Coulter) for all cell cultures. Dilution (1 in 5) was applied after Day 5 (at approximately more than 10^6 cells/million). For the ambr^{TM}15 runs glucose and lactate were determined off-line by YSI 2000 (YSI, Yellow Spring). A more comprehensive metabolite analysis was carried out for the bench scale, including glutamate, ammonia and potassium in addition to glucose and lactate.

TABLE 1 Summary of experimental data used for the PSD analysis and to build the PLS models for LDH release and [solids] with scales, level of shear, harvest day and seeding density conditions.

	Experiment#	Project ID	Vessel ID	Shear	Harvest day	Seeding density[a]
(A)			**LDH release with PSD analysis**			
Training data set	1–12	Project Null	Bench scale 7 L BR 1	No, Low, High	8, 9, 10, and 12	Mid point
	13–24	Project A	Bench scale 7 L BR 2	No, Low, High	8, 9, 10, and 12	Mid point
	25–45	Project A	ambr1	No	11, 13, and 14	Low, Mid, High
	46–63	Project A	ambr2	No	10 and 12	Low, Mid, High
	64	Project B	Bench scale 7 L BR 3	No	13	Mid point
	65–67	Project B	Pilot BR 1	No	12 and 13	Mid point
Test data set	68–72	Project C	Bench scale 7 L BR 4	No, Low, High	12	Mid point
	73–78	Project C	Pilot BR 2	No, Low, High	11 and 12	Mid point
(B)			**Solids concentration with PSD analysis**			
Training data set	1–19	Project A	ambr1	N/A	11, 13, and 14	Low, Mid, High
	20–68	Project A and Project D	ambr3	N/A	8, 13, and 15	Low, Mid, High
	69–87	Project D	ambr4[b]	N/A	12, 13, and 14	Low, Mid, High
	88–107	Project D	ambr5[c]	N/A	13, 14, and 15	Low, Mid, High
Test data set	108–114	Project A	Bench scale 7 L BR 2 & 3	N/A	11, 12, and 13	Mid point
	115–128	Project D	ambr3	N/A	8, 13, and 15	Low, Mid, High

[a]Mid point refers to processes where there was no specific target seeding density;

[b]DoE center composite design (face centered): target T°C of 35.5°C, shift on Day 6 from 35.5°C to 34.5 or to 33.5°C; target pH of 7.0, pH shift on Day 5 from 7.0 to 6.8 or 7.2; harvest Day 13, 14 and 15; target seeding density of 0.5, 0.75 and 1×10^6 cells/mL;

[c]screening DoE: target T°C of 35.5°C and shift on Day 4 from 35.5°C to 33.5°C; target pH of pH 6.8, 7, and 7.2; harvest Day 13, 14, and 15; target seeding density of 0.5, 0.75, and 1×10^6 cells/mL.

2.2 | Centrifugation shear mimic

A capillary shear device (CSD) was used as a centrifugation shear mimic. The CSD consists of a 10 cm long Upchurch Scientific® stainless steel capillary (VWR.co.uk order # 554-3222) attached to a chromatography skid—ÄKTAExplorer™ or ÄKTA™ Pure (Cytiva), which delivers the constant flow through the capillary. The capillary features a very small inner OD of 0.01′ (0.0254 cm). Further details about the CSD are discussed by Westoby et al. (2011) and Joseph et al. (2015, 2017).

2.3 | Analytics for cell lysis and solids concentration

2.3.1 | Cell lysis evaluation

The release of LDH by the CHO cells was measured using the LDH Cytotoxicity Assay Kit II (Abcam®). A fully lysed sample was generated upon incubation with the LDH kit lysing solution for 30 min at room temperature. The amount of LDH reactant products for the lysed and fresh samples were then measured using absorbance readings at 450 nm for 1 h using the Optima plate reader (BMG Labtech). The background for each was measured at 650 nm and subtracted from the 450 nm measurement. Percentage LDH release for harvest sample was, therefore, calculated using Equation (1).

$$R^{\mathrm{LDH}} = \frac{\left(\mathrm{ABS450}^{\mathrm{sample}} - \mathrm{ABS650}^{\mathrm{sample}}\right) - \left(\mathrm{ABS450}^{\mathrm{blank}} - \mathrm{ABS650}^{\mathrm{blank}}\right)}{\left(\mathrm{ABS450}^{\mathrm{sample}}_{\mathrm{lysed}} - \mathrm{AB650}^{\mathrm{sample}}_{\mathrm{lysed}}\right) - \left(\mathrm{ABS450}^{\mathrm{blank}} - \mathrm{ABS650}^{\mathrm{blank}}\right)} \quad (1)$$

where R^{LDH} is the percentage of LDH release, $\mathrm{ABS450}^{\mathrm{sample}}$ is the sample's absorbance at 450 nm, $\mathrm{ABS450}^{\mathrm{sample}}_{\mathrm{lysed}}$ is the absorbance at 450 nm of the fully lysed sample and $\mathrm{ABS450}^{\mathrm{blank}}$ is the absorbance of the blank sample (media) at 450 nm. The background is measured at 650 nm: $\mathrm{ABS650}^{\mathrm{sample}}$ for the sample, $\mathrm{AB650}^{\mathrm{sample}}_{\mathrm{lysed}}$ for the lysed sample and $\mathrm{ABS650}^{\mathrm{blank}}$ for the blank.

2.3.2 | Solids concentration

The solids concentration was determined using packed cell volume (PCV) tubes (TPP®). The bottom of the tubes has a capillary, which is graded to accurately measure the volume of solids in the sample. The tube was filled with 40 µL of cell culture material (in triplicates) and the samples were then spun at 2500 g for 1 min in a swing-out bucket centrifuge (Thermo Scientific Sorvall Legend® T Plus). More details about this protocol can be found in Stettler et al. (2006).

2.4 | Particle size distribution analyses

Each cell count measurement on the Vi-Cell XR generates a CSV file, containing the particle size distribution split into 140 discrete particle size bins in the size range between ~6 µm and 50 µm. A program was built in Matlab® (The MathWorks, Inc.) to import and pretreat the raw PSD data. The pretreatment consists of normalizing and binning the data at different intervals. As part of the data pretreatment, the PSD data was normalized to a frequency distribution as follows:

$$\mathrm{P}_s^{particle} = \frac{N_s^{\mathrm{particle}}}{\sum_{s=1}^{k} N_s^{\mathrm{particle}}} \quad (2)$$

where $\mathrm{P}_s^{\mathrm{particle}}$ is percentage P of particles of channel s (a channel or bin is defined by upper and lower bound for a particle size range) and N_s^{particle} is the number of particles N in channel s and k is the total number of channels.

2.5 | PLS model development

PLS is an advanced MVDA technique that is particularly suited for modeling data with large numbers of correlated predictor variables, which is the case for PSD. The dimensionality of the data is reduced through the projection of the data on small numbers of vectors (latent variables), which are orthogonal by nature and hence non-correlated. The predictor variables for the percentage LDH release consisted of the binned frequency PSD data for the viable cells, while for the solids concentration the predictor variables also included total and viable cell number as well as the PSD data for the viable cells. Before the analysis all variables were centered on the mean and scaled to unit variance.

The NIPALS algorithm was applied to build models with different numbers of latent variables. The detailed description of the approach applied to generate the PLS models can be found in Wold et al. (1987) and more recently applied to a bioprocessing scenario in Goldrick et al. (2017) and Goldrick et al. (2020). K-fold cross validation was applied where the data were split into seven and onefold was taken out of the analysis. The fit of the predictive model to the validation data set was determined using the predicted residual error sum of squares (PRESS). The PLS model with the optimum number of factors (latent variables) was selected based on minimum PRESS and prob> van der Voet $T^2 > 0.1$—the van der Voet T^2 tests shows if a model with different numbers of latent variables is significantly different from the optimum model (van der Voet, 1994). The variables included in the model were based on a minimum variable importance on projection (VIP) of at least 0.8 and hence the VIP score was used as a criteria for the model "pruning"—removing variables with a VIP score below 0.8 to improve the model. The VIP score indicates the importance of each variable and detailed information about its calculation can be found in Wold et al. 2001).

2.6 | Software

The data were extracted from the Vi-Cell and normalized as shown in Section 2.4 using Matlab (The MathWorks, Inc., Version 2015 to 2018b). JMP® Pro (SAS Institute Inc., Version 12.0.1 to 15.0.0) was used to derive the PLS models.

3 | RESULTS AND DISCUSSION

This section presents the different studies and rationale that were used to build the harvest toolbox for enhanced process predictability. Key outputs from the harvest toolbox are the levels of cell lysis prior and post centrifugation and the solids concentration before centrifugation. First, the limitations of current techniques were explored for measuring cell lysis occurring either as a result of environment stress such as nutrient depletion or an increase in by-products that promote natural decline in cell health towards the end of cell culture or due to centrifugation shear. The PSD was also plotted for the final days of cell culture to see to how the PSD changes with the drop in cell culture viability. PLS models were built leveraging the PSD data to determine if the PSD data could be mined to predict the levels of LDH release and the solids concentration.

3.1 | Exploring limitations of current techniques for measuring cell lysis

LDH release assays are often used for detection of cell lysis and yet can be lengthy and cumbersome to perform. On the other hand, the automated cell counters, such as Vi-CELL™ XR, are high throughput, easy to use routinely and also provide viability measurements using a live/dead dye. Therefore, this section investigates if viability as outputted from the Vi-CELL™ XR can be leveraged to predict percentage LDH release. Since percentage LDH release is often applied to characterize the shear in the centrifugation operation, it was also investigated how shear may affect the accuracy of the viability measurements.

3.1.1 | Relationship between LDH release and cell culture viability

The relationship between percentage LDH release and cell culture viability as outputted from the Vi-Cell was investigated to see if the automated Vi-Cell viability measurements can be leveraged to predict percentage LDH release. The data used in these studies was presheared material from across four separate projects—Null, A, B and C and different scales—working volume of 15 mL, 7 L, 50 L and 200 L—with analysis performed on cell culture of different age—between Day 10 and 14. This data set provided a wide distribution of Vi-Cell measured cell culture viabilities ($\mu_{viability}$ = 79% and $\sigma_{viability}$ = 17%) and LDH release measurements ($\mu_{R^{LDH}}$ = 44% and $\sigma_{R^{LDH}}$= 26%). In Figure 1 the viability was plotted against the %LDH release revealing that relationship can be described by a piecewise function. The relationship that was established between percentage LDH release and viability can be expressed as a mathematical function shown below:

$$\%LDH_{release} = \begin{cases} 100\% & Viability < Viability_{Crit} \\ \frac{100\% - Viability}{100\% - Viability_{Crit}} & Viability \geq Viability_{Crit} \end{cases} \quad (3)$$

where Viability is the measured cell culture viability in the Vi-Cell, $Viability_{Crit}$ is the critical viability, below which the % LDH release is 100%. The $Viability_{Crit}$ was calculated through a least squares algorithm that minimized the error of the prediction of the % LDH release at the various viabilities and this was calculated to be equal to 54% with a coefficient of determination (R^2) value equal to 0.80. This analysis demonstrated that a nonlinear relationship can be derived to describe the link between Vi-cell cell culture viability and percentage LDH release. This function can be applied to leverage the cell culture viability data to predict the percentage LDH release (cell lysis) in the cell culture.

As discussed earlier, centrifugation shear can also lead to LDH release due to cell membrane lysis. Cell membrane lysis post shear is an important parameter since mammalian cells are often exposed to shear during the centrifugation-based recovery operation. The following section investigates how shear impacts on the accuracy of the Vi-Cell viability measurements.

3.1.2 | Measuring cell lysis due to shear

As discussed earlier a nonlinear relationship was established between percentage LDH release and cell culture viability (e.g., Vi-Cell). Both methods have been applied in literature to measure the levels of cell lysis due to centrifugation shear (Joseph et al., 2016; Tait et al., 2013 and Westoby et al., 2011). Understanding the level of cell lysis and where it occurs in the processes can shed light on what conditions can lead to product degradation. However, to date there is no detailed investigation of the impact of shearing the cells on the viability measurement as outputted from the Vi-Cell or a comparison between the LDH and trypan blue methods such as the Vi-Cell for evaluating cell lysis post shear (Figure 1).

A scale-down centrifugation mimic was applied to shear various harvested cell culture samples. The centrifugation mimic was developed through comparison of the levels of LDH release between the mimic and a pilot scale centrifugation process (data not shown) similar to work conducted by Joseph et al. (2016). Based on this comparison – low shear (21 mL/min corresponding to ~1.2×10^5 W/kg) and high shear (41 mL/min corresponding to ~7.9×10^5 W/kg) conditions were achieved through the use of different flowrates in the capillary. In Figure 2a the levels of LDH release and cell culture viability (Vi-CELL™ XR) were measured for samples at no shear, low shear and high shear. Samples collected on different harvest days across different

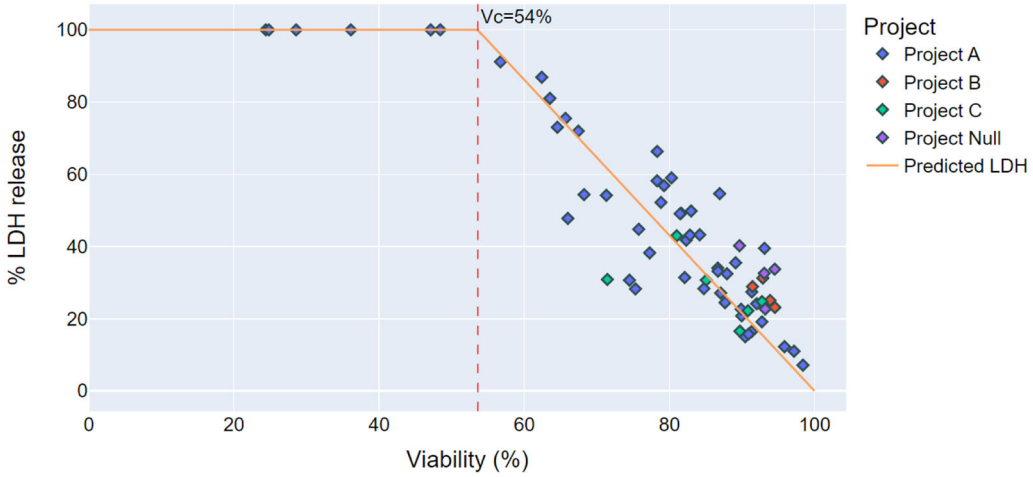


FIGURE 1 Percentage LDH release versus Vi-Cell XR viability measurements for the fresh harvest (no shear) for four different projects: Project A in blue (◆), Project B in red (◆), Project C in green (◆) and Project null in purple. (◆). The solid line (—) shows predicted LDH release using piecewise regression where Vc is the critical viability, below which the LDH is 100%. All samples above 95% viability have been adjusted to 100%.

projects—Null, A, and C—were investigated to obtain a wide range of cell culture material with different properties. The viability and total cell number (TCN) for the cell culture at no shear are summarized in the top graph in Figure 2a; for all cell culture material the viability was relatively high (above 80%) and the TCN was in the range of 25–40 million cells per mL.

The apparent increase in TCN in Figure 2b is calculated as the difference between TCN before and after shear. Figure 2b reveals that applying shear to cell culture samples led to an apparent increase or decrease in the total number of cells. The apparent increase in TCN is likely due to the increase in cellular debris counted as cells, while the decrease in cells can be attributed to disintegration of the cells into smaller particles not counted as cells. The change in apparent TCN questions the reliability of the viability measurements obtained for sheared samples using the trypan blue exclusion method since the viability is calculated as a ratio of viable to total number of cells. It is important to note that the automated cell counter was optimized for routine use in the lab and the parameters were not adapted for each individual run, for example, tuning the instrument's settings for declustering of live cells or altering the criteria for counting a particle as a cell such as size and circularity. Adjusting these parameters would have introduced additional variability in the measurements and would have not suited high throughput experimentation approaches.

This section demonstrated that for sheared samples the automated cell counter will not provide an accurate evaluation of the level of cell culture lysis due to overestimation of the total number of cells linked to the increased levels of debris or the micronisation of the cell debris to sizes below the size recognizable. Therefore, there is the need to develop improved methods for evaluating the level of cell lysis for samples subjected to shear.

3.2 | Investigating the relationship between cell culture particle size and viability

As discussed in the previous section, there is the need to develop tools to measure cell lysis that do not require cumbersome methods such as the LDH assay and that provide accurate measurements that are not susceptible to issues such as cell micronisation. The particle size analysis discussed in this section was performed on an ambr run where the seeding density and harvest days were varied following a DoE design to generate a wide range of cell culture viabilities. The data were imported as described in 2.4 and normalized to facilitate comparison between samples with different total numbers of cells. In Figure 3 the averaged PSD across all runs can be found for different harvest days, described by the solid line with the upper and lower dotted lines indicating the standard deviation for the measurements. It was revealed that, the PSD remains narrow during the first few days of cell culture, with one single peak observed at ~15 µm. However, towards the end of the cell culture, and especially during the stationary and death phase (Day 12), multiple peaks were observed and there was significant increase in the standard deviation (Figure 3). Further investigation into the last harvest day, with the biggest standard deviation, reveals that the PSD changed with cell culture viability. To simplify the analysis, the samples from Day 12 were grouped into discrete ranges of harvests of different levels of cell culture viability. This highlighted that the PSD shifted to the left as viability dropped, moving from a broad distribution with a pronounced peak observed at ~20 µm for high viability samples (>85%) to one with a peak at ~10 µm for low viability samples (24–36%). The following sections investigate whether the PSD can be further leveraged to gain insights into the harvest properties, and in particular the level of LDH release and solids concentration.

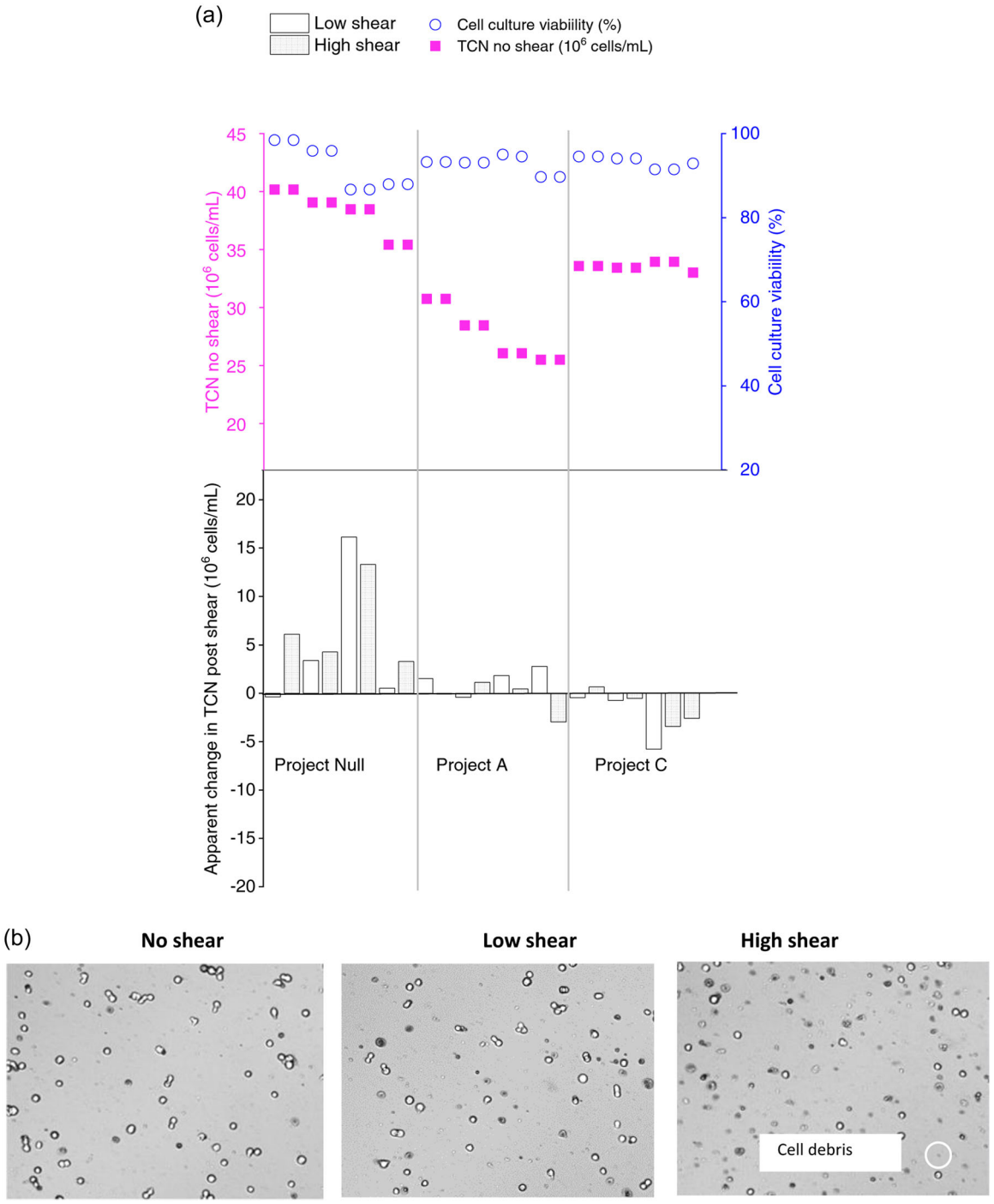


FIGURE 2 Impact of shear on (a) apparent change in total cell number (TCN) as outputted from the Vi-CELL™ XR for various mammalian cell culture harvests for three projects and (b) the appearance of cell debris for high shear samples compared to no and low shear. Harvests from three different bioreactor runs were collected on different days of cell culture (Day 10–14) to represent different viabilities subsequently subjected to low and high shear in the capillary shear device. The apparent TCN change was calculated as a difference between the TCN before and after shear.

3.3 | Leveraging PSD data to predict LDH release using PLS techniques

The previous section demonstrated that the particle size distribution changes with cell culture viability. In this section, the PSD data were outputted from the Vi-Cell for various cell cultures (Table 1) and analyzed to determine if it can be leveraged to estimate the level of cell lysis or LDH release. To date, we are not aware of other studies that have leveraged the PSD data for this purpose. In this section, predictive PLS models were derived to determine LDH release as a

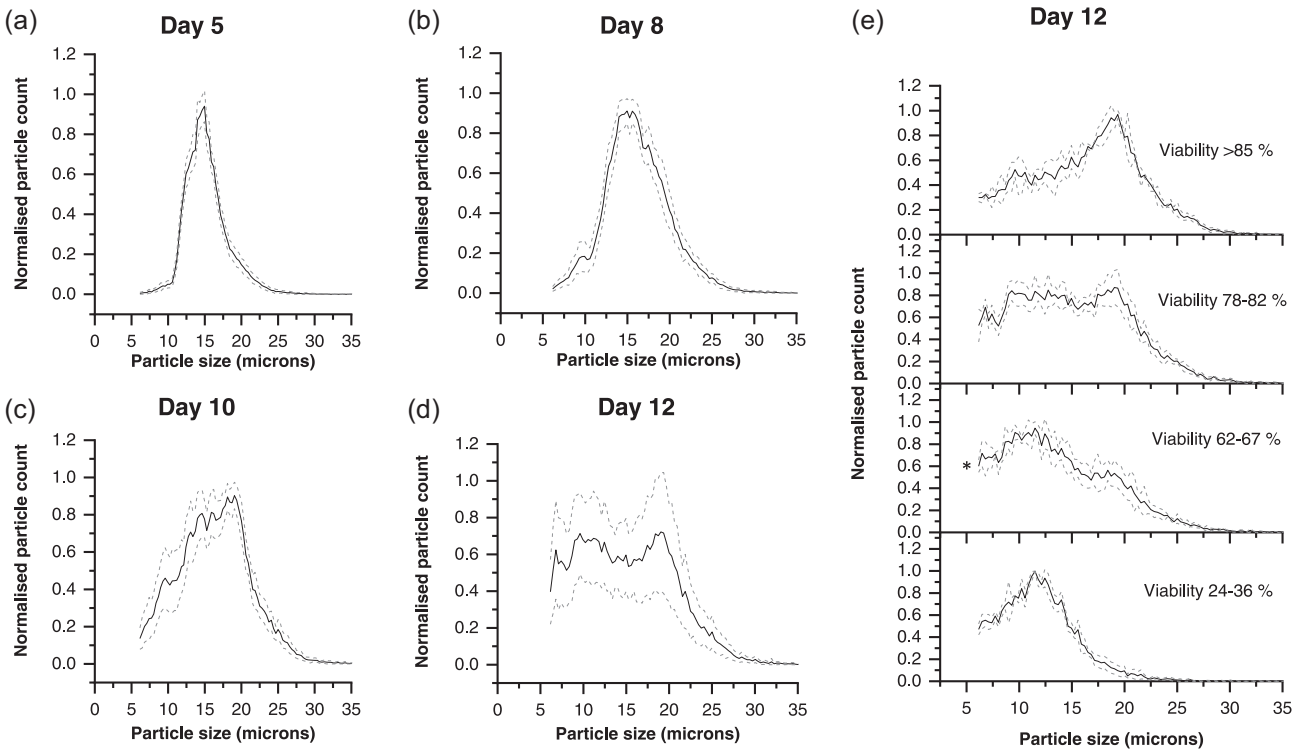


FIGURE 3 Change in particle size distribution over cell culture duration in an ambr set of experiments. The solid line depicts the mean of the PSD, the dotted lines depict the standard deviation across 24 ambr bioreactor runs for (a) Day 5 ($n = 24$), (b) Day 8 ($n = 24$), (c) Day 10 ($n = 24$), (d) Day 12 ($n = 17$) and (e) Day 12 for various viability ranges (>85% viability calculated from $n = 3$; 72% to 82% calculated from $n = 6$; 62% to 67% calculated from $n = 4$; 24% calculated from 36%and $n = 3$). N indicates the number of bioreactors. See Table 1 ambr2 for ambr details.

function of the PSD data for the viable number of cells as inputs (binned at 1 µm intervals). This overall workflow is summarized in Figure 4.

To determine if LDH release can effectively be determined from PSD data, the prediction statistics of the PLS model were analyzed. Figure 5a revealed that the PLS model could accurately predict the % LDH release for an unknown test sample from the PSD data (R^2adj. = 0.90 for a parity plot). This was further reinforced by the model statistics where the R^2Y and R^2X based on the cross-validation results for the model were 82% and 96% respectively indicating that the model is able to predict variances in Y (predicted variable, i.e., LDH) and in X (the predictor, i.e., PSD). Hence these findings demonstrate that PSD data can be mined to get an early predictor of LDH release without need for additional assays that are traditionally used.

To determine which PSD variables were important for prediction of % LDH release, a VIP plot was examined. Figure 5b shows that the highest VIP score is for the lower end of the PSD (between 6 and 12 µm), which could be explained by increase in smaller particles for the lower viability cell cultures, that was observed earlier (Figure 4). The PLS model had four latent variables and that predictors that were used in the PLS model were particle sizes below the expected cell size (21 µm) based on these sizes meeting the minimum VIP score for model inclusion (0.8).

The new approach described above only used data that was available thus eliminating the need for additional analytics to determine the LDH release levels. Since these data were routinely generated with each cell count, it allows for retrospective analysis of the PSD data to estimate the level of LDH release. Furthermore, as seen in Table 1 the analysis included fresh samples from the bioreactors and samples subjected to shear in the CSD as well as for different cell cultures and operating conditions.

3.4 | Leveraging PSD data to predict solids concentration using PLS techniques

The levels of [solids] in the cell culture harvest is a key parameter for the harvest operation as it impacts the centrifugation yield. The approach that was developed to leverage the PSD data to predict the levels of LDH release was also applied here to predict the levels of [solids]. A total of 117 measurements for cell culture [solids] and Vi-Cell cell count data (including PSD, total and viable cell number and cell culture viability) were used for the study; the data were generated from ambr and bench scale bioreactor runs (as described in Table 1). The conditions varied largely between the different projects and that led to cell cultures with a wide range of viabilities ($\mu_{viability} = 85\%$ and $\sigma_{viability} = 14\%$) and densities ($\mu_{TCN} = 30$ million cells/mL and $\sigma_{TCN} = 10$ million cells/mL). The PSD measurements were analyzed using the same Matlab program as for the LDH analysis (described in Section 2.4). The viability, the total and viable

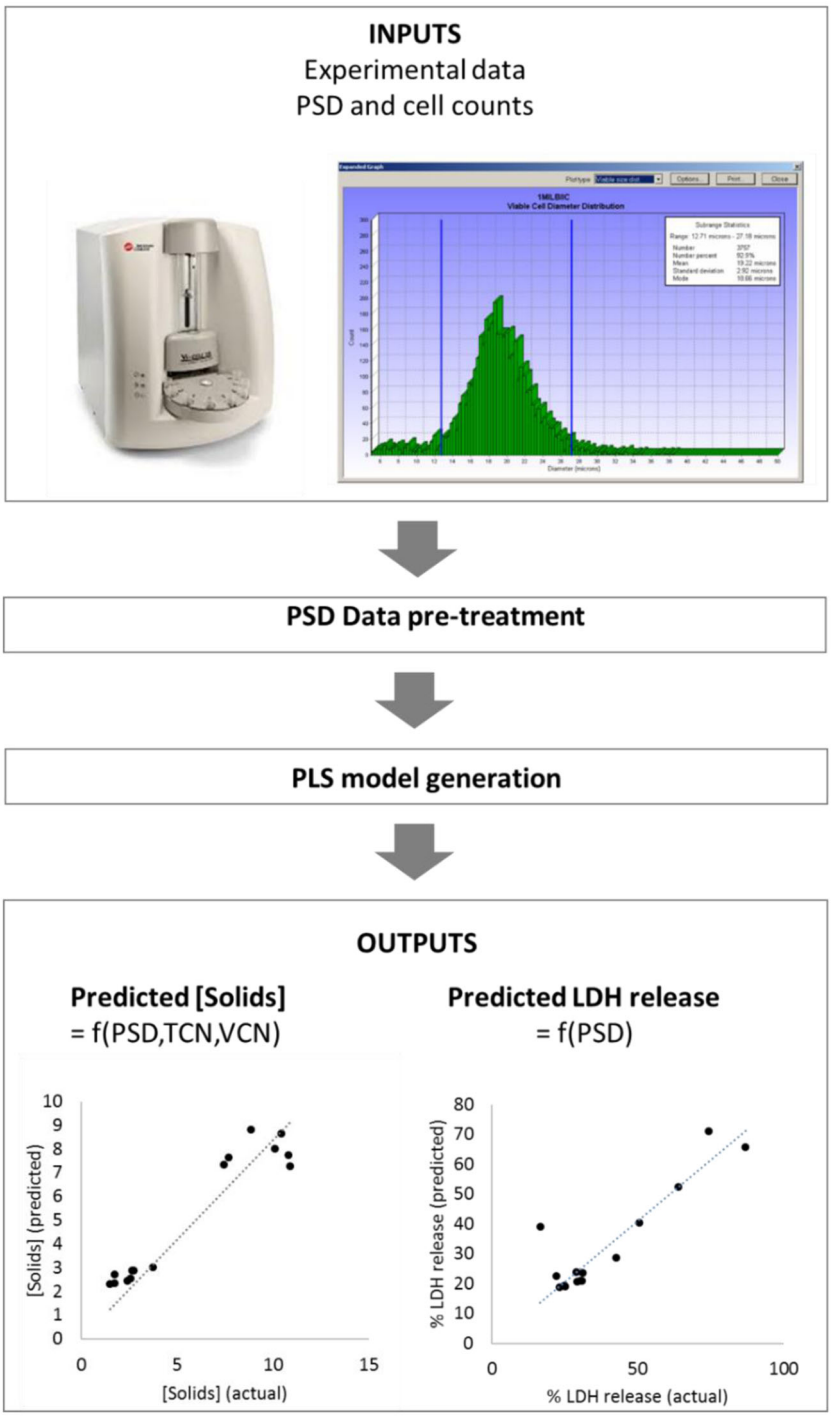


FIGURE 4 Workflow for analyzing the PSD data outputted from Vi-Cell XR. The raw data are extracted from the Vi-Cell in the form of CSV files and an algorithm in Matlab was used to pretreat and structure the data for further analysis. PLS models were derived for the pretreated data to predict LDH release and [solids] as a function of PSD, TCN (total cell number) and VCN (viable cell number). The image of the PSD shown in the figure was extracted from Vi-Cell user manual (Beckman Coulter, 2011).

cell number together with the frequency PSD data (binned at 1 µm) for the viable number of cells were used to build a PLS model that predicts the [solids] in the harvest. The calibration data set used to build the model consists of four independent ambr runs, while the validation data set consists of a bench scale run and a separate culture station from one of the ambr runs. A total of six factors (latent variables) were used to build the model.

Figure 6a revealed that the model provided a good prediction for the [solids] with $R^2Y = 84\%$ and $R^2 = 0.86$ for actual versus predicted (parity plot). The VIP plot in Figure 6b revealed that the most important predictors in the model are the total and viable cell counts and the particle size around 13 µm, which is very close to the actual average CHO cells' size.

Hence, using PLS techniques and the outputs from the automated cell counter it was possible to accurately predict the [solids] without the need for additional sampling. Furthermore, the [solids] can be linked back to the cell culture conditions, so that the screening for optimal conditions also accounts for the potential impact on [solids], and consequently centrifugation yield.

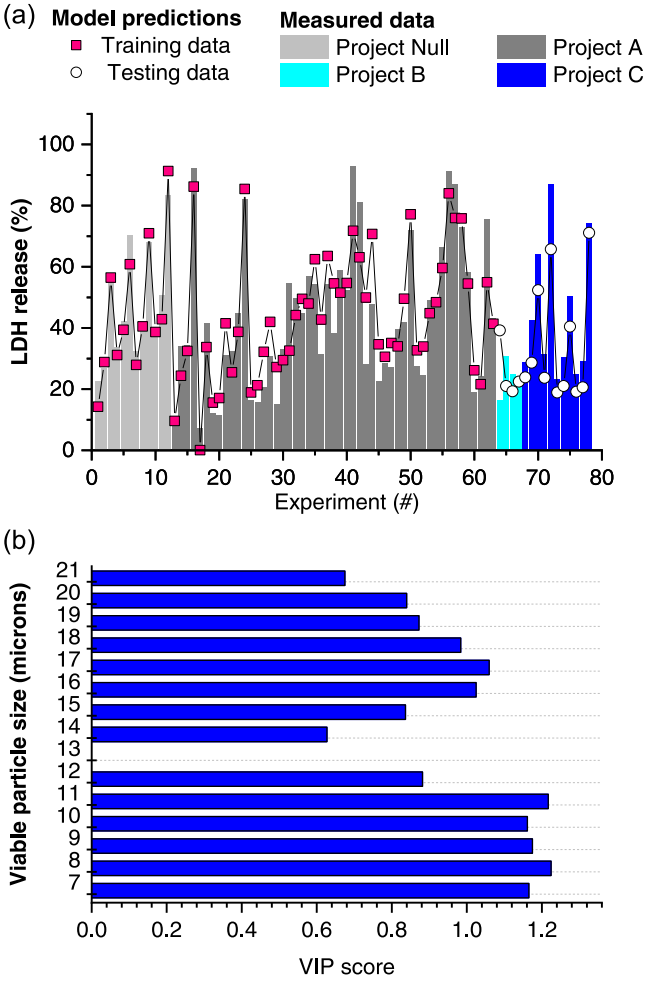


FIGURE 5 PLS model predicting LDH release using PSD data for the viable cells as outputted from the Vi-Cell. In (a) the actual LDH release is shown by the bars, the filled squares depict the predicted LDH release for the data set used to calibrate the model, while the empty circles show the predicted percentage solids concentration for the external testing data set. The different colored bars refer to different projects - project null (■); project A (■); project B (■) and project C (■). In (b) VIP plot is shown for the predictors used in the model. For the PLS model the R^2Y is 83% and the R^2X is 96%.

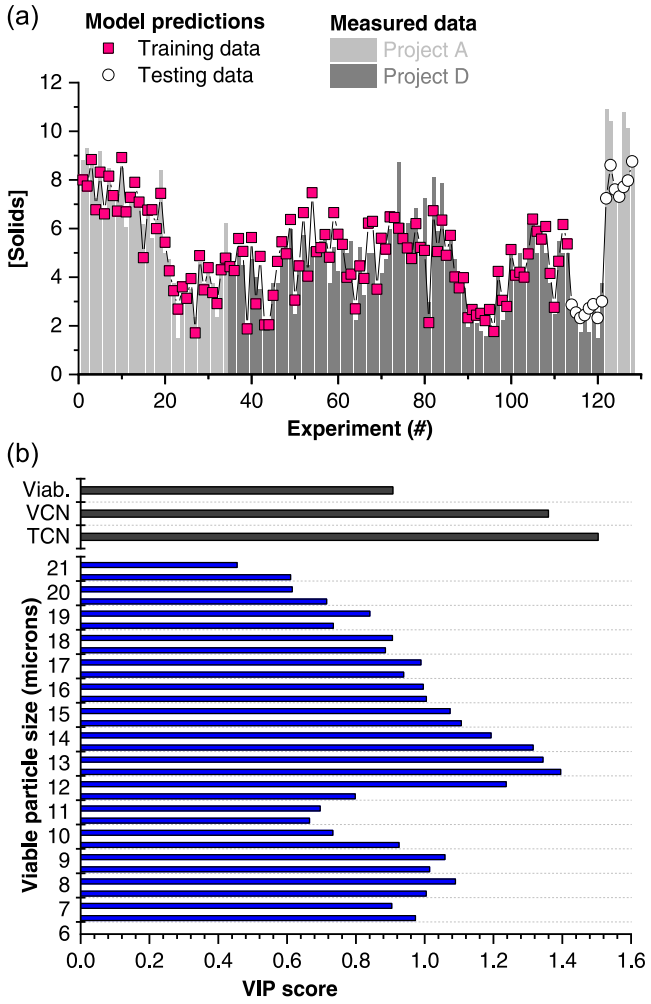


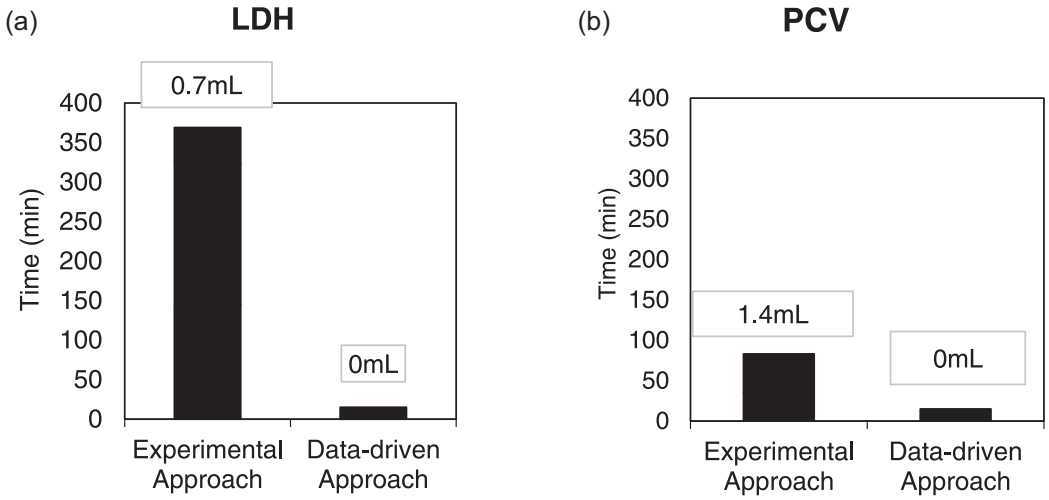
FIGURE 6 PLS model for [solids] with (a) actual versus predicted [solids] and (b) VIP plot with the predictors used in the model. In (a) the actual [solids] are shown by bars, the filled squares depict the predicted solids concentration for the data set that was used to calibrate the model, while the empty circles show the predicted percentage solids concentration for the test data set. The different colored bars refer to different projects (A—light gray and D—dark gray), R^2Y = 81% and R^2X = 86%. TCN, total cell number (million cells/mL); VCN, Viable cell number (million cells/mL); viab., viability.

3.5 | Volume and time savings

During process development, access to material is limited and time is critical. Figure 7 summarizes the savings for the daily analysis of the 48 vessels of ambr^{TM}15 in terms of time applying the novel workflow versus traditional sampling and analytics techniques. The figure highlights that the time for analysis could be reduced to a few minutes resulting in savings of over 6 h for the LDH release measurement and over an hour for the [solids] analysis when leveraging the PSD data instead of the traditional offline analytics. In addition, the volume savings per vessel per run were determined to be over 2 mL (1.4 mL for the [solids] and 0.7 mL for the LDH measurement), which is significant considering the working volume of each ambr vessel (15 mL). Hence, the proposed approach using the PSD data to predict these metrics can be particularly valuable during high throughput experiments.

4 | CONCLUSIONS AND FUTURE WORK

The paper described a novel approach that leveraged routinely collected PSD data from an automated cell counter and PLS models to accurately predict cell culture solids concentration and cell lysis (% LDH release) for the harvest operation. This avoids the need for offline analytics for such measurements that can be laborious and hence often discourage collecting such measurements. Applying the novel approach, it was possible to predict the % LDH release and solids concentration for a test data set with a 90% and 86%



(c)

Activity	Process and Analytics Details
Cell culture	Ambr15 with 48 bioreactors 14 day process Cell counts every other day
LDH release measurement based on a 96-well plate format and an Optima plate reader	100 µL per sample 6 measurements per sample* 30 sec preparation time per sample 60 min instrument running time per 96-well plate 15 min analysis per 96-well plate
PCV measurements using PCV tubes	50µL per sample** Two min per centrifuge spin 24 samples per centrifuge spin Four measurements per sample 6 min analysis per 24 samples 6 sec prep per sample

* Triplicates for lysed and measured sample

** 40µL required per sample plus 10µL overage

FIGURE 7 Comparison in time and volume daily requirement for ambr^{TM}15 with 48 vessels for (a) LDH release measurements using manual sampling and a colorimetric assay versus the novel data-driven PLS—based estimation and (b) PCV measurements using manual sampling versus the novel PLS—based estimation. (c) The assumptions used in the comparison. LDH, lactate dehydrogenase; PCV, packed cell volume; PLS, partial least squares.

goodness-of-fit respectively (based on actual vs. predicted parity plots). The paper highlights the volume and time savings by leveraging the PSD data rather than using offline analytics, with the analytics time reducing from more than 7 h using the conventional approach to a few minutes with the proposed workflow when working with the high throughput ambr15 microbioreactor system with 48 vessels. The speed and resource-efficiency of this tool makes it best suited for early high throughput experimentation. This will enable gaining understanding about cell lysis and solids concentration in the cell culture earlier in the development cycle. Adopting such an approach will facilitate earlier identification of challenging harvests that can exceed the centrifuge capacity and hence inform process modifications that lead to more robust processes and improved facility fit.

ACKNOWLEDGMENTS

The authors would like to thank the following members of the AstraZeneca upstream team with their out-of-hours support and providing the pilot scale (50 and 200 L material)—Steven Ruddock, Anagha Eswar, Ruchika Bandekar and Jennifer Smith. For the recovery experiments, the authors would like to express their gratitude to Aled Charles and David Gruber who performed the large-scale centrifugation experiments with overall support with booking and using the equipment in the downstream suite. I would

also like to thank Dr. Joseph Egan for his useful discussion on the derrivation of the correlation for the percentage LDH release and viability. Financial support from the UK Engineering and Physical Sciences Research Council (EPSRC) and AstraZeneca for the Engineering Doctorate studentship for M. Sebastian is gratefully acknowledged (Grant Ref: EP/L01520X/1). This research is associated with the joint UCL-AstraZeneca Centre of Excellence for predictive decision-support tools in the bioprocessing sector and is aligned with the EPSRC Future Targeted Healthcare Manufacturing Hub hosted by UCL Biochemical Engineering.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Martina Sebastian http://orcid.org/0009-0000-8830-1227
Stephen Goldrick http://orcid.org/0000-0002-8319-1679
Suzanne S. Farid http://orcid.org/0000-0001-8155-0538

REFERENCES

Aucamp, J. P., Davies, R., Hallet, D., Weiss, A., & Titchener-Hooker, N. J. (2014). Integration of host strain bioengineering and bioprocess development using ultra-scale down studies to select the optimum combination: An antibody fragment primary recovery case study. *Biotechnology and Bioengineering*, *111*(10), 1971–1981. https://doi.org/10.1002/bit.25259

Banner, M., Alosert, H., Spencer, C., Cheeks, M., Farid, S. S., Thomas, M., & Goldrick, S. (2021). A decade in review: use of data analytics within the biopharmaceutical sector. *Current Opinion in Chemical Engineering*, *34*, 100758. https://doi.org/10.1016/J.COCHE.2021.100758

Beckman Coulter, I. (2011). *Vi-CELL XR Cell Viability Analyzer Reference Manual*. Retrieved February 22, 2023, from www.beckmancoulter.com

Cummings, B. S., & Schnellmann, R. G. (2004). Measurement of cell death in mammalian cells. *Current Protocols in Pharmacology*, *25*(1). 12.8.1–12.8.22. https://doi.org/10.1002/0471141755.ph1208s25

Galluzzi, L., Bravo-San Pedro, J. M., Vitale, I., Aaronson, S. A., Abrams, J. M., Adam, D., Alnemri, E. S., Altucci, L., Andrews, D., Annicchiarico-Petruzzelli, M., Baehrecke, E. H., Bazan, N. G., Bertrand, M. J., Bianchi, K., Blagosklonny, M. V., Blomgren, K., Borner, C., Bredesen, D. E., Brenner, C., … Kroemer, G. (2015). Essential versus accessory aspects of cell death: Recommendations of the NCCD 2015. *Cell Death & Differentiation*, *22*(1), 58–73. https://doi.org/10.1038/cdd.2014.137

Goldrick, S., Holmes, W., Bond, N. J., Lewis, G., Kuiper, M., Turner, R., & Farid, S. S. (2017). Advanced multivariate data analysis to determine the root cause of trisulfide bond formation in a novel antibody-peptide fusion. *Biotechnology and Bioengineering*, *114*(10), 2222–2234. https://doi.org/10.1002/bit.26339

Goldrick, S., Umprecht, A., Tang, A., Zakrzewski, R., Cheeks, M., Turner, R., Charles, A., Les, K., Hulley, M., Spencer, C., & Farid, S. S. (2020). High-Throughput Raman spectroscopy combined with innovate data analysis workflow to enhance biopharmaceutical process development. *Processes*, *8*(9), 1179. https://doi.org/10.3390/PR8091179

Handlogten, M. W., Wang, J., Ahuja, S., Sciences, A., Bio, V., & Way, M. (2020). Online control of cell culture redox potential prevents antibody interchain disulfide bond reduction. *Biotechnology and Bioengineering*, *117*(5), 1329–1336. https://doi.org/10.1002/BIT.27281

Handlogten, M. W., Zhu, M., & Ahuja, S. (2017). Glutathione and thioredoxin systems contribute torecombinant monoclonal antibody interchain disulfide bond reduction during bioprocessing. *Biotechnology and Bioengineering*, *114*(7), 1469–1477. https://doi.org/10.1002/bit.26278

Hiebl, B., Peters, S., Gemeinhardt, O., Niehues, S. M., & Jung, F. (2017). Impact of serum in cell culture media on in vitro lactate dehydrogenase (LDH) release determination. *Journal of Cellular Biotechnology*, *3*(1), 9–13. https://doi.org/10.3233/jcb-179002

Hsu, W. T., Aulakh, R. P. S., Traul, D. L., & Yuk, I. H. (2012). Advanced microscale bioreactor system: A representative scale-down model for bench-top bioreactors. *Cytotechnology*, *64*(6), 667–678. https://doi.org/10.1007/s10616-012-9446-1

Hutterer, K. M., Hong, R. W., Lull, J., Zhao, X., Wang, T., Pei, R., Flynn, G. C. (2013). Monoclonal antibody disulfide reduction during manufacturing: Untangling process effects from product effects. *MAbs*, *5*(4), 608–613. https://doi.org/10.4161/mabs.24725

Joseph, A. (2015). *Lab-scale methodology for the Generation of Sheared Mammalian Cell Culture Samples*.

Joseph, A., Goldrick, S., Mollet, M., Turner, R., Bender, J., Gruber, D., & Titchener-Hooker, N. (2017). An automated laboratory-scale methodology for the generation of sheared mammalian cell culture samples. *Biotechnology Journal*, *12*(5), 1600730. https://doi.org/10.1002/biot.201600730

Joseph, A., Kenty, B., Mollet, M., Hwang, K., Rose, S., Goldrick, S., Bender, J., Farid, S. S., & Titchener-Hooker, N. (2016). A scale-down mimic for mapping the process performance of centrifugation, depth and sterile filtration. *Biotechnology and Bioengineering*, *113*(9), 1934–1941. https://doi.org/10.1002/bit.25967

Kao, Y. H., Hewitt, D. P., Trexler-Schmidt, M., & Laird, M. W. (2010). Mechanism of antibody reduction in cell culture production processes. *Biotechnology and Bioengineering*, *107*(4), 622–632. https://doi.org/10.1002/bit.22848

Khan, K. H. (2013). Gene expression in mammalian cells and its applications. *Advanced Pharmaceutical Bulletin*, *3*(2), 257–263. https://doi.org/10.5681/apb.2013.042

Li, F., Vijayasankaran, N., Shen, A., Kiss, R., & Amanullah, A. (2010). Cell culture processes for monoclonal antibody production. *mAbs*, *2*(5), 466–479. https://doi.org/10.4161/mabs.2.5.12720

Lu, R. M., Hwang, Y. C., Liu, I. J., Lee, C. C., Tsai, H. Z., Li, H. J., & Wu, H. C. (2020). Development of therapeutic antibodies for the treatment of diseases. *Journal of Biomedical Science*, *27*, 1–30. https://doi.org/10.1186/s12929-019-0592-z

Méry, B., Guy, J.-B., Vallard, A., Espenel, S., Ardail, D., Rodriguez-Lafrasse, C., Rancoule, C., & Magné, N. (2017). In vitro cell death determination for drug discovery: A landscape review of real issues. *Journal of Cell Death*, *10*, 1179670717691251. https://doi.org/10.1177/1179670717691251

Mullard, A. (2021). FDA approves 100th monoclonal antibody product. *Nature Reviews Drug Discovery*, *20*(7), 491–495. https://doi.org/10.1038/D41573-021-00079-7

O'Mara, B., Gao, Z. H., Kuruganti, M., Mallett, R., Nayar, G., Smith, L., Meyer, J. D., Therriault, J., Miller, C., Cisney, J., & Fann, J. (2019). Impact of depth filtration on disulfide bond reduction during downstream processing of monoclonal antibodies from CHO cell cultures. *Biotechnology and Bioengineering*, *116*(7), 1669–1683. https://doi.org/10.1002/bit.26964

Riss, T., Niles, A., Moravec, R., Karassina, N., & Vidugiriene, J. (2004). Cytotoxicity Assays: In Vitro Methods to Measure Dead Cells, *Assay Guidance Manual*. https://www.ncbi.nlm.nih.gov/pubmed/31070879

Stettler, M., Jaccard, N., Hacker, D., De Jesus, M., Wurm, F. M., & Jordan, M. (2006). *New Disposable Tubes for Rapid and Precise Biomass Assessment for Suspension Cultures of Mammalian Cells*. https://doi.org/10.1002/bit.21071

Tait, A. S., Tarrant, R. D. R., Velez-Suberbie, M. L., Spencer, D. I. R., & Bracewell, D. G. (2013). Differential response in downstream processing of CHO cells grown under mild hypothermic conditions. *Biotechnology Progress*, *29*(3), 688–696. https://doi.org/10.1002/btpr.1726

Tebbe, H., Lutkemeyer, D., Gudermann, F., Heidemann, R., & Lehmann, J. (1996). Lysis-free separation of hybridoma cells by continuous disc stack centrifugation. *Cytotechnology*, *22*(1–3), 119–127. https://doi.org/10.1007/BF00353931

van der Voet, H. (1994). Comparing the predictive accuracy of models using a simple randomization test. *Chemometrics and Intelligent Laboratory Systems*, *25*(2), 313–323. https://doi.org/10.1016/0169-7439(94)85050-X

Westoby, M., Rogers, J. K., Haverstock, R., Romero, J., & Pieracci, J. (2011). Modeling industrial centrifugation of mammalian cell culture using a capillary based scale-down system. *Biotechnology and Bioengineering*, *108*(5), 989–998. https://doi.org/10.1002/bit.23051

Wold, S., Geladi, P., Esbensen, K., & Öhman, J. (1987). Multi-way principal components-and PLS-analysis. *Journal of Chemometrics*, *1*(1), 41–56. https://doi.org/10.1002/cem.1180010107

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, *58*(2), 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1

How to cite this article: Sebastian, M., Goldrick, S., Cheeks, M., Turner, R., & Farid, S. S. (2023). Enhanced harvest performance predictability through advanced multivariate data analysis of mammalian cell culture particle size distribution. *Biotechnology and Bioengineering*, 1–13. https://doi.org/10.1002/bit.28571