Variability patterns identification of an α-CD20 monoclonal antibody´s perfusion process by exploratory data analysis: A Case Study

Lisandra Calzadilla[1], Osvaldo Gozá[2], Mayla Echazabal[2], Arturo Toledo[1], and Tammy Boggiano[1]

[1]Center of Molecular Immunology
[2]Instituto Superior Politecnico Jose Antonio Echeverria

June 10, 2020

Abstract

Perfusion processes have gained interest as mammalian cell culture mode. Due to the high complex issues and variability regarding such culture mode, a Principal Component Analysis (PCA) to the data was used to characterize such variable process patterns and to achieve a better understanding. The transfected NS0/1B8 cell line was fermented in a 500 L bioreactor in perfusion culture mode to obtain a desired monoclonal antibody against CD20 molecule. Given the high variability of the process, an exploratory data analysis based on a multivariate analysis technique such as PCA was applied. The variables were selected by a risk model analysis based on the cause-effect matrix, the experience accumulated in the process, over the critical quality attributes, and parameters focus on the fermentation process. As a result, it was obtained that two main components were able to explain more than 95% of the total variance, and it was possible to select between the critical parameters those that have the greatest contribution to the variability of the fermentation process. Furthermore, the practical experiences of the specialists matched with the results and new process recommendations were projected to improve the control strategy for a further Continuous Process Verification. Keywords: Monoclonal antibody, Principal Component Analysis, Fermentation, Critical parameters.

INTRODUCTION

Mammalian cell line culture have emerged as one of the most applied platforms for several biopharmaceutical products, mainly due to their ability to generate species with complex posttranslational modifications such as monoclonal antibodies (mAb) (*Kantardjieff et al., 2014*). Different culture modes have been developed to produce them across time, in order to increase the product titer and cell density as well as to reduce the process complexity and costs, and the most commonly used are fed-batch and perfusion mode (*Farid et al., 2006; Fike, 2009; Xu et al, 2017*). Hence, companies such as Amgen, Merck and Pfizer that were established on Fed-batch mode, also have been exploring perfusion mode as a wise alternative mostly leading to increase volumetric productivity on their processes (*Lin et al., 2017*).

Moreover, for a long time, process validation has been object of greater scrutiny in Good Manufacturing Practice (GMP), and more recently, the Quality by Design (QbD) concept have gained an important role by approaching to an efficient, agile and flexible biopharmaceutical manufacturing sector. While traditional product development and manufacture often involve the use of heuristic approach methods, QbD points out to better process understanding and control by defining the process design space. The use of statistical approaches and Process Analytical Technology (PAT) are some of the stages to take into consideration in product lifecycle, mainly focused on the efficacy and safety as principal target to achieve (*Rathore, 2009; Stosch et al., 2016; FDA, 2011*). On the other hand, the quality risk management (QRM) and the evaluation

and ranking of process parameters (PP) over the quality attributes (QA) and critical quality attributes (CQA) have been other approaches applied to lead the process to a more controlled environment in order to achieve higher product quality. Many companies have been developing the concept of Criticality over PP and QA as well. The main focuses of those approaches are related to define monitoring and operations control, and the continued process verification strategies as the process validation last stage and therefore, assuring a continuum for more consistent and robust processes (*Witcher, 2017; Mitchell, 2013; Mitchell, 2014*).

Data analysis, is carried out after experimental tasks and is mainly addressed to a deeper knowledge of processes for diagnosis on different development steps. In contrast, the processes optimization and prediction models might be reached by data analysis techniques. There are tools concerning this topic most commonly named as Multivariate Data Analysis (MVDA). The application of MVDA techniques have been currently used, focused to a wide number of targets, and Biotechnology is one of those. Partial Least Square Regression (PLSR) and Principal Component Analysis (PCA) are some of the methods associated toward such applications. For instance, the matrix building with missing values and its interpretation could be carried out successfully by the PCA model. On the other hand, the scale-up, process comparability, process characterization and fault diagnosis are some of the possible activities facilitated by the MVDA techniques as an advanced data processing tool. Furthermore, some authors have been combining different MVDA techniques to explain the impact on the Critical Quality Attributes (CQA) of the evaluated mAb depending on culture media. They have showed its effectiveness to analyze multiple correlated objectives and factors for early process development by applying for instance PCA as one of the main MVDA techniques (*Kirdar et al., 2007; Sokolov et al., 2017; Folch-Fortuny et al., 2015; Abdi et al., 2010; Sokolov et al., 2015*). Thus, really useful tools can be found to lead the process to a highly controlled environment in order to reach deeper process parameters knowledge and more safety products (*Thomassen et al.* , 2010;*Suarez-Zuluaga et al.* , 2019).

A high process variability has been observed across the time in the NS0/1B8 cell line fermentation stage, a myeloma cell line that produces a chimeric mAb against the CD20 surface lymphocyte marker. The current work combines the approaches presented above to estimate the influence of PP over CQA by implementing a cause-effect matrix risk model, and afterward, MVDA was applied, specifically PCA to find out the main causes (process parameter or key process parameters) that can trigger process variability. The perfusion mode has been used in the fermentation process, for that reason, novel and interesting results are expected in this case study.

MATERIALS AND METHODS

Cell culture and fermentation process

NSO/1B8 transfected cells were cultivated in stainless steel bioreactor (Bioengineering Biotechnology Company, Switzerland) of 500 L working volume. Inoculum for fermentations were grown in rotatory roller flask which was used to seed the bioreactor. The medium used was a homemade property animal free component medium. Process set points were: temperature 37°C, pH 6.95 ± 0.1, agitation rate 60 rpm, pH control method by CO2 sparging, and DO set point of 30%. The volumetric gas rate was kept in 0.003 - 0.008 vvm of total mix (CO_2, O_2 and Air).

Offline measurement

Viable cell density was obtained by optical microscopy technique using trypan blue dye exclusion method. The specific growth and production rate were calculated by mass balance equations showed below (*Doran, 1995*).

Continuous culture biomass balance with and without cell retention device

$\frac{dx_v}{\mathrm{dt}} = \mu_{\mathrm{ap}} x_v - \alpha \Delta \xi_v$ Eq. 1

$\mu_{\mathrm{ap}} = \mu - k_d$ Eq. 2

$\alpha = \frac{X_H}{X_F}$ Eq. 3

Continuous culture extracellular product balance

$\frac{dP}{dt} = q_P x_v - DP$ Eq. 4

Where:

μ: Specific growth rate (η^{-1}) .

μ_{ap}: Specific growth rate apparent (η^{-1}) .

Xv: Viable cell density in bioreactor (cell/mL).

kd: Specific death rate (h-1).

a: Cell leakage factor

D: Dilution rate (d-1)

XH: Total cell density in the harvest stream (cell/mL)

XF: Total cell density in the bioreactor (cell/mL)

Π: Antibody concentration (μg/μL)

χΠ: Specific product rate (μg/cell/η)

Numerical Assessment of process parameters and cause-effect matrix

The determination of critical parameters was carried out by working together with the specialists and supervisors with most experience in the process, and a Cause-Effect matrix risk model was used.

A first determination of those unit operations and stages of the process that most impact on each of the QA, was carried out. For each of the operations/stages of the process, defined with the biggest impact, the input parameters, operation parameters and output attributes that characterize their status, were established.

First, the effect of QA variability over the next process step and the final product quality was numerically evaluated considering different criticality levels. Then, the numerical assessment is performed for each process parameter, taking into account the effect of its variability on each critical quality attribute (Table 1), and a cause-effect matrix was used as indicated in Table 2

Data preprocessing and Principal Component Analysis (PCA) by The UNSCRAMBLER software

The data preprocessing was carried out by using The UNSCRAMBLER software, version 8.0. THE UNSCRAMBLER is a comprehensive data analysis tool for exploratory statistics, multivariate analysis, classification, prediction and experiment design. It is easy to use and stands out for its excellent ability to graphically represent data and results. This program is widely used by many industries; especially it has been applied in pharmaceutical and biotechnology successfully (*The Unscrambler* , 2006). The multivariate tool applied for the current work was the PCA, commonly used for dimensionality reduction into a small number of uncorrelated variables called principal components (PC's) (*Maadooliat et al., 2014*). The data were transformed by centering and auto-scaling prior to processing.

RESULTS AND DISCUSSION

Critical parameters to assess in the fermentation process

The QA's used in this study were selected by the available information related to the process from the expert criteria, regarding the influences on the next step of the process. Only the fermentation process was taken into consideration for the current work.

The numerical valuation was performed taking into consideration the criticality level of each QA according to Table 3. Afterward, a cause-effect risk matrix was built by combining both, the valuation criteria of

parameters and the criticality levels of QA, and thus, new hypothesis can be generated to find out those parameters with higher impact over QA (*Suarez-Zuluaga et al., 2019*). Then, the index assigned to each parameter was evaluated as a percent value referred to the total amount of parameters index, and hence, the values provide a measure of the greater or lesser impact of each process parameter on CQA's variability.

The cumulative percentage was determined following a Pareto analysis. Subsequently, a bar graph was made with the impact index of each parameter in descending form and the ascending line of the cumulative % was also represented, in the style of a Pareto graph. The results are shown in Table 4 and Figure 1.

Applying the 80/20% Pareto criterion, 26 process parameters have been found as Figure 1 shows. Those parameters need to be taken into account for supervision and control.

In order to confirm the deduction of critical PP through knowledge and experience, PCA was applied to data obtained from routine monitoring, facilitating process variability analysis. This investigation was focused on the bioreactor, for that reason only the parameters highlighted in red in Table 4 were considered. Nevertheless, some of those parameters related to the bioreactor are not currently measured and some other have only one value at the beginning of the run. Hence, eleven variables were taken into account for this analysis, which constitute the Bioreactor critical parameters to be analyzed (Table 5).

Data understanding

The production process of the monoclonal antibody of interest, has shown a low yield and especially the protein concentration exhibits a high variability. PCA modelling was applied to five fermentation runs already available. The process was carried out in continuous operation with perfusion, taking into account that it is the technological production platform in question, and each sample correspond to a fermentation day.

Data pre-processing

For the data pre-processing, the characteristics of each data file were analyzed using the descriptive statistics, determining indicators for each variable, such as: minimum (Min.), maximum (Max), mean, standard deviation (SDev), the coefficient of variation (CV), skewness (S) and kurtosis (K). These indicators were found through histograms and statistical reports provided by THE UNSCRAMBLER, and were corroborated from the Microsoft Excel calculation tool.

Due to the heterogeneity of the data, motivated by the presence of variables of different nature and magnitudes, centering and auto-scaling were applied to provide standardization as part of preprocessing, which facilitated the analysis (*Coffey et al., 2018*).

Principal Component Analysis

As the Figure 2 shows, two principal components (PC) are sufficient to explain more than 80% of data variability at least. The PC1 allows to explain 74,61% of the explained variance, on the other hand, the PC2 allows to explain 23,39%.

In order to classify the samples regarding their leverage and residual X-variance, the influence graphic was elaborated taking into consideration two PC's. After making a look to the graphic (represented in Figure 3), few samples were found as outliers. Those points were not removed from data as they could represent useful information regarding the process, besides, those are not considered as dangerous outliers according to their leverage and residual X-variance.

The Figure 4 shows the Hotelling´s ellipse by considering 95% confidence. As the results show up, multiple samples were found inside the ellipse, which proves they belong to the same process. Nevertheless, some of them were placed outside the Hotelling´s ellipse, such us: m1, m73, m84, m93, m94, m101 and m102. This is usually related to specific changes on the process or to measuring errors and noise [17]. Actually, one of the aims of PCA is to detect deviations from the desired process behavior and state transitions (*Le et al., 2012; Nucci et al., 2010*).

For instance, samples m1, m73, m84, m93 and m94 match the change from discontinuous to continuous mode, which is characterized by low values of cell density and Feeding flow rate; and specifically, sample m93 correspond to the highest CSPR of the campaign.

On the other hand, samples m101 y m102 were related to mechanical issues due to the cell retention device, causing disturbances on both, Xv and Viability.

Also, the correlation loadings graph (see Figure 5) facilitate the determination of each variable influence on each component, depending on how far they are from the origin of coordinates, and furthermore, to identify the correlated variables depending on how close they are to each other.

Figure 5 shows that all variables are in between the two rings, therefore they have a considerable contribution to process variability. In addition, a group of variables can be observed as very close to each other such as: temperature (T), culture time (t), pH, stirring speed (rpm), dissolved oxygen (%DO) and cellular viability (Viab), which might indicate the existence of correlation between these parameters; while the specific rates are further separated from the Feeding flow rates and CSPR, therefore not showing a mathematical correlation each other. Some of parameters found as relevant and the possible existence of correlation were obtained by other authors, and some others not*(Thomassen et al., 2010 Suarez-Zuluaga et al., 2019)* . For instance, Thomassen *et al.* found DO and pH as relevant variables for inactivated polio vaccine (IPV) production process but they were not related each other (*Thomassen et al., 2010)* . On the other hand, Suares-Zuluaga *et al.* found specific growth rate as one of the most important parameters (indirectly controlled and linked to DO), nevertheless, DO, pH and other PP did not change, for that reason they were discarded from their further analysis (*Suarez-Zuluaga et al., 2019)* . Furthermore, Konstantinov *et al.* found a clear relationship between the physiological state variables (Ex: specific rates) and those linked to nutritional state such as CSPR and F, where limitation or inhibition conditions might lead the perfusion process to break (*Konstantinov et al., 2006)* .

PCA analysis was employed by Nucci *et al.* to analyze some of the parameters mentioned before, applying in that case an online verification to lead the process through optimal trajectories (*Nucci et al., 2010)* . From the present work it is determined that all of them should be considered to include in the process control strategy.

CONCLUSION

The approach applied in the present work, based on the proper combination of risk analysis and data advanced processing using MVDA tools, showed its potential facing the complexities of a mammalian cell culture process. Cause-Effect risk matrix model facilitated critical process parameters deduction using process knowledge and experience, subsequently confirmed through PCA modelling on data obtained from process routine monitoring. Likewise, PCA proved to be a useful tool for characterizing process variability patterns and identifying abnormal operational conditions related to different eventualities. Relevant results also found in this case study leads to update parameters in order to elaborate a more solid continued process verification plan, and thus, consolidating the cell culture control strategy.

Acknowledgements:

This study was supported by the Center of Molecular Immunology, Biosimilar Monoclonal Antibodies Division.

Conflict of interest:

Lisandra Calzadilla, Osvaldo Gozá, Mayla Echazabal, Arturo Toledo, and Tammy Boggiano declare that they have no conflict of interest.

REFERENCES

- Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics* , *2* (4), 433–459.*https://doi.org/10.1002/wics.101*

- Coffey, T., & Yang, H. (2018). *Statistics for Biotechnology Process Development* (1st ed.). FL, USA: Chapman and Hall/CRC.
- Doran, P. M. (1995). *Bioprocess Engineering Principles* (1st ed.). Sydney, Australia: Academic Press.
- Farid, S. S. (2006). Established Bioprocesses for Producing Antibodies as a Basis for Future Planning. *Adv Biochem Engin/Biotechnol* , *101* , 1–42.*https://doi.org/10.1007/10_014*
- Fike, R. (2014, August 18). Nutrient Supplementation Strategies for Biopharmaceutical Production. Retrieved from*https://bioprocessintl.com/upstream-processing/biochemicals-raw-materials/nutrient-supplementation-strategies-for-biopharmaceutical-production-185084/*
- Folch-Fortuny, A., Arteaga, F., & Ferrer, A. (2015). PCA model building with missing data: New proposals and a comparative study. *Chemometrics and Intelligent Laboratory Systems* , *146* , 77–88.*https://doi.org/10.1016/j.chemolab.2015.05.006*
- Kantardjieff, A., & Zhou, W. (2014). Mammalian Cell Cultures for Biologics Manufacturing. *Adv Biochem Eng Biotechnol* , 139, 1–9.*https://doi.org/10.1007/10_2013_255*
- Kirdar, A. O., Conner, J. S., Baclaski, J., & Rathore, A. S. (2007). Application of Multivariate Analysis toward Biotech Processes: Case Study of a Cell-Culture Unit Operation. *Biotechnology Progress* , *23* (1), 61–67.*https://doi.org/10.1021/bp060377u*
- Konstantinov, K., Goudar, C., Ng, M., Meneses, R., Thrift, J., Chuppa, S. ... Naveh, D. (2006). The "Push-to-Low" Approach for Optimization of High-Density Perfusion Cultures of Animal Cells. *Adv Biochem Engin/Biotechnol* , *101* , 75–98.*https://doi.org/10.1007/10_016*
- Le, H., Kabbur, S., Pollastrini, L., Sun, Z., Mills, K., Johnson, K. ... Hu, W.-S. (2012). Multivariate analysis of cell culture bioprocess data—Lactate consumption as process indicator. *Journal of Biotechnology* , *162* (2–3), 210–223.*https://doi.org/10.1016/j.jbiotec.2012.08.021*
- Lin, H., Leighty, R. W., Godfrey, S., & Wang, S. B. (2017). Principles and approach to developing mammalian cell culture media for high cell density perfusion process leveraging established fed-batch media. *Biotechnology Progress* , *33* (4), 891–901.*https://doi.org/10.1002/btpr.2472*
- Maadooliat, M., Huang, J. Z., & Hu, J. (2015). Integrating Data Transformation in Principal Components Analysis. *Journal of Computational and Graphical Statistics* , *24* (1), 84–103.*https://doi.org/10.1080/10618600.2014.891461*
- Mitchell, M. (2017, September 29). Determining Criticality-Process Parameters and Quality Attributes Part I: Criticality as a Continuum. Retrieved from*http://www.biopharminternational.com/determining-criticality-process-parameters-and-quality-attributes-part-i-criticality-continuum*
- Mitchell, M. (2018, May 16). Determining Criticality-Process Parameters and Quality Attributes Part III: Process Control Strategies-Criticality throughout the Lifecycle. Retrieved from*http://www.biopharminternational.com/determining-criticality-process-parameters-and-quality-attributes-part-iii-process-control-strategie*
- Nucci, E. R., Cruz, A. J. G., & Giordano, R. C. (2009). Monitoring bioreactors using principal component analysis: production of penicillin G acylase as a case study. *Bioprocess and Biosystems Engineering* , *33* (5), 557–564.*https://doi.org/10.1007/s00449-009-0377-y*
- Process Validation: General Principles and Practices. (2018, August 24). Retrieved from*https://www.fda.gov/regulatory-information/search-fda-guidance-documents/process-validation-general-principles-and-practices*
- Rathore, A. S. (2009). Roadmap for implementation of quality by design (QbD) for biotechnology products. *Trends in Biotechnology* , *27* (9), 546–553.*https://doi.org/10.1016/j.tibtech.2009.06.006*
- Sokolov, M., Ritscher, J., MacKinnon, N., Bielser, J.-M., Brühlmann, D., Rothenhäusler, D. ... Butté, A. (2016). Robust factor selection in early cell culture process development for the production of a biosimilar monoclonal antibody. *Biotechnology Progress* , *33* (1), 181–191.*https://doi.org/10.1002/btpr.2374*
- Sokolov, M., Soos, M., Neunstoecklin, B., Morbidelli, M., Butté, A., Leardi, R. ... Broly, H. (2015). Fingerprint detection and process prediction by multivariate analysis of fed-batch monoclonal antibody cell culture data. *Biotechnology Progress* , *31* (6), 1633–1644.*https://doi.org/10.1002/btpr.2174*
- Suarez-Zuluaga, D. A., Borchert, D., Driessen, N. N., Bakker, W. A. M., & Thomassen, Y. E. (2019).

Accelerating bioprocess development by analysis of all available data: A USP case study. *Vaccine* , *37* (47), 7081–7089.*https://doi.org/10.1016/j.vaccine.2019.07.026*

- The Unscrambler (Version 8.0) [Software]. (2006).
- Thomassen, Y. E., van Sprang, E. N. M., van der Pol, L. A., & Bakker, W. A. M. (2010). Multivariate data analysis on historical IPV production data for better process understanding and future improvements. *Biotechnology and Bioengineering* , *107* (1), 96–104.*https://doi.org/10.1002/bit.22788*
- von Stosch, M., Hamelink, J.-M., & Oliveira, R. (2016). Hybrid modeling as a QbD/PAT tool in process development: an industrial E. coli case study. *Bioprocess and Biosystems Engineering* , *39* (5), 773–784.*https://doi.org/10.1007/s00449-016-1557-1*
- Witcher, M. F. (2017). Vol 15, No 4 (2016/2017) - Quality Risk Management (QRM): Evaluating the Impact of Process Parameters on Critical Quality Attributes for Biopharmaceutical Products. Retrieved from*https://bioprocessingjournal.com/index.php/article-downloads/762-j154-quality-risk-management-qrm-evaluating-the-impact-of-process-parameters*
- Xu, S., Gavin, J., Jiang, R., & Chen, H. (2017). Bioreactor productivity and media cost comparison for different intensified cell culture processes. *Biotechnology Progress* , *33* (4), 867–878.*https://doi.org/10.1002/btpr.2415*

Table 1. Valuation criteria according to the effect of parameters variability on each supernatant quality attribute.

Valuation elements
Variability directly affects the critical quality attribute, even overcoming intermediate stages of the process where there are
Variability indirectly affects the critical quality attribute, and/or intermediate stages of the process, whose performance has
Variability affect the process stages performance, although with little or none influence over the critical quality attribute. T
Variability has little influence on the process stages performance without affecting any of critical quality attributes. There i

Table 2. Cause-effect matrix

Quality Attribute	QA_1	QA_2	...	QA_n	Total	% TT
Criticality level	N_1	N_2	...	N_n		
Operation/stage 1						
Parameter 1	M_1	M_2	...	M_m	T_1	PC_1
Parameter 2	:	:	:	:	:	:
:	:	:	:	:	:	:
Operation/Stage 2						
Parameter 3	:	:	:	:	:	:
:	:	:	:	:	:	:
Overall total:	Overall total:	Overall total:	Overall total:	Overall total:	TT	100

Where:

Nn: Criticality level of critical quality attributes.

Mn: Criticality level of each parameter over critical quality attributes.

Table 3. Valuation criteria according to the variability effect of quality attributes of supernatant over the next process steps and product quality.

Valuation elements
The attribute has a potential impact on the performance of downstream processes and can even be negatively reflected in t

Valuation elements
The attribute has a potential impact on the performance of downstream processes, and can even be negatively reflected in t
The attribute has a potential impact on the performance of downstream processes, although without negative effect on the
The attribute has no impact on the performance of downstream processes.

Table 4. Pareto criteria applied to process parameters

Id.	**Parameter**	**Process step**	**Valuation Index**	**% total**	**% cumulat**
24	Viability	500 L bioreactor	193	4.89	4.89
25	Temperature	500 L bioreactor	191	4.84	9.74
26	Acid-Base equilibrium	500 L bioreactor	185	4.69	14.43
31	Specific growth rate	500 L bioreactor	173	4.39	18.81
33	Metabolic content in culture media	500 L bioreactor	159	4.03	22.84
20	Culture time	500 L bioreactor	157	3.98	26.83
34	Efficiency of the perfusion system device	500 L bioreactor	151	3.83	30.65
36	Viable cell density	500 L bioreactor	151	3.83	34.48
37	CSPR	500 L bioreactor	151	3.83	38.31
32	Specific production rate	500 L bioreactor	131	3.32	41.63
39	Supernatant's holding time	Harvest	131	3.32	44.95
35	Osmolality	500 L bioreactor	128	3.25	48.2
23	Stirrer speed	500 L bioreactor	105	2.66	50.86
3	Biological activity	Thawing	101	2.56	53.42
22	Inoculum Viability	500 L bioreactor	99	2.51	55.93
27	Disolved oxigen	500 L bioreactor	97	2.46	58.39
38	Supernatant's holding temperature	Harvest	97	2.46	60.85
28	Feeding flow rate	500 L bioreactor	89	2.26	63.11
10	Viability	Stationary flasks expansion	85	2.16	65.26
12	Viable cell density	Stationary flasks expansion	85	2.16	67.42
18	Viability	Shaken flasks expansion	85	2.16	69.57
19	Viable cell density	Shaken flasks expansion	85	2.16	71.73
21	Inoculum cell density	500 L bioreactor	85	2.16	73.88
6	IgG concentration	Thawing	77	1.95	75.84
11	Osmolality	Stationary flasks expansion	73	1.85	77.69
30	Pressure	500 L bioreactor	71	1.8	79.49
5	Viable cell density	Thawing	65	1.65	81.14
7	Temperature	Stationary flasks expansion	65	1.65	82.78
8	Inoculum cell density	Stationary flasks expansion	65	1.65	84.43
9	Metabolic content in culture media	Stationary flasks expansion	65	1.65	86.08
13	Temperature	Shaken flasks expansion	65	1.65	87.73
16	Metabolic content in culture media	Shaken flasks expansion	65	1.65	89.38
29	Gas flow	500 L bioreactor	65	1.65	91.02
1	Centrifugation time	Thawing	59	1.5	92.52
2	Centrifugation speed	Thawing	59	1.5	94.02
4	Viability	Thawing	59	1.5	95.51
14	Inoculum cell density	Shaken flasks expansion	59	1.5	97.01
15	Shaker speed	Shaken flasks expansion	59	1.5	98.5
17	Osmolality	Shaken flasks expansion	59	1.5	100

Table 5. Critical parameters of the bioreactor

No.	Critical parameter	Notation	Measurement units
1	Viable cell density	Xv	cel/mL
2	Viability	Viab	%
3	Temperature	T	ºC
4	Acid-base equilibrium (pH)	pH	-
5	Feeding flow rate	F	L/d
6	Stirrer speed	rpm	min^{-1}
7	Dissolved oxygen	%DO	%
8	Specific production rate	qP	µg/cel*h
9	Specific growth rate	Vc	h^{-1}
10	Culture time	t	d
11	Cell specific perfusion rate	CSPR	nL/cel*d

Figure Legends

Figure 1. Process parameters assessment by applying Pareto criteria. On the left axis: Valuation Index represented by the blue bars. On the right axis: % cumulative represented by red line.

Figure 2. Explained variance. Two PC's are sufficient to explain more than 80% of data variability.

Figure 3. Influence graph applied to five fermentation runs for the assayed year. On X-axis: leverage. On Y-axis: Residual X-variance

Figure 4. Scores graph and Hotelling´s ellipse represented for a 95% of confidence. (Applied to five fermentation runs for the assayed year). On X-axis: PC1. On Y-axis: PC2.

Figure 5. Correlation loading applied to five fermentation runs for the assayed year. On X-axis: PC1. On Y-axis: PC2