UNICA

UNIVERSITÀ
DEGLI STUDI
DI CAGLIARI

Ph.D. DEGREE IN

Industrial Engineering

Cycle XXXV

TITLE OF THE Ph.D. THESIS

Development of monitoring and control systems for biotechnological processes

Scientific Disciplinary Sector(s)

ING-IND/26

Ph.D. Student:	Silvia Lisci
Supervisor	Prof. Ing. Massimiliano Grosso
Co-Supervisor	Prof.ssa Ing. Stefania Tronci

Final exam. Academic Year 2021/2022
Thesis defence: April 2023 Session

"Silvia Lisci gratefully acknowledges the Sardinian Regional Government for the financial support of her/his PhD scholarship (P.O.R. Sardegna F.S.E. - Operational Programme of the Autonomous Region of Sardinia, European Social Fund 2014-2020 - Axis III Education and training, Thematic goal 10, Investment Priority 10ii), Specific goal 10.5".

I would like to express my deep gratitude to Prof. Massimiliano Errico for hosting me during my time abroad at the University of Southern Denmark (SDU), for his generous help and constant guidance.

Contents

Abstract

The field of biotechnology represents an important research area that has gained increasing success in recent times. Characterized by the involvement of biological organisms in manufacturing processes, its areas of application are broad and include the pharmaceuticals, agri-food, energy, and even waste treatment. The implication of living microorganisms represents the common element in all bioprocesses. Cell cultivations is undoubtedly the key step that requires maintaining environmental conditions in precise and defined ranges, having a significant impact on the process yield and thus on the desired product quality. The apparatus in which this process occurs is the bioreactor. Unfortunately, monitoring and controlling these processes can be a challenging task because of the complexity of the cell growth phenomenon and the limited number of variables can be monitored in real-time.

The thesis presented here focuses on the monitoring and control of biotechnological processes, more specifically in the production of bioethanol by fermentation of sugars using yeasts. The study conducted addresses several issues related to the monitoring and control of the bioreactor, in which the fermentation takes place. First, the topic concerning the lack of proper sensors capable of providing online measurements of key variables (biomass, substrate, product) is investigated. For this purpose, nonlinear estimation techniques are analyzed to reconstruct unmeasurable states. In particular, the geometric observer approach is applied to select the best estimation structure and then a comparison with the extended Kalman filter is reported. Both estimators proposed demonstrate good estimation capabilities as input model parameters vary. Guaranteeing the achievement of the desired ethanol composition is the main goal of bioreactor control. To this end, different control strategies, evaluated for three different scenarios, are analzyed. The results show that the MIMO system, together with an estimator for ethanol composition, ensure the compliance with product quality.

After analyzing these difficulties through numeric simulations, this research work shifts to testing a specific biotechnological process such as manufacturing bioethanol from brewery's spent grain (BSG) as renewable waste biomass. Both acid pre-treatment, which is necessary to release sugars, and fermentation are optimized. Results show that a glucose yield of 18.12 per 100 g of dried biomass is obtained when the pre-treatment step is performed under optimized conditions (0.37 M H_2SO_4, 10% S-L ratio). Regarding the fermentation, $T = 25°C$, $pH = 4.5$, and inoculum volume equal to $12.25\%\ v/v$ are selected as the best condition, at which an

ethanol yield of 82.67% evaluated with respect to theoretical one is obtained. As a final step, the use of Raman spectroscopy combined with chemometric techniques such as Partial Least Square (PLS) analysis is evaluated to develop an online sensor for fermentation process monitoring. The results show that the biomass type involved significantly affects the acquired spectra, making them noisy and difficult to interpret. This represents a nontrivial limitation of the applied methodology, for which more experimental data and more robust statistical techniques could be helpful.

Chapter 1: Introduction

This chapter provides the motivations for the thesis work here presented. In addition, the outline of the chapters is also shown, with a brief description for each of them. Conference participations and journal papers written during the Ph.D. program are listed at the end.

1.1 Motivations

The term biotechnology refers to the scientific discipline that has been particularly successful in recent decades as the main area of technological research and development, in which microbiology, biochemical, and molecular biology play a key role. As reported by Buchholz (2007), biotechnology (BT) can be defined as the "application of biological organisms or processes to the manufacturing industry". Biotechnological processes, also known as bioprocesses, have gained so much importance that they have become essential to any economy that aims to maintain a competitive status/condition in future markets. The areas of application are several and range from pharmaceuticals to agri-food, food processing, energy, to wastewater and biowaste treatment (Lourenço et al., 2012). Consequently, a substantial amount of different products and services such as antibiotics, therapeutic proteins, vaccines, value-added foods, food additives, vitamins, amino acids, and agricultural products, as well as fuels and industrial chemicals can be manufactured by bioprocesses (Gomes et al., 2019; Lòpez et al., 2020). Obviously, each individual process can be considered as a system by itself, and such as differs from the others even in terms of the goals and price of the final product. However, there is one common element to all bioprocesses, which is the involvement of microorganisms. As stated by Lourenço et al. (2012), the cultivation of microorganisms represents a crucial phase of the bioprocess that must necessarily be carried out under controlled conditions, since it is the result of a complex combination of physical (equilibrium and transport), chemical, and biological phenomena. The bioproduction platform includes several unit operations between upstream and downstream (Rathore et al., 2021), but the most important one is definitely the bioreactor, or fermenter, where cell cultivation takes place (Gomes et al., 2019). Since the yield of bioprocesses depends essentially on the performance of the cells, it is by controlling the bioreactor that the desired process profitability can be achieved. The control objective/goal/aim is generally the achievement of consistent production (yield) at a defined quality of the desired product (Gomes et al., 2019). Therefore, monitoring and maintaining, but also optimizing and efficiently controlling the bioprocess are fundamental and essential steps. This can be achieved by having a good control system that allows for automatic and frequent control of the bioprocess at its optimum point, in terms of efficiency, productivity, and reproducibility. All this would have positive impacts economically with significant cost reductions, but also in terms of maintaining industrial competitiveness (Alford, 2006; Gargalo et al., 2020, 2022; Lourenço et al., 2012). Unfortunately, the bioreactor also represents the most difficult unit operation to monitor and control because of the complex phenomena that occur during cell

growth. In particular, features that make monitoring and controlling bioprocessing, such as fermentation, challenging are the batch or fed-batch nature of various commercial processes, their being multiphase systems characterized by nonlinearity, the high interaction and correlation among process variables, as well as the nonstationary nature of the process (Gargalo et al., 2020; Stanke & Hitzmann, 2012; Veloso & Ferreira, 2017). Real-time monitoring of unit operations such the bioreactor is an essential element for the purpose of control and achievement of those objectives mentioned above. Indeed, it provides information regarding the bioprocess and its evolution that is useful for failure identification, and thus for possible corrective control actions. In this scenario, Process Analytical Technology (PAT) can help in the real-time understanding of the investigated bioprocess. It is defined as a system or tool for the design, analysis and control of production through measurements of quality and performance attributes aimed at ensuring the quality of the final product (Wechselberger et al., 2010). However, the main problem when considering bioreactors concerns the limited number of variables that can be measured online. Generally, only process variables, also known as engineering data, such as temperature, pressure, stirring speed, dissolved oxygen and so on, can be monitored in real-time (Lourenço et al., 2012; Schügerl, 2001). Regarding chemical components such as substrate, products or intermediates, which are usually monitored off-line with well-established but also time-consuming techniques such as HPLC, it is not always possible to get their measurement online due to the lack of adequate sensors. For this reason, offline monitoring is still performed. Samples are regularly taken from the bioreactor medium, and then analyzed. Obviously, offline measurements have unquestionable reliability. However, it is the time required to perform them that makes them unsuitable for process automation purposes (Gomes et al., 2019). Several efforts have been made in recent years to improve strategies for real-time monitoring bioprocesses by developing optical sensor, chemosensors, and biosensors (Gargalo et al., 2022; Goker et al., 2020; Gomes et al., 2019; Ulber & Sell, 2007). Despite the progress in the field of sensor technology, some online sensors for monitoring bioprocesses still have characteristics such as high price, frequent maintenance needs, and single-property analysis capabilities that limit their manufacturing-scale application. A possible and promising solution to overcome this problem is the design of tools known as observers or estimators in order to estimate state variables describing the internal state of the process, which is difficult to access, using the few available measurements. These observers are known also as software sensors or soft-sensors and represent advanced online process monitoring strategies. Although this concept was introduced in the 1990s, it is still a popular topic in both academia and industry, in fact several reviews and papers can be found

in the literature (Cabaneros Lopez et al., 2019; Chéruy, 1997; Gargalo et al., 2022; Gustavsson, 2018; Kadlec et al., 2009; Kadlec & Gabrys, 2009; Luttmann et al., 2012; Randek & Mandenius, 2018; Veloso & Ferreira, 2017). The process parameters or variables that are thus estimated can consequently be used to monitor the process online or even be integrated into control loops (Ajbar & Ali, 2017; Arndt & Hitzmann, 2004; Gargalo et al., 2022; Petre et al., 2021). As reported by Lourenço et al. (2012), the features that an ideal bioprocess online monitoring sensor must exhibit are several, but in particular it should be fast, sensitive, robust, and non-destructive. In this scenario, spectroscopy has attracted considerable interest in the area of bioprocess monitoring and control (Beutel & Henkel, 2011; Claßen et al., 2017; Lourenço et al., 2012). Several papers have been published in recent years concerning the potential applications of spectroscopic techniques such as UV/Vis, Near Infrared (NIR), Mid Infrared (NIR), Fluorescence, and Raman spectroscopy in the context of bioprocesses (Iversen et al., 2014; Ödman et al., 2009; Roychoudhury et al., 2007; Sarraguça et al., 2009; Schenk et al., 2007). The combination of large data sets collected in the form of spectra through these spectroscopic methods, with multivariate data analysis techniques (PCA, PLS, etc.) and other mathematical tools, i.e., chemometrics (Rathore et al., 2011), allow essential information to be extracted for the purpose of process understanding and its monitoring and control. Raman spectroscopy due to significant characteristics such as flexibility to the sample type, no sample preparation, good resolution, and no sensitivity to water presence, has proven to be a valuable tool for obtaining information about the compounds present in the bioreactor medium, and consequently for the development of quantitative analysis. Although encouraging and satisfactory results have been obtained in the monitoring of microbial cultivations as reported by several works in the literature, its application in this research area is still limited (Ávila et al., 2012; Hirsch et al., 2019; Picard et al., 2007; Schalk et al., 2017).

1.2 Thesis outline

The thesis presented here focused on the problem of monitoring and control of biotechnological processes. The manufacturing of bioethanol by fermentation of sugars carried out by was chosen as case study. Different issues related to the monitoring and control of the bioreactor, in which this process takes place, were addressed and investigated during the PhD programme. Particularly, the problem related to the lack of proper sensors to measure key process variables was analzyed and the development of a model-based soft-sensor for estimate the unmeasured

states was proposed as a possible solution. Since the quality of the final product is an essential parameter to control, different control strategies were proposed in order to ensure the achievement of the desired ethanol composition. After exploring these challenges theoretically by means of simulations, the focus shifted to a typical biotechnological process such as the process of bioethanol production from waste biomass, which was conducted experimentally and optimization of both acid pre-treatment and fermentation conditions was realized. Regarding this process, subsequently Raman spectroscopy was investigated, in combination with chemometric tools, in order to develop a quantitative analysis system for ethanol produced and glucose consumed during the process.

The thesis work will be presented in the following order given below, where a brief summary of each chapter is given:

Chapter 2 describes issues related to the monitoring of biotechnological processes. After an overview of key variables generally monitored during such processes, it focuses on state estimators as a possible solution. A fermentation process for bioethanol production for bioethanol production using *Saccharomyces cerevisiae* is taken as case study and the mathematical model of the process taken as reference is described in detail. A state observer to reconstruct the states of the system that cannot be measured is developed and implemented. Different situations were analyzed depending on the available measurements.

Chapter 3 addresses the problem of bioprocess control by first providing a general overview of the state of the art. Always considering the fermentation bioreactor for ethanol production as case study, three different control strategies were investigated: a direct temperature control (SISO), a cascade control where the primary loop exploits delayed ethanol composition measurements, and a MIMO control system with an inferential control for product concentration. The results of the simulations are reported and compared as the operating conditions change.

Chapter 4 focuses on one biotechnological process, i.e., the production of biofuels from waste biomass. The biomass investigated is the brewery's spent grain (BSG), and it is described in detail, both in terms of its chemical composition and possible applications in different areas. Experimental work carried out during the abroad research period is presented, particularly those related to the optimization of the operating conditions of the acid pre-treatment and fermentation phases.

Chapter 5 deals again with the problem of monitoring bioprocesses but proposing the development of a sensor based on Raman spectroscopy as a possible solution. This methodology is briefly explained along with the other main spectroscopic techniques. The case study analyzed is once again the fermentation process of waste biomass to produce second-generation ethanol. Data from Raman spectra of fermentation samples, derived from the process described in the previous chapter, were combined with chemometric techniques to develop a real-time quantitative analysis system.

Chapter 6 summarizes the conclusions.

1.3 Participation in conferences and publications in Journals

Some of the work presented in this thesis has been presented in national and international conferences and published in international journals.

Presentations at national and international conferences

Monitoring and control of a bioreactor for yeast fermentation. GRICU 2022 national conference, Ischia (Italy), July 3-9, 2022.

Brewer's Spent Grain: its Value as Renewable Biomass and its Possible Applications. IConBM2022 International Conference, Naples (Italy), June 5-8, 2022.

Different control strategies for a yeast fermentation bioreactor. ADCHEM2021 International Conference, fully virtual, June 13-16, 2021.

Modeling a Biological Reactor using Sparse Identification Method. ICheaP15 International Conference, fully virtual, May 23-26, 2021.

Publications in international journals

Lisci, S., Tronci, S., Grosso, M., Karring, H., Hajrizaj, R., & Errico, M. (2022). Brewer's Spent Grain: its Value as Renewable Biomass and its Possible Applications. Chemical Engineering Transactions, 92, 259-264. IConBM2022 International Conference, Naples (Italy).

Lisci, S., Grosso, M., & Tronci, S. (2021). Different control strategies for a yeast fermentation bioreactor. IFAC-PapersOnLine, 54(3), 306-311. ADCHEM2021 International Conference.

Lisci, S., Gitani, E., Mulas, M., & Tronci, S. (2021). Modeling a Biological Reactor using Sparse Identification Method. Chemical Engineering Transactions, 86, 901-906. ICheaP15 International Conference.

Lisci, S., Grosso, M., & Tronci, S. (2020). A Robust Nonlinear Estimator for a Yeast Fermentation Biochemical Reactor. In Computer Aided Chemical Engineering (Vol. 48, pp. 1303-1308). Elsevier. ESCAPE30 European Symposium.

Lisci, S., Grosso, M., & Tronci, S. (2020). A geometric observer-assisted approach to tailor state estimation in a bioreactor for ethanol production. Processes, 8(4), 480.

Bibliography

Ajbar, A. H., & Ali, E. (2017). Study of advanced control of ethanol production through continuous fermentation. *Journal of King Saud University - Engineering Sciences*, *29*(1), 1–11. https://doi.org/10.1016/j.jksues.2015.10.005

Alford, J. S. (2006). Bioprocess control: Advances and challenges. *Computers and Chemical Engineering*, *30*(10–12), 1464–1475. https://doi.org/10.1016/J.COMPCHEMENG.2006.05.039

Arndt, M., & Hitzmann, B. (2004). Kalman Filter Based Glucose Control at Small Set Points during Fed-Batch Cultivation of Saccharomyces cerevisiae. *Biotechnology Progress*, *20*(1), 377–383. https://doi.org/10.1021/bp034156p

Ávila, T. C., Poppi, R. J., Lunardi, I., Tizei, P. A. G., & Pereira, G. A. G. (2012). Raman spectroscopy and chemometrics for on-line control of glucose fermentation by Saccharomyces cerevisiae. *Biotechnology Progress*, *28*(6), 1598–1604. https://doi.org/10.1002/btpr.1615

Beutel, S., & Henkel, S. (2011). In situ sensor techniques in modern bioprocess monitoring. *Applied Microbiology and Biotechnology*, *91*(6), 1493–1505. https://doi.org/10.1007/S00253-011-3470-5

Buchholz, K. (2007). Science - Or not? The status and dynamics of biotechnology. In *Biotechnology Journal* (Vol. 2, Issue 9, pp. 1154–1168). https://doi.org/10.1002/biot.200600222

Cabaneros Lopez, P., Feldman, H., Mauricio-Iglesias, M., Junicke, H., Huusom, J. K., & Gernaey, K. v. (2019). Benchmarking real-time monitoring strategies for ethanol production from lignocellulosic biomass. *Biomass and Bioenergy*, *127*, 105296. https://doi.org/10.1016/J.BIOMBIOE.2019.105296

Chéruy, A. (1997). Software sensors in bioprocess engineering. *Journal of Biotechnology*, *52*(3), 193–199. https://doi.org/10.1016/S0168-1656(96)01644-6

Claßen, J., Aupert, F., Reardon, K. F., Solle, D., & Scheper, T. (2017). Spectroscopic sensors for in-line bioprocess monitoring in research and pharmaceutical industrial application. In *Analytical and Bioanalytical Chemistry* (Vol. 409, Issue 3, pp. 651–666). Springer Verlag. https://doi.org/10.1007/s00216-016-0068-x

Gargalo, C. L., Lopez, P. C., Hasanzadeh, A., Udugama, I. A., & Gernaey, K. v. (2022). On-line monitoring of process parameters during fermentation. *Current Developments in Biotechnology and Bioengineering*, 117–164. https://doi.org/10.1016/B978-0-323-91167-2.00003-4

Gargalo, C. L., Udugama, I., Pontius, K., Lopez, P. C., Nielsen, R. F., Hasanzadeh, A., Mansouri, S. S., Bayer, C., Junicke, H., & Gernaey, K. v. (2020). Towards smart biomanufacturing: a perspective on recent developments in industrial measurement and monitoring technologies for bio-based production processes. *Journal of Industrial Microbiology and Biotechnology*, *47*(11), 947–964. https://doi.org/10.1007/s10295-020-02308-1

Goker, M., Kurbanoglu, S., Huri, P. Y., Sezgintürk, M. K., & Ozkan, S. A. (2020). Bioprocess monitoring by biosensor-based technologies. *Commercial Biosensors and Their Applications: Clinical, Food, and Beyond*, 259–285. https://doi.org/10.1016/B978-0-12-818592-6.00010-4

Gomes, J., Chopda, V., & Rathore, A. S. (2019). *Monitoring and Control of Bioreactor: Basic Concepts and Recent Advances* (C. Komives & W. Zhou, Eds). Wiley & Sons.

Gustavsson, R. (2018). *Development of soft sensors for monitoring and control of bioprocesses*. Linköping University Electronic Press.

Hirsch, E., Pataki, H., Domján, J., Farkas, A., Vass, P., Fehér, C., Barta, Z., Nagy, Z. K., Marosi, G. J., & Csontos, I. (2019). Inline noninvasive Raman monitoring and feedback control of glucose concentration during ethanol fermentation. *Biotechnology Progress*, *35*(5). https://doi.org/10.1002/btpr.2848

Iversen, J. A., Berg, R. W., & Ahring, B. K. (2014). Quantitative monitoring of yeast fermentation using Raman spectroscopy. *Analytical and Bioanalytical Chemistry*, *406*(20), 4911–4919. https://doi.org/10.1007/s00216-014-7897-2

Kadlec, P., & Gabrys, B. (2009). Soft sensors: where are we and what are the current and future challenges? *IFAC Proceedings Volumes*, *42*(19), 572–577. https://doi.org/10.3182/20090921-3-TR-3005.00098

Kadlec, P., Gabrys, B., & Strandt, S. (2009). Data-driven Soft Sensors in the process industry. *Computers & Chemical Engineering*, *33*(4), 795–814. https://doi.org/10.1016/J.COMPCHEMENG.2008.12.012

Lòpez, R. A., Pérez, P. A. L., & Femat, R. (2020). *Control in Bioprocessing: Modeling, Estimation and the Use of Soft Sensors* (John Wiley & Sons, Ed.). John Wiley & Sons.

Lourenço, N. D., Lopes, J. A., Almeida, C. F., Sarraguça, M. C., & Pinheiro, H. M. (2012). Bioreactor monitoring with spectroscopy and chemometrics: A review. In *Analytical and Bioanalytical Chemistry* (Vol. 404, Issue 4, pp. 1211–1237). https://doi.org/10.1007/s00216-012-6073-9

Luttmann, R., Bracewell, D. G., Cornelissen, G., Gernaey, K. v., Glassey, J., Hass, V. C., Kaiser, C., Preusse, C., Striedner, G., & Mandenius, C. F. (2012). Soft sensors in bioprocessing: A status report and recommendations. In *Biotechnology Journal* (Vol. 7, Issue 8, pp. 1040–1048). https://doi.org/10.1002/biot.201100506

Ödman, P., Johansen, C. L., Olsson, L., Gernaey, K. v., & Lantz, A. E. (2009). On-line estimation of biomass, glucose and ethanol in Saccharomyces cerevisiae cultivations using in-situ multi-wavelength fluorescence and software sensors. *Journal of Biotechnology*, *144*(2), 102–112. https://doi.org/10.1016/J.JBIOTEC.2009.08.018

Petre, E., Selişteanu, D., & Roman, M. (2021). Advanced nonlinear control strategies for a fermentation bioreactor used for ethanol production. *Bioresource Technology*, *328*. https://doi.org/10.1016/j.biortech.2021.124836

Picard, A., Daniel, I., Montagnac, G., & Oger, P. (2007). In situ monitoring by quantitative Raman spectroscopy of alcoholic fermentation by Saccharomyces cerevisiae under high pressure. *Extremophiles*, *11*(3), 445–452. https://doi.org/10.1007/s00792-006-0054-x

Randek, J., & Mandenius, C. F. (2018). On-line soft sensing in upstream bioprocessing. In *Critical Reviews in Biotechnology* (Vol. 38, Issue 1, pp. 106–121). Taylor and Francis Ltd. https://doi.org/10.1080/07388551.2017.1312271

Rathore, A. S., Bhushan, N., & Hadpe, S. (2011). Chemometrics applications in biotech processes: A review. In *Biotechnology Progress* (Vol. 27, Issue 2, pp. 307–315). https://doi.org/10.1002/btpr.561

Rathore, A. S., Mishra, S., Nikita, S., & Priyanka, P. (2021). Bioprocess control: Current progress and future perspectives. In *Life* (Vol. 11, Issue 6). MDPI AG. https://doi.org/10.3390/life11060557

Roychoudhury, P., O'Kennedy, R., McNeil, B., & Harvey, L. M. (2007). Multiplexing fibre optic near infrared (NIR) spectroscopy as an emerging technology to monitor industrial bioprocesses. *Analytica Chimica Acta*, *590*(1), 110–117. https://doi.org/10.1016/j.aca.2007.03.011

Sarraguça, M. C., Paulo, A., Alves, M. M., Dias, A. M. A., Lopes, J. A., & Ferreira, E. C. (2009). Quantitative monitoring of an activated sludge reactor using on-line UV-visible and near-infrared spectroscopy. *Analytical and Bioanalytical Chemistry*, *395*(4), 1159–1166. https://doi.org/10.1007/s00216-009-3042-z

Schalk, R., Braun, F., Frank, R., Rädle, M., Gretz, N., Methner, F. J., & Beuermann, T. (2017). Non-contact Raman spectroscopy for in-line monitoring of glucose and ethanol during yeast fermentations. *Bioprocess and Biosystems Engineering*, *40*(10), 1519–1527. https://doi.org/10.1007/s00449-017-1808-9

Schenk, J., Marison, I. W., & von Stockar, U. (2007). Mid-Infrared spectroscopy as a tool for pH prediction and control in bioprocesses. *Journal of Biotechnology*, *2*(131).

Schügerl, K. (2001). Progress in monitoring, modeling and control of bioprocesses during the last 20 years. *Journal of Biotechnology*, *85*(2), 149–173. https://doi.org/10.1016/S0168-1656(00)00361-8

Stanke, M., & Hitzmann, B. (2012). Automatic control of bioprocesses. In *Measurement, Monitoring, Modelling and Control of Bioprocesses* (pp. 35–63).

Ulber, R., & Sell, D. (2007). *White Biotechnology* (Vol. 105). Springer.

Veloso, A. C., & Ferreira, E. C. (2017). Online Analysis for Industrial Bioprocesses: Broth Analysis. In *Current Developments in Biotechnology and Bioengineering: Bioprocesses, Bioreactors and Controls* (pp. 679–704). Elsevier Inc. https://doi.org/10.1016/B978-0-444-63663-8.00023-9

Wechselberger, P., Seifert, A., & Herwig, C. (2010). PAT method to gather bioprocess parameters in real-time using simple input variables and first principle relationships. *Chemical Engineering Science*, *65*(21), 5734–5746. https://doi.org/10.1016/j.ces.2010.05.002

Chapter 2: The issue of bioprocess monitoring

This Chapter addresses the problem of monitoring biotechnological processes, first describing an overview of the most measured variables and estimation techniques available in the literature. The estimation problem of a fermentation reactor will be presented as case study, the results of which will be reported and discussed.

The increasing development in recent years of biotechnological processes (BTP) and their industrial applications have confirmed the importance of adequately monitoring and controlling them in order to guarantee profitability and quality. Generally, the success of process monitoring and control depends significantly on proper/appropriate and functioning measurement and monitoring techniques (Sonnleitner, 2012). The development of a fully automated monitoring and control system represents a fundamental element in such processes, particularly in those operating in continuous mode, in which possible deviations/malfunctions must be identified and corrected in time, leading to robust production with constant performance. However, in the context of bioprocesses, the demands on measurements represent a challenging task. Indeed, although physic-chemical parameters such as temperature, pH, aeration, agitation, and dissolved oxygen are easily measurable, the same cannot be said for critical biological variables such as the concentration of biomass, substrate, products, or by-products (Liu et al., 2001). These represent so-called product quality indices that must be measured in order to meet the mandatory requirements, and the reason why it is often difficult to measure these variables in real-time can be found in two main causes. First, biological processes/systems are characterized by a certain complexity resulting from the involvement of living organisms, which have a complex nature that is difficult to analyze and understand. On the other hand, it is also necessary to consider the systematic lack of measuring devices capable of providing measurements for understanding the functioning of the bioprocesses (Holzberg et al., 2018; Lyubenova et al., 2021). This means that the application of the continuous production mode, to which biotechnological processes are also beginning to convert, cannot be considered unless an adequate monitoring and control system can be implemented, and this cannot be satisfied if appropriate sensors are not available.

A possible solution can be the design and the implementation of nonlinear estimators, also known as model-based software sensors and state observers. These are estimation techniques obtained from the combination of a measurement processor and the mathematical model of the process. Indeed, they rely on first principles process models, i.e., balance equations for mass and energy as well as constitutive equations, and also on an estimation algorithm that reconciles the available measurements with the prediction of the model (Cabaneros Lopez et al., 2019a). In this way, it is possible to estimate those variables whose measurements are not simple to obtain or that are too time-consuming (Veloso & Ferreira, 2017), from secondary measurements that are more accessible such as the environmental ones (temperature, pH, dissolved oxygen concentration, ect.) (Lyubenova et al., 2021). Although soft sensor strategy has been used for decades in chemical industries, in biochemical ones it is quite recent and still under

development. In any case, remarkable applications and reviews of soft sensing can be found in the literature, confirming that this is certainly a strategy that generates interest in research, and that is able to reconstruct the evolution of key variables over time, despite the process-model mismatch, nonlinear dynamic and measurement noise, which usually make its development challenging (Gargalo et al., 2022; Cabaneros Lopez et al., 2019;Holzberg et al., 2018; Lopez et al., 2020; Lourenço et al., 2012; Mauricio-Iglesias et al., 2015; Ödman et al., 2009). Among the various estimation techniques proposed in the literature, both for chemical and biochemical processes, those that have proved to have a strong potential for the online estimation of nonlinear systems are the following: extended Kalman filter (EKF) (Jazwinski, 2007); high gain observer (Ciccarella et al., 1993); sliding mode observer (G.-B. Wang et al., 1997); geometric observer (Alvarez & López, 1999; Zeitz, 1987). However, many of the strategies to estimate unmeasurable states and disturbances for partially known systems are based on the extended Kalman filter (EKF) being its design quite simple and its application accepted by relevant industries (Dewasme et al., 2013; Longhi et al., 2002).

In this chapter the problem of real-time monitoring bioprocesses will be addressed/discussed. A brief presentation of the main variables usually monitored in processes like these will be described, as well as the techniques mainly used, also focusing on the most difficult to measure. Later, an overview of the principal estimation techniques is presented, dwelling on the geometric observer and the Kalman filter algorithms since they are the estimation strategies applied in the case study presented here. Regarding this, the problem of state estimation in a bioreactor for ethanol production is addressed. The geometric observer estimation algorithm is used to select the best estimation configuration in terms of states and measured outputs. Then, the Kalman filter algorithm is applied in the validation phase. A series of simulations are carried out in order to evaluate the performance of the designed estimation structures, the results of which are shown and discussed.

2.1 Overview of standard variables monitored in bioprocesses

As stated in the introduction, the greatest difficulty in implementing a proper control system for a bioprocess is the real-time monitoring, since there is a lack of valid sensors providing real-time information about the concentration of products, substrate, and biomass employed (Edwiges et al., 2022). The few variables commonly measured online in bioreactors are temperature, pressure, dissolved oxygen, pH, stirring speed, and liquid flow rate. These are usually referred to as process variables or engineering data as reported by Lourenço et al.

(2012). In recent years, there has been considerable effort in developing efficient monitoring strategies for bioprocesses such as fermentation, including the development of biosensors, optical sensor, and chemosensors (Holzberg et al., 2018). Although such online sensors developed for bioprocess monitoring can boast features such as reliability and advanced performance, their application at the industrial level still seems to be far off due to their high price and frequent maintenance (Mitra & Murthy, 2022).

The techniques commonly used to monitor the variables involved in bioprocesses will be described below, classified into physical, chemical, and biological variables.

2.1.1 Physical variables

Among the physical variables, surely the one that is always monitored and determined directly is temperature which plays a key role in the growth of the microorganisms involved, thus influencing the production rate and the evolution of the bioprocess. The temperature measurement is preferably carried out using resistance thermometers (PT-100 or PT-1000) due to their accuracy and reproducibility (Pörtner et al., 2017). Alternatively, thermocouples can also be used, which are certainly cheaper but also less accurate (A. Doyle & Griffiths, 1998; Ozturk & Hu, 2005). The pressure, which is usually not measured on very small scales, is mainly measured for safety reasons. However, it plays an important role in the media sterilization, contributing to its maintenance. The standard measurement sensors are membrane pressure gauges based on capacitance or strain measurements (Caramihai & Severi, 2013).

2.1.2 Chemical variables

In terms of chemical variables, pH represents definitely the most important one, and for this reason it must be kept with a very constrained range. Usually, the pH is controlled manually by adding proper amounts of acidic or basic solutions, depending on the evolution of the pH trend. The measuring device used for pH consists of a combination in one body of both glass and reference electrodes (Pörtner et al., 2017). The concentration of dissolved oxygen is also critical to monitor, especially for fermentation processes that evolve in an aerobic mode. Indeed, it is essential that the value of this variable does not fall below a specified minimal level. Galvanic electrodes are usually used for its measurement in small fermenters, while in the case of pilot or production bioreactors, polarographic electrodes are employed.

2.1.3 Biological variables

Biological variables refer to biomass, substrates used for the microorganism growth, and finally products and/or by-products. Although these represent the most crucial variables when evaluating the evolution of bioprocesses, their online monitoring still seems to be a very difficult goal to achieve. Standard offline reference methods such as gas chromatography (GC), liquid chromatography (LC), and optical density (OD) measurements appear as the most trusted ones. However, research has made many strides in this area. As reported by Holzberg et al. (2018) in their review, several sensors have been developed on a laboratory scale for real-time monitoring of these variables. These are mainly electrochemical and optical sensors that have been developed to detect analytes such as organic compounds (glucose, glutamine, and lactate) but also cell biomass, products (ethanol), and by-products (ammonia and protein products). These have been classified according to their potential application in continuous production processes, where precisely their presence is essential. From this analysis it was found that electrochemical sensors are not suitable for monitoring in this type of processes as they are unable to provide continuous online measurements, require high-frequency sampling and destructive analysis (Bäcker et al., 2013; Derfus et al., 2010; Mross, Fürst, et al., 2015; Mross, Zimmermann, et al., 2015; Schlueter et al., n.d.; Weltin et al., 2014). However, optical sensors have proven to have good potential. They include different types of spectroscopy, generally couples with optical fibers, which cover different wavelength ranges such as visible-ultraviolet (UV-Vis), near infrared (NIR), mid infrared (MIR), Raman, and fluorescence spectroscopy (Lourenço et al., 2012). Their use has been investigated in various in situ applications (Abu-Absi et al., 2011; Esmonde-White et al., 2017; Lindner et al., 2014; Luttmann et al., 2012; Ohadi et al., 2015; Rowland-Jones et al., 2017; Tamburini et al., 2014; Theuer et al., 2017), and what has emerged is that they represent certainly rapid, non-invasive, non-destructive techniques, with high accuracy and precision, and good stability. Obviously, they have some limits, strictly connected to the performance of the instrumentation available. Indeed, if the signal is weak, it is very likely that it is susceptible to interference. However, this can also be influenced by other factors such as the growth medium of the process which could create problems in terms of robustness and accuracy of the probes.

2.2 Overview of state estimate techniques

The lack of reliable online information about process state variables, and in particular those more informative such as metabolic products or intracellular metabolites that can provide a

deeper understanding of the process, represents a significant limiting factor for an effective control of bioreactors (Venkateswarlu, 2004). This is a very common issue in the process industry in which, in addition to the more common situation where key variables cannot be measured in real time and at frequent intervals, other situations can occur such as the total absence of available measurements on a regular basis, or again when measurements can be taken only very infrequently or when samples are sent to the laboratory for analysis. In these last two cases, the result is a measurement affected by significant delay, which are therefore not indicative of the present time in which one is interested but refer to some undefined instant in the past. In any case what is difficult/challenging is the implementation of feedback control strategy based on measurements (Ogunnaike & Ray, 1994). As explained in the introduction, an efficient monitoring system of the process represents a key element in ensuring the success of the bioprocess operation. Online estimation of process variables and parameters that are not directly accessible is a valuable and interesting solution/tool to fill this gap and improve monitoring and control of the operating units involved. Methods that can meet this task, thus providing an estimate of the value of unmeasurable process variables, are usually known as state estimators. In Figure 1 the block diagram of the estimation scheme is reported. The main idea behind this operation is to reconstruct reliably and in real time the trends of process variables and parameters that define the process, exploiting the process knowledge given by the model together with a limited set of available process measurements (Dochain, 2003; Venkateswarlu, 2004). In other words, the state estimator can thus be defined as an estimation algorithm whose task is to “observe” the value of unmeasured state variables, and that is why it is also known as state “observer” (Ogunnaike & Ray, 1994). More recently, they have also been referred to as soft sensors, highlighting the fact that these are tools composed of a software part because the sensor signal evaluation models are usually computer programs, and these models are carrying information similar to their hardware counterpart (Kadlec et al., 2009; Luttmann et al., 2012).

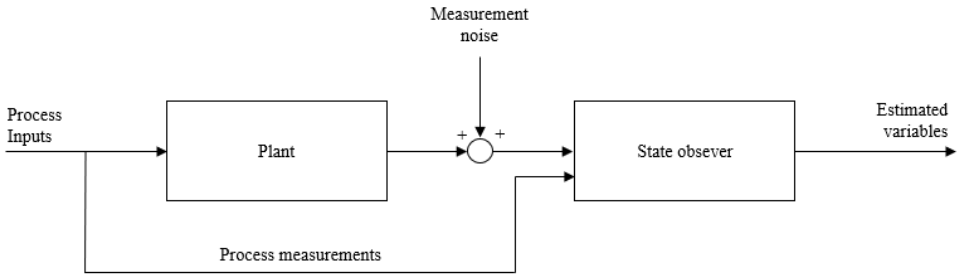

Figure 1. Block diagram of state observer.

State estimation has been a quite active area of research for several decades (F. J. Doyle, 1997). Indeed, suffice it to say that the concept of state observer and Kalman filter-based observer were introduced by Luenberger (Luenberger, 1964, 1966, 1967, 1971) and Kalman (Welch & Bishop, 1995.) in the 1960s. Over the years, several advances have been made in this area but obviously, the implementation of observers has become an increasingly challenging task as a consequence of the higher requirements in terms of accuracy, cost-effectiveness (or convenience), and performance (Mohd Ali, Ha Hoang, et al., 2015). Most of today's observers are merely extended versions and modifications of the classical Luenberger and Kalman approaches/structures. The study and the applications of state estimation strategies continues to be a still-discussed topic. Indeed, several types of observers have recently been developed for applications concerning nonlinear chemical processes as well. In some cases, their application has involved not only the theoretical aspect but also the practical aspect in real plants. Their applications are not limited only to chemical processes, but also to biotechnological processes. Indeed, several examples of applications in this area are available in the literature (Dochain & Rapaport, 2018; García-Mañas et al., 2019; Hernandez et al., 2013; Krämer & King, 2019; Quintero et al., 2008; Semcheddine & Bouchareb, 2019; Simutis et al., 2013) . As clearly reported in Mohd Ali et al. (2015), it is possible to classify observers into six main groups: Luenberger-based observers, finite-dimensional system observers, Bayesan estimators, disturbances and fault detection observers, artificial intelligence (AI)-based observes, and finally hybrid observers.

Table 1. Classification of major classes of observers according to Mohd Ali et al. (2015).

Main Categories	Specific observers	
Luenberger-based observers	1. Extended Luenberger observer (ELO) 2. Sliding mode observer (SMO) 3. Adaptive state observer (ASO) 4. High-gain observer	5. Zeitz nonlinear observer 6. Discrete-time nonlinear recursive observer (DNRO) 7. Geometric observer 8. Backstepping observer
Finite-dimensional system observers	1. Reduced-order observer 2. Low-order observer 3. High gain observer 4. Asymptotic observer (AO)	5. Exponential observer 6. Integral observer 7. Interval observer
Bayesian estimators	1. Particle filter (PF) 2. Extended Kalman filter (EKF) 3. Unscented Kalman filter (UKF) 4. Ensemble Kalman filter (EnKF)	5. Steady state Kalman filter (SSKF) 6. Adaptive fading Kalman filtering (AFKF) 7. Moving horizon estimator (MHE) 8. Generic observer 9. Specific observer
Disturbance and fault detection observers	1. Disturbance observer (DOB) 2. Modified disturbance observer (MDOB) 3. Fractional-order disturbance observer 4. Bode-ideal cut-off observer	5. Unknown input observer (UIO) 6. Nonlinear unknown input observer 7. Extended unknown input observer 8. Modified proportional observer
AI-based observers	1. Fuzzy Kalman filter 2. Augmented fuzzy Kalman filter	3. Differential neural network observer 4. EKF with neural network model
Hybrid observers	1. Extended Luenberger-asymptotic observer 2. Proportional-integral observer 3. Proportional-SMO 4. Continuous-discrete observer	5. Continuous-discrete-interval observer 6. Continuous-discrete-EKF 7. High-gain-continuous-discrete

In this section, the main types of observers will be briefly illustrated following the order shown in Table 1.

2.2.1 Luenberger-based observers

As stated before, the concept of observer was first introduced by Luenberger (Luenberger, 1964, 1966, 1967, 1971) and for this reason these tools are usually known as Luenberger's observers. To this category belong the observers designed based on Luenberger's approach, and thus its extensions. Hence, the extended Luenberger observer (ELO) (Méndez-Acosta et al., 2008), sliding mode observe (SMO) (de Battista et al., 2011; Gonzalez et al., 2001; Hajatipour & Farrokhi, 2010; Picó et al., 2009), adaptive state observer (ASO) (Sheibat-Othman et al., 2008; Zhang & Guay, 2002), and the geometric observer (GO) (López & Alvarez, 2004; Schaum et al., 2019; Tronci et al., 2005), to name only a few, fall into this category.

2.2.2 Finite-dimensional system observers

The category of finite-dimensional systems includes those of reduced order, low order, high gain, asymptotic and exponential. Several examples of their applications to bioprocesses can be found in the literature (Biagiola & Figueroa, 2004; el Assoudi et al., 2002; Kazantzis et al., 2005; Salehi & Shahrokhi, 2008). They are characterized by a quite simple implementation and usually are designed for chemical processes with dynamics described by ordinary differential equations. In addition, they are suitable for nonlinear systems, and it is worth noting that they require high-quality models (Dochain, 2013).

2.2.3 Bayesan estimators

Among the Bayesian estimators, it is worth mentioning the Extended Kalman Filter (EKF), the particle filter (PF), and the moving horizon estimator (MHE). In general, their operation is based on the assumption that all state variables are stochastic, and they allow to obtain the estimation of the probabilistic distribution of these variables by exploiting the available measurements. In this way they represent a consistent and versatile approach that makes possible to achieve an estimation in a short time. In addition, they also represent one of the most widely used methods for determining model parameters, as reported by Ogunnaike & Ray (1994). Several examples of applications concerning this category, and involving biotechnological processes, can be found in the literature (Chitralekha et al., 2010; Raïssi et al., 2005; J. Wang et al., 2010).

2.2.4 Disturbance and fault detection observers

This is a very specific category of observers, which can actually be regarded as the union of two distinct classes incorporated into one since they have the common aim to estimate all the possible irregularities, in terms of disturbances or failures, that may occur in a process. Among these, it is possible to find the disturbance observer (DOB), the modified disturbance observer (MDOB), the unknown input observer (UNIO), and also the nonlinear unknown input observer (NUIO). Therefore, they are useful for detecting disturbances or faults during the variable estimation process or for estimating the disturbances of faults themselves. In both cases they have important roles, but in the last one in particular they play a preventive role by warning/alerting operators to possible problems in process units. Some applications to bioprocesses can be found in literature (Aguilar-López & Martinez-Guerra, 2005; Avilés et al., 2022; Rocha-Cózatl & Wouwer, 2011; Safaeipour et al., 2021).

2.2.5 Artificial intelligence-based observers

As reported by Mohd Ali et al. (2015b), artificial intelligence (AI) can be defined as the ability of computers or other machines to perform activities for which human intelligence is required. It is a widely used method and recently is interest in applications concerning not only to process model and control, but also the estimation of states and parameters that are difficult to measure. So-called artificial intelligence (AI)-based observers or estimators are thus useful computational algorithms that can predict states or parameters that are not accessible (but that are important to develop feedback control law for a system). Algorithms such as artificial neural networks (ANN), fuzzy logic, expert system (ES) and genetic algorithm (GA) fall into this category. Although conventional observers such as the Luenberger observer and the extended Kalman filter provide good estimation performance, AI-based estimators have as advantages the ability to retune more easily in the presence of parameter variations/changes and to avoid time delays. Many papers can be found in the literature concerning the application of these estimators alone (Aziz et al., 2000; Fortuna et al., 2005; Hussain et al., 2002.; Patnaik, 1997) or combined with other estimation approach (Chairez et al., 2007; Porru et al., 2000; Prakash & Senthil, 2008; Senthil et al., 2006) as will be explained in the next section.

2.2.6 Hybrid observers

The last category is that of so-called hybrid observers resulting from the combination of more than one observer. It can be considered as a solution that is usually applied when a single observer has limitations in its operation, so to improve performance in terms of estimation another more suitable observer is combined. Several examples exist in the literature. Hulhoven et al. (2006) developed a hybrid "Luenberger-asymptotic" observer that combines the advantages of the former of converging by having an accurate model, and the latter of providing state estimates having no kinetic model information. Another example is that provided by Sheibat-Othman et al. (2008), where an adaptive observer was coupled/associated with a discrete continuous one to monitor online the evolution of the molecular weight of a polymer over time, identifying model parameters. In particular, the concentration of radicals was estimated through the adaptive observer using the monomer concentration measurement, while the termination rate coefficient using a continuous-discrete observer exploiting the off-line measurements of the polymer molecular weight. Other possible combinations have been developed in the past (Bogaerts, 1999; Hulhoven & Bogaerts, 2002), and all of them have demonstrated that the end result is an improved observer. However, beyond the advantage

overcoming the possible functioning limitations, choosing the most viable and appropriate combination can be a tedious and time-consuming practice (Mohd Ali et al., 2015a).

2.3 A geometric observer-assisted approach to tailor state estimation in a bioreactor for ethanol production

In this section, the design and implementation of a model-based soft sensor for the estimation of non-measurable states in a fermentation process for ethanol production will be shown. The results of different simulations will be reported and analysed.

2.3.1 Process model

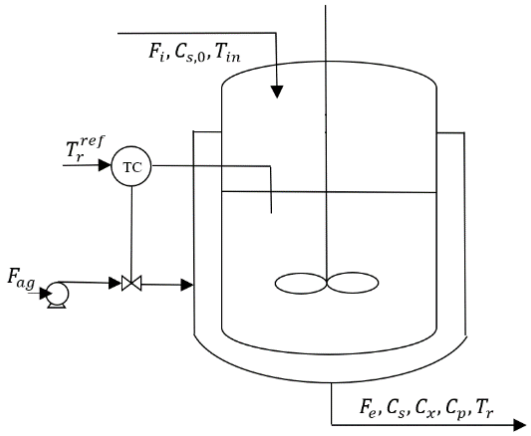

Figure 2. Schematization of a continuous fermentation bioreactor.

The biochemical process considered in this thesis work is a fermentation bioreactor for ethanol production, which is shown in Figure 2. The model has been carefully developed by Nagy (2007) and subsequently extended by other authors (Imtiaz et al., 2013; Ławryńczuk, 2008). The reactor is obtained as a continuous stirred tank (CSTR) with a constant feed rate. The device contains three different components: the biomass (Cx), namely a yeast suspension fed to the system and continuously removed from it; the substrate (Cs), which is the glucose that feeds the microorganisms; the product (i.e., the ethanol C_P) that is removed from the reactor together with other components. Dissolved oxygen is also present in the reactor (C_{O_2}) and it is consumed during the fermentation. A low dilution rate (F_e/V) is necessary in order to have a quasi-stationary state for biomass and the consequence is the quite slow dynamics of the process. Along with microorganisms (*Saccharomyces cerevisiae*), the addition of salts was also considered. These represent the source of inorganic nitrogen and play an important role in yeast

growth, but especially in the formation of coenzymes. Moreover, they have a significant influence on the equilibrium concentration of oxygen in the liquid phase due to the so-called salting-out effect. The balance equations describing this influence, but also those modelling the dynamics of the three main components (C_X, C_S, C_P), reactor temperature (T_r), and coolant temperature (T_{ag}) are given in detail below. The values of the model parameters and nominal operating conditions of the process are given in Table 1 and 2.

Dissolved oxygen model

Mass concentrations of ions in the reaction medium were calculated by the following equations, considering that the chloride ion is contained in the two salts $NaCl$ and $MgCl_2$:

$$C_{Na} = \frac{m_{NaCl}}{M_{NaCl}} \frac{M_{Na}}{V} \tag{1}$$

$$C_{Ca} = \frac{m_{CaCO_3}}{M_{CaCO_3}} \frac{M_{Ca}}{V} \tag{2}$$

$$C_{Mg} = \frac{m_{MgCl_2}}{M_{MgCl_2}} \frac{M_{Mg}}{V} \tag{3}$$

$$C_{Cl} = \left(\frac{m_{NaCl}}{M_{NaCl}} + 2 \frac{m_{MgCl_2}}{M_{MgCl_2}} \right) \frac{M_{Cl}}{V} \tag{4}$$

$$C_{CO_3} = \frac{m_{CaCO_3}}{M_{CaCO_3}} \frac{M_{CO_3}}{V} \tag{5}$$

$$C_H = 10^{-pH} \tag{6}$$

$$C_{OH} = 10^{-(14-pH)} \tag{7}$$

The ionic strength of the generic ion was evaluated by the following equation:

$$I_i = \frac{1}{2} C_i z_i^2 \tag{8}$$

Consequently, it can be applied to all ions present:

$$I_{Na} = 0.5 C_{Na} (1)^2 \tag{9}$$

$$I_{Ca} = 0.5 C_{Ca} (2)^2 \tag{10}$$

$$I_{Mg} = 0.5 C_{Mg} (2)^2 \tag{11}$$

$$I_{Cl} = 0.5 C_{Ca} (-1)^2 \tag{12}$$

$$I_{CO_3} = 0.5 C_{CO_3} (-2)^2 \tag{13}$$

$$I_H = 0.5C_H(1)^2 \quad (14)$$

$$I_{OH} = 0.5C_{OH}(-1)^2 \quad (15)$$

The overall effect of ionic forces was expressed by the following equation:

$$\sum H_i I_i = H_{Na}I_{Na} + H_{Ca}I_{Ca} + H_{Mg}I_{Mg} + H_{Cl}I_{Cl} + .. + H_{CO_3}I_{CO_3} + H_H I_H + H_{OH}I_{OH} \quad (16)$$

Where H_i is the specific ionic constant of the *i*-th ion. The final expression can be derived from the Eq. (16):

$$\begin{aligned}\sum H_i I_i = {} & 0.5H_{Na}\frac{m_{NaCl}}{M_{NaCl}}\frac{M_{Na}}{V} + 2H_{Ca}\frac{m_{CaCO_3}}{M_{CaCO_3}}\frac{M_{Ca}}{V} + 2H_{Mg}\frac{m_{MgCl_2}}{M_{MgCl_2}}\frac{M_{Mg}}{V} \\ & + 0.5H_{Cl}\left(\frac{m_{NaCl}}{M_{NaCl}} + 2\frac{m_{MgCl_2}}{M_{MgCl_2}}\right)\frac{M_{Cl}}{V} + 2H_{CO_3}\frac{m_{CaCO_3}}{M_{CaCO_3}}\frac{M_{CO_3}}{V} + 0.5H_H 10^{-pH} \\ & + 0.5H_{OH}10^{-(14-pH)}\end{aligned} \quad (17)$$

The influence of temperature in distillate water on the equilibrium concentration of oxygen can be described by means of the following empirical equation derived by Sevella (1992):

$$C^*_{O_2,0} = 14.6 - 0.3943T_r + 0.007714{T_r}^2 - 0.0000646{T_r}^3 \quad (18)$$

Since the salts are dissolved in the reaction medium, the equilibrium concentration of oxygen was calculated by the Eq.(19) (Sevella, 1992):

$${C_{O_2}}^* = C^*_{O_2,0} \times 10^{-\sum H_i I_i} \quad (19)$$

Mass transfer coefficient for oxygen can be expressed as a function of temperature (Sevella, 1992):

$$(k_l a) = (k_l a)_0 (1.024)^{T_r - 20} \quad (20)$$

While the rate of oxygen consumption is:

$$r_{O_2} = \mu_{O_2}\frac{1}{Y_{O_2}}C_X\frac{C_{O_2}}{K_{O_2} + C_{O_2}} \quad (21)$$

Finally, the dissolved oxygen concentration in the reaction medium can be considered as the result of the amount entering the reaction medium due to mass transfer, expressed by the first term of Eq. (22), and the amount consumed for the fermentation process (last term) as showed below:

$$\frac{dC_{O_2}}{dt} = k_l a \left(C^*_{O_2} - C_{O_2}\right) - r_{O_2} \tag{22}$$

Where $C^*_{O_2}$ and C_{O_2} are the equilibrium concentration of dissolved oxygen and the oxygen concentration in the liquid phase, respectively. r_{O_2} is the consumption rate of oxygen and $k_l a$ is the product of mass transfer coefficient for oxygen and the gas-phase specific area.

Fermentation process model

Denoting F_i and F_e as the flow rate of substrate entering the reactor and the outlet flow rate respectively, the variation rate of the total reaction volume V was expressed as follows:

$$\frac{dV}{dt} = F_i - F_e \tag{23}$$

Since F_i and F_e are equal and therefore the volume V remains constant, the mass balance for the biomass and the product can be written as follows:

$$\frac{dC_X}{dt} = \mu_X \, C_X \, \frac{C_S}{K_{S1} + C_S} e^{-K_{P1} C_P} - \frac{F_e}{V} C_X \tag{24}$$

$$\frac{dC_P}{dt} = \mu_P C_X \frac{C_S}{K_{S1} + C_S} e^{-K_{P1} C_p} - \frac{F_e}{V} \, C_P \tag{25}$$

The first terms in Eqs. (24) and (25) represent the amount of biomass and ethanol produced during the fermentation reaction, while the second terms describe the amount of yeast and ethanol leaving the reactor, respectively. It is possible to observe that the maximum specific growth rate is present in the mass balance for the biomass, and it is expressed as a temperature function:

$$\mu_X = A_1 e^{-(E_{a1}/R(T_r+273))} - A_2 e^{-(E_{a2}/R(T_r+273))} \tag{26}$$

Whereas μ_P is the maximum specific fermentation rate, C_X is the biomass (yeast) concentration, C_S is the substrate (glucose) concentration, C_P is the product (ethanol) concentration, $K_{S1}$is the constant in the substrate term for ethanol production, and K_{P1} is the constant of fermentation inhibition by ethanol. The mass balance of substrate (glucose) is expressed by the following equation:

$$\frac{dC_S}{dt} = -\frac{1}{R_{SX}}\mu_X C_X \frac{C_S}{K_S + C_S} e^{-K_P C_P} - \frac{1}{R_{SP}}\mu_P C_X \frac{C_S}{K_{S1} + C_S} e^{-K_{P1} C_P} + \frac{F_i}{V} C_{S,in} - \frac{F_e}{V} C_S \quad (27)$$

Where R_{SX} is the ratio of cell produced per glucose consumed for growth, with R_{SP} is indicated the ratio of ethanol produced per glucose consumed for fermentation, K_S is the constant in the substrate term for growth, and $C_{S,in}$ is the substrate (glucose) concentration in the feed flow. The first term of the Eq. (27) represents the substrate amount consumed by biomass for the growth, whereas the second term describes the amount of glucose consumed by biomass for the ethanol production. Instead, the third and the last one represents the glucose amount entering the bioreactor through the fresh substrate feed and the amount of substrate leaving the reactor, respectively.

Temperature model

The energy balance for the bioreactor allows to define the variation rate of the temperature of the reactor by means of the following equation:

$$\frac{dT_r}{dt} = \left(\frac{F_i}{V}\right)(T_{in} + 273) - \left(\frac{F_e}{V}\right)(T_r + 273) + \frac{r_{O_2}\,\Delta H_r}{32\,\rho_r\,C_{heat,r}} - \frac{K_T A_T\,(T_r - T_{ag})}{V\,\rho_r\,C_{heat,r}} \quad (28)$$

The energy balance was also written on the cooling jacket (Eq. (29)):

$$\frac{dT_{ag}}{dt} = \left(\frac{F_{ag}}{V_j}\right)(T_{in,ag} - T_{ag}) + \frac{K_T A_T (T_r - T_{ag})}{V_j \rho_{ag} C_{heat,ag}} \quad (29)$$

Where F_{ag} is the flow of the cooling agent, V_j is the jacket volume, $T_{in,ag}$ is the temperature of the cooling agent entering to the jacket, ρ_{ag} and $C_{heat,ag}$ are the density and the specific heat of the coolant, respectively. The values of the model parameters of the examined bioreactor are

given in Table 2, whereas the initial nominal operating conditions of the process are reported in Table 3 (Imtiaz et al., 2013; Ławryńczuk, 2008; Nagy, 2007).

The model presented here was used to simulate a real process and to develop the model-based soft-sensor (estimator). Since the purpose of this work is to mimic a real situation, the simulation using the model parameters taken from Nagy (2007), and reported in Table 2, was considered as the real plant (hereafter referred to as the virtual plant). On the other hand, in order to simulate what usually occurs in a real situation and make the estimation problem more demanding and more representative of an industrial plant, two sources of error were included. The first is an additive white noise which behaves like a uniformly distributed random number and that corrupts the available measurements. The precision of the respective sensors is reported in Table 4. The second source is responsible for a model mismatch, which implies that the model used in the estimator algorithm is different from the model used to simulate the "virtual plant". In particular, it is assumed that the kinetic parameters used in the estimator (Table 5) are different from the ones taken from Nagy (2007) and reported in Table 2. All other parameters were left unchanged, and the same nominal operating conditions were used as reported in Table 3.

Table 2. Parameters of the bioreactor model.

$A_1 = 9.5 \times 10^8$	$(k_l a)_0 = 38\ h^{-1}$	$M_{Mg} = 24\ g/mol$
$A_2 = 2.55 \times 10^{33}$	$K_{O_2} = 8.86\ mg/l$	$M_{MgCl_2} = 95\ g/mol$
$A_T = 1\ m^2$	$K_p = 0.139\ g/l$	$M_{Na} = 23\ g/mol$
$C_{heat,ag} = 4.18\ J\ g^{-1}K^{-1}$	$K_{p_1} = 0.07\ g/l$	$M_{NaCl} = 58.5\ g/mol$
$C_{heat,r} = 4.18\ J\ g^{-1}K^{-1}$	$K_s = 1.03\ g/l$	$R = 8.31\ J\ mol^{-1}K^{-1}$
$E_{a1} = 55\ J/mol$	$K_{s_1} = 1.68\ g/l$	$R_{SP} = 0.435$
$E_{a2} = 220\ J/mol$	$K_T = 3.6 \times 10^5\ Jh^{-1}m^{-2}K^{-1}$	$R_{SX} = 0.607$
$H_{Ca} = -0.303$	$m_{CaCO_3} = 100\ g$	$V_j = 50\ l$
$H_{Cl} = 0.844$	$m_{MgCl_2} = 100\ g$	$Y_{O_2} = 0.97\ mg/mg$
$H_{CO_3} = 0.485$	$m_{NaCl} = 500\ g$	$\Delta H_r = 518\ kJ/mol\ O_2$
$H_H = -0.774$	$M_{Ca} = 40\ g/mol$	$\mu_{O_2} = 0.5\ h^{-1}$
$H_{Mg} = -0.314$	$M_{CaCO_3} = 90\ g/mol$	$\mu_p = 1.79\ h^{-1}$
$H_{Na} = -0.550$	$M_{Cl} = 35.5\ g/mol$	$\rho_{ag} = 1000\ g/l$
$H_{OH} = 0.941$	$M_{CO_3} = 60\ g/mol$	$\rho_r = 1080\ g/l$

Table 3. Nominal operating conditions of the process.

$C_{O_2} = 2.5$ mg/l	pH = 6
$C_p = 13$ g/l	$T_{ag} = 29$ °C
$C_s = 27$ g/l	$T_{in} = 25$°C
$C_{s,in} = 60$ g/l	$T_{in,ag} = 15$°C
$C_x = 1$ g/l	$T_r = 26$°C
$F_{ag} = 18$ l/h	V = 1000 l
$F_i = F_e = 51$ l/h	

Table 4. Noise for the different measuring sensors with respect to the corresponding nominal values.

C_X	C_S	C_{O_2}	T_r	T_{ag}
±2.5%	±2.5%	±2.5%	±0.1°C	±0.1°C

Table 5. Modified model parameters for estimator sensors.

$\mu_P\ [h^{-1}]$	1.7465
$K_S\ [g/l]$	1.0248
$K_P\ [g/l]$	0.1281
$K_{S1}\ [g/l]$	1.8090
$K_{P1}\ [g/l]$	0.0692
$R_{SX}\ [-]$	0.6274
$R_{SP}\ [-]$	0.4549

For the purpose of this work, it is important to specify that all simulation tests were performed by considering step variations on three inputs of the model. More specifically, in the results concerning the structure of the full-order estimator, the temperature of feed entering the bioreactor (T_{in}) was changed after 100 h of simulation, the concentration of substrate inlet ($C_{S,in}$) after 150 h, and finally the coolant inlet temperature ($T_{in,ag}$) after 250 h (Table 6).

Table 6. Step changes of input variables.

$T_{in}\ [°C]$	25	30
$C_{S,in}\ [g/l]$	60	75
$T_{in,ag}\ [°C]$	15	10

On the other hand, as for the analysis of the reduced-order estimator, step changes on the same three inputs were considered but individually, as described in Table 7.

Table 7. Step changes of input variables.

	Input	t = 0 h	t = 100 h	t = 200 h
T1	$C_{S,in}$ [g/l]	60	45	75
T2	$T_{in,ag}$ [°C]	15	10	20
T3	T_{in} [°C]	25	20	30

2.3.2 State estimation problem

As reported in Cabaneros Lopez et al. (2019), the current real-time monitoring methods used in ethanol production generally consist of secondary measurements such as pH, turbidity, gas composition and temperature. Even if such variables provide important information about the process, they do not directly relate to the state of the system, making it difficult to apply advanced control strategies. Furthermore, even the best process measurements are corrupted by some amount of signal noise and their true values are somewhat uncertain. States estimation technique can be used to improve the output signal of measured process states in the presence of uncertainty and when it is not possible to directly measure all the variables of interest.

The estimation problem consists in jointly designing the estimation structure (i.e., estimator model, sensors, innovated states and data assimilation mechanism), and the estimation algorithm (i.e., the dynamic data processor), to infer some or all the states of the bioreactor on the basis of the available model in conjunction with available measurements, according to a specific estimation objective. In the present fermentation reactor estimation study, the emphasis has been placed on: (i) the detection of the more adequate measured outputs leading to the best performance, (ii) the selection of the innovated states, meaning the states which are updated by using the available measurement. For simplifying the formulation of the problem, the model described in the previous paragraph can be written in compact form as reported in Eqs. (30):

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}), \qquad \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \tag{30a}$$

$$\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) \tag{30b}$$

where $\boldsymbol{x}$ is the n-dimensional state vector, equal to $\boldsymbol{x_0}$ at the initial time t_0, $\boldsymbol{u}$ is the p-dimensional input vector, $\boldsymbol{f}$ is the n-dimensional vector fields, $\boldsymbol{y}$ is the m–dimensional vector of the measured

outputs and $\boldsymbol{h}$ is the map relating states and measurements. The dimension of the measured outputs is less than the number of states, that is $m < n$. As reported in Salas et al. (2019), it is possible to consider a non-linear map ϕ, the components of which are the measured outputs and some of their directional derivatives.

$$\Phi(x,u) = [\Phi_1, \dots, \dots, \Phi_i, \dots, \dots, \Phi_m]^T \tag{31a}$$

$$\Phi_i = \left(h_i(x), L_f^1 h_i(x), \dots, \dots, L_f^{\kappa_i - 1} h_i(x)\right) \tag{31b}$$

where $L_f^j h_i(x)$ are the recursive j^{th} Lie derivatives of the time varying scalar field $h_i(x)$ along the vector field time variant $\boldsymbol{f}\left(x, u(t)\right)$, κ_i is the observability index of the i^{th} output and κ is the estimator order defined by the following Equation:

$$\kappa_1 + \kappa_2 + \cdots + \kappa_m = \kappa = n \tag{32}$$

If this condition is verified, that is the sum of the m observability indices κ_i is equal to the dimension of the state vector, and the map $\boldsymbol{\Phi}(\boldsymbol{x}, \boldsymbol{u})$ is invertible with respect to x, it is possible to relate the measured outputs to the states and to reconstruct the system dynamics. This can be assessed by evaluating the rank of the matrix $\partial_x \boldsymbol{\Phi}(\boldsymbol{x}, \boldsymbol{u})$ for given trajectories, meaning that the system is observable if:

$$rank\left(\partial_x \Phi(x,u)\right) = n \tag{33}$$

In this way the states can be reconstructed using the available model and a proper measurement process algorithm (Dewasme et al., 2013). It is important underline that the robust observability matrix can be detected by evaluating the condition number σ of the observability matrix and its minimum singular value κ. Such metrics are important tools for choosing the best estimator structure (López & Alvarez, 2004; Salas et al., 2019).

Robust estimability and robust detectability

If all states can be fully observable, the observability matrix should be full-rank, but practical observability can be assessed if the condition number of the observability matrix ($\boldsymbol{\Sigma}$) is small (Dochain et al., 1997). Furthermore, a small singular value of the observability matrix implies the worst estimate of the states (Long et al., 2008):

$$rank(\boldsymbol{\Upsilon}) = n, \boldsymbol{\Upsilon} = \partial_x \boldsymbol{\phi}(x,u) \tag{34a}$$

$$\frac{\overline{\sigma}(\boldsymbol{\Upsilon})}{\underline{\sigma}(\boldsymbol{\Upsilon})} = \Sigma < \Xi \tag{34b}$$

$$\underset{t}{avg}\,\underline{\sigma}(\boldsymbol{\Upsilon}) \geq \varepsilon_0 \tag{34c}$$

where Ξ and ε_0 are, respectively, the selected thresholds. On the other hand, if matrix $\boldsymbol{\Upsilon}$ is rank deficient and the unobservable states are stable, it is necessary to distinguish between states that can be innovated (distinguishable states) and states that cannot (undistinguishable states). In this case, the dimension of the map in Equation (31) is equal to the dimension of the distinguishable states, and robust detectability can be assessed if the following conditions are satisfied (Equation (35)):

$$\frac{\overline{\sigma}(\boldsymbol{\Upsilon}_p)}{\underline{\sigma}(\boldsymbol{\Upsilon}_p)} = \Sigma_i < \Xi_p \tag{35a}$$

$$\underset{t}{avg}\,\underline{\sigma}(\boldsymbol{\Upsilon}_p) \geq \varepsilon_{\mathrm{p0}} \tag{35b}$$

$$\boldsymbol{\Upsilon}_p = \partial_x \boldsymbol{\phi}_{\mathrm{p}}(x,u) \tag{35c}$$

$$\boldsymbol{\phi}_p = (h_i, L_f h_i, \dots, L_f^{\kappa_i - 1} h_i\,) \tag{35d}$$

$$\kappa_1 + \kappa_2 + \cdots + \kappa_m = \kappa = p \tag{35e}$$

The constants Ξ_p and $\varepsilon_{\mathrm{p0}}$ are, again, the selected thresholds.

Selection of the Estimator Structure

The performance of an estimator is obviously strongly affected by the model of the process and the quality of the available measurements. Biological processes are complex systems, therefore the presence of model uncertainty in terms of parameters and neglected dynamics are in general to be expected. This means that the complete reconstruction of the states requires, in general, a combination of different measurements (Cabaneros Lopez et al., 2019a). Within this framework, it is important to underline that there is still a gap between the sensors for laboratory use and large-scale monitoring in real-time (Cabaneros Lopez et al., 2019a). The selection of the estimator structure is therefore focused on the choice of the best monitoring strategies, by considering which are the most representative measured outputs and the presence of parameter errors in the model used in the estimator. It is considered that system monitoring can be expensive, in terms of both fixed and operation costs, therefore it could be useful to optimize performance with the least number of sensors. This analysis has been carried out comparing condition number and minimum singular values of the observability matrix. The

performances have been also evaluated by simulating different trajectories, from which the convergence rate, presence of off-set and signal noise have been evaluated.

Geometric observer algorithm

In this work, the geometric observer in the form developed by López and Alvarez (2004) was selected and applied as the estimation algorithm, which is formally connected with the observability properties reported in the previous section. Then, the observer can be constructed in the following way:

$$\frac{d\hat{x}}{dt} = \hat{f}\left(\hat{x}, u(t)\right) + \Phi_x^{-1} K \left(y - h(\hat{x})\right) \tag{36a}$$

$$y = h(\hat{x}, r) \tag{36b}$$

where $\Phi_x^{-1}$is the inverse of the Jacobian Φ_x of the map calculated with respect to system states, K is the observer gain matrix, the components of which are tuning parameters, and their values are calculated using the procedure suggested in Alvarez and Fernández (2009). The block diagonal matrix K is calculated as reported below:

$$\boldsymbol{K} = \begin{pmatrix} \boldsymbol{B_1} & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B_2} & \dots & \boldsymbol{0} \\ \vdots & \vdots & \dots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{B_m} \end{pmatrix}, \quad \boldsymbol{B_1} = \begin{bmatrix} k_{11} \\ \vdots \\ k_{1\nu_1} \end{bmatrix}, \quad \boldsymbol{B_2} = \begin{bmatrix} k_{21} \\ \vdots \\ k_{2\nu_2} \end{bmatrix}, \boldsymbol{B_m} = \begin{bmatrix} k_{m1} \\ \vdots \\ k_{m\nu_m} \end{bmatrix} \tag{37}$$

$$\nu_i = \kappa_{i-1}$$

Where the matrices that compose it are column matrices and whose components are set in such a way that the error dynamics become stable. Tuning guidelines for the geometric observer provided by Álvarez & Fernández (2009) proved that a set of tuning parameters is necessary for each measurement. For observability indexes equal to 1 or 2 ($\kappa_i = 1{,}2$), the proportional gains can be obtained by considering Eq. (38).

$$k_{i1} = 2\zeta\omega_0,\ k_{i2} = {\omega_0}^2 \tag{38a}$$

$$\omega_0 \in [10\omega_c, 30\omega_c],\ \zeta = [1, 3] \tag{38b}$$

Where ω_0 is the characteristic frequency and ζ is the attenuation factor, and they can assume values in the ranges shown in Eq. (38b).

If the observability condition given in Eq. (33) is not verified, and thus the rank of the observability matrix will not be maximum (n), it is not possible to construct a full order observer. In this case, it is possible to evaluate a not-complete observability structure in which there will be unobservable states, which can be consequently deduced from the estimation model in open-loop mode. States that are not corrected by the estimator, usually indicated as $\boldsymbol{x_u}$, are distinguished from observable or innovated states ($\boldsymbol{x_i}$). Hence, the state vector of the system will assume the following form:

$$x = [x_i, x_u]^T \tag{39}$$

Therefore, the geometric observer algorithm will be structured as follows:

$$\dot{\hat{\boldsymbol{x}}}_{\boldsymbol{i}} = \hat{\boldsymbol{f}}_{\boldsymbol{i}}(\hat{\boldsymbol{x}}, \boldsymbol{u}) + \left(\boldsymbol{\partial}_{\boldsymbol{x_i}}\boldsymbol{\phi}(\hat{\boldsymbol{x}}, \boldsymbol{u})\right)^{-\boldsymbol{1}} \boldsymbol{K}\big(\boldsymbol{y} - \boldsymbol{h}(\hat{\boldsymbol{x}})\big), \;\; \boldsymbol{x_{i0}} = \boldsymbol{x_i}(\boldsymbol{t_0}) \tag{40a}$$

$$\dot{\hat{\boldsymbol{x}}}_{\boldsymbol{u}} = \hat{\boldsymbol{f}}_{\boldsymbol{u}}(\hat{\boldsymbol{x}}, \boldsymbol{u}), \;\; \boldsymbol{x_{u0}} = \boldsymbol{x_u}(\boldsymbol{t_0}) \tag{40b}$$

Where it is assumed that some states are not innovated, and so they are only predicted by the model, while for the innovated states the dynamic predicted by the model are adjusted by means of the available measurements. Regarding the construction of the observer gain matrix K, the tuning guidelines reported in Eqs. (37) and (38) are followed.

Kalman filter algorithm

The geometric observer (GO) has been then compared with the Extended Kalman Filter (EKF), which is the most used estimator algorithm in industry, because of its straightforward construction (Álvarez & Fernández, 2009). Even if the EKF is usually applied to complete observability systems, in this investigation it has been used also when the choice of measurements leads to a rank deficient observability matrix. The EKF algorithm has been applied in the continuous form, reported in the following equations (41).

$$\dot{\hat{\boldsymbol{x}}}_{\boldsymbol{i}} = \hat{\boldsymbol{f}}_{\boldsymbol{i}}(\hat{\boldsymbol{x}}, \boldsymbol{u}) + \boldsymbol{K_{EKF}}\big(\boldsymbol{y} - \boldsymbol{h}(\hat{\boldsymbol{x}})\big), \boldsymbol{x_{i0}} = \boldsymbol{x_i}(\boldsymbol{t_0}) \tag{41a}$$

$$\dot{\hat{\boldsymbol{x}}}_{\boldsymbol{u}} = \hat{\boldsymbol{f}}_{\boldsymbol{u}}(\hat{\boldsymbol{x}}, \boldsymbol{u}), \boldsymbol{x_{u0}} = \boldsymbol{x_u}(\boldsymbol{t_0}) \tag{41b}$$

$$\boldsymbol{K_{EKF}} = \boldsymbol{P}(\boldsymbol{t})\boldsymbol{H^T}\boldsymbol{R^{-1}} \tag{41c}$$

$$\dot{\boldsymbol{P}}(\boldsymbol{t}) = \boldsymbol{P}(\boldsymbol{t})\boldsymbol{F}(\boldsymbol{t}) + \boldsymbol{F}^{\boldsymbol{T}}(\boldsymbol{t})\boldsymbol{P}(\boldsymbol{t}) + \boldsymbol{Q}(\boldsymbol{t}) - \boldsymbol{K}_{\boldsymbol{EKF}}\boldsymbol{H}\boldsymbol{P}\,, \;\; \boldsymbol{P}(\boldsymbol{t}_0) = \boldsymbol{P}_0 \tag{41d}$$

$\boldsymbol{F}(t)$ is the Jacobian of the vector field $\hat{\boldsymbol{f}}_{\boldsymbol{i}}$, calculated with respect to the innovated states, $\boldsymbol{P}$ is the error covariance matrix of the innovated states, $\boldsymbol{H}$ is the matrix of the derivative of the map $\boldsymbol{h}$ with respect to the states, $\boldsymbol{Q}$ and $\boldsymbol{R}$ are, respectively, the covariance matrix of the model and measurements errors (Jazwinski, 2007). The constant matrix $\boldsymbol{Q}$, $\boldsymbol{R}$, and $\boldsymbol{P}_0$ are tuning parameters of the estimation model and they have been calculated minimizing the error between the states calculated with the simulated plant and the estimator along a reference trajectory.

2.3.3 Results

Full order estimator

The choice of the estimation structure has been carried out considering: (i) condition number and the minimum singular value of the Jacobian matrix Φ_X for a different choice of measurements and innovated states and (ii) evaluating the responses of the reconstructed states for a given trajectory. As first case, the hypothetical ideal situation in which five measurements ($C_X, C_S, C_{O_2}, T_r, T_{ag}$) are available while the product (C_P) is not measured was analyzed. Temperature and dissolved oxygen measurements have been always considered available, according to the laboratory and industrial practice (Randek & Mandenius, 2018). On the other hand, sensors suited for ethanol measurements as well as substrate and biomass are not always available for large scale real-time applications (Cabaneros Lopez et al., 2019). The choice of considering substrate and biomass as measured variables has been accomplished by considering the literature on sensors for biomanufacturing (Holzberg et al., 2018). In more detail, robustness, stability, and costs have been considered. This first scenario considering five measured outputs is certainly the less demanding because almost all variables can be monitored online. In this case, observability property is satisfied with two configurations:

$$\Phi_1 = \left[C_x, L_f C_x, C_s, C_{O_2}, T_r, T_{ag}\right] \tag{42a}$$

$$\Phi_2 = \left[C_x, C_s, L_f C_s, C_{O_2}, T_r, T_{ag}\right] \tag{42b}$$

The best structure between (42a) and (42b) can be selected by considering the minimum singular value and condition number of the Jacobian matrix for the two maps. The mean values of the selected indexes calculated along a reference trajectory are reported in Table 8, and they

indicate that the second configuration should be the better choice in terms of robustness (lower condition number) and the relationship between measured outputs and states (highest minimum singular value).

Table 8. Minimum singular values and condition number with five measurements.

	Φ_1	Φ_2
κ	133.29	28.9358
σ	0.0076	0.069

The choice based on Table 8 has been confirmed by the dynamic simulation, where it is evident that the second configuration allows a better reconstruction of the product composition, which is the only unmeasured state (Figure 3 and 4). Indeed, the nonlinear estimator (red dash-dotted line) is able to reduce the mismatch between the model without correction (magenta dashed line), indicated as open-loop model, and virtual plant (blue continuous line). The response is highly corrupted by noise, because of the amplification of the measurement error due to the high gain values used to reduce the offset in the ethanol composition estimation.

The same procedure can be used to select the best configuration when only four measured outputs are available. According to the analysis reported in Holzberg et al. (2018), two possible scenarios have been considered: (i) biomass concentration in the reactor is measured on-line or (ii) substrate concentration in the reactor is measured on-line. Using the representation in Cabaneros Lopez et al. (2019), the considered cases are reported in Eqs. (43):

$$\boldsymbol{y} = (C_x, C_{O_2}, T_r, T_{ag}) \tag{43a}$$

$$\boldsymbol{y} = (C_s, C_{O_2}, T_r, T_{ag}) \tag{43b}$$

where $\boldsymbol{y}$ represents the measured output vector.

According to Eq. (34), it is easy to demonstrate that no combination of indexes κ_i satisfies the observability property for the output vector in Eq. (43a). This implies that a full order observer is possible if the substrate concentration is measured online, therefore when using the output configuration reported in Eq. (43b). In this case, the nonlinear estimation maps satisfying Eq. (34a) are reported in Eqs. (50).

$$\Phi_3 = \left[C_S, L_f C_S, C_{O_2}, L_f C_{O_2}, T_r, T_{ag}\right] \tag{44a}$$

$$\Phi_4 = \left[C_S, L_f C_S, C_{O_2}, T_r, L_f T_r, T_{ag}\right] \tag{44b}$$

The means of condition number and minimum singular value for the Jacobian of the maps (44a) and (44b) along the reference trajectory are reported in Table 9. The structure Φ_4 seems to be more robust (lower condition number), but it shows a lower minimum singular value, indicating that changes in the states should affect the outputs to a lesser extent.

Table 9. Condition number and minimum singular value with four measurements.

	Φ_3	Φ_4
κ	1814.8	54.29
σ	0.0842	0.0332

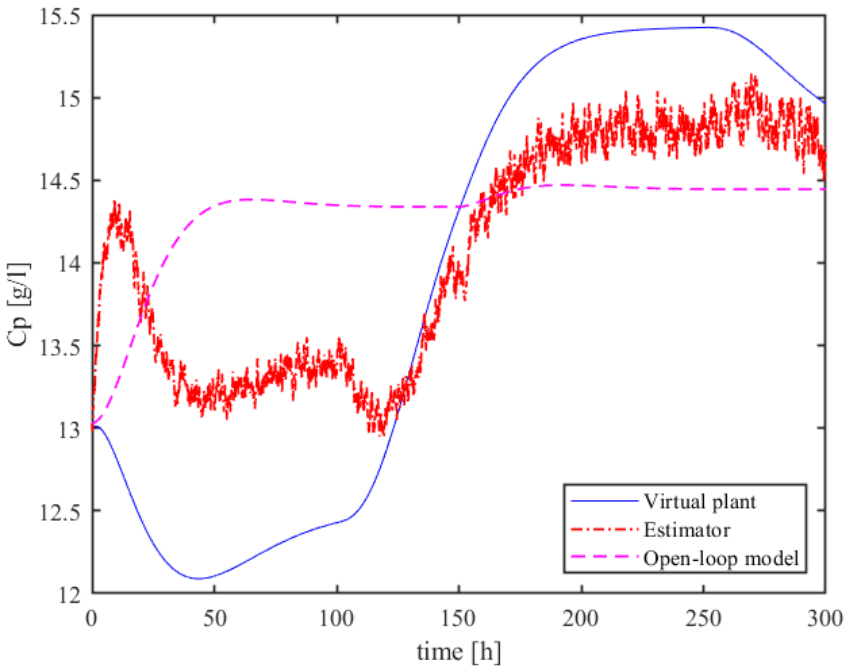

Figure 3. Comparison of the ethanol dynamic behavior between virtual plant (blue continuous line), open loop model (magenta dashed line) and estimator (red dash-dotted line) for structure with map Φ_1.

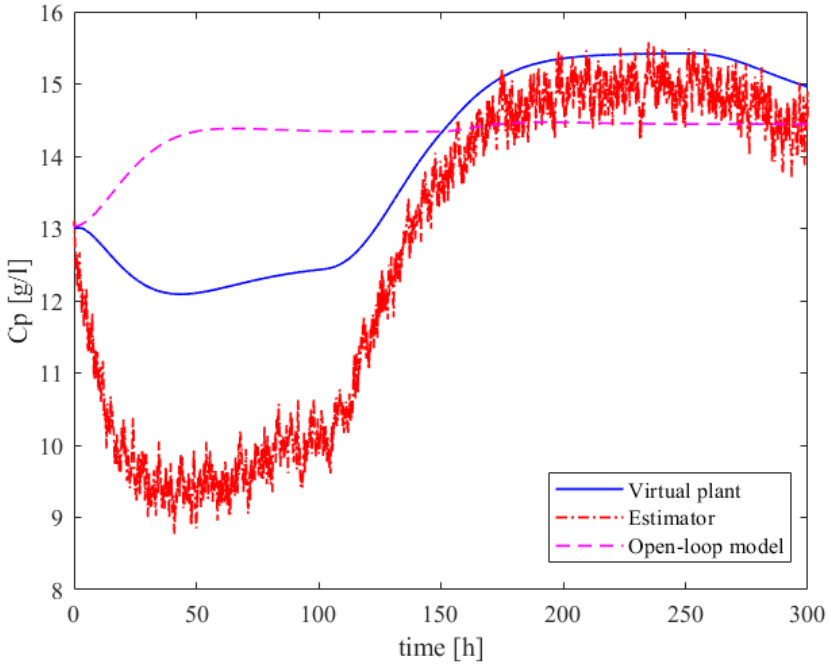

Figure 4. Comparison of the ethanol dynamic behavior between virtual plant (blue continuous line), open loop model (red dashed line) and estimator (magenta dash-dotted line) for structure with map Φ_2.

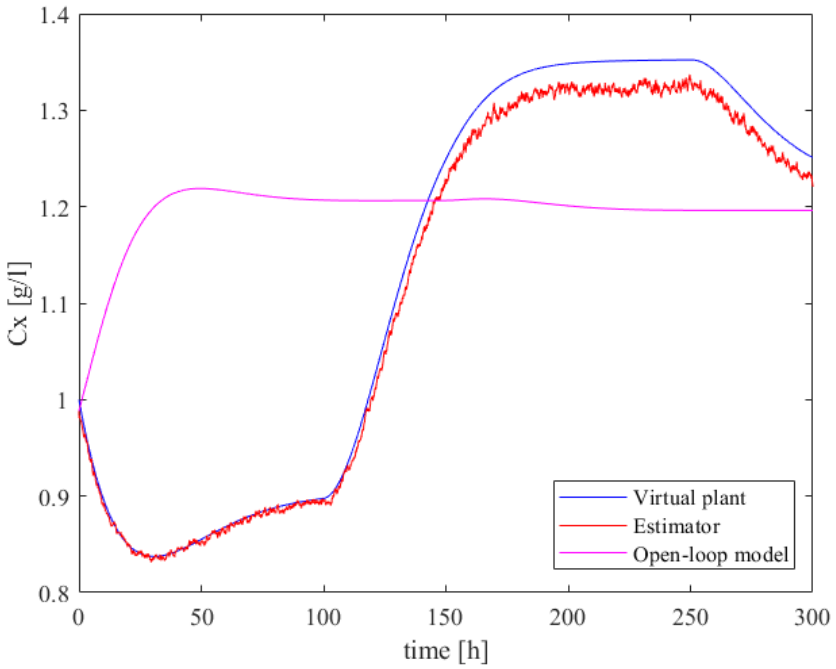

Figure 5. Comparison of the biomass dynamic behavior between virtual plant (blue continuous line), open loop model (magenta dashed line) and estimator (red dash-dotted line) for structure with map Φ_3.

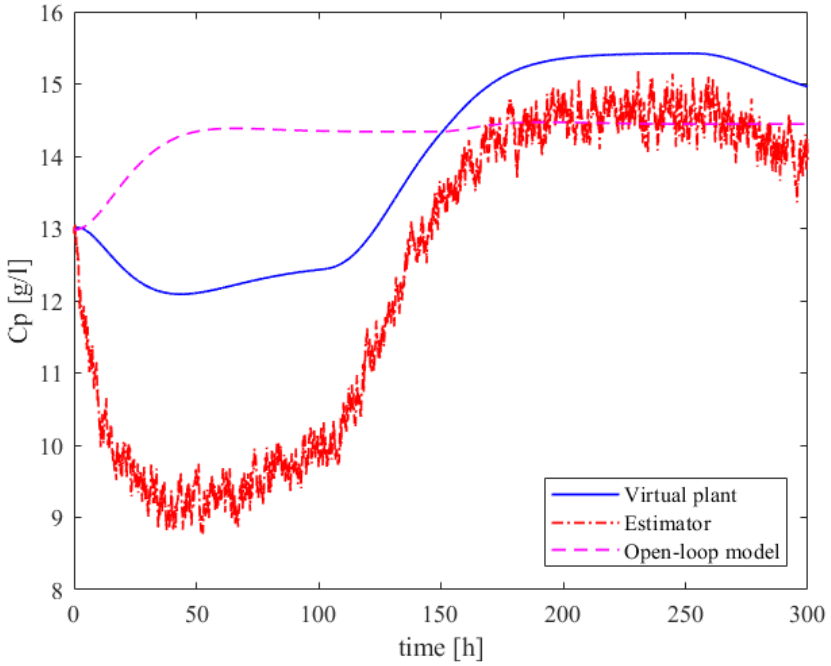

Figure 6. Comparison of the ethanol dynamic behavior between virtual plant (blue continuous line), open loop model (magenta dashed line) and estimator (red dash-dotted line) for structure with map Φ_3.

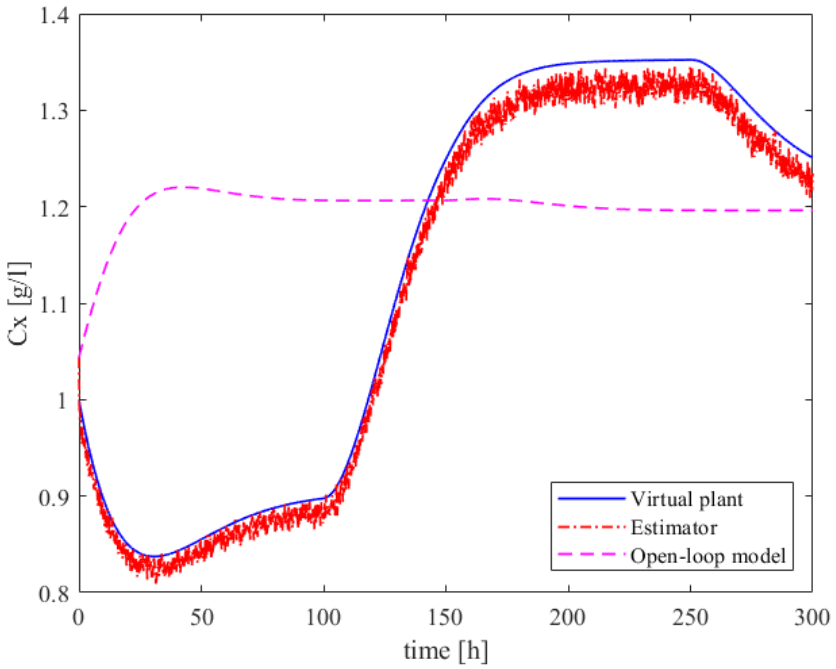

Figure 7. Comparison of the biomass dynamic behavior between virtual plant (blue continuous line), open loop model (magenta dashed line) and estimator (red dash-dotted line) for structure with map Φ_4.

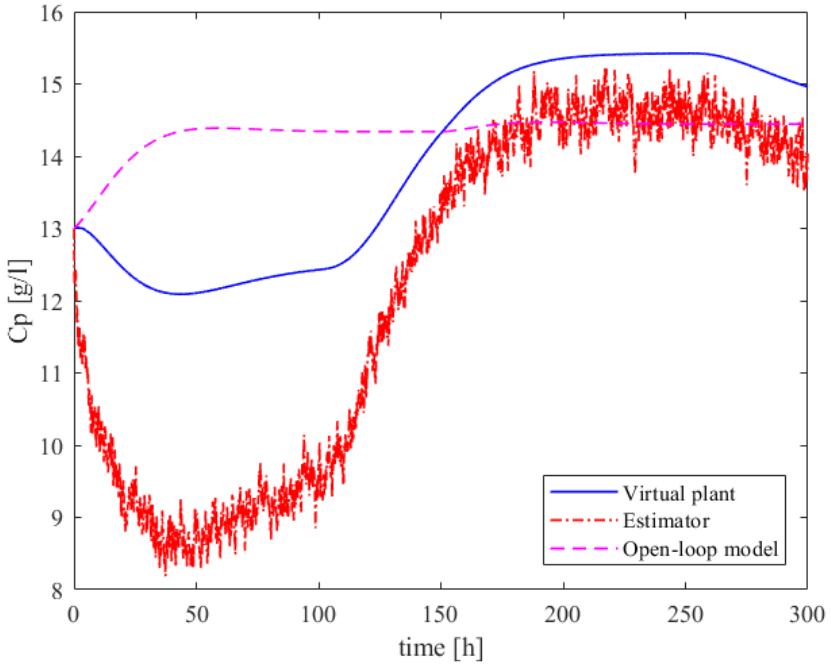

Figure 8. Comparison of the ethanol dynamic behavior between virtual plant (blue continuous line), open loop model (magenta dashed line) and estimator (red dash-dotted line) for structure with map Φ_4.

The reconstructed dynamic behaviour for the two unmeasured states (C_X and C_P) is reported in Figures 5-8. It is worth noticing that also the state values calculated only with the model used in the estimation algorithm (open-loop model), but without innovation are reported in order to better highlight the correction provided by the estimation algorithm.

It is possible to observe that using the map Φ_3, allows a good reconstruction of the biomass behavior (Figure 5), while there is a large mismatch between the ethanol concentration obtained with the virtual plant and the reconstructed one (Figure 6).

When using the second configuration results worsen, both for biomass (Figure 7) that for ethanol (Figure 8) concentration. It is worth noticing that the state's values estimated with map Φ_4 are more corrupted by the measurement noise because in this case a greater observer gain has been used to decrease the offset.

Low-Order Estimator

The two full order structures Φ_3 and Φ_4 are not able to adequately estimate the product of the reactor, therefore a different solution is required to improve ethanol concentration. Using the same measured outputs, it is possible to improve estimation performance by reducing the order of the observer using only one Lie's derivative (Salas et al., 2019). The maps reported in Eqs. (45a-b) lead to five observable states and one only detectable.

$$\Phi_{p5} = \left[C_s, C_{O_2}, T_r, L_f\, T_r, T_{ag} \right] \quad (45a)$$

$$\Phi_{p6} = \left[C_s, L_f C_s, C_{O\,2}, T_r, T_{ag} \right] \quad (45b)$$

The rank of the Jacobian of the maps Φ_i (i=3,4) depends on the choice of the not-innovated state ($\hat{x}_u$) between the two that are not measured, which are ethanol and biomass concentration. It can be verified that the map Φ_{p5} can be inverted only if C_x is innovated and C_P is not. On the other hand, the Jacobian of map Φ_{p6} has always rank equal to five, regardless of the choice of the innovated states. Recalling Eq. (40), the following partitions are considered:

$$\boldsymbol{x}_i = [C_x, C_s, C_{O_2}, T_r, T_{ag}], x_u = [C_p] \quad (46a)$$

$$\boldsymbol{x}_i = \left[C_p, C_s, C_{O_2}, T_r, T_{ag}\right], x_u = [C_x] \quad (46b)$$

The map Φ_{p5} can be used with the partition (46a), while the map Φ_{p6} can be used with both partitions (46a-b). Therefore, two different solutions are identified: $\Phi_{6,1}$ for partition (46b) and $\Phi_{6,2}$ for partition (46a). A first analysis of the possible configurations can be obtained by considering the minimum singular values and condition number reported in Table 10. The indexes' values are comparable; therefore, the evaluation of the best structure has been performed analysing the reconstruction performance. Figures (9-12) represent the estimation of the unmeasured states (ethanol and biomass concentration) for the input step change T1 and T2 described in Table 7. The best reconstruction capabilities are shown by configuration Φ_{p5} for both the states. This result may suggest that conditions calculated with Eqs. (35) are informative when the magnitude between the different configurations is significantly different, otherwise, it is necessary to evaluate the estimations capabilities by evaluating the estimator response for given input changes.

Table 10. Mean condition number and minimum singular value for low order structures.

	Φ_{p5}(C_P open-loop)	$\Phi_{p6,1}$ (C_X open-loop)	$\Phi_{p6,2}$ (C_P open-loop)
Σ	2.15	8.75	1.53
$\underline{\sigma}$	0.47	0.12	0.99

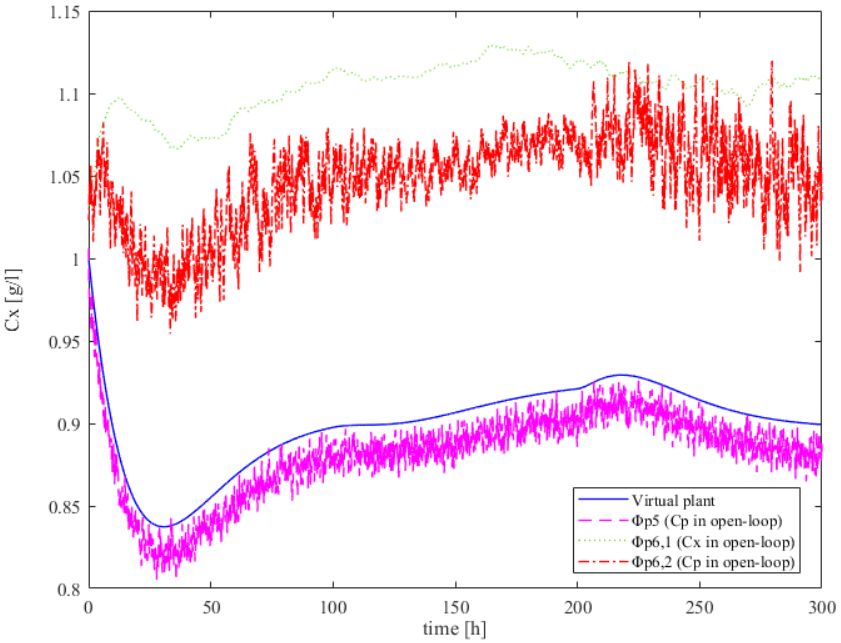

Figure 9. Dynamic response of biomass concentration calculated with the virtual plant (blue continuous line), GO with map Φ_{p5} (magenta dashed line), GO with map $\Phi_{p6,1}$ (green dotted line), GO with map $\Phi_{p6,2}$ (red dashed-dotted line) along the trajectory T1.

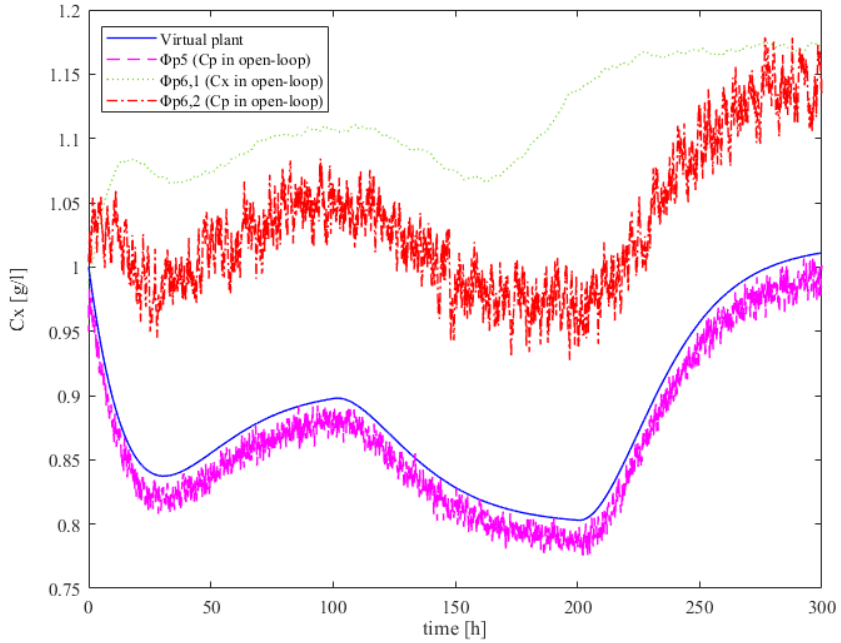

Figure 10. Dynamic response of biomass concentration calculated with the virtual plant (blue continuous line), GO with map Φ_{p5} (magenta dashed line), GO with map $\Phi_{p6,1}$ (green dotted line), GO with map $\Phi_{p6,2}$ (red dashed-dotted line) along the trajectory T2.

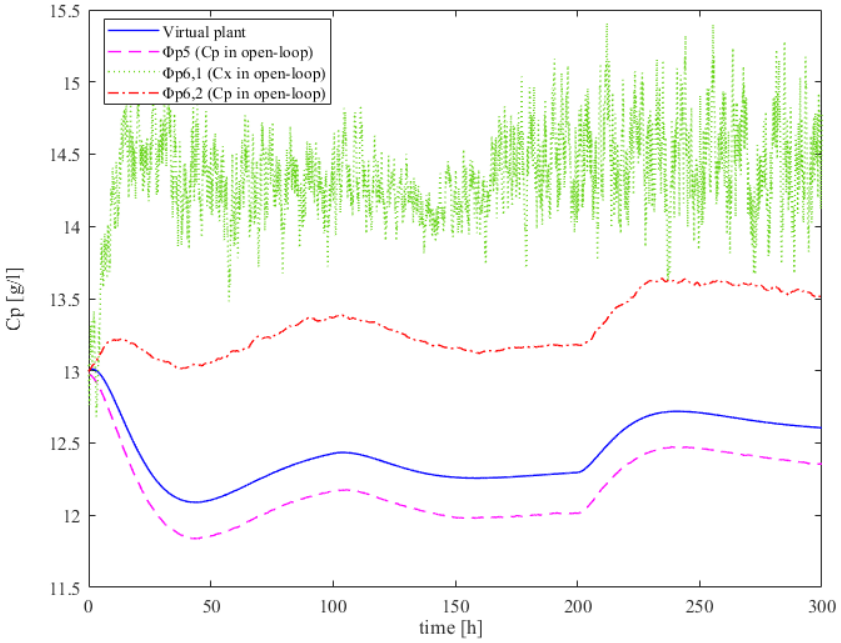

Figure 11. Dynamic response of ethanol concentration calculated with the virtual plant (blue continuous line), GO with map Φ_{p5} (magenta dashed line), GO with map $\Phi_{p6,1}$ (green dotted line), GO with map $\Phi_{p6,2}$ (red dashed-dotted line) along the trajectory T1.

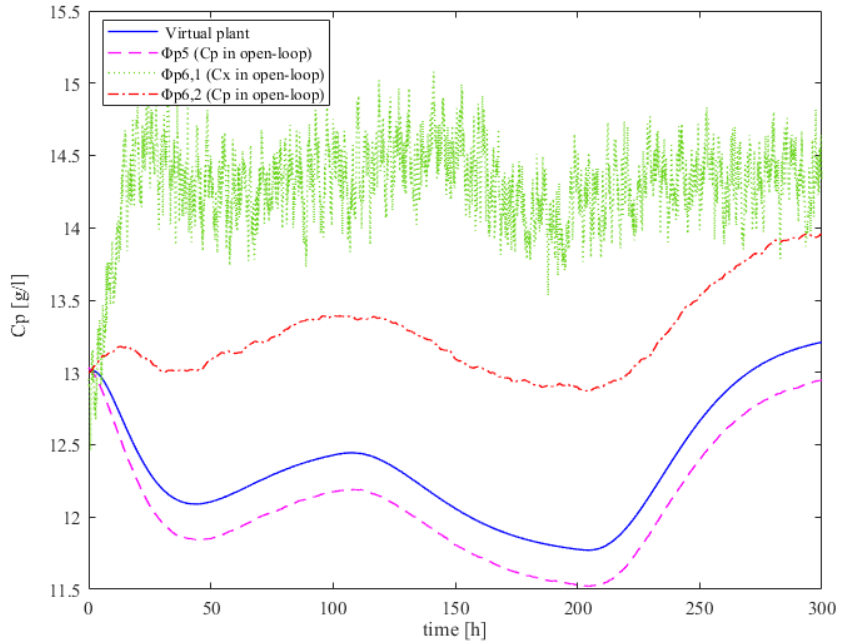

Figure 12. Dynamic response of ethanol concentration calculated with the virtual plant (blue continuous line), GO with map Φ_{p5} (magenta dashed line), GO with map $\Phi_{p6,1}$ (green dotted line), GO with map $\Phi_{p6,2}$ (red dashed-dotted line) along the trajectory T2.

Validation

The analysis carried out in the previous section leads to find the best estimation structure with four measured outputs. In order to validate the obtained results, a new test was carried out considering as reference trajectory the variation of the input temperature (T_{in}) as shown in Table 7 (Case T3). Figures 13 and 14 show the dynamic behavior of biomass and product concentration respectively and confirm that the proposed structure can effectively reconstruct the unmeasured states also with different process conditions. It is worth noticing that the ethanol concentration is not innovated, but the correction of the other states has a positive impact also on its estimation.

Using the same number and choice of measured outputs (43b) and partition between innovated and not innovated states (46a), the estimation task has been addressed using the extended Kalman filter (Figures 15-16). The main reason for using another algorithm as a measurement processor is to demonstrate that the estimator performance depends on the structure selection rather than estimation algorithm. EKF has been preferred for this validation because it is usually preferred in the industrial practice being it easy to implement and robust if adequately calibrated (Leu & Baratti, 2000; Salas et al., 2019). Results show that EKF can effectively reconstruct the unmeasured states, revealing that estimator structure design is the key step for a successful achievement of the estimation goals. The only difference between the two approaches is that the biomass calculated with the geometric observer is more affected by noise. This behavior can be explained by the presence of the Lie derivative in GO, which implies a higher sensitivity to measurement noise with respect to the EKF.

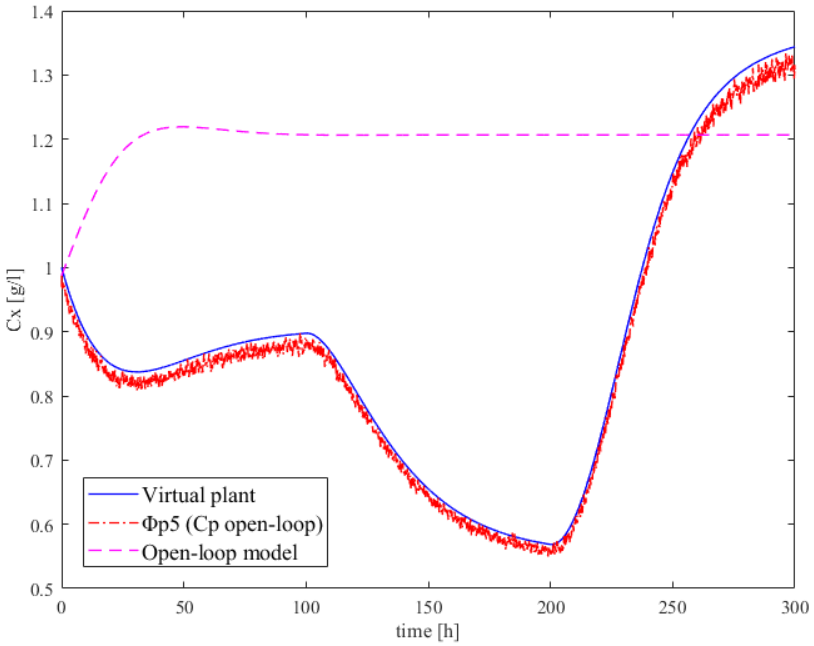

Figure 13. Dynamic response of biomass concentration calculated with the virtual plant (blue continuous line), open-loop model (magenta dashed line) and GO (red dashed-dotted line) for structure Φ_{p5} along trajectory T3.

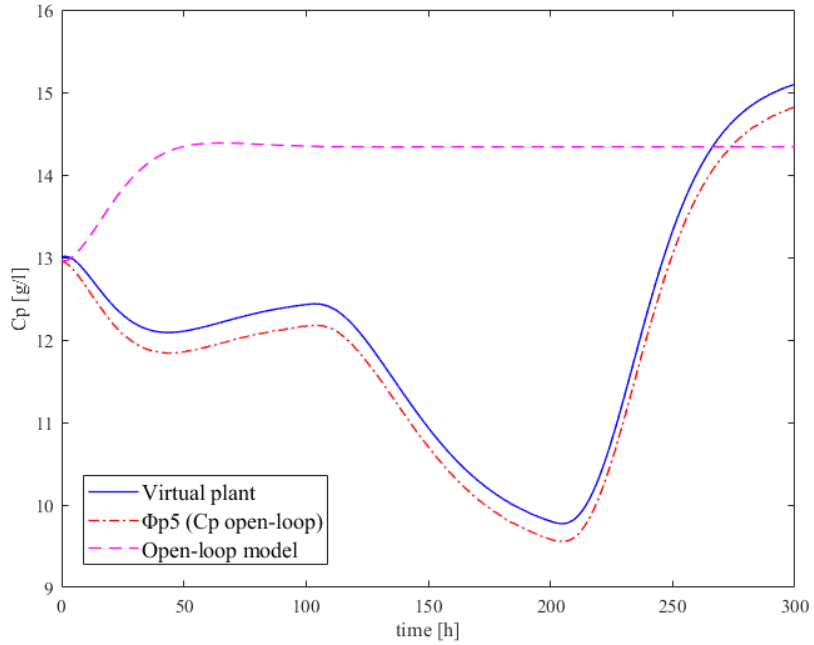

Figure 14. Dynamic response of ethanol concentration calculated with the virtual plant (blue continuous line), open-loop model (magenta dashed line) and GO (red dashed-dotted line) for structure Φ_{p5} along trajectory T3.

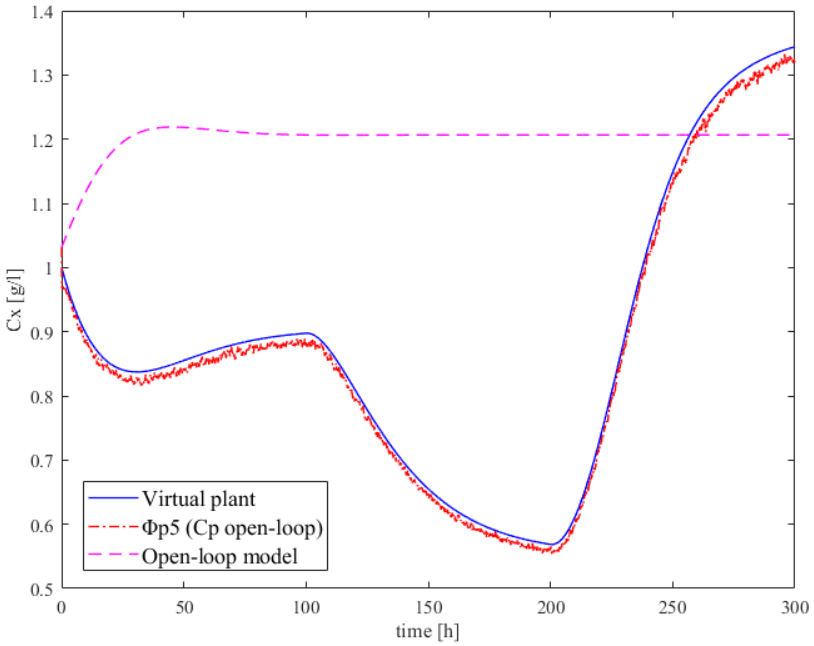

Figure 15. Dynamic response of biomass concentration calculated with the virtual plant (blue continuous line), open-loop model (magenta dashed line) and extended Kalman Filter (EKF) (red dashed-dotted line) for structure Φ_{p5} along trajectory T3.

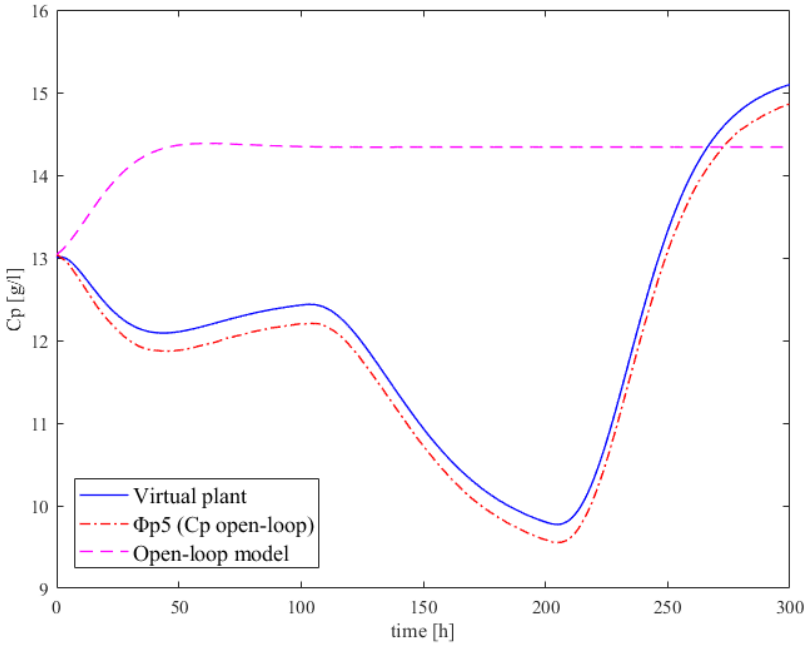

Figure 16. Dynamic response of ethanol concentration calculated with the virtual plant (blue continuous line), open-loop model (magenta dashed line) and extended Kalman Filter (EKF) (red dashed-dotted line) for structure Φ_{p5} along trajectory T3.

2.4 Conclusions

In this chapter, the problem of monitoring biotechnological processes was addressed by providing a quick overview of the state variables generally monitored, and of the corresponding measurements techniques, but also the estimation methods that can be applied to compensate the lack of some of the measurements. Later, a focus was placed on a fermentation process for the production of ethanol, and in particular on the development of a soft sensor for estimating the unmeasurable states of the bioprocess. It was demonstrated that the estimation performance relies on an appropriate structure selection rather than the chosen measurement processor algorithm. An adjustable-structure geometric estimation approach was used, and the estimator structure constituted a design degree of freedom to improve its performance versus robustness behavior. The estimation structure design was based on estimability and detectability properties used together with a geometric approach. The analysis of the estimability measures showed the ill and well-conditioned structures (condition number of the observability matrix), and the poorest estimation performance for the given structure (minimum singular value of the observability matrix). From the implementation stage with simulations, it was found that the results agreed with the ones of the structural assessment when estimability measure values calculated for the different structures were significantly different. The used estimation algorithm was the geometric observer with proportional innovation, which offers simplicity of tuning and implementation. With the aim of showing that the proposed procedure for choosing the estimation structure can be applied to other estimation techniques, the extended Kalman filter was also used as measurement processor algorithm. The obtained results showed that the two estimators lead to good estimation performance, with the only difference that the geometric observer estimation is more sensitive to measurement noise, probably because of the presence of the Lie derivative in the correction term. Summarizing, the systematic geometric approach led to the best solution for the estimation problem, giving a structure that did not depend on the correction algorithm. The latter can be chosen according to the wishes of the personnel of the plant or developer experience. It is worth noticing that the systematic tuning procedure of the geometric approach was very useful for comparing the reconstruction capabilities of the different structures. The results here presented in terms of methodology could be applied to more complex biotechnological processes, as the obtainment of ethanol from cellulosic material, where the measurement devices for real-time application in the industry are still missing. In this case, the proposed approach can be used to detect which are the measurements that lead to the best reconstruction capabilities and invest in them.

Bibliography

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

Aguilar-López, R., & Martinez-Guerra, R. (2005). State estimation for nonlinear systems under model unobservable uncertainties: application to continuous reactor. *Chemical Engineering Journal*, *108*(1–2), 139–144. https://doi.org/10.1016/J.CEJ.2005.01.008

Álvarez, J., & Fernández, C. (2009). Geometric estimation of nonlinear process systems. *Journal of Process Control*, *19*(2), 247–260. https://doi.org/10.1016/J.JPROCONT.2008.04.017

Alvarez, J., & López, T. (1999). Robust dynamic state estimation of nonlinear plants. *AIChE Journal*, *45*(1), 107–123. https://doi.org/10.1002/aic.690450110

Avilés, J. D., Torres-Zúñiga, I., Villa-Leyva, A., Vargas, A., & Buitrón, G. (2022). Experimental validation of an interval observer-based sensor fault detection strategy applied to a biohydrogen production dark fermenter. *Journal of Process Control*, *114*, 131–142. https://doi.org/10.1016/J.JPROCONT.2022.04.012

Bäcker, M., Rakowski, D., Poghossian, A., Biselli, M., Wagner, P., & Schöning, M. J. (2013). Chip-based amperometric enzyme sensor system for monitoring of bioprocesses by flow-injection analysis. *Journal of Biotechnology*, *163*(4), 371–376. https://doi.org/10.1016/J.JBIOTEC.2012.03.014

Biagiola, S. I., & Figueroa, J. L. (2004). A high gain nonlinear observer: application to the control of an unstable nonlinear process. *Computers & Chemical Engineering*, *28*(9), 1881–1898. https://doi.org/10.1016/J.COMPCHEMENG.2004.03.004

Bogaerts, P. (n.d.). *A hybrid asymptotic-Kalman observer for bioprocesses*.

Cabaneros Lopez, P., Feldman, H., Mauricio-Iglesias, M., Junicke, H., Huusom, J. K., & Gernaey, K. v. (2019a). Benchmarking real-time monitoring strategies for ethanol production from lignocellulosic biomass. *Biomass and Bioenergy*, *127*, 105296. https://doi.org/10.1016/J.BIOMBIOE.2019.105296

Cabaneros Lopez, P., Feldman, H., Mauricio-Iglesias, M., Junicke, H., Huusom, J. K., & Gernaey, K. v. (2019b). Benchmarking real-time monitoring strategies for ethanol production from lignocellulosic biomass. *Biomass and Bioenergy*, *127*, 105296. https://doi.org/10.1016/J.BIOMBIOE.2019.105296

Caramihai, M., & Severi, I. (2013). Bioprocess Modeling and Control. In *Biomass Now - Sustainable Growth and Use*. InTech. https://doi.org/10.5772/55362

Chitralekha, S. B., Prakash, J., Raghavan, H., Gopaluni, R. B., & Shah, S. L. (2010). A comparison of simultaneous state and parameter estimation schemes for a continuous fermentor reactor. *Journal of Process Control*, *20*(8), 934–943. https://doi.org/10.1016/J.JPROCONT.2010.06.008

Ciccarella, G., Dalla Mora, M., & Germani, A. (1993). A Luenberger-like observer for nonlinear systems. *International Journal of Control*, *57*(3), 537-556.

de Battista, H., Picó, J., Garelli, F., & Vignoni, A. (2011). Specific growth rate estimation in (fed-)batch bioreactors using second-order sliding observers. *Journal of Process Control*, *21*(7), 1049–1055. https://doi.org/10.1016/J.JPROCONT.2011.05.008

Derfus, G. E., Abramzon, D., Tung, M., Chang, D., Kiss, R., & Amanullah, A. (2010). Cell culture monitoring via an auto-sampler and an integrated multi-functional off-line analyzer. *Biotechnology Progress*, *26*(1), 284–292. https://doi.org/10.1002/btpr.303

Dewasme, L., Goffaux, G., Hantson, A. L., & Wouwer, A. vande. (2013). Experimental validation of an Extended Kalman Filter estimating acetate concentration in E. coli cultures. *Journal of Process Control*, *23*(2), 148–157. https://doi.org/10.1016/J.JPROCONT.2012.09.004

Dochain, D. (2003). State and parameter estimation in chemical and biochemical processes: a tutorial. *Journal of Process Control*, *13*(8), 801–818. https://doi.org/10.1016/S0959-1524(03)00026-X

Dochain, D., & Rapaport, A. (2018). *An asymptotic observer for batch processes with single biogas measurement*. *51*(13), 420–424. https://doi.org/10.1016/j.ifacol.2018.07.315

Dochain, D., Tali-Maamar, N., & Babary, J. P. (1997). On modelling, monitoring and control of fixed bed bioreactors. *Computers & Chemical Engineering*, *21*(11), 1255–1266. https://doi.org/10.1016/S0098-1354(96)00370-5

Doyle, A., & Griffiths, J. B. (1998). *Cell and Tissue Culture: Laboratory Procedures in Biotechnology* (Wiley, Ed.).

Doyle, F. J. (1997). Nonlinear Inferential Control for Process Applications. *IFAC Proceedings Volumes*, *30*(9), 167–178. https://doi.org/10.1016/S1474-6670(17)43156-9

Edwiges, T., Roda-Serrat, M. C., Errico, M., Segovia-Hernández, J. G., Sánchez-Ramírez, E., & Tronci, S. (2022). Sustainable bioalcohol production: Pretreatment, separation, and control strategies leading to sustainable processes. *Biofuels and Biorefining: Volume 2: Intensification Processes and Biorefineries*, 41–85. https://doi.org/10.1016/B978-0-12-824117-2.00004-1

el Assoudi, A., el Yaagoubi, E. H., & Hammouri, H. (2002). Non-linear observer based on the Euler discretization. *International Journal of Control*, *75*(11), 784–791. https://doi.org/10.1080/00207170210141860

Esmonde-White, K. A., Cuellar, M., Uerpmann, C., Lenain, B., & Lewis, I. R. (2017). Raman spectroscopy as a process analytical technology for pharmaceutical manufacturing and bioprocessing. In *Analytical and Bioanalytical Chemistry* (Vol. 409, Issue 3, pp. 637–649). Springer Verlag. https://doi.org/10.1007/s00216-016-9824-1

García-Mañas, F., Guzmán, J. L., Berenguel, M., & Acién, F. G. (2019). Biomass estimation of an industrial raceway photobioreactor using an extended Kalman filter and a dynamic model for microalgae production. *Algal Research*, *37*, 103–114. https://doi.org/10.1016/j.algal.2018.11.009

Gargalo, C. L., Lopez, P. C., Hasanzadeh, A., Udugama, I. A., & Gernaey, K. v. (2022). On-line monitoring of process parameters during fermentation. *Current Developments in Biotechnology and Bioengineering*, 117–164. https://doi.org/10.1016/B978-0-323-91167-2.00003-4

Gonzalez, J., Fernandez, G., Aguilar, R., Barron, M., & Alvarez-Ramirez, J. (2001). Sliding mode observer-based control for a class of bioreactors. *Chemical Engineering Journal*, *83*(1), 25–32. https://doi.org/10.1016/S1385-8947(00)00177-7

Hajatipour, M., & Farrokhi, M. (2010). Chattering free with noise reduction in sliding-mode observers using frequency domain analysis. *Journal of Process Control*, *20*(8), 912–921. https://doi.org/10.1016/J.JPROCONT.2010.06.015

Hernandez, S. C., Bueno, J. A., Sanchez, E. N., & Diaz-Jimenez, L. (2013). State Estimation by artificial neural networks in a continuous bioreactor. *IFAC Proceedings Volumes (IFAC-PapersOnline)*, *12*(PART 1), 215–220. https://doi.org/10.3182/20131216-3-IN-2044.00033

Holzberg, T. R., Watson, V., Brown, S., Andar, A., Ge, X., Kostov, Y., Tolosa, L., & Rao, G. (2018). Sensors for biomanufacturing process development: facilitating the shift from batch to continuous manufacturing. In *Current Opinion in Chemical Engineering* (Vol. 22, pp. 115–127). Elsevier Ltd. https://doi.org/10.1016/j.coche.2018.09.008

Hulhoven, X., & Bogaerts, P. (2002). HYBRID FULL HORIZON-ASYMPTOTIC OBSERVER FOR BIOPROCESSES. *IFAC Proceedings Volumes*, *35*(1), 419–424. https://doi.org/10.3182/20020721-6-ES-1901.01379

Hulhoven, X., vande Wouver, A., & Bogaerts, Ph. (2006). Hybrid extended Luenberger-asymptotic observer for bioprocess state estimation. *Chemical Engineering Science*, 7151–7160.

Imtiaz, U., Assadzadeh, A., Jamuar, S. S., & Sahu, J. N. (2013). Bioreactor temperature profile controller using inverse neural network (INN) for production of ethanol. *Journal of Process Control*, *23*(5), 731–742. https://doi.org/10.1016/j.jprocont.2013.03.005

Jazwinski, A. H. (2007). *Stochastic processes and filtering theory* (Dover Publications, Ed.). Courier Corporation.

Kadlec, P., Gabrys, B., & Strandt, S. (2009). Data-driven Soft Sensors in the process industry. *Computers & Chemical Engineering*, *33*(4), 795–814. https://doi.org/10.1016/J.COMPCHEMENG.2008.12.012

Kazantzis, N., Huynh, N., & Wright, R. A. (2005). Nonlinear observer design for the slow states of a singularly perturbed system. *Computers & Chemical Engineering*, *29*(4), 797–806. https://doi.org/10.1016/J.COMPCHEMENG.2004.09.009

Krämer, D., & King, R. (2019). A hybrid approach for bioprocess state estimation using NIR spectroscopy and a sigma-point Kalman filter. *Journal of Process Control*, *82*, 91–104. https://doi.org/10.1016/j.jprocont.2017.11.008

Ławryńczuk, M. (2008). Modelling and nonlinear predictive control of a yeast fermentation biochemical reactor using neural networks. *Chemical Engineering Journal*, *145*(2), 290–307. https://doi.org/10.1016/j.cej.2008.08.005

Leu, G., & Baratti, R. (2000). An extended Kalman filtering approach with a criterion to set its tuning parameters; application to a catalytic reactor. *Computers & Chemical Engineering*, *23*(11–12), 1839–1849. https://doi.org/10.1016/S0098-1354(00)00298-2

Lindner, P., Bulnes Abundis, D., Lücking, T. H., & Solle, D. (2014). *In situmicroscopy and MIR-spectroscopy as non-invasive optical sensors for cell cultivation process monitoring Polysialic acid View project Target Omics View project*. *2*(2), 157–166. https://doi.org/10.2217/PBP.14.13

Liu, Y.-C., Wang, F.-S., & Lee, W.-C. (2001). On-line monitoring and controlling system for fermentation processes. In *Biochemical Engineering Journal* (Vol. 7).

Longhi, L. G. S., Marcon, S. M., Trierweiler, J. O., & Secchi, A. R. (2002). State estimation of an experimental bioreactor using the extended kalman filtering technology. *IFAC Proceedings Volumes*, *35*(1), 379–382. https://doi.org/10.3182/20020721-6-ES-1901.01372

Lopez, P. C., Udugama, I. A., Thomsen, S. T., Roslander, C., Junicke, H., Mauricio-Iglesias, M., & Gernaey, K. v. (2020). Towards a digital twin: a hybrid data-driven and mechanistic digital shadow to forecast the evolution of lignocellulosic fermentation. *Biofuels, Bioproducts and Biorefining*, *14*(5), 1046–1060. https://doi.org/10.1002/bbb.2108

López, T., & Alvarez, J. (2004). On the effect of the estimation structure in the functioning of a nonlinear copolymer reactor estimator. *Journal of Process Control*, *14*(1), 99–109. https://doi.org/10.1016/S0959-1524(03)00030-1

Lourenço, N. D., Lopes, J. A., Almeida, C. F., Sarraguça, M. C., & Pinheiro, H. M. (2012). Bioreactor monitoring with spectroscopy and chemometrics: A review. In *Analytical and Bioanalytical Chemistry* (Vol. 404, Issue 4, pp. 1211–1237). https://doi.org/10.1007/s00216-012-6073-9

Luenberger, D. G. (1964). Observing the State of a Linear System. *IEEE Transactions on Military Electronics*, *8*(2), 74–80. https://doi.org/10.1109/TME.1964.4323124

Luenberger, D. G. (1966). Observers for Multivariable Systems. *IEEE Transactions on Automatic Control*, *11*(2), 190–197. https://doi.org/10.1109/TAC.1966.1098323

Luenberger, D. G. (1967). Canonical Forms for Linear Multivariable Systems. *IEEE Transactions on Automatic Control*, *AC-12*(3), 290–293. https://doi.org/10.1109/TAC.1967.1098584

Luenberger, D. G. (1971). An Introduction to Observers. In *IEEE Transactions on Automatic Control* (Vol. 16, Issue 6, pp. 596–602). https://doi.org/10.1109/TAC.1971.1099826

Luttmann, R., Bracewell, D. G., Cornelissen, G., Gernaey, K. v., Glassey, J., Hass, V. C., Kaiser, C., Preusse, C., Striedner, G., & Mandenius, C. F. (2012). Soft sensors in bioprocessing: A status report and recommendations. In *Biotechnology Journal* (Vol. 7, Issue 8, pp. 1040–1048). https://doi.org/10.1002/biot.201100506

Lyubenova, V., Kostov, G., & Denkova-Kostova, R. (2021). Model-based monitoring of biotechnological processes-a review. *Processes*, *9*(6). https://doi.org/10.3390/pr9060908

Mauricio-Iglesias, M., Gernaey, K. v., & Huusom, J. K. (2015). State Estimation in Fermentation of Lignocellulosic Ethanol. Focus on the Use of pH Measurements. *Computer Aided Chemical Engineering*, *37*, 1769–1774. https://doi.org/10.1016/B978-0-444-63577-8.50140-6

Méndez-Acosta, H. O., Palacios-Ruiz, B., Alcaraz-González, V., Steyer, J. P., González-Álvarez, V., & Latrille, E. (2008). Robust control of volatile fatty acids in anaerobic digestion processes. *Industrial and Engineering Chemistry Research*, *47*(20), 7715–7720. https://doi.org/10.1021/ie800256e

Mitra, S., & Murthy, G. S. (2022). Bioreactor control systems in the biopharmaceutical industry: a critical perspective. In *Systems Microbiology and Biomanufacturing* (Vol. 2, Issue 1, pp. 91–112). Springer. https://doi.org/10.1007/s43393-021-00048-6

Mohd Ali, J., Ha Hoang, N., Hussain, M. A., & Dochain, D. (2015). Review and classification of recent observers applied in chemical process systems. *Computers & Chemical Engineering*, *76*, 27–41. https://doi.org/10.1016/J.COMPCHEMENG.2015.01.019

Mohd Ali, J., Hussain, M. A., Tade, M. O., & Zhang, J. (2015). Artificial Intelligence techniques applied as estimator in chemical process systems - A literature survey. In *Expert Systems with Applications* (Vol. 42, Issue 14, pp. 5915–5931). Elsevier Ltd. https://doi.org/10.1016/j.eswa.2015.03.023

Mross, S., Fürst, P., Pierrat, S., Zimmermann, T., Vogt, H., & Kraft, M. (2015). Enzyme Sensor With Polydimethylsiloxane Membrane and CMOS Potentiostat for Wide-Range Glucose Measurements. *IEEE Sensors Journal*, *15*(12), 7096–7104. https://doi.org/10.1109/JSEN.2015.2470111

Mross, S., Zimmermann, T., Winkin, N., Kraft, M., & Vogt, H. (2015). Integrated Multi-sensor System for Parallel In-situ Monitoring of Cell Nutrients, Metabolites and Cell Mass in Biotechnological Processes. *Procedia Engineering*, *120*, 372–375. https://doi.org/10.1016/J.PROENG.2015.08.642

Nagy, Z. K. (2007). Model based control of a yeast fermentation bioreactor using optimally designed artificial neural networks. *Chemical Engineering Journal*, *127*(1–3), 95–109. https://doi.org/10.1016/j.cej.2006.10.015

Ödman, P., Johansen, C. L., Olsson, L., Gernaey, K. v., & Lantz, A. E. (2009). On-line estimation of biomass, glucose and ethanol in Saccharomyces cerevisiae cultivations using in-situ multi-wavelength fluorescence and software sensors. *Journal of Biotechnology*, *144*(2), 102–112. https://doi.org/10.1016/J.JBIOTEC.2009.08.018

Ogunnaike, B. A., & Ray, W. H. (1994). *Process Dynamics, Modeling, and Control* (1st ed.). Oxford University Press.

Ohadi, K., Legge, R. L., & Budman, H. M. (2015). Development of a Soft-Sensor Based on Multi-Wavelength Fluorescence Spectroscopy and a Dynamic Metabolic Model for Monitoring Mammalian Cell Cultures. *Biotechnol. Bioeng*, *112*, 197–208. https://doi.org/10.1002/bit.25339/abstract

Ozturk, S. S., & Hu, W.-S. (2005). *Cell culture technology for pharmaceutical and cell-based therapies*. Taylor & Francis.

Picó, J., de Battista, H., & Garelli, F. (2009). Smooth sliding-mode observers for specific growth rate and substrate from biomass measurement. *Journal of Process Control*, *19*(8), 1314–1323. https://doi.org/10.1016/J.JPROCONT.2009.04.001

Pörtner, R., Platas Barradas, O., Frahm, B., & Hass, V. C. (2017). Advanced Process and Control Strategies for Bioreactors. *Current Developments in Biotechnology and Bioengineering: Bioprocesses, Bioreactors and Controls*, 463–493. https://doi.org/10.1016/B978-0-444-63663-8.00016-1

Quintero, L., Amicarelli, A., & di Sciascio, F. (n.d.). *Recursive Bayesian Filtering for States Estimation: An Application Case in Biotechnological Processes*.

Raïssi, T., Ramdani, N., & Candau, Y. (2005). Bounded error moving horizon state estimator for non-linear continuous-time systems: application to a bioprocess system. *Journal of Process Control*, *15*(5), 537–545. https://doi.org/10.1016/J.JPROCONT.2004.10.002

Randek, J., & Mandenius, C. F. (2018). On-line soft sensing in upstream bioprocessing. In *Critical Reviews in Biotechnology* (Vol. 38, Issue 1, pp. 106–121). Taylor and Francis Ltd. https://doi.org/10.1080/07388551.2017.1312271

Rocha-Cózatl, E., & Wouwer, A. vande. (2011). State and input estimation in phytoplanktonic cultures using quasi-unknown input observers. *Chemical Engineering Journal*, *175*(1), 39–48. https://doi.org/10.1016/J.CEJ.2011.09.025

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

Safaeipour, H., Forouzanfar, M., & Casavola, A. (2021). A survey and classification of incipient fault diagnosis approaches. *Journal of Process Control*, *97*, 1–16. https://doi.org/10.1016/J.JPROCONT.2020.11.005

Salas, S. D., Romagnoli, J. A., Tronci, S., & Baratti, R. (2019). A geometric observer design for a semi-batch free-radical polymerization system. *Computers and Chemical Engineering*, *126*, 391–402. https://doi.org/10.1016/j.compchemeng.2019.04.026

Salehi, S., & Shahrokhi, M. (2008). Two observer-based nonlinear control approaches for temperature control of a class of continuous stirred tank reactors. *Chemical Engineering Science*, *63*(2), 395–403. https://doi.org/10.1016/J.CES.2007.09.040

Schaum, A., Mandis, M., Jerono, P., Tronci, S., & Meurer, T. (2019). State estimation of microalgae photobioreactors: Applications to Haematococcus pluvialis. *Chemical Engineering Transactions*, *74*, 1423–1428. https://doi.org/10.3303/CET1974238

Schlueter, K., Johnson, J., Warner, C., & Miller, W. (n.d.). *Glucose Analytical Comparability Evaluation of the YSI 2300 STAT Plus™ and YSI 2900D Biochemistry Analyzers Life Sciences Data for Life. TM*.

Semcheddine, S., & Bouchareb, H. (2019). Robust control and state estimation of a three-stage anaerobic digestion process. *Ecological Engineering and Environment Protection*, 29–38. https://doi.org/10.32006/eeep.2019.2.2938

Sevella, B. (1992). Bioengineering Operations. *Tankonykiado (Ed.)*.

Sheibat-Othman, N., Peycelon, D., Othman, S., Suau, J. M., & Févotte, G. (2008). Nonlinear observers for parameter estimation in a solution polymerization process using infrared spectroscopy. *Chemical Engineering Journal*, *140*(1–3), 529–538. https://doi.org/10.1016/J.CEJ.2007.11.039

Simutis, R., Galvanauskas, V., Levisauskas, D., Repsyte, J., & Vaitkus, V. (2013). Comparative Study of Intelligent Soft-Sensors for Bioprocess State Estimation. *Journal of Life Sciences and Technologies*, 163–167. https://doi.org/10.12720/jolst.1.3.163-167

Sonnleitner, B. (2012). Automated measurement and monitoring of bioprocesses: key elements of the M 3 C strategy. In *Measurement, monitoring and control of bioprocesses* (pp. 1–33).

Tamburini, E., Marchetti, M. G., & Pedrini, P. (2014). Monitoring key parameters in bioprocesses using near-infrared technology. *Sensors (Switzerland)*, *14*(10), 18941–18959. https://doi.org/10.3390/s141018941

Theuer, L., Lehmann, M., Junne, S., Neubauer, P., & Birkholz, M. (2017). Micro-electromechanical affinity sensor for the monitoring of glucose in bioprocess media. *International Journal of Molecular Sciences*, *18*(6). https://doi.org/10.3390/ijms18061235

Tronci, S., Bezzo, F., Barolo, M., & Baratti, R. (2005). Geometric observer for a distillation column: Development and experimental testing. *Industrial and Engineering Chemistry Research*, *44*(26), 9884–9893. https://doi.org/10.1021/ie048751n

Veloso, A. C., & Ferreira, E. C. (2017). Online Analysis for Industrial Bioprocesses: Broth Analysis. In *Current Developments in Biotechnology and Bioengineering: Bioprocesses, Bioreactors and Controls* (pp. 679–704). Elsevier Inc. https://doi.org/10.1016/B978-0-444-63663-8.00023-9

Venkateswarlu, C. (2004). Advances in monitoring and state estimation of bioreactors *. In *Journal of Scientific & Industrial Research* (Vol. 63).

Wang, G.-B., Peng, S.-S., & Huang, H.-P. (1997). A sliding observer for nonlinear process control. In *Chemical Engineering Science* (Vol. 52, Issue 5).

Wang, J., Zhao, L., & Yu, T. (2010). On-line Estimation in Fed-batch Fermentation Process Using State Space Model and Unscented Kalman Filter. *Chinese Journal of Chemical Engineering*, *18*(2), 258–264. https://doi.org/10.1016/S1004-9541(08)60351-1

Welch, G., & Bishop, G. (n.d.). *An Introduction to the Kalman Filter*. http://www.cs.unc.edu/~gb

Weltin, A., Slotwinski, K., Kieninger, J., Moser, I., Jobst, G., Wego, M., Ehret, R., & Urban, G. A. (2014). Cell culture monitoring for drug screening and cancer research: A transparent, microfluidic, multi-sensor microsystem. *Lab on a Chip*, *14*(1), 138–146. https://doi.org/10.1039/c3lc50759a

Zeitz, M. (1987). The extended Luenberger observer for nonlinear systems. *Systems & Control Letters*, *9*(2), 149–156. https://doi.org/10.1016/0167-6911(87)90021-1

Zhang, T., & Guay, M. (2002). Adaptive nonlinear observers of microbial growth processes. *Journal of Process Control*, *12*(5), 633–643. https://doi.org/10.1016/S0959-1524(01)00058-0

Chapter 3: Control strategies for biotechnological processes

After the bioprocess monitoring problem, possible control strategies will be discussed in this chapter. A brief overview of the state of the art is given. Next, simulation results regarding the implementation of control techniques for a fermentation bioreactor will be reported and discussed.

In order to develop a bioprocess that can meet the strict requirements imposed on product quality, along with the various constraints related to performance and productivity, it is necessary to have adequate control strategies. Unfortunately, designing a control system for processes like these is not an easy task for a variety of reasons such as the nonlinear characteristic of the system dynamics, the slow dynamic responses, the complex nature of the process due to the involvement of microorganisms, the lack of online sensors for measuring key process variables (i.e., biomass, substrate, product), and also the model uncertainties. As stated by Stanke & Hitzmann (2012), there are three basic phases that distinguish bioprocesses:

i. Up-streaming, where filling, sterilization, and mixing represent the main operations that must be performed.
ii. Cultivation/Fermentation, where cell growth, bioconversion, and production occur.
iii. Downstreaming, where collection, separation, and concentration are generally required.

Certainly, the step where it is most complicated to implement an automatic control is the cultivation phase, during which a combination of complex transport processes and numerous dynamical biochemical reactions occur. For this reason, it is absolutely critical to make sure that the environmental conditions in terms of pH, temperature, and dissolved oxygen are optimal so that the microorganisms can grow, multiply, and produce the desired product. Control of environmental variables is usually accomplished since there are sensors that proved measurements. Instead, the control of biological variables is difficult to carry out because there are no available measurements or, if there are, they are affected by significant delays. This represents a very discussed issue in literature. Indeed, biological variables represents the most informative ones, the trend of which help to understand if the process is evolving in the right direction and have a great influence on the efficiency of subsequent units, and thus on the target achievement. For this reason, the bioreactor represents undoubtedly the most difficult operational unit to control, but also the most important.

This chapter focuses on the problem of controlling bioprocesses, for this purpose a brief review of the control strategies of the most analyzed/controlled variables among the physical, chemical, and biological ones will be presented. Next, the result of the examined case study will be reported. Indeed, to complete the analysis conducted in the previous chapter, the implementation of different control strategies for a fermentation bioreactor will be evaluated. In addition, the obtained performance will be assessed considering both dynamic state trend and controller performance indexes.

3.1 General aspects of bioprocess control problem

The main purposes of implementing a control strategy in a bioprocess are the following:

- Overcome any form of malfunction
- Maximize the process yield
- Maximize the productivity and the product quality
- Prevent the product inhibition minimizing the formation of unwanted by-products
- Maintain/Ensure an optimal environment for the growth of microorganisms, and thus for product formation
- Ensure that the final product meets the regulatory standards imposed, as well as the necessary levels of quality and safety

Some of these goals can be guaranteed/achieved by implementing standard control algorithms. For example, bioreactors are generally equipped with probes for automatic control of temperature, pH, dissolved oxygen, and even the addition of antifoaming agents (Stanke & Hitzmann, 2012). However, these are not always sufficient solutions. The next section will present a brief state of the art of bioprocess control, along with recent advances reported in the literature.

3.2 Overview of bioprocess control techniques

T control

As already explained, temperature is one of the variables usually controlled in a bioprocess, and through which indirect control of key unmeasurable variables such as product composition can be realized. Typically, the control of bioreactor temperature is realized by manipulating the flow rate of water in the cooling jacket using a pump. The opening of the valve allows the flow to be regulated, maintaining the required temperature. However, since bioprocesses like fermentation exhibit a strongly nonlinear nature, temperature control can become a challenging task. Several efforts have been made over the years in order to obtain improvements, as demonstrated by the following papers on this topic. Nagy (2007) developed a detailed analytical model describing the dynamic performance of a continuous fermentation bioreactor for ethanol production. This model simulating the real process was then used to generate so-called training data to develop a feedforward artificial neural network (ANN). This ANN model was then implemented as an internal model in a MPC control algorithm to predict control actions. The performance of the so-called Neural Network model based predictive control

(NNMPC) was compared with that of Linear Model Predictive Control (LMPC) and proportional-integral-derivative (PID) control, and the results obtained have demonstrated the inadequacy of the traditional controllers. Moreover, the robustness of the proposed NNMPC was evaluated not only against set-point change, but also against temperature measurements affected by noise. Based on the same model, an efficient nonlinear MPC together with Nonlinear Prediction and Linearisation (MPC-NPL) was developed by Ławryńczuk (2008) to control the temperature of the yeast fermentation bioreactor. The MPC-NPL algorithm showed a good control accuracy and good disturbance rejection abilities in presence of noisy measurements and process disturbances. An inverse neural network (INN) controller for the temperature profile in an ethanol production process was studied (Imtiaz et al., 2013). Its performance evaluated by the mean square error (MSE) criterion was found to be better than that of conventional PID controller, which also resulted in a higher amount of ethanol produced. In order to overcome the problem related to the significant offset error that characterizes fractional order IMC-PID (FOIMC-PID) controllers for nonlinear processes such as fermentation, Pachauri et al. (2017) designed a modification. Indeed, an extra proportional feedback loop was added, which increases the overall gain of the system and leads to the modified control structures (MFOIMC-PID). A metaheuristic optimization algorithm (water cycle algorithm, WCA) was subsequently implemented to tune the controller parameters, leading to the final WMFOIMC-PID control structure. Fonseca et al. (2013) developed a fuzzy-PI controller together with a split range control strategy to regulate the fermentation temperature. Its performance was compared with that of a conventional PI, and a reduction in control effort and total demand of utilities were observed. A MIMO control system was proposed by Imtiaz et al. (2014), which involves an auto regressive moving average (NARMA) neuro controller for temperature control, and a two degree of freedom PID (2DPF-PID) was used to control pH and dissolved oxygen. The performance of the controllers was also tested at rapid set-point change, and satisfying results were obtained in terms of response time, residual error reduction and delay time. Instead, a temperature control algorithm based on Takagi-Sugeno approach was designed by Flores-Hernández et al. (2018). They also proposed a new nonlinear representation of the fermentation process, by means of a T-S model, that can better reproduce/describe the nonlinear dynamics, allowing a wider range of applications of the control algorithms. Recently, Kumar et al. (2019) successfully designed and tested an internal model control based proportional-integral-derivative (IMC-PID) controller for a fermentation bioreactor and Bakošová et al. (2019) studied the implementation of a robust model-based predictive control with integral action (RMPC-IA) to control a biochemical reactor. The

performances were compared with those of NN predictive control, and the results showed that it was outperformed by RMPC-IA, which guaranteed maximum product yield and minimum energy consumption.

DO control

The dissolved oxygen (DO) concentration is another important process parameter influencing the cell viability and the process yield. It is one of the most controlled variables and this task is usually accomplished by regulating the mixture agitation in the bioreactor, and in particular by manipulating the stirrer speed. Several studies have been carried out over the years on this topic, trying to design an efficient DO concentration controller. An important contribution is that of Gomes & Menawat (2000), who developed a Model-Based Geometric Algorithm (MGA) to control the DO concentration in fermentation processes. Two different components can be identified in this approach: the predictor, whose task is to estimate the DO concentration one time step ahead based on the profile of state variables in time; the controller, which calculates the control action using the estimated values. Its performance was evaluated and compared with that of IMC and PI controllers, both through simulations and through the online implementation for DO control during the manufacturing process of an antibiotic. Ertunc et al. (2009) implemented conventional PID and self-tuning generalised minimum variance (GMV) control algorithms to track the dissolved oxygen concentration in a batch reactor. Two set-point variations were examined to evaluated performance of the two algorithms in comparison. The results showed that better performance can be obtained with the self-tuning generalised minimum variance (STGMV) algorithm than the conventional PID. Chitra et al. (2018) proposed a Model Reference Adaptive Control (MRAC) scheme to control DO in an aerobic fermentation process. The comparison between the performance of the proposed MRAC strategy and a conventional PI controller revealed that the former provides better tracking performance than the PI controller. Butkus et al. (2021) developed an adaptive control system for controlling the set-point and rejecting disturbances of the dissolved oxygen concentration. The gain scheduling of PID (PI) controller is used, which is based on the controller input and output signals eliminating the need for online measurements of process variables to develop gain scheduling algorithms. The authors also designed control algorithm for set-point tracking and disturbance rejection during DO concentration control for the bioreactor, which operated both in batch and fed-batch mode. The performance of this control system was evaluated

considering extreme operating conditions, and the results showed significant advantages of the proposed control solution compared to conventional PI control.

pH control

As reported by Najafpour (2006), pH is generally kept constant during the fermentation process since it could change with the metabolic product of the microorganism and significantly affect cell growth and product formation. For this reason, it is necessary to control it, and this is usually done by regulating the flow rate of acid/base. Together with the temperature, it belongs to those variables that are chosen first to be controlled during a bioprocess. As explained by Gnoth et al. (2010), the reason for this is that pH and temperature affect both the specific growth rate and the product formation rate. However, the impact that the temperature has on the specific growth rate is greater, and therefore pH control does not receive the same attention (Simutis & Lübbert, 2015). Indeed, only few papers can be found in the literature that have investigated its control. Mészáros et al. (2004) proposed the use of ANN models to identify and control pH, together with DO concentration, for a fermentation process with a *Saccharomyces cerevisiae*-based culture. An adaptive term was added to the control scheme, resulting in more robust regulatory and tracking performance. Another work worth mentioning is that of Gnoth et al. (2010), who proposed an adaptive controller for the pH of the fermentation of a recombinant protein production process. Reductions in the pH value and corresponding decreases in the base consumption signal were observed during the whole process by means of the application of the gain scheduling technique. This represents a PID control technique suitable for nonlinear processes with dynamics influenced by operating conditions such as biotechnological production process and that allows the controller parameters to dynamically adapt to changes due to process dynamics.

Biomass, substrate, and product control

During a bioprocess, in addition to physical and chemical parameters, biological parameters such as biomass, substrate and product concentration must also be collected. However, the lack of reliable biological sensors makes it poorly possible to obtain feedback of biological information. This is undoubtedly the main obstacle to the introduction of advanced controls. As seen in the previous chapter, the issue of lack of measurements, needed to implement control strategies, is usually handled with estimation techniques, which are a cheaper alternative to biosensors. Although bioprocesses with a deficiency of available online sensors represent the

majority, there are still few works in the literature on this topic. Some of them are reported below. Ajbar & Ali (2017) proposed an advanced Nonlinear Model Predictive Control (NLMPC) strategy to control a continuous fermentation process. The Kalman filter estimation algorithm was applied to reconstruct the output states, which were also corrected by the additive disturbances estimates. Ethanol concentration, productivity, and the inverse of productivity were chosen as controlled variables. Simulation results revealed that satisfactory closed-loop performance can be obtained for both servo and regulatory control problems. A standard PI controller was also investigated for comparison and found to be reasonable when the product concentration was used as the controlled variable, but it couldn't work properly when the productivity was the controlled output. In all cases, and regardless of control algorithm used, the closed-loop responses suffered from slow dynamic. The robustness of NLMPC algorithm in the face of model-plant mismatch was also tested and results showed that it is able to reach the control objectives even in presence of parametric errors in the model. Arndt & Hitzmann (2004) presented a substrate control system which demonstrated to be able to control *Saccharomyces cerevisiae* cultivations also at low glucose concentration. The glucose concentration measurements were determined using a special flow injection analysis (FIA). An extended Kalman filter (EKF) was applied for smoothing glucose measurements as well as for the prediction of glucose and biomass concentration, the maximum specific growth rate, and the volume of the culture broth. Then, the predicted values were used for feedforward/feedback control. The results obtained showed that the combination of a rapid glucose measurement method with an estimation technique makes it possible to control glucose composition even at low setpoint levels. Petre et al. (2021) developed an advanced control strategy for a fermentation bioreactor for ethanol production. The proposed control scheme included two control loops, one for temperature and the second for substrate concentration. Moreover, a state observer for reconstructing biomass composition, a sliding mode observer (SMO) for estimating substrate concentration, and an estimator for reconstructing specific reaction rate were also included. Both loops were based on adaptive control laws. The results obtained are encouraging, and despite the disturbances and uncertainties evaluated to test control performance, the adaptive control structure has proved to achieve the control objective, i.e., a substantial amount of ethanol and a low level of residual substrate. In other works, the possibility of having both online and offline measurements of biological variables has been considered. Persad et al. (2013) designed a decoupled input-output linearizing controller (DIOLC) as an alternative advanced control strategy for controlling a fermentation process where *Saccharomyces cerevisiae* were used. In order to achieve defined ethanol and biomass

production, control of substrate and dissolved oxygen was realized by manipulating glucose feed and air flow rate, respectively. It was assumed that only substrate and oxygen concentrations could be measured online, while biomass and ethanol compositions were available as offline measurements. The performance of the proposed controller was then compared with that of a PID controller. The results showed that DIOLC performance was better in tests where an accurate response to simultaneous changes in substrate and dissolved oxygen trajectories was required. Moreover, it has demonstrated better performances even in the presence of perturbations of significant parameters of the process model. Dewasme et al. (2010) presented an adaptive Reference Signal Tracking (RST) control scheme for regulating ethanol concentration in a fed-batch culture of *Saccharomyces cerevisiae*. The main objective of this work was to demonstrate the efficiency and robustness of this controller in experimental applications from laboratory to industrial scales. The basic law of the controller was to regulate ethanol composition by manipulating the inlet flow rate to the fed-batch system. The ethanol concentration was measured online by a specific probe, while other offline measurements by gas chromatography were available for validation reasons. Regardless of the bioreactor scale, the investigated controller showed good performance and reliability under different conditions. Compared with conventional open-loop operations and other closed-loop PID-like control strategies, this particular closed-loop control scheme provided robust ongoing control and significant productivity.

3.3 Different control strategies for a yeast fermentation bioreactor

In this section, the problem of controlling a fermentation bioreactor for ethanol production will be addressed. A classic PI algorithm was applied for its resolution, which was proved to achieve the same results obtainable with more sophisticated techniques such as neural networks. Moreover, it was decided to implement a control system also for the product concentration by following two different approaches. First, a cascade control was developed in which the master is the ethanol composition controller that determines the temperature set-point. In this case, ethanol measurements affected by time delay were considered available, with the aim of improving the control system performance and making sure that it can achieve the target product composition, even in presence of disturbances. Then, the implementation of a MIMO system, where reactor temperature and product composition are both controlled, was evaluated.

An inferential control was used for ethanol concentration, thanks to the presence of an estimator that reconstructed the composition of unmeasurable ethanol.

3.3.1 Bioreactor model

The system here investigated is the same for which, in the previous chapter (2.3.1), the monitoring problem was discussed. The model taken as a reference was developed in detail by Nagy (2007) and it describes the dynamic behaviour of the following six states: biomass concentration (C_X), ethanol concentration (C_P), substrate concentration (C_S), dissolved oxygen concentration $\left(C_{O_2}\right)$, reactor temperature (T_r), and jacket temperature $\left(T_{ag}\right)$. For the sake of clarity, only the mass and thermal energy balances of the analyzed system will be reported (Eqs. 47-52), omitting the values of model parameters and nominal operating conditions, as well as all the other relationships, for which the reader is referred to the previous chapter.

$$\frac{dC_X}{dt} = \mu_X\, C_X\, \frac{C_S}{K_S + C_S} e^{-K_P C_P} - \frac{F_e}{V} C_X \tag{47}$$

$$\frac{dC_P}{dt} = \mu_P C_X \frac{C_S}{K_{S1} + C_S} e^{-K_{P1}\, C_P} - \frac{F_e}{V}\, C_P \tag{48}$$

$$\frac{dC_S}{dt} = -\frac{1}{R_{SX}} \mu_X C_X \frac{C_S}{K_S + C_S} e^{-K_P C_P} - \frac{1}{R_{SP}} \mu_P C_X \frac{C_S}{K_{S1} + C_S}\, e^{-K_{P1} C_P} + \frac{F_i}{V} C_{S,in} - \frac{F_e}{V} C_S \tag{49}$$

$$\frac{dC_{O_2}}{dt} = k_l a \left(C_{O_2}^* - C_{O_2}\right) - \mu_{O_2} \frac{1}{Y_{O_2}} C_X \frac{C_{O_2}}{K_{O_2} + C_{O_2}} \tag{50}$$

$$\frac{dT_r}{dt} = \left(\frac{F_i}{V}\right)(T_{in} + 273) - \left(\frac{F_e}{V}\right)(T_r + 273) - \mu_{O_2} \frac{1}{Y_{O_2}} C_X \frac{C_{O_2}}{K_{O_2} + C_{O_2}} \frac{\Delta H_r}{32\, \rho_r\, C_{heat,r}} - \frac{K_T A_T \left(T_r - T_{ag}\right)}{V\, \rho_r\, C_{heat,r}} \tag{51}$$

$$\frac{dT_{ag}}{dt} = \left(\frac{F_{ag}}{V_j}\right)\left(T_{in,ag} - T_{ag}\right) + \frac{K_T A_T \left(T_r - T_{ag}\right)}{V_j \rho_{ag} C_{heat,ag}} \tag{52}$$

Introducing the state vector in the following form:

$$\boldsymbol{x} = \left[C_X, C_P, C_S, C_{O_2}, T_r, T_{ag}\right] \tag{53}$$

But also output measured vector ***y*** defined later, and the input vector:

$$\boldsymbol{u} = [F_i, F_{ag}] \tag{54}$$

Then, the system dynamics can be compactly written as:

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) \tag{55}$$

$$\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) \tag{56}$$

Where $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u})$ is the vector field of the system and $\boldsymbol{h}(\boldsymbol{x})$ is the vector relating states and measured outputs.

3.3.2 Control system design

As seen so far, the quality of the final products represents an essential parameter to be controlled in the bioreactor. The objective of the controller is to achieve the desired concentration as early as possible in the presence of disturbances and noise. This task has been solved by studying different situations with respect to measured outputs. First, only temperature measurements (reactor and cooling agent) are considered available, then the output vector is the following:

$$\boldsymbol{y}_I = \left[T_r(t), T_{ag}(t)\right] \tag{57}$$

As second possible options, ethanol concentration is available with delay (t_d) due to the time required by the analyser to perform the measurement, with output vector described by Eq. (58):

$$\boldsymbol{y}_{II} = \left[C_P(t - t_d), T_r(t), T_{ag}(t)\right] \tag{58}$$

The final case is when substrate and oxygen concentration can be measured online along with temperature, and the output vector is:

$$\boldsymbol{y}_{III} = \left[C_S(t), C_{O_2}(t), T_r(t), T_{ag}(t)\right] \tag{59}$$

Substrate and dissolved oxygen concentration measurements have been considered online, without delay, according to the precious chapter and a review on the available sensors for biosystems (Holzberg et al., 2018).

Controllability

Controllability of a *n*-dimensional linear system, *m* inputs and *l* outputs in the form (66) can be assessed by considering the controllability matrix $\boldsymbol{L_c}$ (61)

$$\frac{dx}{dt} = \mathbf{A}\boldsymbol{x} + B\boldsymbol{u}, \boldsymbol{y} = C\boldsymbol{x}\,, \boldsymbol{x} \in \Re^n, \boldsymbol{u} \in \Re^m, \boldsymbol{y} \in \Re^l \quad (60)$$

$$L_c = [\mathbf{B}\ \mathbf{AB}\ \mathbf{A}^2\mathbf{B}\ \ldots\ \mathbf{A}^{\mathrm{n}-1}\mathbf{B}] \quad (61)$$

This definition can be used to assess local controllability of a nonlinear system if **A** is the Jacobian matrix calculated at the reference conditions and the coefficients of **B** are the derivative of the functions describing the dynamics of the states with respect to the inputs. Controllability is verified if the rank of matrix $\mathbf{L}_c$ is equal to the dimension of the state vector. By taking the Jacobian matrix of (47-52), it is possible to verify that the system is locally controllable when coolant and inlet flow rates, F_{ag} and F_i, can be both manipulated. The system is still locally controllable when only F_{ag} is manipulated because kinetic parameters and mass transfer coefficient depends on reactor temperature.

Observability

Real-time information about concentration of product, substrate and biomass is the key to controlling and optimizing the bioreactor. When these variables are not measured online, soft sensors can be used to obtain information on their dynamics if observability is satisfied.

Observability of a *n*-dimensional linear system, *m* inputs and *l* outputs in the form (60) can be assessed by considering the observability matrix $\boldsymbol{L_o}$ in Eq. (62):

$$\boldsymbol{L_o} = [\mathbf{C}\ \mathbf{CA}\ \mathbf{CA}^2\ \ldots\ \mathbf{CA}^{\mathrm{n}-1}]^T \quad (62)$$

In case that only temperature measurements are available, the system is not observable therefore it is not possible to reconstruct the dynamics of all the states. Local observability is satisfied when the output vector $\boldsymbol{y}_{III}$ is considered, as demonstrated by the results shown before.

Temperature control

Because of the controllability property, it is theoretically possible to drive the system to the required conditions for all the six states by using only one manipulated variable. When

considering that only temperature measurements are available, the first proposed solution is to design a temperature controller using the coolant flow rate as manipulated variable. Temperature set-point is selected such that the required product composition is obtained. A schematization of the proposed control structure is shown in Figure (17). A PI from IMC algorithm (Skogestad, 2003) is used to control the output and a step-response identification method is used to obtain the input-output model.

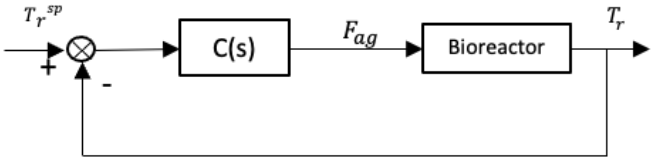

Figure 17. Configuration of feedback temperature control.

Cascade control with concentration delayed measurements

Even if controllability is satisfied, some changes in the process conditions (disturbances) may cause a discrepancy between the desired species concentration and actual values even if temperature is maintained at set-point. In this case it could be useful to add another control loop that guarantees the respect of product quality. Indeed, this variable generally influences the successive separation process. As it can be observed in Figure (18), the delayed ethanol concentration measurement has been used in a cascade arrangement, where the outer loop exploits a discrete regulator to keep the product concentration around a desired value, while the inner loop guarantees that the bioreactor temperature was maintained at a predetermined set-point. The delay in the composition measurements of the product has been included in the simulation so that the updated C_P value was available only after 18 minutes. The use of a cascade control guarantees a faster response because temperature measurement is continuously available, while the outer loop reduces, when necessary, the offset for the ethanol concentration.

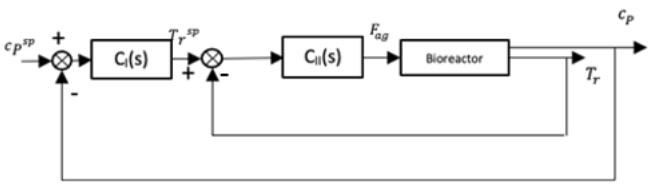

Figure 18. Configuration of cascade control, with ethanol concentration delayed measurements.

Inferential control

When ethanol concentration measurement is not available online, but the system is observable, that is the case of measured output $\boldsymbol{y}_{III}$, inferential control can be used to ensure product quality. This strategy is based on the state estimator reported in the previous chapter. The ethanol concentration has been inferred by applying the extended Kalman filter, which is one of the most widely used estimation techniques for monitoring bioprocesses (Dewasme et al., 2013). The estimated value has been used in a classical feedback control strategy manipulating the inlet flow rate. This control solution involves two loops, one for ethanol composition and the other for reactor temperature (Figure (19)).

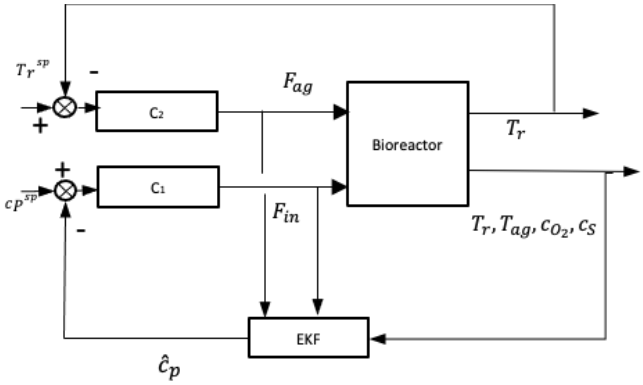

Figure 19. Configuration of MIMO control (2x2 system), with estimation of ethanol composition measurements.

3.3.3 State Estimation

The development of a state estimation for the bioreactor in (53-58) has been discussed in the previous chapter. To facilitate understanding, the obtained estimator structure is reported here again. The estimated states have been partitioned in innovated ($\hat{\boldsymbol{x}}_i$) and not innovated ($\hat{\boldsymbol{x}}_u$) states (69-72), defined in (73), following the procedure reported in Salas et al. (2019). Here it is reminded that the innovated states are dynamic states of the estimation model whose changes are captured by the secondary measurements (69), while the not innovated states are inferred by the estimation model in an open-loop mode (Porru & Özkan, 2017).

$$\frac{d\hat{\boldsymbol{x}}_i}{dt} = \boldsymbol{f}_i(\hat{\boldsymbol{x}}_i, \hat{\boldsymbol{x}}_u, \boldsymbol{u}) + \boldsymbol{K}(\boldsymbol{y} - \hat{\boldsymbol{y}}), \hat{\boldsymbol{x}}_i(t_0) = \hat{\boldsymbol{x}}_{i,0} \quad (69)$$

$$\frac{d\hat{x}_u}{dt} = \boldsymbol{f}_u(\hat{\boldsymbol{x}}_i, \hat{\boldsymbol{x}}_u, \boldsymbol{u}), \hat{\boldsymbol{x}}_u(t_0) = \hat{\boldsymbol{x}}_{u,o=0} \quad (70)$$

$$K = \boldsymbol{P}(t)\boldsymbol{H}^T\boldsymbol{R}^{-1} \quad (71)$$

$$\dot{\boldsymbol{P}}(t) = \boldsymbol{P}(t)\boldsymbol{F}(t) + \boldsymbol{F}^{\boldsymbol{T}}(t)\boldsymbol{P}(t) + \boldsymbol{Q}(t) - \boldsymbol{K}(t)\boldsymbol{H}(t)\boldsymbol{P}(t), \boldsymbol{P}(t_0) = \boldsymbol{P}_0 \quad (72)$$

$$\hat{\boldsymbol{x}}_i = \left[\hat{C}_X, \hat{C}_S, \hat{C}_{O_2}, \hat{T}_r, \hat{T}_{ag}\right], \hat{\boldsymbol{x}}_u = \left[\hat{C}_P\right] \quad (73)$$

$\boldsymbol{F}(t)$ is the Jacobian of the vector field $\boldsymbol{f_i}(\hat{\boldsymbol{x}}_{\boldsymbol{i}}, \hat{\boldsymbol{x}}_{\boldsymbol{u}}, \boldsymbol{u})$, calculated with respect to the innovated states, $\boldsymbol{P}(t)$ is the error covariance matrix of the innovated states, $\boldsymbol{H}(t)$ is the matrix of the derivative of the map $\boldsymbol{h}$ with respect to the states, $\boldsymbol{Q}$ and $\boldsymbol{R}$ are, respectively, the covariance matrix of the model and measurements errors (Jazwinski, 2007). The constant matrix $\boldsymbol{Q}$, $\boldsymbol{R}$, and $\boldsymbol{P}_0$ are tuning parameters of the estimation model and they have been calculated minimizing the error between the states calculated with the simulated plant and the estimator along a reference trajectory.

As demonstrated in the previous chapter, the estimator is more robust and efficient with the selected configuration with respect to the use of a full order structure, in which all states are innovated. It has been proved that ethanol concentration trajectory is well reconstructed by the EKF even if it is not innovated and it can be used to design an inferential control.

3.3.4 Performance Indexes

In order to optimize the choice of the best control structure among those proposed here, the time-integral performance criteria have been used. In particular, the integral of the squared error (ISE), the integral of the absolute value of the error (IAE) and the integral of the time-weighted absolute error (ITAE) have been calculated to compare the performance of the analysed control structures. The indexes have been calculated using the following equations (74-76):

$$ISE = \int_0^\infty e^2(t)dt \quad (74)$$

$$IAE = \int_0^\infty |e(t)|dt \quad (75)$$

$$ITAE = \int_0^\infty t|e(t)|dt \quad (76)$$

Since the goal of the control structure is to obtain the desired ethanol concentration, *e(t)* is the error signal obtained as the difference between the required ethanol concentration and the output signal. As showed in Eq. (74), the ISE error criterion integrates the square of the error over time. It penalizes large errors relative to small errors, since their square is significantly bigger. The ITAE (Eq. (76)) integrates the absolute error, multiplied by time, over time. This criterion obviously weights more significantly the errors that occur after a long period of time than those occurring at the beginning of the system response. Instead, the IAE (Eq. (75)) minimizes the error calculated by the previous two criteria since it integrates it over time, without adding any weight to the errors in the system response.

3.3.5 Results

This section analyses and compares the performance of different designed controllers for disturbance rejection. Three control strategies are compared:

i. Reactor temperature control (SISO system)
ii. Cascade control using ethanol delayed measurements where temperature control is the secondary loop
iii. Inferential control for ethanol concentration and reactor temperature control (2x2 MIMO system).

In the simulations, step variations of the following three inputs have been considered as disturbances: the inlet temperature (T_{in}); the substrate inlet concentration ($C_{s,in}$); the biomass specific growth rate (μ_X), which could change due to a pH variation or the presence of possible inhibitors. The steps used to excite the system are reported in Figure 20.

SISO control configuration

The results of the SISO configuration as the three inputs vary are reported in Figures 21-23. In detail, Figures 21(a) and 21(b) show the simulated closed-loop response of the bioreactor temperature T_r and the product concentration C_p respectively, with only temperature feedback control, along with the manipulated variable F_{ag}, when a step-change in T_{in} is introduced. It can be observed that changes in inlet temperature have a small effect on ethanol composition (Figure 21(b)) compared the other two disturbances. Indeed, when T_{in} varies, the SISO configuration with only feedback temperature controller is able to maintain the ethanol concentration at the required set-point. Figures 22(a)-(b) and 23(a)-(b) represent the simulated

closed-loop trend of the variables T_r and the ethanol concentration C_p, along with the manipulated variable F_{ag}, when varying $C_{s,in}$ and μ_X, respectively.

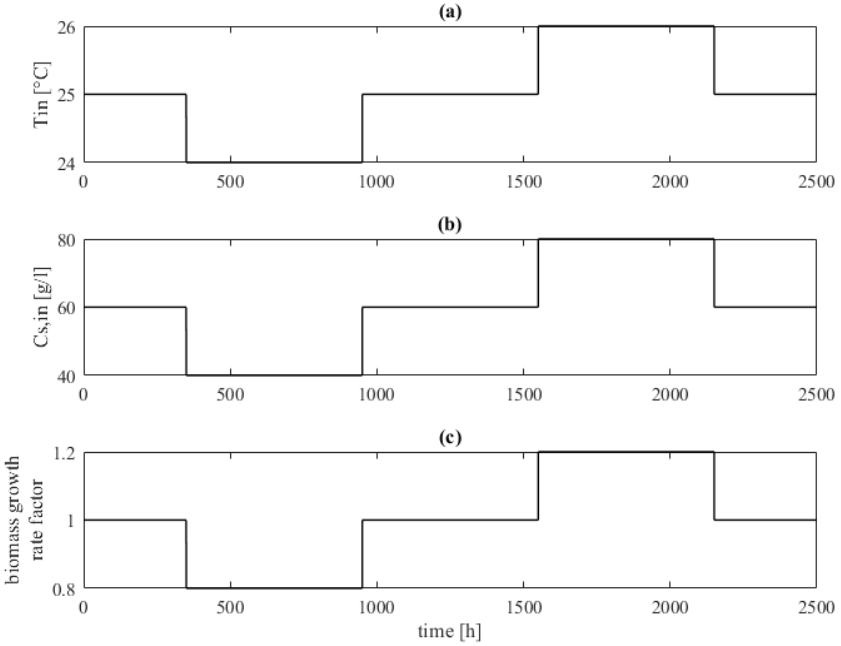

Figure 20. Disturbance trajectories used to analyse the control for three runs: inlet temperature (a), substrate inlet concentration (b), and biomass growth factor (c).

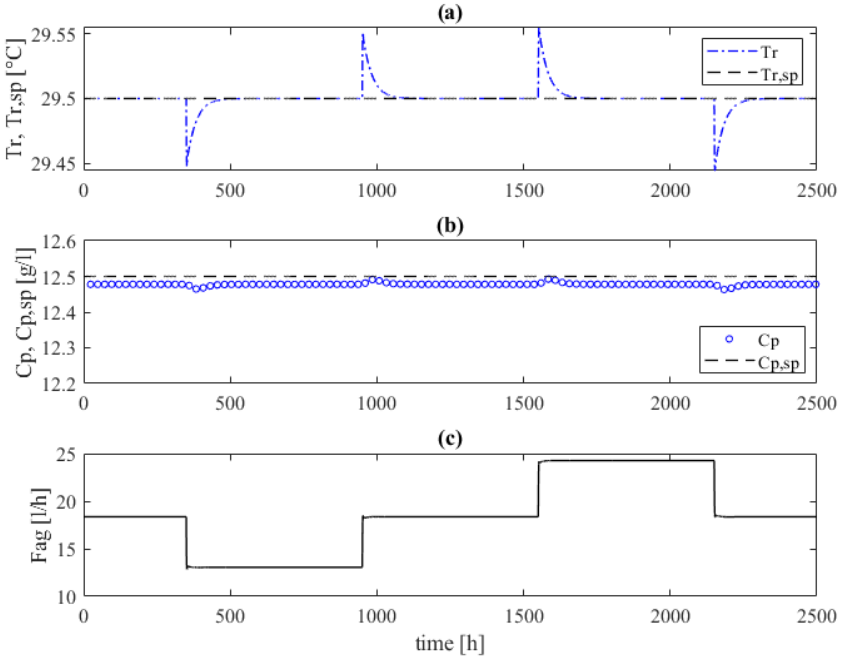

Figure 21. SISO control performances when T_{in} vary: (a) controlled reactor temperature; (b) ethanol concentration (open-loop); (c) manipulated coolant flow rate.

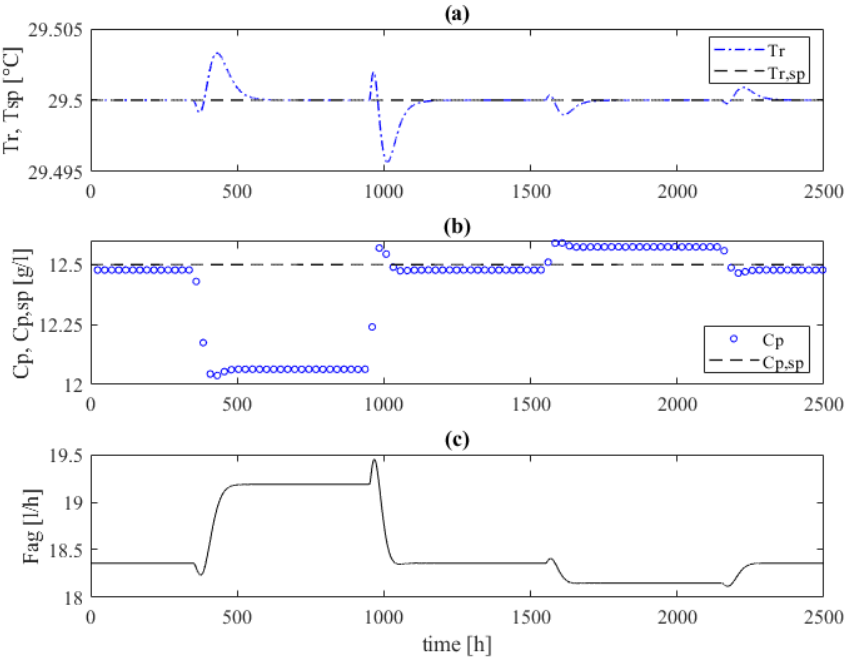

Figure 22. SISO control performance when $C_{S,in}$ vary: (a) controlled reactor temperature; (b) ethanol concentration (open-loop); (c) manipulated coolant flow rate.

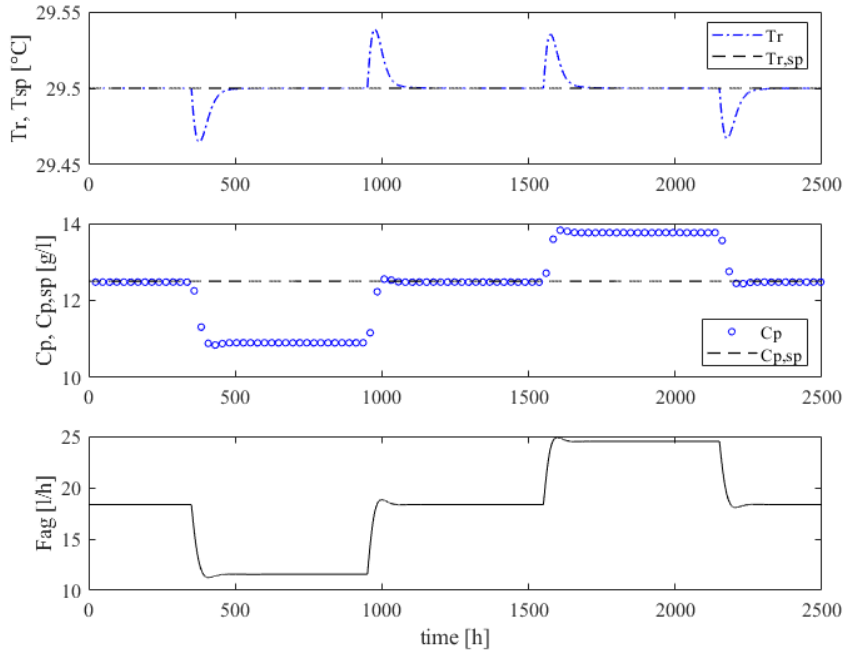

Figure 23. SISO control performance when μ_X vary: (a) controlled reactor temperature; (b) ethanol concentration (open-loop); (c) manipulated coolant flow rate.

In all three cases, results show that temperature controller cannot guarantee the desired value of ethanol composition and an offset is registered (Figs. 21(b), 22(b), and 23(b)). The offset is significantly higher when the reactor conditions imply a change in the biomass growth rate, as it may happen when μ_X varies. These results are confirmed by the performance index values reported in Table 11.

	SISO		
	IAE	**ISE**	**ITAE**
T_{in}	62.87	1.42	9.91E+04
$C_{s,in}$	340.32	115.39	4.62E+05
μ_X	1.76E+03	2.44E+03	2.98E+06

Table 11. Controller performance indexes for SISO control structure.

Cascade control configuration

In Figures 24, 25, and 26 the results obtained when cascade control is used are reported. As expected, this control structure has proven to be more effective despite the time delay of ethanol composition measurements. This can be confirmed by evaluating the error indices shown in Table 11 and 12. The cascade controller is more performing than the SISO system, thus determining a smaller error and a faster achievement of the desired ethanol composition. It is possible to observe, in presence of disturbances, how the ethanol composition can quite well follow the setpoint value by the action of outer loop (Figs. 24(b), 25(b), 26(b)). In this way, the set-point of T_r is modified in order to ensure the required product concentration. The temperature controller is effective to follow the set-point variations.

	Cascade		
	IAE	**ISE**	**ITAE**
T_{in}	5.85	0.04	9.10E+03
$C_{s,in}$	44.12	7.52	6.08E+04
μ_X	216.04	106.79	3.76E+05

Table 12. Controller performance indexes for cascade control structure.

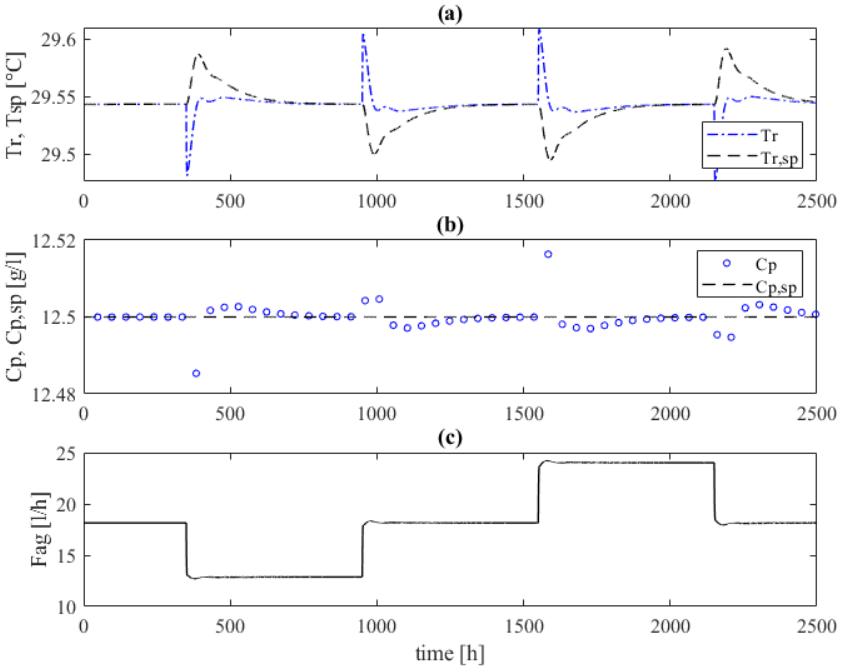

Figure 24. Cascade control performances when T_{in} vary: (a) controlled reactor temperature; (b) ethanol concentration (open-loop); (c) manipulated coolant flow rate.

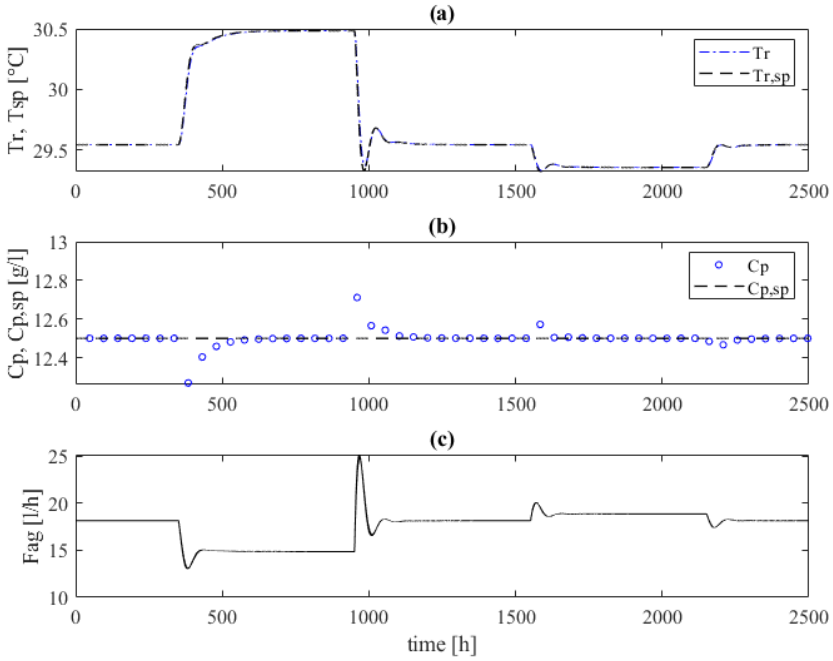

Figure 25. Cascade control performance when $C_{S,in}$ vary: (a) controlled reactor temperature; (b) ethanol concentration (open-loop); (c) manipulated coolant flow rate.

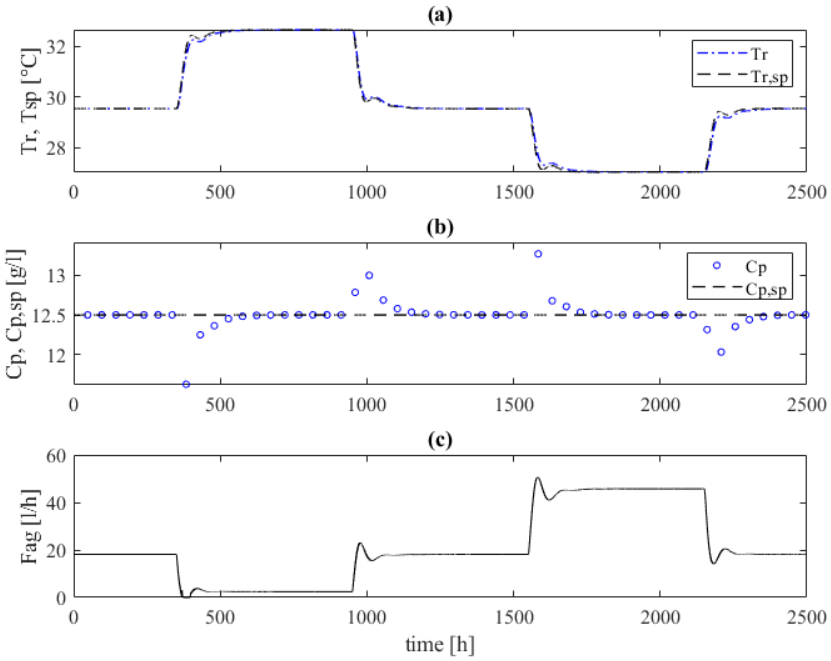

Figure 26. Cascade control performance when μ_X varies: (a) controlled reactor temperature; (b) ethanol concentration (open-loop); (c) manipulated coolant flow rate.

MIMO control configuration

In Figures 27, 28, and 29 the trends obtained with MIMO configuration are reported, where an inferential control is used for the ethanol composition. In this case, the inferred concentration control loop allows to maintain the product close to the set-point. As it can be observed in Figs. 27(b), 28(b), and 29(b), in correspondence to the variations of μ_X factor, it can be observed that C_p moves further away from the setpoint value than when $C_{S,in}$ disturbance is applied. However, the controller brings it back quickly enough to the desired value compared to the cascade control configuration. Therefore, the MIMO control system handles the situation of bioreactor temperature and composition control more effectively compared to other designed structures when disturbances include inlet temperature, substrate concentration, or conditions affecting the growth of microorganisms. The quality indexes reported in Table 13 confirm such considerations.

MIMO			
	IAE	**ISE**	**ITAE**
T_{in}	4.12	0.02	6.62E+03
$C_{s,in}$	34.34	1.73	4.99E+04
μ_X	164.02	27.38	3.04E+05

Table 13. Controller performance indexes for MIMO control structure that used the inferential control for the ethanol concentration.

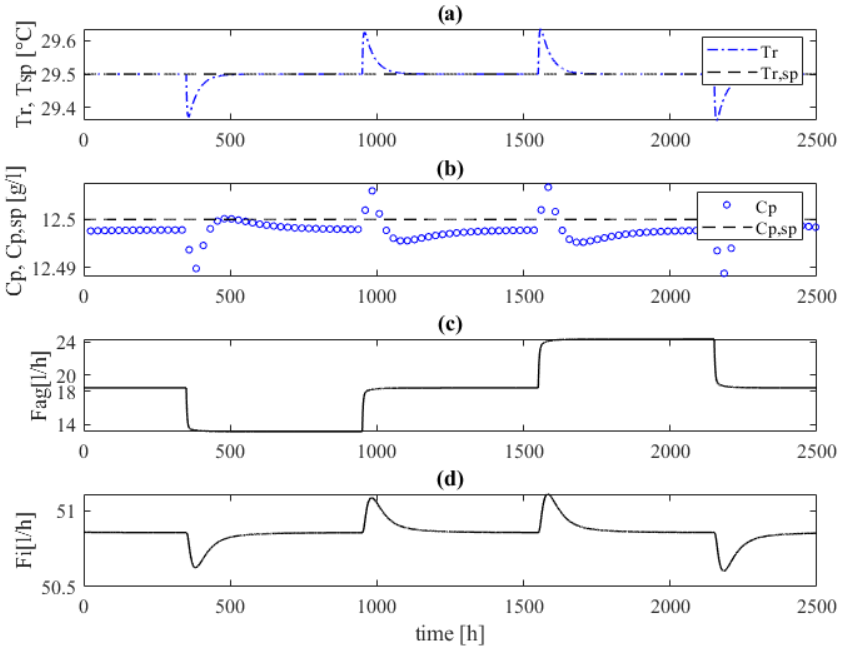

Figure 27. MIMO control performances when T_{in} vary: (a) controlled reactor temperature; (b) ethanol concentration (open-loop); (c) manipulated coolant flow rate; (d) manipulated inlet flow rate.

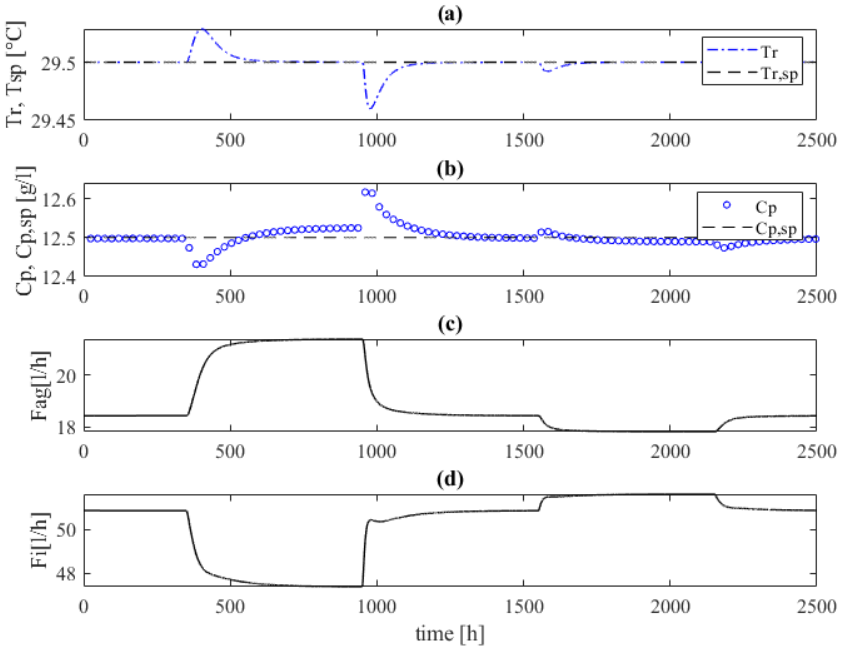

Figure 28. MIMO control performance when $C_{S,in}$ vary: (a) controlled reactor temperature; (b) ethanol concentration (open-loop); (c) manipulated coolant flow rate; (d) manipulated inlet flow rate.

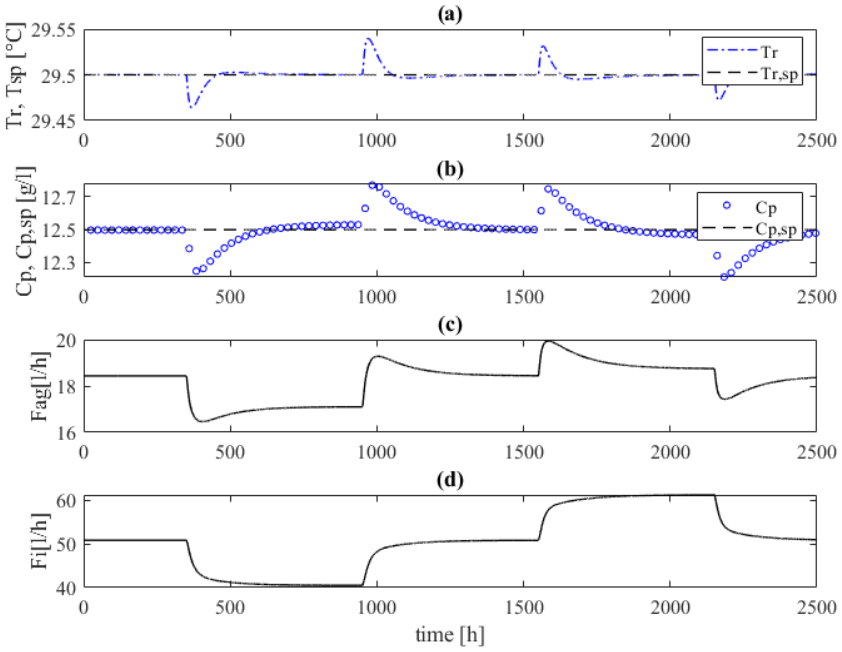

Figure 29. Cascade control performance when μ_X vary: (a) controlled reactor temperature; (b) ethanol concentration (open-loop); (c) manipulated coolant flow rate; (d) manipulated inlet flow rate.

3.4 Conclusions

The implementation of an adequate control system for biotechnological processes represents a particularly discussed topic. In this chapter, it has been shown that its presence is critical in order to meet the standards imposed by the industry in terms of quantity and quality. However, this is a difficult task due to the nonlinear nature of the process and the lack of sensors that provide real-time measurements of process variables. Analysis of the state of the art has made it clear that so-called biological variables are the most difficult to control. Regarding this, the case study presented here focused on the design and implementation of different control strategies for a fermentation bioreactor in order to achieve efficient and precise control of temperature and product composition. Performance comparison of the controllers designed showed that the temperature control was not able to maintain the ethanol concentration at the required set-point when disturbances varied substrate concentration in the bioreactor or when process conditions, such as pH or presence of inhibitors, affected the biomass growth rate. Thus, when significant disturbances are present, as the ones considered here, it is necessary to develop different control strategies that can efficiently eliminate disturbances effect on the product concentration. In order to improve system performance, two different scenarios were considered: (i) online analyser for measuring ethanol concentration with significant delay; (ii) estimation of ethanol by means of secondary available measurements. In the first case, a cascade control was proposed, where temperature control received the set-point from ethanol concentration control loop. In this way, the offset on the product composition was reduced with respect to using only temperature controller. In the second scenario, an inferential control for the ethanol concentration in conjunction with the temperature controller was implemented. The 2x2 MIMO control system outperformed the cascade structure, but the successful of this solution is due to the good performance of the estimation system and its robustness.

Bibliography

Bakošová, M., Oravec, J., Vasičkaninová, A., Mészáros, A., & Artzová, P. (2019). Advanced control of a biochemical reactor for yeast fermentation. *Chemical Engineering Transactions*, *76*, 769–774. https://doi.org/10.3303/CET1976129

Butkus, M., Levišauskas, D., & Galvanauskas, V. (2021). Simple gain-scheduled control system for dissolved oxygen control in bioreactors. *Processes*, *9*(9). https://doi.org/10.3390/pr9091493

Chitra, M., Pappa, N., & Abraham, A. (2018). Dissolved Oxygen Control of Batch Bioreactor using Model Reference Adaptive Control scheme. *IFAC-PapersOnLine*, *51*(4), 13–18. https://doi.org/10.1016/J.IFACOL.2018.06.008

Dewasme, L., Goffaux, G., Hantson, A. L., & Wouwer, A. vande. (2013). Experimental validation of an Extended Kalman Filter estimating acetate concentration in E. coli cultures. *Journal of Process Control*, *23*(2), 148–157. https://doi.org/10.1016/J.JPROCONT.2012.09.004

Ertunc, S., Akay, B., Boyacioglu, H., & Hapoglu, H. (2009). Self-tuning control of dissolved oxygen concentration in a batch bioreactor. *Food and Bioproducts Processing*, *87*(1), 46–55. https://doi.org/10.1016/j.fbp.2008.04.003

Flores-Hernández, A. A., Reyes-Reyes, J., Astorga-Zaragoza, C. M., Osorio-Gordillo, G. L., & García-Beltrán, C. D. (2018). Temperature control of an alcoholic fermentation process through the Takagi–Sugeno modeling. *Chemical Engineering Research and Design*, *140*, 320–330. https://doi.org/10.1016/j.cherd.2018.10.021

Fonseca, R. R., Schmitz, J. E., Fileti, A. M. F., & da Silva, F. V. (2013). A fuzzy-split range control system applied to a fermentation process. *Bioresource Technology*, *142*, 475–482. https://doi.org/10.1016/j.biortech.2013.05.083

Gnoth, S., Kuprijanov, A., Simutis, R., & Lübbert, A. (2010). Simple adaptive pH control in bioreactors using gain-scheduling methods. *Applied Microbiology and Biotechnology*, *85*(4), 955–964. https://doi.org/10.1007/s00253-009-2114-5

Gomes, J., & Menawat, A. S. (2000). Precise control of dissolved oxygen in bioreactors - a model-based geometric algorithm. *Chemical Engineering Science*, *55*(1), 67–78.

Holzberg, T. R., Watson, V., Brown, S., Andar, A., Ge, X., Kostov, Y., Tolosa, L., & Rao, G. (2018). Sensors for biomanufacturing process development: facilitating the shift from batch to continuous manufacturing. In *Current Opinion in Chemical Engineering* (Vol. 22, pp. 115–127). Elsevier Ltd. https://doi.org/10.1016/j.coche.2018.09.008

Imtiaz, U., Assadzadeh, A., Jamuar, S. S., & Sahu, J. N. (2013). Bioreactor temperature profile controller using inverse neural network (INN) for production of ethanol. *Journal of Process Control*, *23*(5), 731–742. https://doi.org/10.1016/j.jprocont.2013.03.005

Imtiaz, U., Jamuar, S. S., Sahu, J. N., & Ganesan, P. B. (2014). Bioreactor profile control by a nonlinear auto regressive moving average neuro and two degree of freedom PID controllers. *Journal of Process Control*, *24*(11), 1761–1777. https://doi.org/10.1016/j.jprocont.2014.09.012

Jazwinski, A. H. (2007). *Stochastic processes and filtering theory* (Dover Publications, Ed.). Courier Corporation.

Kumar, M., Prasad, D., Giri, B. S., & Singh, R. S. (2019). Temperature control of fermentation bioreactor for ethanol production using IMC-PID controller. *Biotechnology Reports*, *22*. https://doi.org/10.1016/j.btre.2019.e00319

Ławryńczuk, M. (2008). Modelling and nonlinear predictive control of a yeast fermentation biochemical reactor using neural networks. *Chemical Engineering Journal*, *145*(2), 290–307. https://doi.org/10.1016/j.cej.2008.08.005

Mészáros, A., Andrášik, A., Mizsey, P., Fonyó, Z., & Illeová, V. (2004). Computer control of pH and DO in a laboratory fermenter using a neural network technique. *Bioprocess and Biosystems Engineering*, *26*(5), 331–340. https://doi.org/10.1007/s00449-004-0374-0

Nagy, Z. K. (2007). Model based control of a yeast fermentation bioreactor using optimally designed artificial neural networks. *Chemical Engineering Journal*, *127*(1–3), 95–109. https://doi.org/10.1016/j.cej.2006.10.015

Najafpour, G. D. (2006). *Biochemical Engineering and Biotechnology* (Elsevier, Ed.).

Pachauri, N., Rani, A., & Singh, V. (2017). Bioreactor temperature control using modified fractional order IMC-PID for ethanol production. *Chemical Engineering Research and Design*, *122*, 97–112. https://doi.org/10.1016/j.cherd.2017.03.031

Porru, M., & Özkan, L. (2017). Monitoring of Batch Industrial Crystallization with Growth, Nucleation, and Agglomeration. Part 2: Structure Design for State Estimation with Secondary Measurements. *Industrial and Engineering Chemistry Research*, *56*(34), 9578–9592. https://doi.org/10.1021/acs.iecr.7b00243

Salas, S. D., Romagnoli, J. A., Tronci, S., & Baratti, R. (2019). A geometric observer design for a semi-batch free-radical polymerization system. *Computers and Chemical Engineering*, *126*, 391–402. https://doi.org/10.1016/j.compchemeng.2019.04.026

Simutis, R., & Lübbert, A. (2015). Bioreactor control improves bioprocess performance. In *Biotechnology Journal* (Vol. 10, Issue 8, pp. 1115–1130). Wiley-VCH Verlag. https://doi.org/10.1002/biot.201500016

Stanke, M., & Hitzmann, B. (2012). Automatic control of bioprocesses. In *Measurement, Monitoring, Modelling and Control of Bioprocesses* (pp. 35–63).

Chapter 4: A typical biotechnological process: the valorization and exploitation of waste biomass to produce bioethanol

In this chapter, the production of second-generation ethanol from a waste biomass is investigated as a case study. Experimental results related to pre-treatment and fermentation processes will be reported and discussed.

The depletion of fossil resources and the growing economic and demographic pressures are major problems facing humankind. One of the most concerning consequences is the large environmental impact caused by significant greenhouse gas (GHG) emissions. Researchers are working to find sustainable solutions that can help reduce this impact, while also providing a constant supply of energy and goods encouraging a more sustainable economic model that moves away from the established fossil-based one (Guerrero et al., 2013; Nizami et al., 2017). In this scenario, lignocellulosic biomass waste, such as forestry residues, agricultural and food processing waste, represents a promising renewable source that can be exploited to produce energy, biofuels, and bioproducts. This is the idea behind the circular economy, which was proposed by the European Community in order to promote the use of biomass-based feedstocks and to create a sustainable model of economic growth. This approach involves integrating the principles of sustainability, circular economy, and bioeconomy, by converting renewable resources and waste into bioenergy and valuable products. In addition, by optimizing the biomass value in an integrated multi-output production chain (biorefinery), the circular economy aims to create a sustainable and more efficient way of revalorizing apparently low-value materials (Nagarajan et al., 2021; Stegmann et al., 2020). At the same time, this can be helpful to the critical problem of waste disposal management. Indeed, although their organic nature, an improper disposal can lead to serious health and environmental damage (Alatzas et al., 2019; Cho et al., 2020). Food wastes and food industry wastes contribute significantly to this concern. As reported by Uisan et al. (2020), 1.3 billion tonnes are the amount of food wastes (FW) generated annually in the world, corresponding to an equivalent amount of 3.3 billion tons of CO_2, and is expected to increase as a result of continuous population growth and economic development. In recent decades the scientific community has made large efforts in improving waste management or otherwise reducing the accumulation of food waste by recycling it as a raw material to produce high value bioproducts. This is made feasible by the particularly attractive chemical composition. Although they can come from different sources, food industry waste boasts a composition rich in carbon and nitrogen. In particular, they show high level of carbohydrates (33%) such as starch, cellulose, hemicellulose, or lignin, proteins (10%), and lipids (15%) (Uisan et al., 2020), but also organic acids and smaller inorganic part (Cecilia et al., 2019). Food waste is generally exploited to produce high value bioproducts trough microbial fermentation (Gmoser et al., 2019; Mensah & Twumasi, 2017a, 2017b; Nair et al., 2015; Negro et al., 2014; Sadh et al., 2018), for the extraction of antioxidant compounds such as phenolics, vitamins, and carotenoids, but also functional groups such as functional lipids, proteins, and starches (Battistella Lasta et al., 2019; Bengardino et al., 2019; Catalkaya

& Kahveci, 2019; de Andrade Lima et al., 2019; Gopinatha Kurup et al., 2019; Phongthai et al., 2016; Wu et al., 2017). In addition, many works can be found in the literature regarding their use for the production of biogas (Xiong et al., 2019) through the more conventional approach of anaerobic digestion (Atelge et al., 2020; Oh et al., 2018; Pramanik et al., 2019; Xu et al., 2018) or through the recent electro-fermentation (Liu et al., 2019; Shanthi Sravan et al., 2018; Sravan et al., 2018), biodiesel and biofuels (Bušić et al., 2018; Dhiman & Mukherjee, 2021; Li & Yang, 2016), biochar (Novak & Johnson, 2018; Pahla et al., 2018), and immobilized enzymatic bioconversion (Feng et al., 2020; Ladole et al., 2018; Ma et al., 2014; Nadar & Rathod, 2019).

Noteworthy industrial food waste certainly includes brewery's spent grain (BSG), which represents the main by-product of beer production and in particular the 85% of the total by-products as reported by Contreras et al. (2021). Several papers and reviews are of interest to this lignocellulosic matrix (Bachmann et al., 2022; Contreras et al., 2021; Emmanuel et al., 2022; Leite et al., 2019; Mathias et al., 2017; Mitri et al., 2022; Outeiriño et al., 2019), especially concerning its sustainable use as substrate for bioethanol production (Akermann et al., 2022; Barampouti et al., 2022; da Silva et al., 2020; Pinheiro et al., 2019; Rojas-Chamorro et al., 2018, 2020a; Wagner et al., 2022). Indeed, the production of second-generation ethanol represents one of the main targets/goals when considering alternative applications of cellulose-hemicellulose-lignin composed feedstocks like this (Chiaramonti et al., 2012; Contreras et al., 2021). Owing to its availability, low cost, renewable characteristics along with low levels of CO_2 released during combustion, bioethanol is considered an attractive alternative to gasoline and has gained worldwide interest in recent decades (Anwar Saeed et al., 2018; Gupta & Verma, 2015). The high moisture content, which requires the involvement of drying processes, and the need to use chemical compounds to carry out a pre-treatment step in order to modify the complex recalcitrant lignocellulosic structure represent the main difficulties in developing a low-cost bioprocess for BSG valorization (Contreras et al., 2021). In this chapter, a way to exploit the brewer's spent grain as a renewable waste biomass will be discussed. In particular, its origin, composition and a general outline of all its possible applications are reported. In more detail, my experimental work carried out during the abroad research period will be presented, which is divided into a first phase of pre-treatment and a second phase of fermentation. Supported by the various examples available in the literature, the chemical approach, and in particular dilute acid hydrolysis, was chosen as the most adopted pre-treatment to increase the digestibility of the starting matrix as reported in the literature (Saravanan et al., 2022). For both steps, a design of experiments (DOE) was carried out, which

allowed to optimize the operating conditions under which to conduct the process. Experimental results for both steps will be shown and analyzed.

4.1 Overview of general characteristics/properties and possible applications of BSG

It is well known that beer is one of the oldest and most widely consumed beverages in the world, in particular is the third after water and tea as reported by Mitri et al. (2022). The brewing process represents one of the most economically significant since its world production has reached 1.94 billion hectolitres, while in Europe it is around 531 million hectolitres (*Barthhaas_report_2020_en,* n.d.). Although this work will not report in detail on the steps involved in brewing beer, for the purposes of the discussion it is enough to know that BSG (Figure 30) is derived from the wort preparation, after its filtration (Mussatto et al., 2006). As reported by Ikram et al. (2017), the amount of BSG produced annually by the brewing industry in the European Union is about 3.4 million tons, while global production reaches 39 million tons (Birsan et al., 2019). 100 kg of malt gives approximately 100-130 kg of fresh BSG, with a moisture content of 70-80%, equivalent to around 20 kg of BSG per 100 L of brewed beer. As reported by Mussatto (2014), this insoluble part generally made up of the husks of the barley malt grain along with the pericarp and the seed coat layer. However, depending on the type of the beer, it is possible that in addition to the barley malt grain there may be other cereal additions such as corn, rice, wheat, oats, rye, or sorghum. The next section will discuss and analyse in detail its chemical composition and possible applications of this valuable raw material.

Figure 30. Fresh brewer's spent grain (BSG).

Table 14. Chemical composition of brewer's spent grain (BSG) on a dry weight basis.

Components [g kg^{-1}]	(Mussatto & Roberto, 2006)	(Xiros et al., 2008)	(Waters et al., 2012)	(Meneses et al., 2013)
Cellulose	167.8	120	260	217.3
Hemicellulose	284.2	402	222	192.7
Xylan	199.4	NR	NR	136.3
Arabinan	84.8	NR	NR	56.4
Lignin	277.8	115	NR	194.0
Proteins	152.5	142	221.3	246.9
Ashes	46.0	33	11	41.8

4.1.1 Chemical composition of BSG

Brewery's spent grain consists of a heterogeneous mixture of cereal grain husks, pericarp, residual amounts of endosperm depending on the brewing regime applied, and the original barley grain coating layer (Steiner et al., 2015a). Table 14 summarises the chemical composition of BSG. Independently of the variations that can be observed in its composition in terms of the individual component concentration, this residue is predominantly composed of fiber (cellulose, hemicellulose and lignin) since barley malt husk is a lignocellulosic material. Hemicellulose and cellulose are fractions composed of sugars, of which xylose, arabinose and glucose are the most abundant mono-glycans as stated by Mussatto & Roberto (2006), comprising around 50% (w/w) of the BSG composition. Lignin and proteins are also significant constituents of BSG. In particular, lignin is a polyphenolic macromolecule responsible for maintaining the structural integrity and rigidity of the cell wall. Moreover, it contains several phenolic compounds, the most important of which are ferulic, p-coumaric, syringic, vanillic and p-hydroxybenzoic (Mussatto et al., 2007). Protein generally comprises 15% to 25% of the composition of BSG. The study conducted by Santos et al. (2003) showed that the protein fraction in oven dried BSG can be 15% to 24.2% and includes globulins, albumins, glutelins and hordeins. BSG also contains significant amounts of minerals, the most abundant being calcium, followed by magnesium, phosphorus and sodium (Meneses et al., 2013). Still, it is possible to find also iron, copper, potassium and manganese (Mussatto, 2009). Vitamins are also present in noteworthy quantities, in particular biotin, niacin, folic acid, choline, thiamine, pantothenic acid, riboflavin and pyridoxine (Chetrariu & Dabija, 2020; Ikram et al., 2017). Finally, there are extractives formed by waxes, resins, tannins, essential oils, and lipids, among which triglycerides and fatty acids are the most abundant (Chetrariu & Dabija, 2020; Ikram et al., 2017; Mussatto, 2009).

4.1.2 Potential applications of BSG

Being a low-cost raw material, many efforts have been made to investigate BSG possible application in the field of animal and human nutrition, as feedstock for biofuels production, and in biotechnological processes through the extraction of value-added compounds. The most relevant studies in the literature are presented and discussed in the following paragraphs.

BSG in animal feed

Due to BSG low cost, availability, high protein and fiber content, spent grain has always been mainly used for different animals (cattle, fish, chickens, pigs and ruminants) feed, either in wet

or dried form (Mussatto, 2014), creating an output for this material and solving the problem of its disposal (Aliyu & Bala, 2011). It has been demonstrated that the combination of BSG with nitrogen sources such as urea can provide ideal nutrition for ruminants with all essential amino acids (Huige, 2006). As reported by Tang et al. (2009), the consumption of BSG or its derivatives by rats is responsible for beneficial effects on intestinal digestion, which are related to the glutamine-rich protein content, to a high content of non-cellulosic polysaccharides and small amounts of β-glucans. Other important benefits have been observed in other animals such as fish. A study conducted by Kaur & Saxena (2004) revealed that the substitution of rice with spent grain resulted in greater body weight gain attributed to the high-quality protein and essential amino acid content of a BSG-based diet. Aliyu & Bala (2011) examined the effect of BSG on the nutritional value of milk yield in dairy cattle observing that, compared to a maise-based diet (45% w/w), an integration of BSG (45% w/w) leads to an increase in current milk yield, milk total solids content and milk fat yield. Although the major nutritional benefits of BSG consumption have been observed especially in dairy cattle, several efforts have been made to exploit all or part of this raw material in mixed forms and in different animals. In fact, as reported in Mussatto (2014), it is used as part of the poultry diet. However, most of the polysaccharides in the BSG cell walls, such as arabinoxylan and β-glucans cannot be digested by these animals. This problem is due to the lack of enzymes required for hydrolysis of the polymer chains and has been solved simply by adding the missing enzymes (xylanase and β-glucanase) to the feed.

BSG in human diet

Several beneficial effects have been observed in the use of BSG or its components in the human diet, including accelerated transit time, increased faecal volume, reduced cholesterol and postprandial glucose levels (Mussatto, 2014). Further positive effects are associated with a significant reduction in contracting type II diabetes, obesity, diverticulitis, cardiovascular problems, or colorectal cancer (Steiner et al., 2015). This is due to the presence of biologically active compounds such as β-glucans which play an important physiological role in the body (Chetrariu & Dabija, 2020). This has led to increased interest in developing BSG fortified foods, which are considered functional food offering health benefits when combined with a balanced diet. The incorporation of spent grain in the manufacturing of bakery products such as bread, biscuits and snacks, was also evaluated in order to increase the fiber content. To make this possible, BSG is usually converted into flour as its grainy nature may alter the final product's physical properties. In this way, the particle size is reduced, which is important to

make the final product acceptable to the consumer (Mussatto, 2014). BSG incorporation in food products was analysed by Lynch et al. (2016) at different levels (10-40% on a dry weight). The main effects observed were an increase in protein, fiber and amino acids content and a decrease in the starch level and calories. However, it is important to consider the possible sensory alterations that may occur. In fact, since it has a brown colour, it is recommended to use BSG only in colored products (bread, biscuits, etc.) to prevent colour alteration of the product. For a BSG addition higher than 20%, the colour is not the only property to be affected, but also the product structure, the volume, and the texture. According to Steiner et al. (2015), these problems can be prevented by replacing only 10% of normal flour with BSG. Besides bakery products, Özvural et al. (2009) investigated the production of Frankfurters sausages using BSG as a fat substitute to produce a meat product with high fiber and low-fat content. Other than a direct application of BSG as a substitute for other ingredients in food preparation, the extraction of phenolic compounds as additives represents a feasible valorisation route. As suggested by McCarthy et al. (2013), the addition of BSG extracts to food drinks such as fruit juices and smoothies can significantly improve the total phenolic content in the case of important differences with the originals. Further studies were conducted on the use of polyphenols and flavonoids extracted from BSG for the preparation of fish burgers using various bioactive powders (Spinelli, Conte, & del Nobile, 2016), and on the possibility of strengthening food products such as pasta and infant formulas using the extracted proteins (Nazzaro et al., 2018).

BSG for ethanol production

Biofuels obtained from the conversion of lignocellulosic materials are a viable alternative to fossil fuels since they are considered carbon neutral. Cellulosic bioethanol has received considerable global attention as a transport fuel for several benefits, among which its potential to reduce greenhouse gas emissions by 86% certainly stands out (M. Wang et al., 2007). Several studies were performed to obtain ethanol from the biological conversion of SBG using different pre-treatment techniques and fermentation conditions, leading to significantly different results in terms of ethanol yield. White et al. (2008) compared the results obtained from acid pre-treatment with HNO3, H2SO4 and HCl at different concentrations and found the former to be the most suitable for performing this step. The hydrolysate pre-treated with 0.16 N HNO3 and subjected to enzymatic hydrolysis, was subsequently fermented for 48 h at 30°C using two different yeast strains, *Pichia stipitis* NCYC 1540 and *Kluyveromyces marxianus* NCYC 1425, obtaining respectively 8.3 and 5.9 g L-1 ethanol, which correspond to 0.32 and 0.23 g ethanol

(g substrate-1) and 63% and 45% in terms of theoretical conversion. Liguori et al. (2015) obtained a higher concentration of ethanol, combining an acid pre-treatment with H_2SO_4 with an alkaline treatment. The fermentation step was performed using the strain *Saccharomyces cerevisiae* NRRL YB 2293 and a comparison between the ethanol yield obtained using the BSG hydrolysate as a growth substrate with and without a supplement of yeast extract was shown. In the first case, an ethanol concentration of 12 gL-1 corresponding to 0.26 g ethanol (g substrate^{-1}) was obtained, compared to 12.79 gL^{-1} ethanol and 0.28 g ethanol (g substrate^{-1}) when the hydrolysate was enriched with nutrients. Therefore, the authors demonstrated that in both cases, similar concentrations of ethanol were achieved. The only difference is that enriching the medium with a nitrogen source improves performance in reducing fermentation time. Mata et al. (2015) optimised the acid and enzyme BSG pre-treatments carried out sequentially by evaluating different combinations of reaction times and enzyme/BSG ratio values but keeping the acid amount and concentration constant. The maximum total sugar conversion obtained was 22.24%. Fermentation was then conducted at 30°C for 72 h using both a synthetic growth medium and BSG hydrolysates in order to evaluate the potential presence of inhibitors from the previous pre-treatment steps. Two different strains were used in this case as well, *Pichia stipitis* NCYC 1541 and *Kluyveromyces marxianus* NCYC 2791. In both cases the results showed that the fermentation efficiency is much higher in the case of synthetic medium (around 80%), and only 45.10% for *P.stipitis* and 36.58% for *K.marxianus*. The theoretical ethanol yield was 0.27 and 0.19 g ethanol (g substrate-1) respectively for *P.stipitis* and *K.marxianus*, while the actual was 0.0856 and 0.0308 g ethanol (g substrate-1). The low values of these yields compared to those obtained by White et al. (2008), were justified as a consequence of the presence of fermentation inhibitors form the pre-treatment steps.
Overall, BSG represents a valuable raw material with potential for ethanol production. However, research efforts are still necessary to optimise the process conditions, develop green technologies for the pre-treatment steps, develop cheap enzymatic routes, and invest in genetically modified strains to improve the sugar-ethanol conversion.

Extraction of value-added compounds from BSG

As already mentioned, BSG represents an important and inexpensive source of value-added components such as carbohydrates, proteins, and lipids but it contains also phenolic compounds that have recently gained considerable interest for their valuable health-benefiting properties. In particular, due to their role in the prevention of chronic disorders, cancer and intracellular oxidative stress their recovery is attractive for manufacturers and scientists (Ikram et al., 2020).

Different levels of phenolic compounds have been found in BSG, but usually the hydroxybenzoic acids (HBA) content of BSG is lower than hydroxycinnamic acids (HCA). The most abundant among HBA is the syringic acid but their quantity can vary depending on the barley variety, the harvesting time, and the characteristics of the growing region. Regarding the HCA, those present in higher quantities are ferulic acid (FA) and p-coumaric acid (p-CA) (Mussatto et al., 2007), which are both contained in the outer layers of grain and remain in BSG after the entire brewery's process. The recovery process of phenolic compounds can be summarised in the following steps: pre-treatment, extraction, isolation, and purification. The pre-treatment can include maceration, homogenisation, grinding and milling leading to the disruption of the cellular structure improving the recovery of bioactive compounds and increasing the mass transfer between the solvent and the biomass. In addition to the pre-treatment techniques already mentioned, other methods such as autohydrolysis, acid, alkaline, and enzymatic pre-treatments can be useful in the breakdown of the biomass structure and access to the cell vacuole where the phenolics are contained. Among the different extraction techniques used to recover phenolic compounds, a differentiation can be made between conventional and non-conventional methods. The solid-liquid extraction (SLE) together with Soxhlet (SE) are the most common and well-established techniques used due to their simplicity, efficiency and wide industrial applicability. The former is usually combined with alkaline hydrolysis to increase extraction efficiency by degrading lignin and cellulose and releasing unbound HCA (FA and p-CA). Although this technique ensures a good recovery of phenolics (Meneses et al., 2013), there are drawbacks such as the important amounts of solvent required, the long extraction time but also the possible need for additional procedures to remove unwanted non-phenolic compounds thus increasing the process cost, which led to consider alternative methods. Alternative extraction methods, also called green techniques because they follow the standards set by the Environmental Protection Agency (USA), have recently been proposed. Supercritical fluid extraction and pressurised fluid extraction belong to this category. In recent decades, they have caught the interest of researchers due to the reduced extraction time, reduced solvent use, higher extraction yield and enhanced quality of extracts. Ultrasound and microwave-assisted extraction can be considered in the same way if combined with a conventional method using inorganic solvents or with pressure liquid extraction. The advantage of their use is the obtainment of cleaner extracts without presence of residues, which would occur when extracting with organic solvents. Although there are still few works involving their application on BSG, a significant progress has been made (Herbst et al., 2021; Spinelli, Conte,

Lecce, et al., 2016; X. J. Wang et al., 2013) and, they are very encouraging, offering an interesting starting point for further research into possible a possible scale-up.

Energy production from BSG

As an agri-food biomass, BSG represents a valid raw material that can be used effectively in a waste-to-energy process due to its ability to produce energy. This approach includes technologies such as aerobic digestion and thermochemical conversion processes (combustion, pyrolysis and gasification). Important requirements for a valid energy valorisation are calorific values over 15 MJ kg-1 and humidity in the range 10-15 % (Gil-Castell et al., 2022). Considering the BSG composition, a drying phase is certainly necessary in order to reach the desired moisture value. Following the circular economy principle, it is worth noting the possibility of reusing the spent grain energy content within the beer industry, allowing a reduction in energy consumption. However, mature strategies are not yet available and further studies in this direction are required.

4.2 Second generation ethanol production from BSG

The aim of the work presented in this section was to evaluate the possibility of exploiting BSG as biomass for the production of second-generation ethanol. The bioconversion to ethanol was carried out by considering an initial step of acid hydrolysis, followed by fermentation of the released sugars performed by *Saccharomyces cerevisiae* yeasts. These two steps were optimized within a specific range of operative conditions by designs of experiments. In particular, a central two-factor composite design was applied for acid pre-treatment, where the sulfuric (0.065-0.37 M) and nitric (0.01-0.5 M) acid concentration, along with the liquid-solid ratio (8-12 w/w %) were selected as independent factors. The fermentation process was also optimized by using the Behnken-box design and considering temperature (25-37°C), inoculum volume (5-15 v/v %), and pH (4.5-6.5) as investigated factors.

4.2.1 Materials and Methods

Chemicals and reagents

Sulfuric (98%) and nitric (65%) acids used for acid hydrolysis were purchased from Merck (Darmstandt, Germany). 1 M NaOH (Reagecon, Co.Clare, Ireland) was used for the neutralization of hydrolysates. For the preparation of yeast growth inoculum, yeast extract from brand, glucose from brand, salts and anti-foam purchased from Sigma-Aldrich were used.

Standard solution for HPLC analysis were prepared using 70% ethanol (Merck, Darmstadt, Germany), glucose (brand), L-(+)-Arabinose (99%, Sigma-Aldrich) and D-(+)-Xylose (99%, Sigma-Aldrich).

BSG, yeast, and inoculum

The BSG used in this work was kindly donated by the Vestfyen brewery (Assens, Denmark). After collection it was stored at -20°C, fresh BSG was oven dried at 60°C for 24 h in a humidity-controlled oven (Memmert HCP 108). To ensure that no further evaporation occurred, BSG samples were weighed three separate times during the end of the 24 h. Before weighing, BSG was placed in a dessicator with silica gel to reach ambient temperature. Weighing was done in 1h intervals and the time spent in the dessicator was not included in the 24h drying time. A commercial strain of Saccharomyces cerevisiae (Malteserkors, Denmark) was grown overnight at 30°C in an orbital shaker at 150 rpm using a growth medium containing the following compounds: 20 gL^{-1} glucose, 6 gL^{-1} yeast extract, 0.23 gL^{-1} $CaCl_2 \cdot 2H_2O$, 4 gL^{-1} $(NH_4)_2SO_4$, 1 gL^{-1} $MgSO_4$, 1.5 gL^{-1} KH_2PO_4 (Sivakesava et al., 2001). The medium was stored at 4°C until use.

Dilute acid hydrolysis

The experiments were carried out by autoclaving (SHP Laboklay ECO 135M, SHP Steriltechnik AG, Germany) the dried BSG with 200 mL of acid solutions in 250 mL flasks, at 120°C for 20 min. The liquid phase was separated from the solid part by centrifugation at 4500 rpm and 25°C for 20 min (Avanti J-HC, Beckman Coulter, United States), and then vacuum filtration (brand), and characterized in terms of sugar compositions.

Fermentation of BSG

In order to evaluate the time required by the yeast to consume the substrate, preliminary fermentation experiments were conducted in duplicate in 1 L fermentation flasks working in anaerobic conditions at 30°C in a orbital shaker (IKA KS 4000 i control, Germany) at 150 rpm for 24 h. The working volume of acid hydrolysate for both flasks was approximately 170 mL. Before being inoculated, the fermentation medium was neutralized using 0.1 M NaOH until a pH of 5.5 was reached, then 10% (v/v) of inoculum was added. The run was conducted for 24 h, during which samples were taken at different times, filtered, and stored for sugars and ethanol determination. Subsequent experiments for process optimization were performed in 500 mL fermentation flasks using approximately 200 mL of hydrolysate. Each of them was

carried out for 9 h under different conditions in terms of temperature, inoculum volume, and pH.

Sugars and ethanol quantification

Liquid samples from acid hydrolysis and fermentation steps were centrifuged and filtered through 0.20 µm membranes (Sartorius, Germany), and then analyzed by HPLC. The HPLC system (Ultimate 3000, Thermo Fisher) was equipped with a refractive index detector (Dionex Softrom GmbHm Germany), and a Phenomenex Rezex RHM-Monosaccharide H+ (8%) analysis column working at 79°C with ultrapure water as mobile phase (0.6 mL/min). Glucose, xylose, arabinose, and ethanol were identified using standard solutions and quantified through calibration curves. The limits of the curves were [0.1-50 g/L] for glucose, [1-10 g/L] for xylose, [1-10 g/L] for arabinose, and [0.15-81 g/L] for ethanol, with a $R^2_{Glucose}$=0.9993, R^2_{Xylose}=0.9917, $R^2_{Arabinose}$=0.9997, and $R^2_{Ethanol}$=0.9995.

4.2.2 Design of experiments

The use of waste biomass for the obtainment of valuable products is strongly dependent on the raw material. For this reason, it is crucial to properly design the experimental campaign in such a way that the description and understanding of the variation in the data is simplified and it is statistically evident the impact of each considered factors. Different procedures for an optimal design of experiment (DoE) have been used for the hydrolysis and fermentation of the BSG and they are described in the following.

Acid hydrolysis

In order to determine the optimal conditions at which to carry out the acid pre-treatment, and experimental design based on Central Composite Design (CCD) methodology was applied. The acid concentration and liquid-solid ratio were selected as design factors, while the temperature and process time were kept constant. In particular, the following ranges were investigated: [8-12 w/w %] for the L-S ratio, [0.065-0.37 M] and [0.01-0.5 M] as the concentration range for H2SO4 and HNO3, respectively. Table (number) shows the design of experiments obtained as a function of the two independent variables investigated. One noteworthy aspect is that the axial conditions were removed from the experimental plan as well as the combination between the mildest treatment in terms of acid concentration and the highest L-S ratio. The reason for this choice is related to an expected low glucose yield under such conditions. Therefore, only 8 experimental conditions were evaluated which is equivalent to

16 runs since each condition was carried out in duplicate. The efficiency of the pre-treatment performed with the two acids was assessed by determining the glucose yield defined as g of glucose per 100 g of dry BSG present in the liquors.

Table 15. DOE for sulfuric and nitric acid pre-treatment.

	H_2SO_4				**HNO_3**			
	C_{acid} [M]		L-S ratio [w/w]		C_{acid} [M]		L-S ratio	
run	real	coded	real	coded	real	coded	real	coded
1	0.065	-1	12	-1	0.01	-1	12	-1
2	0.37	1	12	-1	0.5	1	12	-1
3	0.37	1	8	1	0.5	1	8	1
4	0.18	0	10	0	0.1	0	10	0

Fermentation of the hydrolysate

The Box-Benken experimental design was applied to identify the conditions under which to conduct the fermentation process. Temperature (25, 30, 37 °C), inoculum volume (5, 10, 15% v/v) and pH (4.5, 5.5, 6.5) are the three key process variables chosen as factors for the experimental design. A total of 14 experimental combinations were obtained, including two points in the center of the experimental domain. The uncoded and coded values of these factors are shown in Table 6 (number).

4.2.3 Results

Acid hydrolysis

The first experimental campaign was aimed at investigating the effect of acid concentration and liquid-solid ratio on the recovery of sugars from BSG. Figure 31 shows the hydrolysates for each experimental run set by the DoE (Table 15) performed in duplicate. The efficiency of acid pre-treatment was evaluated by considering the glucose release via the hydrolysis of cellulose and hemicellulose and the glucose yield has been evaluated as grams of glucose on dry biomass basis ($g_{glucose}/g_{dry\ BSG}$). Glucose recovery mean values and standard deviations of each experimental run conducted in duplicate are reported in Tables 7 and 8 when used, respectively, H_2SO_4 and HNO_3. This comparative analysis allows the identification of the acid that leads to the highest amount of glucose. It is worth noting that fractions of other monomeric

sugars such as xylose and arabinose are contained in the biomass, but glucose yield was chosen as process response because the Saccharomyces Cerevisiae's strain used for the fermentation step is highly selective toward such sugar.

Table 16. Experimental conditions for the fermentation process of BSG.

	Temperature (°C)		**Inoculum volume (% v/v)**		**pH**	
run	real	coded	real	coded	real	coded
1	25	-1	5	-1	5.5	0
2	25	-1	15	1	5.5	0
3	37	1	5	-1	5.5	0
4	37	1	15	1	5.5	0
5	25	-1	10	0	4.5	-1
6	25	-1	10	0	6.5	1
7	37	1	10	0	4.5	-1
8	37	1	10	0	6.5	1
9	30	0	5	-1	4.5	-1
10	30	0	5	-1	6.5	1
11	30	0	15	1	4.5	-1
12	30	0	15	1	6.5	1
13	30	0	10	0	5.5	0
14	30	0	10	0	5.5	0

Figure 31. Acid hydrolysis samples for different DOE runs.

Table 17. Results of sulfuric acid pre-treatment in terms of glucose yield and its standard deviation.

		H_2SO_4	
run	C_{acid} [M]	L-S ratio [g/g]	GR [$g_{glucose}$/100 g dry BSG]
1	0.065	12	10.10 ± 0.19
2	0.37	12	15.98 ± 0.076
3	0.37	8	14.71 ± 0.322
4	0.18	10	15.14 ± 0.37

Table 18. Results of nitric acid pre-treatment in terms of glucose yield and its standard deviation.

		HNO_3	
run	C_{acid} [M]	L-S ratio [g/g]	GY [$g_{glucose}$/100 g dry BSG]
1	0.01	12	2.53 ± 0.04
2	0.5	12	15.19 ± 0.26
3	0.5	12	11.14± 0.18
4	0.1	10	12.04 ± 0.015

Surface response methodology has been applied to mathematically describe the effects of process conditions on the glucose recovery and to identify the optimum condition of acid pre-treatment. From the experimental data obtained for H_2SO_4 (Table 17) and HNO_3 (Table 18), the following second-order regression models which describe the interaction between the factors have been derived (Eqs. (77) and (78)):

$$GY_{H_2SO_4} = -50.91 + 12.98\,LS + 15.103\,C_{H_2SO_4} - 0.6639\,LS^2 \quad (77)$$

$$GY_{HNO_3} = -127.23 + 28.517\,LS + 17.565\,C_{HNO_3} - 1.4766\,LS^2 \quad (78)$$

where LS is the liquid-solid ratio expressed as gram of liquid per gram of BSG, $C_{H_2SO_4}$ and C_{HNO_3} are the concentration of sulfuric and nitric acid, and GY is the glucose yield expressed as gram of glucose recovered per 100 g of BSG. Figures 32 and 33 show the resulting surface plot for H_2SO_4 (Fig. 32) and HNO_3 (Fig. 33), compared with the experimental results (blue dots). It is possible to observe that both regression models well describe the experimental data obtained and in both cases the glucose yield exhibits a positive linear trend with the acid concentration and a quadratic relationship with L-S ratio, as it can be also observed from Eqs. (77) and (78). The models predicted the optimal pre-treatment conditions for BSG in terms of concentration and liquid solid ratio at 0.37 M and 10 g/g for sulfuric acid, and 0.5 M and 10 g/g for nitric acid, resulting in an estimated optimal glucose recovery of 18.12 g/g in the former and 19.24 g/g in the latter case. Although the highest glucose recovery can be achieved by conducting acid hydrolysis with HNO_3, this also represents the condition under which a higher acid concentration is required. This would mean a greater possibility of formation of inhibitory compounds such as furfural and hydroxymethylfurfural (HMF), which result from the decomposition of pentose and hexose sugars, and that could affect the fermentation reaction (Rojas-Chamorro et al., 2020b). In addition, analyzing the monomeric sugars released during this preliminary step, it was found that sulfuric acid shows greater selectivity to glucose than nitric acid. This is a relevant detail since the yeast used to ferment the hydrolysates prefers glucose over other monomeric sugars in solution.

Figure 34 shows a comparison between the surface plot of the two acids, and it can be easily observed that the surface of H_2SO_4 is above that of HNO_3 in its concentration range (0.065 M-0.37 M). For the mentioned reasons, the experimental condition involving 0.37 M as H_2SO_4 concentration and 10 % (g/g) as liquid-solid ratio was selected as the optimal one for carrying out the acid hydrolysis pre-treatment on the BSG. According to these conditions, the identified model was validated in order to compare the predicted response with the experimental one. The

experimental recovery value was $YG_{H_2SO_4}^{experimental} = 18.3 \pm 0.1\ g/g$, that is quite close to the glucose recovery calculated with the model reported in Eq. (78), $YG_{H_2SO_4}^{predicted} = 18.1\ g/g$, indicating excellent performance of the model within the investigated conditions.

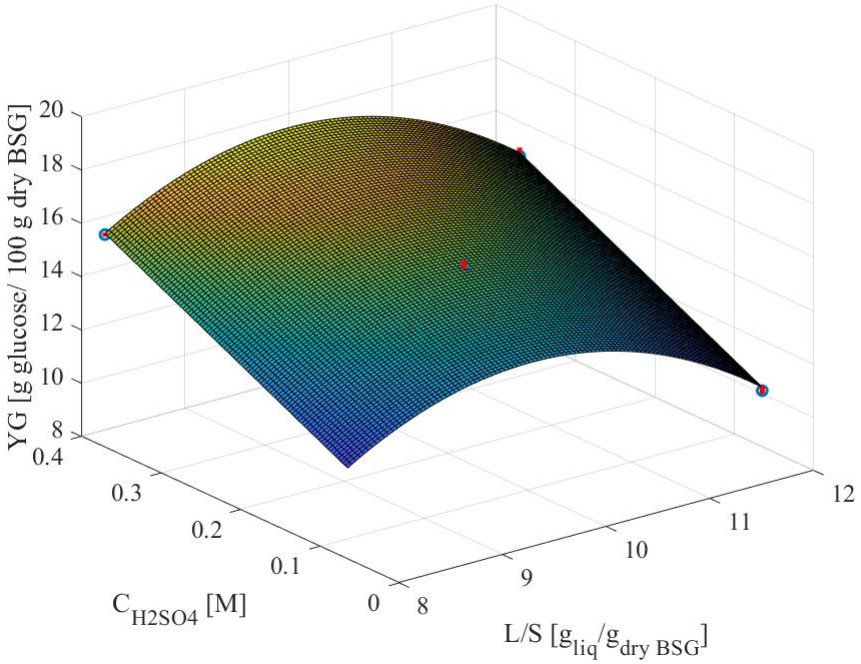

Figure 32. Glucose yield as a function of H_2SO_4 concentration and liquid-solid ratio. The blue dots represent the experimental conditions, and the red bars the corresponding standard deviations.

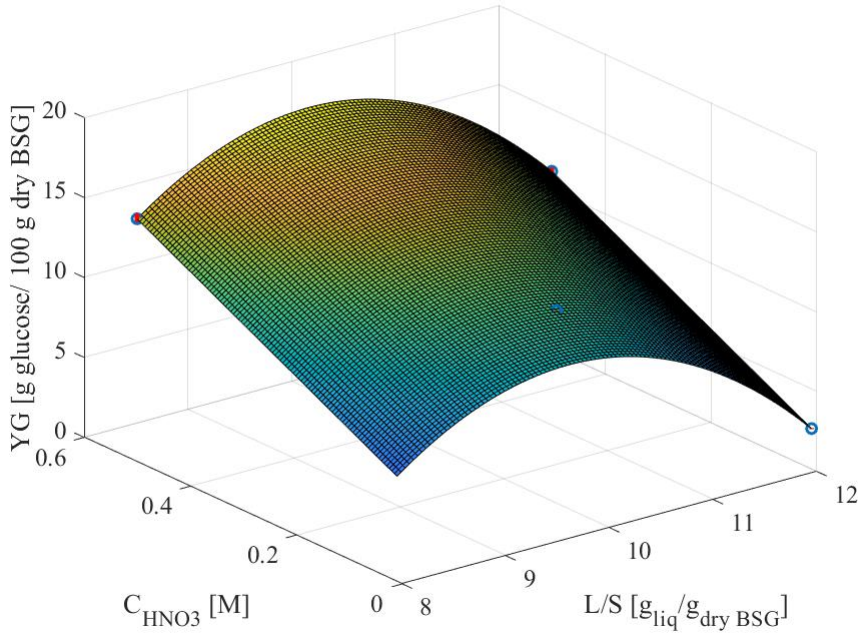

Figure 33. Glucose yield as a function of HNO_3 concentration and liquid-solid ratio. The blue dots represent the experimental conditions, and the red bars the corresponding standard deviations.

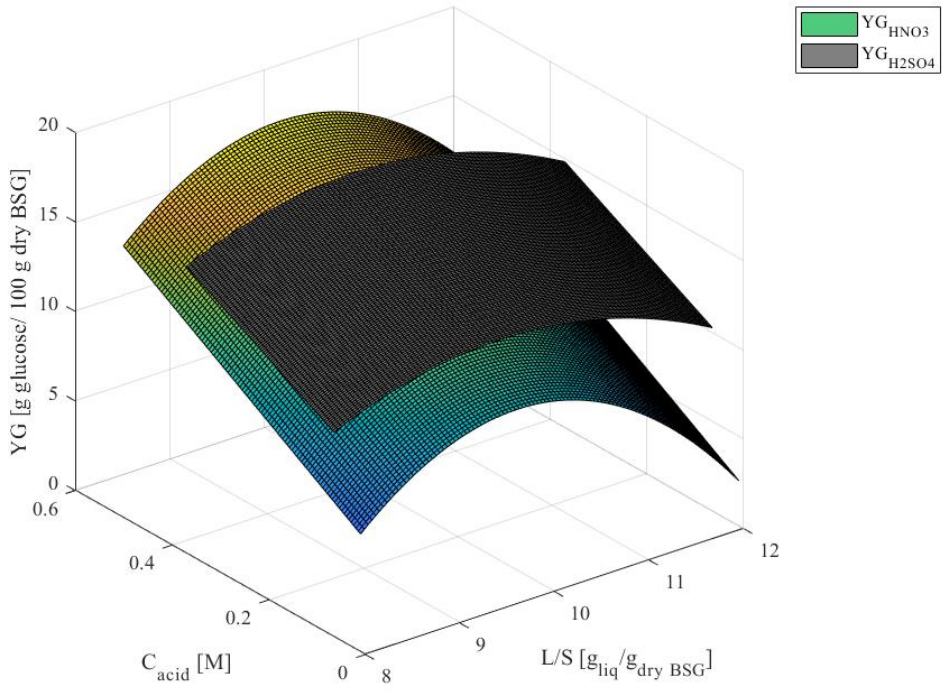

Figure 34. Comparison of the surface plot for H_2SO_4 and HNO_3.

Fermentation of optimized hydrolysate

Two different types of experiments were carried out for studying the fermentation of the hydrolysis product. The first set of experiments was aimed at evaluating the time required for obtaining the maximum ethanol production from the fermentable sugars, that means that it corresponds to monitor the process until steady state is detected for ethanol concentration. The profile of sugars and ethanol concentration during the fermentation is reported in Figures 35-38, using logarithmic scale for the time coordinate in order to facilitate the visualization of the changes before and after glucose depletion. It is worth noting that both fermentation runs had been conducted at the same conditions, but the amount of pentoses at the beginning of the experiments is slightly different. This depends on the previous hydrolysis step which did not lead to the same production of arabinose and xylose. This aspect was not further investigated because the yeast used in this work is highly selective toward glucose. Indeed, the yeast completely consumed glucose after almost 4 h (Figure 34), whereas arabinose and xylose were slightly consumed, and the consumption rate increased only after the total depletion of glucose. The evidence that ethanol is obtained also from pentoses is given by the increase in ethanol concentration after 9 h, that is almost 13% in run 1 and circa 10% in run 2. Furthermore, the final ethanol concentration is higher when starting from a reactant mixture containing a higher amount of pentoses. It is worth noting that arabinose and xylose had not been completely transformed to ethanol, because the amount theoretically obtainable from the consumption of pentoses is higher than the quantity produced in the time interval 9 – 24 hours. Anyway, the production of ethanol from xylose and arabinose is quite scarce in presence of Saccharomyces Cerevisiae's strain, confirming previous results (Rojas-Chamorro et al., 2020b). For this reason, experiments aimed at finding the best condition for glucose fermentation was chosen equal to 9 h, because sufficient for a complete depletion of glucose consumption by *S. Cerevisiae* yeasts.

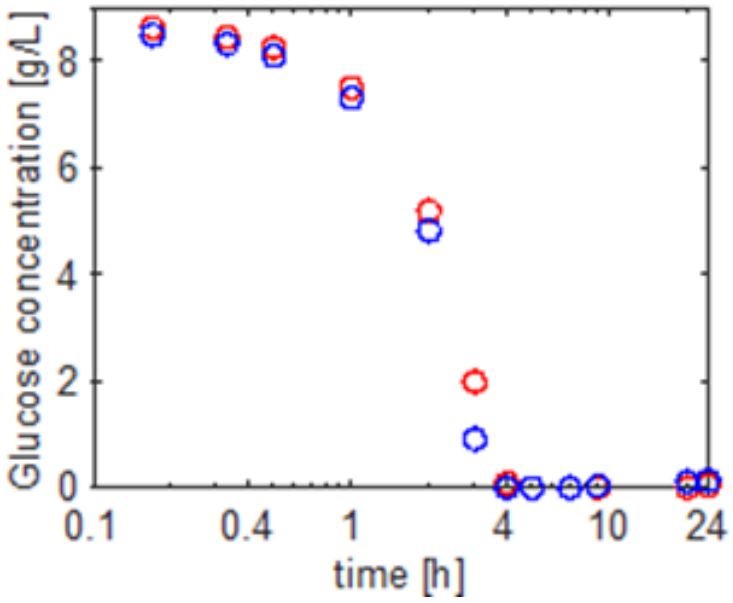

Figure 35. Glucose consumption during the fermentation process for run 1 (red circle) and run 2 (blue circle). Time is in logarithmic scale.

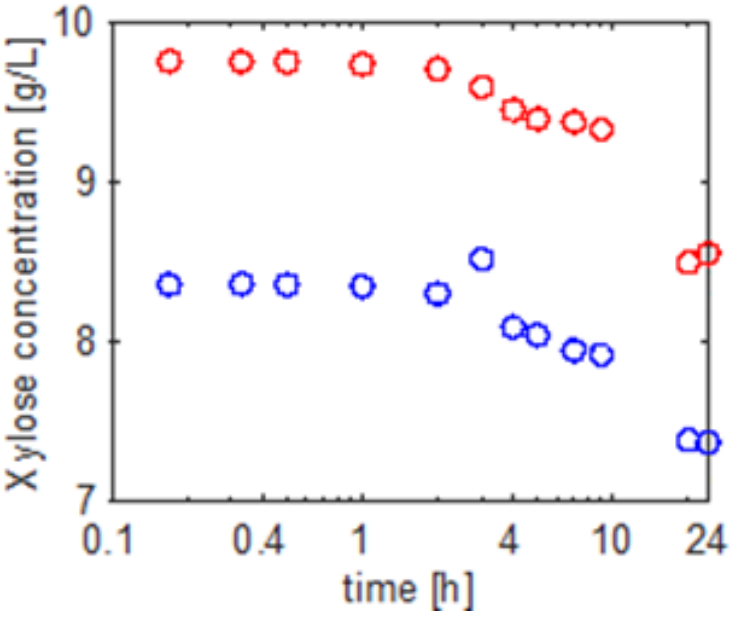

Figure 36. Xylose consumption during the fermentation process for run 1 (red circle) and run 2 (blue circle). Time is in logarithmic scale.

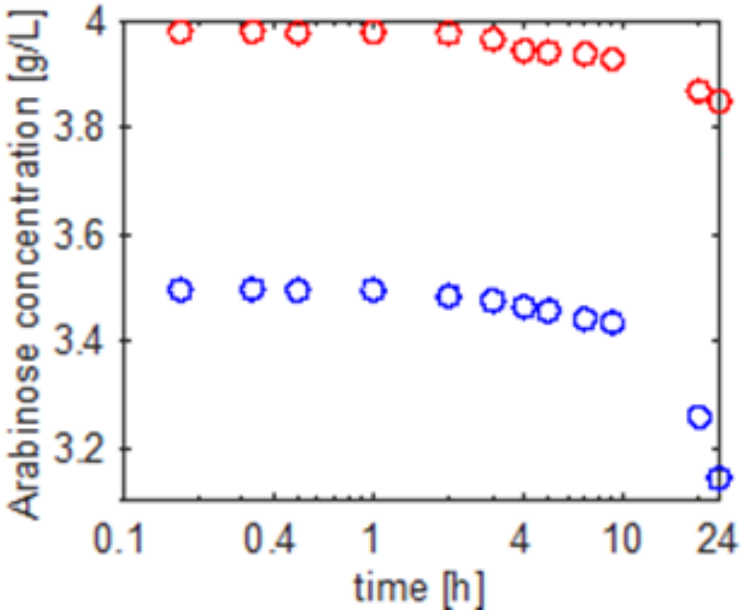

Figure 37. Arabinose consumption during the fermentation process for run 1 (red circle) and run 2 (blue circle). Time is in logarithmic scale.

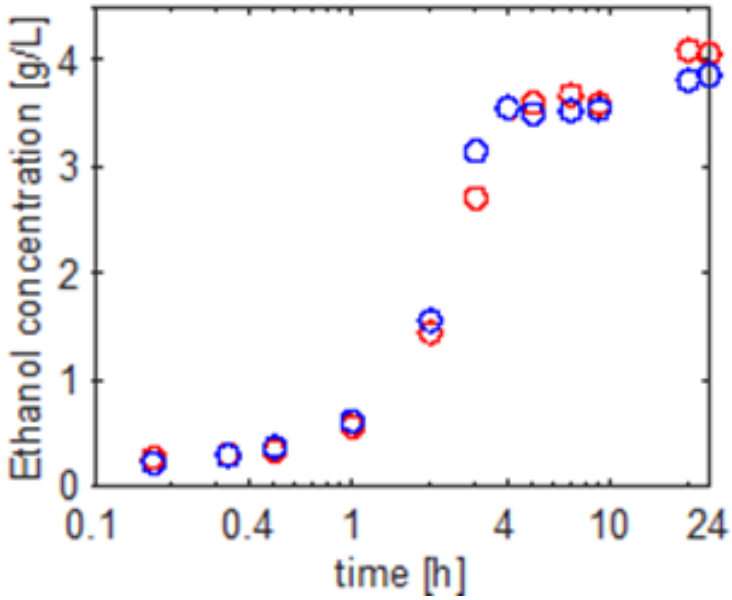

Figure 38. Ethanol production during the fermentation process for run 1 (red circle) and run 2 (blue circle). Time is in logarithmic scale

Optimal condition for ethanol production

Figures 39 and 40 show the hydrolysate samples for the different experimental runs provided by the DoE (Table 16). Table 19 reports the results of the experimental campaign for the fermentation in terms of the calculated ethanol yield (Y^*_{EtOH}) with respect to the theoretical one ($Y_{EtOH,theoretical}$ = 51 %) obtained varying the independent factors: temperature (T), inoculum volume (iv) and pH. A response surface relating the yield to the factors have been estimated

by exploiting a backward elimination procedure with a significance level α=0.05. The final second-order regression model is reported in Eq.79 and exhibits a determination coefficient R^2=88.3% and R^2_{pred}=65.6%. Model adequacy is confirmed in Figure 41, which shows that experimental measurements and computational predictions are in excellent agreement.

Table 19. Results of DoE for the fermentation step.

Run	T [°C]	iv [% v/v]	pH	Y^*_{EtOH} [%]
1	37	15	5.5	78
2	30	10	5.5	83
3	30	10	5.5	79
4	37	10	4.5	78
5	30	15	4.5	80
6	30	15	6.5	79
7	37	5	5.5	68
8	30	5	4.5	71
9	25	15	5.5	80
10	30	5	6.5	66
11	37	10	6.5	81
12	25	10	4.5	86
13	25	10	6.5	73
14	25	5	5.5	59

$$Y^*_{EtOH} = 152.7 - 3.42\,T + 7.21\,iv - 21.64\,pH - 0.2942\,iv^2 + 0.644\,T \cdot pH \quad (79)$$

As can be deduced from the coefficients of Eq. (79), temperature and pH exerts a negative influence on ethanol yield even though these influences are slightly attenuated by the positive interaction between both factors. The influence of inoculum volume is positive for low values, but it adversely affects ethanol production when it is greater than 24.51 % v/v. The condition required for the obtainment of the maximum yield can be predicted by the model (79) and it is $pH = 4.5, T = 25°C$ and $iv = 12.25\ \%\ v/v$, meaning that the optimal condition is at the lowest extreme of the investigate pH and temperature ranges. The effect of the investigated parameters on the ethanol yield can be better appreciated

by observing contour plots evaluated at the optimal iv value (Figure 42) and at the optimal T value (Figure 43). It is interesting to note that the production of ethanol is favored by low pH and T or high pH and T and that maintaining $T = 25°$, the maximum yield decreases as pH increases. Considering the ethanol yield versus the inoculum volume (Figure 44) at different values of temperature and pH, it is evident that there is a small difference in ethanol production when considering one extreme of the range ($pH = 4.5, T = 25°C$) and the other ($pH = 6.5, T = 37°C$). On the other hand, there is a significant decrease of the yield if at high pH, temperature is decreased, obtaining a minimum when $T = 25°C$. Such results are extremely important from the operative point of view, because they evidenced that the interaction between temperature and pH is not trivial and that effects of eventually not controlled pH could be partially compensated by proper temperature variation.

The validation of the regression model has been also carried out by performing a further experiment at intermediate conditions i.e., $T = 33.4°C$, $pH = 6$ and $iv = 12\ \%v/v$. The measured yield is $Y_{EtOH,exp}$ =82.67% that is quite close to the model prediction $Y_{EtOH,pred}$=81.84%, confirming the model effectiveness. It is worth noting that the obtained results agree with previous investigation on valorization of BSG (Rojas-Chamorro et al., 2020b), where a theoretical yield equal to 78% had been obtained when using a simultaneous saccharification and fermentation in presence of microorganism able to convert xylose to ethanol.

Figure 39. Fermentation samples for DoE.

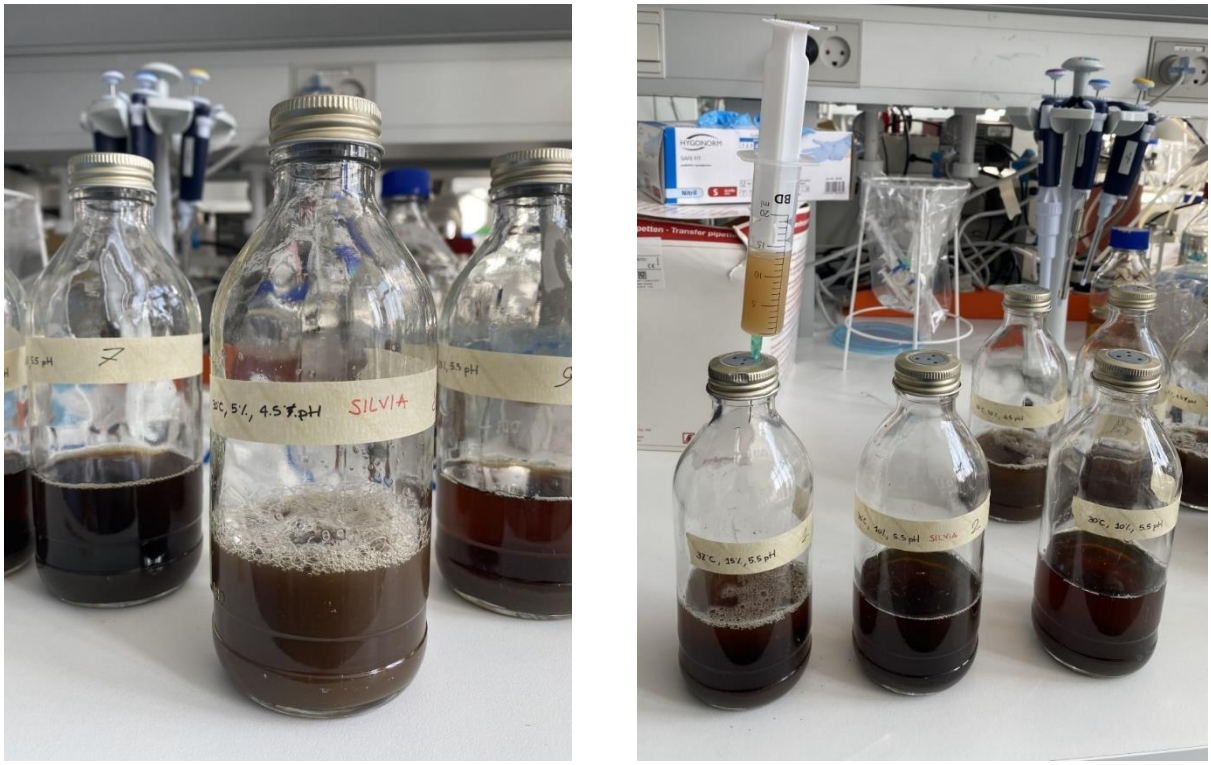

Figure 40. Sample detail for fermentation DoE (left panel) and inoculation moment (right panel).

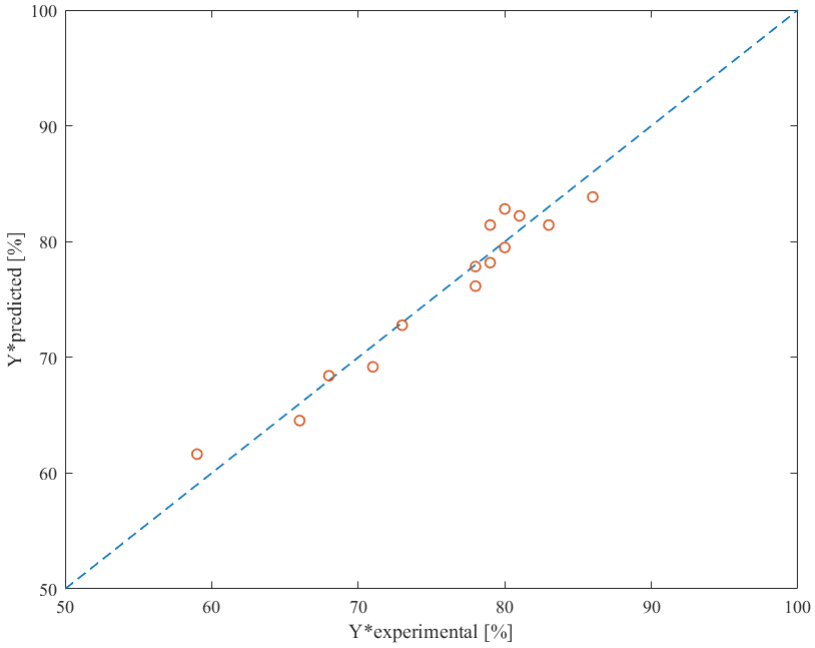

Figure 41. Comparison between experimental values and values predicted from the model.

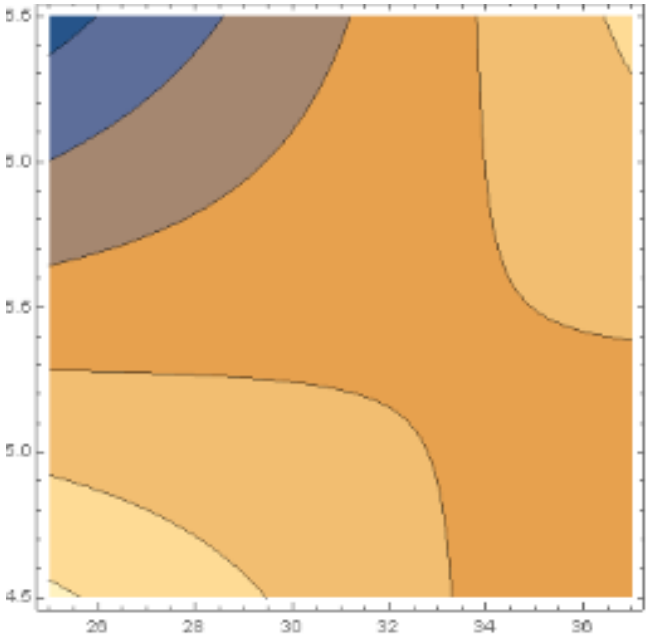

Figure 42. Contour plot of the yield response at the optimal conditions ($pH = 4.5$, $T = 25°C$, and $iv = 12.25$). Isolevel curves with respect to T and pH, inoculum volume is set to 12.879.

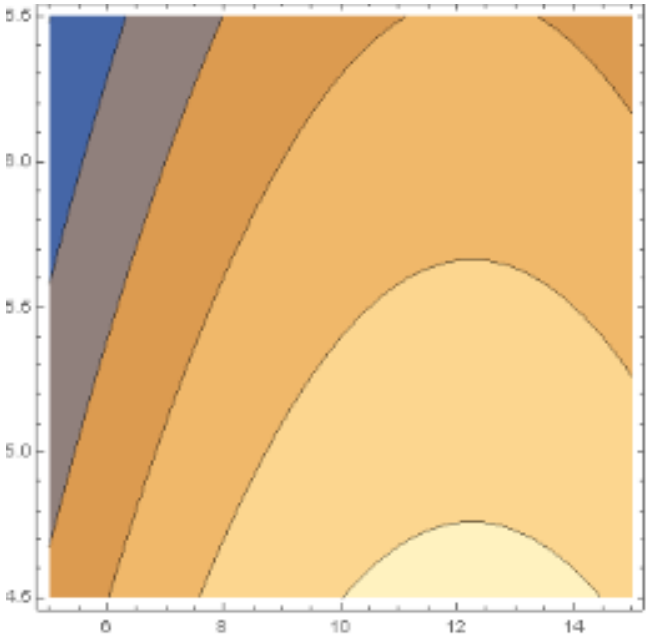

Figure 43. Contour plot of the yield response at the optimal conditions ($pH = 4.5$, $T = 25°C$, and $iv = 12.25$). Isolevel curves with respect to pH and iv, T is set to 25°C.

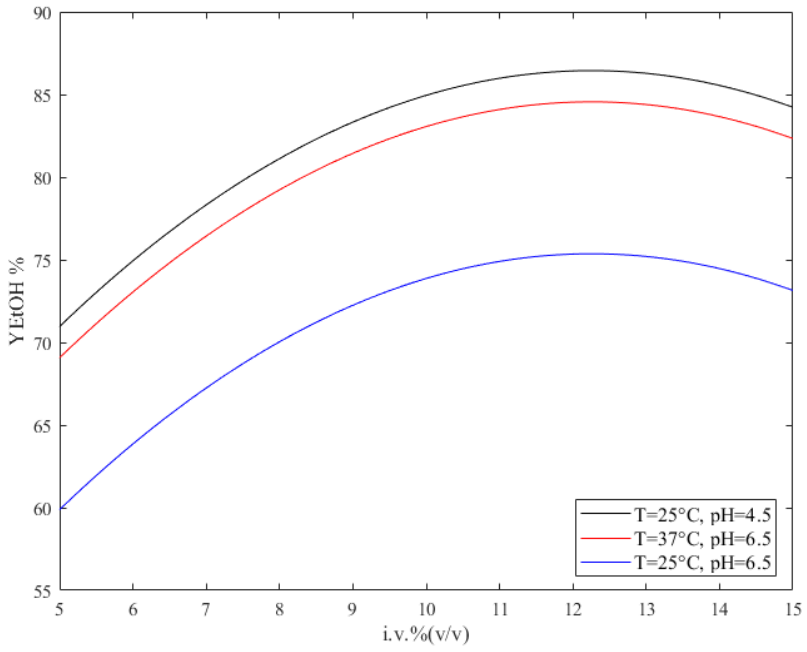

Figure 44. Ethanol yield vs inoculum volume at $pH = 4.5$, $T = 25°C$ (red), at $pH = 6.5$, $T = 37°C$ (blue), at $pH = 6.5$, $T = 25°C$ (black).

4.2.4 Conclusions

Waste recycling has become a popular practice as it addresses management problems while exploiting waste materials as valuable resources. This chapter focused on the brewer 'spent grain (BSG), which represents a valuable by-product of the brewing industry due to its chemical composition and low-cost availability. Its composition and possible applications were reported and analyzed. Recently, efforts have been made to use BSG in alternative ways other than as animal feed. Indeed, successful results have been found from incorporating it into human diets, highlighting the value of its chemical composition. BSG proved to have a potential in the chemical and biotechnology area, particularly as substrate for biofuel production and for the extraction of valuable bioactive compounds for use in the chemical, pharmaceutical, and cosmetic production. The study here reported specifically examined the potential use of BSG as a fermentation substrate for ethanol production. An initial phase of acid hydrolysis was performed to maximize the sugars release, i.e., finding the condition where the maximum released amount was realized. A glucose yield of 18.12 g per 100 g of dried BSG was obtained performing the acid pre-treatment under optimized conditions (0.37 M H_2SO_4, 10% S-L ratio). The hydrolysed biomass was then fermented by *Saccharomyces cerevisiae* yeasts, and the results achieved showed that the best condition was found at $T = 25°C$, $pH = 4.5$ (lowest extremes in the investigated intervals), and $i.v. = 12.25\%\ v/v$. At the optimal condition, a yield of 81.03% evaluated with respect to theoretical one was obtained.

Bibliography

Akermann, A., Weiermüller, J., Chodorski, J. N., Nestriepke, M. J., Baclig, M. T., & Ulber, R. (2022). Optimization of bioprocesses with Brewers' spent grain and Cellulomonas uda. *Engineering in Life Sciences*, *22*(3–4), 132–151. https://doi.org/10.1002/elsc.202100053

Alatzas, S., Moustakas, K., Malamis, D., & Vakalis, S. (2019). Biomass potential from agricultural waste for energetic utilization in Greece. *Energies*, *12*(6). https://doi.org/10.3390/en12061095

Aliyu, S., & Bala, M. (2011). Brewer's spent grain: A review of its potentials and applications. *African Journal of Biotechnology*, *10*(3), 324–331. https://doi.org/10.5897/AJBx10.006

Anwar Saeed, M., Ma, H., Yue, S., Wang, Q., & Tu, M. (2018). Concise review on ethanol production from food waste: development and sustainability. In *Environmental Science and Pollution Research* (Vol. 25, Issue 29, pp. 28851–28863). Springer Verlag. https://doi.org/10.1007/s11356-018-2972-4

Atelge, M. R., Krisa, D., Kumar, G., Eskicioglu, C., Nguyen, D. D., Chang, S. W., Atabani, A. E., Al-Muhtaseb, A. H., & Unalan, S. (2020). Biogas Production from Organic Waste: Recent Progress and Perspectives. *Waste and Biomass Valorization*, *11*(3), 1019–1040. https://doi.org/10.1007/s12649-018-00546-0

Bachmann, S. A. L., Calvete, T., & Féris, L. A. (2022). Potential applications of brewery spent grain: Critical an overview. In *Journal of Environmental Chemical Engineering* (Vol. 10, Issue 1). Elsevier Ltd. https://doi.org/10.1016/j.jece.2021.106951

Barampouti, E. M., Grammatikos, C., Stoumpou, V., Malamis, D., & Mai, S. (2022). Emerging Synergies on the Co-treatment of Spent Coffee Grounds and Brewer's Spent Grains for Ethanol Production. *Waste and Biomass Valorization*, *13*(2), 877–891. https://doi.org/10.1007/s12649-021-01543-6

barthhaas_report_2020_en. (n.d.).

Battistella Lasta, H. F., Lentz, L., Gonçalves Rodrigues, L. G., Mezzomo, N., Vitali, L., & Salvador Ferreira, S. R. (2019). Pressurized liquid extraction applied for the recovery of phenolic compounds from beetroot waste. *Biocatalysis and Agricultural Biotechnology*, *21*. https://doi.org/10.1016/j.bcab.2019.101353

Bengardino, M. B., Fernandez, M. v., Nutter, J., Jagus, R. J., & Agüero, M. v. (2019). Recovery of bioactive compounds from beet leaves through simultaneous extraction: Modelling and process optimization. *Food and Bioproducts Processing*, *118*, 227–236. https://doi.org/10.1016/j.fbp.2019.09.013

Birsan, R. I., Wilde, P., Waldron, K. W., & Rai, D. K. (2019). Recovery of polyphenols from brewer's spent grains. *Antioxidants*, *8*(9). https://doi.org/10.3390/antiox8090380

Bušić, A., Kundas, S., Morzak, G., Belskaya, H., Mardetko, N., Šantek, M. I., Komes, D., Novak, S., & Šantek, B. (2018). Recent trends in biodiesel and biogas production. In *Food Technology and Biotechnology* (Vol. 56, Issue 2, pp. 152–173). University of Zagreb. https://doi.org/10.17113/ftb.56.02.18.5547

Catalkaya, G., & Kahveci, D. (2019). Optimization of enzyme assisted extraction of lycopene from industrial tomato waste. *Separation and Purification Technology*, *219*, 55–63. https://doi.org/10.1016/j.seppur.2019.03.006

Cecilia, J. A., García-Sancho, C., Maireles-Torres, P. J., & Luque, R. (2019). Industrial Food Waste Valorization: A General Overview. In *Biorefinery* (pp. 253–277). Springer International Publishing. https://doi.org/10.1007/978-3-030-10961-5_11

Chetrariu, A., & Dabija, A. (2020). Brewer's spent grains: Possibilities of valorization, a review. In *Applied Sciences (Switzerland)* (Vol. 10, Issue 16). MDPI AG. https://doi.org/10.3390/app10165619

Chiaramonti, D., Prussi, M., Ferrero, S., Oriani, L., Ottonello, P., Torre, P., & Cherchi, F. (2012). Review of pretreatment processes for lignocellulosic ethanol production, and development of an innovative method. *Biomass and Bioenergy*, *46*, 25–35. https://doi.org/10.1016/j.biombioe.2012.04.020

Cho, E. J., Trinh, L. T. P., Song, Y., Lee, Y. G., & Bae, H. J. (2020). Bioconversion of biomass waste into high value chemicals. In *Bioresource Technology* (Vol. 298). Elsevier Ltd. https://doi.org/10.1016/j.biortech.2019.122386

Contreras, M. del M., Lama-Muñoz, A., Romero-García, J. M., García-Vargas, M., Romero, I., & Castro, E. (2021). Production of renewable products from brewery spent grains. In *Waste Biorefinery: Value Addition through Resource Utilization* (pp. 305–347). Elsevier. https://doi.org/10.1016/B978-0-12-821879-2.00011-9

da Silva, E. G., Borges, A. S., Maione, N. R., Castiglioni, G. L., Suarez, C. A. G., & Montano, I. D. C. (2020). Fermentation of hemicellulose liquor from Brewer's spent grain using Scheffersomyces stipitis and Pachysolen tannophilus for production of 2G ethanol and xylitol. *Biofuels, Bioproducts and Biorefining*, *14*(2), 127–137. https://doi.org/10.1002/bbb.2072

de Andrade Lima, M., Kestekoglou, I., Charalampopoulos, D., & Chatzifragkou, A. (2019). Supercritical fluid extraction of carotenoids from vegetable waste matrices. *Molecules*, *24*(3). https://doi.org/10.3390/molecules24030466

Dhiman, S., & Mukherjee, G. (2021). Present scenario and future scope of food waste to biofuel production. In *Journal of Food Process Engineering* (Vol. 44, Issue 2). Blackwell Publishing Inc. https://doi.org/10.1111/jfpe.13594

Emmanuel, J. K., Nganyira, P. D., & Shao, G. N. (2022). Evaluating the potential applications of brewers' spent grain in biogas generation, food and biotechnology industry: A review. *Heliyon*, *8*(10), e11140. https://doi.org/10.1016/j.heliyon.2022.e11140

Feng, K., Huang, Z., Peng, B., Dai, W., Li, Y., Zhu, X., Chen, Y., Tong, X., Lan, Y., & Cao, Y. (2020). Immobilization of Aspergillus niger lipase onto a novel macroporous acrylic resin: Stable and recyclable biocatalysis for deacidification of high-acid soy sauce residue oil. *Bioresource Technology*, *298*. https://doi.org/10.1016/j.biortech.2019.122553

Gil-Castell, O., Mascia, N., Primaz, C., Vásqez-Garay, F., Baschetti, M. G., & Ribes-Greus, A. (2022). Brewer's spent grains as biofuels in combustion-based energy recovery processes: Evaluation of thermo-oxidative decomposition. *Fuel*, *312*. https://doi.org/10.1016/j.fuel.2021.122955

Gmoser, R., Sintca, C., Taherzadeh, M. J., & Lennartsson, P. R. (2019). Combining submerged and solid state fermentation to convert waste bread into protein and pigment using the edible filamentous fungus N. intermedia. *Waste Management*, *97*, 63–70. https://doi.org/10.1016/j.wasman.2019.07.039

Gopinatha Kurup, G., Adhikari, B., & Zisu, B. (2019). Recovery of proteins and lipids from dairy wastewater using food grade sodium lignosulphonate. *Water Resources and Industry*, *22*. https://doi.org/10.1016/j.wri.2019.100114

Guerrero, L. A., Maas, G., & Hogland, W. (2013). Solid waste management challenges for cities in developing countries. *Waste Management*, *33*(1), 220–232. https://doi.org/10.1016/j.wasman.2012.09.008

Gupta, A., & Verma, J. P. (2015). Sustainable bio-ethanol production from agro-residues: A review. In *Renewable and Sustainable Energy Reviews* (Vol. 41, pp. 550–567). Elsevier Ltd. https://doi.org/10.1016/j.rser.2014.08.032

Herbst, G., Hamerski, F., Errico, M., & L. Corazza, M. (2021). Pressurized liquid extraction of brewer's spent grain: Kinetics and crude extracts characterization. *Journal of Industrial and Engineering Chemistry*, *102*, 370–383. https://doi.org/10.1016/j.jiec.2021.07.020

Huige, N. J. (2006). Brewery By-Products and Effluents . In *Handbook of Brewing* (pp. 670–729). CRC Press.

Ikram, S., Huang, L. Y., Zhang, H., Wang, J., & Yin, M. (2017). Composition and Nutrient Value Proposition of Brewers Spent Grain. In *Journal of Food Science* (Vol. 82, Issue 10, pp. 2232–2242). Blackwell Publishing Inc. https://doi.org/10.1111/1750-3841.13794

Ikram, S., Zhang, H., Ming, H., & Wang, J. (2020). Recovery of major phenolic acids and antioxidant activity of highland barley brewer's spent grains extracts. *Journal of Food Processing and Preservation*, *44*(1). https://doi.org/10.1111/jfpp.14308

Kaur, V. I., & Saxena, P. K. (2004). Incorporation of brewery waste in supplementary feed and its impact on growth in some carps. *Bioresource Technology*, *91*(1), 101–104. https://doi.org/10.1016/S0960-8524(03)00073-7

Ladole, M. R., Nair, R. R., Bhutada, Y. D., Amritkar, V. D., & Pandit, A. B. (2018). Synergistic effect of ultrasonication and co-immobilized enzymes on tomato peels for lycopene extraction. *Ultrasonics Sonochemistry*, *48*, 453–462. https://doi.org/10.1016/j.ultsonch.2018.06.013

Leite, P., Silva, C., Salgado, J. M., & Belo, I. (2019). Simultaneous production of lignocellulolytic enzymes and extraction of antioxidant compounds by solid-state fermentation of agro-industrial wastes. *Industrial Crops and Products*, *137*, 315–322. https://doi.org/10.1016/j.indcrop.2019.04.044

Li, S., & Yang, X. (2016). Biofuel production from food wastes. In *Handbook of Biofuels Production: Processes and Technologies: Second Edition* (pp. 617–653). Elsevier Inc. https://doi.org/10.1016/B978-0-08-100455-5.00020-5

Liguori, R., Soccol, C. R., de Souza Vandenberghe, L. P., Woiciechowski, A. L., & Faraco, V. (2015). Second generation ethanol production from brewers' spent grain. *Energies*, *8*(4), 2575–2586. https://doi.org/10.3390/en8042575

Liu, S., Deng, Z., Li, H., & Feng, K. (2019). Contribution of electrodes and electric current to process stability and methane production during the electro-fermentation of food waste. *Bioresource Technology*, *288*. https://doi.org/10.1016/j.biortech.2019.121536

Lynch, K. M., Steffen, E. J., & Arendt, E. K. (2016). Brewers' spent grain: a review with an emphasis on food and health. In *Journal of the Institute of Brewing* (Vol. 122, Issue 4, pp. 553–568). John Wiley and Sons Inc. https://doi.org/10.1002/jib.363

Ma, H. zhi, Xing, Y., Yu, M., & Wang, Q. (2014). Feasibility of converting lactic acid to ethanol in food waste fermentation by immobilized lactate oxidase. *Applied Energy*, *129*, 89–93. https://doi.org/10.1016/j.apenergy.2014.04.098

Mata, T. M., Tavares, T. F., Meireles, S., & Caetano, N. S. (2015). Bioethanol from brewers' spent grain: Pentose fermentation. *Chemical Engineering Transactions*, *43*, 241–246. https://doi.org/10.3303/CET1543041

Mathias, T. R. dos S., de Aguiar, P. F., de Almeida e Silva, J. B., de Mello, P. P. M., & Sérvulo, E. F. C. (2017). Brewery waste reuse for protease production by lactic acid fermentation. *Food Technology and Biotechnology*, *55*(2), 218–224. https://doi.org/10.17113/ftb.55.02.17.4378

McCarthy, A. L., O'Callaghan, Y. C., Neugart, S., Piggott, C. O., Connolly, A., Jansen, M. A. K., Krumbein, A., Schreiner, M., FitzGerald, R. J., & O'Brien, N. M. (2013). The hydroxycinnamic acid content of barley and brewers' spent grain (BSG) and the potential to incorporate phenolic extracts of BSG as antioxidants into fruit beverages. *Food Chemistry*, *141*(3), 2567–2574. https://doi.org/10.1016/j.foodchem.2013.05.048

Meneses, N. G. T., Martins, S., Teixeira, J. A., & Mussatto, S. I. (2013). Influence of extraction solvents on the recovery of antioxidant phenolic compounds from brewer's spent grains. *Separation and Purification Technology*, *108*, 152–158. https://doi.org/10.1016/J.SEPPUR.2013.02.015

Mensah, J. K. M., & Twumasi, P. (2017a). Use of pineapple waste for single cell protein (SCP) production and the effect of substrate concentration on the yield. *Journal of Food Process Engineering*, *40*(3). https://doi.org/10.1111/jfpe.12478

Mensah, J. K. M., & Twumasi, P. (2017b). Use of pineapple waste for single cell protein (SCP) production and the effect of substrate concentration on the yield. *Journal of Food Process Engineering*, *40*(3). https://doi.org/10.1111/jfpe.12478

Mitri, S., Salameh, S. J., Khelfa, A., Leonard, E., Maroun, R. G., Louka, N., & Koubaa, M. (2022). Valorization of Brewers' Spent Grains: Pretreatments and Fermentation, a Review. In *Fermentation* (Vol. 8, Issue 2). MDPI. https://doi.org/10.3390/fermentation8020050

Mussatto, S. I. (2009). Biotechnological Potential of Brewing Industry By-Products. In Springer (Ed.), *Biotechnology for Agro-Industrial Residues Utilisation* (pp. 313–327).

Mussatto, S. I. (2014). Brewer's spent grain: A valuable feedstock for industrial applications. *Journal of the Science of Food and Agriculture*, *94*(7), 1264–1275. https://doi.org/10.1002/jsfa.6486

Mussatto, S. I., Dragone, G., & Roberto, I. C. (2006). Brewers' spent grain: Generation, characteristics and potential applications. In *Journal of Cereal Science* (Vol. 43, Issue 1, pp. 1–14). https://doi.org/10.1016/j.jcs.2005.06.001

Mussatto, S. I., Dragone, G., & Roberto, I. C. (2007). Ferulic and p-coumaric acids extraction by alkaline hydrolysis of brewer's spent grain. *Industrial Crops and Products*, *25*(2), 231–237. https://doi.org/10.1016/j.indcrop.2006.11.001

Mussatto, S. I., & Roberto, I. C. (2006). Chemical characterization and liberation of pentose sugars from brewer's spent grain. *Journal of Chemical Technology and Biotechnology*, *81*(3), 268–274. https://doi.org/10.1002/jctb.1374

Nadar, S. S., & Rathod, V. K. (2019). A co-immobilization of pectinase and cellulase onto magnetic nanoparticles for antioxidant extraction from waste fruit peels. *Biocatalysis and Agricultural Biotechnology*, *17*, 470–479. https://doi.org/10.1016/j.bcab.2018.12.015

Nagarajan, D., Lee, D. J., & Chang, J. S. (2021). Circular Bioeconomy: An Introduction. In *Biomass, Biofuels, Biochemicals: Circular Bioeconomy-Current Developments and Future Outlook* (pp. 3–23). Elsevier. https://doi.org/10.1016/B978-0-12-821878-5.00006-4

Nair, R. B., Lundin, M., Brandberg, T., Lennartsson, P. R., & Taherzadeh, M. J. (2015). Dilute phosphoric acid pretreatment of wheat bran for enzymatic hydrolysis and subsequent ethanol production by edible fungi Neurospora intermedia. *Industrial Crops and Products*, *69*, 314–323. https://doi.org/10.1016/j.indcrop.2015.02.038

Nazzaro, F., Fratianni, F., Ombra, M. N., d'Acierno, A., & Coppola, R. (2018). Recovery of biomolecules of high benefit from food waste. In *Current Opinion in Food Science* (Vol. 22, pp. 43–54). Elsevier Ltd. https://doi.org/10.1016/j.cofs.2018.01.012

Negro, M. J., Alvarez, C., Ballesteros, I., Romero, I., Ballesteros, M., Castro, E., Manzanares, P., Moya, M., & Oliva, J. M. (2014). Ethanol production from glucose and xylose obtained from steam exploded water-extracted olive tree pruning using phosphoric acid as catalyst. *Bioresource Technology*, *153*, 101–107. https://doi.org/10.1016/j.biortech.2013.11.079

Nizami, A. S., Rehan, M., Waqas, M., Naqvi, M., Ouda, O. K. M., Shahzad, K., Miandad, R., Khan, M. Z., Syamsiro, M., Ismail, I. M. I., & Pant, D. (2017). Waste biorefineries: Enabling circular economies in developing countries. In *Bioresource Technology* (Vol. 241, pp. 1101–1117). Elsevier Ltd. https://doi.org/10.1016/j.biortech.2017.05.097

Novak, J. M., & Johnson, M. G. (2018). Elemental and spectroscopic characterization of low-temperature (350°C) lignocellulosic- and manure-based designer biochars and their use as soil amendments. In *Biochar from Biomass and Waste: Fundamentals and Applications* (pp. 37–58). Elsevier. https://doi.org/10.1016/B978-0-12-811729-3.00003-0

Oh, J. I., Lee, J., Lin, K. Y. A., Kwon, E. E., & Fai Tsang, Y. (2018). Biogas production from food waste via anaerobic digestion with wood chips. *Energy and Environment*, *29*(8), 1365–1372. https://doi.org/10.1177/0958305X18777234

Outeiriño, D., Costa-Trigo, I., Pinheiro de Souza Oliveira, R., Pérez Guerra, N., & Domínguez, J. M. (2019). A novel approach to the biorefinery of brewery spent grain. *Process Biochemistry*, *85*, 135–142. https://doi.org/10.1016/j.procbio.2019.06.007

Özvural, E. B., Vural, H., Gökbulut, I., & Özboy-Özbaş, Ö. (2009). Utilization of brewer's spent grain in the production of Frankfurters. *International Journal of Food Science and Technology*, *44*(6), 1093–1099. https://doi.org/10.1111/j.1365-2621.2009.01921.x

Pahla, G., Ntuli, F., & Muzenda, E. (2018). Torrefaction of landfill food waste for possible application in biomass co-firing. *Waste Management*, *71*, 512–520. https://doi.org/10.1016/j.wasman.2017.10.035

Phongthai, S., Lim, S. T., & Rawdkuen, S. (2016). Optimization of microwave-assisted extraction of rice bran protein and its hydrolysates properties. *Journal of Cereal Science*, *70*, 146–154. https://doi.org/10.1016/j.jcs.2016.06.001

Pinheiro, T., Coelho, E., Romaní, A., & Domingues, L. (2019). Intensifying ethanol production from brewer's spent grain waste: Use of whole slurry at high solid loadings. *New Biotechnology*, *53*, 1–8. https://doi.org/10.1016/j.nbt.2019.06.005

Pramanik, S. K., Suja, F. B., Zain, S. M., & Pramanik, B. K. (2019). The anaerobic digestion process of biogas production from food waste: Prospects and constraints. In *Bioresource Technology Reports* (Vol. 8). Elsevier Ltd. https://doi.org/10.1016/j.biteb.2019.100310

Rojas-Chamorro, J. A., Cara, C., Romero, I., Ruiz, E., Romero-García, J. M., Mussatto, S. I., & Castro, E. (2018). Ethanol Production from Brewers' Spent Grain Pretreated by Dilute Phosphoric Acid. *Energy and Fuels*, *32*(4), 5226–5233. https://doi.org/10.1021/acs.energyfuels.8b00343

Rojas-Chamorro, J. A., Romero, I., López-Linares, J. C., & Castro, E. (2020a). Brewer's spent grain as a source of renewable fuel through optimized dilute acid pretreatment. *Renewable Energy*, *148*, 81–90. https://doi.org/10.1016/j.renene.2019.12.030

Rojas-Chamorro, J. A., Romero, I., López-Linares, J. C., & Castro, E. (2020b). Brewer's spent grain as a source of renewable fuel through optimized dilute acid pretreatment. *Renewable Energy*, *148*, 81–90. https://doi.org/10.1016/J.RENENE.2019.12.030

Sadh, P. K., Kumar, S., Chawla, P., & Duhan, J. S. (2018). Fermentation: A boon for production of bioactive compounds by processing of food industries wastes (By-Products). In *Molecules* (Vol. 23, Issue 10). MDPI AG. https://doi.org/10.3390/molecules23102560

Santos, M., Jiménez, J. J., Bartolomé, B., Gòmez-Cordovés, C., & del Nozal, M. J. (2003). Variability of brewer's spent grain within a brewery. *Food Chemistry*, 17–21.

Saravanan, A., Senthil Kumar, P., Jeevanantham, S., Karishma, S., & Vo, D. V. N. (2022). Recent advances and sustainable development of biofuels production from lignocellulosic biomass. In *Bioresource Technology* (Vol. 344). Elsevier Ltd. https://doi.org/10.1016/j.biortech.2021.126203

Shanthi Sravan, J., Butti, S. K., Sarkar, O., Vamshi Krishna, K., & Venkata Mohan, S. (2018). Electrofermentation of food waste – Regulating acidogenesis towards enhanced volatile fatty acids production. *Chemical Engineering Journal*, *334*, 1709–1718. https://doi.org/10.1016/j.cej.2017.11.005

Sivakesava, S., Irudayaraj, J., & Demirci, A. (2001). Monitoring a bioprocess for ethanol production using FT-MIR and FT-Raman spectroscopy. In *Journal of Industrial Microbiology & Biotechnology* (Vol. 26). www.nature.com/jim

Spinelli, S., Conte, A., & del Nobile, M. A. (2016). Microencapsulation of extracted bioactive compounds from brewer's spent grain to enrich fish-burgers. *Food and Bioproducts Processing*, *100*, 450–456. https://doi.org/10.1016/j.fbp.2016.09.005

Spinelli, S., Conte, A., Lecce, L., Padalino, L., & del Nobile, M. A. (2016). Supercritical carbon dioxide extraction of brewer's spent grain. *Journal of Supercritical Fluids*, *107*, 69–74. https://doi.org/10.1016/j.supflu.2015.08.017

Sravan, J. S., Butti, S. K., Sarkar, O., & Mohan, S. V. (2018). Electrofermentation: Chemicals and Fuels. In *Biomass, Biofuels, Biochemicals: Microbial Electrochemical Technology: Sustainable Platform for Fuels, Chemicals and Remediation* (pp. 723–737). Elsevier. https://doi.org/10.1016/B978-0-444-64052-9.00029-7

Stegmann, P., Londo, M., & Junginger, M. (2020). The circular bioeconomy: Its elements and role in European bioeconomy clusters. In *Resources, Conservation and Recycling: X* (Vol. 6). Elsevier B.V. https://doi.org/10.1016/j.rcrx.2019.100029

Steiner, J., Procopio, S., & Becker, T. (2015a). Brewer's spent grain: source of value-added polysaccharides for the food industry in reference to the health claims. In *European Food Research and Technology* (Vol. 241, Issue 3, pp. 303–315). Springer Verlag. https://doi.org/10.1007/s00217-015-2461-7

Steiner, J., Procopio, S., & Becker, T. (2015b). Brewer's spent grain: source of value-added polysaccharides for the food industry in reference to the health claims. In *European Food Research and Technology* (Vol. 241, Issue 3, pp. 303–315). Springer Verlag. https://doi.org/10.1007/s00217-015-2461-7

Tang, D. S., Yin, G. M., He, Y. Z., Hu, S. Q., Li, B., Li, L., Liang, H. L., & Borthakur, D. (2009). Recovery of protein from brewer's spent grain by ultrafiltration. *Biochemical Engineering Journal*, *48*(1), 1–5. https://doi.org/10.1016/j.bej.2009.05.019

Uisan, K., Wong, J., & Kaur, G. (2020a). Bioproducts From Food Waste. In *Current Developments in Biotechnology and Bioengineering: Sustainable Food Waste Management:*

Resource Recovery and Treatment (pp. 279–304). Elsevier. https://doi.org/10.1016/B978-0-12-819148-4.00010-5

Uisan, K., Wong, J., & Kaur, G. (2020b). Bioproducts From Food Waste. In *Current Developments in Biotechnology and Bioengineering: Sustainable Food Waste Management: Resource Recovery and Treatment* (pp. 279–304). Elsevier. https://doi.org/10.1016/B978-0-12-819148-4.00010-5

Wagner, E., Sierra-Ibarra, E., Rojas, N. L., & Martinez, A. (2022). One-pot bioethanol production from brewery spent grain using the ethanologenic Escherichia coli MS04. *Renewable Energy*, *189*, 717–725. https://doi.org/10.1016/j.renene.2022.03.014

Wang, M., Wu, M., & Huo, H. (2007). Life-cycle energy and greenhouse gas emission impacts of different corn ethanol plant types. *Environmental Research Letters*, *2*(2). https://doi.org/10.1088/1748-9326/2/2/024001

Wang, X. J., Qi, J. C., Wang, X., & Cao, L. P. (2013). Extraction of polyphenols from barley (Hordeum vulgare L.) grain using ultrasound-assisted extraction technology. *Asian Journal of Chemistry* , *25*(3), 1324–1330.

Waters, D. M., Jacob, F., Titze, J., Arendt, E. K., & Zannini, E. (2012). Fibre, protein and mineral fortification of wheat bread through milled and fermented brewer's spent grain enrichment. *European Food Research and Technology*, *235*(5), 767–778. https://doi.org/10.1007/s00217-012-1805-9

White, J. S., Yohannan, B. K., & Walker, G. M. (2008). Bioconversion of brewer's spent grains to bioethanol. *FEMS Yeast Research*, *8*(7), 1175–1184. https://doi.org/10.1111/j.1567-1364.2008.00390.x

Wu, W. C., Ng, H. S., Sun, I. M., & Lan, J. C. W. (2017). Single step purification of bromelain from Ananas comosus pulp using a polymer/salt aqueous biphasic system. *Journal of the Taiwan Institute of Chemical Engineers*, *79*, 158–162. https://doi.org/10.1016/j.jtice.2017.04.001

Xiong, X., Yu, I. K. M., Tsang, D. C. W., Bolan, N. S., Sik Ok, Y., Igalavithana, A. D., Kirkham, M. B., Kim, K. H., & Vikrant, K. (2019). Value-added chemicals from food supply chain wastes: State-of-the-art review and future prospects. *Chemical Engineering Journal*, *375*. https://doi.org/10.1016/j.cej.2019.121983

Xiros, C., Topakas, E., Katapodis, P., & Christakopoulos, P. (2008). Hydrolysis and fermentation of brewer's spent grain by Neurospora crassa. *Bioresource Technology*, *99*(13), 5427–5435. https://doi.org/10.1016/j.biortech.2007.11.010

Xu, F., Li, Y., Ge, X., Yang, L., & Li, Y. (2018). Anaerobic digestion of food waste – Challenges and opportunities. In *Bioresource Technology* (Vol. 247, pp. 1047–1058). Elsevier Ltd. https://doi.org/10.1016/j.biortech.2017.09.020

Chapter 5: Development of a Raman spectroscopy-based soft-sensor for monitoring a fermentation process

In this chapter, the development of a spectroscopic sensor for monitoring a fermentation process will be investigated. Results from the combination of Raman spectroscopy with chemometric methods will be reported and analyzed.

The goal of Industry 4.0 is to create a connection between the digital word and traditional industrial processes in order to develop a more efficient manufacturing system, achieving fundamental goals such as greater flexibility, sustainability, and competitiveness. In this scenario, the development of smart technologies seems even more necessary for biomanufacturing. Process monitoring is a key tool for achieving the goals set by the smart industry, as it guarantees safe operations, product quality control and minimization of energy consumption (Gargalo et al., 2020; Udugama et al., 2020). However, in several processes there are still serious difficulties in measuring key variables, and this happens particularly in the biomanufacturing sector. As discussed in the previous chapters, most biotechnological processes include the cultivation of microorganism, during which different combined phenomena occur involving several phases and concentration gradients. Although it is easy to monitor in situ common parameters such as temperature, oxygen partial pressure and pH, this cannot be said for information concerning cell count, substrate and product concentration, or metabolic activity. Thus, there is the necessity for a reliable and consistent analytical system to monitor process conditions at all bioprocess stages (Pachauri et al., 2017). The process of bioethanol production by biomass fermentation is certainly a good example of bioprocess where the presence of an analytical method, capable of providing real-time system status information, is required (Ávila et al., 2012). Among the various innovative methods proposed in the literature, spectral analysis technologies are the ones that can meet the basic requirements for implementing rapid monitoring and control of fermentation processes. These optical methods, such as near-infrared spectroscopy or multiwavelength fluorescence, have been widely applied to develop new techniques for monitoring processes like these (Claßen et al., 2017). In particular, Raman spectroscopy represents a promising methodology in monitoring biotechnological systems due to important features such as speed, high signal-to-noise ratio, good resolution, ability to provide a stable signal, and also low water interference. In addition, no destructivity and no sample pre-treatment requirements must be mentioned (Ávila et al., 2012; Claßen et al., 2017; Esmonde-White et al., 2017; S. K. Oh et al., 2013).

This chapter investigated the development of a Raman spectroscopy-based sensor for the quantitative analysis of products obtained from a fermentation process. Specifically, the analyses reported here were performed on fermentation samples from the previously reported case study (Chapter 4). As proved in Chapter 3, the rejection of disturbances was possible only in presence of a product composition controller. One possible solution involved available online measurements of ethanol concentration, while as an alternative the presence of a nonlinear estimator to infer ethanol composition was evaluated. In either scenario, achieving

the desired product quality depends on knowledge of the reacting medium composition. The first experimental results of the development of a fully automated monitoring system will be presented and analyzed. For this purpose, offline Raman spectra were collected for several samples having known ethanol and glucose concentration, but also of fermentation samples related to the Case Study reported in Chapter 4. Combining Raman spectra with chemometric tools, what is intended to be developed are predictive models for product and substrate concentration, and to accomplish this task spectra of samples of ideal and non-ideal fermentation mixtures were recorded off-line. A data-driven modelling approach was applied in order to correlate the information contained in the Raman spectra to the concentration of investigated compounds. The Variable Importance in Projection (VIP) method was used to identify and select the most significant spectral wave regions, then partial least squares (PLS) regression models for both ethanol and glucose were developed and validated.

5.1 Overview of spectroscopic sensors

As reported by Lourenço et al. (2012), spectroscopy was originally defined as the study of the interaction between electromagnetic radiation and matter as a function of wavelength. Successively, this definition was expanded to include measurements of any property considered a function of either wavelength or frequency. In the review proposed by Claßen et al. (2017), it is clearly explained that there are three different types of effects arising from the interaction between matter and light: absorption, emission, and scattering. These three different types of interactions can be investigated by different spectroscopic methods operating in different wavelength ranges. However, for the purpose of the discussion, only those techniques operating in the spectral ranges between UV and mid-infrared (MIR) will be discussed. Indeed, these include Raman spectroscopy, the exploitation of which was investigated in the work here presented.

5.1.1 UV/Vis Spectroscopy

UV/Vis spectroscopy represents a sensitive molecular spectroscopy that uses ultraviolet and visible light in the wavelength range 200-780 nm. As reported by Beutel & Henkel (2011), new developments involving this technique have resulted in improved applicability in biotechnological process monitoring. Indeed, UV/Vis spectrophotometers show excellent measurement performance not only in the UV region, but also in the range between UV and near-infrared (NIR). Sensitivity, high scanning speed, low cost, and robustness are the main features that make these devices valuable analytical tools for different applications such as the

bioprocess monitoring (Langergraber, Fleischmann, et al., 2004; Langergraber, Gupta, et al., 2004; Noui et al., 2002; Pons et al., 2004; Roberts et al., 2018; Sarraguça et al., 2009).

5.1.2 IR Spectroscopy

Based on the wavelength range, IR spectroscopy can be differentiated into Far-Infrared (FIR) (10-400 cm^{-1} or 25-1000 μm), Mid-Infrared (MIR) (200-4000 cm^{-1} or 2500-25000 nm), and Near Infrared (NIR) (4000-12500 cm^{-1} or 800-2500 nm). FIR is generally not applied in this area, unlike NIR, which seems to be more widely adopted. Indeed, as reported by Lourenço et al. (2012), recently NIR has been widely used with the pharmaceutical industry for testing raw materials, product quality control, and process monitoring. Several examples of its application can be found in the literature (Arnold et al., 2002, 2003; Nordon et al., 2008; Roychoudhury et al., 2007; Tosi et al., 2003). This growing interest can be attributed to its interesting characteristics such as the absence of sample pre-treatment, no alteration of samples (which can be used for further analysis), as well as the rapidity in acquiring spectra. MIR spectroscopy has also been used for quantitative analysis of fermentative process (Fayolle et al., 2000; Schenk et al., 2007). In general, the greatest drawback of this spectroscopy is the significant absorbance of water in the infrared region. This represents a non-trivial problem, since the water absorption could cover the characteristic peaks of other compounds, reducing the information level of the spectra.

5.1.3 Fluorescence Spectroscopy

Fluorescence spectroscopy, also known as fluorimetry, exploits the characteristics of particular chemical compounds called fluorophores. These, when subjected to light excitation, are able to re-emit light. This occurs at specified wavelength, in correspondence of which fluorophores absorb visible or UV light. Thus, electrons are excited, and they relax to their ground electronic state emitting light of lower energy (Claßen et al., 2017). In the context of bioprocesses, some molecules such as proteins, certain amino acids, co-enzymes, and even vitamins can exhibit the same behaviour. Fluorescence measurements can be carried out both in situ and online using spectrophotometers. Several examples of applications concerning this spectroscopic technique are available in the literature, since it has been used for more than 30 years for monitoring bioprocesses, and in particular for the analysis of key compounds such as biomass, glucose, and ethanol (Eliasson Lantz et al., 2006; Jain et al., 2011; Ödman et al., 2009; Rhee & Kang, 2007; Surribas et al., 2006).

5.1.4 Raman Spectroscopy

Unlike IR spectroscopy, Raman spectroscopy is a technique that relies on the inelastic scattering of light onto interaction with the sample (Claßen et al., 2017). It represents a useful tool since it provides information regarding vibrational and rotational transitions in molecules whose polarizability varies upon excitation. When monochromatic light, generally from a laser, is incident on molecules, and thus interacts with their vibrational frequencies, most of the photons are scattered elastically (Raylegh scattering). Instead, a small amount of the scattered light shifts in energy from its original wavelength to on o more different wavelengths (Raman scattering). The wavelength shift, i.e., the difference in energy between the laser (incident) light and the scattered light, is a function of the chemical bounds that caused the Raman scattering. By plotting the intensity of the scattered light as a function of the frequency (Raman shift wavenumber), it is possible to obtain the Raman spectrum of the analyzed sample. The scattered Raman photons carry with them information about the chemical structure and the nature of the material. In addition, from the spectra obtained it is possible to gain both qualitative and quantitative information (Lourenço et al., 2012). Indeed, the band areas are proportional to the concentration and for this reason this technique can be implemented for the purpose of quantitative analysis. As stated by Claßen et al. (2017) and Lourenço et al. (2012), Raman spectroscopy has several advantages such as the flexibility of the sample type that can be analyzed (solid, liquid, gases, or combinations), no sample preparation, a high signal-to-noise ratio, sufficient resolution, and no sensitivity of the spectra to water. This makes it suitable for monitoring operating units such as bioreactors, of which useful information can be obtained. In general, the Raman signal tends to be weak, and the possible presence of biological molecules could cause fluorescence in the Raman scattering band, creating interference in the signal. As explained by Beutel & Henkel (2011), when this phenomenon occurs, the strong fluorescent activity of the possible biological molecules could overlap with the Raman bands. The intensity recorded is very high even though in absolute terms it may not be, but this is because Raman spectrometers are constructed to read only weak signals. In this case, what is obtained is a noisy spectrum difficult to interpret, where the Raman bands appear as small and narrow peaks on a high-intensity background. Although Raman spectroscopy represents a valuable analytical tool that can be applied in various fields, its application in bioprocess monitoring is still circumscribed. In several cases, its potential in monitoring microbial cultivations has been evaluated and demonstrated, along with the use of statistical

tools (Ávila et al., 2012; Hirsch et al., 2019; Iversen et al., 2014; S.-K. Oh et al., 2012.; Picard et al., 2007; Schalk et al., 2017).

5.2 Materials and Methods

This section will report only the chemical and reagents used for the preparation of ideal mixtures, regarding those related to pre-treated BSG fermentation, including yeast cultivation and acid hydrolysis, refer the reader to chapter 4.

5.2.1 Chemicals and Reagents

Samples of ideal mixtures

Ethanol (absolute for analysis, Carlo Erba reagents), glucose (highly pure, Microbial Diagnostic), acetic acid (≥99%, Sigma-Aldrich) and formic acid (85%, Carlo Erba reagents) were used for the preparation of standard aqueous solutions for calibration of PLS models.

5.2.2 Fermentation

The fermentation process was performed in duplicate in 2 L reactors (Applikon Biotechnology BV, Nieuwpoortweg, The Netherlands) anaerobically at 50 rpm, T=33.6 °C, pH=5.3, and i.v.=12.25% for 9 h. These do not represent the optimized conditions identified in the previous chapter, but they are intermediate conditions of the investigated domain. The working volume of the acid hydrolysate for each reactor was about 1.7 L. Before being inoculated, the fermentation medium was neutralized using 0.1 M NaOH until a pH of 5.3 was reached, then 12.25% (v/v) of inoculum was added. During the experimental run, samples were taken at different times, filtered and stored at 4°C for determination of ethanol and sugars by HPLC, and Raman spectroscopy analysis. Figures 45-46 show the bioreactors used to conduct the run.

Figure 45. The two bioreactors used to conduct the experiments.

Figure 46. Detail of the second bioreactor.

5.2.3 Raman spectra

Raman spectra of ideal and not-ideal fermentation solutions were collected/recorded by using a Cora 5X00 Raman spectrophotometer from Anton Paar, with a 532 nm laser probe. The Raman equipment available in the laboratory covers the wavelengths between 197.7 cm^{-1} and 3500 cm^{-1}. Since Raman spectroscopy appears to be very sensitive to light, the measurements were conducted in a dark environment. In Figure 47 it is possible to observe the dark room/camera for the laser probe, and in which the sample to be analyzed was placed. Instead, Figure 48 shows the entire apparatus.

Figure 47. Detail of the dark camera for the laser probe.

Figure 48. Cora 5X00 Raman spectrophotometer equipment.

The values of the measurement parameters at which the samples were analyzed are reported in the following Table 20. The only parameters that could be changed were the integration time and the number of averaged spectra, while the laser power remained constant at 50 mW for the 532 nm laser probe. In addition, the spectra acquisitions were made by choosing the background subtraction mode, which helps eliminate the background noise present, the smoothing mode, useful in the case of turbid samples such as the not-ideal ones analyzed here, and finally the baseline correction. It is important to highlight that baseline correction was carried out in all spectra. Indeed, this represents a crucial and necessary step to eliminate background signals generated by residual Rayleigh scattering or fluorescence, which would obscure the acquired Raman spectrum.

Table 20. Values of measurement parameters.

Laser power	50 mW
Integration time	3000 ms
Average spectra	30

5.2.4 Design of experiments (DOE) for ideal fermentation solutions

As stated by de Luca et al. (2009), the prediction capabilities of a regression model are significantly affected by the design of the mixtures for the construction of the calibration set. For this reason, it is appropriate to carefully select the reference solutions by means of an experimental campaign, in order to maximize the effectiveness of the model by improving the information that can be derived from the calibration set. For this purpose, the design of statistical experiments represents a useful tool to choose the concentrations of the components of interest in order to minimize the total number of samples considering the entire experimental range. In this work, the Simplex-lattice design was considered and implemented on Minitab®. This software returned as result a table in which the combinations of the components considered (ethanol, glucose, acetic and formic acids) are shown, the proportions of which are indicated with values between 0 and 1. It is worth noting that each combination shown represents the composition of a specific mixture, the total sum of which must equal to 1. The composition values of each component returned by DOE in coded form were multiplied by the corresponding maximum concentration value selected, considering this limit as the concentration of the component when the mole fraction is equal to 1. The maximum values

considered are the following: 15% (v/v) for ethanol, 20 g/L for glucose, 0.5 g/L for formic acid, and 2 g/L for acetic acid. The table 21 reported below shows all the combinations of the analyzed mixture, together with those of ethanol and glucose only evaluated in a primary phase of sensor calibration, and for which the following composition ranges were considered: $C_{EtOH} = [1 - 15\%\ (v/v)]$, $C_{glucose} = [1 - 30\ g/L]$.

Table 21. Compositions of the analyzed solutions.

	Glucose [g/L]	Ethanol [%(v/v)]	Formic Acid[g/L]	Acetic Acid [g/L]
1	0.5	-	-	-
2	1	-	-	-
3	1.5	-	-	-
4	2.5	-	-	-
5	5	-	-	-
6	5	3.75	0.25	-
7	5	3.75	-	1
8	5	11.25	-	-
9	7	-	-	-
10	10	-	-	-
11	10	-	-	1
12	10	-	0.25	-
13	10	7.5	-	-
14	15	-	-	-
15	15	3.75	-	-
16	20	-	-	-
17	30	-	-	-
18	-	1	-	-
19	-	2.5	-	-
20	-	5	-	-
21	-	7.5	-	1
22	-	7.5	0.25	-
23	-	10	-	-
24	-	15	-	-
25	-	-	0.5	-
26	-	-	-	2
27	-	-	0.25	1

5.2.5 Statistical analysis

The complexity of the samples analyzed often results in a complex spectrum. In complicated systems such as fermentation, it can be challenging to identify which bands belong to specific components. This task becomes even harder when other substances are present, such as water or substrate, that have a higher concentration and they may overpower the signal of the interest compounds. In order to overcome these difficulties and to perform an accurate and robust analysis, it is necessary to use statistical tools such as chemometric methods. This term refers to the use of mathematical or statistical methods to analyze data about a chemical system, to extract information about a system state, to characterize the behavior of a system, and also to facilitate in the identification of processes occurring in an analyzed system, as reported by Lourenço et al. (2012). Generally, monitoring of bioprocesses by spectroscopic techniques provides a set of spectra that constitute consistent data sets, characterized by less information content than data volume, and from which significant information should be extracted quickly. The application of multivariate data analysis, which combines spectroscopy with statistical techniques, to spectral data enables to reduce data complexity and to get significant information out of the spectra. This information is usually distributed over an entire spectroscopic data set and is not restricted to a specific part of the spectrum or to just one spectrum (Rathore et al., 2011). Among the most widely used chemometric methods for the analysis of spectral data in the bioprocess monitoring is possible include/mention dimension reduction, latent variable methods such as principal component analysis (PCA), principal component regression (PCR), and partial least-square (PLS). The reason for their widespread use is their efficiency in extracting information, their simplicity, and the stability over time and interpretability of their models (Lourenço et al., 2012). In this work, the PLS regression technique was applied to extract information regarding ethanol and glucose composition from the collected Raman spectra. A brief overview of this method is provided below.

Partial Least Square (PLS) regression method

As reported by Lourenço et al. (2012), regression methods are usually applied to obtain a relationship between the collected spectra and the quantifiable properties of the analyzed samples. More specifically, the regression problem consists of modeling one or more independent variables, known as responses and denoted by **Y**, by means of a set of predictor variables, **X**. It represents undoubtedly the most common data-analytic problem. As stated by Wold et al. (2001), modeling of **Y** by **X** is usually accomplished using multiple linear regression (MLR), whose performance is good as long as **X** variables are few and uncorrelated.

However, in data derived from modern measurement instruments such as spectrometers or chromatographs, the **X** variables tend to be too many and highly correlated. In these situations, PLS-regression (PLSR) allows more complex problems to be investigated and the available data to be analized more realistically. This approach represents a variation of traditional multiple linear regression (MLR), from which it differs in its ability to analyze data having strong collinearity (correlated), noise, and numerous independent variables (**X**). Additionally, it allows for the simultaneous modeling of several dependent variables (**Y**) (Wold et al., 2001). It was developed in 1975 by Herman Wold for analyzing complex data sets, grouped in chains of matrices (blocks), using a method called NIPALS (Non.linear Iterative PArtial Least Squares), from which the acronym PLS was derived for these models.

Also known as projection to latent structures, the PLSR method is one of the most widely applied multivariate calibration methods that allows to construct a mathematical model though matrices of **X** (e.g., spectra) and **Y** (sample properties such as concentration) from a set of reference samples (de Luca et al., 2009). In other words, it can be defined as a guided decomposition model in which the dependent variables directly intervene in the decomposition of the independent ones (Lourenço et al., 2012). The aim of this method is to establish a small number of latent variables that are able to predict the properties of the samples (e.g., compounds concentration) through the spectral data that should be used in an efficient way. There are several methods for calculating the terms of the PLSR model, but the one most widely used is the NIPALS algorithm. The next section will discuss the development of the PLSR model.

PLSR model

The PLS regression model can be obtained from a training set composed of N observations with L X-variables, denoted by x_l ($l=1,\ldots, L$) and M Y-variables denoted by y_m ($m=1,\ldots, M$). As showed in Figure 49, these data form two matrices X and Y of dimension $(N \times L)$ and $(M \times L)$.

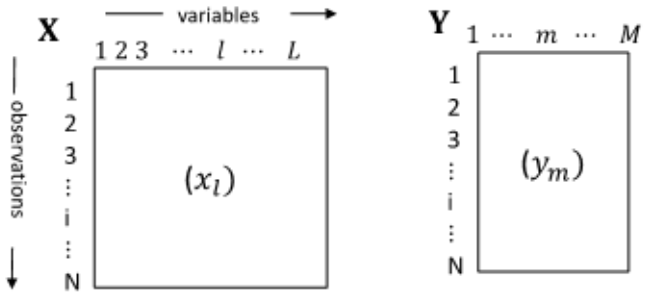

Figure 49. Graphical representation of descriptor matrices (**X**-variables, Raman spectra) and measurements (**Y**-variables, compositions).

The following Equations 80-82 show how the PLS regression model is developed, more specifically how the matrices **X** and **Y** are decomposed into the smaller score **T**, loadings **P** and **Q**, and error **E** and **F** matrices. Consequently, a linear relationship between the scores can be established as reported below.

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} = \sum_{i=1}^{A} t_i p_i^T + \mathbf{E} \tag{80}$$

$$\mathbf{Y} = \mathbf{T}\mathbf{Q}^T + \mathbf{F} = \sum_{i=1}^{A} t_i q_i^T + \mathbf{F} \tag{81}$$

$$\mathbf{T} = \mathbf{X}\mathbf{W}(\mathbf{P}^T\mathbf{W})^{-1} \tag{82}$$

As already said, **E** and **F** are error matrices of size $(N \times L)$ and $(N \times M)$, respectively. They contain the part of **X** and **Y** which the model does not explain. The score matrix **T** $(N \times A)$ is composed of the column vector $\mathbf{t_i}$, and $\mathbf{p_i}$ and $\mathbf{q_i}$ are the column vectors, so-called the loadings, that compose the matrices **P** $(L \times A)$ and **Q** $(M \times A)$. Moreover, **W** $(L \times A)$ is the weight matrix obtained by the PLS model regression and A is the number of latent variables chosen to explain the significative variance of the data. It is worth noting that, during the matrix **X** decomposition, information contained in matrix **Y** is considered in order to increase the variance between these two matrices. Usually, the NIPALS algorithm is responsible for the decomposition step.

In this work, PLS regression models were obtained through the Matlab ® function `plsregress` that uses the Statistically Inspired Modification of the Partial Least Squares (SIMPLS) algorithm (de Jong, 1993). NIPALS and SIMPLS algorithms represent the most widely used algorithms for PLS analysis. However, the latter differs from the first one because it is not an

iterative algorithm, and the latent variables calculation is performed directly on the basis of singular value decomposition (SVD) of the data. From a computational point of view, this algorithm is certainly faster than NIPALS, but it is also less robust and for this reason NIPALS algorithm is usually used. Nevertheless, SIMPLS is particularly useful for the analysis of data particularly affected by noise or when the number of predictors is greater than the samples number. Indeed, in these situations the NIPALS algorithm could be too time consuming (Alin, 2009).

Variable selection methods

As already stated, the complexity of the samples analyzed results in complexity of the spectrum acquired, and thus this has a significant impact on the data to be analyzed. When talking about samples derived from online monitoring, one must think of the raw, untreated samples, and particularly if they derive from fermentation mixtures as in the case discussed here, they will surely also contain other compounds in addition to those of interest. Among these is certainly biomass, the presence of which tends to bring/add considerable noise into the spectrum, complicating its interpretation. Therefore, in order to minimize the influence of noisy variables such as this, data reductions are usually necessary, which can be realized either through projection or variable selection methods, or through a combination of both (Mehmood et al., 2012).

Variable selection methods turn out to be highly adopted in Partial Least Squares regression (PLSR), as they do not only reduce the noise generated by irrelevant variables as mentioned above, but also improve estimation/prediction performance, model interpretation and understanding of the investigated system. There are numerous methods for variable selection in the literature, which can be grouped according to the classification proposed by Mehmood et al. (2012), and later also adopted by Cocchi et al. (2018). It involves the distinction into three main groups: filter methods, wrapper methods, and embedded methods. In the first one, it is possible to find those methods according to which the variable selection takes place through two fundamental steps: firstly, the PLSR model is fitted to the data, and secondly the variable is selected through the introduction of a threshold value on some measure of relevance obtained from the PLS model. These techniques are usually quick and easy to implement. In the wrapper methods, the feature selection takes place using the filter methods in an iterative way. These are techniques based on some supervised learning approach, where the refitting of the model is included in the variable search algorithm. In addition, a measure of model performance is

used to determine the best subset of the variables. Finally, the embedded methods are those in which variable selection is included with model derivation in a single one-step procedure. In this case, the optimal subset of variable is sought for each component of PLS model. Furthermore, relying on a single iterative procedure, these techniques are usually faster than the wrapper methods.

Among the categories of variable selection methods listed above, one of the most popular is certainly the variable importance in PLS projections (VIP) which was introduced by Wold et al. (1993) as "Variable influence on projection" and known also as "Variable importance in projection score" or VIP scores (Eriksson et al., 2006). In this work it was used to identify the wavelengths at which the absorption signal appears to have a greater influence on the **Y** variations (i.e. concentrations), thus discriminating them with respect to those without any discrimination power. VIP scores are useful in understanding the **X** space predictor variables that best explain **Y** variance. For each j-th variable of **X**, VIP scores can be defined as the sum, over the latent variables, of its PLS-weight value (w_j) weighted by the percentage of explained **Y** variance (SSY) by each specific latent variable, as reported in the following formula:

$$VIP_j = \sqrt{\frac{\sum_{a=1}^{A}\left(SSY_a\left(\frac{w_{ja}}{\|w_a\|}\right)^2\right)L}{SSY_{tot}A}} \tag{83}$$

Where SSY_a is the percentage of variance explained by the ath component, w_{ja} are the elements of the weight matrix found by the model, and L is the number of variables in **X**. The VIP criterion, as it was first suggested, had a set range for its parameters since the total of the squared VIP for all variables equals the number of variables. For this reason, it is common to consider a threshold of VIP greater than 1, which indicates that the chosen variable has a more influence on the model predicting **Y** than the average of the squared VIP values. This technique is useful for eliminating unimportant variables, but it may have limitation when determining the significance of features. This selection criterion has been shown to be a suitable option, even when the **X** variables are spectral data as reported by study of Chong & Jun (2005). Other possible options include applying a threshold of 2/3 or using the average of VIP values as the threshold (Cocchi et al., 2018).

In this work, the VIP variable selection method was also implemented on Matlab®. After performing the PLS regression on the complete data, as shown in the previous section, a

reduced number of variables are selected by means of the VIP technique. On the new data matrix, decomposed only by the variables that are considered significant and thus having a VIP index greater than 1, the PLS regression is again carried out.

Another important statistical parameter that was calculated at each simulation, in order to assess the quality of the model, is the coefficient of determination R^2. As explained by de Luca et al. (2009) in their work, R^2 can be defined as a significant quality index in fitting all data to a straight line and represents the fraction of the total variance explained by the model. It can be calculated as follows (Eq. 84).

$$R^2 = 1 - \frac{(\hat{y}_i - y_i)^2}{(y_i - \overline{y})^2} \tag{84}$$

Where y_i are the observed data, $\overline{y}$ their mean, and $\hat{y}$ are the values predicted by the regression model.

PLSR model validation

It is common for the training set to produce accurate predictions, but this does not guarantee that the method can predict properly unknown data. To confirm the prediction quality, it is usually recommended to use a test set. This represents a set of samples that were not included in the original calculations, thus can be defined as a "blind test". These samples are initially considered as unknown. The evaluation of the prediction quality by using a test set can be considered as a form of validation. This step can be accomplished in several ways, and cross-validation represents one of them. This method requires only one training set and it consists of removing one or a group of samples, and then determining the model using the remaining samples. In the work presented here, it was decided to opt for leave-one-out cross-validation, which represents the most common approach. According to this technique, one sample is left out at a time, and then the same sample is predicted using the calibration constructed with the other reference samples. The process is repeated until all samples have been left out in turn (Brereton, 2007; de Luca et al., 2009).

5.3 Results

This section will report the results obtained. Since the study of the development of a Raman-based spectroscopy sensor to monitor ethanol and glucose composition was divided into two

stages, the results will be shown in the following order: firstly, ethanol and glucose regression models obtained from the statistical analysis of Raman spectra recorded for ideal solutions, prepared ad hoc at known concentration of both investigated compounds, will be analyzed. In this case, the presence of secondary compounds and their possible influence on the predictive capabilities of the models was also evaluated. Secondly, the results related to the analysis of samples obtained from the fermentation of pre-treated BSG will be discussed. In this case, the results of the model calibration step will be followed by the cross-validation step using the leave-one-out (LOO) technique.

5.3.1 Ideal fermentation solutions

All spectra recorded in this first phase of analysis were pretreated by normalizing them with respect to a reference wavelength value, i.e., by subtracting from each Raman shift value the one corresponding to the peak at 2446 cm^{-1}. This value was selected because it did not correspond to any significant peak, and since its corresponding intensity value was very low, its influence on other peaks was assumed to be unremarkable. Moreover, starting from a full Raman shift range between 197.7 cm^{-1} and 3500 cm^{-1}, it was decided to evaluate a reduced range between 400 cm^{-1} and 3020 cm^{-1} being the lower wavelengths less informative and more affected by noise.

Ethanol

Figure 50 shows the spectra of ethanol solutions at different concentration values [1-15% (v/v)], where it is possible to clearly identify its characteristic peaks:

- Stretching vibrations C-C at 886 cm^{-1}.
- Librations of the CH_3 groups at 1100-1116 cm^{-1}.
- Torsion and rotational vibrations of the CH_2 groups near to 1280 cm^{-1}.
- Bending vibration of CH_3 and CH_2 groups at 1454 cm^{-1}.
- Stretching symmetric and asymmetric vibrations of the CH_3 groups between 2880 and 2980 cm^{-1}, with the most evident peak approximately at 2934 cm^{-1}.

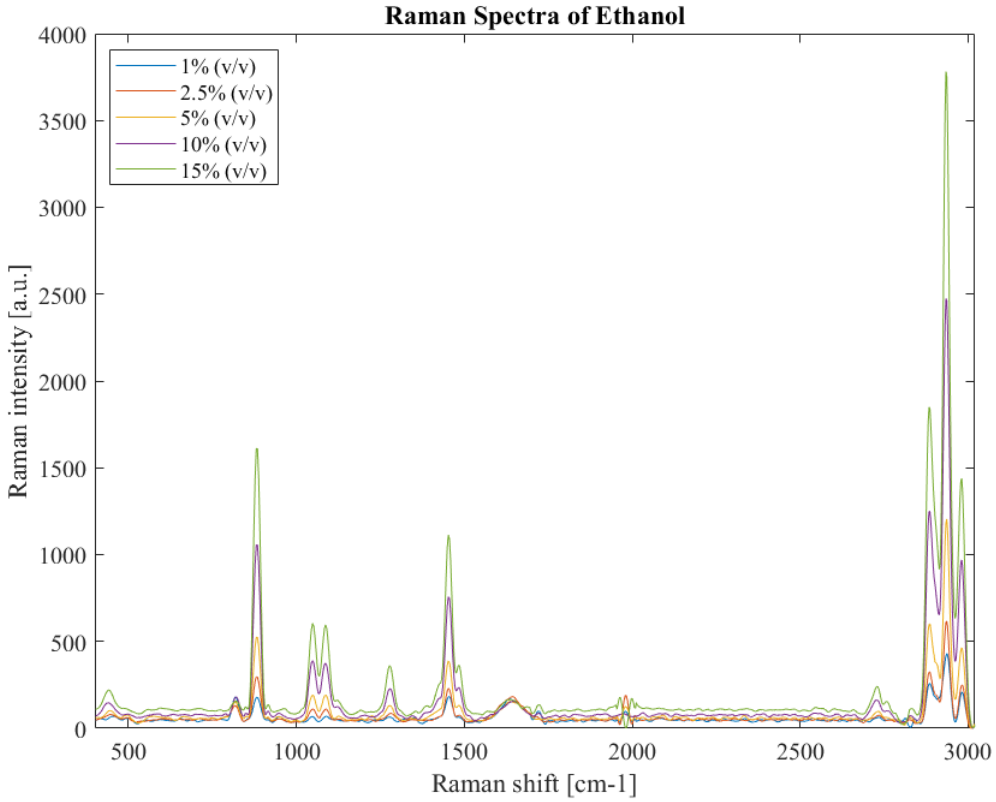

Figure 50. Raman spectra of aqueous ethanol solutions at different concentrations.

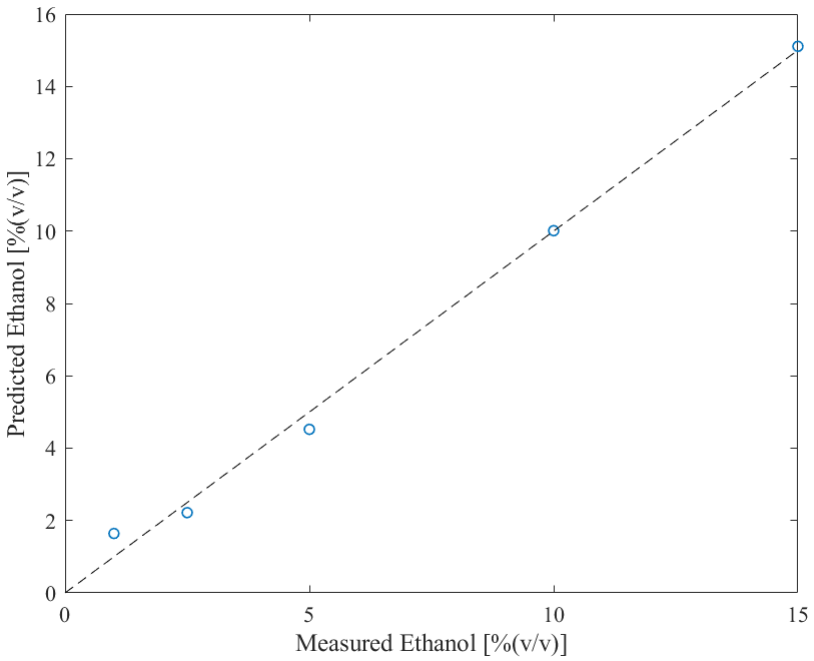

Figure 51. Comparison of predicted and measured values of ethanol with 1 LV PLSR model.

Figure 51 represents the result of the PLS regression model calculated with only one latent variable (LV) and using all available data. In this case, an R^2 value of 0.9945 is obtained.

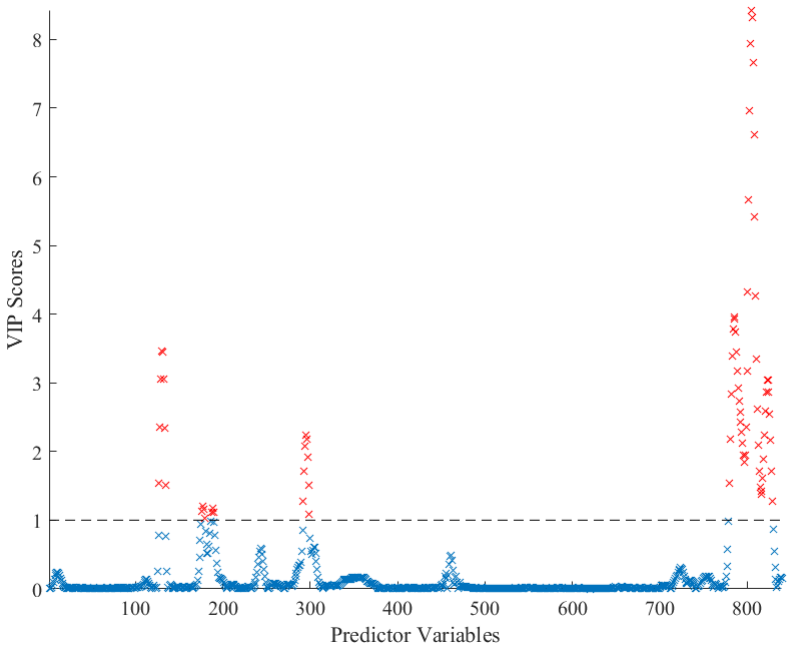

Figure 52. VIP scores for ethanol (73 selected variables).

Application of the VIP method revealed that 73 predictor variables (shown in red in Figure 52) had a VIP score greater than 1, and thus had a greater influence on ethanol composition (y). These selected data constituted the new **X** data matrix, which was used to build a new PLSR model as shown in Figure 53. It can be observed that the result obtained with the reduced data set was equal to previous one calculated with the whole data set. This was also confirmed by the R^2 value obtained (0.9945). Therefore, in the case of ethanol, the selection of the most significant predictor variables did not bring any improvement in the predictive capabilities of the model. In any case, this represents a very satisfactory result, since the Raman spectra used confirmed to have a high informative content for the purpose of evaluating ethanol composition and obtaining an excellent predictive model.

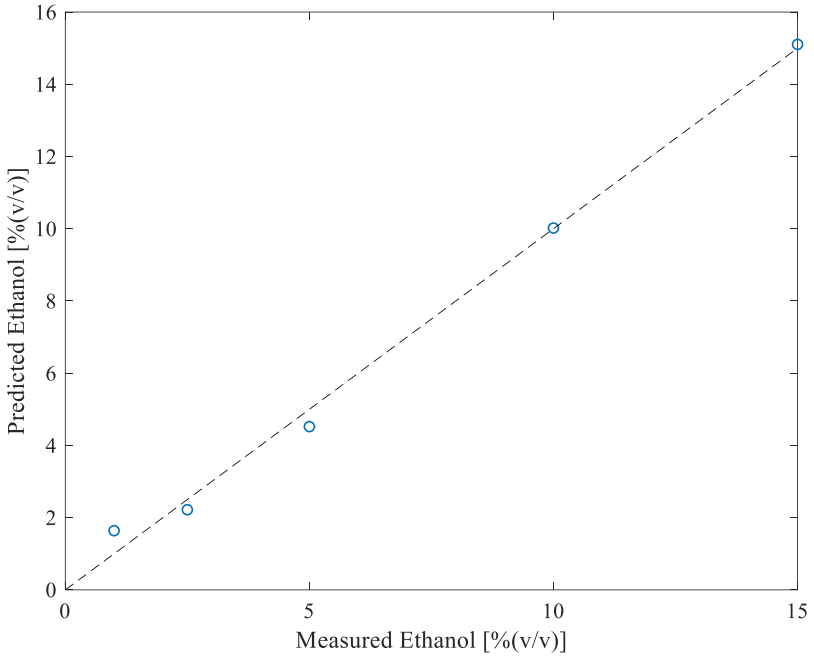

Figure 53. Comparison of predicted and measured values of ethanol with 1 LV PLSR model obtained with the reduced data set.

Glucose

Regarding glucose, the Raman spectra for different concentration values are reported in Figure 54. As can be observed, its characteristic peaks are significantly weaker in terms of intensity and also overwhelmed by those of water, especially at 1633 cm^{-1}, which represents the characteristic peak of H-O-H bending. However, some of its characteristic peaks can be identified, such as the one at 451 cm^{-1} of the C-CO bending, and the one at 1063 cm^{-1} of the C-O bending. An anomalous and particularly intense peak was recorded around 2000 cm^{-1}. The literature was not helpful, but it interesting to note that this peak is present in all the solutions analyzed.

Figure 55 shows the comparison between measured and predicted glucose compositions from the PLS regression model obtained with 3 latent variables. In this case, the R^2 coefficient was 0.9807. Figure 56 represents the distribution of selected predictor variables from the VIP method. Those considered significant were 131, which were then used to build the new PLSR model (Figure 57). With the reduction of the data set, the R^2 coefficient decreased to 0.9567. This means that, unlike what was observed for ethanol, for glucose composition the variable selection resulted in a loss of information that led in loss of predictive ability of the model, as confirmed by the R^2 value. Therefore, although a decrease in the total variance explained by the model of about 2.4% was recorded, the results obtained showed that the information

contained in the evaluated spectra was significant and allowed for the construction of a statistical model with satisfactory performance.

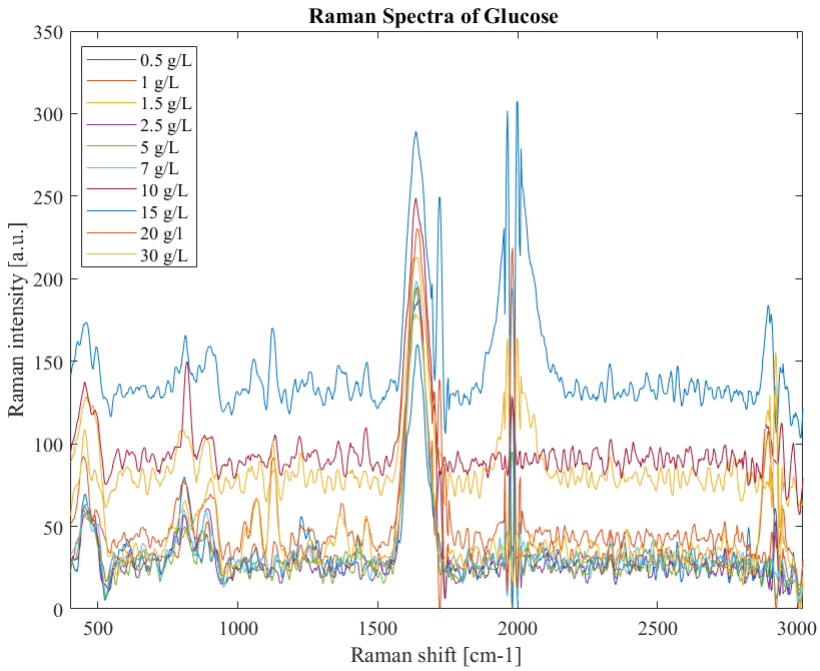

Figure 54. Raman spectra of aqueous glucose solutions at different concentrations.

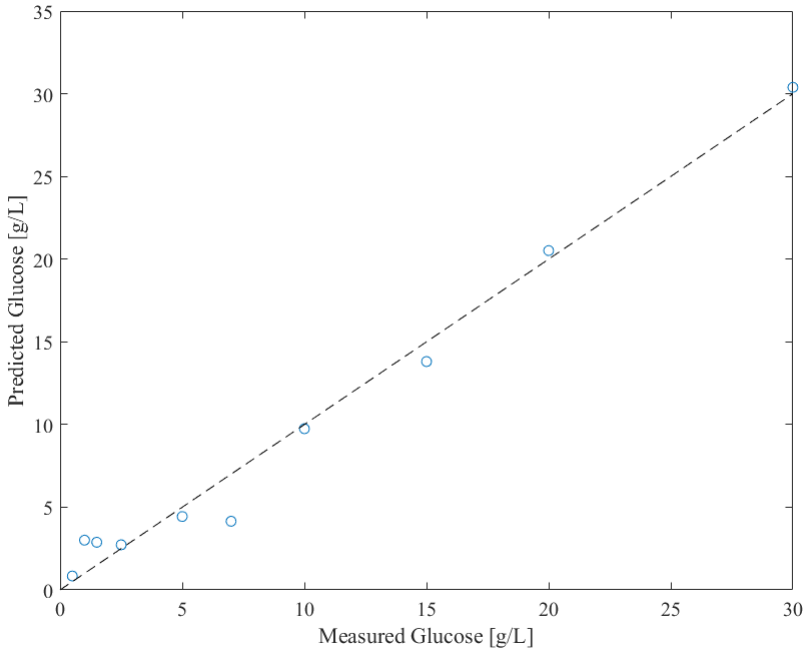

Figure 55. Comparison of predicted and measured values of glucose with 3 LV PLSR model.

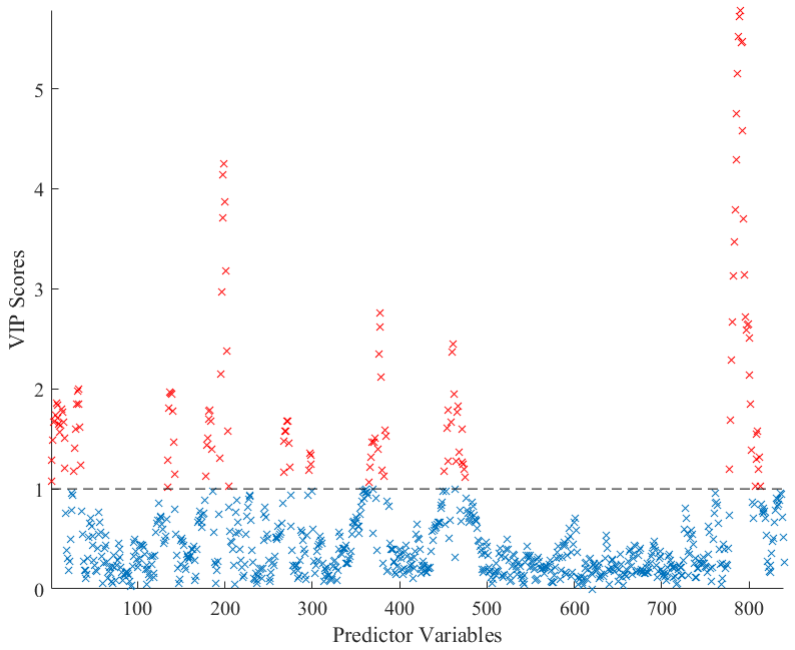

Figure 56. VIP scores for glucose (131 selected variables).

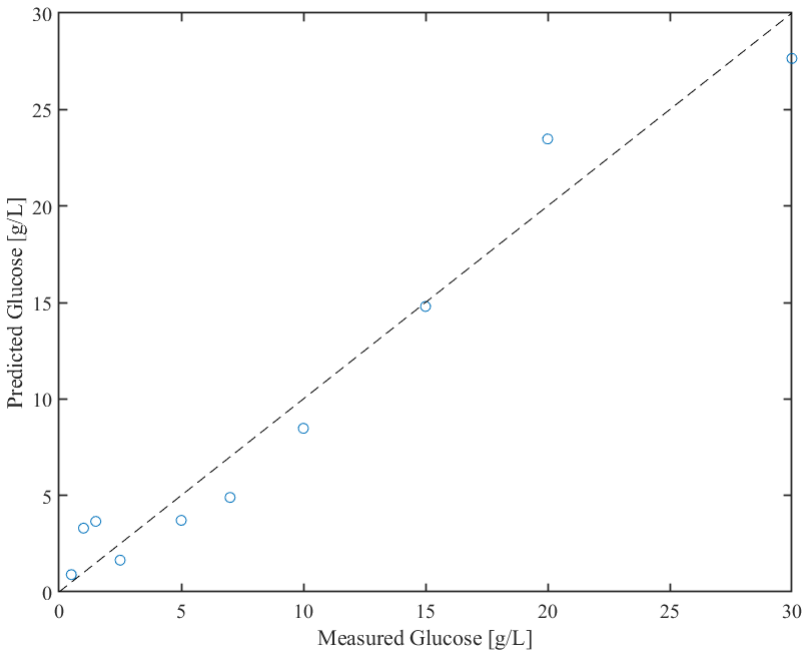

Figure 57. Comparison of predicted and measured values of glucose with 3 LV PLSR model obtained with the reduced data set.

Ethanol in presence of glucose and secondary substances

In this section, the construction of a PLSR model for ethanol was evaluated in the case where glucose and secondary substances, such as acetic acid and formic acid are present in the analyzed mixture. Therefore, in addition to the Raman spectra acquired for the previously investigated ethanol solutions, the not pure ethanol solutions shown in Table 21 were also included in this case. Thus, a total of 12 samples were analyzed. Figure 58 shows the corresponding Raman spectra recorded for these solutions. The evaluation of a model with only one latent variable (Figure 59) returned an R^2 value of 0.8827, which changed to 0.8819 (Figure 61) after the data set was reduced due to the selection of significant predictor variables (Figure 60). By adding a latent variable, improvements were achieved. Indeed, the total variance explained by the model reached a value of 0.9930 (Figure 62). Then, the application of VIP method allowed for the selection of 79 significant variables (Figure 63), compared to 73 with only one latent variable. The model built with the reduced data set enabled a further improvement to be achieved, in fact the R^2 value was 0.9939 (Figure 64). The results obtained demonstrate once again that Raman spectra are highly informative for estimating the composition of ethanol. A performant model with more input data, but still a low value of components, was developed, which allowed for performance almost equal to the first one obtained (with only aqueous ethanol solutions). It is worth noting that compared to the case just mentioned, the model with 2 latent variables identified 6 more predictor variables. These included wavelengths between 810 and 820 cm^{-1} where characteristic peaks can be found due to C-C stretch, around 1090 cm^{-1} where peaks can be found due to CH_3 oscillations, around 2868 cm^{-1} for CH_2 vibrations, and finally approximately at 2957 cm^{-1} for CH_3 stretch. These 6 variables proved to be significant, it is true that R^2 increased by only 0.09%, but still it represents an improvement.

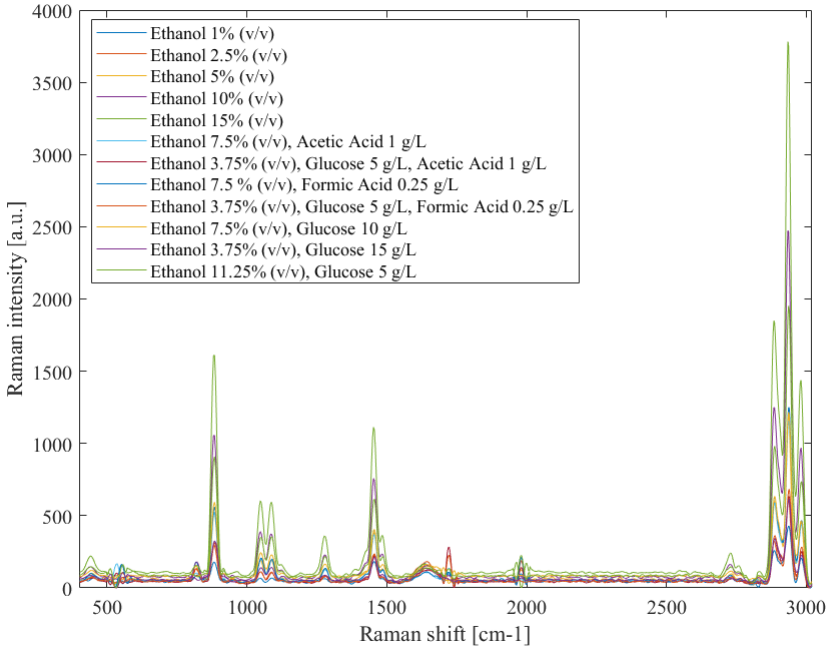

Figure 58. Raman spectra of aqueous ethanol solutions in presence of glucose, acetic and formic acid at different concentrations.

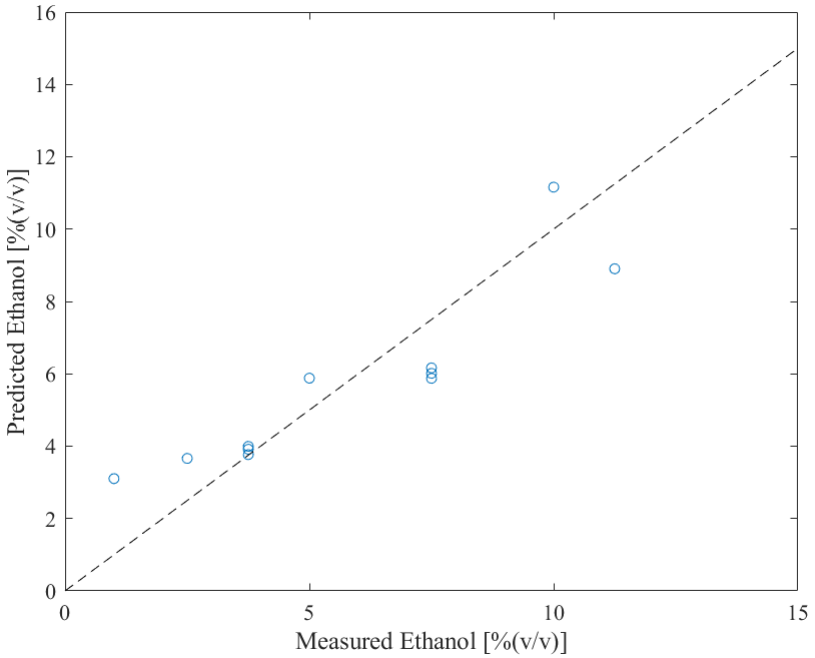

Figure 59. Comparison of predicted and measured values of ethanol with 1 LV PLSR model.

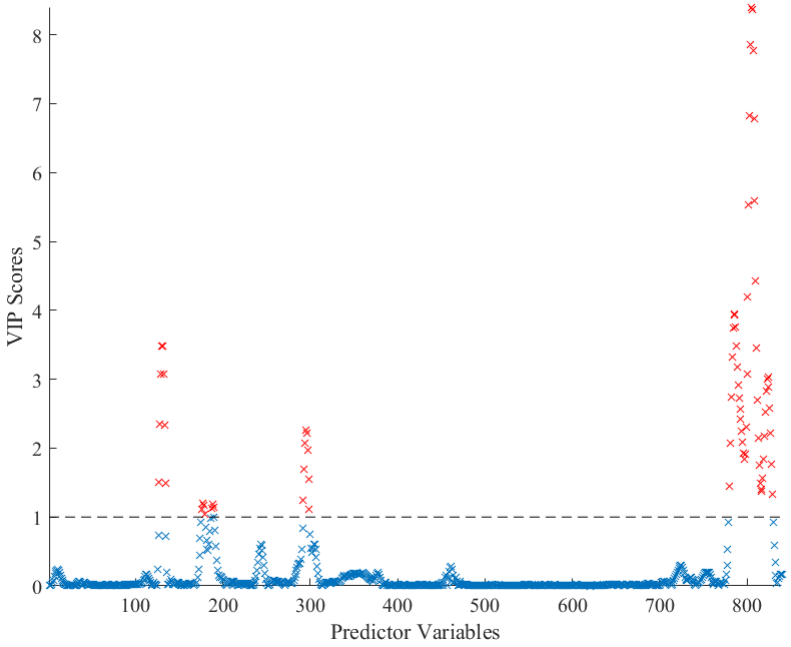

Figure 60. VIP scores for ethanol (73 selected variables).

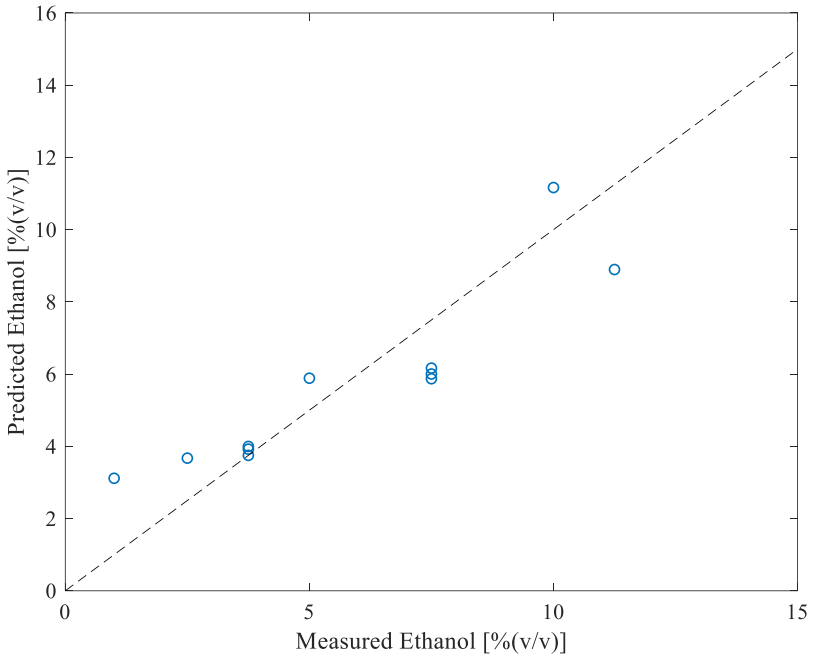

Figure 61. Comparison of predicted and measured values of ethanol with 1 LV PLSR model obtained with the reduced data set.

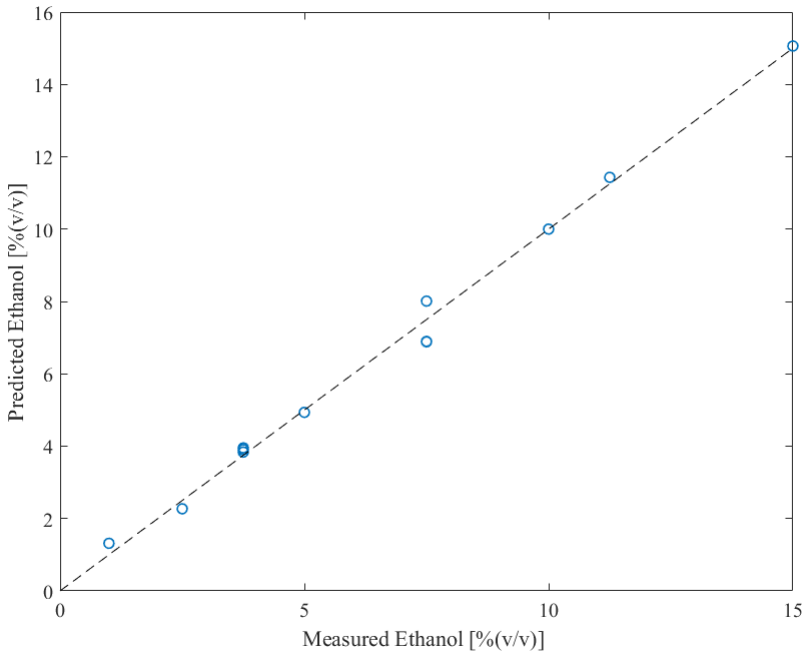

Figure 62. Comparison of predicted and measured values of ethanol with 2 LV PLSR model.

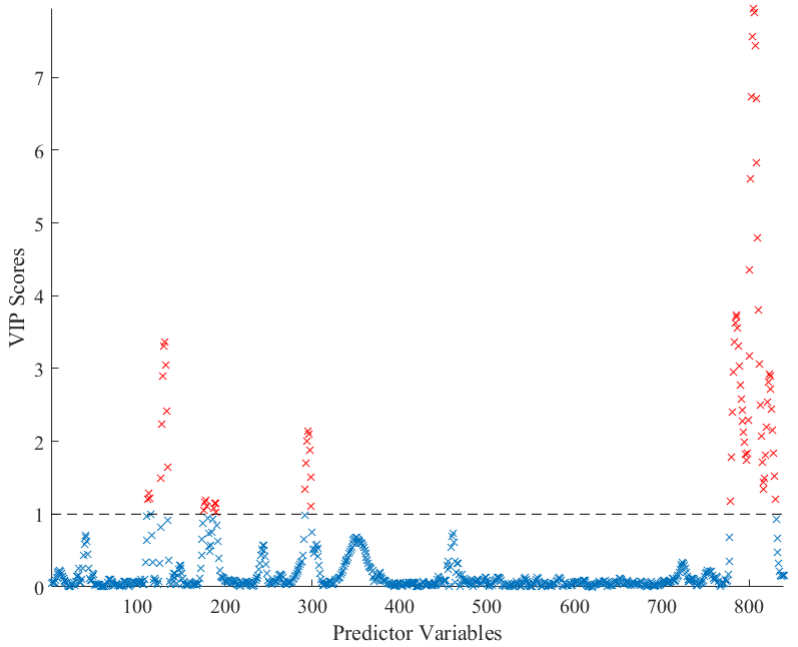

Figure 63. VIP scores for ethanol (79 selected variables).

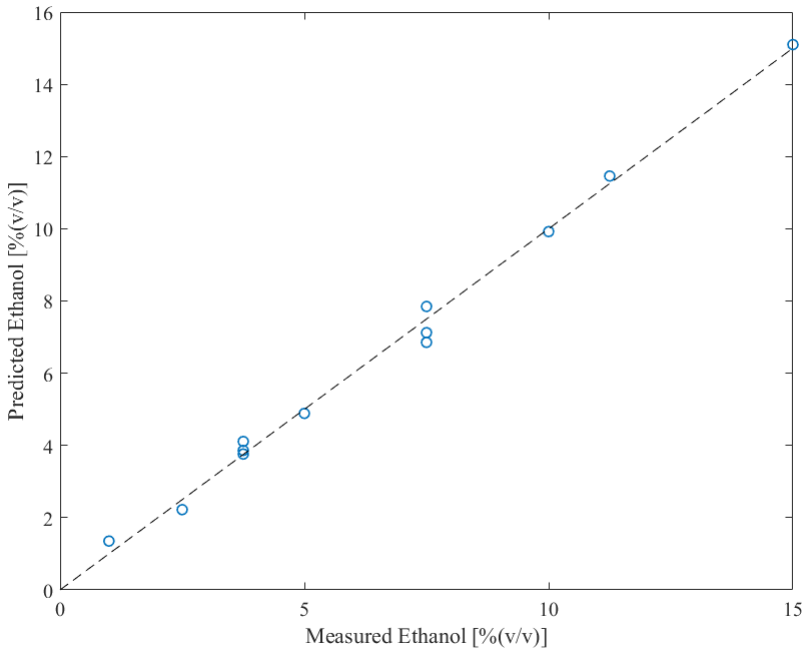

Figure 64. Comparison of predicted and measured values of ethanol with 2 LV PLSR model obtained with the reduced data set.

Glucose in presence of ethanol and secondary substances

In this section, the development of the glucose prediction model was investigated in the case where, in addition to glucose-only solutions, combined solutions of the same were included along with ethanol, acetic acid, and formic acid. Again, the reader considers the combinations shown in Table 21. In this case, there were 17 total solutions analyzed.

Figure 66 shows the comparison of the measured glucose with that predicted by the model obtained with 4 latent variables. In this case, an R^2 coefficient of 0.9587 was obtained. In Figure 67, the trend of the most significant predictor variables is reported. The identification of 147 variables, and the consequent model built on them, resulted in a decreasing in terms of the total variance explained by the model (Figure 68). Similar result was obtained in the case of glucose-only solutions. Considering one more latent variable, the model construction resulted in an increase of 0.01 of R^2 (Figure 69). In this case, 148 significant variables were identified (Figure 70), and the regression model calculated on the reduced data set gave an R^2 coefficient of 0.9574 (Figure 71). Compared with the case of glucose-only solutions (Figure 54), 17 more significant variables were identified. Looking at the graph in Figure 70, it is possible to observe that the differences are especially around 600 and 700 cm-^{1}, and then

between 2800-2900 cm^{-1}. In the first mentioned area it is possible that the method considered the presence of acids as significant, while in the second area it selected the characteristic peaks of ethanol due to CH_3 stretch, which in contrast to those of glucose are particularly intense.

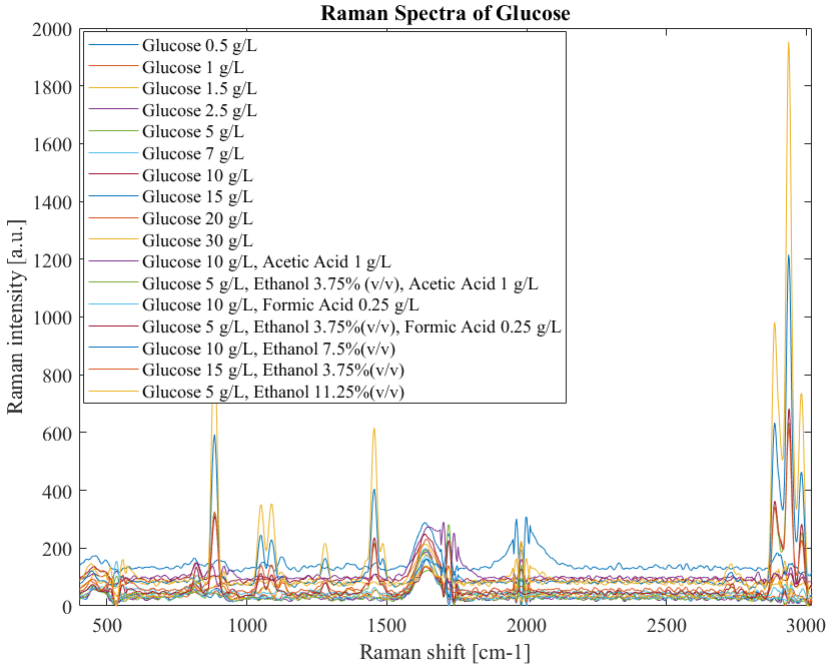

Figure 65. Raman spectra of aqueous glucose solutions in presence of ethanol, acetic and formic acid at different concentrations.

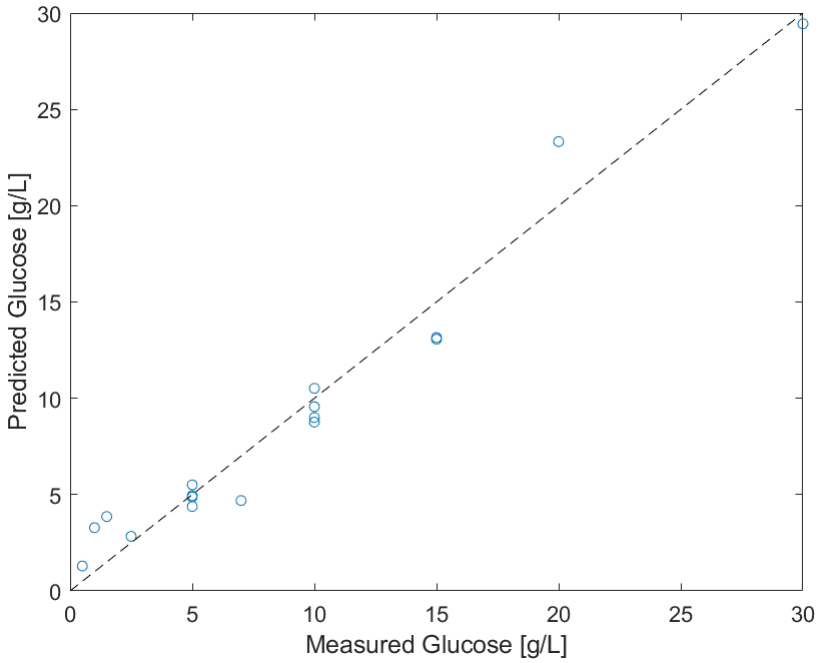

Figure 66. Comparison of predicted and measured values of glucose with 4 LV PLSR model.

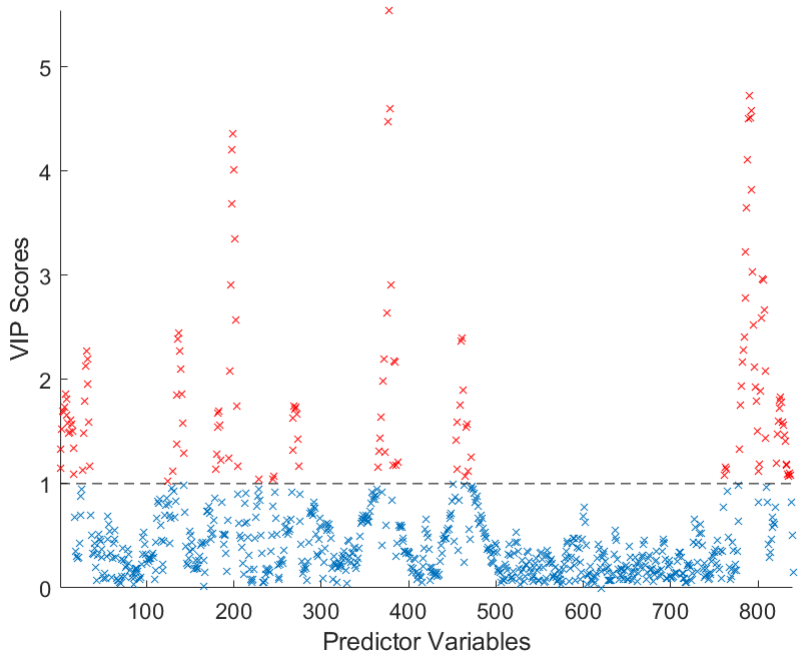

Figure 67. VIP scores for glucose (147 selected variables).

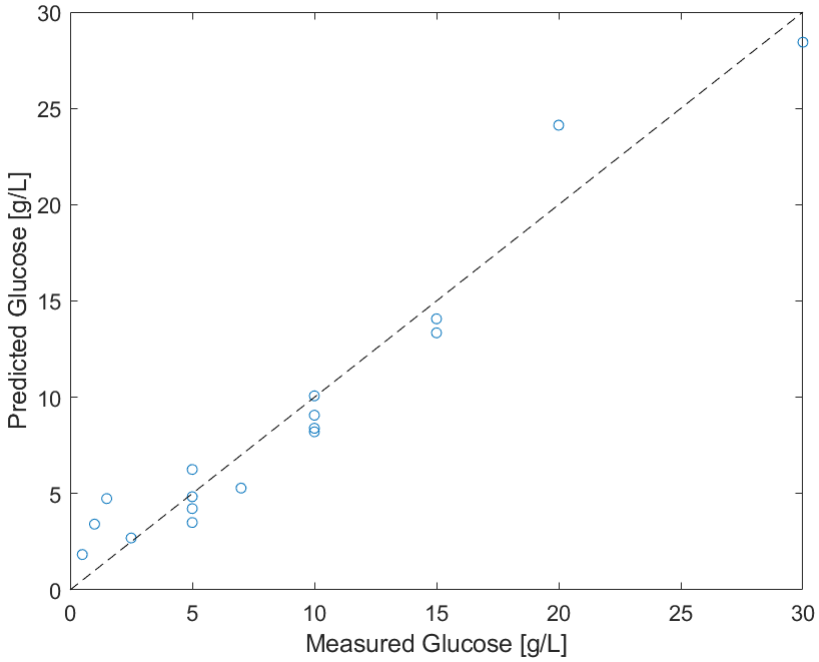

Figure 6*8*. Comparison of predicted and measured values of glucose with 4 LV PLSR model obtained with the reduced data set.

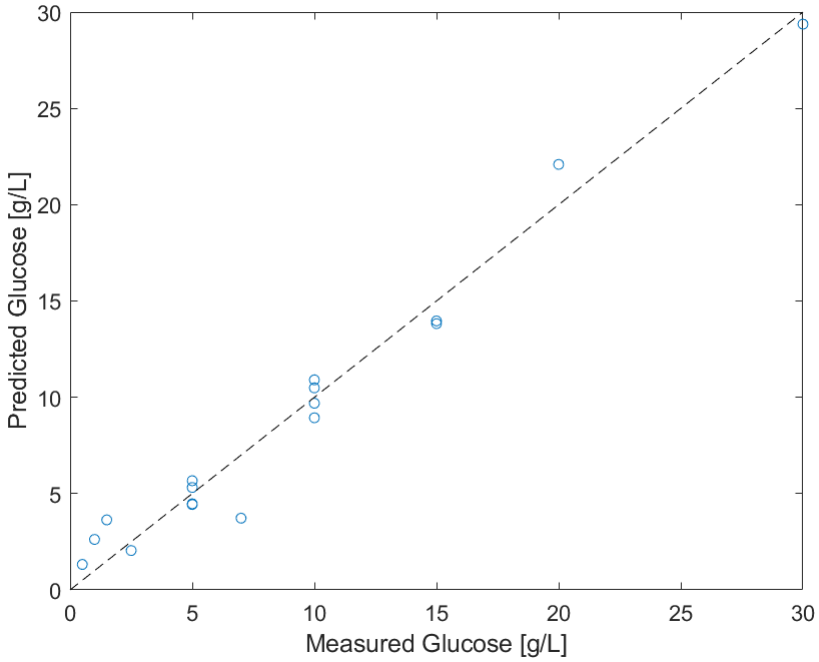

Figure 69. Comparison of predicted and measured values of glucose with 5 LV PLSR model.

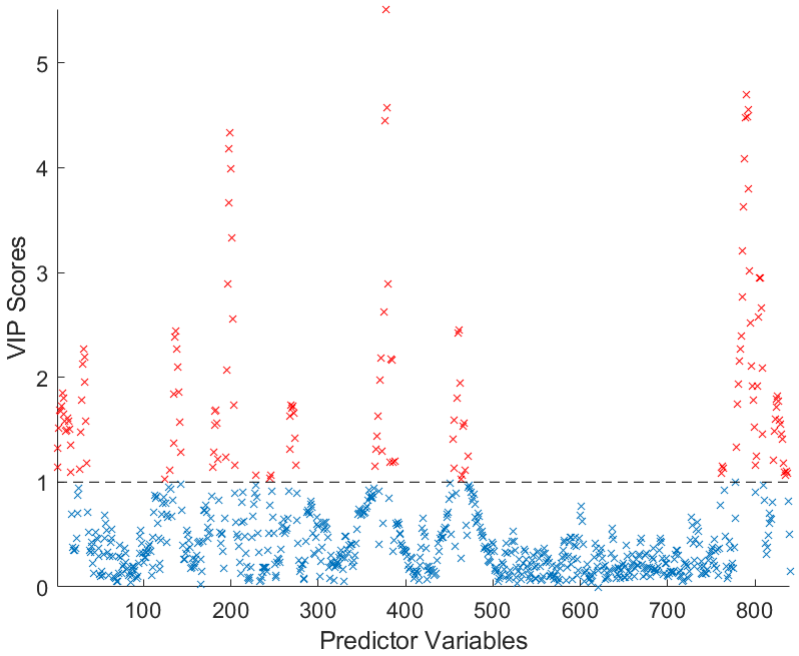

Figure 70. VIP scores for glucose (148 selected variables).

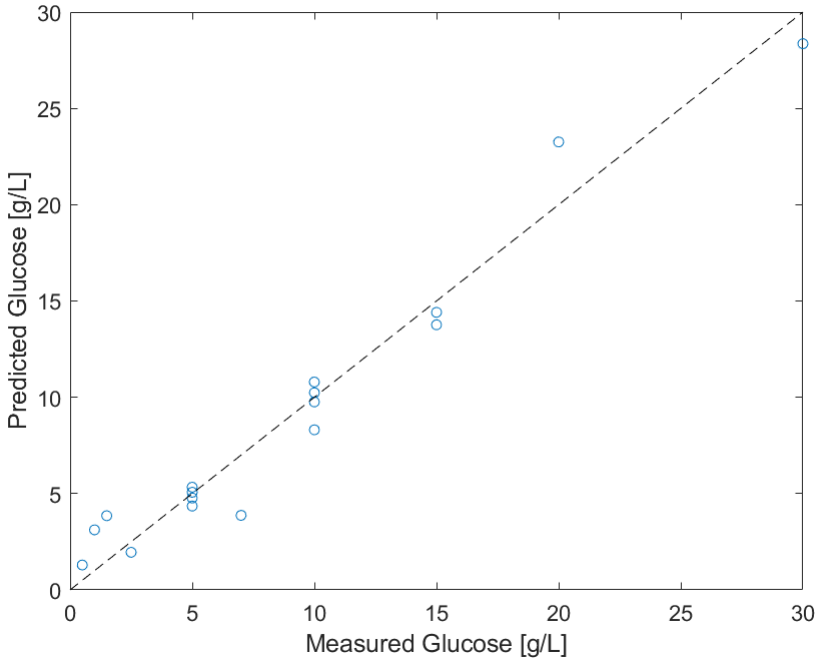

Figure 71. Comparison of predicted and measured values of glucose with 5 LV PLSR model obtained with the reduced data set.

5.3.2 Not-ideal fermentation solutions

This section will discuss the results obtained from the Raman analysis of the fermentation samples. In particular, it is important to note that samples taken from both bioreactors were used to develop the regression model. Therefore, a total of 20 samples were used to develop the PLSR model of both ethanol and glucose. The Raman spectra of these samples are reported in Figure 72. As can be seen, the spectra were considered in the full Raman shift range, as a reduction of it proved to have a negative effect on the predictive capabilities of the models.

Ethanol (with only fermentation solutions)

First, the model development for ethanol composition with 5 latent variables was evaluated. Figure 73 shows the comparison between the measured ethanol compositions and those predicted by the model. An R^2 coefficient of 0.6567 was obtained. Then, 149 variables were selected as significant (Figure 74). The regression conducted on the reduced data set improved the performance of the model by 7.63%. Indeed, the new R^2 value was 0.7330. This means that the VIP method selected precisely the most significant variables. This is an encouraging result, because as can be observed in Figure 75, the points fall close to the straight line. However, the same cannot be said in validation. The results regarding this step are shown in Figure 76. The

validation was carried out on the previously reduced data set by means of VIP method. The new regression model obtained exhibited poor predictive ability, having a negative R^2 coefficient and equal to -0.3544. Although from the graph some points appear to be quite close to the straight line, the negative R^2 value may be due to the presence of a negative predicted value of composition, which certainly affects it significantly.

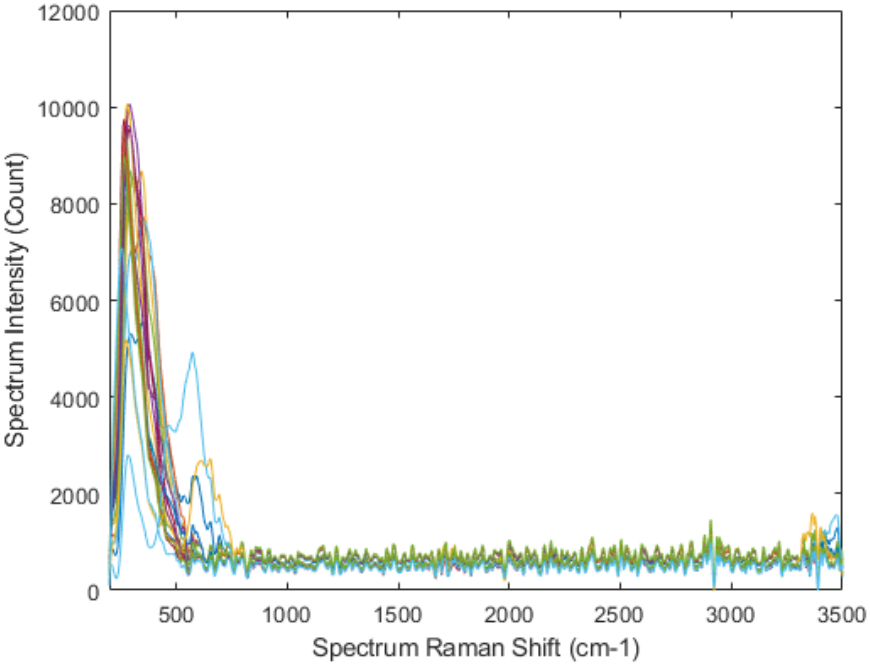

Figure 72. Raman spectra recorded for the 20 fermentation samples.

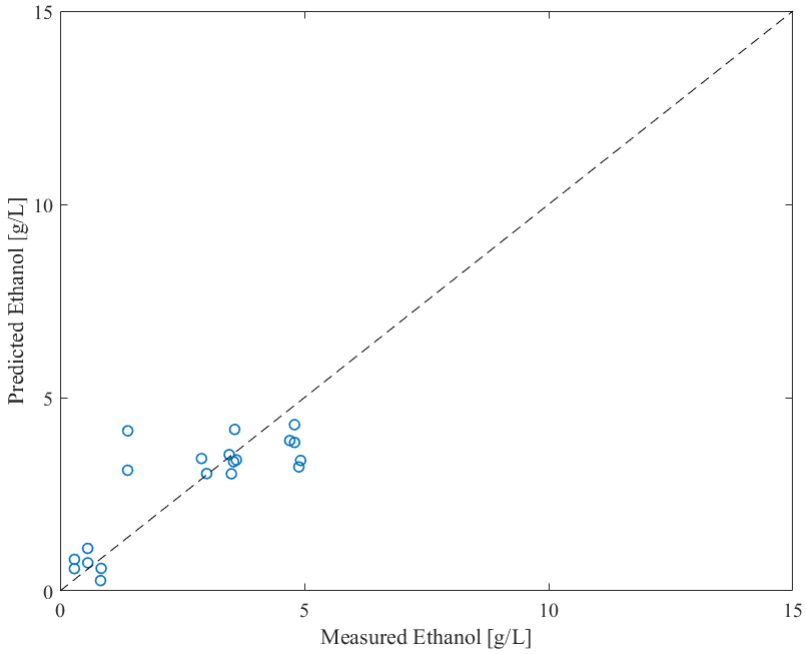

Figure 73. Comparison of predicted and measured values of ethanol with 5 LV PLSR model.

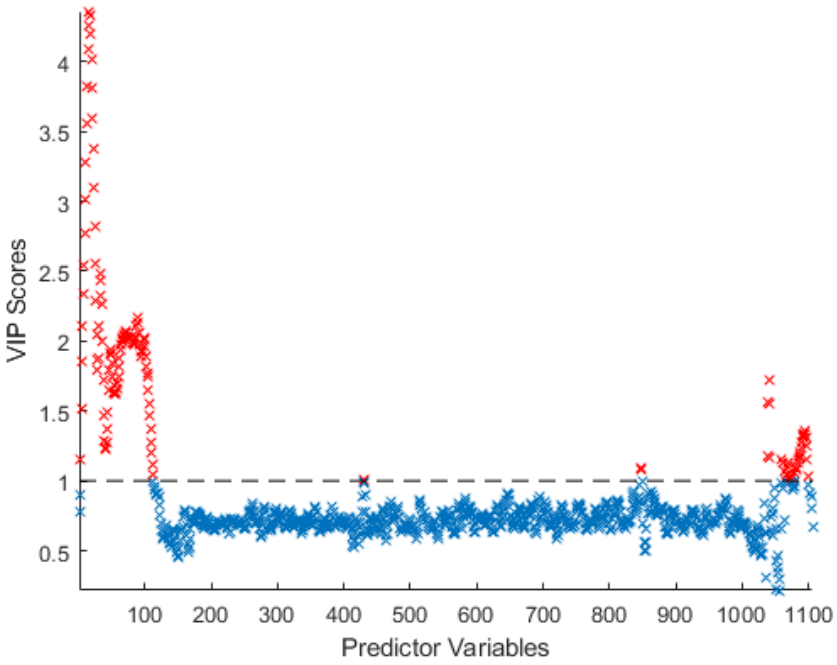

Figure 74. VIP scores for ethanol (149 selected variables).

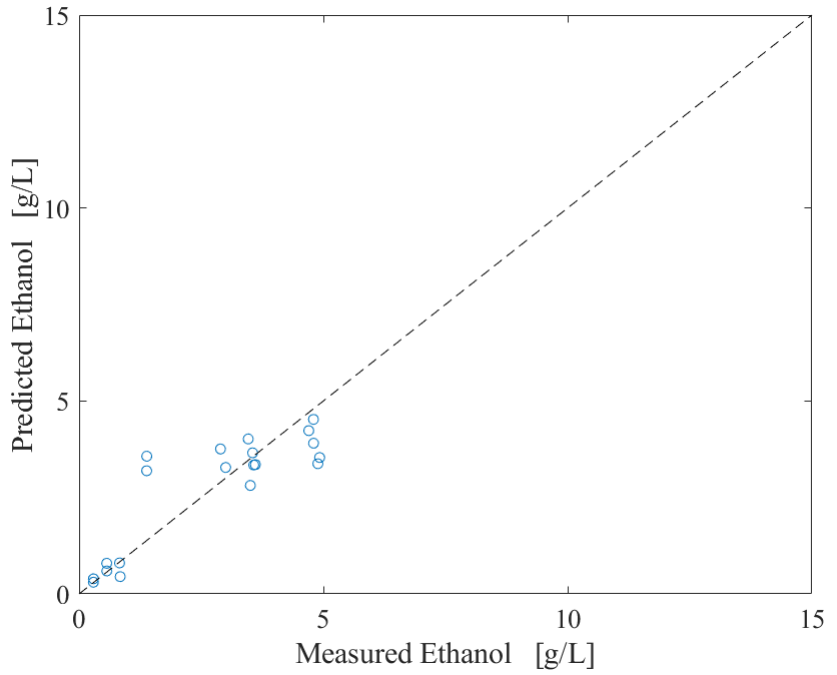

Figure 75. Comparison of predicted and measured values of ethanol with 5 LV PLSR model obtained with the reduced data set.

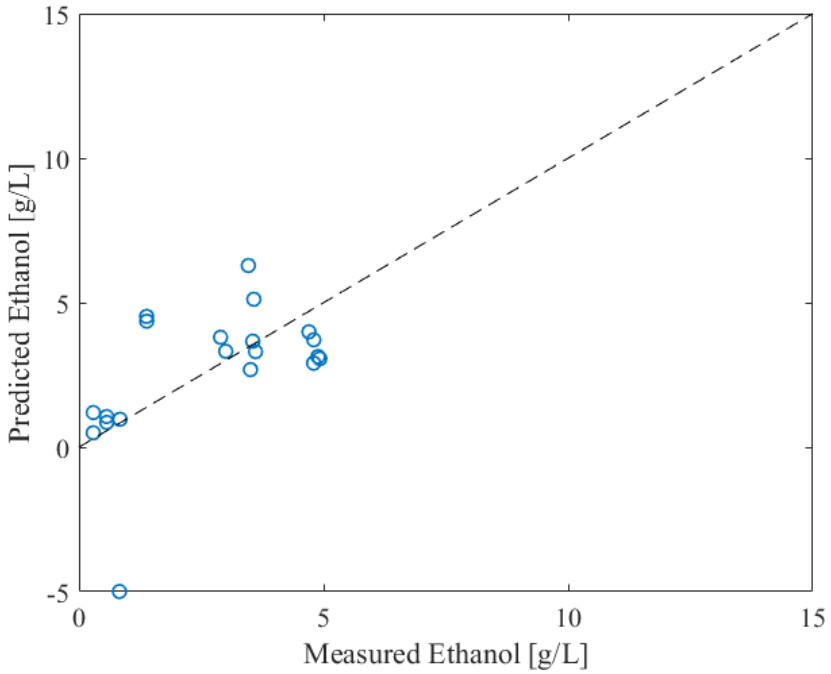

Figure 76. Comparison of predicted and measured values of ethanol with 5 LV PLSR model obtained with the reduced data set in validation phase.

Because of the results obtained with 5 variables, the regression model for ethanol was developed considering one more latent variable. The results obtained are shown in Figure 77-80. In this case, the PLSR model provided a slightly higher R^2 coefficient in calibration (0.6890) than in the previous case (0.6567). As also noted above, the regression performed on the reduced data set as a result of selecting the most significant variables (Figure 78), determined an improvement in R^2 value of about 7.6%. The comparison of measured values and those predicted by this model are depicted in Figure 79. Again, moving on to the validation step, substantial decreases were recorded in terms of model performance. The total variance explained by the model was found to be 0.3141. In comparison with the case of 5 latent variables, it is possible to state that the model appears to be more capable of predicting ethanol composition values (Figure 80). The information contained in the reduced data set used in this validation phase allowed for a better, but still not satisfactory result.

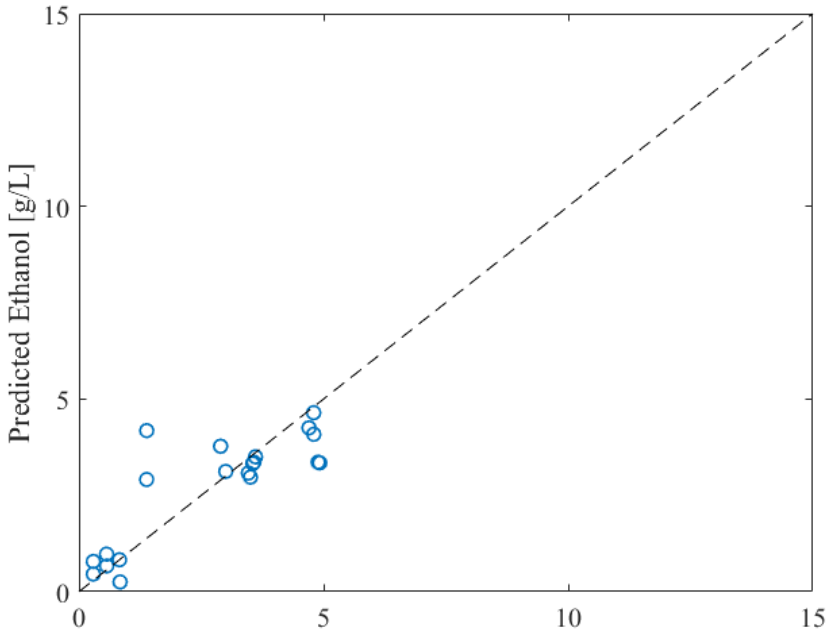

Figure 77. Comparison of predicted and measured values of ethanol with 6 LV PLSR model.

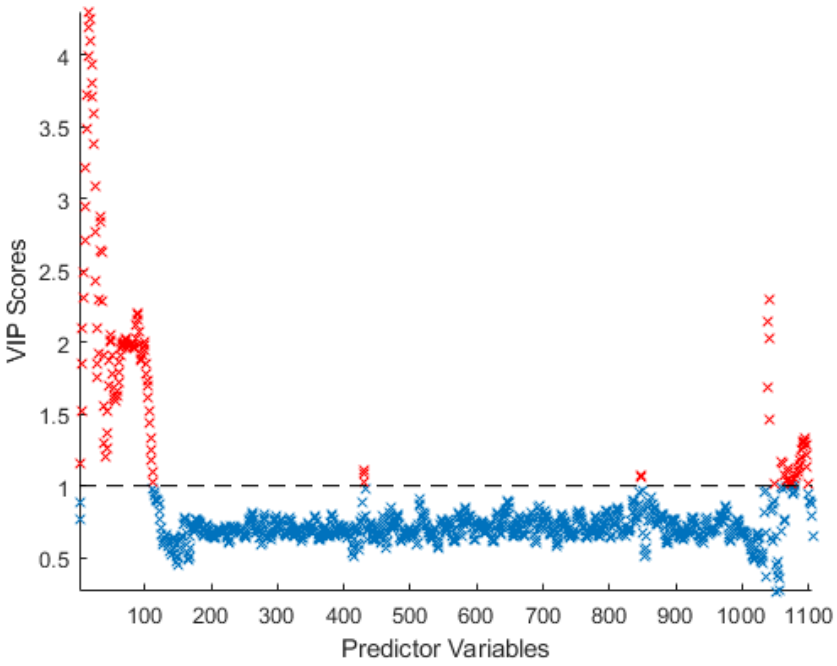

Figure 78. VIP scores for ethanol (152 selected variables).

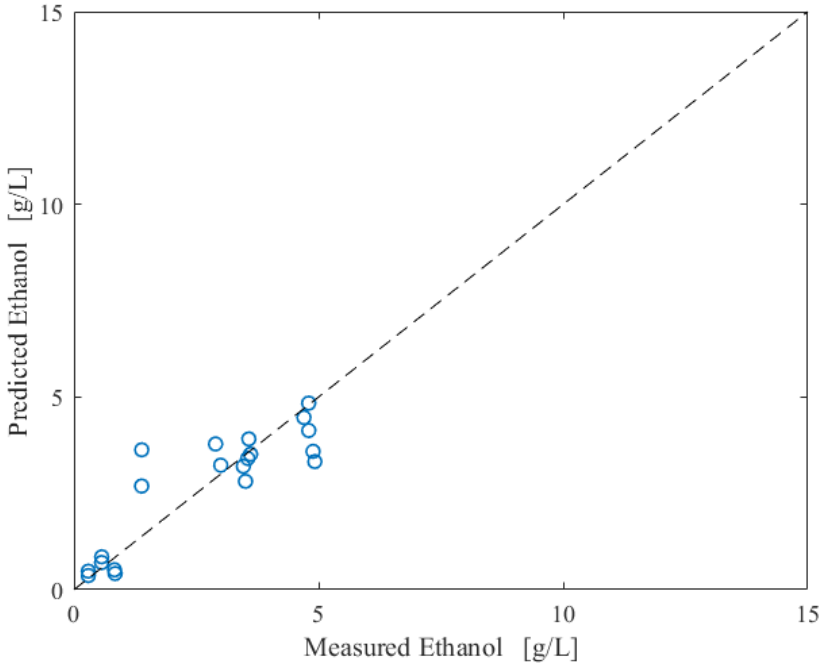

Figure 79. Comparison of predicted and measured values of ethanol with 6 LV PLSR model obtained with the reduced data set.

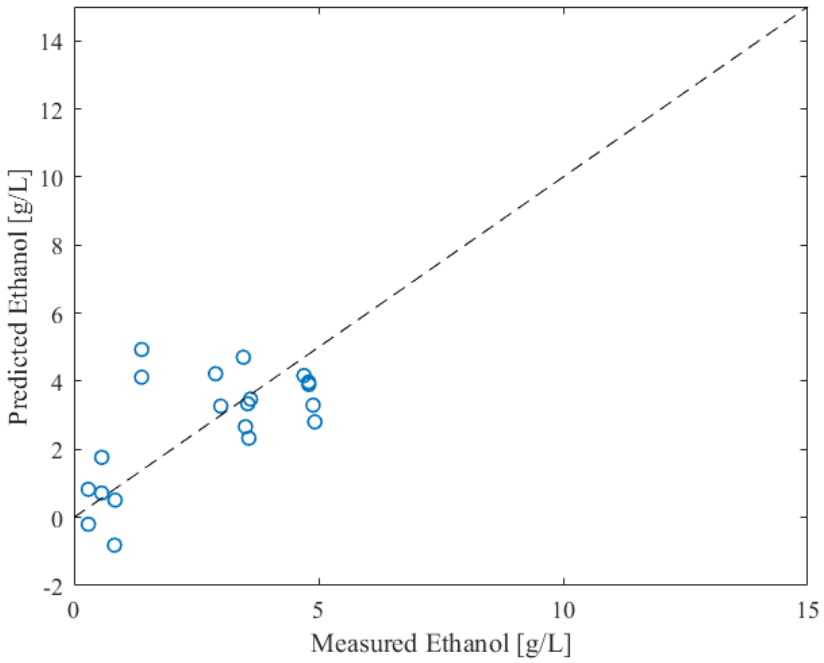

Figure 80. Comparison of predicted and measured values of ethanol with 6 LV PLSR model obtained with the reduced data set in validation phase.

Glucose (with only fermentation solutions)

Similarly to the case of ethanol, a model was first developed for glucose by evaluating 5 latent variables. The R^2 obtained at the calibration phase (0.7187), both before and after VIP selection (120 selected variables), was lower than the one obtained from the analysis of ideal solutions. This can be clearly observed from graphs reported in Figure 81 and 83, where the comparison of measured and predicted values, with and without the reduced data set, is reported. Although the selection of 120 significative variables (Figure 82) resulted in an increase in R^2 (0.7664), this is still insufficient. As a confirmation of this, the comparison between measured and predicted values obtained in the validation step is reported in Figure 84. In this case, R^2 coefficient was significantly negative (-0.9844).

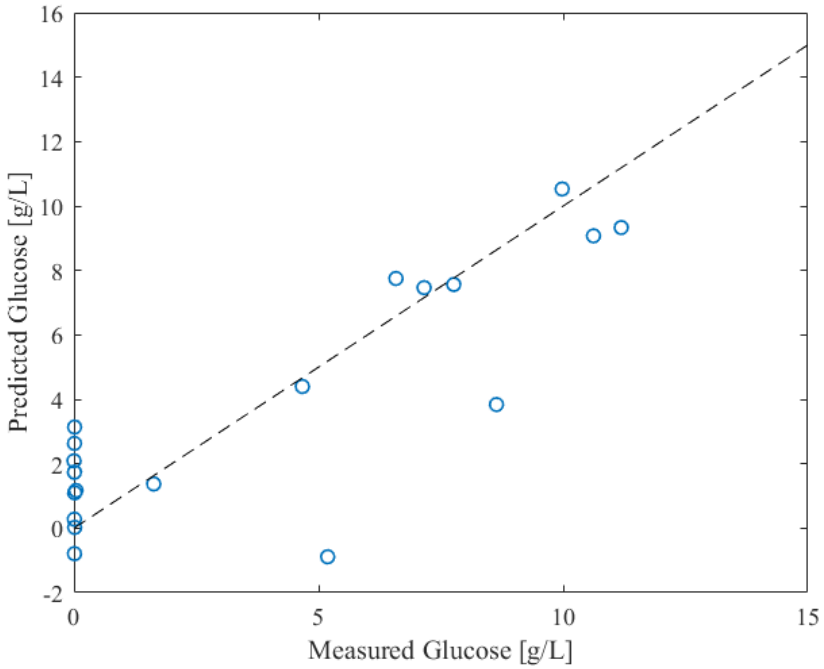

Figure 81. Comparison of predicted and measured values of glucose with 5 LV PLSR model.

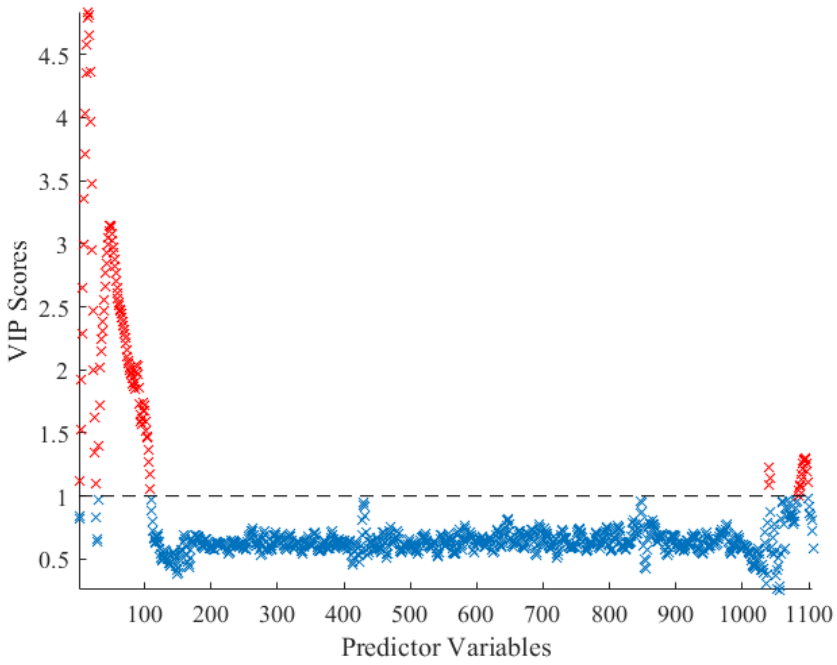

Figure 82. VIP scores for glucose (120 selected variables).

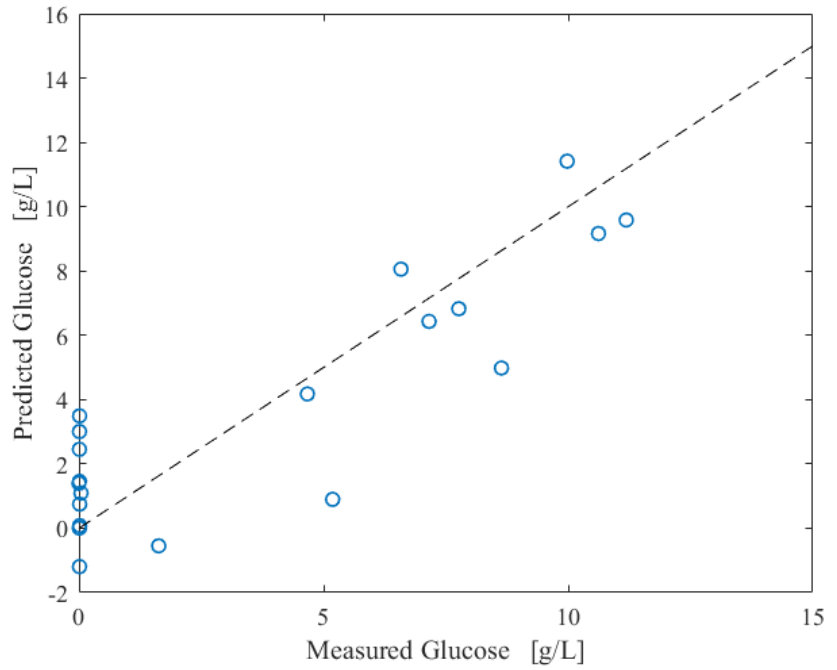

Figure 83. Comparison of predicted and measured values of glucose with 5 LV PLSR model obtained with the reduced data set.

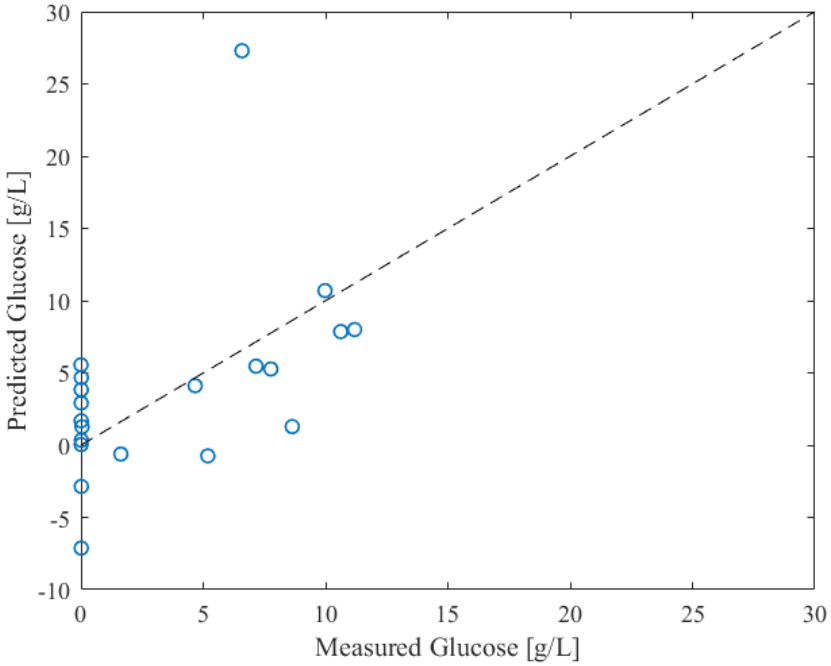

Figure 84. Comparison of predicted and measured values of glucose with 5 LV PLSR model obtained with the reduced data set in validation phase.

Due to the results obtained with 5 latent variables, a regression model with 6 latent variables was also developed for glucose. The results in calibration and validation are reported below. R^2 coefficient under calibration assumed a slightly higher value than the previous case (0.7324). The model still showed difficulties in predicting glucose composition. Indeed, as can be seen in Figure 85, only some values fall close to the straight line, while others deviate significantly assuming even negative values. 123 significant variables were selected as shown in Figure 86. The regression conducted on the reduced data set allowed for a better R^2 and equal to 0.8056 (Figure 87). The validation conducted in the same reduced data set did not give the hoped-for results. An R^2 coefficient of 0.0256 was obtained. Although compared to the previous case, the improvement is evident, from Figure 88 it is possible to see that the glucose model still has significant limitations.

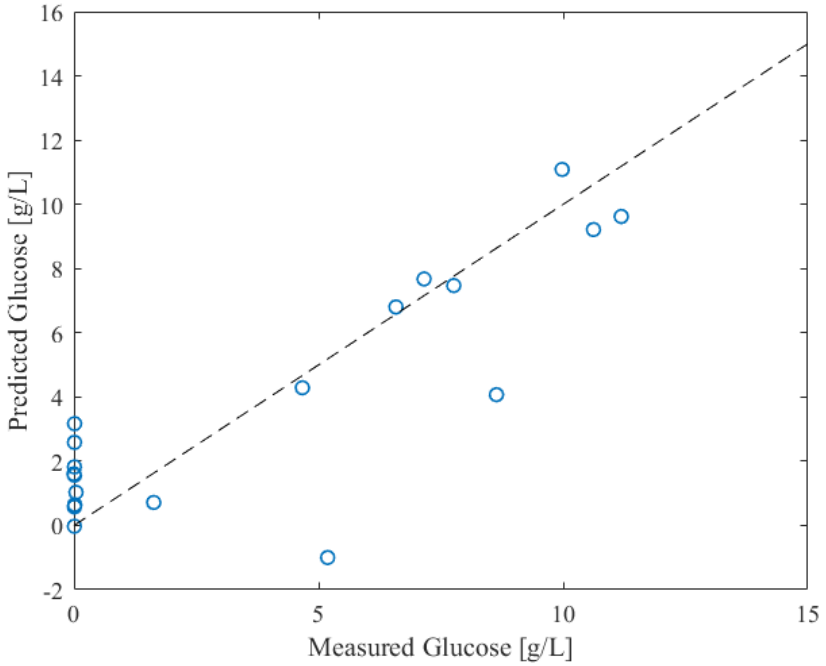

Figure 85. Comparison of predicted and measured values of glucose with 6 LV PLSR model.

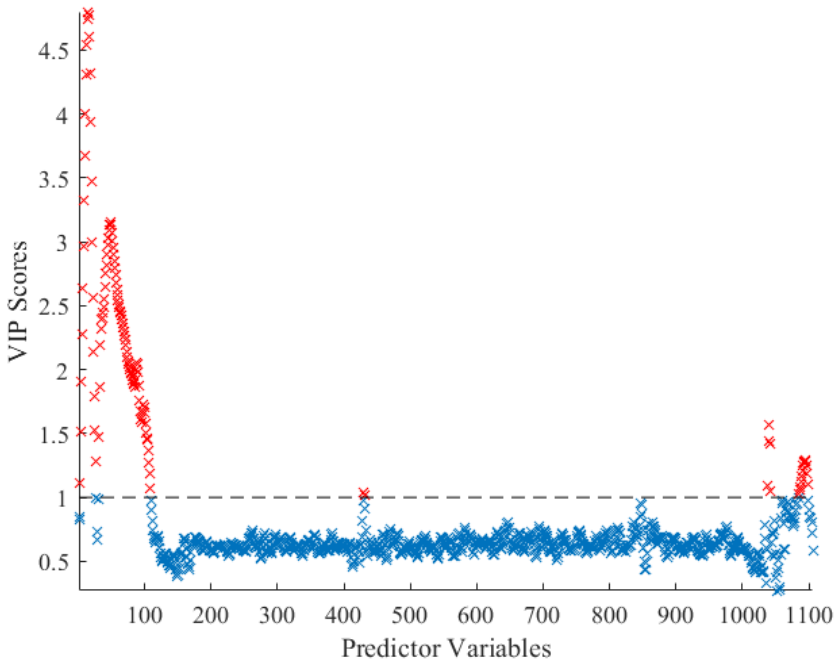

Figure 86. VIP scores for glucose (123 selected variables).

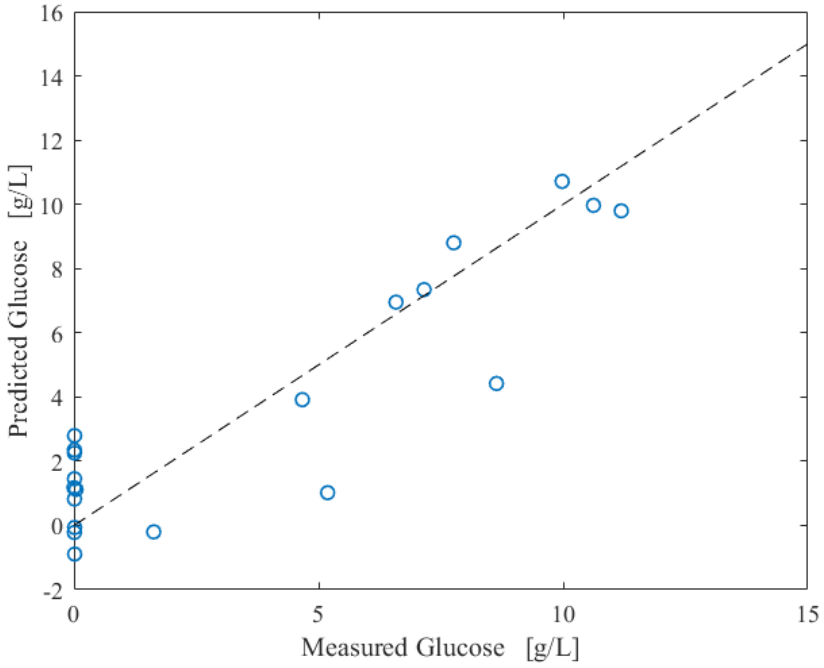

Figure 87. Comparison of predicted and measured values of glucose with 6 LV PLSR model obtained with the reduced data set.

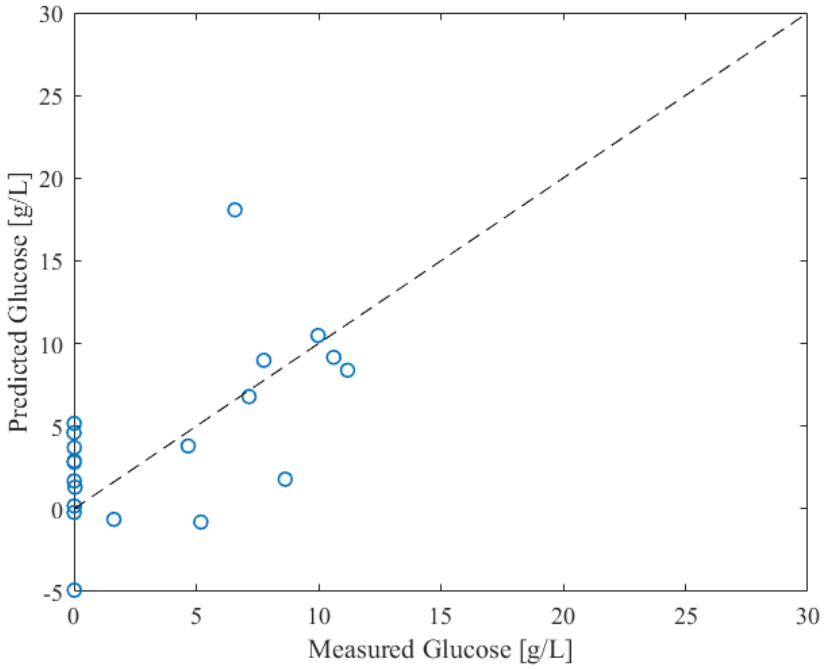

Figure 88. Comparison of predicted and measured values of glucose with 6 LV PLSR model obtained with the reduced data set in validation phase.

5.4 Conclusions

Monitoring of key variables in a biotechnological process is certainly a key step in control. As outlined in previous chapters, this represents a nontrivial task. The benefits to be gained from implementing a monitoring system are diverse, such as improvements in product quality, improved safety, and lower operating costs. The lack of sensors capable of providing real-time information represents the greatest limitation to this goal. In recent decades, the use of spectroscopic analyzers has become increasingly popular in the context of bioprocess monitoring. The ability to provide near-real-time results, the fact that no chemicals are needed to operate, and the low maintenance required, along with the fact that they do not require time for sample preparation represent the main advantages of these instruments. In this chapter, the application of Raman spectroscopy was investigated and discussed as a monitoring tool for bioethanol production in fermentation bioreactors. The goal was to develop a reliable and accurate sensor to measure ethanol and glucose composition. Chemometric techniques were evaluated and applicated in order to determine a relationship between the recorded Raman spectra and the concentrations of the investigated compounds. The study was first conducted on ideal solutions of ethanol and glucose at known concentrations. Additional combined solutions of ethanol, glucose, and acetic and formic acid were added to these in order to assess the presence of possible undesirable compounds. Subsequently, samples of fermentation solutions were analyzed using BSG as biomass. Raman data from all considered samples were statistically analyzed using the PLS technique. With regard to the ad hoc prepared solutions, the results are encouraging. Even in the presence of undesirable compounds that might disturb the reading of the peaks characteristic of those investigated, both the ethanol and glucose models were shown to have good predictive capabilities. In the case of nonideal fermentation solutions, however, the evaluation was far from trivial. Regarding ethanol, its predictive model with 5 latent variables recorded an R^2 value of 0.6567 at the calibration stage. The application of the regression on the range of wavelengths considered most significant (179) by the VIP method did not provide the expected results. Indeed, the R^2 obtained was slightly higher (0.7330). In validation, as expected, the predictive performance of the model decreased significantly reaching a negative value of R^2 (-0.3544). Considering one more latent variable, it was possible to obtain better results but still not satisfactory. At the calibration phase, a slightly higher R^2 (0.6890) was obtained. By performing the regression on the reduced dataset, selected by the VIP method, an R^2 of 0.7656 was achieved. However, at the validation, the model continued to show difficulties reconstructing the ethanol concentration. Indeed, the

value of R^2 is still low (0.3141) meaning that the model has limitations and beyond it cannot go. Glucose analysis proved to be even more complicated. The regression model built with fermentation data has significant limitations in predicting glucose concentration, particularly when only 5 latent variables were evaluated. Indeed, in the validation phase an R^2 of -0.9844 was found. Considering one more latent variable, the predictive capabilities of the model were improved a little. The R^2 was found to be equal to 0.7324 in calibration with the full dataset, 0.8056 with the reduced dataset, but in validation it dropped to 0.0256. Thus, the analysis of these data shows that the predictive ability of the sensor is strongly related to the solutions analyzed. Limpid solutions such as those evaluated in the first phase yielded satisfactory results. Torpid solutions such as those derived from a real process generate noisy spectra that are difficult to interpret, and from which it is complicated to extract useful information. A possible solution to this problem would be to have more experimental data and to use more robust statistical analysis tools, but also add a further filtering step of the analyzed solutions in order to obtain more interpretable spectra.

Bibliography

Alin, A. (2009). Comparison of PLS algorithms when number of objects is much larger than number of variables. *Statistical Papers*, *50*(4), 711–720. https://doi.org/10.1007/s00362-009-0251-7

Arnold, S. A., Crowley, J., Woods, N., Harvey, L. M., & McNeil, B. (2003). In-situ near infrared spectroscopy to monitor key analytes in mammalian cell cultivation. *Biotechnology and Bioengineering*, *84*(1), 13–19. https://doi.org/10.1002/bit.10738

Arnold, S. A., Gaensakoo, R., Harvey, L. M., & McNeil, B. (2002). Use of at-line and in-situ near-infrared spectroscopy to monitor biomass in an industrial fed-batch Escherichia coli process. *Biotechnology and Bioengineering*, *80*(4), 405–413. https://doi.org/10.1002/bit.10383

Ávila, T. C., Poppi, R. J., Lunardi, I., Tizei, P. A. G., & Pereira, G. A. G. (2012). Raman spectroscopy and chemometrics for on-line control of glucose fermentation by Saccharomyces cerevisiae. *Biotechnology Progress*, *28*(6), 1598–1604. https://doi.org/10.1002/btpr.1615

Beutel, S., & Henkel, S. (2011). In situ sensor techniques in modern bioprocess monitoring. *Applied Microbiology and Biotechnology*, *91*(6), 1493–1505. https://doi.org/10.1007/S00253-011-3470-5

Brereton, R. G. (2007). *Applied Chemometrics for Scientists* (1st ed.). John Wiley & Sons.

Chong, I. G., & Jun, C. H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems*, *78*(1), 103–112. https://doi.org/10.1016/j.chemolab.2004.12.011

Claßen, J., Aupert, F., Reardon, K. F., Solle, D., & Scheper, T. (2017). Spectroscopic sensors for in-line bioprocess monitoring in research and pharmaceutical industrial application. In *Analytical and Bioanalytical Chemistry* (Vol. 409, Issue 3, pp. 651–666). Springer Verlag. https://doi.org/10.1007/s00216-016-0068-x

Cocchi, M., Biancolillo, A., & Marini, F. (2018a). Chemometric Methods for Classification and Feature Selection. In *Comprehensive Analytical Chemistry* (Vol. 82, pp. 265–299). Elsevier B.V. https://doi.org/10.1016/bs.coac.2018.08.006

Cocchi, M., Biancolillo, A., & Marini, F. (2018b). Chemometric Methods for Classification and Feature Selection. In *Comprehensive Analytical Chemistry* (Vol. 82, pp. 265–299). Elsevier B.V. https://doi.org/10.1016/bs.coac.2018.08.006

de Jong, S. (1993). SIMPLS: an alternative approach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, *18*, 251–263.

de Luca, M., Oliverio, F., Ioele, G., & Ragno, G. (2009). Multivariate calibration techniques applied to derivative spectroscopy data for the analysis of pharmaceutical mixtures. *Chemometrics and Intelligent Laboratory Systems*, *96*(1), 14–21. https://doi.org/10.1016/j.chemolab.2008.10.009

Eliasson Lantz, A., Jørgensen, P., Poulsen, E., Lindemann, C., & Olsson, L. (2006). Determination of cell mass and polymyxin using multi-wavelength fluorescence. *Journal of Biotechnology*, *121*(4), 544–554. https://doi.org/10.1016/j.jbiotec.2005.08.007

Eriksson, L., Kettaneh-Wold, N., Trygg, J., Wikström, C., & Wold, S. (2006). Multi-and megavariate data analysis: Part I: basic principles and applications. . *Computer and Information Science*, *1*.

Esmonde-White, K. A., Cuellar, M., Uerpmann, C., Lenain, B., & Lewis, I. R. (2017). Raman spectroscopy as a process analytical technology for pharmaceutical manufacturing and bioprocessing. In *Analytical and Bioanalytical Chemistry* (Vol. 409, Issue 3, pp. 637–649). Springer Verlag. https://doi.org/10.1007/s00216-016-9824-1

Fayolle, P., Picque, D., & Corrieu, G. (2000). *On-line monitoring of fermentation processes by a new remote dispersive middle-infrared spectrometer*. www.elsevier.com/locate/foodcont

Gargalo, C. L., Udugama, I., Pontius, K., Lopez, P. C., Nielsen, R. F., Hasanzadeh, A., Mansouri, S. S., Bayer, C., Junicke, H., & Gernaey, K. v. (2020). Towards smart biomanufacturing: a perspective on recent developments in industrial measurement and monitoring technologies for bio-based production processes. *Journal of Industrial Microbiology and Biotechnology*, *47*(11), 947–964. https://doi.org/10.1007/s10295-020-02308-1

Hirsch, E., Pataki, H., Domján, J., Farkas, A., Vass, P., Fehér, C., Barta, Z., Nagy, Z. K., Marosi, G. J., & Csontos, I. (2019). Inline noninvasive Raman monitoring and feedback control of glucose concentration during ethanol fermentation. *Biotechnology Progress*, *35*(5). https://doi.org/10.1002/btpr.2848

Iversen, J. A., Berg, R. W., & Ahring, B. K. (2014). Quantitative monitoring of yeast fermentation using Raman spectroscopy. *Analytical and Bioanalytical Chemistry*, *406*(20), 4911–4919. https://doi.org/10.1007/s00216-014-7897-2

Jain, G., Jayaraman, G., Kökpinar, Ö., Rinas, U., & Hitzmann, B. (2011). On-line monitoring of recombinant bacterial cultures using multi-wavelength fluorescence spectroscopy. *Biochemical Engineering Journal*, *58–59*(1), 133–139. https://doi.org/10.1016/j.bej.2011.09.005

Langergraber, G., Fleischmann, N., Hofstaedter, F., & Weingartner, A. (2004). Monitoring of a paper mill wastewater treatment plant using UV/VIS spectroscopy. *Water Science and Technology*, *49*(1), 9–14. https://doi.org/10.2166/wst.2004.0004

Langergraber, G., Gupta, J. K., Pressl, A., Hofstaedter, F., Lettl, W., Weingartner, A., & Fleischmann, N. (2004). On-line monitoring for control of a pilot-scale sequencing batch reactor using a submersible UV/VIS spectrometer. *Water Science and Technology*, *50*(10), 73–80. https://doi.org/10.2166/wst.2004.0613

Lourenço, N. D., Lopes, J. A., Almeida, C. F., Sarraguça, M. C., & Pinheiro, H. M. (2012). Bioreactor monitoring with spectroscopy and chemometrics: A review. In *Analytical and Bioanalytical Chemistry* (Vol. 404, Issue 4, pp. 1211–1237). https://doi.org/10.1007/s00216-012-6073-9

Mehmood, T., Liland, K. H., Snipen, L., & Sæbø, S. (2012). A review of variable selection methods in Partial Least Squares Regression. In *Chemometrics and Intelligent Laboratory Systems* (Vol. 118, pp. 62–69). https://doi.org/10.1016/j.chemolab.2012.07.010

Nordon, A., Littlejohn, D., Dann, A. S., Jeffkins, P. A., Richardson, M. D., & Stimpson, S. L. (2008). In situ monitoring of the seed stage of a fermentation process using non-invasive NIR spectrometry. *Analyst*, *133*(5), 660–666. https://doi.org/10.1039/b719318a

Noui, L., Hill, J., Keay, P. J., Wang, R. Y., Smith, T., Yeung, K., Habib, G., & Hoare, M. (2002). Development of a high resolution UV spectrophotometer for at-line monitoring of bioprocesses. *Chemical Engineering and Processing: Process Intensification*, *41*(2), 107–114. https://doi.org/10.1016/S0255-2701(01)00122-2

Ödman, P., Johansen, C. L., Olsson, L., Gernaey, K. v., & Lantz, A. E. (2009). On-line estimation of biomass, glucose and ethanol in Saccharomyces cerevisiae cultivations using in-

situ multi-wavelength fluorescence and software sensors. *Journal of Biotechnology*, *144*(2), 102–112. https://doi.org/10.1016/j.jbiotec.2009.08.018

Oh, S. K., Yoo, S. J., Jeong, D. H., & Lee, J. M. (2013). Real-time estimation of glucose concentration in algae cultivation system using Raman spectroscopy. *Bioresource Technology*, *142*, 131–137. https://doi.org/10.1016/j.biortech.2013.05.008

Oh, S.-K., Yoo, S. J., & Lee, J. M. (n.d.). *Predicting concentrations of a mixture in bioreactor for on-line monitoring using Raman spectroscopy*.

Pachauri, N., Rani, A., & Singh, V. (2017). Bioreactor temperature control using modified fractional order IMC-PID for ethanol production. *Chemical Engineering Research and Design*, *122*, 97–112. https://doi.org/10.1016/j.cherd.2017.03.031

Picard, A., Daniel, I., Montagnac, G., & Oger, P. (2007). In situ monitoring by quantitative Raman spectroscopy of alcoholic fermentation by Saccharomyces cerevisiae under high pressure. *Extremophiles*, *11*(3), 445–452. https://doi.org/10.1007/s00792-006-0054-x

Pons, M. N., Bonté, S. le, & Potier, O. (2004). Spectral analysis and fingerprinting for biomedia characterisation. *Journal of Biotechnology*, *113*(1–3), 211–230. https://doi.org/10.1016/j.jbiotec.2004.03.028

Rathore, A. S., Bhushan, N., & Hadpe, S. (2011). Chemometrics applications in biotech processes: A review. In *Biotechnology Progress* (Vol. 27, Issue 2, pp. 307–315). https://doi.org/10.1002/btpr.561

Rhee, J. il, & Kang, T. H. (2007). On-line process monitoring and chemometric modeling with 2D fluorescence spectra obtained in recombinant E. coli fermentations. *Process Biochemistry*, *42*(7), 1124–1134. https://doi.org/10.1016/j.procbio.2007.05.007

Roberts, J., Power, A., Chapman, J., Chandra, S., & Cozzolino, D. (2018). The use of UV-Vis spectroscopy in bioprocess and fermentation monitoring. In *Fermentation* (Vol. 4, Issue 1). MDPI AG. https://doi.org/10.3390/fermentation4010018

Roychoudhury, P., O'Kennedy, R., McNeil, B., & Harvey, L. M. (2007). Multiplexing fibre optic near infrared (NIR) spectroscopy as an emerging technology to monitor industrial bioprocesses. *Analytica Chimica Acta*, *590*(1), 110–117. https://doi.org/10.1016/j.aca.2007.03.011

Sarraguça, M. C., Paulo, A., Alves, M. M., Dias, A. M. A., Lopes, J. A., & Ferreira, E. C. (2009). Quantitative monitoring of an activated sludge reactor using on-line UV-visible and near-infrared spectroscopy. *Analytical and Bioanalytical Chemistry*, *395*(4), 1159–1166. https://doi.org/10.1007/s00216-009-3042-z

Schalk, R., Braun, F., Frank, R., Rädle, M., Gretz, N., Methner, F. J., & Beuermann, T. (2017). Non-contact Raman spectroscopy for in-line monitoring of glucose and ethanol during yeast fermentations. *Bioprocess and Biosystems Engineering*, *40*(10), 1519–1527. https://doi.org/10.1007/s00449-017-1808-9

Schenk, J., Marison, I. W., & von Stockar, U. (2007). Mid-Infrared spectroscopy as a tool for pH prediction and control in bioprocesses. *Journal of Biotechnology*, *2*(131).

Sivakesava, S., Irudayaraj, J., & Demirci, A. (2001). Monitoring a bioprocess for ethanol production using FT-MIR and FT-Raman spectroscopy. In *Journal of Industrial Microbiology & Biotechnology* (Vol. 26). www.nature.com/jim

Surribas, A., Montesinos, J. L., & Valero, F. F. (2006). Biomass estimation using fluorescence measurements in Pichia pastoris bioprocess. *Journal of Chemical Technology and Biotechnology*, *81*(1), 23–28. https://doi.org/10.1002/jctb.1352

Tosi, S., Rossi, M., Tamburini, E., Vaccari, G., Amaretti, A., & Matteuzzi, D. (2003). Assessment of In-Line Near-Infrared Spectroscopy for Continuous Monitoring of Fermentation Processes. *Biotechnology Progress*, *19*(6), 1816–1821. https://doi.org/10.1021/bp034101n

Udugama, I. A., Gargalo, C. L., Yamashita, Y., Taube, M. A., Palazoglu, A., Young, B. R., Gernaey, K. v., Kulahci, M., & Bayer, C. (2020). The Role of Big Data in Industrial (Bio)chemical Process Operations. *Industrial and Engineering Chemistry Research*, *59*(34), 15283–15297. https://doi.org/10.1021/acs.iecr.0c01872

Wold, S., Johansson, E., & Cocchi, M. (1993). *PLS: partial least squares projections to latent structures* . 523550.

Wold, S., Sjostrom, M., Eriksson, L., & Sweden˚˚, S. (n.d.). PLS-regression: a basic tool of chemometrics. In *Chemometrics and Intelligent Laboratory Systems* (Vol. 58). www.elsevier.comrlocaterchemometrics

Chapter 6: Conclusions

This chapter summarizes the main conclusions.

In this thesis, the problem of monitoring and control of biotechnological processes was addressed. Particularly, the focus was on the fermentation process in which living microorganisms (yeasts) use the substrate present in the bioreactor medium to produce ethanol. The involvement of microbial cultures represents the factor in common to all bioprocesses and is also the main reason why the bioreactor is the most difficult unit operation to control. The issue may become more intricate when considering the production of biochemicals or biofuels derived from waste biomass, since the type of feedstock and pre-treatment selected to obtain fermentable sugars have a significant effect on the growth of microorganisms involved, and consequently on the desired product quality. The problem had been first addressed considering the detailed mathematical model of a continuous fermentation bioreactor as a virtual plant. This model, previously developed by other researchers, was selected for the purpose of the work here presented because it contained all the nonlinear and complex characteristics of the process, offering a useful tool to test nonlinear estimation and control methods. The monitoring problem was addressed by means of an adjustable-structure geometric estimation approach, where the estimator structure represented a design degree of freedom to improve its performance versus robustness behavior. The investigated estimation algorithm was the geometric observer with proportional innovation, which offers simplicity of tuning and implementation. In order to demonstrate that the proposed procedure for choosing the estimation structure could be applied to other estimation techniques, the extended Kalman filter was also implemented as measurement processor algorithm. The results analysis showed good estimation performance for both estimators, although regarding the geometric observer a higher sensitivity to measurement noise, attributable to the presence of the Lie derivatives in the correction term, was noted. In summary, the geometric method demonstrated to be the most effective solution to address the estimation problem. It provided a structure independent of the correction algorithm, allowing flexibility in the choice of the latter based on the preferences of the plant personnel or developer experience. The results of this study, in terms of methodology, can be applied to more advanced biotechnological processes. Moreover, it can be helpful to identify the most crucial measurements for optimal reconstruction and invest in them.

Subsequently, the research work focused on the control of the fermentation bioreactor, using the detailed mathematical model with the estimators previously developed to accomplish the task. Analysis of the control performance indices and the evaluation of the dynamic trends pointed out that when product composition measurements were not available, the only temperature controller was not able to ensure the achievement of product composition set point.

Evaluating the case where ethanol measurements were affected by delay, the proposed cascade structure performed better even in presence of variations in model input parameters. However, it was the configuration of the MIMO system, together with the estimator for ethanol concentration, that guaranteed compliance with product quality. Results from dynamic trend analysis was also demonstrated by the performance index values calculated. It worth noting that the excellent result achieved with this configuration was mainly due to the good capabilities of the estimator, and its robustness.

The proposed solutions for monitoring and controlling the fermentation process were initially developed using numeric simulations. However, these were based on the assumption that the fermentation was performed in an ideal reactor, using a pure glucose solution without considering the presence of possible inhibitory compounds. This does not reflect the conditions of a real plant, especially when using waste biomass as a renewable source of fermentable sugars. An experimental study was conducted to gain a deeper understanding of the fermentation process evaluating its sensitivity to process conditions and design an effective monitoring tool. The exploitation and valorization of waste biomass such as brewery's spent grain (BSG) to produce second-generation biofuels was investigated as a valid and concrete example of a biotechnological process. Using proper statistical techniques for the design of experiments, the bioconversion of BSG to ethanol was investigated in a wide range of operating conditions, and the main variables affecting the process were identified. Raman spectroscopy in combination with chemometric tools were applied for the obtainment of real-time quantitative analysis of products during fermentation of sugars for ethanol production. The results indicated that the proposed approach was successful when using sample specially prepared containing only glucose, ethanol, and traces of acids. When turbid solutions obtained from the fermentation experiments were analysed, it was not possible to obtain informative and interpretable spectra, evidencing a limit in this technology.

Acknowledgements

I am deeply grateful to my supervisor, Prof. Eng. Massimiliano Grosso, and my co-supervisor, Prof. Eng. Stefania Tronci, for their unwavering trust in me and their precious guidance, constant support, patience, and availability throughout my PhD. Their passion and devotion to their research activities have been a source of inspiration, and their valuable teachings and advice have been essential in my growth and development.

I would also like to extend special thanks to Prof. Eng. Massimiliano Errico, who generously hosted me at the University of Southern Denmark (SDU) and gave me with the opportunity to carry out the experimental part of my research work. I will always be grateful to him for his helpfulness and the knowledge he shared.

Lastly, I would like to express my sincere gratitude to Dr. Eng. Laura Mais for her support in my laboratory activities and her valuable contribution to my research work.