**JESSICA RUBIRA GAMBA**

# Experimental improvement by audit and data smoothing: application in biopolymer production

**São Paulo**
**2023**

**JESSICA RUBIRA GAMBA**

# Experimental improvement by audit and data smoothing: application in biopolymer production

Revised version

Dissertation submitted in fulfillment of the requirements for the degree of Master of Science in Chemical Engineering.

Advisor: Prof. Dr. Galo Carrillo Le Roux

Co-advisor: Prof. Dr. José Gregório Gomez

Revised version. The original version is available at Biblioteca Central da EPUSP.

**São Paulo**
**2023**

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 25 de Julho de 2023

Assinatura do autor: Jessica Rubira Gamba

Assinatura do orientador:

Catalogação-na-publicação

Gamba, Jessica Rubira
Experimental improvement by audit and data smoothing: application in biopolymer production / J. R. Gamba -- versão corr. -- São Paulo, 2023.
65 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Química.

1.Specific growth rate 2.Smoothing methods 3.PHB production I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia Química II.t.

GAMBA, J. R. **Experimental improvement by audit and data smoothing: application in biopolymer production.** 2023. 62 p. Masters thesis (Chemical Engineering) - Polytechnic School, University of Sao Paulo, Sao Paulo, 2023.

Aprovado em:

Banca Examinadora

Prof. Dr.: .....
Instituição: .....
Julgamento: .....

Prof. Dr.: .....
Instituição: .....
Julgamento: .....

Prof. Dr.: .....
Instituição: .....
Julgamento: .....

Prof. Dr.: .....
Instituição: .....
Julgamento: .....

*To my mother,*

*for a lifetime of dedication to my studies.*

## ACKNOWLEDGEMENT

To Professors Dr. Galo Carillo Le Roux and Dr. José Gregório Cabrera Gomez for their guidance, confidence, encouragement and support at all times.

To all the friends I made at the Chemical Engineering Department and in the Bioproducts Laboratory, especially Matheus Guedes, Priscila, Maria Clara, Rafael, Andréa, Pedro, Edmar, Sofia, Vânia, Franz, Jung, Ingrid and many others.

To my family and friends who supported me at all times.

To the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) - Financing Code 001, the National Council for Scientific and Technological Development (CNPq).

## ABSTRACT

GAMBA, J. R. **Experimental improvement by audit and data smoothing: application in biopolymer production.** 2023. 62 p. Masters thesis (Chemical Engineering) - Polytechnic School, University of Sao Paulo, Sao Paulo, 2023.

Researchers often perceive the complexity of bioprocesses, the limited number of models applicable to them and the measures available to study them. In order to interpret and reduce the variability of bioprocesses, it is necessary to evaluate the accuracy and consistency of the collected data, since these are usually corrupted by deviations, failures and human errors, in addition to poor equipment calibration. Within a bioprocess, the application of such data has a variety of objectives, such as parameter calculation, formulation of kinetic relationships, analysis of metabolic flow, monitoring and others. An important parameter is the specific growth rate, a property that does not depend on the size of the system under investigation, facilitating comparisons between different experimental platforms. There are several approaches for estimating such parameter (Ferraz et al., 1995; Le Duy and Zajic, 1973; Oner et al., 1986) and the present work aims to explore the Loess method as smoothing method followed by differentiation in order to obtain this rate, in addition to evaluating the experimental variability and reliability of data in a biopolymers production process. Therefore, we proposed improvements in sampling procedures to increase data reliability and compared a traditional method with the Loess-based smoothing method in calculating specific growth rate. Both the proposed improvements and the smoothing method proved to be satisfactory in their performance.



**Keywords**: specific rate, smoothing methods, PHB production.

## RESUMO

GAMBA, J. R. **Melhoria experimental por auditoria e alisamento de dados: aplicação em produção de biopolímero.** 2023. 62 f. Dissertação (Mestrado em Engenharia Química) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2023.

Os pesquisadores muitas vezes percebem a complexidade dos bioprocessos, o número limitado de modelos aplicáveis a eles e as medidas disponíveis para estudá-los. Para interpretar e reduzir a variabilidade dos bioprocessos, é necessário avaliar a precisão e a consistência dos dados coletados, pois estes geralmente são corrompidos por desvios, falhas e erros humanos, além da má calibração dos equipamentos. Dentro de um bioprocesso, a aplicação desses dados tem diversos objetivos, como cálculo de parâmetros, formulação de relações cinéticas, análise de fluxo metabólico, monitoramento e outros. Um parâmetro importante é a velocidade de crescimento específica, propriedade que independe do tamanho do sistema sob investigação, facilitando comparações entre diferentes plataformas experimentais. Existem várias abordagens para estimar tal parâmetro (Ferraz et al., 1995; Le Duy e Zajic, 1973; Oner et al., 1986) e o presente trabalho visa explorar o método de Loess para suavização desses dados seguido de diferenciação para obter essa taxa, além de avaliar a variabilidade experimental e a confiabilidade dos dados em um processo de produção de biopolímeros. Portanto, propusemos melhorias nos procedimentos de amostragem para aumentar a confiabilidade dos dados e comparamos um método tradicional com o método de suavização baseado em Loess no cálculo da taxa de crescimento específica. Tanto as melhorias propostas quanto o método de suavização mostraram-se satisfatórios em seu desempenho.



**Palavras-chaves**: velocidade específica, métodos de alisamento, produção de PHB.

## LIST OF FIGURES

## LIST OF TABLES

## LIST OF ABBREVIATIONS AND ACRONYMS

| | |
|---|---|
| D | Dilution rate; |
| DSM | Deutsche Sammlung vor Mikrorganismen und Zellkulturen; |
| EFM | Elementary Flux Mode |
| EKF | Extended Kalman Filter; |
| FBA | Flux Balance Analysis; |
| LB | Lysogeny broth; |
| OD | Optical density; |
| P(3HB) | Poly(3-hydroxybutyrate); |
| PHA | Polyhydroxyalkanoate. |

## CONTENTS

# 1. INTRODUCTION

## 1.1. Background and motivation

Fluctuations detected when distinct replicates are compared with each other are due to unavoidable experimental fluctuations (BOX et. al 2005). For this reason, the study of experimental errors is of paramount importance when the objective of the study is to analyze the performance of biochemical processes and/or build models. As it is well known, the quality of the experimental data can only be specified when experimental errors are known. However, the proper characterization of experimental errors and the analysis of how unavoidable experimental errors affect process analysis and model building are seldom performed in the biotechnological field (HIMMELBLAU, 1970 and LARENTIS et al. 2003). The proper characterization of experimental errors is important in bioprocesses when there is interest in designing optimum operation conditions, analyzing product recovery, designing bioreactors, improving process control, and reducing process variability (ABUD, 2005). In particular, the control of biotechnological processes differs from traditional control systems due to the many issues related to the metabolism of living organisms. For example, the proper control of the inoculum and of the inoculation is a very complex task, as it may not be possible to characterize precisely whether the microorganisms have the same level of activity from one batch to the other (WEBB and ATKINSON, 1992). Therefore, characterization of process variability may greatly facilitate the proper analysis of process data. Nevertheless, the influence of experimental errors on the operation of bioprocesses has been largely overlooked in the literature.

Still in this context, the accurate estimation of cell growth or the substrate consumption rate in bioprocesses is crucial for the understanding of the current state of a bioprocess. Rates unveil the actual cell status, making them valuable for quality-by-design concepts. However, in bioprocesses, the real rates are commonly not accessible due to analytical errors (BAYER, 2020). Extracellular metabolites, accumulating in or disappearing from the growth medium, are particularly interesting. Their time profiles, while being relatively easy to measure, provide a footprint of intracellular physiology (KELL et al., 2005). In particular, the time-varying concentrations of extracellular metabolites allow the computation of uptake and

excretion rates that can provide a kinetic study of the process and also to assess metabolic phenotypes of cultured cells..

The specific uptake rate is important to identify possible consumption of the limiting substrates for maintenance, and the specific growth and excretion rates are the first step for a good formulation and adjustment of a mathematical model of processes. Measured extracellular rates, therefore, are critical inputs for:

- Formulation of kinetic relationships which, together with mass balances, are the basis for a model construction;
- Obtaining preliminary estimates of the parameters through simplifications and linearization of the model to be used later as a starting point in the methodologies for parameter adjustment (HIMMELBLAU, 1972);
- Basis for intracellular flux calculations using flux balance analysis (FBA) or metabolic flux analysis (MFA) (QUEK et al., 2010);
- Online monitoring and control of biotechnological processes.

Regarding flux analysis, the cell rates constrain the solution space of feasible intracellular fluxes. Thus, reconstruction of accurate metabolic flux maps is totally associated with an accurate estimation of cell specific extracellular rates and their uncertainties. This can be illustrated in the work of CINQUEMANI *et al.* (2017), in which they compared the results from two specific rate estimators with a flux balance model of *Escherichia coli*. They were able to conclude that, for one set of estimated rates, the partitioning of the incoming flux of glucose over the glycolysis, pentose-phosphate, and glycogenolysis pathways was not appropriately accounted for, evidencing that the specific rate estimations were not accurate and providing a not reliable metabolic flux map.

Having said that, the importance of carefully calculating the specific growth and production rates of metabolic products from experimental data is characterized, which is hampered by the strong influence that small changes in the variables put forth on the calculation of their rate. However, there are some challenges on performing that calculation.

Off-line or online extracellular metabolites measurements are necessary to calculate the time-varying uptake and excretion rates and this is a challenging problem for a number of reasons. First, the available data are noisy, even when taking into account progress in such measurement methods. Second, different extracellular

metabolites time-course profiles are often strongly correlated. For instance, one obvious source of correlation is the proportionality of uptake and excretion rates to the size of the growing population of cells consuming or producing the metabolites (STEPHANOPOULOS et al., 1998). Third, the time-course profiles of extracellular metabolites are subject to discontinuities, due to sudden changes in the functioning of metabolism, such as seen in many bacteria catabolite repression that leads to the sequential utilization of carbon sources, generally favoring carbon sources that sustain a higher growth rate (KREMLING et al., 2015).

In a fed-batch culture, the addition of nutrients modifies the volumes and consequently the concentrations in a way that is not due to the metabolic activity. This issue can be solved by converting to mass (g) and not concentrations. Mass determination in fed-batch processes, however, generates a new problem associated with the sample volume, as specific rate measurements are normally made considering the values before or after sampling, but the real transformation is the mass that remained in the reactor right after the sample and the one that was in the reactor immediately before the next sample. Additionally, in a fed-batch culture, the specific rates are derivative quantities. In other words, it is necessary to differentiate the batch biomass and product concentration data to estimate the specific growth rate and specific product formation rate. However, the differentiation of measured concentrations amplifies the experimental error. In order to reduce the noise, data pre-treatment can be carried out, aiming to reduce undesirable variations that may influence the final results, and carefully, so as not to distort or render the analysis results unusable. These variations (noise) make up the stochastic portion of the signal and can be reduced using smoothing methods (FERREIRA, 2015). The most common techniques, in addition to the spline (BAYER et al., 2020), consider a segment of the sample, called a window, which moves across the entire signal, as is the case with the moving average method – MA, the weighted average - SG (SAVITZKY-GOLAY, 1964) and local weighted regression - Lowess/Loess (CLEVELAND, 1979). This work aims to explore the application of the Loess method in dry biomass measurements and specific growth rate.

The chosen bioprocess to apply such study is a particular bioprocess that has attracted much attention in recent years, concerning the production of polyhydroxyalkanoates (PHAs), which are considered green plastics and have a positive social and environmental impact when compared to conventional plastics in

terms of production and and re-integration into biological cycles. PHAs are biopolymers produced and accumulated intracellularly by various bacteria in the form of granules as a carbon and energy reserve material (ANDERSON AND DAWES, 1990). Poly-3-hydroxybutyrate (P(3HB)) is the most studied polymer of this class and can be produced by several bacteria, *Ralstonia eutropha* being a reference model due to its great efficiency in the production of P3HB and its copolymers like P3HB-*co*-3HV and others (STEINBÜCHEL AND SCHLEGEL, 1991; REINECKE AND STEINBUCHEL, 2009). Its production is generally carried out in a fed batch reactor in two stages: in the first, there is a supply of carbon and mineral nutrients sources to favor growth; and in the second, an essential nutrient (nitrogen, phosphorous etc) was exhausted and favor the production of P3HB as reserve of carbon and energy (ANDERSON AND DAWES, 1990).

## 1.2. Project objectives

The main objective of this work was to explore the Loess method as smoothing method followed by differentiation in order to obtain this rate. The specific objectives of the project are as follow:

- evaluate the experimental variability and reliability of data in a biopolymers production process;
- make a correlation between the different methods used to determine cell mass concentration;
- deliver some improvement actions in those procedural conditions to cut down experimental errors and deviations.

# 2. TRANSFORMATION PARAMETERS IN BIOPROCESSES

## 2.1. Quantifying cell concentration

The quantification of cell concentration in a culture medium is essential for the determination of the kinetics and stoichiometry of microbial growth and, as a consequence, growth, product and substrate specific rates. The methods used for this purpose can be classified in two categories: direct and indirect. In many cases, the direct methods are not feasible due to the presence of suspended solids or other interfering compounds in the medium. Either cell number or cell mass can be quantified depending on the type of information needed and the properties of the system. Cell mass concentration is often preferred due to the ease in manipulating this variable for further analysis.

Determination of cellular dry weight is the most commonly used direct method for determining cell mass concentration and is applicable only for cells grown in solids-free medium. Typically, samples of culture broth are centrifuged or filtered and washed with a buffer solution or water. The washed wet cell mass is then dried at 80$^{o}$C for 24 hours; then dry cell weight is measured. This measurement can also be carried out in different ways, such as by the difference in the empty recipient and the same recipient with the dried cells, and also removing all the dried cell mass and weighting it.

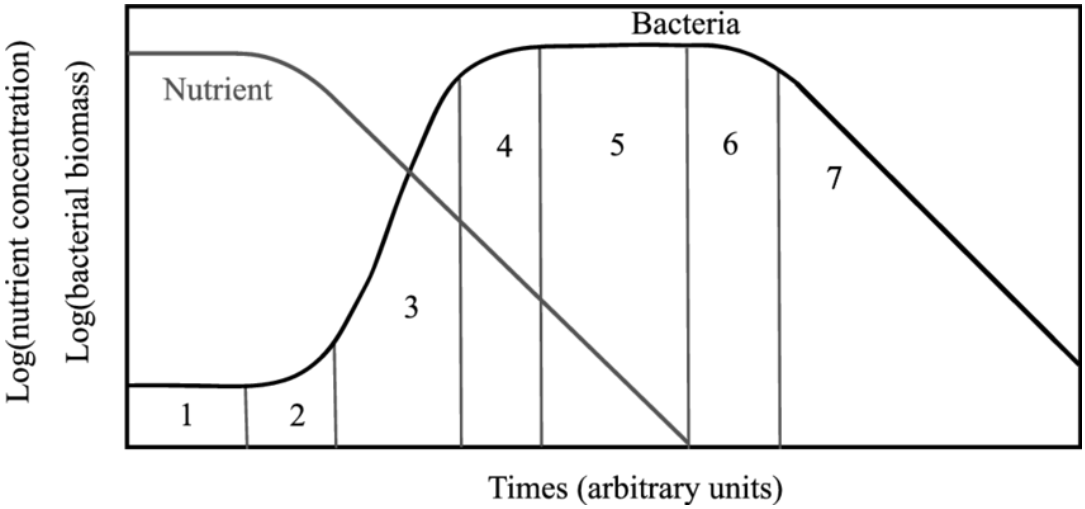
Another rapid method is based on the absorption of light by suspended cells in sample culture media. The intensity of the transmitted light is measured using a spectrophotometer. The extent of light transmission in a sample chamber is a function of cell density and the thickness of the chamber. It is also a function of the size/volume of the particles (cells), which varies with the osmotic pressure of the medium, an effect that can be problematic for estimating cell density based on absorbance/scattering. Light transmission is modulated by both absorption and scattering. Pigmented cells give different results than unpigmented ones. Background absorption by components in the medium must be considered, particularly if absorbing dissolved species are taken into cells. The medium should be essentially particle free. Proper procedure entails using a wavelength that minimizes absorption by medium components (600- to 700-nm wavelengths are often used), “blanking” against medium, and the use of a calibration curve. The calibration curve relates optical density (OD) to dry-weight

measurements. Such calibration curves can become nonlinear at high OD values (> 0.3) and depend to some extent on the physiological state of the cells.

## 2.2. Microbial growth curve

After inoculation of a culture medium favorable to the development of the microorganism under study, under controlled temperature and appropriate agitation, a behavior in the cell concentration values is observed, as shown in Figure 1 (BUCHANAN, 1918).

Figure 1.Phases of bacterial growth.




(Source: ATEKWANA *et al.* 2012)

The following stages in growth are observed:

- Phase 1: known as lag or latency phase, which follows immediately after inoculation of the medium with the microorganism in question. This is a period of adaptation during which the cell synthesizes the enzymes necessary for the metabolism of the components present in the medium. During this first phase, there is no cell reproduction and, thus, $X = X_0$ = constant. The duration of this phase varies mainly with the inoculum concentration (and therefore with the value of X0), with the age of the microorganism (pre-cultivation time) and with its physiological state.
- Phase 2 - acceleration of the bacterial growing rate (acceleration phase).There is a gradual increase in both the reproduction rate and the specific growth rate, in

which not all microorganisms complete the previous phase simultaneously. At the end of this phase, the entire population begins to divide into an average regular interval of time.

- Phase 3: exponential increase (maximal and constant exponential growth rate; μ is given by the slope of the ln line in this phase). Under these circumstances, the growth rate is directly proportional to the concentration X.
- Phase 4: decreasing growth rate due to the gradual decrease in substrate concentration and the increased accumulation of toxic metabolites (the substrate is typically exhausted at the end of this phase).
- Phase 5: stationary phase (the nutrient is exhausted at the end of this phase)
- Phase 6: acceleration of the bacterial decay.
- Phase 7: exponential decay phase (phase with endogenous metabolism, high death rate, and cell lysis).

### 2.3. Methods for growth rate estimation

For microbes, growth is their most essential response to their physiochemical environment, being a result of both replication and change in cell size. Microorganisms can grow under a variety of physical, chemical, and nutritional conditions. In a suitable nutrient medium, organisms extract nutrients from the medium and convert them into biological compounds. Part of these nutrients are used for energy production and part are used for biosynthesis and product formation. As a result of nutrient utilization, microbial mass increases with time and can be described simply by:

$$substrates + cell \rightarrow extracellular\ products + more\ cells$$

$$\sum S + X \rightarrow \sum P + nX$$

The rate of microbial growth is characterized by the specific growth rate, defined as

$$\mu = \frac{1}{X}\frac{dX}{dt}$$

where $X$ is cell mass concentration (g/L), $t$ is time (h), and μ is specific growth rate (g/g.h or $h^{-1}$) (GADEN, 1955). Methods to estimate specific growth rate will be presented in the next sections, as an attempt to cover their use and drawbacks.

Table 1. Traditional methods of estimating specific growth velocity and their disadvantages.

| Method | Disadvantage | Reference |
|---|---|---|
| Exponential phase detection with linear adjustment (lnX vs t); adjustment of X values according to the trend line; polynomial fit on the new curve obtained | μ is calculated at each time interval; if the time interval is decreased, noise is increased in biomass measurements. | FERRAZ et al. (1995) |
| Tangent tracing methods | Manual tracing, directly connected to the ability and variability of operators. | MICKLEY et al., 1957 |
| Geometric method | Approximation of the derivative by finite differences | LE DUY e ZAJIC, 1973 |
| Adjustments to exponential or sigmoidal curves | They assume that the specific growth rate is constant; does not determine changes | BORZANI, 1986 |
| Polynomial fits or splines | The order of the polynomial function is unknown; each measurement is considered as exact, disregarding the precision and noise characteristics of biomass measurements | ONER et al. 1989<br>WOLD, 1974 |

(Source: own elaboration)

### 2.3.1. Traditional methods

FERRAZ *et al.* (1995) describe a methodology that could be used to calculate specific growth rate which can be divided in three steps. In the first step, the exponential growth phase is detected, plotting (ln $X$) vs. (t) for different initial and final time limits, determining, by visual inspection, the beginning and duration of the exponential phase of growth; the angular coefficient of the interval will give the value of maximum specific speed of growth. In the second step, improvement of the curve $X$ vs. (t) is made. Recovering the $X$ values that satisfy the linear regression chosen in the previous step improves the $X$ vs. (t) during the exponential phase. The last step is the calculation of specific growth rate. With the new curve (X) vs. (t) the specific growth rate curve is obtained using a polynomial adjust.

Other procedures consist on graphical methods for tangent tracing (MICKLEY et al., 1957) and geometric method (LE DUY and ZAJIC, 1973). The latter work proposes a differentiation method based on geometry through a script in Fortran. SCHMIDELL *et al.* (2001) applies this method in an Excel spreadsheet.

Besides dividing the total growth rate by the average biomass concentration between two measurements, biotechnology textbooks present the following frequently used function to calculate the specific growth rate between two consecutive biomass measurements:

$$\mu = \frac{\ln\left(\frac{C_{X,k+1}}{C_{X,k}}\right)}{t_{k+1} - t_k} \qquad 1(1)$$

With this approach, the specific growth rate is calculated over each time interval; however, with a decreasing time interval, noise on the biomass measurements plays an increasing role. Other approaches seen regularly are fitting exponential curves or sigmoid. Main disadvantage of fitting such equations through data is the assumption that the specific growth rate is constant and so these methods are not appropriate to determine changes.

Polynomials or splines solve that problem. Spline functions are piecewise polynomials of degree $n$that canbe used for smoothing and differentiation of data (ONER et al. 1986). Carbon and electron balances can be used to evaluate the

consistency of the estimated specific rates and also to select parameters that are employed in smoothing the data. REINSCH (1967) developed the cubic spline function as:

$$g(t) = a_i + b_i(t - t_i) + c_i(t - t_i)^2 + d_i(t - t_i)^3 \quad (2)$$

where $ti < t < ti + 1$. A spline function of degree $n$ has $n - 1$ continuous derivatives. The following equation which controls the smoothness of the curve and fidelity of the fit should be minimized:

$$G = \sum_{i=1}^{N} \left[\frac{g(t) - y_i}{dy_i}\right]^2 + \lambda \int_{t_0}^{t_n} \left(\frac{d^2 g}{dt^2}\right)^2 dt \quad (3)$$

where $y_i$ represents a measured value at time $t_n$, parameters $dy_i, i = 1,2,\ldots N$, are input parameters which weight the corresponding data points, and $\lambda$ is a smoothing parameter. The first term of Equation (3) measures the fidelity of the spline fit to the data and the second term determines the smoothness of the curve. WOLD (1974) provides a more detailed formulation for spline functions in data analysis.

The problem of using spline and polynomial functions is to fix the specific growth rate to change according to a polynomial of lower degree. The order of this polynomial function is a crucial and unknown parameter. Therefore, the main disadvantage of these traditional techniques is that they consider each measurement to be exact by disregarding the accuracy and noise characteristics of the biomass measurements.

### 2.3.2. Recursive estimation

Several authors described recursive estimators. SAN and STEPHANOPOULOS (1984) reported on the subject and CLAES and VAN IMPE (1999) validated the performance of an observer online. The model necessary for specific growth rate estimation describes changes of biomass as directly related to the biomass concentration, net specific growth rate and the dilution rate, D:

$$\frac{dC_X}{dt} = \mu C_X - D C_x \quad (4)$$

The behavior of the specific growth rate during cultivation is not known a priori, and so the change of the specific growth rate is assumed to be a random walk process:

$$\frac{d\mu}{dt} = 0 + zz \sim (0, \gamma) \tag{5}$$

with z as a white-noise process with zero mean and spectral density$\gamma$. The expected variability of the specific growth rate in time is reflected by the choice of the variance of $\gamma$ in the algorithm. Transferring this model (Eq. 4 and Eq. 5) into state representation for the estimator, using the state vector $x = [C_x \mu]^T$, input $u = [D]$ and output $y = C_X$

$$\begin{aligned} \frac{dx}{dt} &= f(x,u) + \ ww \sim (0, Q^S) \\ y_k &= Hx_k + \ v_k; \quad v_k \sim (0, R_k) \end{aligned} \tag{6}$$

where $f$ is the non-linear function of x and $u$. $w$ is the zero mean white noise with spectral density $Q^S$, $y_k$ is the sampled output with $H$ as the output matrix and $v_k$ as zero mean white noise with covariance $R_k$. Biomass is determined by measuring the optical density of the cells and, due to the limited range of the analyzer, samples with high biomass concentration are diluted either with water or medium. This results in a standard deviation of the biomass measurements, which is proportional to biomass: $\sigma = \lambda C_X$. Thus, the covariance $R_k$ of these measurements equals $\sigma^2$. The parameter $\lambda$ can easily be determined from actual measurements. Since all unpredicted changes of the biomass have to be reflected by the specific growth rate, the model description for biomass growth is assumed perfect. Thus, the spectral density of the biomass prediction equals zero. The remaining single tuning parameter is $\gamma$.

$$Q^S = \begin{bmatrix} 0 & 0 \\ 0 & \gamma \end{bmatrix} \tag{(7}$$

### 2.3.3. Software used to determine specific rates in bioprocesses

Table 2 presents a list of the available software with its methods and price used to determine specific rates in bioprocesses. Two of those methods will be described in

more detailed further in this section, since they were the ones available for our group at the time this work was carried out.

Table 2. Available software, and its methods, used to determine specific rates in bioprocesses.

| Software | Included methods | Availability | Comments |
|---|---|---|---|
| Prism 8 | Curve fitting; interpolations; spline | Paid (30 days trial) | - |
| OriginPro Graphing & Analysis | Adjacent mean,; Savitzky-Golay; Loess/Lowess | Paid (21 days trial) | It interacts with Matlab and LabView |
| SigmaPlot Exact Graphs and Data Analysis | -Média móvel Mediana móvel Loess | Paid (30 days trial) | - |
| Matlab (toolbox) | Média móvel Savitzky-Golay Loess/Lowess Loess/Lowess robustos | Paid (30 days trial) | Spline available without need of a additional toolbox |
| Lissage (UFPE, 1994) | Spline | Not commercial | - |
| Emerson (IPT, 2004) | Interpolações Spline | Not commercial | It calculates specific rates |

(Source: own elaboration)

*Emerson*

Application for calculating rates and specific rates of reaction in biotechnological processes (BRUNA, W.; BARRAL, M.F.; PRADELLA, J.G.C.; 2004) and it is used for the treatment of experimental data. This software makes it possible to discard or correct experimental points that present gross errors, to elaborate growth, substrate consumption and product formation curves. It reads a text file containing the points to be treated and creates the corresponding graphs containing the experimental points and the curve that passes through these points. This curve can be calculated by different types of interpolation. One interpolation method is the spline method, which fits a polynomial of degree n to an interval of two points of X, incorporating a number of points “ahead” of the interval to be defined, in addition, the method forces that the derivative of the polynomial adjusted in the previous interval be equal to the derivative of the adjusted polynomial in the new interval, at the point of intersection (characteristic of spline methods). Through visual tests, the degree of the polynomial to be adjusted is defined, as well as the number of points “ahead” included in the adjustment (BONOMI; SCHIMIDELL, 2001).

Graphs in this software can be plotted on a logarithmic scale and linear regression values can be obtained. By selecting the experimental points, it is possible to position them in the desired location or remove them. By selecting the curve, it is possible to add new points to it. By selecting curves, the program also performs calculations such as specific, derivative and integral velocity. As an output, the software provides a text file containing the original data and the adjusted data, the amount of data to be added in this file can be defined by the user.

*Lissage*

It is a software developed at INSA-Toulouse by Diogo A. Simões (ARDAILLON-SIMÕES; ARROYO; URIBELARREA, 1994). This program is based on a variation of the spline method, which employs the adjustment of a polynomial of degree n (3 or 4) to an interval of two points, incorporating a previously defined number of points ahead.

### 2.3.4. Differentiation methods in Loess/Lowess

LOESS is one of many "modern" modeling methods that build on "classical" methods, such as linear and nonlinear least squares regression, which are designed

to address situations in which the classical procedures do not perform well or cannot be effectively applied without undue labor. LOESS combines the simplicity of linear least squares regression with the flexibility of nonlinear regression by fitting simple models to localized subsets of the data to build up a function that describes the deterministic part of the variation in the data, point by point (CLEVELAND, 1979).

This method was originally proposed by Cleveland (1979) and further developed by Cleveland and Devlin (1988), and specifically denotes a method that is more descriptively known as locally weighted polynomial regression. At each point in the data set a low-degree polynomial is fit to a subset of the data, with explanatory variable values near the point whose response is being estimated. The polynomial is fit using weighted least squares, giving more weight to points near the point whose response is being estimated and less weight to points further away. The value of the regression function for the point is then obtained by evaluating the local polynomial using the explanatory variable values for that data point. The LOESS fit is complete after regression function values have been computed for each of the *n* data points. (CLEVELAND, 1979). Many of the details of this method, such as the degree of the polynomial model and the weights, are flexible.

The subsets of data used for each weighted least squares fit in LOESS are determined by a nearest neighbors' algorithm. A user-specified input to the procedure called the "bandwidth" or "smoothing parameter" determines how much of the data is used to fit each local polynomial. The smoothing parameter, *q*, is a number between *(d+1)/n* and 1, with *d* denoting the degree of the local polynomial. The value of *q* is the proportion of data used in each fit. The subset of data used in each weighted least squares fit is comprised of the nq (rounded to the next largest integer) points whose explanatory variables values are closest to the point at which the response is being estimated.

The local polynomials fit to each subset of the data are almost always of first or second degree; that is, either locally linear (in the straight-line sense) or locally quadratic. Using a zero-degree polynomial turns LOESS into a weighted moving average. Such a simple local model might work well for some situations but may not always approximate the underlying function well enough. LOESS is based on the ideas that any function can be well approximated in a small neighborhood by a low-order polynomial and that simple models can be fit to data easily. High-degree polynomials

would tend to overfit the data in each subset and are numerically unstable, making accurate computations difficult.

As mentioned above, the weight function gives the most weight to the data points nearest the point of estimation and the least weight to the data points that are furthest away. The use of the weights is based on the idea that points near each other in the explanatory variable space are more likely to be related to each other in a simple way than points that are further apart. The weight for a specific point in any localized subset of data is obtained by evaluating the weight function at the distance between that point and the point of estimation, after scaling the distance so that the maximum absolute distance over all of the points in the subset of data is exactly one.

The biggest advantage LOESS has over many other methods is the fact that it does not require the specification of a function to fit a model to all of the data in the sample (CLEVELAND, 1979). Instead, only the smoothing parameter value and the degree of the local polynomial has to be provided. In addition, LOESS is very flexible, making it ideal for modeling complex processes for which no theoretical models exist. These two advantages, combined with the simplicity of the method, make LOESS a very attractive regression method for applications that fit the general framework of least squares regression but which have a complex deterministic structure, like bioprocesses.

## 2.4. Methods for uptake and excretion rate estimation

GADEN (1955) defines specific substrate consumption ($\mu_S$) and product formation ($\mu_P$) rates as:

$$\mu_S = \frac{1}{X}\left(-\frac{dS}{dt}\right) \tag{8}$$

$$\mu_P = \frac{1}{X}\left(\frac{dP}{dt}\right) \tag{9}$$

with X being the microbial concentration and P the product concentration in a certain time.

Considering a bioprocess with $t$ duration, the corresponding values of X, P and S can be related to each other through the factors, expressing their instantaneous values by:

$$Y_{X/S} = \left(-\frac{dX}{dS}\right) = \frac{\mu_X}{\mu_S} \tag{10}$$

$$Y_{X/P} = \left(\frac{dX}{dP}\right) = \frac{\mu_X}{\mu_P} \tag{11}$$

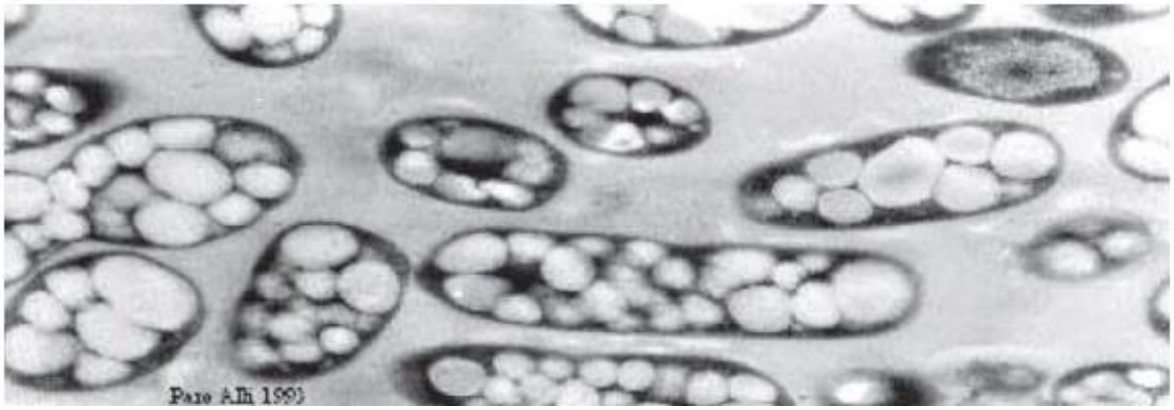
The factors can be estimated through a linear approximation ($X = Y_{X/S}.S \; or \; X = Y_{X/S}.P$) in which their values are the slope. Replacing them in Equations (8) and (9) will provide the specific rates $\mu_S$ and $\mu_P$.

# 3. CASE STUDY: POLYHYDROXYALKANOATES PRODUCTION

## 3.1. Bioprocess description

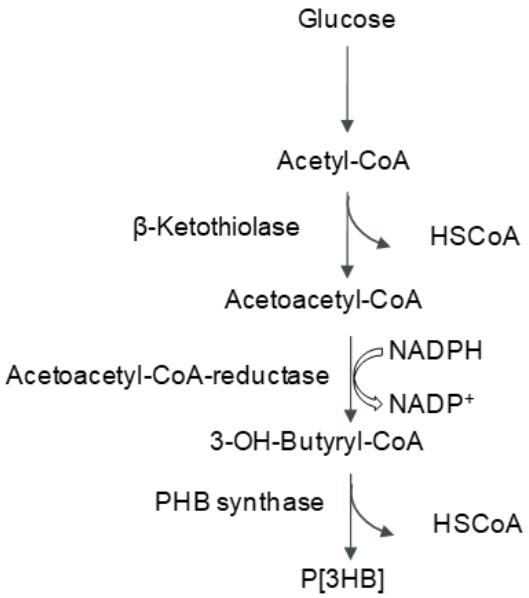
Polyhydroxyalkanoates (PHAs) are biopolyesters stored within cells as carbon and energy storage materials by various bacteria (ANDERSON and DAWES, 1990). More than 150 monomers have been detected in PHA produced by bacteria. The monomeric composition of PHA determines its physical-mechanical properties and, consequently, the range of applications. Due to their biocompatibility and biodegradability, PHAs have a wide range of applications in various industries such as biomedical sector including tissue engineering, bio-implant patches, drug delivery, surgery and wound dressing (RAZAA et *al*. 2018). PHAs are green plastics and they have positive social and environmental impact when compared with conventional plastics in terms of life expectancy. In contrast to traditional plastics, which have a life expectancy of between 100 and 1000 years, PHA-based bioplastic may degrade into water (H2O) and carbon dioxide (CO2) in 20 to 45 days if there is adequate humidity, oxygen, and a sufficient number of microbes in the open environment (MOSHOOD et *al*., 2022).

Figure 2. Intracellular accumulation of biopolymer granules.



(Source: SILVA et *al*., 2007).

Figure 3. PHB Biosynthesis.




(Source: Adapted from MADISON and HUISMAN,1999).

Poly-3-hydroxybutyrate (P3HB) is the most studied polymer from the PHA class and can be produced by bacteria belonging to quite diverse taxonomic groups, with *Ralstonia eutropha* being a model of study due to its great efficiency in the production of P3HB and its copolymers (STEINBÜCHEL and SCHLEGEL, 1991; REINECKE AND STEINBUCHEL, 2009). Its production is usually carried out in a two-stage fed-batch process. In the first stage the bacteria are provided with both the carbon and mineral nutrient sources to favor the growth and obtain a large concentration of biomass. In the second stage, taking advantage of the positive effect it occurs in the absence of nitrogen (ANDERSON and DAWES, 1990), the bioreactor is only fed with carbon source. When there is too much carbon supply and limitation of an essential nutrient, such as nitrogen, phosphorus, magnesium or sulfate, for example, the cell deviate acetyl-CoA to produce P3HB. In this stage, called the production phase, there is no cell proliferation. Figure 3 shows the biosynthesis of P3HB.

### 3.2. Relevance of the chosen bioprocess

Biodegradability is the ability of materials to decompose, especially in innocuous products, by the action of living beings such as microorganisms. Bacteria and fungi are the main participants in the process of biodegradation in nature. The decomposition of

materials provides microorganisms with precursors for cellular components and energy. Thus, biodegradability is nothing more than a catabolic process (BRAUNEGG et *al*., 1998).

The property that distinguishes PHAs from plastics of petrochemical origin is their biodegradability, becoming an important commercial feature. The PHAs are degraded in various environments, such as: soils, sewage, sea water and lakes. The biodegradation depends on several factors, such as: microbial activity in the environment, humidity, temperature, pH and molar mass. It was found that the nature of the monomer units also affect degradation. Microorganisms excrete enzymes that break down polymers in molecular blocks, called hydroxy acids, which are used as a carbon source for growth. The main enzyme for the degradation of P3HB and oligomers derived from polymers is PHB depolymerase (KHANNA and SRIVASTAVA, 2005a).

PHAs are considered of great industrial interest, as polymers biodegradable and/or biocompatible, for different areas of application. The biodegradable plastic must complement the petrochemical, conquering specific spaces where its purity and biodegradability characteristics are required.

Due to the characteristics of P(3HB) and P(3HB-*co*-3HV) polymers, initially these polymers were used in the manufacture of bottles, films and fibers for packaging biodegradable and protective bags for plants. In addition to these uses, PHAs films can be applied to paper or cardboard forming an impermeable film and producing a completely biodegradable composite material, which is an alternative to non-biodegradable compounds prepared from, for example, cardboard plus polyethylene or aluminum (STEINBÜCHEL and FÜCHTENBUSCH, 1998). Applications similar to the ones of conventional plastic products include disposable items such as razors, utensils, diapers, feminine hygiene products, cosmetic containers-shampoo bottles and cups (KHANNA and SRIVASTAVA, 2005a).

PHAs also have their applications in the medical field; many studies have shown clearly that biopolymers possess the biodegradability, biocompatibility and thermoprocessability for not only applications in implants, but also in the use of controlled release of drugs, being an attractive material for application in drug engineering. P(3HB) has been used as a suture thread, grafts cardiovascular devices, orthopedic pins, tissue regeneration and repair projects, bone repair articular cartilage, tendon repair and wound dressings (CHEN and WU, 2005).

## 3.3. Experimental conditions

### 3.3.1. Microorganism

*Ralstonia eutropha* was selected for this study due to its potentially high cell P(3HB) content and simple nutritional and culture requirement (SCHLEGEL et *al*. 1961). The strain LFM026 (reference DSM545) was used, which consumes glucose as carbon source being a H1 strain mutant.

### 3.3.2. Medium and culture conditions

Lysogeny broth (LB) complex culture medium was used for cell recovery and inoculum preparation (SAMBROOK et al., 1989). Mineral medium was used in cultures aiming growth and P(3HB) production. The media compositions are detailed in Tables 1, 2 and 3.

Table 3. LB medium composition

| Component | Concentration (g/L) |
|---|---|
| Tryptone | 10 |
| Yeast Extract | 5 |
| NaCl | 5 |
| Agar (only for solid media) | 20 |

(Source: own elaboration).

Table 4. Mineral medium composition used in pre-inoculum.

| Component | Concentration |
|---|---|
| $(NH_4)_2SO_4$ | 1 g/L |
| $Na_2HPO_4$ | 3.5 g/L |
| $KH_2PO_4$ | 1.5 g/L |
| $MgSO_4.7H_2O$ | 0.2 g/L |
| $CaCl_2.2H_2O$ | 0.01 g/L |
| $(NH_4)_5[Fe(C_6H_4O_7)_2]$ | 0.06 g/L |
| Trace elements solution* | 1mL/L |

(Source: RAMSAY et al., 1990).

* Trace element solution: $H_3BO_3$(0.30 g/L); $CoCl_2$.6H.O (0.20 g/L); $ZnSO_4H_2O$(0.10 g/L); $MnCl_2.4H_2O$ (0.03 g/L); $NaMoO_4.2H_2O$ (0.03 g/L); $NiCl_2.6H_2O$ (0.02 g/L); $CuSO_4.5H_2O$ (0.01 g/L).

Table 5. Mineral medium composition used in bioreactor experiments.

| Component | Concentration |
|---|---|
| $(NH_4)_2SO_4$ | 2 g/L |
| $KH_2PO_4$[a] | 0.5 g/L |
| $MgSO_4.7H_2O$ [a] | 0.2 g/L |
| $CaCl_2.2H_2O$ [a] | 0.01 g/L |
| NaCl | 4.0 g/L |
| $(NH_4)_5[Fe(C_6H_4O_7)_2]$ | 0.06 g/L |
| Glucose (inicial) [a] | 20 g/L |
| Trace elements solution* | 2 mL/L |

(Source: RAMSAY et al., 1990).

* Trace element solution: $H_3BO_3$(0.30 g/L); $CoCl_2$.6H.O (0.20 g/L); $ZnSO_4H_2O$(0.10 g/L); $MnCl_2.4H_2O$ (0.03 g/L); $NaMoO_4.2H_2O$ (0.03 g/L); $NiCl_2.6H_2O$ (0.02 g/L); $CuSO_4.5H_2O$ (0.01 g/L).
[a]Autoclaved separately

### 3.3.3 Fed-batch cultivation in the bioreactor

Four sets of experiments in bioreactors were carried out, in which two of them used the mineral medium and the other two had their medium modified in ammonium ferric citrate concentration (1.25 mg/L and 0 mg/L, respectively). The concentration of ammonium ferric citrate was varied in order to evaluate its effect on the production of the biopolymer and oxygen intake, which studies were carried out by another graduate student. Table 4 summarizes the experiments and the differences amongst them.

Table 6. Summary of the bioreactor experiments.

| Experiment | $(NH_4)_5[Fe(C_6H_4O_7)_2]$ | Fed-batch extent | Feeding glucose |
|---|---|---|---|
| R1 | 60 mg/L | 52 h | 500 g/L |
| R2 | 1.25 mg/L | 52 h | 500 g/L |
| R3 | 60 mg/L | 51.5 h | 600 g/L |
| R4 | 0 mg/L | 51.5 h | 600 g/L |

(Source: own elaboration).

The experiments were performed according to the following steps:

- Pre-inoculum: cultures stored in a freezer at -80°C were inoculated in solid LB medium and kept for 48 hours in an oven at 30°C. Isolated colonies from the solid medium were inoculated into 25 mL of LB medium and grown for 12 hours at 30°C and stirring at 150 rpm.
- Inoculum: the total volume of the pre-inoculum was transferred to an Erlenmeyer flask containing 225 mL of Mineral Medium containing 6 g/L of carbon source (glucose) and 3 g/L of nitrogen source (ammonium sulfate), and kept under stirring (150 rpm) at 30°C for 12 hours.

After growth, the culture was inoculated into the bioreactor, representing a fraction of 10% of the total culture volume. The experiment was conducted in fed-batch mode with pulses of glucose, keeping its concentration above 7 g/L in the bioreactor. For this purpose, a solution of 500 g/L glucose was prepared and added throughout cultivation.

During the culture, pH and temperature were controlled at 7.00 ± 0.05 and 30.0 ± 0.5°C, respectively. For pH control, solutions of NaOH (2 M) and $H_2SO_4$ (1.0 M) were used.

Dissolved oxygen was maintained above 30% of saturation, ensuring that there was no limitation of this compound. Cultivations were performed in 7-liter Ez-Control Applikon® Biotechnology bioreactors, using 3L as the total volume of cultivation. Periodic sampling was performed for further analysis of cell biomass, P(3HB) content and concentration of carbon source. $CO_2$ and $O_2$ were also analyzed in the reactor exaust gas. Table 5 summarizes the operational conditions of all the experiments.

Table 7. Operational conditions of all the experiments.

| Parameter | Value |
|---|---|
| Total volume (mL) | 3000 |
| pH | 7.00 ± 0.05 |
| Dissolved oxygen (%) | > 30 |
| Aeration (Lpm) | 2.0 |
| Stirring (rpm) | Variable |
| Initial glucose (g/L) | 20.0 |
| Temperature (°C) | 30.0 ± 0.5 |

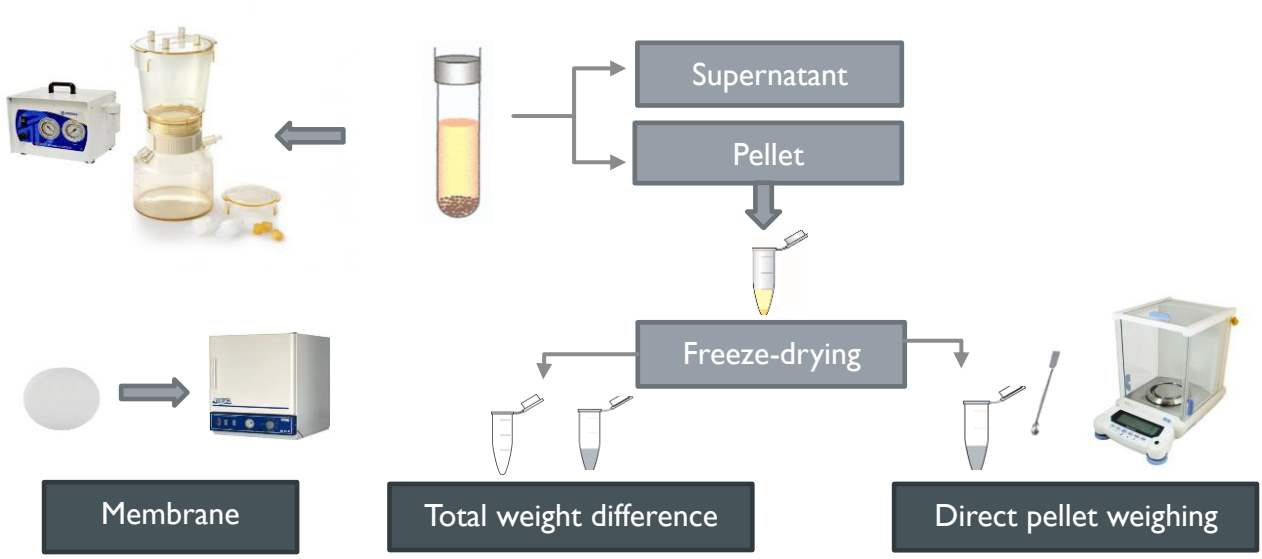
(Source: own elaboration).

## 3.4. Measured and calculated variables

### 3.4.1. Cellular dry mass

Total cell dry mass was determined by three different approaches, which are represented in Figure 2. In all methods, the first step consisted on centrifuging 5 or 10 mL of the culture (8500x*g*, 10 min, 4°C). The cells were then resuspended with physiological solution (0.85% NaCl), washed with the same physiological solution and transferred either to microtubes (2 mL) or to a membrane (0.45 µm porosity). In the membrane filter technique, sample was passed through the membrane using a filter funnel and vacuum system. The membrane was then dried until stable weight in a drying oven and weighted. Prior to the experiments, membrane humidity was determined and considered in the final dry cell mass weight. In the other methods, the microtubes were centrifuged for 10 minutes (13000 rpm), cells were resuspended with physiological solution (0.85% NaCl), washed with the same physiological solution and transferred to microtubes (2 mL). The supernatant was discarded and the microtubes frozen for later freeze drying (Triad Labconco, Missouri, USA). The mass was obtained either removing and weighing the dry cell mass from the microtubes (direct pellet weighing in this work) or by the difference in the microtubes without and with the cell mass (total weight difference in this work). The residual dry mass corresponds to the total dry mass, discounting the accumulated P(3HB) mass.

Figure 4. Cell dry mass determination methods used in this work.




(Source: own elaboration)

### 3.4.2. Glucose concentration

Glucose concentration was determined on Ultimate 3000 Liquid Chromatograph (HPLC) (Dionex, Sunnyvale, CA, United States) with Refractive Index Detector (Shodex, Kawasaki, Kanagawa, Japan) at 35°C, a column Aminex HPX- 87H (Bio-Rad, Hercules, CA, United States) at 45°C, mobile phase 5 mM $H_2SO_4$ solution with a flow rate of 0.6 mL/min, and injection volume of 20 µL. Carbohydrates were detected by refractive index. The standard curve was built at concentrations up to 20 g/L and samples were diluted when necessary, so that their concentration could be within the curve.

### 3.4.3. PHA amount and composition

The amount and composition of PHA was determined by propyl ester gas chromatography (RIIS and MAI, 1988). About 10 to 15 mg of freeze-dried cells were transferred to tubes to which 100 µL of benzoic acid (40 mg/mL in propanol), 2 mL of 1,2-dichloroethane and 2 mL of hydrochloric acid in propanol solution (1:4 v/v) were added. The tubes were in a heated bath (100ºC) for 3 hours and agitated after the first 30 minutes. After, they were kept out of the bath to cool and then 4 mL of distilled water

were added to separate the aqueous phase. This phase was discarded and to the remaining mixture was added anhydrous sodium sulfate for adsorption of remaining water in the organic phase containing the propyl esters. The resulting mixture was transferred to vials. A volume of 0.2 µL of the organic phase was analyzed after sample fractionation (fractionation 1:5) in a HP7890A GC System gas chromatograph (Agilent Technologies, Santa Clara, CA, United States), equipped with an HP-1 capillary column (length 30 m diameter 0.25 mm and film thickness 0.25 µm). The analysis was conducted under the following conditions: carrier gas - Nitrogen (1.0 mL/min); injector temperature - 250°C; detector temperature - 300°C; detection system - flame ionization (FID); oven temperature program: 100°C for 3 minutes (6°C / min), temperature rise to 180°C for 5 minutes (6°C / min) and up to 240°C for 1 minute.

### 3.4.4. Maximum specific growth rate

The specific growth rate was calculated through adjustment of experimental data by an exponential curve and applying Loess smoothing method followed by differentiation. The latter was done using a script developed in MATLAB by PAZ et al. 2021 and the specific growth rate was the average of the values obtained during the exponential growth phase. In both cases, the exponential growth phase was defined through the identification of a straight line interval in a plot of lnX vs t.

### 3.4.5. Carbon dioxide and oxygen in bioreactor outlet gas

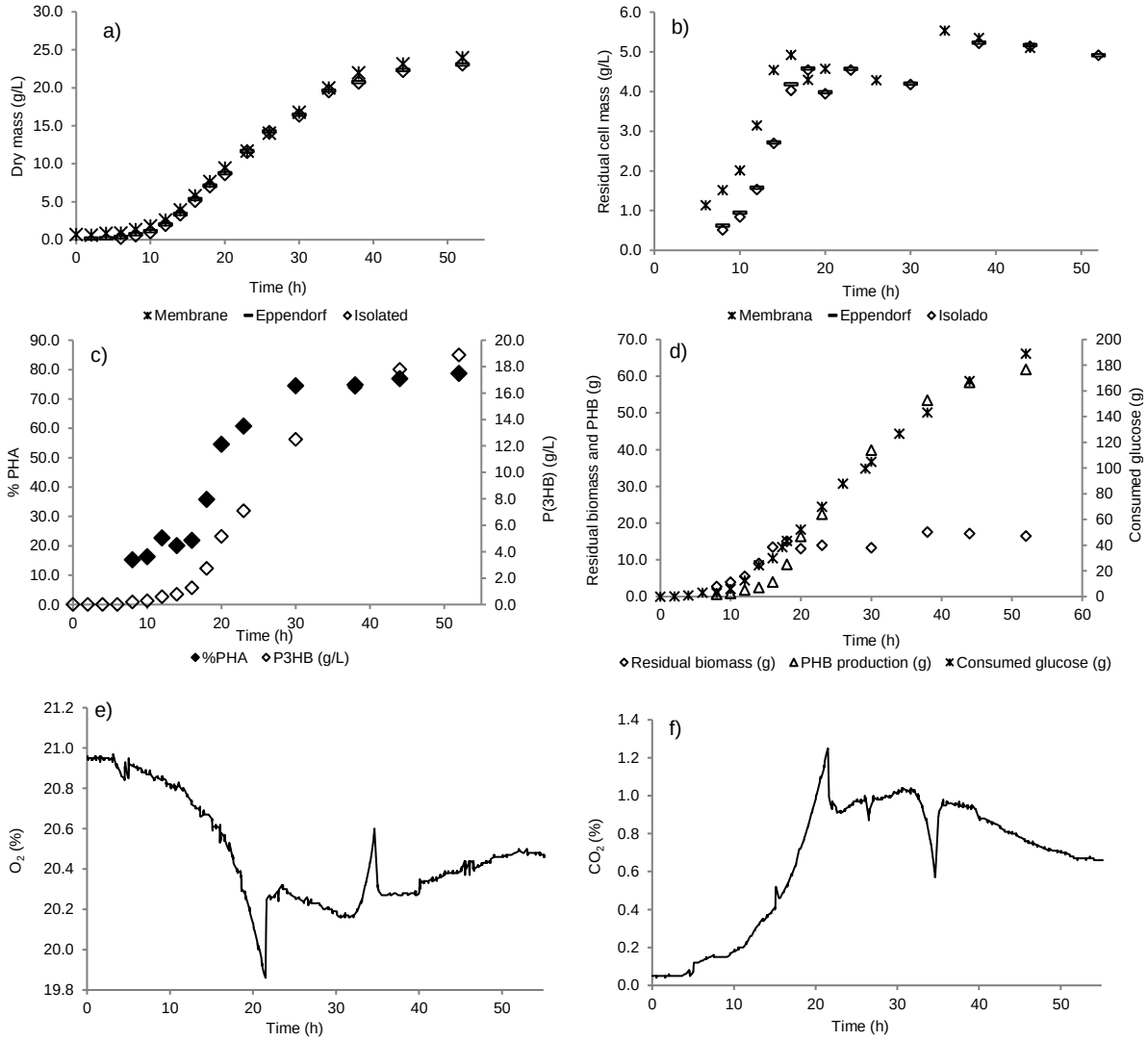
Partial $CO_2$ and $O_2$ pressures in the bioreactor outlet gas were determined using $CO_2$ and $O_2$ gas sensors (BlueSens - Applikon, Netherlands). The principle of measurement of $CO_2$ and $O_2$ was infrared (dual wavelength) and zirconium dioxide, respectively.

# 4. RESULTS AND DISCUSSION

## 4.1. Fed-batch cultivations in bioreactor

The results from the experiment R1 can be seen in Figure 6. The production of P(3HB) is carried out in two stages. In the first stage, called the growth period, the bacteria are provided with both the carbon and mineral nutrient sources to favor the growth and obtain a large concentration of biomass. In the second stage, when the bioreactor is only fed with carbon source, there is too much carbon supply and limitation of nitrogen, there is production of P(3HB) and no cell proliferation (ANDERSON and DAWES, 1990). As seen in Figure 6b, the growth period lasted about 16h (when cells reached a plateau of growth after that period) and P(3HB) production reached 20 g/L approximately (80% of cell composition) at the end of the experiment. However, some discrepancies amongst the cell mass determination methods occurred. Analyzing the residual dry mass graph, there is a decrease in the mass during 20 and 30h, which can be associated with analytical errors during P(3HB) quantification analysis.

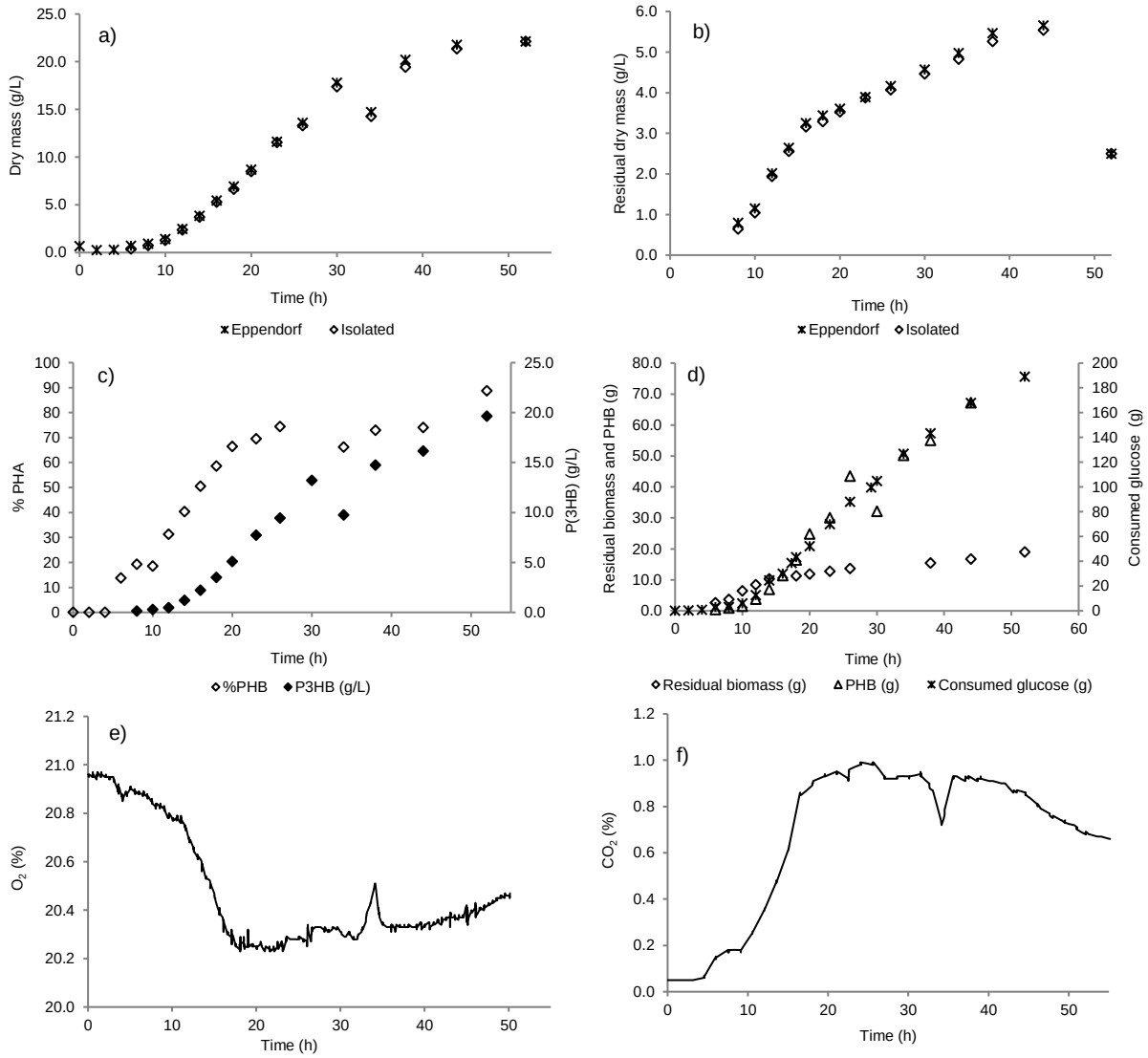
Figure 5. Results from experiment R1 - mineral medium with 60mg/L of ammonium ferric citrate: a) Total dry cell mass and b) Residual dry cell mass with different methods for cell concentration determination; c) PHB amount and concentration throughout the cultivation; d) Curves for residual biomass, PHB production and consumed glucose. e) Oxygen and f) Carbon dioxide during the experiment in outlet gas.




(Source: own elaboration)

Comparing the results with experiment R2 (Figure 7), the latter presented a growth period of about 20h (Figure 7b) reaching the same P(3HB) concentration of the previous experiment at the end of it, although with a higher cell mass composition (90%). Less discrepancies (standard deviations) among the cell mass determination methods occurred that could be due to having the same operator manipulating the samples.

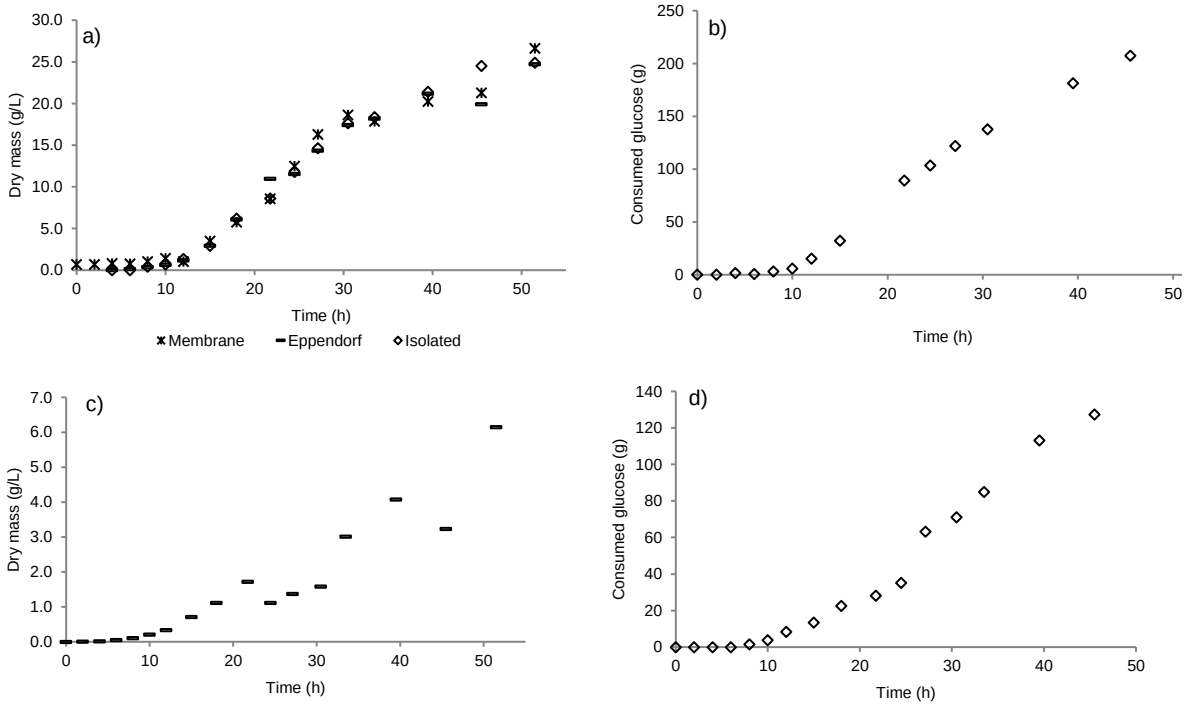
Figure 6. Results from experiment R2 - mineral medium with 1.25 mg/L of ammonium ferric citrate: a) Total dry cell mass and b) Residual dry cell mass with different methods for cell concentration determination; c) PHB amount and concentration throughout the cultivation; d) Curves for residual biomass, PHB production and consumed glucose. e) Oxygen and f) Carbon dioxide during the experiment in outlet gas. Some peaks can be observed in the outlet gases curve and they are due to filter changes during the experiment.




(Source: own elaboration)

Another set of experiments (R3 and R4) were carried out, which results are shown in Figure 8. Differently from the previous experiment sets, the methods of cell concentration presented themselves very similar to one another.

Figure 7. Results from experiments R3 (mineral medium with 60mg/L of ammonium ferric citrate - first row) and R4 (mineral medium without ammonium ferric citrate - second row): a) and c) Total dry cell mass applying different methods for cell concentration determination; b) and d)Consumed glucose during the bioprocess.

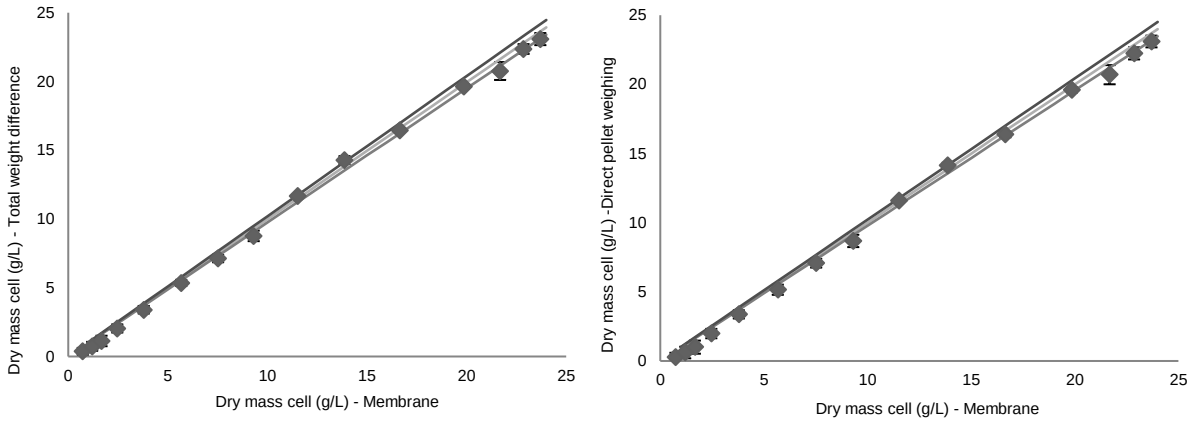



(Source: own elaboration)

## 4.2. Comparison of cell mass determination

One of the specific objectives of the current work is to make a correlation between the different methods used to determine cell mass concentration. In order to do so, data from experiment R1 were analyzed and can be seen in Figure 9. The membrane method was considered to be the golden procedure and the correlation with the other two methods was drawn. For small values of concentration, a higher difference is noticed (higher standard deviation) when compared to the dry mass cell results obtained using a membrane. The direct pellet weighing method presented the greatest difference, resulting in the lowest values of measured dry cell mass. However, attention is needed concerning the membrane method, since the membrane can absorb humidity easily after being dried in the oven.

Figure 8. Comparison of cell mass determination methods. The line in the center represents a linear correlation between the two variables accompanied by the standard deviations (upper and lower lines).




(Source: own elaboration)

## 4.3. Carbon mass balance

A major reason for applying mass balance to experimental data is to check the consistency of the data, since experiments can “go wrong” for many different reasons (products thought to remain in the liquid phase are partly stripped to the gas phase; the measurement instruments are wrongly calibrated, or they suddenly present malfunction).

Typically, chemical reactions are written as a single or a system of stoichiometric equations into which elements should be conserved. For an element such as C, it can shift from one compound to another due to chemical reaction, but the total amount is conserved. Therefore, the system can be written as:

$$rate\ of\ input\ of\ carbon\ = rate\ of\ accumulation\ of\ carbon + rate\ of\ output\ of\ carbon \quad (12)$$

Considering the aspects of the current work, in which samples were removed (labelled as “sampled”) and pulses of glucose were added (labelled as “added”) in some periods of time, Eq 12 becomes:

$$S_{added} + X_{added} = S_{(sampled+bioreactor)} + X_{(sampled+bioreactor)} + P_{(sampled+bioreactor)} + CO_2 \quad (13)$$

where S is the substrate (glucose), X is the residual biomass and P is the product (P(3HB)). The label “bioreactor” refers to the amount present in the bioreactor at the moment. Since it is a fed-batch bioreactor, all values were in grams.

For the carbon balance in this work, it was considered the information contained in Table 6, regarding carbon percentage in the molecules. Applying this information in Eq. 13, in terms of mass of carbon (g):

$$0.4S_{added} + 0.48X_{added} = 0.4S_{(sampled+bioreactor)} + 0.48X_{(sampled+bioreactor)} + 0.56P_{(sampled+bioreactor)} + 0.27CO_2 \quad (14)$$

The results of carbon balance are presented in Tables 7 and 8 and they expressed in grams of carbon. The difference in percentage presented in the last column of the tables is related to the initial and final amount of carbon. The carbon balance was carried out for sets R1 and R2. Up to time 6h, no P(3HB) production was quantified. The amount of $CO_2$ produced was calculated through the integration of the $CO_2$ curve and the area under the curve was given by the MATLAB function *trapz*, which performs discrete integration by using the data points to create trapezoids, so it is well suited to handling data sets with discontinuities.

Table 8. Carbon amount in some molecules.

| Molecule | g/mol | # of C in the molecule | % C weight |
|---|---|---|---|
| Glucose | 180 | 6 | 40.00 |
| P(3HB) | 86 | 4 | 55.81 |
| $CO_2$ | 44 | 1 | 27.27 |
| Biomass* | 100 | 4* | 48.00 |

(Source: own elaboration.)

* Hypothesis: biomass has 48%wt of carbon (NEIDHARDT, 1990).

From the results, it can be said that the experimental data for R2 is more consistent than data from R1, since it presents lower values of differences between the initial and output values of carbon, strongly suggesting that all the carbon introduced and present at the beginning of the period was measured as output. After 30h, carbon balance for R1 was higher than 17%, indicating measurement inaccuracy in P(3HB) or dry cell mass quantification.

Regarding 14h in R2, it was expected a lower amount of glucose in the bioreactor than what was found (12.6 g/L). This is supported by the fact that 2 hours later, glucose concentration in bioreactor was 7.2 g/L and a pulse of glucose was needed. The higher value found can be explained by contamination of a more concentrated glucose solution.

The differences in carbon balance suggest the presence of experimental errors and possible noise, highlighting the importance of this analysis prior to specific rate estimations, assuring the use of reliable data. As seen further in this work, some procedure improvements were identified and a data smoothing was applied before estimating the specific growth rate.

Table 9. Carbon balance for experiment R1 expressed in grams of C. Input (whether in the form of added substrate or biomass) and output (substrate in the bioreactor, biomass that was sampled and left in the bioreactor, and product that was samples and what was present in the bioreactor) are expressed in grams of carbon. The difference is between the total input and output in terms of grams of carbon.

| Time (h) | Input | | Output | | | | | | | Total input | Total output | Δ% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | S | X | $S_{sampled}$ | $S_{biorreactor}$ | $X_{sampled}$ | $X_{biorreactor}$ | $P(3HB)_{sampled}$ | $P(3HB)_{biorreactor}$ | $CO_2$ | | | |
| 0-8 | 24.06 | 1.08 | 0.85 | 22.05 | 0.01 | 1.72 | 0.003 | 0.360 | 0.968 | 25.14 | 25.96 | 3% |
| 0-10 | 24.06 | 1.08 | 1.00 | 20.69 | 0.03 | 2.28 | 0.006 | 0.515 | 1.432 | 25.14 | 25.95 | 3% |
| 0-12 | 24.06 | 1.08 | 1.13 | 17.86 | 0.05 | 3.01 | 0.014 | 1.025 | 2.045 | 25.14 | 25.13 | 0% |
| 0-14 | 24.06 | 1.08 | 1.23 | 13.06 | 0.09 | 4.68 | 0.024 | 1.365 | 2.903 | 25.14 | 23.35 | 7% |
| 0-16 | 24.06 | 1.08 | 1.32 | 10.78 | 0.14 | 6.75 | 0.043 | 2.190 | 4.088 | 25.14 | 25.30 | 1% |
| 0-18 | 42.22 | 1.08 | 1.56 | 24.47 | 0.20 | 7.46 | 0.082 | 4.834 | 5.903 | 43.30 | 44.51 | 3% |
| 0-20 | 42.22 | 1.08 | 1.71 | 20.73 | 0.25 | 6.45 | 0.151 | 9.013 | 8.256 | 43.30 | 46.57 | 8% |
| 0-23 | 42.22 | 1.08 | 1.82 | 13.54 | 0.30 | 6.81 | 0.244 | 12.276 | 12.184 | 43.30 | 47.18 | 9% |
| 0-26 | 42.22 | 1.08 | 1.87 | 6.27 | 0.38 | 9.54 | 0.345 | 12.933 | 16.027 | 43.30 | 47.37 | 9% |
| 0-30 | 60.66 | 1.08 | 2.01 | 18.96 | 0.43 | 6.45 | 0.505 | 21.873 | 21.985 | 61.74 | 72.20 | 17% |
| 0-34 | 60.66 | 1.08 | 2.08 | 10.12 | 0.51 | 10.64 | 0.678 | 22.330 | 26.812 | 61.74 | 73.17 | 19% |
| 0-38 | 79.04 | 1.08 | 2.19 | 22.87 | 0.55 | 8.49 | 0.810 | 29.313 | 32.520 | 80.12 | 96.74 | 21% |
| 0-44 | 79.04 | 1.08 | 2.24 | 13.06 | 0.58 | 8.24 | 0.945 | 31.874 | 40.234 | 80.12 | 97.18 | 21% |
| 0-52 | 79.04 | 1.08 | 2.28 | 4.52 | 0.64 | 7.85 | 1.058 | 33.765 | 49.443 | 80.12 | 99.56 | 24% |

(Source: own elaboration.)

Table 10. Carbon balance for experiment R2 expressed in grams of C. Input (whether in the form of added substrate or biomass) and output (substrate in the bioreactor, biomass that was sampled and left in the bioreactor, and product that was samples and what was present in the bioreactor) are expressed in grams of carbon. The difference is between the total input and output in terms of grams of carbon.

| Time (h) | Input | | Output | | | | | | | Total input | Total output | Δ% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | S | X | $S_{sampled}$ | $S_{biorreactor}$ | $X_{sampled}$ | $X_{biorreactor}$ | $P(3HB)_{sampled}$ | $P(3HB)_{biorreactor}$ | $CO_2$ | | | |
| 0-8 | 22.87 | 1.07 | 0.78 | 20.92 | 0.01 | 1.21 | 0.002 | 0.224 | 1.070 | 23.94 | 24.21 | 1% |
| 0-10 | 22.87 | 1.07 | 0.93 | 19.45 | 0.02 | 1.74 | 0.005 | 0.483 | 1.672 | 23.94 | 24.30 | 1% |
| 0-12 | 22.87 | 1.07 | 1.04 | 16.69 | 0.04 | 3.04 | 0.011 | 0.800 | 2.508 | 23.94 | 24.14 | 1% |
| 0-14 | 22.87 | 1.07 | 1.17 | 15.72 | 0.07 | 3.96 | 0.027 | 2.095 | 3.845 | 23.94 | 26.89 | 12% |
| 0-16 | 22.87 | 1.07 | 1.24 | 9.00 | 0.11 | 4.84 | 0.057 | 3.816 | 5.893 | 23.94 | 24.96 | 4% |
| 0-18 | 48.14 | 1.07 | 1.45 | 23.73 | 0.15 | 5.31 | 0.104 | 6.292 | 8.515 | 49.20 | 45.56 | 7% |
| 0-20 | 48.14 | 1.07 | 1.64 | 18.90 | 0.21 | 5.53 | 0.192 | 9.070 | 11.018 | 49.20 | 46.55 | 5% |
| 0-23 | 48.14 | 1.07 | 1.73 | 12.12 | 0.25 | 5.92 | 0.296 | 13.646 | 14.940 | 49.20 | 48.90 | 1% |
| 0-26 | 48.14 | 1.07 | 1.78 | 6.31 | 0.30 | 6.27 | 0.422 | 16.551 | 18.904 | 49.20 | 50.53 | 3% |
| 0-30 | 73.60 | 1.07 | 1.92 | 19.21 | 0.35 | 7.10 | 0.596 | 23.908 | 24.723 | 74.66 | 77.81 | 4% |
| 0-34 | 73.60 | 1.07 | 2.00 | 11.61 | 0.41 | 7.68 | 0.723 | 17.451 | 29.215 | 74.66 | 69.08 | 7% |
| 0-38 | 98.99 | 1.07 | 2.09 | 23.71 | 0.44 | 8.74 | 0.830 | 27.387 | 34.818 | 100.06 | 98.02 | 2% |
| 0-44 | 98.99 | 1.07 | 2.15 | 14.71 | 0.47 | 9.10 | 0.938 | 30.192 | 42.673 | 100.06 | 100.24 | 0% |
| 0-52 | 98.99 | 1.07 | 2.19 | 5.29 | 0.50 | 4.04 | 1.212 | 36.909 | 51.532 | 100.06 | 101.67 | 2% |

(Source: own elaboration.)

## 4.4. Improvements in carrying out the experiments

While carrying out the experiments, some improvements in sampling or procedures were taken in order to reduce the variability and fluctuations in data points. The first improvement was related to the initial samples taken from the bioreactor. We noticed that the initial samples were difficult to quantify the production of P(3HB) because of the biomass presented. So, taking higher sample volume would increase the mass available to further P(3HB) quantification analysis. The other improvement was related to the preparation of the glucose feed solution in grams of glucose per gram of solution. This step eliminates having to calculate the density of the solution, since the solution is very viscous and it can lead to error in calculation. Knowing the mass % of glucose in the feed and using a weighing scale to determine the mass of the glucose flask would allow good accuracy in measuring the glucose supply, without having to measure the density of the solution.

Applying the improvement conditions suggested above, two new bioreactor experiments were done following the conditions described in experiments R1, with each sampling being done in duplicate in each experiment. Figures 10 and 11 display the results for dry cell mass and P(3HB). The improvements were able to reduce the standard deviation between the duplicate (Figures 10 and 11), and both reactors showed similar behavior in dry cell mass and P(3HB) production yields. This was expected since both reactors operated under the same conditions. However, discrepancies in dry cell mass results for 40h were observed (very different values for the same point comparing reactor 1 and 2, even though they operated under the same conditions), consequently influencing the amount of P(3HB). Although unknown, a probable cause could be specific operator sample handling for this specific point.

Figure 9. Dry cell mass results for new experiments applying improvements in in sampling and procedures following the conditions described in experiments R1.

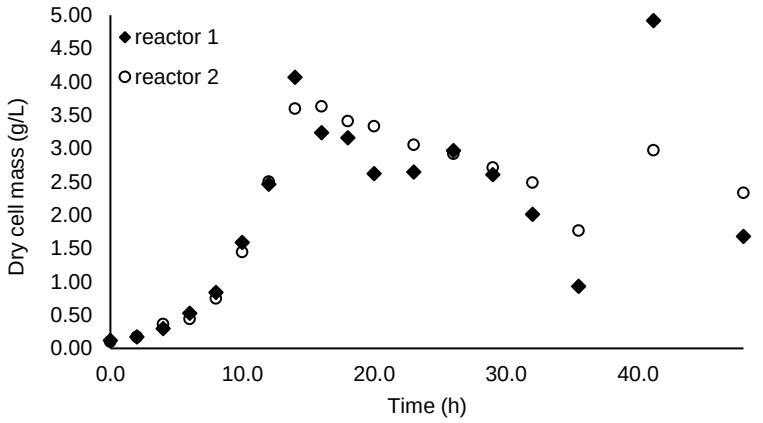
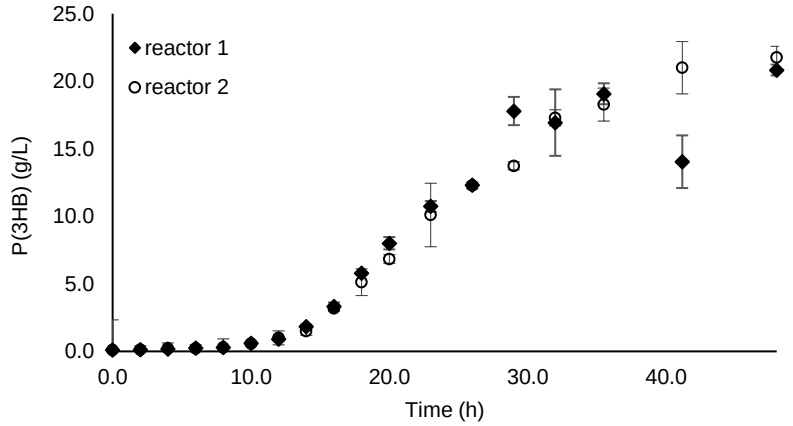



Figure 10. P(3HB) results for new experiments applying improvements in in sampling and procedures following the conditions described in experiments R1.




## 4.5. Comparison of different methods of specific rate determination

To compare the maximum specific growth rate ($\mu_x$), two methods were used: adjustment of experimental data by an exponential curve; and the Loess smoothing method, which transforms the experimental data into a smoothed curve using successive quadratic regressions, determined within certain intervals or windows, and which uses the weighted least squares method as an estimator (CLEVELAND, 1979). After smoothing, $\mu_x$ was estimated by the derivative of the polynomial function obtained ($\mu_X = \frac{1}{X}\frac{dX}{dt}$).

The data used for such comparison is the new set of experiments described and which results are displayed in item 4.6. The data comprises 18 experimental points,

and it was also evaluated the influence of the smoothing parameter *q* in the smoothing curve and in the specific growth rate estimation. The latter was obtained through the specific growth rate profile provided by the script used, and an average of the values during exponential growth was taken to account for the final estimate. As shown in Figure 10, the exponential growth interval in both reactors occurred between 2 and 12h.

The analyzes of the experimental samples had *q* varying between 22% and 83%, approximately (Figures 12 and 13) and large values of *q* produced the smoothest functions that wiggle the least in response to fluctuations in the data and that do not follow the experimental data; while small values of *q* is the closer the regression function will conformed to the data. This expected result is in accordance with Cleveland (1979) and shows that the parameter *q* controls the flexibility of the LOESS regression function. However, despite small values of *q* conforming better to the data, using too small a value of the smoothing parameter is not desirable, since the regression function will eventually start to capture the random error in the data. From the results obtained, when *q* is 44%, a satisfactory curve is displayed, capturing the biological behavior in P(3HB) production as well as reducing the fluctuations from the experimental data. This value of *q* is also in accordance with typical values for the smoothing parameter for most LOESS applications, which lie in the range 25% to 50% (Cleveland, 1979). The specific growth rate for each scenario was also estimated and is shown in Table 9.

Table 11. Specific growth rate values for different values of *q*.

| Smoothing parameter *q* | $\mu_x$ (g/g.h) reactor 1 | $\mu_x$ (g/g.h) reactor 2 |
|---|---|---|
| 22% | 0.2722 | 0.2730 |
| 33% | 0.2705 | 0.2693 |
| 44% | 0.2655 | 0.2681 |
| 56% | 0.2423 | 0.2424 |
| 67% | 0.3282 | 0.3113 |
| 83% | 0.6787 | 0.9006 |

Figure 11. Smoothing curves for dry cell mass in reactor 1 using the Loess method while varying the smoothing parameter *q* (continuous line shows smoothed points and the other points are experimental ones). Chosen values for *q* were: a) 22%, b) 33%, c) 44%, d) 56%, e) 67% and f) 83%.

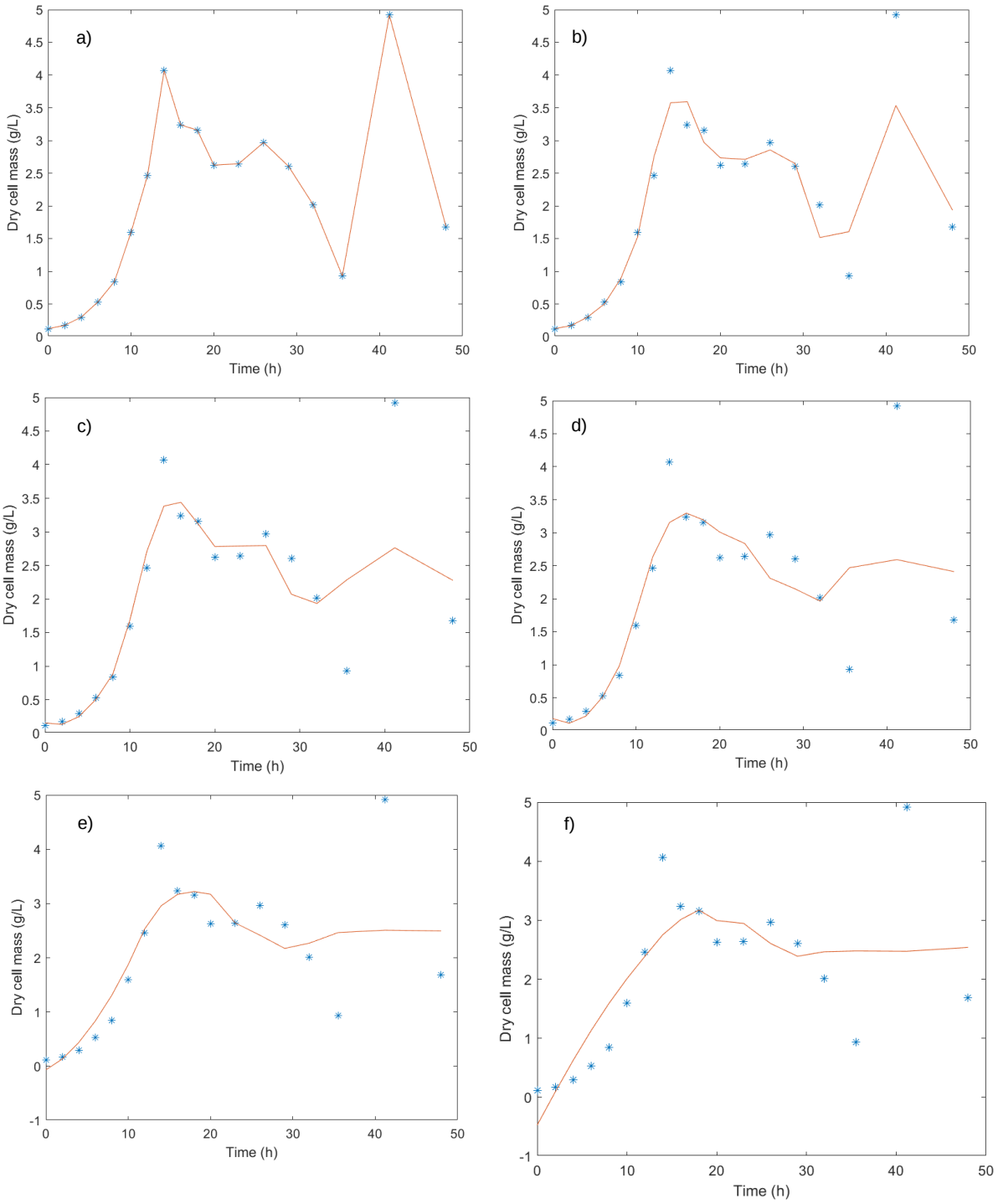
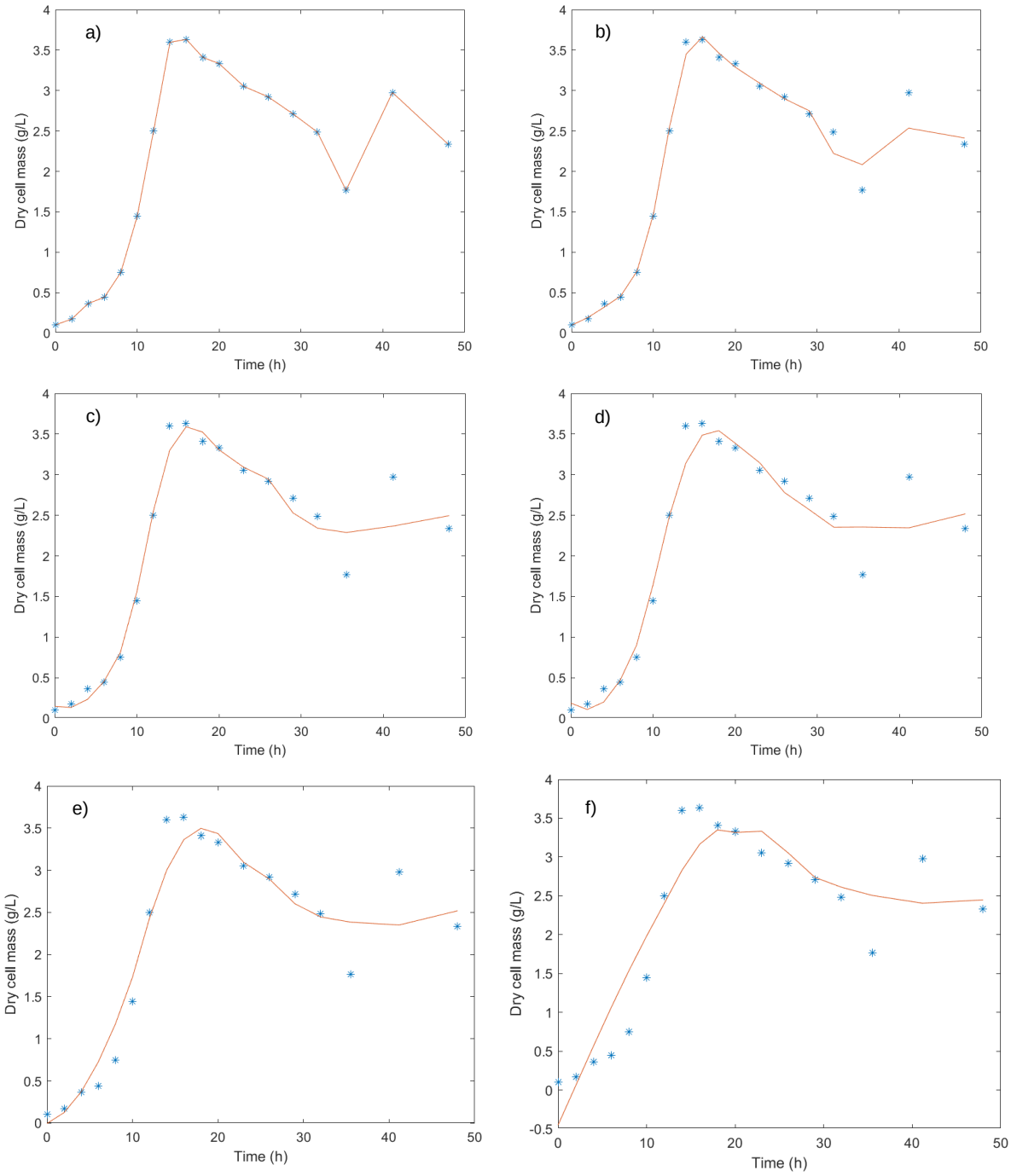


Figure 12. Smoothing curves for dry cell mass in reactor 2 using the Loess method while varying the smoothing parameter *q* (continuous line shows smoothed points and the other points are experimental ones). Chosen values for *q* were: a) 22%, b) 33%, c) 44%, d) 56%, e) 67% and f) 83%.



The exponential adjustment and the specific growth rate profile when *q* is 44% can be seen in Figures 14 and 15. It can be noted that the exponential adjustment resulted in very similar rate values (0.2577 g/g.h and 0.2696 g/g.h) to those obtained by the Loess method when *q* is 44% (0.2655 g/g.h and 0.2681 g/g.h for reactors 1 and 2, respectively). Previous studies with the same microorganism obtained 0.26 g/g.h, validating the application of Loess method for estimation of growth rates in the discussed bioprocess. It is worth mentioning that the exponential adjustment does not soften any possible noise present, unless any adjustment is made before, while the Loess method is capable of such fact. Even though further applications of Loess are needed to completely validate it, this study strongly indicates its suitability.

Figure 13. Specific growth rate by exponential adjustment: left is reactor 1 (0.2577 g/g.h), and right is reactor 2 (0.2696 g/g.h).

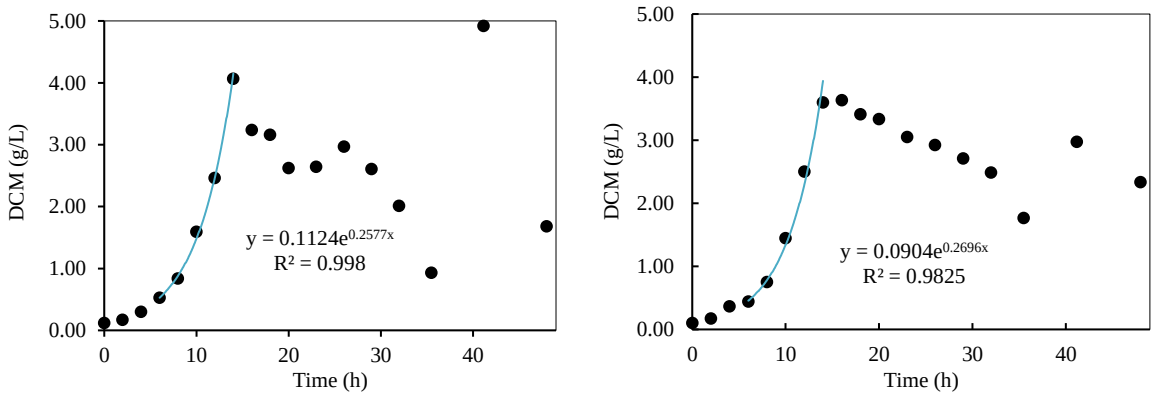
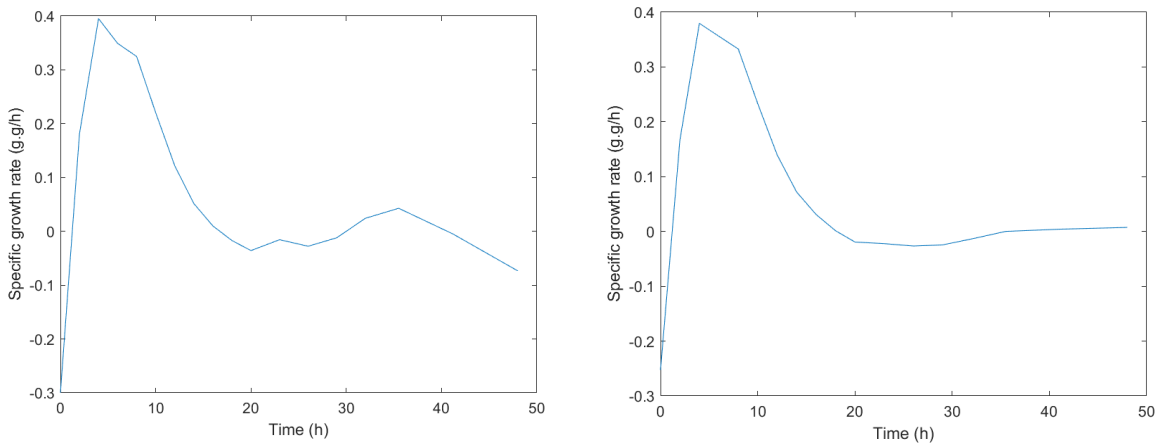



Figure 14. Specific growth rate profile when *q* is 44% for reactor 1 (left) and reactor 2 (right). Average of the values during exponential growth phase results in values for specific growth rate of 0.2655 g/g.h and 0.2681 g/g.h for reactors 1 and 2, respectively.



## 5. CONCLUSIONS

Production of P(3HB) by *Ralstonia eutropha* was selected to this project due its well established knowledge of the process and due to the fact that the microorganism is a model in P(3HB) production. Experiments were conducted in which was noted the impact of different methods for cell mass determination, being the membrane method the most suitable one with careful use to not absorb humidity. The carbon balance showed that the data are not consistent, indicating laboratory errors when performing some quantifications. The experimental improvements resulted in data that are more reliable and this indicates the need to consider the sampling procedure when determining the accuracy of the measurements. The Loess method proved to be satisfactory in calculating the specific growth rate for *q* being 44%, even for small data sets, indicating the importance of the choice of q in influencing these results, as well as the choice of methods for smoothing and determination of $\mu_x$.

In conclusion, having reliable and consistent data to estimate specific rates is essential to represent accurately the biological process. So far, it was seen that not only improvements in analytical procedures have to be made (such as taking larger amounts of sample in the first points of the experiment to increase cell mass values), but also applying smoothing methods to the experimental data before any further analysis.

## 6. REFERENCES

ABUD, A.K.S. Estudo do controle de Qualidade da Produção de L-Asparaginase por Zymomonas mobilis. PhD thesis. PEQ/ COPPE/UFRJ, Rio de Janeiro, 2005.

AKMAL, D; AZIZAN, M.N.; MAJID, M.I.A. Biodegradation of microbial polyesters P(3HB) and P(3HB-*co*-3HV) under the tropical climate environment. *Polymer Degradation and Stability*, v. 80, p. 513-518, 2003. https://doi.org/10.1016/S0141-3910(03)00034-X.

ANDERSON, A.J. and DAWES, E.A. Occurrence, metabolism, metabolic role, and industrial uses of bacterial polyhydroxyalkanoates. *Microbiologial Reviews*, v. 54, p. 450-472, 1990.

ATEKWANA, E.; ZHANG, C.; JARDANI, A.; SMITH, S. A new model for the spectral induced polarization signature of bacterial growth in porous media. *Water Resources Research*, v. 48, 2012.

BASTIN, G. and DOCHAIN, D. On-line estimation of microbial specific growth rates. *Automatica*, v. 22, p.705e9, 1986.

BASTIN, G. and DOCHAIN, D. On-line estimation and adaptive control of bioreactors. *Elsevier Science Publishers*, Amsterdam, 1990.

BAYER, B.; SISSOLAK, B.; DUERKOP, M.; VON STOSCH, M.; STRIEDNER, G. The Shortcomings of Accurate Rate Estimations in Cultivation Processes and a Solution for Precise and Robust Process Modeling. *Bioproc. Biosys. Eng.*, v. 43, p. 169-178, 2020.

BIAGIOLA, S.I.; FIGUEROA, J.L. State estimation in nonlinear processes. application to pH process control. *Ind Eng Chem Res,* v. 41, p. 4777–85, 2002.

BORZANI, W. Cinética de processos fermentativos. *Revista Brasileira de Engenharia*,v.3, p.1-51, 1986.

BOX, G.E.P., HUNTER J.S., HUNTER, W.G.Statistics for experimenters: design, innovation and discovery. Wiley, New York, 2005.

BRUNA, W.; BARRAL, M. F.; PRADELLA, J. G. C. Emerson. Aplicativo para cálculo de velocidades e velocidades especificas de reação em processos biotecnológicos. Relatório técnico IPT. São Paulo. 2004.

BUCHANAN, R. E. Life phases in a bacterial culture, *J. Infect. Dis*., v. 23, p. 109–125, 1918.

CINQUEMANI, E,; LAROUTE, V.; COCAIGN-BOUSQUET, M.; DE JONG, H.; ROPERS, D. Estimation of time-varying growth, uptake and excretion rates from dynamic metabolomics data. *Bioinformatics*, v. 33, p. 301–310, 2017.

CLAES, J.E.; IMPE, J.F.V. On-line estimation of the specific growth rate based on viable biomass measurements: experimental validation. *Bioprocess Eng,* v. 21, p. 389-395, 1999.

CLEVELAND WS, Robust Locally Weighted Regression and Smoothing Scatterplots. *J. A. Stat. Ass.*, v. 74, p. 829-836, 1979.

DE BATTISTA, H.; PICÓ, J.; GARELLI, F.; VIGNONI, A. Specific growth rate estimation in (fed-)batch bioreactors using second-order sliding observers. *J. Process Control,* v. 21, p. 1049–1055, 2011.

DOCHAIN, D. State and parameter estimation in chemical and biochemical processes: a tutorial. *Journal of Process Control*, v. 13, p. 801–818, 2003.

FARZA, M.; HAMMOURI, H.; OTHMAN, S.; BUSAWON, K. Nonlinear observers for parameter estimation in bioprocesses. *Chem Eng Sci,* v. 52, p. 4251–67, 1997.

FARZA, M.; OUEDER, M.; ABDENNOUR, R.B.; SAAD, M. High gain observer with updated gain for a class of MIMO nonlinear systems. *Int J Control,* v. 84, p. 270–80, 2011.

FERRAZ, L., FREITAS, E., AUGUSTO, E.F.P., BONOMI, A. Estratégias para cálculo de velocidades em processos fermentativos. estudo de métodos. In: Encontro de Áreas Técnicas da Divisão de Química, Publicação IPT 2.443, p.ll3-116, 1995.

GADEN, E.L.Jr. Fermentation kinetics and productivity. *Chemistry and Industry*, v. 12, p.154-9, 1955.

GLACKEN, M.W.; ADEMA, E.; SINSKEY, A.J. Mathematical descriptions of hybridoma culture kinetics: I. Initial metabolic rates. *Biotechnol. Bioeng.,* v. 32, p. 491–506, 1988.

GOUDAR, C.T. Computer programs for modeling mammalian cell batch and fed-batch cultures using logistic equations. *Cytotechnology*, v. 64, p. 465–475, 2012.

GOUZÉ, J.L.; RAPAPORT, A., HADJ-SADOK, M.Z. Interval observers for uncertain biological systems. *J. Ecological Modelling*, v. 133, p. 45-56, 2000.

HIMMELBLAU, D.M. Process analysis by statistical methods. Wiley, New York, 1970.

HIMMELBLAU, D.M. Applied nonlinear programing, McGraw Hill, Inc., Nova York, 1972.

JAMILIS, M.; GARELLI, F.; MOZUMDER, M. S. I.; VOLCKE, E.;DE BATTISTA, H. Specific growth rate observer for the growing phase of a polyhydroxybutyrate production process. *Bioprocess and Biosystems Engineering*, v. 38, p. 557–567, 2015.

JOBE, A.M.; HERWIG, C.; SURZYN, M.; WALKER, B.; MARISON, I.; VON STOCKAR, U. Generally applicable fed-batch culture concept based on the detection of metabolic state by on-line balancing. *Biotechnol Bioeng,* v. 82, p. 627–639, 2003.

KELL,D.; BROWN, M.; DAVEY, H.M.; DUNN, W.B.; SPASIC, I.; OLIVER, S.G. Metabolic footprinting and systems biology: the medium is the message. *Nat. Rev. Microbiol*., v. 3, p. 557–565, 2005.

KIM, B.J.; FORBES, N.S. Flux analysis shows that hypoxia-inducible- factor-1-alpha minimally affects intracellular metabolism in tumor spheroids. *Biotechnol. Bioeng.,* v. 96, p. 1167–1182, 2007.

KREMLING,A.; GEISELMANN, J.; ROPERS, D.; DE JONG, H.Understanding carbon catabolite repression in Escherichia coli using quantitative models. *Trends Microbiol*., v. 23, p. 99–109, 2015.

KRICHEN, E.; RAPAPORT, A.; FOUILLAND, E. About frame estimation of growth functions and robust prediction in bioprocess modeling. hal-02095933v3, 2019.

LARENTIS, A.L., BENTES, A.M.P. Jr, RESENDE, N.S., SALIM, V.M.M., PINTO, J.C. Analysis of experimental errors in catalytic tests for production of synthesis gas. *Appl Catal A Gen* 242:365–379, 2003.

LE DUY, A.; ZAJIC, J.E. A geometrical approach for differentiation of an experimental function at a point: applied to growth and product formation. *Biotechnology and Bioengineering*, v.15, p.SOS-810, 1973.

LEMESLE, V. and GOUZ, J. Hybrid bounded error observers for uncertain bioreactor models. *Bioprocess and Biosystems Engineering,* v. 27, p. 311-8, 2005.

LUBENOVA, V.; ROCHA, I.; FERREIRA, E.C. Estimation of multiple biomass growth rates and biomass concentration in a class of bioprocesses. *Bioprocess Biosyst Eng*, v. 25, p. 395–406, 2003.

LYUBENOVA, V.; IGNATOVA, M.; NOVAK, M.; PATARINSKA, T. Reaction Rate Estimators of Fed-Batch Process for Poly-β-Hydroxybutyrate (PHB) Production by Mixed Culture. *Biotechnology & Biotechnological Equipment*, v. 21, p. 113-116, 2007.

MADISON, L.L. and HUISMAN, G.W. Metabolic engineering of poly(3-hydroxyalkanoates): from DNA to plastic. *Microbiol Mol Biol Rev*., v. 63, p. 21-53, 1999.

MARANGONI, C.; FURIGO JR., A.; ARAGAO, G.M.F. The influence of substrate source on the growth of Ralstonia eutropha, aiming at the production of polyhydroxyalkanoate. *Braz. J. Chem. Eng***.**, v. 18, p. 175-180, 2001.

MICKLEY,H.S.; SHERWOOD,T.K.; REED, C.E. Applied mathematics in chemical engineering. 2.ed. Nova York, Toronto, Londres, McGraw-Hill Book Company,1957.

MOSER, A. Kinetics of batch fermentation. In: Rehm, H.J. Reed, G. Volume Editor: Brauer, H. Biotechnology- A comprehensive treatise in 8 volumes. Weinheim, V.C.H. Verlagsgesellschaft, v. 2, p. 243-283, 1985.

MOSHOOD, T. D., NAWANIR, G., MAHMUD, F., MOHAMAD, F., AHMAD, M. H., & ABDULGHANI, A. Sustainability of biodegradable plastics: New problem or solution to solve the global plastic pollution?. *Current Research in Green and Sustainable Chemistry*, v. 100273, 2022. https://doi.org/10.1016/j.crgsc.2022.100273

MURPHY,T. and YOUNG,J. ETA: robust software for determination of cell specific rates from extracellular time courses. *Biotechnol. Bioeng*., v. 110, p. 1748–1758, 2013.

NEELEMAN, R. and VAN BOXTEL, A.J.B. Estimation of specific growth rate from cell density measurements. *Bioprocess Biosys. Eng.*, v. 24, p. 179–185, 2001.

NEIDHARDT, F.C.; INGRAHAM J.L.; SCHAECHTER, M. Physiology of the Bacterial Cell: A Molecular Approach, 20th ed, Sinauer Associates Sunderland, 1990.

OLIVEIRA, R.; FERREIRA, E.C.; OLIVEIRA, F.; FEYO DE AZEVEDO, S. A study on the convergence of observer-based kinetics estimators in stirred tank bioreactors. *J Process Control,* v. 6, p. 367–71, 1996.

ONER, M. D.; ERICKSON, L. E., YANG, S. S. Utilization of Spline Functions for Smoothing Fermentation Data and for Estimation of Specific Rates. *Biotechnology and Bioengineering*, v. 28, p. 902-918, 1986.

PAZ, P. M.; GAMBA, J. R.; LE ROUX, G. A. C. Desenvolvimento de técnica de alisamento de dados de processo biotecnológico. In: ANAIS DO 23° CONGRESSO BRASILEIRO DE ENGENHARIA QUÍMICA, 2021, Gramado. Anais eletrônicos. Campinas, Galoá, 2021. Available at: <https://proceedings.science/cobeq/cobeq-2021/trabalhos/desenvolvimento-de-tecnica-de-alisamento-de-dados-de-processo-biotecnologico?lang=pt-br> Acesso em: 02 jan. 2023.

PICÓ, J.; DE BATTISTA, H.; GARELLI, F. Smooth sliding-mode observers for specific growth rate and substrate from biomass measurement. *Journal of Process Control,* v. 19, p. 1314-23, 2009.

QUEK, L.E.; DIETMAIR, S.; KROMER, J.O.; NIELSEN, L.K. Metabolic flux analysis in mammalian cell culture. *Metab. Eng.*, v. 12, p. 161–171, 2010.

RAMSAY, J.A.; BERGER, E.; RAMSAY, B. A.; CHAVARIE, C. Recovery of poly-3-hydroxyalkanoic acid granules by a surfactant-hypochlorite treatment. *Biotechnology Techniques*, v. 4, p. 221-226, 1990.

RAPAPORT, A. and DOCHAIN, D. Interval observers for biochemical processes with uncertain kinetics and inputs. *Mathematical Biosciences*, v. 193, p. 235–253, 2005.

RAZAA, Z. A.; ABIDA, S.; BANATB, I. M. Polyhydroxyalkanoates: Characteristics, production, recent developments and applications. *International Biodeterioration & Biodegradation*, v. 126, p. 45–56, 2018.

REINECKE, F.; STEINBÜCHEL, A. *Ralstonia eutropha* Strain H16 as Model Organism for PHA Metabolism and for Biotechnological Production of Technically Interesting Biopolymers. *Microbial Physiology*, v. 16, n. 1–2, p. 91–108, 2009.

REINSCH, C. H. Smoothing by spline functions. *Nurnerische Marhernatik*, v. 10, p. 177, 1967.

RIIS, V.; MAI, W. Gas chromatography determination of poly-hydroxybutyric acid in microbial biomass-esther hydrochloric acid propanolisis. *J. Chromatogr***.**, v. 445, p. 285-289, 1988.

SALAZAR, M. C. and BOTERO, H. Specific Growth Rate Estimation in Fed-Batch Bioreactor Using Super Twisting Sliding Mode Observer. *Springer International Publishing*, CCIS 657, p. 403–411, 2016.

SAMBROOK, J.; FRITSCHI, E.F.; MANIATIS, T. Molecular cloning: a laboratory manual. *Cold Spring Harbor Laboratory Press*, New York, 1989.

SAN, K-Y.; STEPHANOPOULOS, G. Studies on on-line bioreactor identification. II. numerical and experimental results. *Biotechnol Bioeng*, v. 26, p. 1189-1197, 1984.

SAVITZKY A, GOLAY S, Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.,* v. 36, p. 1627-1639, 1964.

SCHLEGEL, H.; GOTTSCHALK, G.; VON BARTHA, R. Formation and utilization of polyhydroxybutyric acid by Knallges Bacteria (Hydrogenomonas).*Nature*, v. 191,p. 463–5, 1961.

SCHMIDELL,W.; LIMA, U.A.; AQUARONE, E; BORZANI, W. Biotecnologia industrial - volume 2: Engenharia Bioquímica. *Edgard Blücher*, 1st edition, 2001.

SILVA, L.F. and GOMEZ, J.G.C. Produção biotecnológica de poli-hidroxialcanoatos para a geração de polímeros biodegradáveis no Brasil. *Revisão: Quim. Nova*, v. 30, p. 1732-1743, 2007.

STEINBÜCHEL, A. and SCHLEGEL, H. G. Physiology and molecular genetics of poly(p-hydroxy-alkanoic acid) synthesis in Alcaligenes eutrophus. *Molecular Microbiology*, v. 5, p 535-542, 1991.

STEPHANOPOULOS, G.; ARISTIDOU, A.; NIELSEN, J.Metabolic Engineering: Principles and Methodologies. Academic Press, San Diego, 1998.

ZUPKE, C.;SINSKEY, A.J.;STEPHANOPOULOS, G. Intracellular flux analysis applied to the effect of dissolved oxygen on hybridomas. *Applied Microbiology  Biotechnology*, v. 44, p.27–36, 1995.

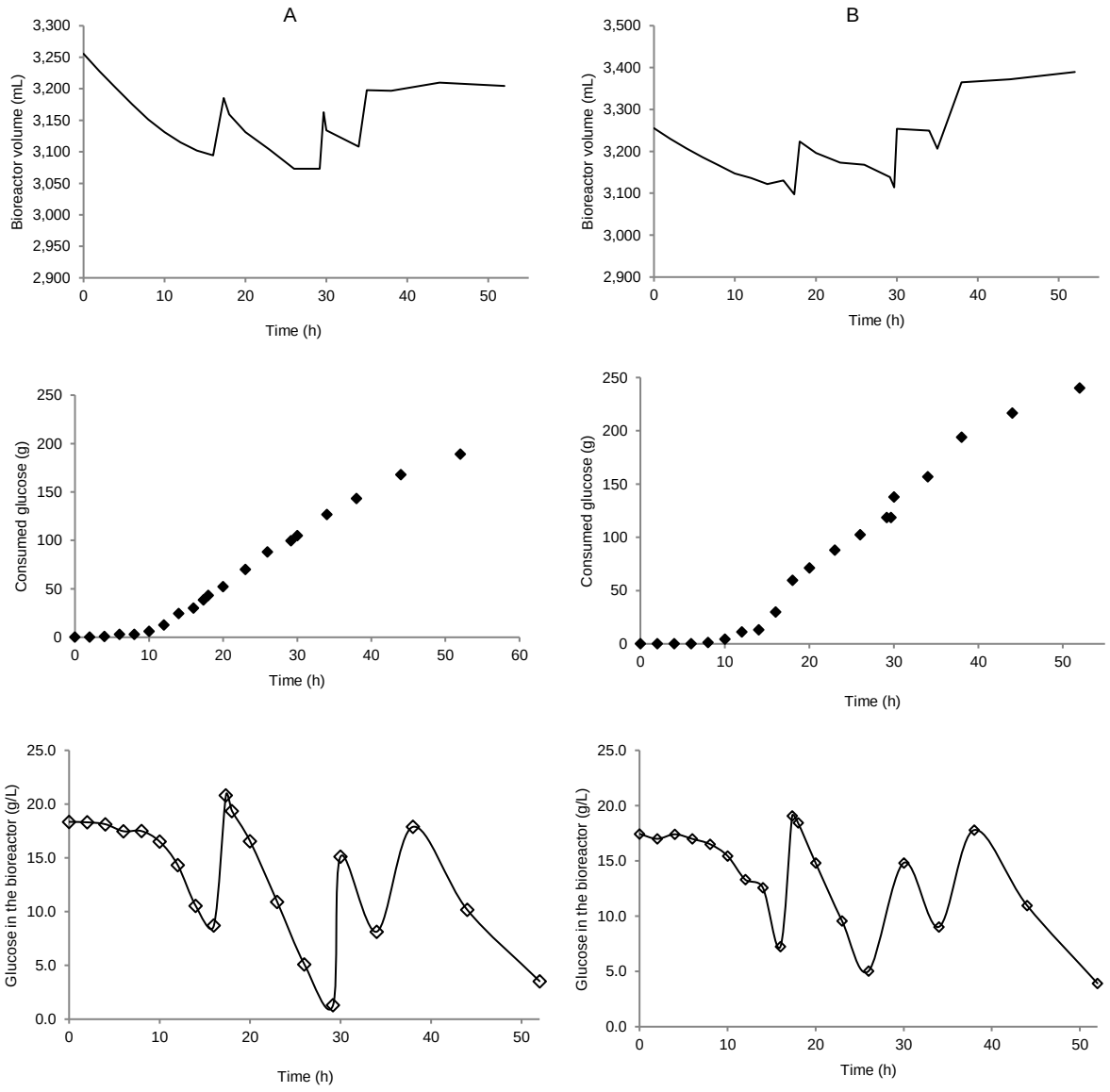
WEBB C., ATKINSON, B. The role of chemical engineering in biotechnology. *Chem Eng J* 50:9–16, 1992.

WOLD, S. Spline Functions in Data Analysis, *Technometrics*, v.16, p. 1-11, 1974.

# 7. APPENDIX

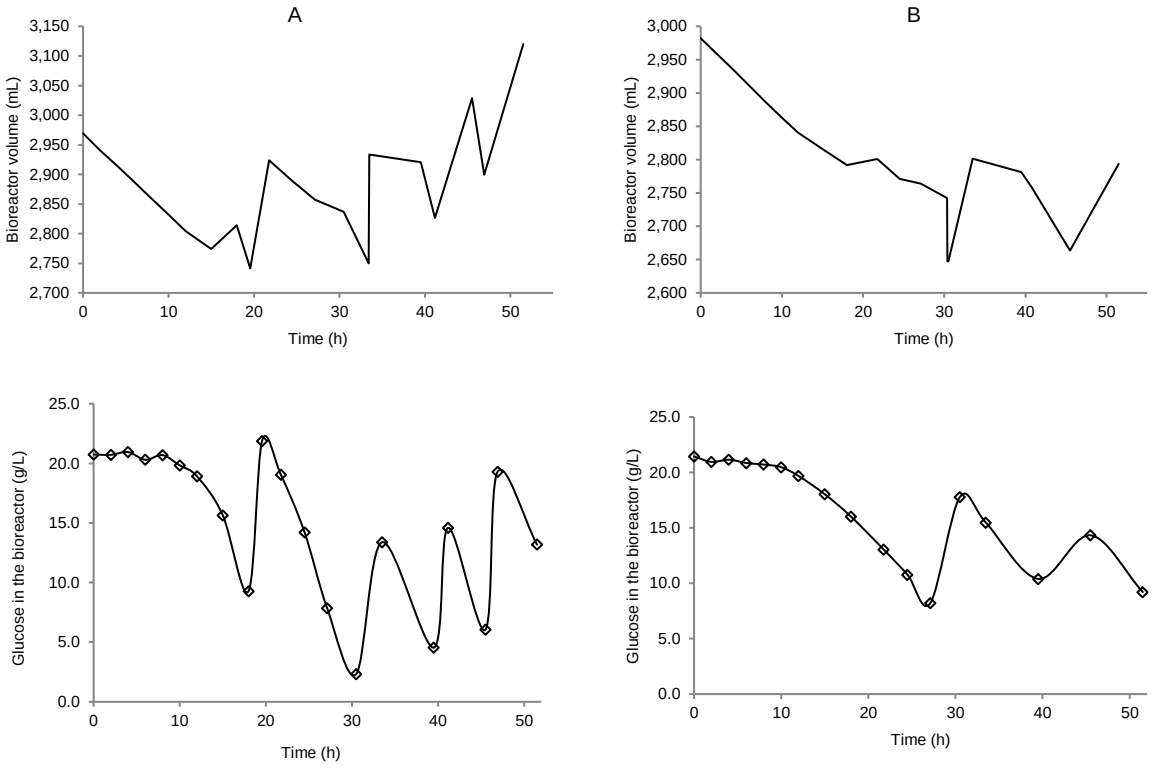
## 7.1. Additional graphs of bioreactor experiments

Figure 15. Additional graphs for the first two sets of experiments. R1 results are represented in column A, and R2 in B. From top to bottom: bioreactor volume, consumed glucose and glucose concentration in the bioreactor.




The peaks in the total reactor volume and also in the glucose graphs are the feeding pulses during the experiments. Glucose was added to the bioreactor after 20h, 30h and 40h approximately after the beginning of the experiments. This was necessary to keep glucose concentration in the system higher than 7g/L.

Figure 16. Additional graphs for the second two sets of experiments. R3 results are represented in column A, and R4 in B. From top to bottom: bioreactor volume and glucose concentration in the bioreactor.




Glucose was fed four times R3 and two time in R4.