Computers and Chemical Engineering 104 (2017) 151–163

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Iterative modeling and optimization of biomass production using experimental feedback

Martin F. Luna, Ernesto C. Martínez*

INGAR (CONICET-UTN), Avellaneda 3657, Santa Fe S3002 GJC, Argentina

ARTICLE INFO

Article history:
Received 19 December 2016
Received in revised form 18 April 2017
Accepted 20 April 2017
Available online 23 April 2017

Keywords:
Model-based optimization
Experimental design
Biomass production
Fed-batch reactor
Baker's yeast

ABSTRACT

Models of cultures of microorganisms are widely used for analysis, control and optimization of bioreactors in order to enhance productivity and performance. Typically, model-based optimization approaches may have acceptable convergence rates to a local optimum, but they are negatively affected by modeling errors when extrapolating to unknown operating conditions. In this work, a model-based optimization methodology that uses experimental feedback is applied to a fed-batch bioreactor. Experimental feedback is used to solve the extrapolation problem. After the model has been (re)parameterized, an optimized experiment is designed to maximize the performance of the bioprocess. Data gathered in this experiment is used to correct the model, and the cycle continues until no further improvement is found. The method is tested in the production of baker's yeast biomass. Results obtained demonstrate the capability of the proposed approach to find an improved feeding profile that leads to better performance with minimum experimental effort.

1. Introduction

Drastically improving the productivity of a bioreactor has been a major concern in the biotechnology industry since its very beginning. Penicillin production optimization is a well-known example (Shuler and Kargi, 2002): despite penicillin was discovered in 1928, it was not until the process was optimized two decades later which made the drug commercially available, thus changing the life of millions of people. Another representative example, even older, is the production of baker's yeast. In the beginning of the twentieth century, yeast producers started to notice that under low carbohydrates concentration (and with sufficient aeration), the biomass yield increases. This led to the development of the *Zulauf-Verfahren* or fed-batch process (Jorgensen, 1948; Rose and Harrison, 2012). With today's recombinant DNA techniques using *Pichia Pastoris* (a species of methylotrophic yeast) for protein production, biomass productivity is of paramount importance due its direct correlation with protein expression.

Nowadays, the biotechnological industry and the academic sector have created an amazing amount of knowledge, merging topics of different areas, from biochemistry to chemical engineering. Product and process efficiency are mandatory to survive in a highly competitive and innovative industry (Pisano, 1997). Nevertheless, there is still plenty of room for improvement since process system engineering (PSE) tools are yet not fully embraced in the biotechnological sector, where top-notch techniques coexist with outdated industrial practices (Gernaey, 2015). Some initiatives like the FDA's Quality by Design (QbD) (FDA, 2006) aim to address this issue, in order to increase the industry output, in a world that demands more and more food and medicine (Tilman et al., 2002; OECD Indicators, 2015). According to QbD, the use of advanced tools such as mathematical modeling is very useful to develop efficient, safe and clean processes. However, some difficulties prevent this approach to be widely used. First, it requires a body of specific knowledge about the biochemical process in order to obtain a model. While there is an enormous bibliography related to bioprocess models and how to develop them, important factors in industrial practice such as unexpected day-to-day contingencies or short development times drive toward simpler approaches, such as trial and error methods (Royle et al., 2013). Besides that, first-principles mathematical models may accurately predict the process response only under conditions close to those used to fit them, but usually fail when extrapolating away to more distant conditions. This may lead the bioreactor to be operated in suboptimal conditions or, in even worse, to unsafe or unprofitable operation (Mandur and Budman, 2015). This is especially true for novel bioprocesses, due to their complex dynamic

* Corresponding author.
E-mail addresses: ecmarti@santafe-conicet.gov.ar, ecmarti@gmail.com (E.C. Martínez).

http://dx.doi.org/10.1016/j.compchemeng.2017.04.020
0098-1354/

Nomenclature

Process variables

F	Inflow rate, [l/h]
Glc	Glucose concentration, [g/l]
Glc_{in}	Glucose concentration in the feed, [g/l]
r_s	Glucose supply rate, [h^{-1}]
r_d	Glucose demand rate, [h^{-1}]
t	Time [h]
t_f	Final time or duration of the experiment, [h]
u	Process parameter vector (or policy vector)
$\tilde{u}$	Process parameter distribution vector
V	Liquid volume, [l]
X	Biomass concentration, [g/l]
Y	Glucose-to-biomass yield, [g/g]
y	Vector of model predictions
θ	Model parameter vector
$\tilde{\theta}$	Model parameter distribution vector
μ	Growth rate, [g/l h]

Optimization problem variables

a	Relative price of glucose
E_r	Error function
J	Performance index
J_I	Objective function for the information gain problem
Q	Sensitivity matrix
S_{ij}	Sensitivity index of the *ith* element at the *jth* time
ts	Sampling schedule vector
V_{ij}	Conditional variance of the *ith* element at the *jth* time
V_j	Total variance at the *jth* time
$\$_{Glc}$	Price of glucose, [$/g]
$\$_x$	Price of biomass, [$/g]

Sub-indices

E	Ethanol oxidation mode
end	Final element of the vector
f	Fermentative mode
max	Maximum
min	Minimum
r	Respiratory mode
o	Initial element of the vector

Hyper parameters

m	Experiments per iteration counter
m_{MAX}	Maximum number of experiments per iteration
n	Iteration counter
sf	Shrinking factor
ε	Stopping criterion

behavior and the uncertainty regarding the best handles to achieve optimal operation (Kiparissides et al., 2011).

In order to address above drawbacks due to imperfect models, some approaches have been proposed. The *modeling for optimization* approach (Bonvin et al., 2016) combines mathematical modeling with experimental feedback with the main objective of improving the process performance. This goal is sensibly different from the *modeling for description* approach, where detailed mathematical models are created to describe data gathered in the experiments, without any special concern for process optimization. When the modeling goal is iterative optimization, models do not need to be excessively detailed (which relieve the burden of parametric precision in the modeling stage), but they have to capture the tendency of the process, i.e. how the process reacts to changes in its controlled inputs. The use of experimental feedback allows iteratively updating model parameters based on data gathered in designed experiments where information content is mainly related to predicting optimal operating conditions.

A benchmark problem in the biotechnology industry is the production of biomass. Microorganisms are used as a catalyst in bioreactors in order to obtain a wide range of high-value products (food and beverage, complex proteins, enzymes, etc) which are directly correlated to biomass production. While biomass usually grows in the bioreactor, an initial seed is needed to start the process. Thus, biotechnological industries have replicating or "seed" reactors which operate in optimal conditions to ensure the initial amount of biomass (which may be different to conditions needed to produce the final product at the industrial scale). This is the case of baker's yeast (*Saccharomyces cerevisiae*). It is one of the most used microorganisms since it can be genetically engineered to produce the desired metabolites (Randez-Gil et al., 1999; Nielsen, 2013). It is worth noting that Baker's yeast uptakes nutrients through different metabolic pathways depending on the operating conditions in the bioreactor. It is a facultative microorganism, which means that it could grow under aerobic (respiration) or anaerobic (fermentation) conditions (Van Dijken and Scheffers, 1986; Rodrigues et al., 2006). However, in the presence of a high concentration of carbohydrates, the anaerobic pathway prevails even with sufficient aeration. This operating mode is not optimal for biomass production, since the yield of this metabolic pathway is lower than the one for the aerobic pathway. Thus, it is desirable that the reactor operates maintaining the carbohydrate concentration low enough to favor respiration, but with a high carbohydrate feed to favor biomass production (measured as mass per unit of time). The fed-batch operation favors this, but the carbohydrates feeding profiles must be optimized to achieve high productivity conditions. Since each yeast strain presents its own kinetic behavior, the optimal profile will vary among strains and cannot be duplicated directly from similar processes. Accordingly, optimization methods must be applied in the development stage to pinpoint optimal conditions for biomass production and protein expression.

In this work, a modeling for optimization methodology is applied to a bench scale bioreactor used to produce baker's yeast biomass from glucose. In Section 2, the problem is presented and the experimental set up for the bench scale bioreactor is described. In Section 3, a mathematical model is proposed and analyzed. In Section 4, the model-based optimization approach used to find the optimal experimental conditions is briefly explained. The results presented in Section 5 demonstrate how model-based optimization methods combined with experimental feedback are very useful to increasingly improve biomass production using a simple model. Section 6 ends the paper with conclusions and ideas for further research.

2. Materials and methods

2.1. Experimental setup and process description

The experiments were performed in a BioFlo 110 Benchtop Fermenter® (New Brunswick Scientific). The reactor was charged initially with a nutrient medium and was then inoculated with baker's yeast at the beginning of the experiment. After an initial lag phase operating in batch mode, the fed-batch mode was started, where a solution of glucose was used as the carbon source for the growth of the microorganism. After the fed-batch mode, the reactor is shortly operated in a second batch mode, in order to consume any glucose left in solution. Samples were taken several times along the experiments (in order to obtain the model parameters off-line). The performance of the process was measured using the following

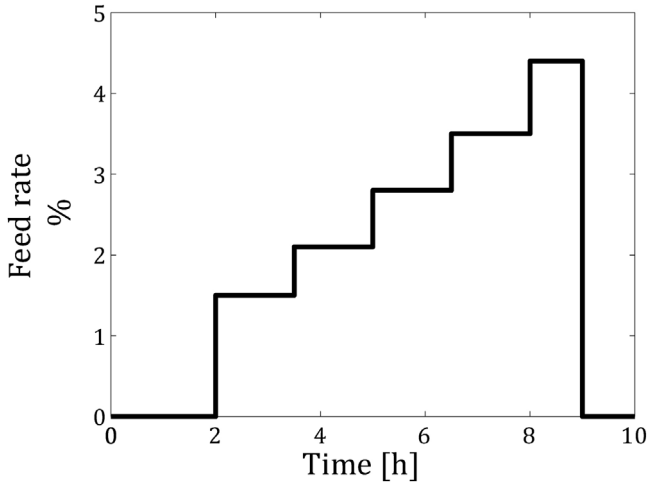

Fig. 1. Shape of the feeding profile to be optimized in the experiments.

objective function, which represents the economic benefit of the run, according to Eq. (1).

$$J = \frac{\Delta X - aGlc_{add}}{t_f} \tag{1}$$

Here, ΔX is the biomass produced in the experiment, Glc_{add} is the total glucose added to the reactor and a is a parameter which is set here to 0.3. The meaning of Eq. (1) will be discussed later in Section 3. The feeding profile will be described according to a vector parameterization approach (Goh and Teo 1988) for incrementally stepping up the inflow rate. A sequence of steps was chosen for simplicity but other functionalities for the feeding profile may be used as well. The objective is to find, after a series of experiments, the optimal parameterization of the feeding profile that maximizes the objective function J presented in Eq. (1). Other controlled inputs such as temperature, pH and the duration of the experiment (or each one of the phases) may be optimized as well, but they were not considered in this work.

2.2. Culture medium and experimental conditions

The initial liquid volume in the reactor was 1.25 l. The aqueous media used was the traditional for yeast growth: 10 g/l of yeast extract, 20 g/l of peptone and an initial charge of glucose of 1.2 g/l, with an addition of 0.1 ml/l of an antifoaming agent. The inoculum for each experimental run consisted of 6 g/l of baker's yeast, prepared from a rehydrated commercial product. The feeding stream has a concentration of 100 g/l of glucose. Aeration was supplied using an air pump with a flow of 2500 cm^3/min, while agitation was performed using a turbine impeller. Temperature was controlled using a heating jacket and a cooling coil, in order to maintain it at the set point value of 30 °C.

2.3. Feeding profile

The feeding profile F is parameterized as a succession of 5 positive step changes &tpcheck;F_i at given times. The fed-batch mode is preceded and followed by a batch mode operating period. The profile for the fed-batch operating phase is described by Eq. (2):

$$F_i = F_{i-1} + \Delta F_i \tag{2}$$

where the sub-index i stands for the *ith* step and ΔF represents the increment to the inflow rate. The duration of each step is fixed *a priori*, and the process variables are the entries in the vector u. A graphical description of the feeding profile is shown in Fig. 1, and its parameterization related to the vector u is detailed in Table 1, where the second column indicates the time at which each stepping up change to the inflow is made. The maximum value for the inflow rate is given by the feeding pump capacity, which in the bench scale bioreactor corresponds to 1.3 l/h. The feeding profile is expressed as the percentage of this maximum value. The concentration of glucose in the inlet flow is set to 100 g/l.

Table 1
Description of the feeding profile parameterization.

Step	Time [h]	F [l/h]	Calculated as
0	0	0	0
1	2	F_1	$F_1 = u_1$
2	3.5	F_2	$F_2 = F_1 + u_2$
3	5	F_3	$F_3 = F_2 + u_3$
4	6.5	F_4	$F_4 = F_3 + u_4$
5	8	F_5	$F_5 = F_4 + u_5$
6	9	0	0
–	10	0	0

2.4. Measurement methods

Samples were taken to assess the concentration of biomass and glucose in the reactor. The sampling schedule was defined as part of the experimental design, and will be discussed in Section 3. Samples were taken from the reactor using a sampling probe, and they were analyzed in a Spectro 20D Plus Spectrophotometer® (Labomed Inc), according to the following techniques.

Biomass was measured by turbidimetry, at a wave length of 600 nm. The samples were diluted and their optical density was determined, in order to calculate the biomass concentration using a calibration curve method.

Glucose was measured using the glucose-oxidase method (Bergmeyer, 1974). The samples were centrifuged using a Combi – 514R Multi Purpose Centrifugue® (Hanil Science Inustrial Co., Ltd), working at 10000 rpm for 5 min, and the supernatant was diluted and analyzed. The method is based on the reaction between glucose and an enzymatic kit (glucose oxidase, 4-aminophenazone, phenol and peroxidase) in order to give a colored solution that is analyzed using a spectrophotometer at 505 nm, which allows estimating the glucose concentration using a calibration curve.

Besides these determinations, in order to control or to register the other operating conditions set in the bioreactor, the following variables were measured online: pH was measured using a pH probe; temperature was measured using a RTD sensor; oxygen concentration was measured using a DO probe and agitation speed was measured using a tachometer.

3. Modeling and analysis

3.1. Baker's yeast

Production of baker's yeast biomass from glucose is a traditional problem in the biotechnological industry that has been studied by several authors (Vemuri and Palanki, 2006; Valentinotti et al., 2003; Kasperski and Miśkiewicz, 2008; Richelle and Bogaerts, 2014). In the biomass production step, the microorganism grows in suspension in an aqueous medium which constitutes the abiotic cell environment. After a batch is ended, the bioreactor content is transferred to the separation step, where biomass can be recovered from the suspension to be later used in the primary process on-site, or to be sold for industrial or personal consumption.

Baker's yeast is a facultative microorganism, i.e. it may live and grow in both the presence and absence of oxygen. This feature comprises the main difficulty regarding the process: depending on the medium composition and conditions, the microorganism

may present different metabolic modes, which uptake glucose with different yields. The macroscopic balances of the main modes are usually described as (Sonnleitner and Käppeli, 1986):

$$Glc + s_r O_2 \rightarrow Y_r X \tag{3}$$

$$Glc \rightarrow Y_f X + s_f EtOH \tag{4}$$

$$EtOH + s_E O_2 \rightarrow Y_E X \tag{5}$$

where *Glc* stands for glucose, O_2 stands for oxygen, *X* stands for biomass, *EtOH* stands for ethanol, *s* stand for the stoichiometric factors and each *Y* stands for the glucose-to-biomass yield of the corresponding mode. Eq. (3) describes the respiratory mode; Eq. (4) describes the fermentative mode whereas Eq. (5) describes the ethanol oxidation mode. The respiratory mode has a glucose-to-biomass yield three to five times greater than the one for the fermentative mode, but the yeast would only prefer this mode if oxygen concentration is high and if the glucose concentration is low enough. If glucose concentration is high, the fermentative mode will prevail even with a high concentration of oxygen, a behavior known as the Crabtree effect (De Deken, 1966). This is an important issue because the high glucose throughput necessary to achieve a high volumetric production (biomass produced with regards to reactor volume or plant capacity) may give rise to a fermentative behavior which lowers the glucose-to-biomass yield. To overcome this difficulty, the bioreactor is operated in fed-batch mode, keeping the glucose concentration in the culture medium in a low level while having a high glucose throughput. To this aim, oxygen concentration should be kept high and complementary nutrients must be added to the medium to achieve a healthy grow of the microorganism. It is worth noting that in big vessels with dense medium and high biomass concentration, mass transport phenomena may limit oxygen diffusion from the air inlet to the cells. The ethanol uptake mode is most likely to occur when the ethanol concentration is high and the glucose concentration is low.

3.2. Mathematical model

The proposed mathematical model should be capable of properly describing the effect of glucose excesses on metabolic modes, and at the same time, keeping the model size small enough to be parameterized with a small dataset having information only from biomass and glucose concentrations. The bench scale reactor is relatively small (2 l), hence it is plausible to assume that there no exist oxygen diffusion limitations. The ethanol oxidation mode it is not accounted for in the model since it is expected that a low quantity of ethanol will be produced due to a good process performance in the considered operating region. With all these considerations in mind, the proposed model is described by Eq. (6)–(12).

$$\frac{dX}{dt} = \mu X - \frac{F}{V} X \tag{6}$$

$$\frac{dGlc}{dt} = -\frac{\mu}{Y} X + \frac{F}{V}(Glc_{in} - Glc) \tag{7}$$

$$\frac{dV}{dt} = F \tag{8}$$

$$\mu_f = \theta_1 \frac{Glc}{\theta_2 + Glc} \tag{9}$$

$$\mu_r = \theta_3 \frac{Glc}{\theta_4 + Glc} \tag{10}$$

$$\mu = \mu_r + \frac{(\mu_f - \mu_r)}{1 + exp\left(\frac{\theta_5 - Glc}{\sigma}\right)} \tag{11}$$

Table 2
Model parameters used in the analysis in Section 2.

Parameter	Value	Units
θ_1	0.1391	h^{-1}
θ_2	0.0017	gr l^{-1}
θ_3	0.4690	h^{-1}
θ_4	0.0170	gr l^{-1}
θ_5	0.0375	gr l^{-1}
θ_6	0.4820	gr gr^{-1}
θ_7	0.1255	gr gr^{-1}

$$Y = \theta_7 + \frac{(\theta_6 - \theta_7)}{1 + exp\left(\frac{\theta_5 - Glc}{\sigma}\right)} \tag{12}$$

Here, σ is a constant whose value is set to 5.10^{-4}. This simple model is easy to parameterize and it is inspired by experimental data presented in the literature (Kaspar von Meyenburg, 1969; Wenger, 1994; Van Hoek et al., 1998). More complex models can be found elsewhere (Sonnleitner and Käppeli, 1986; Lei et al., 2001; Serio et al., 2001; Richelle et al., 2014). In the proposed model, two growth rates (one for each mode) are presented and described by Monod-type kinetics. The resulting growth rate is a combination of both rates, according to Eq. (11). Parameter θ_5 represents the threshold concentration of glucose at which the yeast metabolism changes from one mode to another. The same principle is applied to the glucose-to-biomass yield in Eq. (12). The behavior of the microorganism (especially growth rates and yields) varies from one strain to another, which gives rise to different values for the model parameters. Thus, using either model parameters or data directly from the literature is typically a significant cause for prediction errors.

Based on the proposed model, the basic aspects of the baker's yeast bioprocess can be analyzed. Using the nominal parameterization presented in Table 2, plots of μ and *Y* vs *Glc* concentration are shown in Fig. 2. As can be seen in Fig. 2a, when *Glc* tends to zero, the denominator in Eq. (11) tends to a high number, thus the second summand tend to zero and the growth rate stabilizes to the value for the respiration mode. When *Glc* tends to a high number, the denominator in Eq. (11) tends to one, thus the second summand partially cancels the first summand, and the growth rate stabilizes to the value of the fermentative mode. The threshold value for glucose, represented by the parameter θ_5, is located at the middle point of the curve. A similar type of reasoning can be made for the glucose-to-biomass yield described by Eq. (12) and presented in Fig. 2b. As can be seen, an excess of glucose implies a lower glucose-to-biomass yield.

For a typical fed-batch bioreactor it is expected that, if the substrate concentration is held constant, the amount of biomass increases exponentially due to cellular growth (some other requirements need to be met, i.e. no metabolite or nutrient inhibition, no oxygen diffusion limitation, no contact-dependent growth inhibition due to high cellular density, etc). Such a growth pattern requires that the substrate feed rate increases exponentially as well. However, for the case of yeast, if the glucose concentration increases beyond a threshold or critical value for the respiratory mode, the glucose-to-biomass yield will decrease and the performance index will be lower (and even growth inhibition may occur if too much glucose is added and a high ethanol concentration is present). It has been proven (Srinivasan et al., 2001) that the optimal feeding profile must maintain the glucose concentration near such a threshold value, in order to maximize the amount of biomass obtained with a high yield. To illustrate this, the parameters from

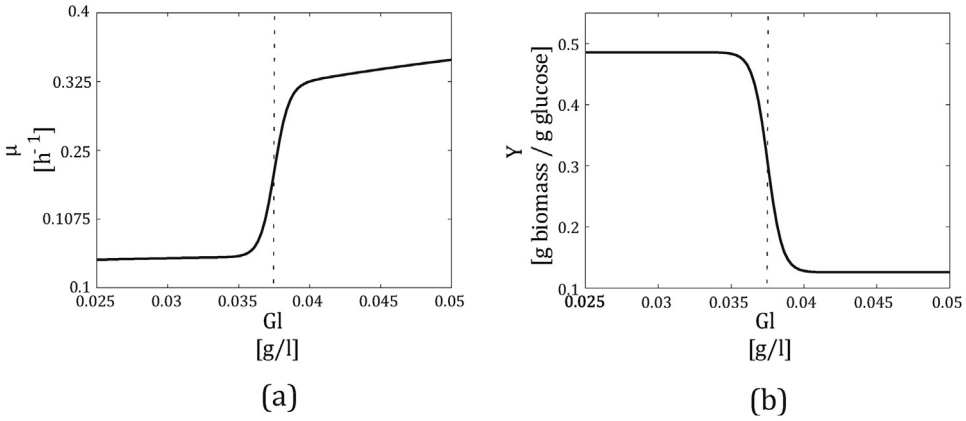

Fig. 2. Dependence of growth rate (a) and glucose-to-biomass yield (b) with respect to glucose concentration.

Table 2 are used to simulate a fed-batch process where the shape of the feeding profile is defined as follows:

$$F = \begin{cases} F_{max} \text{ if } t < u_1 \\ F_C \text{ otherwise} \end{cases} \tag{13}$$

where

$$F_C = u_2 \exp(u_3 (t-2)) \tag{14}$$

u_i, are the profile parameters to be optimized and $F_{\max}$ is the maximum value for the feeding rate that is limited by the pump capacity, i.e. $F = 100\%$. The feeding rate is kept at its maximum capacity until the threshold concentration is achieved, and from that point on, exponential feeding is used. For the two batch phases discussed earlier (the initial two hours and the final hour) there is no substrate feeding, i.e. $F = 0\%$. The performance index and the rest of the experimental conditions are the ones presented in Section 2. The resulting feeding profile is depicted in Fig. 3a, and the corresponding glucose concentration along a simulated run is shown in Fig. 3b (where the threshold concentration has been highlighted using a dashed line).

Based on the analysis above, in this work the shape of the feeding profile is defined by a sequence of step-up changes. Using a feeding profile discretized as a sequence of steps is appealing for two reasons. Firstly, for simplicity, since it is easier to implement. Secondly, it is common in industrial practice resorting to simpler discretizations of feeding profiles to fed-batch units. Even though a sequence of step-up changes does not necessarily imply an exponential increase in the feeding rate, the feeding profile is biased to avoid less efficient patterns. An alternative approach is to optimize the profile as presented in Eq. (13)–(14), and then approximate the solution by a sequence of steps. Again, there are two main reasons for not doing so. First of all, the parameters in Eq. (14), even though they are linked to physical quantities, are less insightful compared to simple step increases. On the other hand, and more importantly, by not giving the feeding profile a fixed shape, its optimization has more degrees of freedom, which may be advantageous especially during re-optimization steps.

3.3. Objective function

Before ending this section, it is important to analyze the objective function to be optimized in the process. The objective function or performance index is the numerical quantity that measures the utility of each experimental outcome. For the case of biomass production, several approaches can be taken, and depending on which is chosen, the operating conditions that optimize it may differ. The more "obvious" performance indices are the ones presented in Eqs. (15) and (16), that represent the total biomass production and the biomass produced per unit of glucose and time (i.e., the inverse of the overall glucose to biomass yield divided by the time of the experiment), respectively:

$$J = \frac{X_{(tf)}V_{(tf)} - X_{(t0)}V_{(t0)}}{tf} \tag{15}$$

$$J = \frac{X_{(tf)}V_{(tf)} - X_{(t0)}V_{(t0)}}{Glc_{add}tf} \tag{16}$$

In the case of the performance index presented in Eq. (15), the amount of glucose used during the experiment is not taken into account. Thus, there is no incentive in using it economically, and favors operating the reactor with high glucose concentration, which leads to a low glucose-to-biomass yield. This solution would be of interest for biomass production at industrial scale only for the case of negligible glucose prize, which is not the case for baker's yeast (both yeast, and the sugars that are used in its production, are both standard commodities). Eq. (16) takes glucose consumption into account, but since there is a biochemical relationship between the glucose consumed and the biomass produced (the glucose-to-biomass yield), a small amount of glucose added would lead to a small amount of biomass production (even when the yield is high). Thus, the objective function in Eq. (16) would favor operating conditions with high glucose-to-biomass yield, without necessarily considering the total amount of biomass produced. This may lead, for a given volumetric capacity, to low productivity levels of biomass per liter of the bioreactor. This would only be of industrial interest if the cost of fixed capital (reactor, separation train, etc.) is negligible. Again, this is not true for case of yeast production at industrial scale.

In order to take into account more realistic scenarios, the performance index must have the form:

$$J = \frac{Benefit - Cost}{tf} \tag{17}$$

where cost and benefit of a production run should be measured in monetary units ($). A similar approach has been proposed elsewhere (Jia et al., 2007). The benefit would be the market value of the biomass to be sold (or the amount of biomass that would replace a supplier's, in the case of captive use). The overall cost may include raw material purchase cost (sugar, medium, inoculum), energy cost (aeration, agitation), separation cost per gram of biomass produced, labor cost and the annualized cost of equipment. Since some of these costs are not dependent on the process variables, they may be not considered. In the case of this work, and

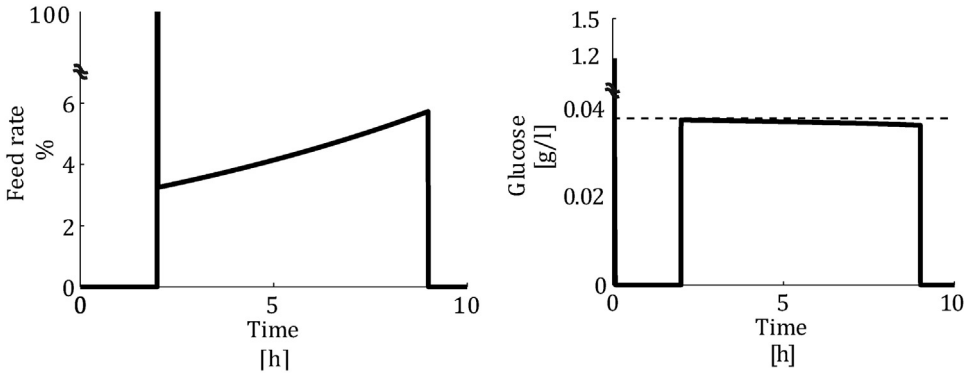

Fig. 3. (a) Optimal exponential feeding profile and (b) its corresponding glucose concentration for a simulated run.

for the sake of simplicity, the costs and benefits are considered as follows:

$$J = \frac{\$_x \left(X_{(tf)}V_{(tf)} - X_{(to)}V_{(to)}\right) - \$_{Glc}Glc_{add}}{tf} \tag{18}$$

Normalizing by the price of yeast and defining:

$$a = \frac{\$_{Glc}}{\$_x} \tag{19}$$

$$\Delta X = X_{(tf)}V_{(tf)} - X_{(to)}V_{(to)} \tag{20}$$

$$Glc_{add} = Glc_{(to)}V_{(to)} + \int_{to}^{tf} FGlc_{in}dt \tag{21}$$

then, the performance index is the one presented in Eq. (1) in Section 2.1, and according to what is exposed in this section, it takes into account the use of glucose from a cost/benefit point of view. Since the overall glucose to biomass yield can be defined as the biomass produced at the end of the batch by the glucose added along the whole run:

$$Y_{overall} = \frac{\Delta X}{Glc_{add}} \tag{22}$$

Then, it can be introduced in Eq. (1) in the following way:

$$J = \frac{Glc_{add}\left(Y_{overall} - a\right)}{tf} \tag{23}$$

According to Eq. (23), the optimal feeding profile is a tradeoff between high glucose addition and low glucose concentration in the bioreactor (which is necessary to achieve a high overall glucose-to-biomass yield). The overall yield should be above the threshold value a in order to generate economic benefits.

4. The modeling for optimization approach

4.1. Iterative optimization

When a mathematical model is developed, it has to be done with an objective in mind (Bonvin et al., 2016). In modeling for optimization, the goal of modeling is the optimization of the process. The main features of the modeling for optimization approach are: i) iterative update of simple models using data gathered in experiments designed so as to improve the performance of the process, ii) tight integration of the modeling and optimization steps with experimental feedback in an iterative scheme. The model of the process is thus used as a guideline for defining a direction or "tendency" for performance improvement and its structure is simple enough to be parameterized with scarce data. The use of experimental feedback allows updating model parameters as operating conditions are changed seeking to improve process performance.

The model is thus used as a tool to improve the performance of the process and its predictions about outcomes unrelated to the objective function, while useful, are not critical for optimization. Moreover, the model may have errors, but they are acceptable as long as its prediction leads to an improvement of the performance index. Model predictions are iteratively updated using evaluative feedback from designed experiments which generates data that are informative for performance improvement. As a result, the model will predict better in the vicinity of the new operating conditions. Using experimental data and first-principles knowledge about the process, the mathematical model is developed and parameterized. Each time the model is re(parameterized), it is used to solve the optimization problem and a new experiment is designed. The solution to the problem is tested experimentally and data gathered can be used to update the model and continue with the iterative scheme (Luna and Martínez, 2014).

The modeling for optimization approach has given rise to a number of different methods for process optimization in biotechnology processes. Some methods resort to models based on first principles, while others mostly rely on data-driven models, but all of them integrate experimental feedback. A thorough account and discussion of the different proposals for iterative learning control and optimization is beyond the scope of this paper. The reader is referred to some works from the existing literature related the optimization and control of bioreactors (Srinivasan et al., 2003; Georgakis, 2013; Bonné et al., 2014; Mandur and Budman, 2015).

Since models are developed using scarce data and simple structures, their predictions usually present high levels of uncertainty and deviations from the observed real behavior. One way to take this into account and use this information to enhance the method efficiency is the implementation of probabilistic tendency models (PTM) (Martínez et al., 2013). PTM use experimental data to obtain distributions of model parameters (instead of unique values), and use these distributions to predict the probabilities of alternative process responses.

The methodology requires the solution of two optimization problems: the model parameterization problem and the experimental design problem. The design of the experiment actually has two objectives and involve two separated optimization subproblems: i) to find the operating condition that maximizes the performance index, and ii) to maximize the information content in data obtained and which will later be used in the model (re)parameterization step. These two objectives in the design of experiment problem may conflict with each other, since operating conditions that improve the performance index may be very different to conditions that are most informative about the pro-

cess optimum, and vice versa. This is known in the literature as the conflict between *exploration* and *exploitation*. In experimental design, exploration is about choosing operating conditions with the objective of maximizing the information content in data gathered (here information content is related to the quality of model predictions regarding the optimal operating conditions). Exploitation, on the other hand, aims to maximize the process performance in the next experiment, without any concern for the information content of data sampled. A tradeoff between these two objectives can be achieved in different ways. It worth noting that some variables may affect the information content while not affecting the performance index (for example the ones related to sampling) while the opposite is usually not true (changes in process variables such as the feeding profile changes the information content of data gathered). In the present work, the experimental design problem is separated in two sub-problems. Firstly, model-based optimization of the performance index is carried out so that the levels for process variables for next experiment are set. Secondly, the design variables that affect only information content are chosen in a second sub-problem related to optimal sampling. In doing so, exploration is subordinated to exploitation. This is consistent with the philosophy of modeling for optimization where the ultimate goal of the method is to improve the performance of the process.

4.2. Model parameterization

Model development in the present work is limited to the (re)parameterization of the model presented in Eqs. (6)–(12). The optimization problem is posed as the typical least squared error problem. The objective function used is:

$$E_{r(\theta,u,y_{exp})} = \sum_i w_i \left(\frac{y_{i(\theta,u)} - y_{exp_i}}{y_{exp_i}} \right)^2 \tag{24}$$

where the sub-index *exp* stands for experimental data, *w* stand for the weighting factors and the summation is performed over each data point *i*. The process variables u are the ones used to obtain the experimental data. The optimization problem is presented as follow:

$$min_\theta E_{r(\theta)} \tag{25}$$

$$y = f\left(\theta, u\right) \tag{26}$$

$$\theta_{min} \leq \theta \leq \theta_{max} \tag{27}$$

where f represents the mathematical model.

In order to define the PTM, the probability distribution functions (*pdfs*) of model parameters must be obtain from available data, in order to take into account the uncertainty about the model predictions. There are different methods available for obtaining parameter distributions for nonlinear models. A common method is to obtain the mean and the variance of the parameters by assuming they follow normal distributions. To obtain them, a linearization of the model with respect to its parameters should be made. For highly nonlinear models, as it is the case of dynamic model of bioreactors, such linearization gives rise to confidence intervals that are usually underestimated. Also, nonlinear models may present parameter distribution functions that are non-normal and skewed (Joshi et al., 2006). This may lead to poor model predictions and suboptimal experimental designs. For this kind of models, the bootstrapping method is preferred, since no assumption is made regarding the type of distribution (Efron and Tibshirani, 1993). Instead, a histogram-based technique is used. Histograms are obtained using data sets that are artificially generated using re-sampling with replacement, and the parameterization problem is solved for each new data set which provides a parameter vector θ. The resulting vectors are compiled in a vector of parameter distributions $\tilde{\theta}$, in which each entry corresponds to a probability distribution function (*pdf*) of the respective parameter.

4.3. Design of the experiment

The design of the experiment is divided in the two aforementioned sub-problems. The performance maximization sub-problem is presented in Eq. (28)–(31).

$$max_u J = \frac{\Delta X - aGlc_{add}}{t_f} \tag{28}$$

$$y = f\left(\theta, u\right) \tag{29}$$

$$V_{(tf)} \leq 1.75l \tag{30}$$

$$u_{min} \leq u \leq u_{max} \tag{31}$$

Since the PTM predicts distributions of the process responses, the following procedure is used to obtain a distribution of optimal operating conditions $\tilde{u}$. A sample is taken from each distribution in $\tilde{\theta}$ and the resulting parameterization is used to solve Eq. (28)–(31). This procedure is repeated and the resulting u vectors are compiled in a vector $\tilde{u}$ whose entries are the distributions for process variables characterizing optimal operation. Using this vector, the operating conditions to be used in the next experiment can be defined in several ways. Here, the mean value of each distribution is chosen.

As was stated before, there is a tradeoff between process improvement and information generation for performance optimization. Here, and following the philosophy of modeling for optimization that states that the optimization of the performance index is more important than the other aspects of modeling, exploitation is deemed the dominant objective for iteratively modeling and optimization. Thus, the optimization problem from Eq. (28)–(31) is solved without regards to the information content. After the process variables are defined, the information-related variables, i.e. the variables that only affect the information content of the experiment can be optimized. In our case, these variables are the sampling times that are collected in the sampling schedule vector *ts*. The information of the experiment is expressed using the sensitivity matrix *Q*:

$$Q = \begin{pmatrix} S_{11} & \dots & S_{1j} \\ \vdots & \ddots & \vdots \\ S_{i1} & \dots & S_{ij} \end{pmatrix} \tag{32}$$

The elements of the matrix are sensitivity indexes. The global sensitivity indexes are used, and are calculated as follows:

$$S_{ij} = \frac{V_{ij}}{V_j} \tag{33}$$

Here, V_{ij} is the conditional variance of the performance index with regards to the i-th element of u at the j-th sampling time, and V_j is the total variance at that sampling time. The global sensitivity analysis is used to measure the variance created by the distribution of u in J, which is used to calculate the sensitivity indexes S_{ij} (Saltelli et al., 2005; Plischke, 2010). In the sensitivity analysis, the most probable parameterization of θ is chosen to be used in the model.

The objective function for the information-content problem is calculated using the well-known D-criterion for design of experiments (Franceschini and Macchietto, 2008):

$$J_I = \det\left(Q.Q^{-1}\right) \tag{34}$$

For maximizing the relevance of data sampled throughout an experiment, the information-content problem is then defined by Eq. (35)–(40):

$$max_{ts} J_I = \det\left(Q.Q^{-1}\right) \tag{35}$$

$$y = f\left(\theta, u\right) \tag{36}$$

$$Q = g\left(\theta, u, ts\right) \tag{37}$$

$$ts_{min} \leq ts_o \tag{38}$$

$$\Delta ts_{min} \leq ts_{i+1} - ts_i \tag{39}$$

$$ts_{end} \leq ts_{max} \tag{40}$$

Here $ts_{\min}$, $\Delta ts_{\min}$ and $ts_{\max}$ are minimal and maximal values for the sampling times (or sampling intervals) to be taken, related to constraints for meaningful sampling and proper use of analytical techniques; g represents the system of equations related to the calculation of Q. As was stated before, the process variables used in the experiment are the ones defined by solving the performance optimization sub-problem which generates model-optimized operating conditions.

4.4. Implementation of the new operating conditions

Once the model has been parameterized and the experiment has been designed, the operating conditions are tested experimentally. Data are gathered in this new run according to the sampling schedule, and the performance of the processes is compared to the performance of current optimal operating conditions. If the performance improves, the model is updated using data from this run and a new experiment is designed. If the performance does not improve, this is an indication that the extrapolation capabilities of the model worsened when moving away from the operating conditions used to fit it. In order to avoid convergence problems (Srinivasan and Bonvin, 2003), the following approach is taken: the optimization region (defined by the lower and upper bonds) is shrunk around the current optimal operating point, where the model is supposed to predict well. In the reduced region, a new experiment is designed without updating the model. If the performance improves with this design, a new iteration may begin by updating the model. If it does not improve, a new shrinking operation may be applied.

A simply algorithm is used here to shrink the feasible region. It is worth defining first the proposed optimal policy in each experiment, u^*, and the first policy tried in the current iteration (which is the best policy tried up to this point, namely the one found in the last iteration), u_0. For each element j, if the corresponding element of u^* is less than the corresponding element of u_0, set the new lower bound as follows:

$$u_{min_j} = u_{0_j} + \left(u^* j - u_{0_j}\right) sf \tag{41}$$

Otherwise, set the new upper bound as follows:

$$u_{max_j} = u_{0_j} + \left(u^* j - u_{0_j}\right) sf \tag{42}$$

Here sf is the shrinking factor, which is one of the hyper-parameters of the methodology. Using this algorithm, the operating region where the tendency of the model is not good enough (the model predict an improvement that is not real) is excluded from the next optimization step.

The methodology may be repeated until convergence, i.e. the current optimal operating conditions u_0 and the predicted one u^* fulfill the following criterion:

$$\frac{u^* - u_0}{u_0} \leq \varepsilon \tag{43}$$

where ε is a hyper-parameters that has to be chosen by the user. When the optimization region is repeatedly shrunk, it is reasonable to expect that the method will converge, since the difference among operating conditions will be small enough to fulfill Eq. (43). Because experimentation is costly in terms of time and money, an additional stopping criterion is added. If the optimization region is shrunk without improving for an arbitrary number of times, the procedure stops. If the number of times the region is shrunk within an iteration is labeled m and the maximum number m_{MAX}, then the stopping criterion is:

$$m = m_{MAX} \tag{44}$$

Finally, since the budget for experimentation is usually limited, a maximum number of experiments may be specified. After any of these criterions is fulfilled, all the experiments performed are checked, and the operating conditions that lead to better performance are chosen as the optimal ones. The proposed method is presented in Fig. 4 (Luna and Martinez, 2015).

Every time an iteration is completed, and an improvement of process performance is achieved, the bounds for the optimization problem are reset to their original values. This may be experimentally expensive, since operating conditions with low probabilities of high performances may be selected again. However, resetting bounds adds exploration to the optimization methodology and circumvent premature convergence due to the reduction of the optimization region. It is worth noting that if the process (actual) optimum is located in the region excluded from optimization problems within a given iteration, it will be unreachable in posterior iterations, and the search would end with suboptimal operating conditions.

Table 3
Process variables and sampling time schedule for experiment #1.

Process variables					
Variable	u_1	u_2	u_3	u_4	u_5
Value [%]	1.7	0.2	0.4	0.5	0.5
$u_{\min}$ [%]	1	0	0	0	0
$u_{\max}$ [%]	3	3	3	3	3

Sampling time schedule						
Sampling time	ts_1	ts_2	ts_3	ts_4	ts_5	ts_6
Time [h]	2.50	3.5	5	6.5	8	9

5. Experimental results

The method presented in Section 4 is used to solve the optimization problem discussed in Section 2. The model used here is the one presented in Section 3. The hyper-parameters for the optimization method are chosen as 0.5 for sf, 2 for m_{MAX} and 5.10^{-2} for ε. The initial feeding profile is designed based on *a priori* knowledge taken from the bibliography and from batch experiments carried out at a smaller scale. The initial sampling schedule is chosen arbitrarily. Sampling times are presented in Table 3, along with the bounds for the process variables. The performance of the first experiment is $J = 0.2977$ g/h. Data gathered during the experiment along with the model predictions for biomass and glucose are shown in Fig. 5. Parameter distributions are shown in Fig. 6. As can be seen, the distributions are non-normal, in part due to the bounds enforced in the parameterization problem that tend to increase the frequency of the parameters near their minimum and maximum allowable values.

The model is then used to design another optimization experiment. The process variables and the sampling schedule are shown in Table 4. Using this design, a new experiment is carried out. The model predictions for this design are shown in Fig. 7, along with data gathered for biomass and glucose concentrations. The performance of the process decreases to $J = 0.2635$ g/h, thus a shrinking

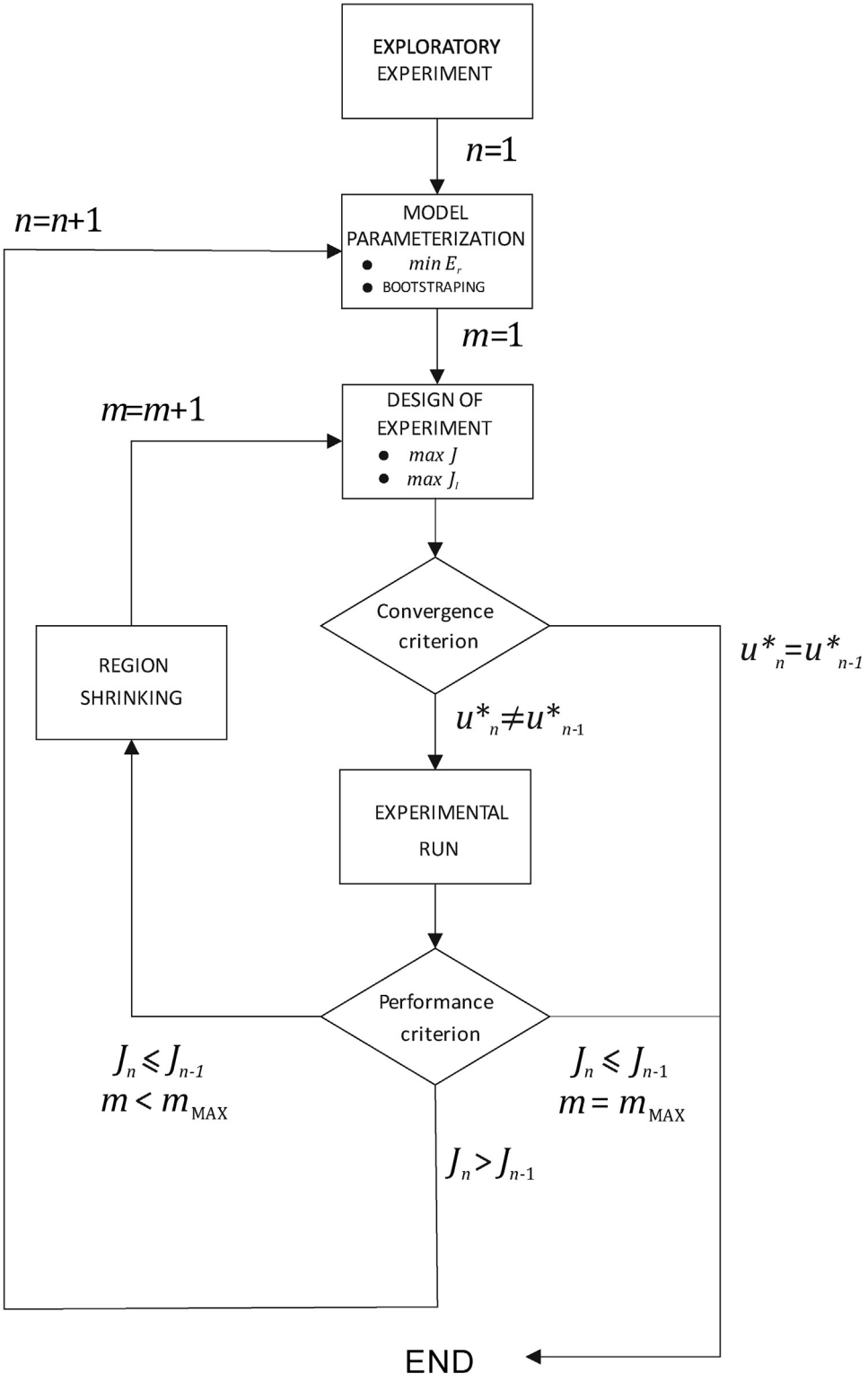

Fig. 4. Block diagram for the optimization methodology proposed in this paper.

of the optimization region is made, followed by a new experimental design. The corrected bounds and the re-optimized design are presented in Table 5.

Experiment #3 is carried out with a performance of $J = 0.5557$ g/h, an improvement regarding the best performance previously found (the one for the first experiment). Thus, data from this experiment is used to update the model, and its predictions for this experiment are shown in Fig. 8 (along with data for biomass and glucose concentrations).

The bounds for the process variables are reset to their initial values and a new iteration begins. Experiment #4 is first designed and carried out but it fails to improve the performance of the process ($J = 0.3276$ g/h). A re-optimization step is then carried out in a smaller feasible region for optimization but it fails again to improve

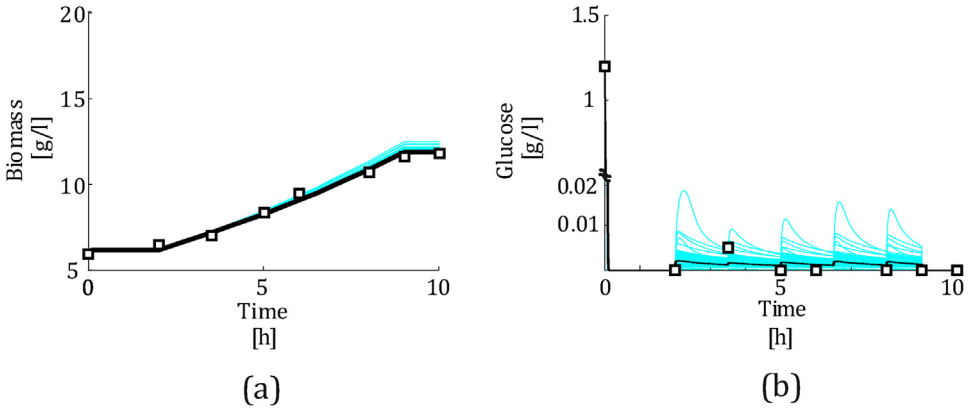

Fig. 5. Biomass and glucose concentration data points and model predictions for experiment #1.

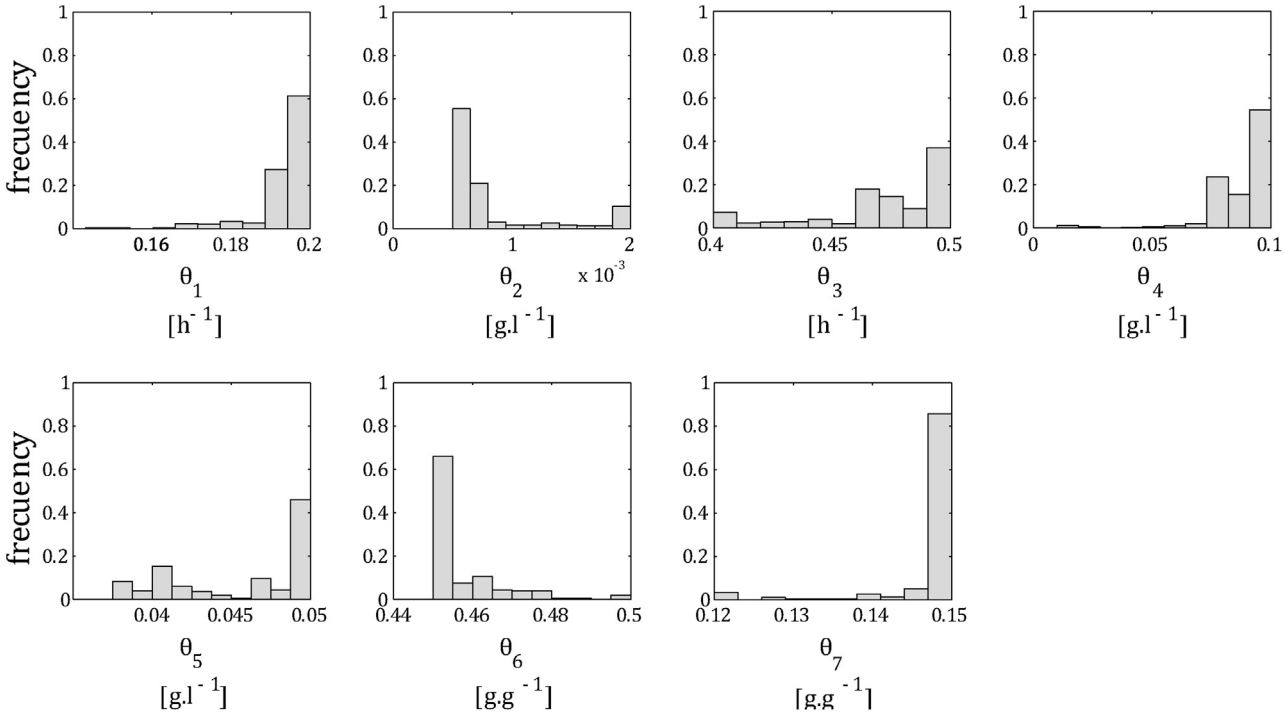

Fig. 6. Distributions (histograms) of model parameters obtained using data from experiment #1.

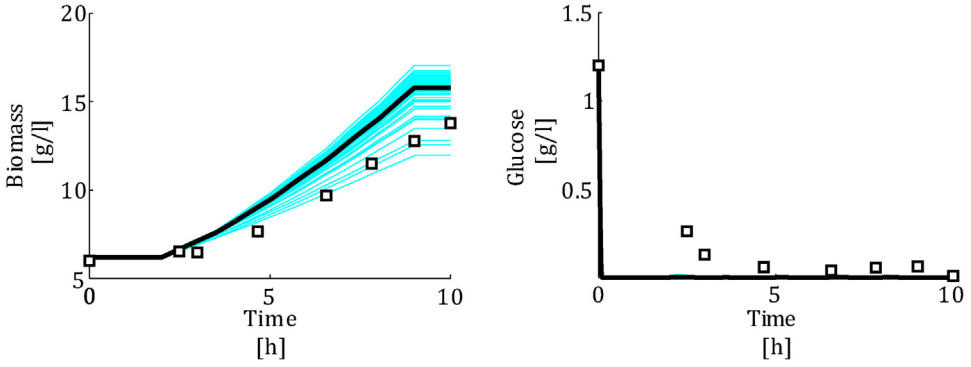

Fig. 7. Biomass and glucose concentration data points for experiment #2 and model predictions using the parameterization from experiment #1 data.

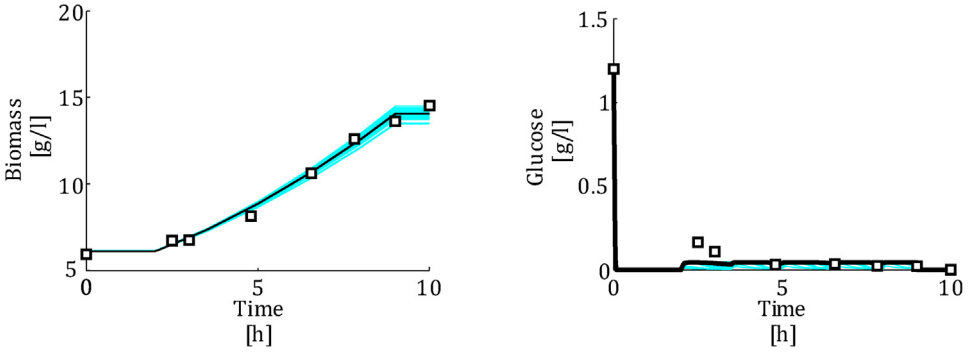

Fig. 8. Biomass and glucose concentration data points and model predictions for experiment #3.

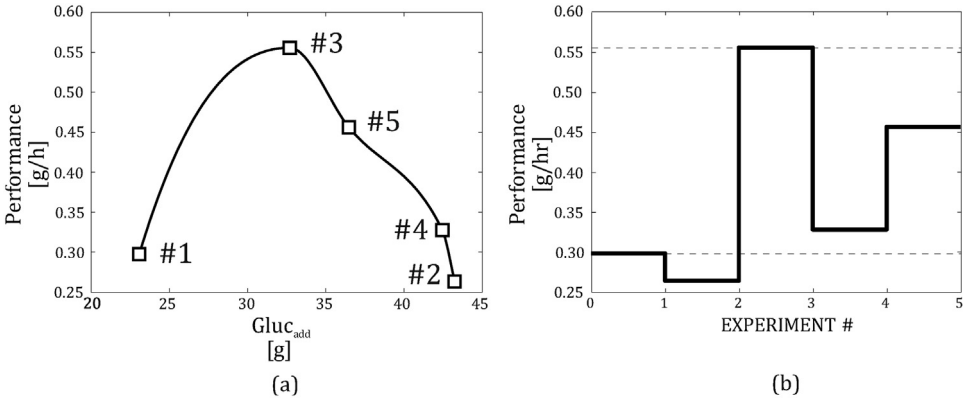

Fig. 9. (a) Performance of the different experiments related to the total glucose addition. (b) Learning curve of the implementation.

Table 4
Process variables and sampling time schedule for experiment #2.

Process variables					
Variable	u_1	u_2	u_3	u_4	u_5
Value [%]	2.4	1.2	0.9	1.4	1.6
u_{min} [%]	1	0	0	0	0
u_{max} [%]	3	3	3	3	3
Sampling time schedule					
Sampling time	ts_1	ts_2	ts_3	ts_4	ts_5	ts_6
Time [h]	2.49	2.99	4.66	6.56	7.82	9

Table 5
Process variables and sampling time schedule for experiment #3.

Process variables					
Variable	u_1	u_2	u_3	u_4	u_5
Value [%]	2	0.7	0.7	0.9	1.1
u_{min} [%]	1	0	0	0	0
u_{max} [%]	2	0.7	0.7	0.9	1.1
Sampling time schedule					
Sampling time	ts_1	ts_2	ts_3	ts_4	ts_5	ts_6
Time [h]	2.5	3	4.8	6.56	7.82	9

the bioreactor performance (J = 0.4563 g/h). At this point, the maximum number of experiments per iteration has been achieved without an improvement, thus the methodology is stopped and the process variables used in experiment #3 are chosen as the optimal operating conditions.

The summary of the implementation is presented in Table 6, where the experiment with the best performance has been highlighted in bold. As can be seen, the performance of the process significantly improves during the implementation. After an initial experiment where the performance decreases, a new operating point is achieved. After that, two new experiments are made. Even though better performance levels are found, they are still lower than the one found in experiment #3, thus the method stops. The amount of glucose fed to the reactor can be used to represent the feeding profile in order to depict the relationship between it and the process performance. Of course, Glc_{add} is a partial indicator since the way the glucose is fed is important as well, but this is done only to represent the aggregate effect of the feeding profile in a real number. The relationship is shown in Fig. 9a, along with the learning curve for the implementation of proposed iterative scheme, Fig. 9b. Complete data from all the experiments can be found in the Supporting Information.

6. Conclusions

In this paper, a method that combines mathematical modeling with experimental feedback was applied to the optimization of biomass production in a fed-batch reactor. The performance of the process was analyzed with an economic approach (which is incorporated in the objective function), describing a tradeoff between high biomass production and low glucose consumption.

The mathematical model used for optimization was presented and analyzed. It captures the metabolic modes of glucose consumption of yeast, which depend on process conditions, specially the excess and scarcity of the carbohydrate fed to the bioreactor. The model structure is simple enough for parameterizing it with few data points and still it is suitable for iterative optimization. The prediction errors due to extrapolation to untried operating conditions are corrected using experimental feedback.

The proposed methodology is briefly described and tried experimentally. After a few experiments, a sensibly improved operating

Table 6
Summary of experiments for the yeast biomass production optimization.

Iteration #	Experiment #	$F_{(1)}$ [%]	$F_{(2)}$ [%]	$F_{(3)}$ [%]	$F_{(4)}$ [%]	$F_{(5)}$ [%]	Glc_{add} [gr]	$Y_{overall}$ [gr/gr]	J [g/hr]
0	1	1.7	1.9	2.3	2.8	3.3	23.08	0.429	0.2977
1	2	2.4	3.6	4.5	5.9	7.5	43.23	0.361	0.2635
	3	***2.0***	***2.7***	***3.4***	***4.3***	***5.4***	***32.70***	***0.470***	***0.5557***
2	4	2.4	4.1	4.8	5.5	6.3	42.45	0.377	0.3276
	5	2.1	3.2	3.8	4.9	5.9	36.47	0.425	0.4563

point is found. Other experiments are tried but they fail to improve the performance, thus the best operating point found is considered to be the optimal operating condition. This experimental implementation of iterative modeling and optimization is useful to shown the capability of the method to obtain improved operating conditions using simple models and experimental feedback. Furthermore, the full implementation of modeling for optimization approaches presented in this paper is important as a proof of concept, since experimental implementations of optimization with imperfect models are scarce in the related literature.

Modeling for optimization is gaining momentum in process system engineering mainly motivated by innovative products and processes. Regarding the best integration of modeling and optimization, there is still plenty of room for improvement and new concepts. In the proposed approach, further research efforts will be put in solving the conflict between model exploitation and efficient exploration of the feasible region. As an example, using information from previous experiments may help reducing the search region in a more efficient manner so as to avoid operating conditions with little chances of including the actual optimum. Also, a tighter integration of active learning with process model exploitation is needed to generate highly informative data in few experiments. To this aim, methods and algorithms involving reinforcement learning and Bayesian optimization may be used.

Acknowledgment

The authors would like to thank the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) for their financial support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.compchemeng.2017.04.020.

References

Bergmeyer, H.U., 1974. Methods of Enzymatic Analysis. Academic Press, New York.
Bonné, D., Alvarez, M.A., Jorgensen, S.B., 2014. Data driven modeling for monitoring and control of industrial fed-batch cultivations. Ind. Eng. Chem. Res. 53 (18), 7365–7381.
Bonvin, D., Georgakis, C., Pantelides, C.C., Barolo, M., Rodrigues, D., Schneider, R., Dochain, D., Grover, M.A., 2016. Linking models and experiments. Ind. Eng. Chem. Res. 55 (25), 6891–6903.
De Deken, R.H., 1966. The crabtree effect: a regulatory system in yeast. J. Gen. Microbiol. 44 (2), 149–156.
Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman & Hall, London.
FDA, 2006. Guidance for Industry: Q8 Pharmaceutical Development [Internet]. Guid. Ind.
Franceschini, G., Macchietto, S., 2008. Model-based design of experiments for parameter precision: state of the art. Chem. Eng. Sci. 63 (19), 4846–4872.
Georgakis, C., 2013. Design of dynamic experiments: a data-driven methodology for the optimization of time-varying processes. Ind. Eng. Chem. Res. 52 (35), 12369–12382.
Gernaey, K.V., 2015. A perspective on PSE in fermentation process development and operation. Comput. Aided Chem. Eng. 37, 123–130.
Goh, C.J., Teo, K.L., 1988. Control parametrization: a unified approach to optimal control problems with general constraints. Automatica 24 (1), 3–18.
Jia, L., Luo, N.S., Yuan, J.Q., 2007. A model-based estimation of the profit function for the continuous fermentation of baker's yeast. Chem. Eng. Process. 46 (11), 1215–1222.
Jorgensen, A., 1948. Micro-organisms and Fermentation. Charles Griffin, London.
Joshi, M., Seidel-Morgenstern, A., Kremling, A., 2006. Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems. Metab. Eng. 8 (5), 447–4455.
Kaspar von Meyenburg, H., 1969. Energetics of the budding cycle of Saccharomyces cerevisiae during glucose limited aerobic growth. Arch. Microbiol. 66 (4), 289–303.
Kasperski, A., Miśkiewicz, T., 2008. Optimization of pulsed feeding in a Baker's yeast process with dissolved oxygen concentration as a control parameter. Biochem. Eng. J. 40 (2), 321–327.
Kiparissides, A., Koutinas, M., Kontoravdi, C., Mantalaris, A., Pistikopoulos, E.N., 2011. Closing the loop in biological systems modeling — from the in silico to the in vitro. Automatica 47 (6), 1147–1155.
Lei, F., Rotboll, M., Jorgensen, S.B., 2001. A biochemically structured model for Saccharomyces cerevisiae. J. Biotechnol. 88 (3), 205–221.
Luna, M.F., Martínez, E.C., 2014. A Bayesian approach to run-to-run optimization of animal cell bioreactors using probabilistic tendency models. Ind. Eng. Chem. Res. 53 (44), 17252–17266.
Luna, M.F., Martinez, E.C., 2015. Run-to-run optimization of biodiesel production using probabilistic tendency models – a simulation study. Can. J. Chem. Eng. 93 (9), 1613–1623.
Mandur, J.S., Budman, H.M., 2015. Simultaneous model identification and optimization in presence of model-plant mismatch. Chem. Eng. Sci. 129, 106–115.
Martínez, E.C., Cristaldi, M.D., Grau, R.J., 2013. Dynamic optimization of bioreactors using probabilistic tendency models and Bayesian active learning. Comput. Chem. Eng. 49, 37–49.
Nielsen, J., 2013. Production of biopharmaceutical proteins by yeast: advances through metabolic engineering. Bioengineered 4 (4), 207–211.
OECD Indicators, 2015. Health at a Glance. OECD Publishing, Paris, pp. 2015.
Pisano, G.P., 1997. The Development Factory: Unlocking the Potential of Process Innovation. Harvard Business Press, Boston.
Plischke, E., 2010. An effective algorithm for computing global sensitivity indices (EASI). Reliab. Eng. Syst. Saf. 95 (4), 354–360.
Randez-Gil, F., Sanz, P., Prieto, J.A., 1999. Engineering baker's yeast: room for improvement. Trends Biotechnol. 17 (6), 237–244.
Richelle, A., Bogaerts, P., 2014. Off-line optimization of baker's yeast production process. Chem. Eng. Sci. 119, 40–52.
Richelle, A., Fickers, P., Bogaerts, P., 2014. Macroscopic modelling of baker's yeast production in fed-batch cultures and its link with trehalose production. Comput. Chem. Eng. 61, 220–233.
Rodrigues, F., Ludovico, P., Leão, C., 2006. Sugar metabolism in yeasts: an overview of aerobic and anaerobic glucose catabolism. Biodivers. Ecophysiol. Yeasts, 101–121.
Rose, A.H., Harrison, J.S., 2012. The Yeasts: Yeast Technology, Vol. 5. Academic press Inc., San Diego.
Royle, K.E., Jimenez, I., Kontoravdi, C., 2013. Integration of models and experimentation to optimise the production of potential biotherapeutics. Drug Discov. Today 18 (23–24), 1250–1255.
Saltelli, A., Ratto, M., Tarantola, S., Campolongo, F., 2005. Sensitivity analysis for chemical models. Chem. Rev. 105 (7), 2811–2828.
Serio, M. Di, Alteriis, E. De, Parascandola, P., Santacesaria, E., 2001. A general kinetic and mass transfer model to simulate the baker's yeast growth in bioreactors. Catal. Today 66, 437–445.
Shuler, M.L., Kargi, F., 2002. Bioprocess engineering: basic concepts. J. Control. Release, 293.
Sonnleitner, B., Käppeli, O., 1986. Growth of Saccharomyces cerevisiae is controlled by its limited respiratory capacity: formulation and verification of a hypothesis. Biotechnol. Bioeng. 28 (6), 927–937.
Srinivasan, B., Bonvin, D., 2003. Convergence analysis of iterative identification and optimization schemes. American Control Conference 3, 1956–1961.
Srinivasan, B., Sedghi, B., Bonvin, D., 2001. Characterization of the optimal solution for a class of hybrid systems. European Control Conference (ECC), 1455–1460.
Srinivasan, B., Bonvin, D., Visser, E., Palanki, S., 2003. Dynamic optimization of batch processes II. Role of measurements in handling uncertainty. Comput. Chem. Eng. 27 (1), 27–44.
Tilman, D., Cassman, K.G., Matson, P.A., Naylor, R., Polasky, S., 2002. Agricultural sustainability and intensive production practices. Nature 418 (6898), 671–677.

Valentinotti, S., Srinivasan, B., Holmberg, U., Bonvin, D., Cannizzaro, C., Rhiel, M., Von Stockar, U., 2003. Optimal operation of fed-batch fermentations via adaptive control of overflow metabolite. Control Eng. Pract. 11 (6), 665–674.
Van Dijken, J.P., Scheffers, W.A., 1986. Redox balances in the metabolism of sugars by yeasts. FEMS Microbiol. Lett. 32 (3–4), 199–224.
Van Hoek, P.I.M., Van Dijken, J.P., Pronk, J.T., 1998. Effect of specific growth rate on fermentative capacity of baker's yeast. Appl. Environ. Microbiol. 64 (11), 4226–4233.
Vemuri, J., Palanki, S., 2006. Optimal operation of baker's yeast fermentation in the presence of uncertainty. Asia-Pac. J. Chem. Eng. 1 (1–2), 92–103.
Wenger, K.S., 1994. Mixing and the Glucose Effect in Saccharomyces Cerevisiae. Colorado State University, Ft. Collins, CO (Ph.D. Thesis).