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