Mario R. Eden, Marianthi Ierapetritou and Gavin P. Towler(Editors)Proceedings of the
13thInternational Symposium on Process Systems Engineering – PSE 2018
July 1-5, 2018,San Diego,California, USA

Goal-directed Design of Dynamic Experiments for Cybernetic Models of Bioreactors

Martin F. Luna, Ernesto C. Martínez*

INGAR (CONICET-UTN), Avellaneda 3657, Santa Fe S3002 GJC, Argentina
ecmarti@santafe-conicet.gov.ar

Abstract

The distinctive feature of cybernetic models of bioreactors is their capacity to account for regulatory mechanisms in a cell metabolism by modeling the synthesis of enzymes and their activities. From a process engineering viewpoint, to guarantee its predictive capabilities regarding one or more process objective or goals (e.g. optimization, controllability, etc.), experimental data used to fit a cybernetic model parameters should be the most informative bearing in mind the adequacy of the resulting model to describe the specific objective of interest. To excite the most relevant metabolic modes in the cybernetic model, a dynamic experiment is optimally designed by accounting for the sensitivity of the chosen objective to operating conditions. The bioreactor feeding profile and sampling times are selected to maximize a global sensitivity index. As a case study, biomass production or fermentation to ethanol conversion in the fed-batch cultivation of *Saccharomyces Cerevisiae* are considered as alternative objectives.

Keywords: Bioreactor; Cybernetic Model; Design of experiments; Optimization.

1. Introduction

Mathematical models of bioreactors are gaining widespread acceptance in both academic and industrial practice since it is vividly clear that, in order to be competitive and to keep pace with the new developments in the biopharmaceutical industry, bioprocesses need to be designed, optimized, monitored and controlled using process engineering methods (van Daele, et al., 2017). Cell metabolism is well-known for its dynamic complexity in comparison with pure reaction networks. The microorganisms used as catalyst are living beings that shown different metabolic behaviors depending on the operating conditions. Thus, simple kinetic and transport phenomena are usually not enough to describe the system dynamics. Some insight into the regulatory system of the cell must be considered in order to have a useful model that can predict the response of the microorganism to changes in the abiotic environment. The supplied nutrients undergo different metabolic pathways within the cells, and these alternative uptake modes are regulated by the microorganism enzymes. In this regard, cybernetic models (Ramkrishna and Song, 2016) are capable of explicitly modeling the metabolic behavior by resorting to matching and proportional control laws for the production and activities of enzymes.

However, the complex nature of cell metabolism is translated into a complex formulation of the mathematical models. Usually, cybernetic models involve several differential equations with many parameters. Thus, experimental information from dynamic experiments is needed in order to parameterize them. But if a metabolic mode is not excited during an experimental run, it may be impossible to obtain meaningful values of the parameters related to it. This may hamper model prediction capabilities which may lead the industrial process to suboptimal operation. Thus, a method to optimize the Design

of Dynamic Experiments (DoDE) is needed in order to obtain the maximum information from each experimental run (Franceschini and Macchietto, 2008; Martinez et al., 2013). The design will depend on the way the information gain is measured. In this work, the sensitivity of the process objective will be evaluated as a scalar of the Global Sensitivity Matrix (GSM). Thus, the process objective chosen (e.g. biomass production, metabolite expression, etc.) will determine which metabolic modes are excited during the experiment so that all the available resources are used to account for them properly in the cybernetic model. Here, two possible objective functions are considered for a bioreactor involving glucose consumption by yeast in a fed-batch bioreactor. Both goal-directed optimal designs are compared and the adequacy of each is discussed. It is shown that explicitly accounting for a given objective in the DoDE is advantageous for data generation in fitting a cybernetic model that properly describes regulatory mechanisms among different metabolic pathways.

2. Cybernetic modeling

In the cybernetic approach for bioreactors, the regulatory system of the microorganism is modeled by describing the enzymes production rates and activities. In order to do so, it uses the matching and proportional laws to link the uptake rates to the different metabolic modes. Here, the term metabolic mode refers to the preferential uptake of a carbon source and the subsequent metabolic pathway that it undergoes within the cell. As a case study here, the cultivation of baker's yeast (*Saccharomyces Cerevisiae*) with a feed of glucose as a carbon source, there are three main metabolic modes: i) the uptake of glucose by a fermentative pathway (fermentation), ii) the uptake of glucose by an oxidative pathway (respiration) and iii) the uptake of ethanol (produced by fermentation) by an oxidative pathway. Each mode has different rates and yields, and the switch from one mode to another depends on the cybernetic laws (Ramkrishna and Song, 2016). In this work, the model used is presented in Eq. (1) through Eq. (8). The matching and proportional laws are described by Eq. (7) and Eq. (8).

$$\frac{dX}{dt} = \left(\sum_i r_i v_i - \frac{F}{V}\right) X \tag{1}$$

$$\frac{dGl}{dt} = (Gl_F - Gl)\frac{F}{V} + \left(\frac{r_1 v_1}{Y_1} - \frac{r_2 v_2}{Y_2}\right) X \tag{2}$$

$$\frac{dEt}{dt} = -Et\frac{F}{V} + \left(p\frac{r_1 v_1}{Y_1} - \frac{r_3 v_3}{Y_3}\right) X \tag{3}$$

$$\frac{dV}{dt} = F \tag{4}$$

$$r_i = e_i \mu_i \frac{S_i}{S_i + K_i}\frac{K_{EI}}{Et + K_{EI}} \tag{5}$$

$$\frac{de_i}{dt} = \alpha u_i \frac{S_i}{S_i + K_i} - \left(\sum_j r_j v_j - \beta \right) e_i + \alpha^* \tag{6}$$

$$v_i = \frac{fc_i\, r_i}{\max(fc_j r_j)} \tag{7}$$

$$u_i = \frac{fc_i\, r_i}{\sum_j fc_j\, r_j} \tag{8}$$

Here, *X* stands for biomass concentration, *Gl* stands for glucose concentration (subscript *F* indicates that it is the concentration in the feed) and *Et* stands for ethanol concentration. Eq. (5) through Eq. (8) are repeated for *i*=1, 2, 3, which correspond to the three modes: fermentation, respiration and ethanol oxidation. In Eq. (5) and Eq. (6) S_i is equal to *Gl* for *i*=1,2 and to *Et* for *i*=3. *V* stands for volume, *F* stands for the feeding profile, *e* stands for the enzymes concentration and *v* and *u* for the cybernetic variables. In Eq. (7) and Eq. (8) fc_i is the number of carbon atoms per molecule of substrate, which is 6 for *i*=1, 2 and 2 for *i*=3. The rest of the symbols shown correspond to the model parameters to be fitted by the method.

The bioreactor is operated in fed-batch mode with complete mixing and it is assumed that no diffusion limitations are present (thus oxygen balance is not taken into account). The feeding profile is parameterized as a sequence of steps. The cybernetic model is parameterized by regression of experimental data generated in one or more dynamic experiments. Distributions (histograms) of the model parameters are generated using bootstrapping and the least-square error method is used to fit them.

3. Design of Dynamic Experiments

The complexity of the proposed model and the number of parameters involved to optimize the information content *J* of each experiment in order to minimize experimental costs and time constraints. Based on an a priori chosen distribution model parameters, the DoDE optimization problem is posed here as follow:

$$\begin{aligned} &\max\ J = det(Q'.Q) \\ &\text{s.t: } \varphi_{LB} \le \varphi \le \varphi_{UB} \end{aligned} \tag{9}$$

In this work, the information content of a dynamic experiment is measured as the determinant of the matrix *Q* defined below. The design vector φ includes the feeding profile parameters and the sampling schedule. Each entry in the vector φ has lower and upper limits φ_{LB} and φ_{UB}, respectively. Each entry of the matrix *Q* corresponds to the global sensitivity index SI_{ij} of the process objective function with respect to the model parameters *i* at sampling time *j*. Accordingly, the Global Sensitivity Analysis (GSA) is defined as follows:

$$Q = \begin{pmatrix} SI_{11} & \dots & SI_{1j} \\ \vdots & \ddots & \vdots \\ SI_{i1} & \dots & SI_{ij} \end{pmatrix} \tag{10}$$

This definition of the information content J is key in the formulation of the mathematical program (9). Since the cybernetic model can be used for different process objectives (e.g. biomass or ethanol production), different experimental designs can be used in order to gather more relevant information about the influence of model parameters in better predicting mode switching and its influence on the process objective function of interest.

It is worth noting that the evaluation of the information content J is computationally expensive due to GSA. A Bayesian optimization (BO) method is used here to find the global solution to the mathematical program in Eq. (9). The details for the implementation of the BO algorithm used can be found elsewhere and references therein (Shahriari et al., 2016).

4. Results

The proposed method is used to design dynamic experiments to parameterize a cybernetic model that describes the operation of a bench-scale fed-batch bioreactor containing baker's yeast, with glucose as carbon source. The overall duration of the experiment is set to 9.5 h, including a batch phase of 2 h after which the feeding of glucose begins. The feeding profile is made up of 5 successive steps of fixed duration (1.5 h), whose inflow rates are the entries of the design vector φ. The concentration of glucose in the bioreactor feed is 100 g/L. The sampling schedule consists of six sampling times, of which five has to be designed and the remaining one corresponds to the final time (9.5 h). The concentrations of biomass, glucose and ethanol are determined at each sampling time. The initial volume of the bioreactor is 1 L, with an inoculum of yeast of 6 g/L and initial glucose concentration of glucose of 1.2 g/L (no ethanol is present at the beginning of each experiment).

The initial parameterization of the cybernetic model is fitted with data gathered from the experiments performed in the work of Luna and Martinez (2017). In the available experimental data, the ethanol concentrations at sampling times were not available. Furthermore, the process objective function was the production of biomass, thus the glucose oxidative pathway, which has the higher biomass yield, was of special concern. As a result, the distribution of parameters related to the production and uptake of ethanol has a significantly higher variance.

Two different process objective functions J_p are proposed here to compare the influence of these goals on the design of the experiments. First, the process objective function is chosen to be the production of biomass, measured as the mass of yeast in the reactor. Alternatively, the production of ethanol is selected.

4.1. Biomass production

Using the cybernetic model fitted with the aforementioned data, the DoDE is performed using the production of biomass as the process objective function:

$$J_p = X\,V \tag{11}$$

The feeding profile is presented in Fig. 1(a) along with optimal sampling times. In Fig. 1(b), the cybernetic variables v for the three main metabolic modes are shown. Time-varying profiles of each cybernetic variable is an indicator of the prevalence of each mode over the experiment (a value of 1 indicates that the enzyme of the mode has maximum activity). As can be seen, during the first 2 h when no glucose is fed, the initial amount of glucose is consumed by the fermentative pathway (producing ethanol). When the

concentration of glucose decreases the importance of the glucose oxidative mode increases and even surpass the fermentation mode, but it is rapidly overtaken by the ethanol oxidative mode. This metabolic mode prevails until fed-batch operation begins and the cell favors the glucose fermentative mode again. However, the glucose oxidation pathway is not at all absent and gains momentum as the glucose concentration decreases again (due to an increase in biomass). When the inlet flow rate drops after the fifth hour, the glucose oxidative mode becomes prevalent.

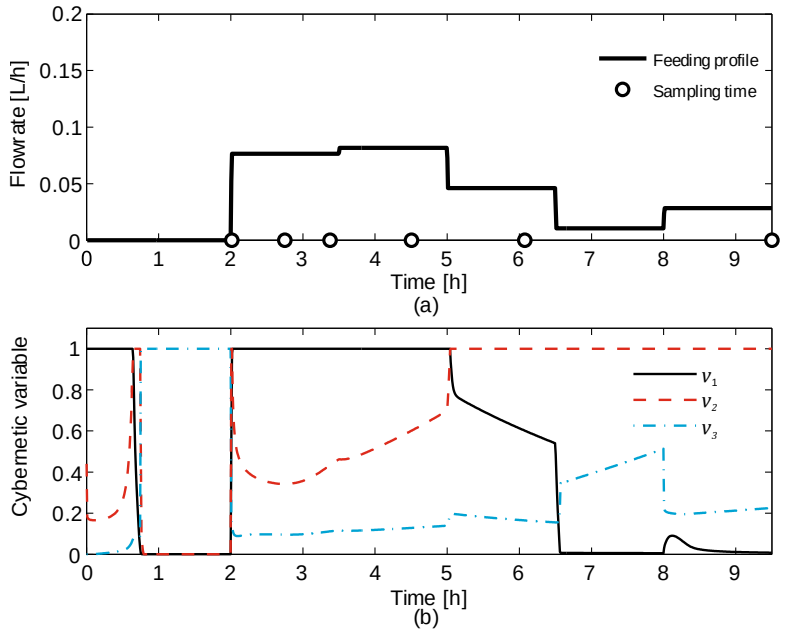

Figure 1. (a) Optimal feeding profile and sampling schedule for the case of biomass production. (b) Values of cybernetic variables for the optimal experimental design.

4.2. Ethanol production

A new experiment is designed using the same model parameters, but using the production of ethanol as the objective function:

$$J_p = Et\,V \tag{12}$$

The feeding profile and the sampling schedule are shown in Fig. 2(a), whereas the cybernetic variables *v* are depicted in Fig. 2(b). As can be seen, all metabolic modes are conveniently excited which generates highly informative data.

It is worth cross-comparing the values of *J* for both optimal designs. The value of *J* for the first design (the one obtained section 4.1), using the biomass production goal is 4.75 10^{-11}, whereas for the second optimal design (the one obtained in this section), using the same process objective the value for the information content *J* is 1.85 10^{-11}. The value for the information content *J* in the first and second DoDE using the ethanol production goal are 1.96 10^{-13} and 1.54 10^{-10}, respectively. The results obtained from cross comparison among optimal designs and information contents are rather expected: the optimal DoDE is dependent on the process objective the model is helping to optimize.

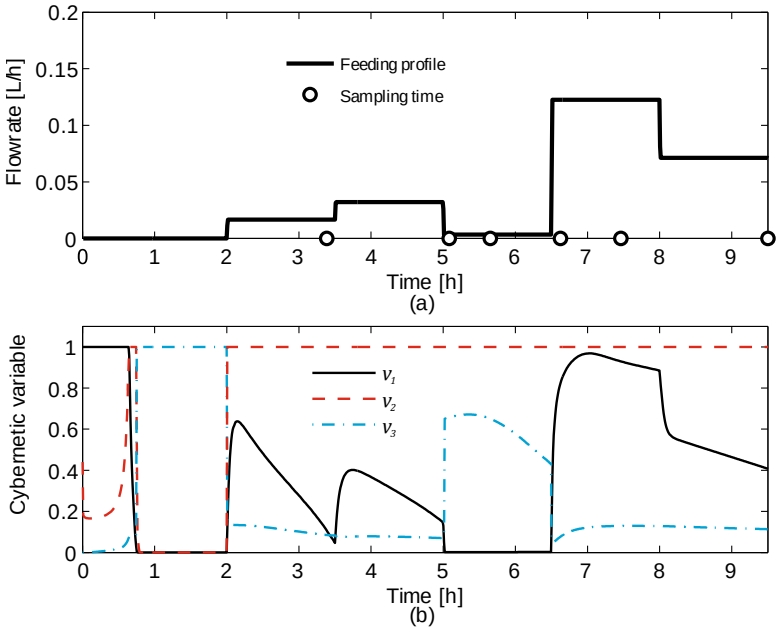

Figure 2. (a) Optimal feeding profile and sampling schedule for the case of ethanol production. (b) Values of cybernetic variables for the optimal experimental design.

5. Conclusions

In this work, a DoDE method for goal-directed parameterization of cybernetic models has been proposed. The importance of the process objective to be modeled is explicitly accounted for in the design of a dynamic experiment, since the same model may need different types of information to better describe the process objective that is being pursued. The method has proven to be useful in the maximization of the information content. However, the high computational burden of the GSA gives rise to the need of speeding up simulation-intensive algorithms used to deal with parameter distributions.

References

T. van Daele et al., 2017, Application of Iterative Robust Model-based Optimal Experimental Design for the Calibration of Biocatalytic Models, Biotechnology Progress, 33, 5, 1278–1293.

G. Franceschini, S. Macchietto, 2008, Model-based design of experiments for parameter precision: State of the art, Chemical Engineering Science, 63, 19, 4846-4872.

M.F. Luna, E.C. Martínez, 2017, Iterative modeling and optimization of biomass production using experimental feedback, Computers & Chemical Engineering, 104, 151-163.

E. C. Martínez, M.D. Cristaldi, R.J. Grau, 2013, Dynamic optimization of bioreactors using probabilistic tendency models and Bayesian active learning, Computers & Chemical Engineering, 49, 37-49.

D. Ramkrishna, H.S. Song, 2016, Analysis of Bioprocesses. Dynamic Modeling is a Must, Materials Today: Proceedings, 3, 10, 3587-3599.

B. Shahriari, K. Swersky, Z. Wang, R.P. Adams, N. de Freitas, 2016, Taking the human out of the loop: A review of Bayesian optimization, Proceedings of the IEEE, 104, 1, 148-175.