Biochemical Engineering Journal 100 (2015) 41–49

Contents lists available at ScienceDirect

Biochemical Engineering Journal

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

Regular article

Systematic methodology for bioprocess model identification based on generalized kinetic functions

Anne Richelle*, Philippe Bogaerts

3BIO-BioControl, Université Libre de Bruxelles, Avenue F. Roosevelt, 50, CP 165/61, 1050 Brussels, Belgium

ARTICLE INFO

Article history:
Received 14 January 2015
Received in revised form 27 March 2015
Accepted 6 April 2015
Available online 9 April 2015

Keywords:
Dynamic modeling
Fed-batch culture
Hybridoma cultures
Kinetic parameters
Production kinetics
Substrate inhibition

ABSTRACT

This study presents a new systematic methodology for kinetic model identification on the basis of available experimental measurements. In a first step, a general kinetic model [Syst. Anal. Model. Simul. 35 (1999) 87–113] is identified, which allows, on the one hand, capturing the activation and/or inhibition effects of any component involved in the reaction and, on the other hand, identifying all the parameters based on a simple linear regression. This circumvents the tedious problems of choosing the kinetic structure and providing initial parameter values on a trial-and-error basis. In a second step, the general kinetic model can be easily transformed into the general extended Monod formalism. The global identification of the nonlinear model is finally performed based on the results of the previous steps. The model and the experimental field (fed-batch culture experiments using hybridoma cell line HB-58) presented in Amribt et al. [Biochem. Eng. J. 70 (2013) 196–209] are used as case study to underline the advantages of this strategy for the global identification of nonlinear models. The proposed systematic procedure leads to the identification of a model structure with similar complexity but whose parameter values present lower variation coefficients. The identified model successfully reproduces the dynamics associated with substrates consumption (glucose and glutamine), metabolites production (lactate and ammonia) and cell growth.

1. Introduction

The issue of bioprocess modelling from extracellular measurements (macroscopic modelling) has been considered for a long time in the literature [1]. This kind of model requires the determination of a macroscopic reaction scheme and appropriate kinetic functions inspired from the available experimental data in order to obtain a general formulation of the bioprocess dynamics. These choices are based on a priori knowledge of the underlying phenomena influencing the biosystem behavior. However, the selection of adequate model structures often relies on a trial-and-error method.

Some methods for the objective determination of macroscopic reaction scheme and stoichiometric identification have been developed [2–8] while the determination of the kinetic structure seems to remain based on arbitrary choices. Indeed, there exists in the literature a profusion of apparently equivalent laws allowing the description of specific kinetic phenomena (e.g., activation, saturation and inhibition) and each of them has specific mathematical properties. A general kinetic formalism using power-law equations has been proposed by Savageau [9,10] to represent any nonlinear function. These canonical modelling forms (S- and Volterra-systems, generalized mass actions, etc.), developed in the framework of biochemical system theory, allow a general and systematic representation of an enormous variety of differential equations and were applied to numerous modelling case studies to describe the overall reaction rates of biological complex systems [11–16]. However, the power-law formalism does not allow the description of a double component effect (i.e., when a macroscopic species is activator and inhibitor of a reaction, depending on its concentration range). Recently, Mailier and Vande Wouwer [17] have developed a decision algorithm for the selection of the most likely kinetic structure among candidate models. While this kinetic determination procedure is objective, it is not really systematic as it depends on the arbitrary choice of appropriate candidate kinetic functions taken into account in the decision graph used by the algorithm. Note that it exists also the overall framework of black-box modelling, whose most famous models are based on the concept of neural networks [18–21]. They benefit from a high flexibility for reproducing several nonlinear kinetic phenomena but they do not keep the physical (biological) meaning of the model.

* Corresponding author. Tel.: +32 2 650 66 08.
E-mail addresses: arichell@ulb.ac.be (A. Richelle), philippe.bogaerts@ulb.ac.be (P. Bogaerts).

http://dx.doi.org/10.1016/j.bej.2015.04.003
1369-703X/

In addition to the difficulty associated with the kinetic model selection, the parameter estimation related to the kinetic structure identification is far from an easy task because these functions often involve a large number of highly correlated parameters. In this context, Haag et al. [22] have proposed an extension of the Michaelis–Menten kinetics in order to avoid identifiability problems related to a potential over-parameterization with the use of the classical extended Monod law. While this general formalism allows either activation/saturation or inhibition effects (but not both) with a reduced number of parameters, the kinetic expression presented in Haag et al. [22] leads to the usual parameter identification problems of nonlinear models. Indeed, for nonlinear models, the identification cost function presents many minima that can be local (minimum of the function in a limited region of the parameter space) or global (minimum throughout the entire parameter space). Actually, the convergence towards a (local or global) minimum strongly depends on the initial values of parameters that are provided to the optimization algorithm. This initial estimation of the parameter values is often a central problem of nonlinear model identification. Techniques such as multistart strategy (random generation of the initial parameter estimation) are currently used in order to cover the broadest spectrum of parameter values as possible. However, up to now, there is no general algorithm that could ensure the convergence to the global minimum in the case of nonlinear models.

This study presents a systematic methodology in three steps for the determination of kinetic structures based on the generalized kinetic formalism introduced by Bogaerts et al. [23] (and extended by Grosfils et al. [24]) which does not suffer from the previously cited model identification problems. Indeed, this general kinetic function allows the description of specific activation and/or inhibition effects of each component involved in the reaction scheme and its use within the proposed procedure permits the objective determination of all potential influencing factors (macroscopic components) involved in a reaction. Moreover, while this kinetic function remains nonlinear, it is linearizable with respect to its parameters thanks to a logarithmic transformation. Hence, based on a simple linear regression, initial estimates of all kinetic parameters can be obtained. In this contribution, it is shown how the identified general kinetic function can easily be transposed into the widely accepted extended Monod formalism, based on an efficient parameter identification procedure. Finally, the global identification of the model can be performed based on the results obtained in the previous steps. The whole procedure is illustrated and validated on real experimental data concerning fed-batch cultures of hybridoma cells. The gains of using this systematic procedure instead of a more classical macroscopic modelling approach [25] are highlighted.

The first part of this paper presents the theory related to generalized kinetic functions [23] in a macroscopic modelling context (Section 2) and their use in the development of a new systematic model identification procedure (Section 3). Section 4.1 presents the overflow metabolism model and the associated experimental field [25] used as case study to compare the results obtained with a usual identification strategy (Section 4.2) and the proposed systematic procedure (Section 4.3).

2. Macroscopic modelling approach and general kinetic structure

A general approach to describe the dynamics of a bioprocess from a macroscopic point of view has been proposed in Bastin and Dochain [1]. In this context, the bioprocess can be considered as a set of M reactions involving N macroscopic components ξ (substrates, products, or biomass). The expression of this kind of reaction scheme is the following:

$$\Sigma_{i\in R_k}\eta_{i,k}\xi_i \xrightarrow{\nu_k} \Sigma_{j\in P_k}\eta_{j,k}\xi_j k \in [1, M] \quad (1)$$

where ν_k is the rate of reaction k, $\eta_{i,k}$ and $\eta_{j,k}$ are pseudo-stoichiometric coefficients (yield coefficients), R_k and P_k represent the sets of components which are, respectively, substrates (or activators) and products of reaction k.

Based on the definition of the reaction scheme (1), the general dynamical model of the bioprocess is defined by the system of mass balances for each of the N components ξ_i, which can be written in the following matrix form [1]:

$$\frac{d\xi(t)}{dt} = K\nu(\xi, t) - D(t)\xi(t) + F(t) - Q(t) \quad (2)$$

where $\xi \in \Re^N$ is the vector of macroscopic species concentrations, $K \in \Re^{N\times M}$ is the pseudo-stoichiometric coefficients matrix ($N \geq M$), $\nu \in \Re^M$ is the vector of reaction rates, $D \in \Re$ is the dilution rate, $F \in \Re^N$ is the vector of external feed rates and $Q \in \Re^N$ is the vector of gaseous outflow rates.

In this formalism, the vector $\nu(\xi, t)$ of Eq. (2) describes the kinetics of the considered reactions by using phenomenological laws. In the development of a model, the formulation of theses kinetic structures is usually a very difficult task. Indeed, it is often not easy to clearly define the main influencing factors and their effects on a specific reaction. Moreover, there is always more than one possibility in the choice of an acceptable analytic description as it is always possible to describe the same effect with another equivalent analytic structure. Hence, several functions can be used to describe the influence of the component ξ_i, which could be substrates (S), products (P), or biomass (X) on the reaction rates $\nu_k(\xi)$.

The most famous kinetic model is probably the Monod's law which can be used more generally in its extended version, when more than one substrate and/or more than one product influence a reaction rate:

$$\nu_{k\_monod}(\xi) = \mu_{max,k}\Pi_{m\in R_k}\frac{\xi_m}{\xi_m + K_{k\_\xi m}}\Pi_{l\in P_k}\frac{K_{k\_I\xi_l}}{K_{k\_I\xi_l} + \xi_l}X \quad (3)$$

where $\mu_{max,k}$ is the maximum specific rate of reaction k, $K_{k\_\xi m}$ and $K_{k\_I\xi l}$ are respectively a saturation and an inhibition constant of reaction k, R_k and P_k are, respectively, the sets of indices of the components which activate and inhibit the reaction k.

As cited above, numerous equivalent kinetic functions, which are able to characterize the same behavior, are available in the literature and each of them has specific mathematical properties. In this context, Bogaerts et al. [23] have developed a kinetic model structure allowing the representation of the activation and/or the inhibition of the reaction by any component of the bioprocess. The reaction rate $\nu_k(\xi)$ with the generalized expression is the following:

$$\nu_{k\_gen}(\xi) = \alpha_k\Pi_{m\in R_k}\xi_m{}^{\gamma k\_m}\Pi_{l\in P_k}e^{-\beta_{k\_l}\xi_l} \quad (4)$$

where $a_k > 0$ is a kinetic constant, $\gamma_{k\_m\geq 0}$ the activation coefficient of component m (activators, e.g., substrates) in the reaction k, $\beta_{k\_1\geq 0}$ the inhibition coefficient of component l (inhibitors, e.g., products) in reaction k. Note that the expression Eq. (4) is not able to describe a saturation effect by a macroscopic component (as described in a Monod factor which saturates at $\mu_{max,k}$). Indeed, as underlined in the work of Grosfils et al. [24], the saturation effect in this structure results from the compensation for the activation by some inhibition. It has therefore been generalized to a structure describing the three effects: activation, saturation and inhibition [24].

The main advantage of structure Eq. (4) is the possibility to develop a systematic identification procedure. Indeed, many kinetic structures such as the extended Monod's law are nonlinear. Hence, the parameter identification of these models usually leads to