Research Article

Mammalian Cell Culture Process for Monoclonal Antibody Production: Nonlinear Modelling and Parameter Estimation

Dan Selişteanu,[1] Dorin Şendrescu,[1] Vlad Georgeanu,[2] and Monica Roman[1]

[1]*Department of Automation, Electronics and Mechatronics, University of Craiova, A.I. Cuza No. 13, 200585 Craiova, Romania*
[2]*Faculty of General Medicine, Carol Davila University of Medicine and Pharmacy, Eroilor Sanitari No. 8, 050474 Bucharest, Romania*

Correspondence should be addressed to Monica Roman; monica@automation.ucv.ro

Monoclonal antibodies (mAbs) are at present one of the fastest growing products of pharmaceutical industry, with widespread applications in biochemistry, biology, and medicine. The operation of mAbs production processes is predominantly based on empirical knowledge, the improvements being achieved by using trial-and-error experiments and precedent practices. The nonlinearity of these processes and the absence of suitable instrumentation require an enhanced modelling effort and modern kinetic parameter estimation strategies. The present work is dedicated to nonlinear dynamic modelling and parameter estimation for a mammalian cell culture process used for mAb production. By using a dynamical model of such kind of processes, an optimization-based technique for estimation of kinetic parameters in the model of mammalian cell culture process is developed. The estimation is achieved as a result of minimizing an error function by a particle swarm optimization (PSO) algorithm. The proposed estimation approach is analyzed in this work by using a particular model of mammalian cell culture, as a case study, but is generic for this class of bioprocesses. The presented case study shows that the proposed parameter estimation technique provides a more accurate simulation of the experimentally observed process behaviour than reported in previous studies.

1. Introduction

As the market demand for monoclonal antibodies is increasing, there is significant interest in developing proper models for mammalian cell culture processes, due to the fact that these are commonly used as production platforms for mAbs, which are the fastest growing segment of the biopharmaceutical industry [1–6]. For mAb production, various mammalian cell lines are usually exploited, such as murine myeloma (NS0), murine hybridomas, Chinese hamster ovary (CHO), and PER.C6 human cells. The selection of expression system is determined by its capability to deliver high productivity with suitable product quality attributes [7]. Medical applications for mAbs are quite extensive: diagnostic tools, therapies for various cancers, rheumatoid arthritis, cardiovascular conditions, and so on [4, 6–9].

Typically, the industrial operation for mammalian cell culture mAb platforms relies on empirical knowledge [2, 3, 10] and the improvements are achieved by using trial-and-error experiments and precedent practices. Consequently,

process improvements have generally been time-consuming and costly, with a high degree of specificity. To assist these laboratory experiments and, in practical terms, to achieve high productivity and better quality products, it is of obvious interest to develop model-based applications and to achieve accurate dynamical models. However, the specific characteristics of these processes, such as complexity, nonlinearity, and absence of cheap and reliable instrumentation, require an enhanced modelling effort and advanced kinetic parameter estimation strategies.

In order to surmount the above-mentioned limitations of trial-and-error process development, the so-called predictive models for mammalian cell culture processes are quite attractive [4]. Generically speaking, cell culture modelling techniques are classified based upon whether a dynamic or a pseudo-steady-state interpretation of cellular metabolism is used [2, 4, 11, 12]. Being well-known in control systems, the pseudo-steady-state approach has a biochemical interpretation in cell culture processes. It is assumed that all metabolites within the cell culture process are accumulated

or depleted at a rate considerably faster than the overall cell growth rate. Consequently, the concentration of each system metabolite and the rate of each metabolic reaction are all considered time-invariant [4]. This approach is simple and the obtained models are linear systems, which can be easily computed regardless of the model size (complexity). The information gathered in such pseudo-steady-state models concerns the metabolic configuration of cell culture. However, mammalian cells have a complicated internal structure, with several interconnected biochemical processes and with phenomena on multiple time scales. Thus, the pseudo-steady-state models cannot describe in detail the changes that occur over a continuous time-horizon (intracellular concentration profiles, changes in reaction rate due to gene regulation, etc.). Therefore, the dynamic modelling is more appropriate for these complex (and dynamical) processes. In this case, a system of differential equations will describe the bioprocess model. In many cases, the difficulty that arises is related to the computational problems, especially for large and stiff systems. No matter what modelling method is chosen, the complexity together with the nonlinearity of these processes is a limiting factor in model building.

In this paper, which is an extended work of [13, 14], an essential problem in dynamic modelling of cell culture systems is analysed, the so-called parameter estimation. The model of such bioprocesses can be obtained by using dynamic classical modelling (based on mass balance) or alternative approaches such as pseudo-bond-graph method (a version of bond graph method introduced by Paynter in 1961 and further developed in [15–26]). However, regardless of the modelling method, in order to obtain a dynamical model useful for process development (including the design of some control strategies), the nonmeasurable parameters of the mammalian cell culture system must to be estimated. However, any parameter in a cell culture model could [4] have physical meaning and be measurable by experiment, have clear physical meaning but be experimentally inaccessible, or have no clear physical meaning (e.g., be purely mathematical in nature). Typically, optimization-based techniques are used for the estimation of nonmeasurable parameters of such biological processes [4, 27, 28]. For example, a quadratic programming technique was used by Gao et al. [27], and a simple discretization scheme combined with a filtered interior point primal-dual line-search algorithm [29] was proposed by Baughman et al. [4]. Other nonlinear optimization-based techniques are the genetic algorithms, orthogonal collocation, and particle swarm optimization, PSO, which have been applied mostly on chemical processes (see, e.g., [30, 31]) or in gene regulatory networks modelling [32, 33]. In order to obtain accurate solutions in the case of the mAb production process, in this paper, a particle swarm-based multistep nonlinear optimization algorithm is proposed [34–36].

Concerning the applications of PSO for identification of biological systems, some results were reported for the process of glycerol fermentation by *Klebsiella pneumoniae* in batch, fed-batch, and continuous cultures [37–41]. The estimation approach used in these works is in most cases a parallel PSO technique, which requires a considerable computational effort. Another trend is related to an indirect use of PSO technique for estimation, more precisely for the training of a neural network, which models the bioprocess [42].

During the last decade, PSO algorithms have gained much attention and wide applications in different fields due to their effectiveness in performing difficult optimization issues, as well as simplicity of implementation and ability of fast converge to a reasonably good solution. PSO is a population-based heuristic global optimization technique, first introduced by Kennedy and Eberhart [43] and referred to as a swarm-intelligence technique. It is motivated from the simulation of social behaviour of animals such as bird flocking, fish schooling, and swarm. In this algorithm, the population is called a swarm and the trajectory of each particle in the search space is controlled through the medium of a term called "velocity," according to its own flying experience and swarm experience in the search space.

This paper proposes a multistep PSO version that uses time-varying acceleration coefficients [35], which is developed to solve the nonconvex optimization problem, ensuring fast convergence and very good performance. Finally, the obtained solution is an optimal set of the kinetic parameter values.

The proposed nonlinear modelling and estimation approaches are analyzed in this work by using a particular model of mammalian cell culture, as a case study, but they are generic for this class of bioprocesses. A previously published dynamic model of mammalian cell culture by Gao et al. in [27] is used as a case study. More precisely, a process of an Immunoglobulin G-secreting murine hybridoma cultured in a growth medium supplemented with proline, L-asparagine, and L-aspartic acid is taken into consideration.

2. Methods

2.1. MAb Synthesis by Mammalian Cell Culture: Process Description and Modelling Issues. In order to model the mammalian cell culture processes, first it is necessary to analyze the reconstruction of metabolic activities. However, the reconstruction generally includes only a subset of the highly active metabolic units found in proliferating mammalian cells [4, 34]. After the choice of this key subset, the next step consists in the modelling of the reactions' rates in the frame of reconstruction. This process is a very difficult one, and the modelling of reactions as single-enzyme processes by using *in vitro* kinetic parameters is possible only for simple and small reconstructions. Often, *in vitro* kinetic parameters do not compare well against *in vivo* observations [4]. The reconstruction in complex processes will frequently combine a number of discrete processes into a single lumped process and will then apply kinetic parameters to the lumped process. Consequently, some kinetic parameters that appear in the model may have small or no physical significance and usually their values are not experimentally measurable [4].

Therefore, if all of the interaction of metabolites and cell physiology are included in the modelling process, then the size of the obtained model is very large and it is not appropriate for model-based optimization and control purposes. The usual solution is to select a priori the elementary

TABLE 1: Macroreactions of the mAb production process [4, 27].

Reaction number	Macroreaction scheme
1	GLC $\rightarrow$ 2LAC
2	GLC + 2GLU $\rightarrow$ 2ALA + 2LAC
3	GLC + 2GLU $\rightarrow$ 2ASP + 2LAC
4	GLU $\rightarrow$ PRO
5	ASN $\rightarrow$ ASP + NH$_3$
6	GLN + ASP $\rightarrow$ ASN + GLU
7	0.0508GLC + 0.0577GLN + 0.0133ALA + 0.006ASN + 0.0201ASP + 0.0016GLU + 0.081PRO $\rightarrow$ BM
8	0.0104GLN + 0.011ALA + 0.072ASN + 0.082ASP + 0.0107GLU + 0.0148PRO $\rightarrow$ MAb
9	GLN $\rightarrow$ GLU + NH$_3$

reaction schemes and to relate the major macroscopic species such as biomass, essential substrates, and products by a set of so-called macroreactions [27]. Thus, a simplified model is obtained, which is suitable for optimization and control. As was mentioned before, the next step in the modelling is related to the determination of reaction kinetics and the final model is obtained based on mass balance equations of the macroscopic species involved in the reactions.

Next, a particular model of mammalian cell culture, published by Gao et al. [27], will be described and used as a case study. Gao et al. [27] provided a detailed description of an Immunoglobulin G- (IgG-) secreting murine hybridoma (130-8F *Sanofi Pasteur*) cultured in a D-MEM (Dulbecco's Modified Eagle Medium) growth medium supplemented with proline, L-asparagine, and L-aspartic acid. In this process, batch cultures of the organism were allowed to grow for a minimum of 7 days. The infrequent measured concentration data for glucose, lactate, and ammonia, as well as for 20 amino acids and the monoclonal antibody, were obtained from the collected samples via proper techniques. By using the measured data, the average rates of transmembrane fluxes were calculated for each metabolite for both the initial exponential growth phase and for the postexponential (decline) phase. Gao et al. [27] used the metabolic flux analysis (MFA) in order to calculate the unknown intracellular fluxes from measured extracellular fluxes by applying steady-state mass balance equations. The obtained metabolic network was constructed based on some preliminary studies [44–47], and it represents the significant metabolic pathways in proliferating animal cells. Gao et al. [27] determined that 16 reactions (a half of the total number) in the chosen reconstruction did not function significantly, and consequently these reactions, with an activity of about 1% of the total, were eliminated. The remaining subset of 16 reactions of the reduced metabolic reconstruction was further reduced by using a technique that combines reactions that share common metabolites [48]. Finally, the reduced reaction scheme for this mAb bioprocess contains a number of only 11 extracellular compounds and it consists of nine macroreactions, presented in Table 1 [4, 27].

The dynamical model of a generic bioprocess inside a bioreactor can be obtained by using the mass balance of the component and it is given by the following set of differential equations [49]:

$$\frac{d\xi}{dt} = K \cdot \varphi(\xi) + D \cdot \xi + F - Q, \qquad (1)$$

where $\xi = [\xi_1 \quad \xi_2 \quad \cdots \quad \xi_n]^T$ is the n-dimensional vector of the instantaneous concentrations (the concentrations of extracellular metabolites in our particular case), $\varphi = [\varphi_1 \quad \varphi_2 \quad \cdots \quad \varphi_m]^T$ is the vector of the reaction rates, and K is the $n \times m$ dimensional matrix of stoichiometric (or yield) coefficients, with $K = [K_{ij}]$, $i = \overline{1, n}$; $j = \overline{1, m}$, where $K_{ij} = (\pm)k_{ij}$ if $j \sim i$. The notation $j \sim i$ indicates that the sum is done in accordance with the reactions j that involve the components i. The sign of the yield coefficients k_{ij} is given by the type of the component ξ_i: plus (+) when the component is a reaction product and minus (−) otherwise. D is the specific volumetric outflow rate (h^{-1}), usually called dilution rate. In (1), $F = [F_1 \quad F_2 \quad \cdots \quad F_n]^T$ is the vector of rates of liquid supply and $Q = [Q_1 \quad Q_2 \quad \cdots \quad Q_n]^T$ is the vector of rates of removal of the components in gaseous form.

Model (1) describes in fact the behaviour of an entire class of bioprocesses and is referred to as the general dynamical state-space model of this class [49, 50]. In (1), the term $K \cdot \varphi(\xi)$ is in fact the rate of consumption and/or production of the components in the reactor, that is, the reaction kinetics. The term $-D\xi + F - Q$ represents the exchange with the environment. The strongly nonlinear character of this model is given by the reaction kinetics. In many practical situations the structure and the parameters of the reaction rates are partially known or even completely unknown.

Typically, in a batch process the reactor is filled with the reactant mixture: substrates and microorganisms. Then, the reactions occur inside the reactor for a time period; after that, the products are removed from the tank. Because the studied bioprocess takes place inside a batch reactor, model (1) becomes

$$\frac{d\xi}{dt} = K \cdot \varphi(\xi); \qquad (2)$$

that is, the term $-D\xi + F - Q$ (which represents the exchange with the environment) is zero in this particular batch mode.

For the mAb production process, the concentrations of the 11 extracellular metabolites (given in the reaction scheme from Table 1) constitute the elements of the state vector from model (1) and are denoted as follows:

$$
\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \\ \xi_9 \\ \xi_{10} \\ \xi_{11} \end{bmatrix} = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \\ P_6 \end{bmatrix} = \begin{bmatrix} \text{GLC} \\ \text{GLN} \\ \text{GLU} \\ \text{ASN} \\ \text{ASP} \\ \text{LAC} \\ \text{ALA} \\ \text{PRO} \\ \text{MAb} \\ \text{BM} \\ \text{NH}_3 \end{bmatrix}, \qquad (3)
$$

where GLC = glucose, GLN = glutamine, GLU = glutamate, ASN = asparagine, ASP = aspartate, LAC = lactate, ALA = alanine, PRO = proline, MAb = monoclonal antibody, BM = biomass, NH_3 = ammonia are the metabolites given in Table 1 (and for simplicity, the concentrations of the corresponding elements in model (1)).

However, in order to complete the model of the mAb production process, it is necessary to add the evolution of the viable cell concentrations of the culture, because the metabolite mass balances depend on the amount of viable cells. Gao et al. [27] noticed the typical behaviour of the batch culture, with exponential growth and postexponential decline, senescence phase (which occurs after the first phase of evolution, due to the aging of the cells and the accumulation of autoinhibitory metabolites). Therefore, another two concentrations enter in the complex model of the bioprocess, the viable cell concentration X and the dead cell concentration X_d. The dynamics of these concentrations will be modelled separately, depending of the phase (growth or decay).

Remark 1. To be exact, for the mAb production process the exchange with environment is zero except the CO_2 gaseous flow, but this flow is not measured and CO_2 is not predicted in the final model, as it is considered in [27].

In the following, the dynamical model (2) of the mAb production process will be presented, starting with the reaction scheme given in Table 1. Afterward, the problem of kinetic rates is addressed, together with the parameter estimation problem, via PSO-based techniques.

The dynamical model of the form (2) can be particularized for the mAb production process described by the reaction scheme from Table 1 by using the mass balance of the components (via classical methods [4, 27] or bond graph approach [13]) inside the batch reactor. The following dynamical model is obtained:

$$
\frac{dS_1}{dt} = -\varphi_1 - \varphi_2 - \varphi_3 - k_{1,7}\varphi_7,
$$

$$
\frac{dS_2}{dt} = -\varphi_6 - k_{2,7}\varphi_7 - k_{2,8}\varphi_8 - \varphi_9,
$$

$$
\frac{dS_3}{dt} = -k_{3,2}\varphi_2 - k_{3,3}\varphi_3 - \varphi_4 + \varphi_6 - k_{3,7}\varphi_7 - k_{3,8}\varphi_8 + \varphi_9,
$$

$$
\frac{dS_4}{dt} = -\varphi_5 + \varphi_6 - k_{4,7}\varphi_7 - k_{4,8}\varphi_8,
$$

$$
\frac{dS_5}{dt} = k_{5,3}\varphi_3 + \varphi_5 - \varphi_6 - k_{5,7}\varphi_7 - k_{5,8}\varphi_8,
$$

$$
\frac{dP_1}{dt} = k_{6,1}\varphi_1 + k_{6,2}\varphi_2 + k_{6,3}\varphi_3,
$$

$$
\frac{dP_2}{dt} = k_{7,2}\varphi_2 - k_{7,7}\varphi_7 - k_{7,8}\varphi_8,
$$

$$
\frac{dP_3}{dt} = \varphi_4 - k_{8,7}\varphi_7 - k_{8,8}\varphi_8,
$$

$$
\frac{dP_4}{dt} = \varphi_8,
$$

$$
\frac{dP_5}{dt} = \varphi_7,
$$

$$
\frac{dP_6}{dt} = \varphi_5 + \varphi_9.
$$

$$(4)$$

Model (4) can be written in a compact form [13]:

$$
\frac{d}{dt}\underbrace{\begin{bmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \\ P_6 \end{bmatrix}}_{\xi} = \underbrace{\begin{bmatrix} -1 & -1 & -1 & 0 & 0 & 0 & -k_{1,7} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -k_{2,6} & -k_{2,7} & -k_{2,8} & -1 \\ 0 & -k_{3,2} & -k_{3,3} & -1 & 0 & 1 & -k_{3,7} & -k_{3,8} & 1 \\ 0 & 0 & 0 & 0 & -1 & 1 & -k_{4,7} & -k_{4,8} & 0 \\ 0 & 0 & k_{5,3} & 0 & 1 & -1 & -k_{5,7} & -k_{5,8} & 0 \\ k_{6,1} & k_{6,2} & k_{6,3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & k_{7,2} & 0 & 0 & 0 & 0 & -k_{7,7} & -k_{7,8} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -k_{8,7} & -k_{8,8} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}}_{K} \underbrace{\begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \varphi_4 \\ \varphi_5 \\ \varphi_6 \\ \varphi_7 \\ \varphi_8 \\ \varphi_9 \end{bmatrix}}_{\varphi(\xi)},
$$

$$(5)$$

where the values of stoichiometric coefficients are given in the reaction schemes from Table 1 and are as follows [4, 27]: $k_{1,7} = 0.0508$, $k_{2,6} = 1$, $k_{2,7} = 0.0577$, $k_{2,8} = 0.0104$, $k_{3,2} = 2$, $k_{3,3} = 2$, $k_{3,7} = 0.0016$, $k_{3,8} = 0.0107$, $k_{4,7} = 0.006$, $k_{4,8} = 0.072$, $k_{5,3} = 2$, $k_{5,7} = 0.0201$, $k_{5,8} = 0.082$, $k_{6,1} = 2$, $k_{6,2} = 2$, $k_{6,3} = 2$, $k_{7,2} = 2$, $k_{7,7} = 0.0133$, $k_{7,8} = 0.011$, $k_{8,7} = 0.081$, and $k_{8,8} = 0.0148$.

The nonlinear dynamical model (5) is obvious from the general form (2). However, in order to complete the model of mAb production process, it is necessary to add the submodels corresponding to the dynamics of viable cell concentration and dead cell concentration, respectively. Here it should be noted that Gao et al. [27] have obtained from the experimental observations that the model describing viable cell growth changes at $t = 54$ h to reflect the transition from exponential growth to the decline phase. With this remark, the dynamical model of the viable and cell concentrations evolutions is as follows [27]:

$$\frac{dX}{dt} = \mu X - k_d X X_d, \quad \text{for } t < t_{\exp},$$

$$\frac{dX}{dt} = -k_d X X_d, \quad \text{for } t \geq t_{\exp}, \tag{6}$$

$$\frac{dX_d}{dt} = k_d X X_d,$$

where μ is the specific growth rate of the viable cells, k_d is a kinetic (decay) parameter, and $t_{\exp}$ is the time period of the exponential growth phase.

The most difficult modelling problem for the system of differential equations (5), (6) is related to the model of nonlinear reaction kinetics. Gao et al. [27] suggested that a generalized form of saturable kinetics (i.e., compound Monod kinetics) is suitable to describe the rate of each macroreaction from the reaction scheme given in Table 1. Rates for each of these macroreactions were expressed in the next compact form [4, 31]:

$$\varphi_i = \varphi_i^* \cdot X \cdot \prod_{S_j \in S_i} \frac{S_j}{K_{S_{j,i}} + S_j}, \quad i = \overline{1, 9}. \tag{7}$$

In the kinetic rates expression (7), φ_i is the reaction rate for reaction i, φ_i^* is the maximum reaction rate for reaction i, S_j is the concentration of substrate j within the set S_i of substrates for reaction i, and $K_{S_{j,i}}$ is a kinetic half-saturation constant for substrate j in reaction i. The specific rate expressions for each macroreaction are given in Table 2 [4, 27]. As Baughman et al. [4] noticed, the rate expressions for macroreactions 7 and 8 do not rigorously conform to the general format (7). More precisely, it was assumed that the principal rate-limiting substrate for both biomass and antibody synthesis is glutamine, and the kinetic contributions of any other substrates were thus omitted.

In conclusion, the full dynamical model of mAb production process is given by (5), where the kinetic rates are of the form presented in Table 3, together with the dynamical models (6) of viable and dead cell evolution in the batch reactor.

Table 2: Kinetics expressions for the macroreactions [4, 27].

Reaction number	Kinetic rate
1	$\varphi_1 = \varphi_1^* X \dfrac{S_1}{K_{S_{1,1}} + S_1}$
2	$\varphi_2 = \varphi_2^* X \dfrac{S_1}{K_{S_{1,2}} + S_1} \dfrac{S_3}{K_{S_{3,2}} + S_3}$
3	$\varphi_3 = \varphi_3^* X \dfrac{S_1}{K_{S_{1,3}} + S_1} \dfrac{S_3}{K_{S_{3,3}} + S_3}$
4	$\varphi_4 = \varphi_4^* X \dfrac{S_3}{K_{S_{3,4}} + S_3}$
5	$\varphi_5 = \varphi_5^* X \dfrac{S_4}{K_{S_{4,5}} + S_4}$
6	$\varphi_6 = \varphi_6^* X \dfrac{S_2}{K_{S_{2,6}} + S_2} \dfrac{S_5}{K_{S_{5,6}} + S_5}$
7	$\varphi_7 = \varphi_7^* X \dfrac{S_2}{K_{S_{2,7}} + S_2}$
8	$\varphi_8 = \varphi_8^* X \dfrac{S_2}{K_{S_{2,8}} + S_2}$
9	$\varphi_9 = \varphi_9^* X \dfrac{S_2}{K_{S9} + S_2}$

The state variables within the dynamical model (5)–(7) are associated with components of the macroreactions from the reaction scheme given in Table 1. While these components represent biological variables (concentrations of some substances or compounds), the kinetic parameters do not have always clear measurable physical representations.

The problem that remains to be solved now is related to the estimation of the unknown (inaccessible) kinetic parameters of the dynamical model (5), (6) of the mammalian cell culture. Therefore, it is necessary to estimate the experimentally inaccessible parameter values for the model that provide the best approximation to the measured culture concentrations data.

2.2. PSO-Based Technique Parameter Estimation

2.2.1. Problem Statement and Basic PSO Algorithms.
At the beginning of parameter estimation, the input and output data are known and the real system parameters are assumed as unknown. The identification problem is formulated in terms of an optimization problem in which the error between an actual physical measured response of the system and the simulated response of a parameterized model is minimized. The estimation of the system parameters is achieved as a result of minimizing the error function by the PSO algorithm.

Consider that the nonlinear system (2) that describes the dynamical behaviour of a class of bioprocesses is written as the following n-dimensional nonlinear system:

$$\frac{d\xi}{dt} = K \cdot \varphi(\xi) = f(\xi, t; \theta), \tag{8}$$

where $\xi \in \mathfrak{R}^n$ is the state vector (i.e., the vector of concentrations), $\theta \in \mathfrak{R}^q$ is the unknown parameters vector (i.e., the vector of unknown kinetic parameters), and f is a given nonlinear vector function.

TABLE 3: Experimental concentration measurements [4, 27].

Time	0 h	28 h	54 h	76 h	101 h	124 h	147 h
GLC [mM]	3.59 ± 0.04	2.59 ± 0.09	1.88 ± 0.15	1.77 ± 0.05	1.70 ± 0.03	1.67 ± 0.04	1.68 ± 0.05
GLN [mM]	2.85 ± 0.04	1.27 ± 0.30	0.42 ± 0.31	0.11 ± 0.10	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
ASN [mM]	0.46 ± 0.00	0.39 ± 0.02	0.35 ± 0.02	0.32 ± 0.03	0.28 ± 0.02	0.25 ± 0.02	0.22 ± 0.02
ASP [mM]	0.27 ± 0.02	0.18 ± 0.02	0.10 ± 0.04	0.07 ± 0.04	0.04 ± 0.04	0.03 ± 0.04	0.03 ± 0.04
LAC [mM]	0.51 ± 0.01	1.33 ± 0.06	1.66 ± 0.19	1.74 ± 0.01	1.71 ± 0.01	1.72 ± 0.01	1.73 ± 0.05
ALA [mM]	0.33 ± 0.03	0.72 ± 0.11	1.15 ± 0.17	1.32 ± 0.12	1.46 ± 0.07	1.48 ± 0.07	1.51 ± 0.08
PRO [mM]	0.30 ± 0.01	0.27 ± 0.02	0.42 ± 0.04	0.53 ± 0.02	0.56 ± 0.02	0.60 ± 0.01	0.60 ± 0.01
MAB [10^{-4} mM]	0.34 ± 0.12	1.02 ± 0.06	1.58 ± 0.16	2.31 ± 0.24	2.66 ± 0.41	3.09 ± 0.60	3.41 ± 0.75
BM [mM]	2.01 ± 0.20	11.61 ± 0.46	16.51 ± 0.85	17.98 ± 0.84	19.41 ± 2.21	18.67 ± 2.49	17.97 ± 1.33
X [10^6 cells/mL]	0.09 ± 0.01	0.58 ± 0.02	0.79 ± 0.05	0.72 ± 0.01	0.47 ± 0.06	0.17 ± 0.03	0.06 ± 0.02
X_d [10^6 cells/mL]	0.02 ± 0.01	0.05 ± 0.01	0.11 ± 0.01	0.25 ± 0.05	0.58 ± 0.05	0.85 ± 0.12	0.91 ± 0.07

To estimate the unknown parameters in (8), a parameter identification system is defined as follows:

$$\frac{d\widehat{\xi}(t)}{dt} = f\left(\widehat{\xi}, t; \widehat{\theta}\right), \tag{9}$$

where $\widehat{\xi} \in \mathfrak{R}^n$ is the estimated state vector and $\widehat{\theta} \in \mathfrak{R}^q$ is the estimated parameters vector.

The objective function defined as the mean squared errors between real and estimated responses for a number N of given samples is considered as fitness of estimated model parameters [14]:

$$W = \frac{1}{N + M} \sum_{j=1}^{M} \sum_{k=1}^{N} \left(\xi_j^k - \widehat{\xi}_j^k\right)^2, \tag{10}$$

where M is the number of measurable states and N is the data length used for parameter identification, whereas ξ_j^k and $\widehat{\xi}_j^k$ are the real and estimated values of state j at time k, respectively.

This objective function is a function difficult to minimize because there are many local minima and the global minimum has a very narrow domain of attraction. Our goal is to determine the system parameters, using particle swarm optimization algorithms in such a way that the value of W is minimized, approaching zero as much as possible.

Mathematical description of basic PSO and some important variants is presented in the following.

Candidate solutions of a population called particles coexist and evolve simultaneously based on knowledge sharing with neighbouring particles. Each particle represents a potential solution to the optimization problem and it has a fitness value decided by optimal function. Supposing search space is M-dimensional, each individual is treated as a particle in the M-dimensional search space. The position and rate of position change for ith particle can be represented by M-dimensional vector, $x_i = (x_{i1}, x_{i2}, \ldots, x_{iM})$ and $v_i = (v_{i1}, v_{i2}, \ldots, v_{iM})$, respectively. The best position previously visited by the ith particle is recorded and represented by $p_i = (p_{i1}, p_{i2}, \ldots, p_{iM})$, called *pbest*. The swarm best position previously visited by all the particles in the population is represented by $p_g = (p_{g1}, p_{g2}, \ldots, p_{gM})$, called *gbest*.

Then particles search their best position, which are guided by swarm information p_g and their own information p_i. Each particle modifies its velocity to find a better solution (position) by applying its own flying experience (i.e., memory of the best position found in earlier flights) and the experience of neighbouring particles (i.e., the best solution found by the population). Each particle position is evaluated by using fitness function and updates its position and velocity according to the following equations:

$$v_i^{k+1} = \omega \cdot v_i^k + c_1 r_1 \left(pbest_i^k - x_i^t\right) + c_2 r_2 \left(gbest_i^k - x_i^k\right),$$
$$x_i^{k+1} = x_i^k + v_i^{k+1}, \tag{11}$$

where k is iteration number, ω is inertia weight, c_1 and c_2 are two acceleration coefficients regulating the relative velocity toward local and global best position, and r_1 and r_2 are two random numbers from the interval $[0, 1]$. Many effects have been made over the last decade to determinate the inertia weight. Various studies have shown that under certain conditions convergence is guaranteed to a stable equilibrium point [51]. These conditions include $\omega > (c_1 + c_2)/2 - 1$ and $0 < \omega < 1$. The technique originally proposed was to bound velocities so that each component of v_i is kept within the range $[V_{\min}, V_{\max}]$.

Unfortunately, this simple form of PSO suffers from the premature convergence problem, which is particularly true in complex problems since the interacted information among particles in PSO is too simple to encourage a global search. Many efforts have been made to avoid the premature convergence. One solution is the use of a constriction factor to insure convergence of the PSO, introduced in [45]. Thus, the expression for velocity has been modified as

$$v_i^{k+1} = h \cdot \left[v_i^k + c_1 r_1 \left(pbest_i^k - x_i^t\right) + c_2 r_2 \left(gbest_i^k - x_i^k\right)\right],$$
$$x_i^{k+1} = x_i^k + v_i^{k+1}, \tag{12}$$

where h represents the constriction factor and is defined as

$$h = \frac{2}{\left(\left|2 - \alpha - \sqrt{\alpha^2 - 4\alpha}\right|\right)}, \quad \alpha = c_1 + c_2 > 4. \tag{13}$$

In this variant of the PSO algorithm, h controls the magnitude of the particle velocity and can be seen as a dampening factor. It provides the algorithm with two important features [52]. First, it usually leads to faster convergence than standard PSO. Second, the swarm maintains the ability to perform wide movements in the search space, even if convergence is already advanced but a new optimum is found. Therefore, the constriction PSO has the potential to avoid being trapped in local optima while possessing a fast convergence. It was shown to have superior performance compared to a standard PSO [53].

It is shown that a larger inertia weight tends to facilitate the global exploration and a smaller inertia weight achieves the local exploration to fine-tune the current search area. The best performance could be obtained by initially setting ω to some relatively high value (e.g., 0.9), which corresponds to a system where particles perform extensive exploration, and gradually reducing ω to a much lower value (e.g., 0.4), where the system would be more dissipative and exploitative and would be better at homing into local optima. In [54], a linearly decreased inertia weight ω over time is proposed, where ω is given by the following equation:

$$\omega = \left(\omega_i - \omega_f \right) \cdot \frac{\left(k_{\max} - k \right)}{k_{\max}} + \omega_f, \tag{14}$$

where ω_i, ω_f are starting and final values of inertia weight, respectively, $k_{\max}$ is the maximum number of the iteration, and k is the current iteration number. It is generally taken that starting value is $\omega_i = 0.9$ and final value is $\omega_f = 0.4$.

On the other hand, in [35] PSO was introduced with time-varying acceleration coefficients. The improvement has the same motivation and similar techniques as the adaptation of inertia weight. In this case, the cognitive coefficient c_1 is decreased linearly and the social coefficient c_2 is increased linearly over time as follows:

$$c_1 = \left(c_{1f} - c_{1i} \right) \cdot \frac{\left(k_{\max} - k \right)}{k_{\max}} + c_{1i},$$
$$c_2 = \left(c_{2f} - c_{2i} \right) \cdot \frac{\left(k_{\max} - k \right)}{k_{\max}} + c_{2i}, \tag{15}$$

where c_{1i} and c_{2i} are the initial values of the acceleration coefficients c_1 and c_2 and c_{1f} and c_{2f} are the final values of the acceleration coefficients c_1 and c_2, respectively. Usually, $c_{1i} = 2.5$; $c_{2i} = 0.5$; $c_{1f} = 0.5$; and $c_{2f} = 2.5$ [14, 35, 36].

The dynamical model of mAb production process given by the relations (5), (6), associated with the expressions of the kinetic rates presented in Table 3, contains a number of 23 kinetic parameters (maximum reaction rates and kinetic half-saturation constants). In order to estimate these unknown (inaccessible) kinetic parameters of the complex dynamical model of mammalian cell culture, the measured concentrations are used and a PSO-based algorithm is implemented. The goal is to obtain a model that approximates as well as possible the behaviour of the process (expressed by means of the experimentally obtained data). The model of mAb production process under investigation is in fact based on

several macroreactions; therefore it results in the fact that the kinetic parameters do not have always clear measurable physical representations. Thus, an optimization-based estimation technique is suitable for this set of kinetic parameters.

2.2.2. Implementation of PSO-Based Technique. In the following, a multistep PSO-based version that uses time-varying acceleration coefficients is implemented and an optimal set of kinetic parameter values of the mAb production process is obtained.

In order to implement the PSO-based technique, the model of mAb production process (5), (6) obtained from the macroreactions schemes is used, translated into the generic parameter identification system represented in (9).

The experimental concentration values for all the involved extracellular metabolites are provided in the work of Gao et al. [27]. The batch cultures of the organism were allowed to grow for 147 h, with 54 h the exponential growth phase and 93 h the postexponential (decline) phase. The infrequent measured concentration data for the metabolites were obtained from the collected samples via proper techniques. The set of concentrations measurements are given in Table 3 [4, 27]. Each data point is the average of measurements taken from three independent experiments, with standard deviation [4, 27].

To facilitate the application of the proposed PSO-based parameter estimation strategy, the time derivatives of the states from model (5)–(7) must be reconstructed. Because the measured data are very few (only 7 experimental measurements for each parameter, see Table 3), an interpolation method is necessary to find intermediate values of the states, which are actually the biological parameters of the process, that is, the concentrations of metabolites. Such situations with a small number of experimental measurements are typical for many bioprocesses. Ideally speaking, the online measurements (in each sampling moment, at every 6 min., e.g.) for each concentration are necessary. However, these online measurements are achieved with expensive instrumentation, or there are no such fast sensors for some concentrations. Thus, the infrequent offline measurements are preferred. To facilitate the achievement of an accurate estimation of model parameters of mAb process, we need the interpolation of these measured data, which allows us to obtain the unavailable data between adjacent measurement points (i.e., to estimate the unavailable data needed to calculate model predictions between these measurement points).

Remark 2. From mathematical point of view, a discussion about the interpolation technique can be done. Many authors use the linear interpolation, with advantages related to rapidity and simple implementation [4]. However, the linear interpolation is not very precise. Another disadvantage is that the interpolant is not differentiable at the points where the value of the function is known. Therefore, we propose a cubic interpolation method that is the simplest method that offers true continuity between the measured data. A cubic Hermite spline or cubic Hermite interpolator is a spline where each piece is a third-degree polynomial specified in Hermite form, that is, by its values and first derivatives at the end
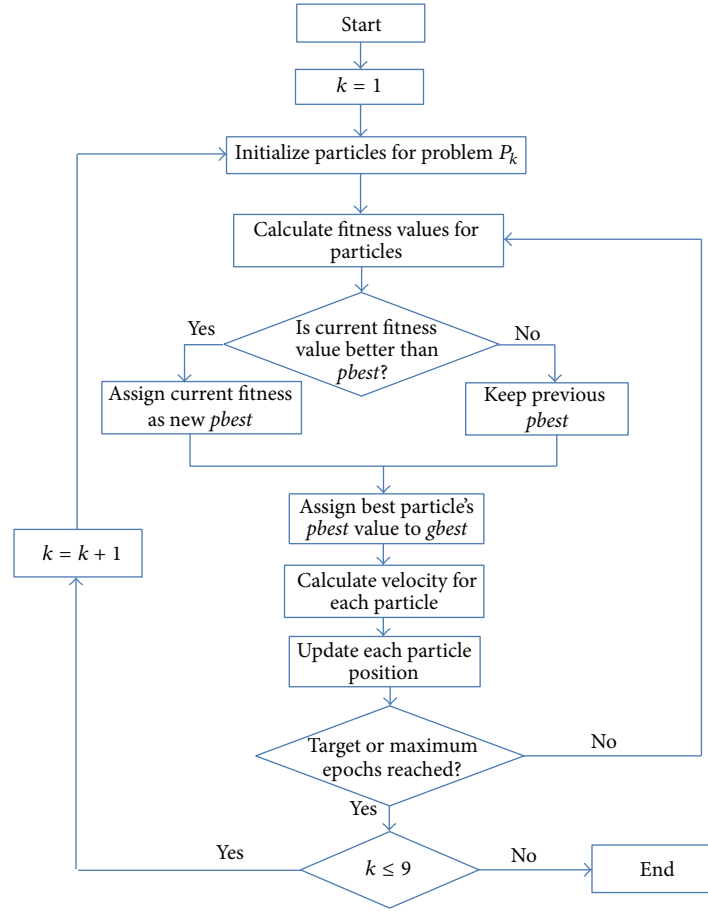
FIGURE 1: The flowchart of the multistep PSO algorithm.

points of the corresponding domain interval. Cubic Hermite splines are typically used for interpolation of numeric data specified at given argument values $t_1, t_2, \ldots, t_n$, to obtain a smooth continuous function. The Hermite formula is applied to each interval (t_k, t_{k+1}) separately. The resulting spline will be continuous and will have continuous first derivative.

The time derivatives of the states are approximated using forward differences:

$$\frac{d\xi(t_k)}{dt} \approx \frac{\xi(t_k) - \xi(t_k + T_s)}{T_s}, \qquad (16)$$

where T_s represents the sampling period. In this approximation, the error is proportional with the sampling interval (a smaller sampling period will give a smaller approximation error).

Because a 23-dimensional optimization problem that must be solved for simultaneously estimation of all unknown parameters requires great computational resources, a multistep approach was used. So, the problem was split in nine simpler problems that are solved sequentially until all 23 parameters are found. These problems are noted with $P1, P2, \ldots, P9$ and the corresponding resulted parameters are presented in Table 4. A flowchart of the multistep PSO algorithm is given in Figure 1.

TABLE 4: The subproblems solved by using the multistep PSO-based approach.

Subproblem	Parameters
P1	$\varphi_7^*, K_{S_{2,7}}$
P2	$\varphi_8^*, K_{S_{2,8}}$
P3	$\varphi_4^*, K_{S_{3,4}}$
P4	$\varphi_2^*, K_{S_{1,2}}, K_{S_{3,2}}$
P5	$\varphi_6^*, K_{S_{2,6}}, K_{S_{2,9}}, \varphi_9^*, K_{S_{5,6}}$
P6	$\varphi_5^*, K_{S_{4,5}}$
P7	$\varphi_3^*, K_{S_{1,3}}, K_{S_{3,3}}$
P8	$\varphi_1^*, K_{S_{1,1}}$
P9	μ, k_d

For example, the problem $P1$ corresponds to the 10th equation from system (5)-(6) (that represents time evolution of the biomass) and only two parameters must be estimated in this case: φ_7^* and $K_{S_{2,7}}$. The PSO algorithm is used to minimize the sum of the square errors between measured and estimated data:

$$W_{P1} = \sum_{k=1}^{N} \left(\xi_{10}(k \cdot T_s) - \widehat{\xi}_{10}(k \cdot T_s) \right)^2$$

Table 5: Kinetic parameter estimates obtained via the multistep PSO approach.

Kinetic parameters	Estimated values
φ_1^* [pmol/(cell h)]	8.443×10^{-4}
φ_2^* [pmol/(cell h)]	2.481×10^{6}
φ_3^* [pmol/(cell h)]	3.968×10^{5}
φ_4^* [pmol/(cell h)]	1.090×10^{2}
φ_5^* [pmol/(cell h)]	7.283
φ_6^* [pmol/(cell h)]	3.337×10^{5}
φ_7^* [pmol/(cell h)]	3.977×10^{3}
φ_8^* [pmol/(cell h)]	6.697×10^{-6}
φ_9^* [pmol/(cell h)]	3.261×10^{4}
$K_{S_{1,1}}$ [mM]	8.989×10^{5}
$K_{S_{1,2}}$ [mM]	6.495×10^{4}
$K_{S_{1,3}}$ [mM]	3.723×10^{4}
$K_{S_{3,2}}$ [mM]	7.076×10^{2}
$K_{S_{3,3}}$ [mM]	2.782×10^{3}
$K_{S_{3,4}}$ [mM]	0.019
$K_{S_{2,6}}$ [mM]	2.719×10^{4}
$K_{S_{2,7}}$ [mM]	9.324×10^{3}
$K_{S_{2,8}}$ [mM]	0.537
$K_{S_{2,9}}$ [mM]	6.683×10^{5}
$K_{S_{4,5}}$ [mM]	1.920×10^{3}
$K_{S_{5,6}}$ [mM]	4.488×10^{4}
μ	0.043
k_d	0.067

Figure 2: Simulation results, profile of the biomass concentration.

$$\text{s.t.} \quad \widehat{\xi}_{10}\left((k+1)\cdot T_s\right)$$

$$= \widehat{\xi}_{10}\left(k\cdot T_s\right) + T_s \cdot \frac{\varphi_7^* \cdot \xi_2\left(k\cdot T_s\right)\cdot \xi_{12}\left(k\cdot T_s\right)}{K_{S_{2,7}} + \xi_2\left(k\cdot T_s\right)}.$$

(17)

3. Results and Discussion

3.1. Optimal Set of Kinetic Parameters. The optimization problem formulated in the previous section is nonlinear and nonconvex with many local minima. The estimated parameters of one subproblem are then considered known in the subsequent equations. In order to be clear, the already estimated parameters are not updated between solutions of subproblems. For example, in the frame of problem $P1$ two parameters are estimated: φ_7^* and $K_{S_{2,7}}$. These parameters will be available in the next problem ($P2$), and so on. In this study, all the computations were achieved with a sampling period $T_s = 6$ min (0.1 h). As example, for problem $P1$ a number of 150 particles randomly initialized were used. The algorithm stops if the square error is smaller then $10e-6$ or after 300 iterations. The optimal set of kinetic parameter values of the mAb production process obtained via this multistep PSO-based approach is given in Table 5.

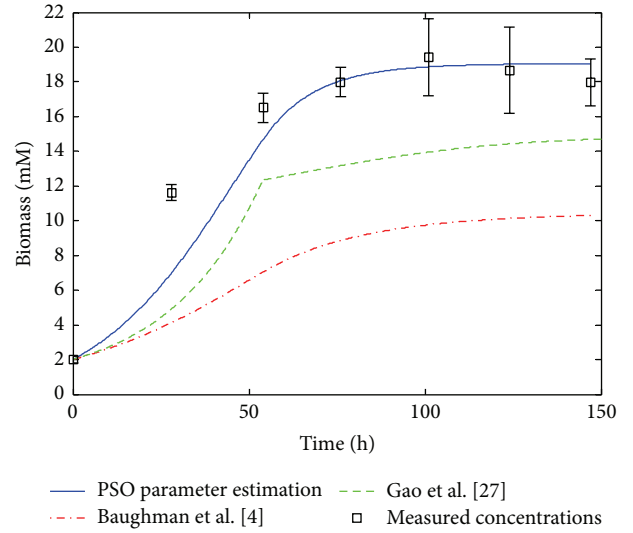The partition of the multidimensional optimization problem proposed within our PSO algorithm not only ensures the decrease of necessary computational resources by comparison with Gao et al. [27] and Baughman et al. [4] approaches but additionally offers a solution for the reported problems concerning the stiffness of estimated parameter set. More precisely, as Baughman et al. [4] noticed, there are some concentrations such as the mAb concentration that are several orders of magnitude less than other metabolites. This fact leads to stiffness problems in the optimization procedure, which are partially solved in [4] by using an alternative error objective for the mAb concentration. In our approach, the partition in simpler PSO problems solves uniformly this issue, using the same error objective for the entire set of parameters.

Another important problem approached and solved by using the proposed PSO method is related to the expressions of reaction kinetics. To simplify the optimization problem, Gao et al. [27] considered that the values of half-saturation constants are sufficiently small and consequently the kinetic rates given in Table 2 can be simplified such that the relation (7) becomes $\varphi_i = \varphi_i^* \cdot X, i = \overline{1, 9}$. This simplification eliminates the necessity of half-saturation constants estimation, and only the maximum reaction rates need to be estimated. However, as was mentioned in [4], half-saturation constants can be significantly smaller than the corresponding substrate concentration in processes controlled by a single enzyme (e.g., glucose transport), but this fact is not necessarily true for the macroreactions in mAb production processes. Therefore, Gao et al. assumption is unwarranted and it can affect the reliability of the model. This is one of the reasons because the proposed PSO method yield good fitting results compared with Gao et al. [27], as it can be seen in Figures 2–5.

Remark 3. There are necessary some comments concerning the overfitting problems. Overfitting arises when a statistical model describes noise instead of the underlying relationship; it usually occurs when a model is very complex, such as
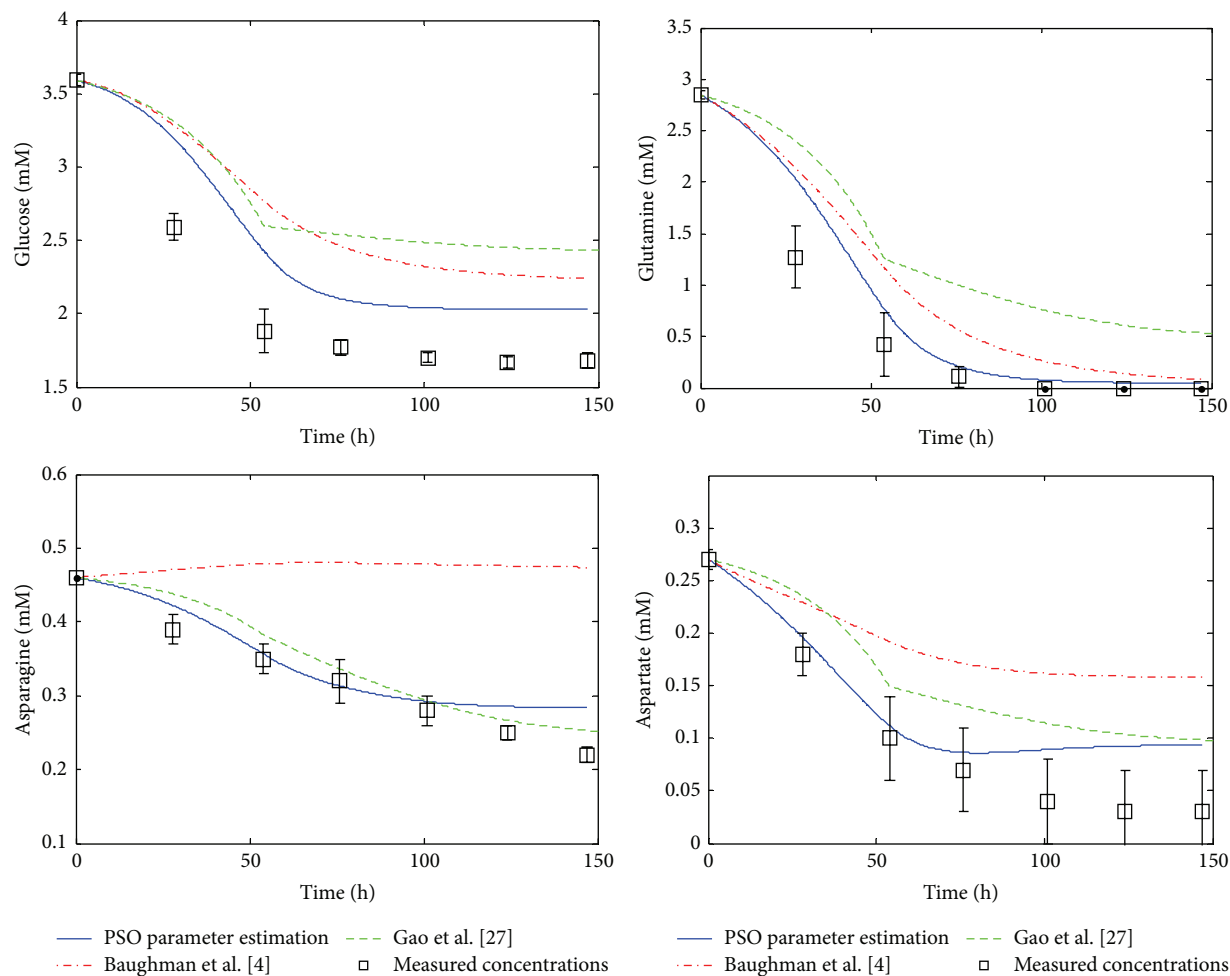
FIGURE 3: Simulation results, profiles of substrate concentrations: GLC, GLN, ASN, and ASP.

having many parameters relative to the number of observations. Even though the approached mAb production model is quite complex, it is not a statistical model. Also, even if the number of measured concentration samples is relatively small by comparison with the number of parameters, the PSO technique is an optimization procedure, which is much less sensitive to overfitting than the methods that are based on model training, such as neural network techniques (see Tetko et al. [55]). The potential for overfitting depends not only on the number of parameters and measured data but also on the conformability of the model structure with the data shape.

Some comparisons of the proposed PSO approach with other PSO applications to bioprocesses can be done. Most of the reported works were focused on the process of glycerol fermentation by *Klebsiella pneumoniae* in batch, fed-batch, and continuous cultures [37–41]. Shen et al. [37] studied a mathematical model of *Klebsiella pneumoniae* in a continuous culture. An eight-dimensional nonlinear dynamical system was obtained, and a parallel PSO technique was implemented in order to identify 19 parameters. The identification results are compared only with experimental steady-state values. The reported mean relative errors between the computational values and the experimental data are quite large (between 8% and 13%). A similar model of bio-dissimilation of glycerol by *K. pneumoniae* in a continuous culture was widely analyzed by Zhai et al. in [38]. Here a parallel PSO pathways identification algorithm was constructed to find the optimal pathway and 21 parameters under various conditions. The combined estimation of pathways and process parameters leads to a vast identification model, solved on a cluster server with 16 nodes (each node with 4 Core, 64-bit, 2.5 GHz processor), in over 130 hours. Comparable PSO approaches were used in [39, 41] in the case of the same fermentation process, but in a batch culture. For example, Yuan et al. [39] used a parallel migration PSO algorithm to estimate pathways and 12 parameters of the eight-dimensional model. The identification problem was split into two subproblems (one for pathways and one for parameters) and solved on the above cluster in about 18 hours. Another work addressed the PSO identification in the case of the fermentation of glycerol by *K. pneumoniae* in a fed-batch culture [40]. A nonlinear hybrid system was developed (with seven differential equations plus a switching mechanism and 8 parameters).
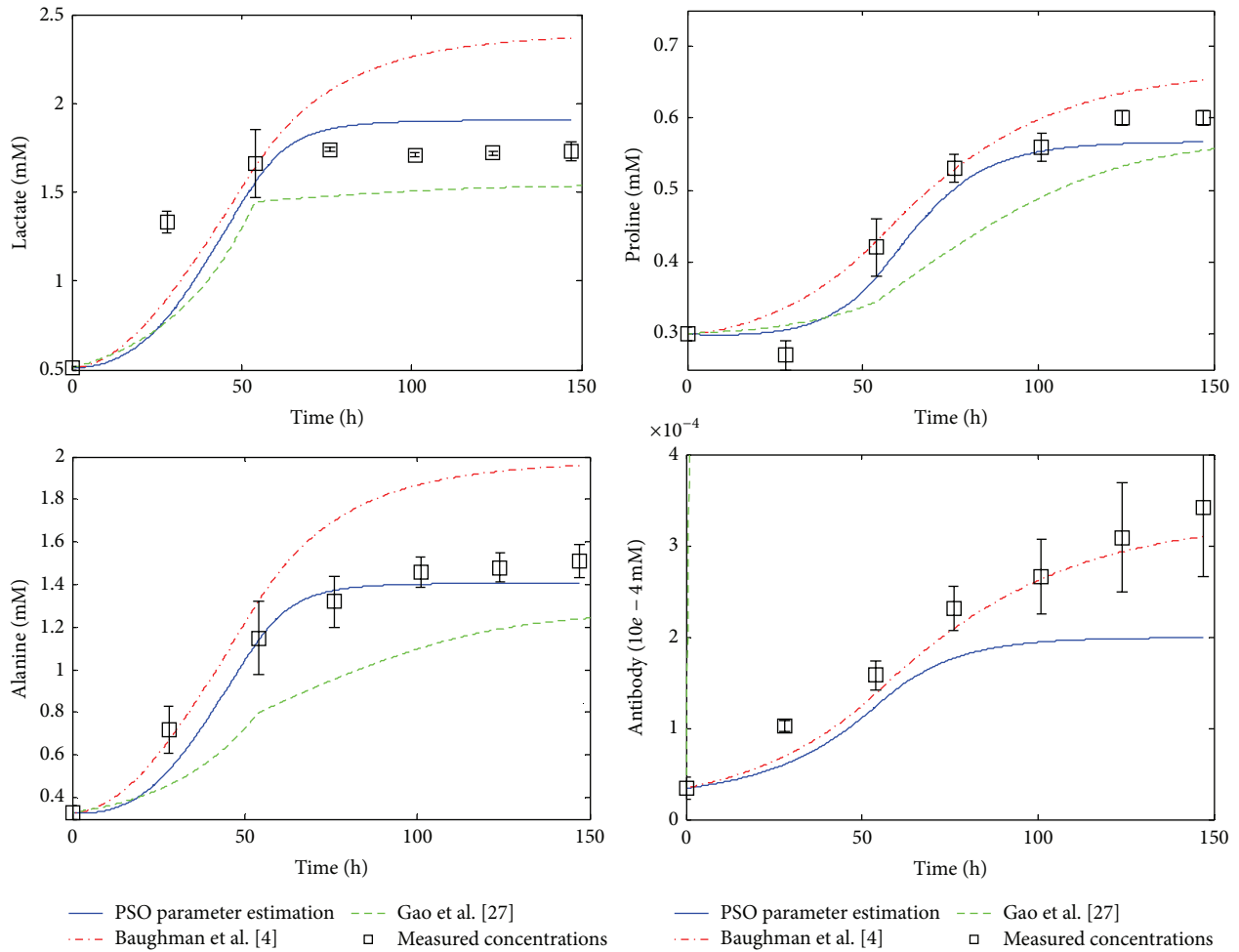
FIGURE 4: Simulation results, profiles of concentrations: LAC, PRO, ALA, and MAb.

The proposed technique was an asynchronous parallel PSO, and the reported averaged computational time was about 3.26 h, on the above-mentioned cluster server with 16 nodes. The reported results in [38–40] are very good, even the accuracy of the algorithms cannot be fully assessed (statistical reports were not provided). However, the computational effort is considerable, given the fact that simultaneous pathways and parameters identification was approached.

The particle swarm-based multistep nonlinear optimization algorithm proposed in the present work was used for the estimation of 23 parameters of an eleven-dimensional nonlinear system (the pathways identification was not considered). By using the multistep approach, the computational effort was quite small (about 20 min. on a computer with Intel Core i5, 64-bit, 3.3 GHz processor). The obtained results and the statistical analysis show a good accuracy of the identification results.

3.2. Simulation Results. The performance of the proposed estimation technique was analyzed by using numerical simulations. All these simulations are achieved by using the development, programming, and simulation environment

MATLAB (registered trademark of The MathWorks, Inc., USA). For comparison, the simulated profiles based on the kinetic parameters obtained via PSO technique (Table 5) are represented together with the original system measurements [27] and with the profiles obtained by Gao et al. [27] and Baughman et al. [4], respectively. The concentration profiles based on the results of Gao et al. and on the results of Baughman et al. [4] approach, respectively, are simulated and plotted using the kinetic parameter values given in Table 6 [4, 27].

The simulated concentration profiles are presented in Figures 2–5. First, in Figure 2 the time evolution of the biomass concentration is depicted. As can be seen, the best matching with the measured data (the values of measured data plotted in all figures are taken from Table 3) is given by the PSO approach. Figure 3 presents the simulated concentration profiles of glucose, glutamine, asparagine, and aspartate. In all cases, the PSO proposed technique ensures the best estimates (in the case of asparagine, the results of Gao et al. [27] are comparable with those obtained using PSO). In Figure 4, the concentration profiles of lactate, proline, alanine, and monoclonal antibody are plotted. The best matching with
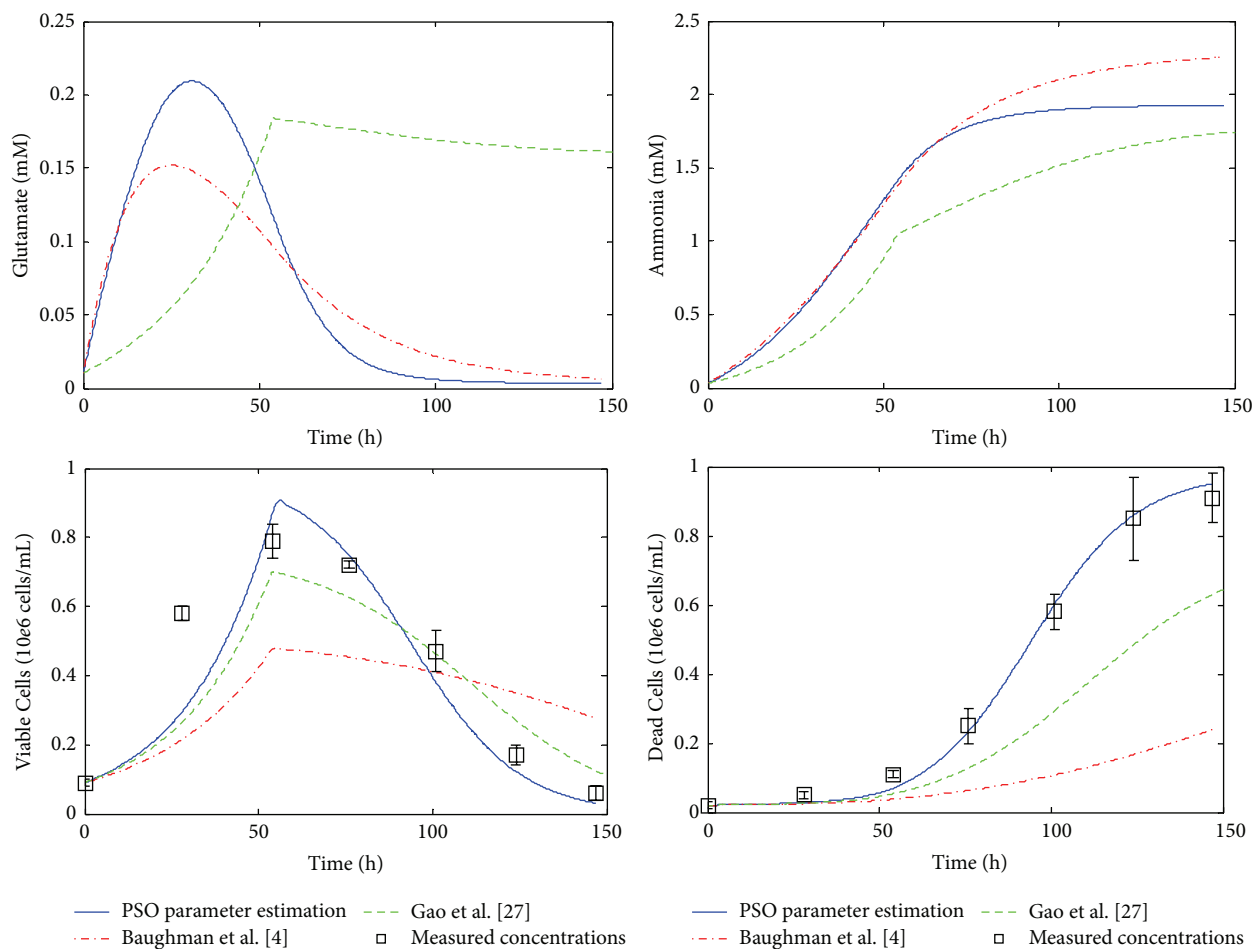
FIGURE 5: Simulation results, profiles of concentrations: GLU, NH$_3$, X, and Xd.

the measured data is obtained with the PSO estimation technique for lactate, proline, and alanine. In the case of antibody, the best results were obtained by Baughman et al. [4]. Finally, Figure 5 shows the concentration evolutions of glutamate, ammonia, viable, and dead cells, respectively. The glutamate and ammonia concentrations are not measured; the time profiles of these variables were obtained from the dynamic model simulation. The viable cells and dead cells evolution obtained via PSO estimation match very well the measured data.

Since in industrial practice the measured data are affected by various disturbances, one explored the extent to which noisy measurements affects the estimated parameter values. For this reason, a Monte Carlo simulation approach was used. First, normal (Gaussian) distributions were constructed for every measured data set in Table 3, subject to the known mean and standard deviation of each point.

A set of 150 simulated measurement sets were generated. Finally, using each randomized data set, a new cubic interpolation of the data was generated for our standard condition and the parameter estimation problems ($P1 - P9$) were solved for each case.

Outlying solutions were identified and excluded using a basic quartile classification method. The quartile values are chosen in the following manner. First, use the median to divide the ordered data set into two halves. The median is not included into the halves. Then, the lower quartile value is the median of the lower half of the data (q_{low}) and the upper quartile value is the median of the upper half of the data (q_{up}). Namely, the lower and upper quartile (q_{low}, q_{up}) positions were found for the set of 150 objective values and the interquartile range $q_{up} - q_{low}$ was calculated. Solutions with objective values lying outside the interval [$q_{low} - 1.5(q_{up} - q_{low}), q_{up} + 1.5(q_{up} - q_{low})$] were considered to be outlying cases. Using the remaining solutions, mean values and associated standard deviations were calculated for each estimated parameter.

The standard deviations were then converted to percentages of their associated mean value. These means and standard deviations are listed in Table 7.

Certain parameter estimates are much more susceptible to variability induced through perturbations in measured data than are others. It can be seen that certain parameters ($\mu, k_d, K_{S_{1,3}}$) are less sensitive to noisy measurements than are certain others ($\varphi_7^*, K_{S_{2,6}}, K_{S_{2,8}}, K_{S_{5,6}}$).

TABLE 6: Kinetic parameter estimates obtained by Gao et al. [27] and by Baughman et al. [4], respectively.

Kinetic parameters	Values (Gao et al. [27]) (exponential phase)	Values (Gao et al. [27]) (decline phase)	Values (Baughman et al. [4])
φ_1^* [pmol/(cell h)]	0.008	−0.0033	8.85×10^{-4}
φ_2^* [pmol/(cell h)]	0.0191	0.0058	1.12×10^6
φ_3^* [pmol/(cell h)]	0.0023	−0.0014	1.1×10^5
φ_4^* [pmol/(cell h)]	0.0081	0.0057	1.97×10^{-2}
φ_5^* [pmol/(cell h)]	−0.01	0.0056	4.95
φ_6^* [pmol/(cell h)]	−0.011	0.0029	1.34×10^5
φ_7^* [pmol/(cell h)]	0.6429	0.0573	1.36×10^3
φ_8^* [pmol/(cell h)]	0.0046	0.0077	1×10^{-5}
φ_9^* [pmol/(cell h)]	0.0731	0.0113	1.83×10^4
$K_{S_{1,1}}$ [mM]	0.01	0.01	1.63×10^5
$K_{S_{1,2}}$ [mM]	0.01	0.01	1.08×10^4
$K_{S_{1,3}}$ [mM]	0.01	0.01	1.44×10^4
$K_{S_{3,2}}$ [mM]	0.001	0.001	9.64×10^2
$K_{S_{3,3}}$ [mM]	0.001	0.001	1.04×10^3
$K_{S_{3,4}}$ [mM]	0.001	0.001	5.42×10^{-2}
$K_{S_{2,6}}$ [mM]	0.01	0.01	3.03×10^3
$K_{S_{2,7}}$ [mM]	0.01	0.01	6.39×10^3
$K_{S_{2,8}}$ [mM]	0.01	0.01	3.73×10^{-1}
$K_{S_{2,9}}$ [mM]	0.01	0.01	3.23×10^5
$K_{S_{4,5}}$ [mM]	0.001	0.001	7.42×10^3
$K_{S_{5,6}}$ [mM]	0.001	0.001	4.45×10^3
μ	0.0399	0.0399	3.22×10^{-2}
k_d	0.06	0.06	4.99×10^{-2}

The proposed modelling and parameter estimation method can be applied to cellular processes described by the general form (1), and it is not yet applicable to all classes of bioprocesses. To be more specific, it is hard to be applied to processes characterized by phenomena such as propagation reactions, transport processes, latency and short intercellular phases (in epidemics), and spread (propagation) of infections, that is, processes with large heterogeneity and delays. A typical class of such nonlinear delay biosystems is represented by the dynamics models describing cell-to-cell spread mechanisms, encountered, for example, in HIV infections [56, 57].

4. Conclusions

In order to develop accurate models for mammalian cell culture processes and to overcome some of the specific problems of mAb production processes such as the nonlinearity, the absence of instrumentation, and the kinetics uncertainties, a multistep nonlinear particle swarm optimization-based technique for the estimation of experimentally unavailable kinetic parameters was designed and implemented. The proposed approach was tested by using a particular dynamical model of mammalian cell culture, as a case study, but is generic for this class of bioprocesses. We have established the capability of proposed technique to identify model parameters that provide an accurate simulation of experimentally observed mAb production process behaviour. The performed statistical analysis demonstrates that the proposed estimation method is robust against normal distributed noisy measurements. The simulations showed that the PSO parameter estimation technique provides more accurate results than those reported in previous studies.

The obtained dynamical model of the mAb production process is accurate and can contribute to the development of model-based applications, which lead to high productivity and better quality products. The performed simulations represent one of the possibilities of model validation. The results show that the proposed model offers good predictions not only of the cell culture, for instance predictions of concentrations of energy sources such as glutamine and glucose, but also of the main amino acids and products. The proposed estimation approach can be also applied to other bioprocesses belonging to the nonlinear class considered in the present study.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Table 7: Monte Carlo parameter means and standard deviation over complete measurement set.

Kinetic parameters	Mean value	Standard deviation
φ_1^* [pmol/(cell h)]	8.127×10^{-4}	44
φ_2^* [pmol/(cell h)]	2.868×10^{6}	51
φ_3^* [pmol/(cell h)]	2.441×10^{5}	63
φ_4^* [pmol/(cell h)]	1.330×10^{2}	71
φ_5^* [pmol/(cell h)]	7.873	23
φ_6^* [pmol/(cell h)]	3.923×10^{5}	35
φ_7^* [pmol/(cell h)]	3.157×10^{3}	93
φ_8^* [pmol/(cell h)]	4.067×10^{-6}	24
φ_9^* [pmol/(cell h)]	2.991×10^{4}	36
$K_{S_{1,1}}$ [mM]	8.936×10^{5}	17
$K_{S_{1,2}}$ [mM]	7.030×10^{4}	29
$K_{S_{1,3}}$ [mM]	3.684×10^{4}	5
$K_{S_{3,2}}$ [mM]	7.31×10^{2}	29
$K_{S_{3,3}}$ [mM]	2.565×10^{3}	14
$K_{S_{3,4}}$ [mM]	0.026	13
$K_{S_{2,6}}$ [mM]	2.295×10^{4}	87
$K_{S_{2,7}}$ [mM]	9.273×10^{3}	51
$K_{S_{2,8}}$ [mM]	0.649	116
$K_{S_{2,9}}$ [mM]	6.224×10^{5}	39
$K_{S_{4,5}}$ [mM]	1.814×10^{3}	44
$K_{S_{5,6}}$ [mM]	4.657×10^{4}	75
μ	0.048	2
k_d	0.070	0.73

Acknowledgments

References

[1] D. F. Slezak, C. Suárez, G. A. Cecchi, G. Marshall, and G. Stolovitzky, "When the optimal is not the best: parameter estimation in complex biological models," *PLoS ONE*, vol. 5, no. 10, Article ID e13283, 2010.

[2] C. Kontoravdi, S. P. Asprey, E. N. Pistikopoulos, and A. Mantalaris, "Development of a dynamic model of monoclonal antibody production and glycosylation for product quality monitoring," *Computers & Chemical Engineering*, vol. 31, no. 5-6, pp. 392–400, 2007.

[3] C. Kontoravdi, E. N. Pistikopoulos, and A. Mantalaris, "Systematic development of predictive mathematical models for animal cell cultures," *Computers & Chemical Engineering*, vol. 34, no. 8, pp. 1192–1198, 2010.

[4] A. C. Baughman, X. Huang, S. T. Sharfstein, and L. L. Martin, "On the dynamic modeling of mammalian cell metabolism and mAb production," *Computers & Chemical Engineering*, vol. 34, no. 2, pp. 210–222, 2010.

[5] C. M. Batista, L. C. S. Medeiros, I. Eger, and M. J. Soares, "MAb CZP-315.D9: an antirecombinant cruzipain monoclonal antibody that specifically labels the reservosomes of *Trypanosoma cruzi* epimastigotes," *BioMed Research International*, vol. 2014, Article ID 714749, 9 pages, 2014.

[6] G. Sautto, N. Mancini, G. Gorini, M. Clementi, and R. Burioni, "Possible future monoclonal antibody (mAb)-Based Therapy against arbovirus infections," *BioMed Research International*, vol. 2013, Article ID 838491, 21 pages, 2013.

[7] F. Li, N. Vijayasankaran, A. Y. Shen, R. Kiss, and A. Amanullah, "Cell culture processes for monoclonal antibody production," *mAbs*, vol. 2, no. 5, pp. 466–479, 2010.

[8] R. J. Pantazes and C. D. Maranas, "MAPs: a database of modular antibody parts for predicting tertiary structures and designing affinity matured antibodies," *BMC Bioinformatics*, vol. 14, article 168, 2013.

[9] A. K. Pavlou and M. J. Belsey, "The therapeutic antibodies market to 2008," *European Journal of Pharmaceutics and Biopharmaceutics*, vol. 59, no. 3, pp. 389–396, 2005.

[10] S. Dhir, K. J. Morrow, R. R. Rhinehart, and T. Wiesner, "Dynamic optimization of hybridoma growth in a fed-batch bioreactor," *Biotechnology and Bioengineering*, vol. 67, no. 2, pp. 197–205, 2000.

[11] C. Giersch, "Mathematical modelling of metabolism," *Current Opinion in Plant Biology*, vol. 3, no. 3, pp. 249–253, 2000.

[12] F. R. Sidoli, A. Mantalaris, and S. P. Asprey, "Modelling of mammalian cells and cell culture processes," *Cytotechnology*, vol. 44, no. 1-2, pp. 27–46, 2004.

[13] M. Roman, D. Selisteanu, E. Bobasu, and D. Sendrescu, "Modeling of culture cells for pharmaceutical industry applications," in *Proceedings of the 17th International Conference on System Theory, Control and Computing (ICSTCC '13)*, pp. 459–464, Sinaia, Romania, October 2013.

[14] D. Sendrescu and E. Bobasu, "Parameter identification of bacterial growth bioprocesses using heuristics for global optimization," in *Proceedings of the 17th International Conference on System Theory, Control and Computing (ICSTCC '13)*, pp. 485–490, Sinaia, Romania, October 2013.

[15] D. Karnopp and R. Rosenberg, *System Dynamics: A Unified Approach*, John Wiley & Sons, New York, NY, USA, 1974.

[16] J. Thoma, *Introduction to Bond Graphs and Their Applications*, Pergamon Press, Oxford, UK, 1975.

[17] P. Gawthrop and L. Smith, *Metamodelling: Bond Graphs and Dynamic Systems*, Prentice Hall, Hemel, UK, 1996.

[18] W. Borutzky, *Bond Graph Methodology. Development and Analysis of Multidisciplinary Dynamic System Models*, Springer, London, UK, 2009.

[19] G. Dauphin-Tanguy, Ed., *Les bond graphs*, Hermes Sciences, Paris, France, 2000.

[20] F. Couenne, C. Jallut, B. Maschke, P. C. Breedveld, and M. Tayakout, "Bond graph modelling for chemical reactors," *Mathematical and Computer Modelling of Dynamical Systems*, vol. 12, no. 2-3, pp. 159–174, 2006.

[21] C. Heny, D. Simanca, and M. Delgado, "Pseudo-bond graph model and simulation of a continuous stirred tank reactor," *The Journal of the Franklin Institute*, vol. 337, no. 1, pp. 21–42, 2000.

[22] J. Thoma and B. Ould Bouamama, *Modelling and Simulation in Thermal and Chemical Engineering. A Bond Graph Approach*, Springer, Berlin, Germany, 2000.

[23] V. Díaz-Zuccarini, D. Rafirou, J. LeFevre, D. R. Hose, and P. V. Lawford, "Systemic modelling and computational physiology: the application of Bond Graph boundary conditions for 3D cardiovascular models," *Simulation Modelling Practice and Theory*, vol. 17, no. 1, pp. 125–136, 2009.

[24] D. Selişteanu, M. Roman, and D. Şendrescu, "Pseudo bond graph modelling and on-line estimation of unknown kinetics for a wastewater biodegradation process," *Simulation Modelling Practice and Theory*, vol. 18, no. 9, pp. 1297–1313, 2010.

[25] M. Roman and D. Selisteanu, "Pseudo bond graph modeling of wastewater treatment bioprocesses," *Simulation*, vol. 88, no. 2, pp. 233–251, 2012.

[26] M. Roman and D. Selişteanu, "Enzymatic synthesis of ampicillin: nonlinear modeling, kinetics estimation, and adaptive control," *Journal of Biomedicine and Biotechnology*, vol. 2012, Article ID 512691, 14 pages, 2012.

[27] J. Gao, V. M. Gorenflo, J. M. Scharer, and H. M. Budman, "Dynamic metabolic modeling for a MAb bioprocess," *Biotechnology Progress*, vol. 23, no. 1, pp. 168–181, 2007.

[28] I. Petric and V. Selimbašić, "Development and validation of mathematical model for aerobic composting process," *Chemical Engineering Journal*, vol. 139, no. 2, pp. 304–317, 2008.

[29] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical Programming: A Publication of the Mathematical Programming Society*, vol. 106, no. 1, pp. 25–57, 2006.

[30] M. Schwaab, E. C. Biscaia, Jr., J. L. Monteiro, and J. C. Pinto, "Nonlinear parameter estimation through particle swarm optimization," *Chemical Engineering Science*, vol. 63, no. 6, pp. 1542–1552, 2008.

[31] M. C. A. F. Rezende, C. B. B. Costa, A. C. Costa, M. R. W. Maciel, and R. M. Filho, "Optimization of a large scale industrial reactor by genetic algorithms," *Chemical Engineering Science*, vol. 63, no. 2, pp. 330–341, 2008.

[32] I. A. Maraziotis, A. Dragomir, and D. Thanos, "Gene regulatory networks modelling using a dynamic evolutionary hybrid," *BMC Bioinformatics*, vol. 11, article 140, 2010.

[33] A. Sîrbu, H. J. Ruskin, and M. Crane, "Comparison of evolutionary algorithms in gene regulatory network model inference," *BMC Bioinformatics*, vol. 11, article 59, 2010.

[34] M. Clerc and J. Kennedy, "The particle swarm-explosion, stability, and convergence in a multidimensional complex space," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 1, pp. 58–73, 2002.

[35] A. Ratnaweera, S. K. Halgamuge, and H. C. Watson, "Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 240–255, 2004.

[36] D. Sendrescu, "Parameter identification of anaerobic wastewater treatment bioprocesses using particle swarm optimization," *Mathematical Problems in Engineering*, vol. 2013, Article ID 103748, 8 pages, 2013.

[37] B. Shen, C. Liu, J. Ye, E. Feng, and Z. Xiu, "Parameter identification and optimization algorithm in microbial continuous culture," *Applied Mathematical Modelling*, vol. 36, no. 2, pp. 585–595, 2012.

[38] J. Zhai, J. Ye, L. Wang, E. Feng, H. Yin, and Z. Xiu, "Pathway identification using parallel optimization for a complex metabolic system in microbial continuous culture," *Nonlinear Analysis. Real World Applications*, vol. 12, no. 5, pp. 2730–2741, 2011.

[39] J. Yuan, X. Zhang, X. Zhu, E. Feng, H. Yin, and Z. Xiu, "Modelling and pathway identification involving the transport mechanism of a complex metabolic system in batch culture," *Communications in Nonlinear Science and Numerical Simulation*, vol. 19, no. 6, pp. 2088–2103, 2014.

[40] J. Ye, Y. Zhang, E. Feng, Z. Xiu, and H. Yin, "Nonlinear hybrid system and parameter identification of microbial fed-batch culture with open loop glycerol input and pH logic control," *Applied Mathematical Modelling*, vol. 36, no. 1, pp. 357–369, 2012.

[41] X. Li, S. Zhang, Z. Xiu, and E. Feng, "Parameter identification model with the control term in batch anaerobic fermentation," *Applied Mechanics and Materials*, vol. 217-219, pp. 1535–1540, 2012.

[42] Z. Liao and C. Mei, "Estimation of biochemical variables using quantumbehaved particle swarm optimization (QPSO)-trained radius basis function neural network: a case study of fermentation process of L-glutamic acid," *African Journal of Biotechnology*, vol. 10, no. 26, pp. 5203–5208, 2011.

[43] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, pp. 1942–1948, Perth, Australia, December 1995.

[44] A. Gambhir, R. Korke, J. Lee, P. C. Fu, A. Europa, and W. S. Hu, "Analysis of cellular metabolism of hybridoma cells at distinct physiological states," *Journal of Bioscience and Bioengineering*, vol. 95, no. 4, pp. 317–327, 2003.

[45] H. P. J. Bonarius, V. Hatzimanikatis, K. P. H. Meesters, C. D. de Gooijer, G. Schmid, and J. Tramper, "Metabolic flux analysis of hybridoma cells in different culture media using mass balances," *Biotechnology and Bioengineering*, vol. 50, no. 3, pp. 299–318, 1996.

[46] F. D. Follstad, R. R. Balcarcel, G. Stephanopoulos, and D. I. C. Wang, "Metabolic flux analysis of hybridoma continuous culture steady state multiplicity," *Biotechnology and Bioengineering*, vol. 63, no. 6, pp. 675–683, 1999.

[47] K. K. Frame and W.-S. Hu, "Kinetic study of hybridoma cell growth in continuous culture. I. A model for non-producing cells," *Biotechnology and Bioengineering*, vol. 37, no. 1, pp. 55–64, 1991.

[48] A. Provost, G. Bastin, S. N. Agathos, and Y. J. Schneider, "Metabolic design of macroscopic bioreaction models: application to Chinese hamster ovary cells," *Bioprocess and Biosystems Engineering*, vol. 29, no. 5-6, pp. 349–366, 2006.

[49] G. Bastin and D. Dochain, *On-Line Estimation and Adaptive Control of Bioreactors*, Elsevier, Amsterdam, The Netherlands, 1990.

[50] D. Dochain, Ed., *Automatic Control of Bioprocesses*, ISTE and John Wiley & Sons, London, UK, 2008.

[51] H. Hasanvand, B. B. Zad, B. Mozafari, and H. Maskani, "Fuzzy logic controller design based SVC for improving power system damping," *International Review of Automatic Control*, vol. 4, no. 5, pp. 740–748, 2011.

[52] Y. B. Amlashi and H. Afrakhte, "Determination of wind plant output capacity using discrete Markov chains and PSO methods in comparison with FCM," *International Review on Modelling & Simulations*, vol. 4, no. 2, pp. 819–823, 2011.

[53] R. C. Eberhart and Y. Shi, "Comparing inertia weights and constriction factors in particle swarm optimization," in *Proceedings of the Congress on Evolutionary Computation*, pp. 84–88, La Jolla, Calif, USA, July 2000.

[54] Y. Shi and R. C. Eberhart, "Empirical study of particle swarm optimization," in *Proceedings of the Congress on Evolutionary Computation*, pp. 1945–1950, Washington, DC, USA, July 1999.

[55] I. V. Tetko, D. J. Livingstone, and A. I. Luik, "Neural network Studies. 1. Comparison of overfitting and overtraining," *Journal of Chemical Information and Computer Sciences*, vol. 35, no. 5, pp. 826–833, 1995.

[56] R. V. Culshaw, S. Ruan, and G. Webb, "A mathematical model of cell-to-cell spread of HIV-1 that includes a time delay," *Journal of Mathematical Biology*, vol. 46, no. 5, pp. 425–444, 2003.

[57] S.-I. Niculescu, C.-I. Morărescu, W. Michiels, and K. Gu, "Geometric ideas in the stability analysis of delay models in biosciences," in *Biology and Control Theory: Current Challenges*, I. Queinnec, S. Tarbouriech, G. Garcia et al., Eds., vol. 357 of *Lecture Notes in Control and Information Sciences*, pp. 217–259, Springer, Berlin, Germany, 2007.