

Research Article

Distribution-Based Identification of Yield Coefficients in a Baker's Yeast Bioprocess

Dorin Sendrescu

Department of Automatic Control, University of Craiova, A.I. Cuza 13, 200585 Craiova, Romania

Correspondence should be addressed to Dorin Sendrescu, dorins@automation.ucv.ro

Received 24 August 2011; Revised 16 November 2011; Accepted 30 November 2011

Academic Editor: Alexander Pogromsky

Copyright © 2012 Dorin Sendrescu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A distribution-based identification procedure for estimation of yield coefficients in a baker's yeast bioprocess is proposed. This procedure transforms a system of differential equations to a system of algebraic equations with respect to unknown parameters. The relation between the state variables is represented by functionals using techniques from distribution theory. A hierarchical structure of identification is used, which allows obtaining a linear algebraic system of equations in the unknown parameters. The coefficients of this algebraic system are functionals depending on the input and state variables evaluated through some test functions from distribution theory. First, only some state equations are evaluated throughout test functions to obtain a set of linear equations in parameters. The results of this first stage of identification are used to express other parameters by linear equations. The process is repeated until all parameters are identified. The performances of the method are analyzed by numerical simulations.

1. Introduction

A process carried out in a bioreactor can be defined as a set of m biochemical reactions involving n components (with $n \geq m$). The global dynamics can be represented by following the dynamical state-space model [1]:

$$\frac{d\xi}{dt} = K \cdot \varphi(\xi, t) - D\xi + F - Q, \quad (1.1)$$

where $\xi \in \mathcal{R}^{n \times 1}$.

This model describes the behaviour of an entire class of biotechnological processes and is referred to as the general dynamical state-space model of this class of bioprocesses [1, 2]. In (1.1), the term $K \cdot \varphi(\xi, t)$ is the rate of consumption and/or production of the components

in the reactor, that is, the reaction kinetics and the term $-D\xi + F - Q$ represents the exchange with the environment, that is, the dynamics of the component transportation through the bioreactor. The strongly nonlinear character of the model (1.1) is given by the reaction kinetics. In many situations, the yield coefficients, the structure, and the parameters of the reaction rates are partially known or unknown. Many of the evolved control methods for these kind of systems—like model predictive control [3], robust or adaptive control [4, 5]—are based on good initial estimates of the yield parameters. This paper deals with the yield coefficients identification of baker's yeast bioprocess.

In recent years, a progress has been made in the area of continuous-time system identification. Even if the most physical systems are naturally continuous, a much more attention has been paid to parameter estimation of discrete-time systems, mainly because they are better suited for numerical implementations. Continuous-time identification makes possible a more direct link to the physical properties and operation of the underlying systems [6, 7] and the direct estimation of physical parameters which have a clear significance [8]. One important direction in continuous-time system identification is to transform the system of differential equations to an algebraic system that reveals the unknown parameters. By using some measures, the direct computation of the input-output data derivatives can be completely avoided. For linear system identification, several methods are reported on this direction: identification based on the Laplace transformation and then use the Laguerre filter [9] or transforming the continuous-time system to the frequency domain [10].

A novel approach to identify the continuous-time system is the distribution-based approach, using deterministic distributions [11] or random distributions [12]. Through these approaches, derivatives are described in the sense of distribution theory and one constructs the input-output algebraic relationships using differential information produced in the distribution sense.

Identification of nonlinear continuous-time systems is far away more complicated. The traditional procedures are based on the Volterra functional series, expressed in time domain or frequency domain [7]. The parameter identification of deterministic nonlinear continuous-time systems (NCTSs), modelled by polynomial-type differential equations has been considered by numerous authors [13, 14]. They use the modulating functions as a sum of sinusoids, whose fundamental period has a fixed length, to convert differential equations into an algebraic form in the frequency domain. These methods can be applied only if the NCTS is exactly integrable. This means that all the terms can be written as pure derivatives of some computable function of the measured signals. The computational burden in applying the method of modulating functions is overtaken by using Fast Fourier Transform algorithms.

In this paper we propose an identification method for a class of NCTS considering that the unknown parameters can appear in rational relations with measured variables. Using techniques utilized in distribution approach the measurable functions and their derivatives are represented by functionals on a fundamental space of testing functions. Such systems are common in biotechnology [1, 15]. It is supposed that the system is described by state equations and all state variables are measurable (or they can be estimated using some state estimators [16]).

The most important approaches for the yield coefficients estimation make use of the state transformations based on the general structure [17–19]. The main idea of this paper is to use a hierarchical structure of identification to estimate the yield coefficients in the baker's yeast bioprocess. First, some state equations are utilized to obtain a set of linear equations in some parameters. The results of this first stage of identification are used for expressing other parameters by linear equations. This process is repeated until all parameters are identified.

The paper is organized in the following way. The nonlinear dynamical model of a baker's yeast process is given in Section 2. Section 3 presents the problem statement of continuous-time systems identification based on distribution approach and the hierarchical structure of yield coefficients identification. Section 4 is dedicated to the analysis of properties of the identification algorithm. Some experimental results are presented in Section 5, and conclusions in Section 6.

2. Nonlinear Dynamical Model of Baker's Yeast Process

The yeast growth has already been the subject of numerous studies and papers. The reaction scheme of baker's yeast bioprocess is [20]



In the reaction schemes (2.1)–(2.3), S is the glucose, X the biomass, C the dissolved oxygen, E the ethanol, and G the dissolved carbon dioxide. The first reaction scheme represents the respiratory growth on glucose; the second reaction scheme represents the fermentative growth on glucose, and finally the third reaction represents the respiratory growth on ethanol. μ_s^o , μ_s^r , and μ_e^o are three specific growth rates that reflect the capacity of the baker's yeast to exploit three catabolic pathways for energy and material sources.

The dynamical model of the fed-batch bioprocess with the reaction schemes (2.1)–(2.3) can be obtained by using the mass balance of the components, considering that the bioreactor is well mixed, the yield coefficients, are constant and the dynamic of the gas phase can be neglected [20]:

$$\frac{dX}{dt} = (\mu_s^o + \mu_s^r + \mu_e^o)X - DX, \quad (2.4)$$

$$\frac{dS}{dt} = (-k_1\mu_s^o - k_2\mu_s^r)X + D(S_{in} - S), \quad (2.5)$$

$$\frac{dE}{dt} = (k_3\mu_s^r - k_4\mu_e^o)X - DE, \quad (2.6)$$

$$\frac{dC}{dt} = (-k_5\mu_s^o - k_6\mu_e^o)X - DC + OTR, \quad (2.7)$$

$$\frac{dG}{dt} = (k_7\mu_s^o + k_8\mu_s^r + k_9\mu_e^o)X - DG + CTR, \quad (2.8)$$

$$\frac{dV}{dt} = F_{in} = DV. \quad (2.9)$$

The dynamical model is expressed as a set of differential equations, in terms of the concentrations of components, and with the additional (2.9), which represents the dynamics

of the solution volume V in the bioreactor. In (2.4)–(2.9), k_i , $i = \overline{1,9}$ are the positive yield coefficients, S_{in} is the glucose concentration on the feed, OTR is the oxygen rate, defined as $\text{OTR} = K_L a (C^* - C)$ (where $K_L a$ is the mass transfer coefficient and C^* the equilibrium concentration of the dissolved oxygen), CTR is the carbon dioxide transfer rate (defined as $\text{CTR} = K_V K_L a G$, with K_V a transfer coefficient), F_{in} is the input feed rate, and $D = F_{\text{in}}/V$ is the so-called dilution rate.

The model (2.4)–(2.8) can be expressed in the matrix form:

$$\begin{bmatrix} \dot{X} \\ \dot{S} \\ \dot{E} \\ \dot{C} \\ \dot{G} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ -k_1 & -k_2 & 0 \\ 0 & k_3 & -k_4 \\ -k_5 & 0 & -k_6 \\ k_7 & k_8 & k_9 \end{bmatrix}}_K \begin{bmatrix} \mu_s^o \\ \mu_s^r \\ \mu_e^o \end{bmatrix} X - D \begin{bmatrix} X \\ S \\ E \\ C \\ G \end{bmatrix} + \begin{bmatrix} 0 \\ DS_{\text{in}} \\ 0 \\ \text{OTR} \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \text{CTR} \end{bmatrix}. \quad (2.10)$$

The next notations will be used: the state vector is $\xi = [X \ S \ E \ C \ G]^T = [\xi_1 \ \xi_2 \ \xi_3 \ \xi_4 \ \xi_5]^T$, the vector of reaction rates (the reaction kinetics) is $\Phi(\xi) = [\mu_s^o X \ \mu_s^r X \ \mu_e^o X]^T$; K is the matrix of yield coefficients, $F = [0 \ DS_{\text{in}} \ 0 \ \text{OTR}]^T$ is the vector of feeding rates, and $Q = [0 \ 0 \ 0 \ 0 \ \text{CTR}]^T$ is the vector of rates of removal of the components in gaseous form.

The most difficult task for the construction of the dynamical model is the modelling of the reaction kinetics. The form of kinetics is complex, nonlinear, and in many cases unknown. The reaction rates can be expressed as

$$\Phi(\xi, t) = H(\xi) \cdot \mu(\xi, t), \quad (2.11)$$

where $H(\xi)$ is a state-dependent diagonal matrix, whose elements correspond to the reactions' reactants.

The specific growth rates $\mu(\xi, t)$ are considered as a vector of time-varying parameters. For the baker's yeast bioprocess the form of matrix $H(\xi)$ is very simple:

$$H(\xi) = \begin{bmatrix} X & 0 & 0 \\ 0 & X & 0 \\ 0 & 0 & X \end{bmatrix}. \quad (2.12)$$

Then the model (2.10) becomes

$$\frac{d\xi}{dt} = K \cdot H(\xi) \cdot \mu(\xi, t) - D \cdot \xi + F - Q. \quad (2.13)$$

For the reaction rates of the baker's yeast bioprocess there are many possible models [21, 22], and so forth. However, the process of yeast growth on glucose with ethanol production is generally described by the following three metabolic reactions, widely analyzed by Sonnleitner and Kappeli [20].

First, the reaction rate of the respiratory growth on glucose and the associated specific growth rate are

$$\begin{aligned}\varphi_1(\xi) &= \mu_1(\xi) \cdot X, \\ \mu_1(\xi) &= \min\left(\mu_s, \frac{\mu_{c \max}}{k_5}\right),\end{aligned}\quad (2.14)$$

where the kinetic terms associated with the glucose consumption μ_s and with the respiratory capacity $\mu_{c \max}$ are modelled by Monod-type laws: $\mu_s = q_{s \max} S / (S + K_s)$, $\mu_{c \max} = q_{c \max} C / (C + K_c)$, with $q_{s \max}$ and $q_{c \max}$ the maximal values of the specific growth rates of glucose and oxygen, and K_s and K_c the saturation parameters for glucose and oxygen uptake, respectively.

Second, the reaction rate of the fermentative growth on glucose and the associated specific growth rate are

$$\begin{aligned}\varphi_2(\xi) &= \mu_2(\xi) \cdot X, \\ \mu_2(\xi) &= \max\left(0, \mu_s - \frac{\mu_{c \max}}{k_5}\right).\end{aligned}\quad (2.15)$$

Finally, if the oxidation capacity is sufficiently high to oxidize both ethanol and glucose, then their consumption is possible. Then, the reaction rate of the respiratory growth on ethanol and the associated specific growth rate are

$$\begin{aligned}\varphi_3(\xi) &= \mu_3(\xi) \cdot X, \\ \mu_3(\xi) &= \max\left(0, \min\left(\mu_e, \frac{\mu_{c \max} - k_5 \mu_s}{k_6}\right)\right),\end{aligned}\quad (2.16)$$

where μ_e is the potential ethanol oxidative rate, modelled by Monod-type law: $\mu_e = q_{e \max} E / (E + K_e)$, with $q_{e \max}$ the maximal value of the specific growth rates of ethanol, and K_e the saturation parameter for growth on ethanol.

The above kinetic model is based on the well-known bottleneck hypothesis developed by Sonnleitner and Kappeli [20]. This model assumes a limited capacity of yeast, leading to the production of ethanol under conditions of oxygen limitation and/or high glucose concentration. The glucose concentration that corresponds to the oxidative capacity is reached when $\mu_s = \mu_{c \max} / k_5$, and it is called critical concentration. Depending on this critical concentration, two operating regimes can be observed. At glucose concentration lower than its critical value, the system is in the respirative regime. The glucose consumption rate is smaller than the maximal respiratory capacity $\mu_s \leq \mu_{c \max} / k_5$ and the rate of the oxidative glucose metabolism is given by the glucose consumption rate. Ethanol can be oxidized in parallel with glucose and the rate of oxidative ethanol metabolism is determined by the excess of respiratory capacity and the available ethanol. At glucose concentration higher than its critical value, the system is in the so-called respirofermentative regime. Then the glucose consumption rate is larger than the maximal respiratory capacity $\mu_s > \mu_{c \max} / k_5$ and the respiratory capacity establishes the rate of the oxidative glucose metabolism. The excess of glucose is processed by the fermentative metabolism. Under lack of oxygen conditions, the fermentative pathway is predominating.

Since there are three unknown kinetics, the implementation of some state observers (such as asymptotic observers) require the online measurements of three state variables. Also, the online estimation algorithms for kinetic parameters (e.g., observer-based techniques) necessitate the knowledge of minimum three states. From the technological point of view, E , C , and G are the online available state variables, so it is important to design the estimators by using these measurements. Unfortunately, according to the yield coefficient values, E , C , and G are linearly dependent, with the corresponding yield matrix being ill conditioned [23]. This would result in sensitive algorithms to numerical errors, thus having performances degraded. To solve this problem, Pomerleau and Perrier [23] suggested a reformulation of model (2.10). This reformulation is based on the division of the complete process model (2.10) into two partial models, corresponding to the above-described mechanisms: the respirofermentative partial model for the ethanol production state of the process and the respirative partial model, corresponding to the ethanol consumption state of the process.

The fermentative partial model is stated as

$$\begin{bmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \\ \dot{\xi}_3 \\ \dot{\xi}_4 \\ \dot{\xi}_5 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -k_1 & -k_2 \\ 0 & k_3 \\ -k_5 & 0 \\ k_7 & k_8 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \xi_1 - D \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \end{bmatrix} + \begin{bmatrix} 0 \\ DS_{in} \\ 0 \\ OTR \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ CTR \end{bmatrix}. \quad (2.17)$$

The respirative partial model is stated as

$$\begin{bmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \\ \dot{\xi}_3 \\ \dot{\xi}_4 \\ \dot{\xi}_5 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -k_1 & 0 \\ 0 & -k_4 \\ -k_5 & -k_6 \\ k_7 & k_9 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_3 \end{bmatrix} \xi_1 - D \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \end{bmatrix} + \begin{bmatrix} 0 \\ DS_{in} \\ 0 \\ OTR \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ CTR \end{bmatrix}. \quad (2.18)$$

3. Distribution-Based Identification of Yield Coefficients

3.1. Problem Statement

In this approach the set of nonlinear differential equations describing the state evolution is mapped into a set of linear algebraic equations with respect to the model parameters. Using techniques utilized in distribution approach, the measurable functions and their derivatives are represented by functionals on the fundamental space of testing functions [14]. The main advantages of this method are that a set of algebraic equations with real coefficients results and the formulations are free from boundary conditions. The idea of utilizing test functions in system identification was proposed by Pearson and Lee [13] in terms of modulating functions in order to ameliorate the noise handling for deterministic least-squares identification based on time-limited data.

Let us denote by Ω_n the fundamental space from the distribution theory [24] of the real functions $\omega : \mathfrak{R} \rightarrow \mathfrak{R}$, $t \rightarrow \omega(t)$ with compact support T , having continuous derivatives at least up to the order n . Let $q : \mathfrak{R} \rightarrow \mathfrak{R}$, $t \rightarrow q(t)$ be a function which admits a Riemann

integral on any compact interval T from \mathfrak{R} . Using this function, a unique distribution (or generalized function)

$$F_q : \Omega_n \longrightarrow \mathfrak{R}, \quad \omega \longrightarrow F_q(\omega) \in \mathfrak{R} \quad (3.1)$$

can be built by the relation:

$$F_q(\omega) = \int_{\mathfrak{R}} q(t)\omega(t)dt, \quad \forall \omega \in \Omega_n. \quad (3.2)$$

In distribution theory, the notion of k -order derivative is introduced. If $F_q \in \Omega_n$, then its k -order derivative is a new distribution $F_q^{(k)} \in \Omega_n$ uniquely defined by the relations:

$$\begin{aligned} F_q^{(k)}(\omega) &= (-1)^k F_q(\omega^{(k)}), \quad \forall \omega \in \Omega_n, \\ \omega \longrightarrow F_q^{(k)}(\omega) &= (-1)^k \int_{\mathfrak{R}} q(t)\omega^{(k)}(t)dt \in \mathfrak{R}, \end{aligned} \quad (3.3)$$

where

$$\omega^{(k)} : \mathfrak{R} \longrightarrow \mathfrak{R}, \quad t \longrightarrow \omega^{(k)}(t) = \frac{d^k \omega(t)}{dt^k} \quad (3.4)$$

is the k -order time derivative of the test function.

When $q \in C^k(\mathfrak{R})$, then

$$F_q^{(k)}(\omega) = \int_{\mathfrak{R}} q^{(k)}(t)\omega(t)dt = (-1)^k \int_{\mathfrak{R}} q(t)\omega^{(k)}(t)dt \quad (3.5)$$

that means the k -order derivative of a distribution generated by a function $q \in C^k(\mathfrak{R})$ equals the distribution generated by the k -order time derivative of the function q . So, in place of the states and their time derivatives of a system one utilize the corresponding distributions and, in some particular cases, it is possible to obtain a system of equations linear in parameters. If the system is compatible the model parameters are structurally identifiable.

In our study, it were utilized three types of test functions characterized by a bounded support $T = (t_a \ t_b)$, $t_a < t_b$, all of these accomplishing the condition:

$$\omega(t) = 0; \quad \forall t \in (-\infty, t_a] \cup [t_b, +\infty). \quad (3.6)$$

The nonzero restriction is one of the following three types.

(1) Exponential:

$$\omega(t) = \exp\left(\frac{|t_a t_b|}{(t - t_a)(t - t_b)}\right). \quad (3.7)$$

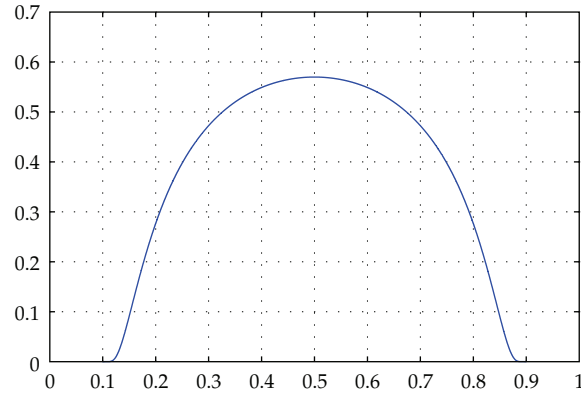


Figure 1: Exponential type test function.

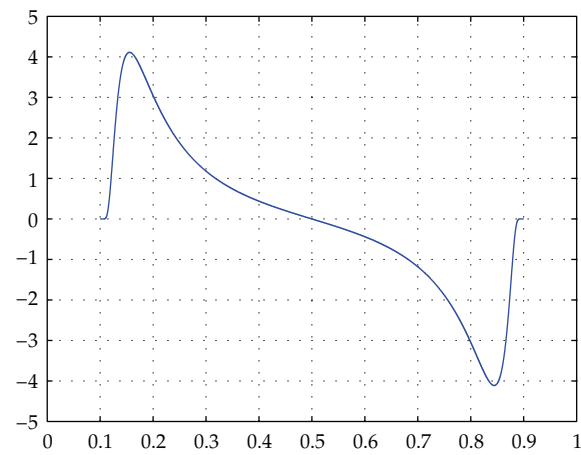


Figure 2: First-order derivative of exponential-type test function.

(2) Sinusoidal:

$$\omega(t) = \sin^p\left(\frac{(t-t_b)\pi}{t_b-t_a}\right). \quad (3.8)$$

(3) Polynomial:

$$\omega(t) = (t-t_a)^p(t-t_b)^p, \quad (3.9)$$

where $p \geq 2$ is an integer.

Figures 1 and 2 present the exponential type test function and its first-order derivative for $T = [0.1 \ 0.9]$.

3.2. The Hierarchical Structure of the Identification Algorithm

Consider all state variables accessible for measurements. This assumption can be considered unrealistic but the unknown states can be estimated using different approaches. From the multitude of proposed solutions one can mention methods based on balance equations, observers (asymptotic observers, high-gain observers) [5], Kalman filters [25], neural networks [26], and fuzzy reasoning [27].

The dynamical systems (2.17) and (2.18) contain rational dependences between parameters and measured variables. To obtain linear equations in unknown parameters, the identification problem is split in several simpler problems [28]. Based on the specific structure of these systems, it is possible to group the state equations in such way to determine some interconnected relations. They are organized in a hierarchical structure. In the first stage some state equations are utilized to obtain a set of linear equations in some parameters. If these parameters are identifiable then they can be used as known parameters in the following stages. This process is repeated in the other stages until all the parameters are identified.

Because the dilution rate D can be externally modified, it will be considered the first component of the input vector and the other component of u is the concentration S_{in} : $u = [u_1 \ u_2] = [D \ S_{in}]$.

From the third and fourth equations of the fermentative partial model (2.17) one gets the following relations for the reaction rates:

$$\begin{aligned}\mu_1 \xi_1 &= \frac{1}{k_5} (-\dot{\xi}_4 - D\xi_4 + \text{OTR}), \\ \mu_2 \xi_1 &= \frac{1}{k_3} (\dot{\xi}_3 + D\xi_3).\end{aligned}\tag{3.10}$$

These relations are replaced in the other three equations to obtain linear equations with respect to yield coefficients. So, one can split the identification problem in the following steps.

Step 1 (identification of k_3 and k_5). Replacing (3.10) in the first equation of the system (2.17) one get

$$(\dot{\xi}_1 + D\xi_1) = \theta_3(\dot{\xi}_3 + D\xi_3) + \theta_5(-\dot{\xi}_4 - D\xi_4 + \text{OTR}),\tag{3.11}$$

where $\theta_3 = 1/k_3$ and $\theta_5 = 1/k_5$.

One denotes

$$\begin{aligned}v(t) &= \dot{\xi}_1(t) + D\xi_1(t), \\ w_1(t) &= \dot{\xi}_3(t) + D\xi_3(t), \\ w_2(t) &= -\dot{\xi}_4(t) - D\xi_4(t) + \text{OTR}.\end{aligned}\tag{3.12}$$

Multiplying these functions with a test function $\omega_i(t)$ and integrating over \mathfrak{R} one gets the following distributions:

$$\begin{aligned} F_v^i(\omega_i) &= \int_{\mathfrak{R}} v(t) \cdot \omega_i(t) dt \\ &= \int_{\mathfrak{R}} \xi_1(t) \cdot \omega_i(t) dt + \int_{\mathfrak{R}} D\xi_1(t) \cdot \omega_i(t) dt = \int_{\mathfrak{R}} \xi_1(t) \cdot \omega_i^{(1)}(t) dt + \int_{\mathfrak{R}} D\xi_1(t) \cdot \omega_i(t) dt, \end{aligned} \quad (3.13)$$

$$\begin{aligned} F_{w_1}^i(\varphi_i) &= \int_{\mathfrak{R}} w_1(t) \cdot \omega_i(t) dt \\ &= \int_{\mathfrak{R}} \xi_3(t) \cdot \omega_i(t) dt + \int_{\mathfrak{R}} D\xi_3(t) \cdot \omega_i(t) dt = - \int_{\mathfrak{R}} \xi_3(t) \cdot \omega_i^{(1)}(t) dt + \int_{\mathfrak{R}} D\xi_3(t) \cdot \omega_i(t) dt, \\ F_{w_2}^i(\omega_i) &= \int_{\mathfrak{R}} w_2(t) \cdot \omega_i(t) dt = \int_{\mathfrak{R}} -\xi_4(t) \cdot \omega_i(t) dt + \int_{\mathfrak{R}} (-D\xi_4(t) + \text{OTR}) \cdot \omega_i(t) dt \\ &= \int_{\mathfrak{R}} \xi_4(t) \cdot \omega_i^{(1)}(t) dt + \int_{\mathfrak{R}} (-D\xi_4(t) + \text{OTR}) \cdot \omega_i(t) dt. \end{aligned} \quad (3.14)$$

As one must identify two parameters it is enough to choose a finite number $N \geq 2$ of test functions ω_i , $i = 1 : N$ and to build an algebraic system of equations,

$$F_w \cdot \theta = F_v, \quad (3.15)$$

where, $\theta = [\theta_3 \ \theta_5]^T$ is the vector of unknown parameters, F_w is a $N \times 2$ matrix of real numbers:

$$F_w = \begin{bmatrix} F_{w_1}^1(\omega_1), \dots, F_{w_1}^i(\omega_i), \dots, F_{w_1}^N(\omega_N) \\ F_{w_2}^1(\omega_1), \dots, F_{w_2}^i(\omega_i), \dots, F_{w_2}^N(\omega_N) \end{bmatrix}^T \quad (3.16)$$

and F_v denotes an N -column real vector build from distributions calculated as in (3.13):

$$F_v = \left[F_v^1(\omega_1), \dots, F_v^i(\omega_i), \dots, F_v^N(\omega_N) \right]^T. \quad (3.17)$$

If $\text{rank}(F_w) = 2$ then a unique solution is obtained:

$$\hat{\theta} = \left(F_w^T \cdot F_w \right)^{-1} F_w^T \cdot F_v. \quad (3.18)$$

One obtains $\hat{k}_3 = 1/\hat{\theta}_3$ and $\hat{k}_5 = 1/\hat{\theta}_5$.

Step 2 (identification of k_1 and k_2). Considering known $k_3 = \hat{k}_3$ and $k_5 = \hat{k}_5$ from *Step 1* and substituting (3.10) in the second equation of (2.17) one gets

$$(\dot{\xi}_2 + D\xi_2 - DS_{in}) = \theta_1 \left(\frac{1}{\hat{k}_5} (-\dot{\xi}_4 - D\xi_4 + OTR) \right) + \theta_2 \left(\frac{1}{\hat{k}_3} (\dot{\xi}_3 + D\xi_3) \right), \quad (3.19)$$

where $\theta_1 = -k_1$ and $\theta_2 = -k_2$.

The procedure of identification from *Step 1* is repeated and one obtains $\hat{k}_1 = -\hat{\theta}_1$ and $\hat{k}_2 = -\hat{\theta}_2$.

Step 3 (identification of k_7 and k_8). Considering known $k_3 = \hat{k}_3$ and $k_5 = \hat{k}_5$ from *Step 1* and substituting (3.10) in the last equation of (2.17) one gets

$$(\dot{\xi}_5 + D\xi_5 + CTR) = \theta_7 \left(\frac{1}{\hat{k}_3} (\dot{\xi}_3 + D\xi_3) \right) + \theta_8 \left(\frac{1}{\hat{k}_5} (-\dot{\xi}_4 - D\xi_4 + OTR) \right), \quad (3.20)$$

where $\theta_7 = k_7$ and $\theta_8 = k_8$.

The procedure of identification from *Step 1* is repeated and one obtains $\hat{k}_7 = \hat{\theta}_7$ and $\hat{k}_8 = \hat{\theta}_8$.

Since the respirative partial model (2.18) has a similar structure with the fermentative partial model (2.17), in a similar manner one obtains the parameters k_4 , k_6 , and k_9 .

4. Analysis of the Algorithm Properties

In the following one presents some consistency and numerical aspects related to the presented algorithm and a short overview of the classical identification procedure widely used for this kind of biotechnological systems.

4.1. Identifiability

Identifiability is a necessary prerequisite for model identification; it concerns uniqueness of the model parameters determined from the input-output data, under ideal conditions of noise-free observations and error-free model structure. A remarkable feature of distribution-based identification procedure is that it provides a linear reparameterization of the input-output relation of the nonlinear system. This reparameterization of the system involves a very simple identifiability condition to be accomplished, that is, the existence of the matrix $(F_w^T \cdot F_w)^{-1}$ or, equivalently, F_w is of full rank.

4.2. Consistency

Obviously, consistency of the estimates is directly influenced by the precision of numerical integration used to compute the value of the distributions. There are available numerous integration methods (often called numerical quadrature) with various degrees of precision.

One presents shortly the techniques used in the simulations to the approximate calculation of a definite integral $I_f = \int_a^b f(x)dx$ where $f(x)$ is a given function and $[a, b]$ a finite interval. Interpolatory quadrature formulas, where the nodes are constrained to be equally spaced, are called Newton-Cotes formulas. These are especially suited for integrating a tabulated function (such is our case). Newton-Cotes numerical integration rule I_f is a weighted sum of function values: $I_f = \sum_{i=0}^{N-1} c_i \cdot f(x_i)$, where coefficients $\{c_i\}$ are derived from interpolating polynomial fitted for points $\{x_i, f_i\}$. A Newton-Cotes formula of any degree n can be constructed. One of the simplest integration methods is the trapezoidal rule. The trapezoidal rule is based on linear interpolation of $f(x)$ at $x_1 = a$ and $x_2 = b$, that is, $f(x)$ is approximated by

$$p(x) = f(a) + (x - a) \frac{f(b) - f(a)}{b - a}. \quad (4.1)$$

The integral of $p(x)$ equals the of trapezoid with base $(b - a)$ times the average height

$$\frac{1}{2}(f(a) + f(b)) \quad (4.2)$$

hence

$$\int_a^b f(x)dx = \frac{b - a}{2}(f(a) + f(b)). \quad (4.3)$$

To increase the accuracy one subdivides the interval $[a, b]$ and assumes that $f_i = f(x_i)$ is known on a grid of equidistant points $x_0 = a$, $x_i = x_0 + ih$, $x_n = b$, where $h = (b - a)/n$ is the step length. The trapezoidal approximation for the i th subinterval is

$$\int_{x_i}^{x_{i+1}} f(x)dx = \frac{h}{2}(f_i + f_{i+1}) + R_i, \quad (4.4)$$

where

$$R_i = -\frac{h^3}{12}f''(\zeta_i), \quad \zeta_i \in [x_i, x_{i+1}]. \quad (4.5)$$

Summing the contributions for each subinterval $[x_i, x_{i+1}]$, $i = 0 : n$, gives

$$\int_a^b f(x)dx = \frac{h}{2}(f_0 + f_n) + h \sum_{i=2}^{i=n-1} f_i + R_T. \quad (4.6)$$

The global truncation error is

$$R_T = -\frac{h^3}{12} \sum_{i=0}^{i=n-1} f''(\zeta_i) = -\frac{h^2}{12}(b - a)f''(\zeta), \quad \zeta \in [a, b]. \quad (4.7)$$

This shows that by choosing h small enough we can make the truncation error arbitrarily small. In other words, we have asymptotic convergence when $h \rightarrow 0$.

In the composite Simpson's rule one divides the interval $[a, b]$ into an even number $n = 2m$ steps of length h and uses the formula

$$\int_a^b f(x)dx = \frac{h}{3}(f_0 + 4S_1 + 2S_2 + f_n) + R_T, \quad (4.8)$$

where

$$S_1 = f_1 + f_3 + \cdots + f_{n-1}, \quad S_2 = f_2 + f_4 + \cdots + f_{n-2} \quad (4.9)$$

are sums over odd and even indices, respectively. The remainder is

$$R_T = \frac{h^4}{180}(b-a)f^{IV}(\zeta), \quad \zeta \in [a, b]. \quad (4.10)$$

This shows that one have gained two orders of accuracy compared to the trapezoidal rule, without using more function evaluations.

Let us study properties of Newton-Cotes formulas in frequency domain (or spectral properties). Newton-Cotes rules are symmetric (hence linear phase) digital filters with finite impulse response. One of the most important characteristic of digital filter is magnitude/frequency response-function which shows how much filter damps or amplifies magnitude of particular frequency contained in input data.

So, for trapezoidal rule and Simpson's rule one obtains

$$H(z) = \frac{h}{2} \cdot \frac{1+z^{-1}}{1-z^{-1}}, \quad H(z) = \frac{h}{3} \cdot \frac{1+4z^{-1}+z^{-2}}{1-z^{-1}}, \text{ respectively.} \quad (4.11)$$

These transfer functions have the amplitude-frequency Bode characteristics from Figure 3.

As one can see classical Newton-Cotes formulas suppress high frequencies (noise) in the input data. In that sense they can be considered low-pass filters. From Figure 3 one observes that trapezoidal rule offers a better suppression of noise than the Simpson's rule, so in the algorithm implementation one used the trapezoidal rule for numerical integration.

4.3. "Classical" Identification Procedure

The "classical" identification approach of yield coefficients is a two-step procedure under the assumption that full state measurements are available [1]. This method is based on a state transformation that allows reformulating the dynamical model into separate submodels. The first submodel depends only on the reaction structure and is independent of the kinetics. It can be linearly reparametrized and used for the identification of the yield coefficients by means of linear regressions, provided suitable identifiability conditions are satisfied. We present briefly this method for estimating the yield coefficients and we use it for comparison in the next section.

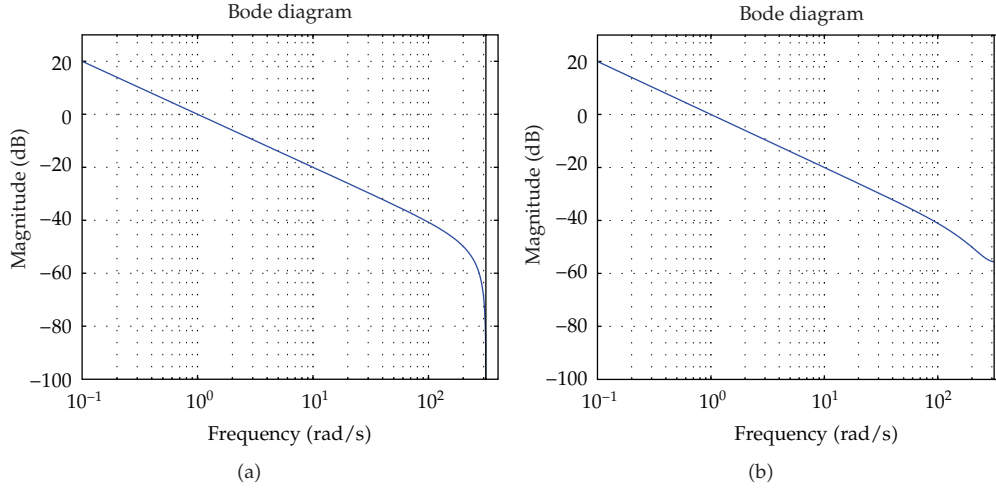


Figure 3: Amplitude-frequency bode characteristics: (a) trapezoidal rule; (b) Simpson's rule.

The general dynamical model given in (2.1) represents a particular class of nonlinear state-space models. The nonlinearity lies in the reaction rates $\varphi_i(\xi)$ that are (nonlinear) functions of the state variables. These functions enter the model in the form $K\varphi_i(\xi)$ (where K is a constant matrix), which is a set of linear combinations of the same nonlinear functions $\varphi_i(\xi)$. This particular feature can be exploited to separate the nonlinear part from the linear part of the model by an adequate linear state transformation. More precisely, one chooses a nonsingular partition

$$\begin{pmatrix} K_a \\ K_b \end{pmatrix} = TK, \quad (4.12)$$

with $K_a \in R^{p \times m}$ of full row rank matrix (i.e., $p = \text{rank}(K)$), $K_b \in R^{(n-p) \times m}$ and T a permutation matrix. The induced partitions of the vectors ξ and u are

$$\begin{pmatrix} \xi_a \\ \xi_b \end{pmatrix} = T\xi \quad \begin{pmatrix} u_a \\ u_b \end{pmatrix} = Tu. \quad (4.13)$$

Model (2.1) is then partitioned into two submodels:

$$\begin{aligned} \frac{d\xi_a}{dt} &= K \cdot \varphi(\xi_a, t) - D\xi_a + u_a, \\ \frac{d\xi_b}{dt} &= K \cdot \varphi(\xi_b, t) - D\xi_b + u_b. \end{aligned} \quad (4.14)$$

Then with the state transformation

$$\begin{aligned} \xi_a &= \xi_a, \\ z &= C\xi_a + \xi_b, \end{aligned} \quad (4.15)$$

one transform the initial model into

$$\begin{aligned}\dot{\xi}_a &= K_a \varphi(\xi_a, z - C\xi_a) - D\xi_a + u_a, \\ \dot{z} &= -Dz + Cu_a + u_b,\end{aligned}\quad (4.16)$$

where the $(n - p) \times p$ matrix C is the unique solution of

$$CK_a + K_b = 0 \quad (4.17)$$

that is,

$$C = -K_b K_a^* \quad (4.18)$$

where K_a^* is a generalized or pseudoinverse of K_a . This subsystem (4.16) can be augmented with an equation derived from (2.7) as follows:

$$\begin{aligned}\dot{z} &= -Dz + Cu_a + u_b, \\ \xi_b &= z - C\xi_a.\end{aligned}\quad (4.19)$$

It can be considered as a linear time-varying (if d varies in the course of time) model with state z , input (ξ_a, u_a, u_b) , and output ξ_b . It is nonlinearly parametrized by the yield coefficients but linearly reparametrized by the nonzero entries of C . When data of the signals ξ_a , u_a , u_b , and ξ_b are available, the auxiliary model (4.19) can be used to identify the yield coefficients independently of the knowledge of the reaction rates. The model (4.19) can be used to perform the identification of the nonzero entries of C by a linear regression technique, with the yield coefficients k_i recovered afterwards from (4.17). The identification makes sense only if the yield coefficients k_i are structurally uniquely identifiable with the auxiliary model (4.19).

5. Simulation Results

The efficacy of our distribution-based approach is shown by simulations, comparing it with this "classical" approach reviewed in the previous section. Also, the influence of sampling period, test functions type, initial conditions, and input type is analysed. To compare statistical performances of the different approaches and simulation conditions the empirical normalized mean square error (NMSE) was used, which is defined as

$$\text{NMSE} = \frac{1}{N} \sum_{j=1}^N \text{NMSE}(\hat{\theta}_j), \quad \text{with } \text{NMSE}(\hat{\theta}_j) = \left(\frac{\hat{\theta}_j - \theta_j^*}{\theta_j^*} \right)^2, \quad (5.1)$$

where N is the number of estimated parameters, $\hat{\theta}_j$ is the j th element of the estimated parameter vector, while the "*" superscript denotes the true value of the parameter.

The model given by (2.17) and (2.18) was integrated using a fourth-order Runge-Kutta routine with a sampling period of 1 minute.

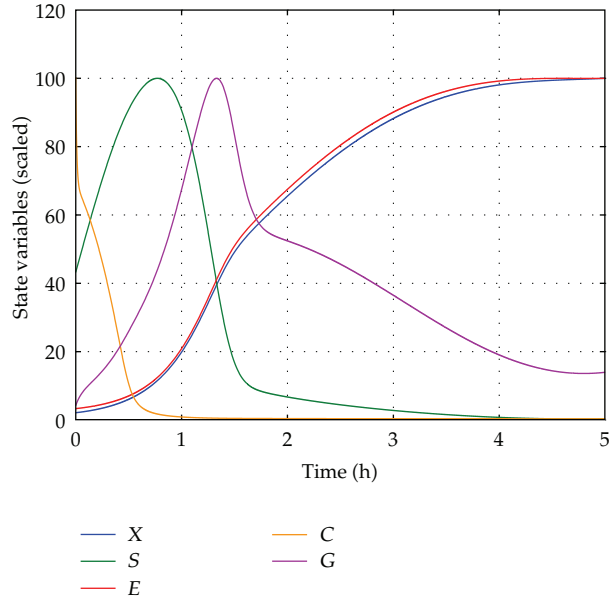


Figure 4: Time response for noise-free system.

In order to study the sensitivity of the estimation method to the sampling period T_s , to the initial conditions, to the test functions type, to the input type, and to the noise, the following parameters were used:

- (a) sampling period: $T_s \in \{1 \text{ min}, 10 \text{ min}, 20 \text{ min}\}$,
- (b) input signals: $u \in \{0.1 + 0.05 * \sin(t), 0.1 + 0.005 * t, 0.15 * \exp(9 * 10^{-3}t)\}$,
- (c) initial conditions:

$$\text{IC}_1: [1 \ 0.2 \ 0.15 \ 0.0066 \ 0.008]$$

$$\text{IC}_2: [0.1 \ 0.3 \ 0 \ 0.002 \ 0.01 \ 2]$$

$$\text{IC}_3: [0.5 \ 0.5 \ 3 \ 0.01 \ 0.001 \ 1],$$

- (d) test functions:

Type1: exponential

Type2: sinusoidal

Type3: polynomial,

- (e) noise: {noise-free, zero-mean white (SNR = 50 dB, 40 dB, 30 dB, 20 dB, 10 dB)}.

Some of the signals used in the identification algorithm are presented in Figures 4–6. Figure 4 presents the noise-free system time response for the fermentative partial model (2.17). The measured states are represented in percent with respect to their maximal values. Three test function of sinusoidal type are presented in Figure 5. In Figure 6 some noisy measurements are presented.

The simulation results presented in Tables 1–5 leads to the following remarks.

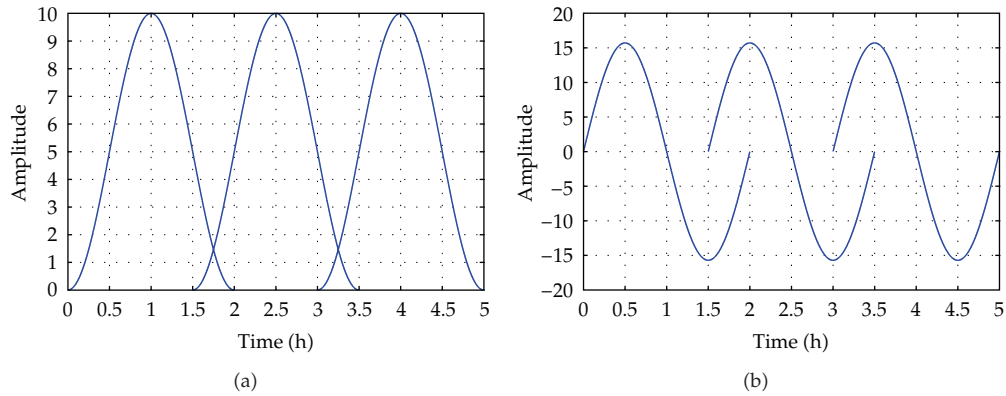


Figure 5: Test functions (a) and their derivatives (b) used for identification.

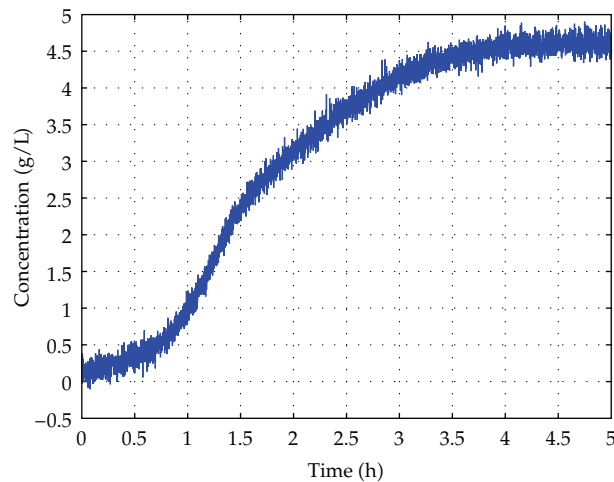


Figure 6: Noisy measurements of state variable E : SNR = 20 dB.

- (1) Sensitivity to the sampling period: results given in Table 1 show that the performance of the algorithm deteriorates with the increasing sampling period.
- (2) Sensitivity to the inputs type: results given in Table 2 conclude that the algorithm is not too sensitive to the type of input. However, as was expected, the best performances are obtained for a sinusoidal input type.
- (3) Sensitivity to the initial conditions and test function type: as shown in Tables 3 and 4, initial conditions and type of the test functions have very little influence on the accuracy of the estimates.
- (4) Sensitivity to the noise: Tables 5 and 6 present the simulation results for distribution-based method and for “classical” method, respectively. The results are still very good in the distribution-based case even for the high level of noise (SNR = 10 dB), while the “classical” approach is less robust to increasing noise level.

Table 1: Influence of the sampling period.

Ts	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	NMSE
1 min	0.4901	0.0500	0.1000	0.7199	1.2001	0.6401	0.8101	0.1100	1.1101	1.6046e-008
10 min	0.5022	0.0497	0.1000	0.7079	1.2038	0.6508	0.8357	0.1094	1.1287	2.8414e-004
20 min	1.2137	0.0381	0.0994	0.0101	1.4980	1.2710	2.3205	0.0840	2.2044	0.9722
"Real" value	0.49	0.05	0.1	0.72	1.2	0.64	0.81	0.11	1.11	

Table 2: Influence of the input type.

Input type	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	NMSE
0.1 + 0.05 * sin(t) (noise free)	0.4901	0.0500	0.1000	0.7199	1.2001	0.6401	0.8101	0.1100	1.1101	1.6046e-008
0.1 + 0.05 * sin(t) (SNR = 20 dB)	0.4531	0.0517	0.1000	0.6926	1.2459	0.6644	0.7969	0.1112	1.1523	0.0015
0.1 + 0.005 * t (noise free)	0.4902	0.0500	0.1000	0.7198	1.2000	0.6402	0.8096	0.1100	1.1103	6.9366e-008
0.1 + 0.005 * t (SNR = 20 dB)	0.3842	0.0531	0.1002	0.6319	1.0730	0.7183	0.6754	0.1128	1.2459	0.0150
0.15 * exp($9 * 10^{-3}t$) (noise free)	0.4902	0.0500	0.1000	0.7198	1.2000	0.6402	0.8097	0.1100	1.1103	6.3148e-008
0.15 * exp($9 * 10^{-3}t$) (SNR = 20 dB)	0.4882	0.0488	0.1004	0.6660	1.0738	0.6880	0.6900	0.1111	1.1932	0.0056
"Real" value	0.49	0.05	0.1	0.72	1.2	0.64	0.81	0.11	1.11	

Table 3: Influence of the initial conditions.

Initial conditions	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	NMSE
IC ₁	0.4901	0.0500	0.1000	0.7199	1.2001	0.6401	0.8101	0.1100	1.1101	1.6046e-008
IC ₂	0.4901	0.0500	0.1000	0.7197	1.1999	0.6401	0.8098	0.1100	1.1104	3.1454e-008
IC ₃	0.4908	0.0500	0.1000	0.7178	1.2002	0.6419	0.8158	0.1099	1.1134	9.1629e-006
"Real" value	0.49	0.05	0.1	0.72	1.2	0.64	0.81	0.11	1.11	

Table 4: Influence of the test functions type.

Test function	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	NMSE
type 1	0.4902	0.0499	0.1000	0.7198	1.1999	0.6401	0.8102	0.1100	1.1102	3.2506e-008
type 2	0.4901	0.0500	0.1000	0.7199	1.2000	0.6401	0.8102	0.1100	1.1101	1.5545e-008
type 3	0.4908	0.0500	0.1000	0.7195	1.1993	0.6405	0.8096	0.1100	1.1108	5.4343e-007
"Real" value	0.49	0.05	0.1	0.72	1.2	0.64	0.81	0.11	1.11	

Table 5: Influence of the noise (distribution-based method).

SNR	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	NMSE
noise free	0.4901	0.0500	0.1000	0.7199	1.2001	0.6401	0.8101	0.1100	1.1101	1.6046e-008
50 dB	0.4824	0.0502	0.1000	0.7147	1.1965	0.6447	0.8037	0.1102	1.1182	5.4504e-005
40 dB	0.4393	0.0509	0.1002	0.6803	1.1358	0.6753	0.7568	0.1102	1.1712	0.0030
30 dB	0.4707	0.0494	0.1007	0.6861	1.0953	0.6701	0.8540	0.1065	1.1622	0.0022
20 dB	0.4531	0.0517	0.1000	0.6926	1.2459	0.6644	0.7969	0.1112	1.1523	0.0015
10 dB	0.5567	0.0529	0.0993	0.5587	1.7758	0.7834	0.9887	0.1119	1.3587	0.0502
"Real" value	0.49	0.05	0.1	0.72	1.2	0.64	0.81	0.11	1.11	

Table 6: Influence of the noise ("classical" method).

SNR	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	NMSE
noise free	0.4903	0.0500	0.1000	0.7195	1.2001	0.6404	0.8109	0.1100	1.1105	3.1151e-07
50 dB	0.4876	0.0489	0.1003	0.7037	1.1748	0.6533	0.8199	0.1072	1.1254	3.2795e-04
40 dB	0.4501	0.0574	0.0986	0.6440	1.2432	0.7024	0.8474	0.1195	1.1821	0.0072
30 dB	0.3923	0.0659	0.0966	0.0527	3.2290	1.1875	0.8393	0.1338	1.7431	0.5517
20 dB	0.0133	0.0711	0.0983	0.0683	3.2732	1.2868	0.5392	0.1303	1.8578	0.7699
10 dB	0.3115	0.0776	0.0977	0.9807	6.1195	2.0355	0.7625	0.1254	2.7235	3.5841
"Real" value	0.49	0.05	0.1	0.72	1.2	0.64	0.81	0.11	1.11	

6. Conclusions

The paper presents a distribution-based identification procedure for offline estimation of yield coefficients in a baker's yeast bioprocess. This procedure is a functional type method, which transforms a differential system of equations to an algebraic system in unknown parameters. The relation between the state variables of the system is represented by functionals using techniques from distribution theory based on test functions from a finite-dimensional fundamental space. The identification algorithm has a hierarchical structure, which allows obtaining a linear algebraic system of equations in the unknown parameters. The coefficients of this algebraic system are functionals depending on the input and state variables and are evaluated through some testing functions from distribution theory.

According to the proposed procedure, first, only some state equations are evaluated throughout testing functions to obtain a set of linear equations in some parameters. The results of this first stage of identification are utilized for expressing other parameters by linear equations. This process is repeated until all parameters are identified.

The influence of the sampling period, initial conditions, test functions type, input type, and noise on the parameters estimates was empirically analyzed.

The proposed method presents two main advantages: there is necessary no information about the parameters of the reaction rates to identify the yield coefficients of the baker's yeast bioprocess and the algorithm provides very good results even the measurements are noise-contaminated because the evaluation of states derivatives is completely avoided.

Nomenclature

C :	Dissolved oxygen concentration (g/L)
C^* :	Equilibrium concentration of the dissolved oxygen (g/L)
CTR:	Carbon dioxide transfer rate ($\text{h}^{-1}\text{g/L}$)
D :	Dilution rate (h^{-1})
E :	Ethanol concentration (g/L)
F :	Vector of feeding rates
F_{in} :	Input feed rate (Lh^{-1})
G :	Dissolved carbon dioxide concentration (g/L)
K :	Matrix of yield coefficients
$k_i, i = \overline{1,9}$:	Yield coefficients
$K_L a$:	Mass transfer coefficient (h^{-1})
K_e :	Saturation parameter for growth on ethanol (g/L)

K_s, K_c :	Saturation parameters for glucose/oxygen uptake, respectively (g/L)
K_V :	Transfer coefficient
OTR:	Oxygen transfer rate ($\text{h}^{-1}\text{g/L}$)
Q :	Vector of rates of removal of the components in gaseous form
$q_{c\max}$:	Maximal oxygen specific growth rate ($\text{g O}_2/(\text{g biomass} \cdot \text{h})$)
$q_{e\max}$:	Maximal ethanol specific growth rate ($\text{g ethanol}/(\text{g biomass} \cdot \text{h})$)
$q_{s\max}$:	Maximal glucose specific growth rate ($\text{g glucose}/(\text{g biomass} \cdot \text{h})$)
S :	Glucose concentration (g/L)
S_{in} :	Glucose concentration on the feed (g/L)
V :	Culture volume in the reactor (L)
X :	Biomass concentration (g/L)
Φ :	Vector of reaction rates (reaction kinetics)
$\mu_s^o, \mu_s^r, \mu_e^o$:	Specific growth rates of baker's yeast bioprocess (h^{-1})
μ :	Vector of specific growth rates
$\mu_{c\max}$:	Kinetic term associated with the respiratory capacity
μ_e :	Potential ethanol oxidative rate
μ_s :	Kinetic term associated with the glucose consumption
ξ :	State vector
Ω_{η} :	The fundamental space from the distribution theory
\mathfrak{R} :	The set of real numbers
$\omega(t)$:	Test function
F_q :	Distribution (generalized function)
θ :	The vector of unknown parameters
$\theta_i, i = \overline{1,9}$:	The unknown parameters.

Acknowledgment

This work was supported by CNCSIS-UEFISCSU, Project number PN II-RU TE 106, 9/04.08.2010.

References

- [1] G. Bastin and D. Dochain, *On-Line Estimation and Adaptive Control of Bioreactors*, Elsevier, New York, NY, USA, 1990.
- [2] M. Roman and D. Selişteanu, "Modeling issues and structural properties of a class of nonlinear bioprocesses," *International Review of Automatic Control*, vol. 3, no. 6, pp. 578–587, 2010, ISSN 1974-6059.
- [3] M. Henson, "Dynamic modelling and control of yeast cell populations in continuous biochemical reactors," *Computers & Chemical Engineering*, vol. 27, no. 8-9, pp. 1185–1199, 2003.
- [4] F. Renard and A. Vande Wouwer, "Robust adaptive control of yeast fed-batch cultures," *Computers & Chemical Engineering*, vol. 32, no. 6, pp. 1238–1248, 2008.
- [5] D. Selişteanu, E. Petre, and V. B. Răsvan, "Sliding mode and adaptive sliding-mode control of a class of nonlinear bioprocesses," *International Journal of Adaptive Control and Signal Processing*, vol. 21, no. 8-9, pp. 795–822, 2007.
- [6] H. Unbehauen and G. P. Rao, *Identification of Continuous Systems*, vol. 10, North-Holland, Amsterdam, The Netherlands, 1987.
- [7] L. M. Li and S. A. Billings, "Continuous time non-linear system identification in the frequency domain," *International Journal of Control*, vol. 74, no. 11, pp. 1052–1061, 2001.
- [8] I. D. Landau, B. D. O. Anderson, and F. De Bruyne, "Recursive identification algorithms for continuous-time nonlinear plants operating in closed loop," *Automatica*, vol. 37, no. 3, pp. 469–475, 2001.

- [9] B. R. J. Haverkamp, C. T. Chou, M. Verhaegen, and R. Johansson, "Identification of continuous-time MIMO state space models from sampled data, in the presence of process and measurement noise," in *Proceedings of the 35th IEEE Conference on Decision and Control*, pp. 1539–1544, December 1996.
- [10] P. Overschee and B. De Moor, *Subspace System Identification for Linear Systems*, Kluwer Academic, Boston, Mass, USA, 1996.
- [11] C. Marin, "System identification based on distribution theory," in *Proceedings of the IASTED International Conference on Applied Simulation and Modelling (ASM '02)*, pp. 456–462, Crete, Greece, 2002.
- [12] A. Ohsumi, K. Kameyama, and K.-I. Yamaguchi, "Subspace identification for continuous-time stochastic systems via distribution-based approach," *Automatica*, vol. 38, no. 1, pp. 63–79, 2002.
- [13] A. E. Pearson and F. C. Lee, "On the identification of polynomial input-output differential systems," *IEEE Transactions on Automatic Control*, vol. 30, no. 8, pp. 778–782, 1985.
- [14] A. Patra and H. Unbehauen, "Identification of a class of nonlinear continuous-time systems using Hartley modulating functions," *International Journal of Control*, vol. 62, no. 6, pp. 1431–1451, 1995.
- [15] D. Selisteanu, E. Petre, and D. Şendrescu, "Modeling and identification of wastewater biodegradation process," in *Proceedings of the 12th International Symposium on Modeling, Simulation and System's Identification (SIMSIS '04)*, pp. 149–154, Galati, Romania, 2004.
- [16] D. Beluhan, D. Gosak, N. Pavlovic, and M. Vampola, "Biomass estimation and optimal control of the baker's yeast fermentation process," *Computers & Chemical Engineering*, vol. 19, supplement 1, pp. 387–392, 1995.
- [17] L. Chen, *Modelling, identifiability and control of complex biotechnological systems*, Ph.D. thesis, Université Catholique de Louvain, Belgium, 1992.
- [18] C. Rocha and E. C. Ferreira, "Design of optimal experiments for identification of yield coefficients in a baker's yeast model," in *Proceedings of the 1st European Symposium on Biochemical Engineering Science*, B. Glennon, P. M. Kieran, and K. Ch. A. M. Luyben, Eds., pp. 99–100, Dublin, Ireland, 1996.
- [19] C. Pertev, M. Türker, and R. Berber, "Dynamic modeling, sensitivity analysis and parameter estimation of industrial yeast fermenters," *Computers and Chemical Engineering*, vol. 21, no. 1, pp. S739–S744, 1997.
- [20] B. Sonnleitner and O. Kappeli, "Growth of *Saccharomyces cerevisiae* is controlled by its respiratory capacity: formulation and verification of a hypothesis," *Biotechnology and Bioengineering*, vol. 28, no. 6, pp. 927–937, 1986.
- [21] S. Feyo De Azevedo, B. Dahm, and F. R. Oliveira, "Hybrid modelling of biochemical processes: a comparison with the conventional approach," *Computers and Chemical Engineering*, vol. 21, no. 1, pp. S751–S756, 1997.
- [22] R. Oliveira, E. C. Ferreira, F. Oliveira, and S. Feyo De Azevedo, "A study on the convergence of observer-based kinetic estimators in fed-batch fermentations," *Journal of Process Control*, vol. 6, no. 6, pp. 367–371, 1996.
- [23] Y. Pomerleau and M. Perrier, "Estimation of multiple specific growth rates in bioprocesses," *AIChE Journal*, vol. 36, no. 2, pp. 207–215, 1990.
- [24] L. Schwartz, *Théorie des Distributions*, Paris, France, 1951.
- [25] P. Bogaerts, "A hybrid asymptotic-Kalman observer for bioprocesses," *Bioprocess and Biosystems Engineering*, vol. 20, no. 3, pp. 249–255, 1999.
- [26] M. N. Karim and S. L. Rivera, "Artificial neural networks in bioprocess state estimation," in *Advances in Biochemical Engineering/Biotechnology*, A. Fiechter, Ed., vol. 46 of *Modern Biochemical Engineering*, pp. 1–34, Springer, Berlin, Germany, 1992.
- [27] P. Ascencio, D. Sbarbaro, and S. F. de Azevedo, "An adaptive fuzzy hybrid state observer for bioprocesses," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 5, pp. 641–651, 2004.
- [28] D. Sendrescu, M. Roman, C. Marin, E. Bobasu, and D. Selisteanu, "Off-line identification of yield coefficients in a Baker's yeast bioprocess," in *Proceedings of the 29th Chinese Control Conference*, pp. 1254–1260, IEEE catalog no. CFP1040A-CDR, ISBN 978-7-89463-104-6, Beijing, China, July 2010.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

