STRUCTURAL IDENTIFIABILITY OF A MODEL FOR THE ACETIC ACID FERMENTATION PROCESS

Jorge E. Jiménez-Hornero,[(a)*] Inés M. Santos-Dueñas,[(b)] Isidoro García-García[(b)]

[(a)] Computing and Numerical Analysis Department, University of Córdoba, PO Box 14071 Córdoba, Spain

[(b)] Chemical Engineering Department, University of Córdoba, PO Box 14071 Córdoba, Spain

Abstract

Modelling has proved an essential tool for addressing research into biotechnological processes, particularly with a view to their optimization and control. Parameter estimation via optimization approaches is among the major steps in the development of biotechnology models. In fact, one of the first tasks in the development process is to determine whether the parameters concerned can be unambiguously determined and provide meaningful physical conclusions as a result. The analysis process is known as “identifiability" and presents two different aspects: structural or theoretical identifiability and practical identifiability. While structural identifiability is concerned with model structure alone, practical identifiability takes into account both the quantity and quality of experimental data. In this work, we discuss the theoretical identifiability of a new model for the acetic acid fermentation process and review existing methods for this purpose.

Keywords: Modelling; Structural Identifiability; Parameter Estimation; Acetic Acid Fermentation; Vinegar.

1. Introduction

Parameter identification —or parameter estimation from experimental data— is one of the critical steps in developing models [1–3]. Some biotechnological models are especially difficult to solve because of the high-order non-linear multi-parametric equation systems and some experimental data being subject to considerable errors. This entails conducting a preliminary analysis in order to ascertain whether the specific parameters to be estimated can be uniquely determined by using a given optimization algorithm. Such an analysis must include the model structure, and also the quality and quantity of experimental data available. This study is known as identifiability and presents two different aspects:

a) Structural or theoretical identifiability, which is exclusively concerned with the complexity of the model structure for a given input–output behaviour [4].
b) Practical identifiability, which considers the quality and quantity of experimental data in addition to model structure.

This paper discusses the structural identifiability of a new model for the acetic acid fermentation process. Identifiability in biotechnological processes has been the subject of much literature. Such processes include wastewater treatments [5–8] and denitrification [9], which have been examined for structural identifiability. Some authors have also examined the theoretical identifiability of yield coefficients in generic wastewater treatment models [10] and the structural identifiability of a model for an activated sludge bioreactor [11]. Others have analysed a known model for endogenous respiration of sludge in a batch reactor under no dissolved oxygen constraints [12]. Still others have examined several models for aerobic and anaerobic processes used in wastewater treatment [13]. Structural identifiability in a microbial growth model has also been studied [14]. Also, a reparametrization technique proposed elsewhere [15] was applied to two models of a high biological and pharmacokinetic interest; the models were also dealt with in another study [16], but using differential algebra,

and the same approach was used for immunological and epidemiological models elsewhere [17]. Some authors have conducted theoretical identifiability studies on robotic models for movement control [18], as well as on mechanical systems in general [19]. A method for transforming chromatographic columns models intended to avoid overparametrization was proposed in [20] and the generating series method, which is described below, used to check for structural identifiability in the resulting model. Finally, key issues regarding the use of various statistical tools in modelling procedures were discussed in [21] and a new approach to verifying model structural identifiability involving semi-infinite programming and max–min problems was reported.

1.1. Structural identifiability

The concept of structural identifiability has been under investigation ever since it was first defined [22]. Formally, structural identifiability relies on the following definitions [23]:

- *Definition 1:* parameter θ_i of a model is structurally globally identifiable for an input class U if, and only if, for almost all $\theta \in P$ (P being the parametric space),

$$\left.\begin{array}{c} \hat{\theta} \in P \\ \hat{y}\left(\hat{\theta},t\right) = \hat{y}\left(\theta,t\right),\ \forall t > 0,\ \forall u \in U \end{array}\right\} \Rightarrow \hat{\theta}_i = \theta_i \qquad (1)$$

- *Definition 2:* parameter θ_i of a model is structurally locally identifiable for an input class U if, and only if, for almost all $\theta \in P$ a neighbourhood $V\left(\theta\right)$ exists such that

$$\left.\begin{array}{c} \hat{\theta} \in V\left(\theta\right) \subset P \\ \hat{y}\left(\hat{\theta},t\right) = \hat{y}\left(\theta,t\right),\ \forall t > 0,\ \forall u \in U \end{array}\right\} \Rightarrow \hat{\theta}_i = \theta_i \qquad (2)$$

- *Definition 3:* model $M\left(\theta\right)$ is structurally globally identifiable if, and only if, all parameters θ_i are structurally globally identifiable. Therefore, this definition also holds with local analyses.

While conceptually relevant, these definitions are scarcely useful for studying structural identifiability in practice. Also, identifiability can be examined in various ways. Thus, theoretical identifiability in linear models can be checked with several reliable tests including Laplace transforms, Taylor series expansion of observations, the Markov parameter matrix approach and methods determining the ranks of matrices A and B from the state–space representation of the model [23–25]. Confirming identifiability in non-linear models is much more complex and relatively few methods exist for verifying their theoretical identifiability. Also, such methods usually require software tools affording symbolic computation [26] and substantial computational resources.

1.2. Review of existing methods

1.2.1. Analytical methods

One of the earliest methods for structural identifiability analysis was based on series expansions of the input–output map of the model. For a specific input function or a zero-input experiment, this involves Taylor series expansion [27]. This approach assumes a state–space model for the process of the following form:

$$\begin{aligned} \frac{dx}{dt} &= f\left(x,u,\theta\right),\ \ x\left(0\right)=x_0\left(\theta\right) \\ y\left(t,\theta\right) &= h\left(x,\theta\right) \end{aligned} \tag{3}$$

where $x \in \mathbb{R}^n$ is the state vector, $\theta \in \Omega$ (Ω being an open subset of $\mathbb{R}^q$) a constant parameter vector, $u \in U$ the input vector and $y \in \mathbb{R}^r$ the output vector. This formulation assumes that, for each $\theta \in \Omega$, a connected open subset of $\mathbb{R}^n$ [denoted by $M\left(\theta\right)$) and containing $x_0\left(\theta\right)$] exists such that $f\left(\cdot,u,\theta\right)$ and $h\left(\cdot,\theta\right)$ are analytical functions on $M\left(\theta\right)$ and $x \in M\left(\theta\right)$ for any $t \geq 0$.

Outputs $y(t)$ are expanded as power series and derivatives are evaluated at a specific time (usually $t = 0$) in order to obtain mathematical expressions as simple as possible:

$$y(t) = y(0) + t\frac{dy}{dt}(0) + \frac{t^2}{2!}\frac{d^2y}{dt^2}(0) + \ldots \tag{4}$$

The derivatives in equation (4) can be calculated as model parameter functions from equations (3):

$$\begin{aligned} y(0) &= \gamma_0(\theta) \\ \frac{dy}{dt}(0) &= \gamma_1(\theta) \\ &\ldots \end{aligned} \tag{5}$$

If one assumes all derivatives in (5) to be known, then the next step involves solving the model parameters as functions of the derivatives and inputs:

$$\begin{aligned} \theta_1 &= \beta_1\left(y(0), \frac{dy}{dt}(0), \ldots, u\right) \\ \theta_2 &= \beta_2\left(y(0), \frac{dy}{dt}(0), \ldots, u\right) \\ &\ldots \\ \theta_q &= \beta_q\left(y(0), \frac{dy}{dt}(0), \ldots, u\right) \end{aligned} \tag{6}$$

Analysing the solutions of (6) allows one to draw revealing conclusions as regards structural identifiability. Thus, if a single solution exists, then the model is theoretically globally identifiable. On the other hand, if a countable set of solutions exists, then the model is locally identifiable. Finally, with an uncountable set of solutions the model is structurally unidentifiable [23,27–29].

Despite its apparent simplicity, this method is subject to two major problems, namely:

(a) Initially, the number of derivatives used for computation should be at least equal to that of parameters of the model so that equation system (6) can be solved. However, if some equations are not independent, the number of derivatives must be even

greater; therefore, no upper limit for the number of power series expansion coefficients of the outputs necessarily exists [30]. Also, finding no solution does not necessarily mean that the model is truly unidentifiable since additional independent equations may exist in the system. An upper limit has indeed been shown to exist for linear [29], bilinear [31] and polynomial systems [32,33].

(b) Provided equation system (6) can be solved, working out every parameter value can be excessively labour-intensive, even with the aid of dedicated software for symbolic computation [8,13,23,28] such as Matlab (Mathworks, Inc.), Maple (Maplesoft, Inc.) or Mathematica (Wolfram Research, Inc.)

Alternatively, one can test for local identifiability by using derivatives calculated in (5) [34]. The test is based on the jacobian matrix of the derivatives with respect to the parameters. If the rank of this matrix is less than the number of parameters, then the model is unidentifiable. The generating series approach is one other way of solving the problem that shares a number of characteristics with the previous one. This approach is based on non-linear control theory and uses the relationship between Lie derivatives and non-linear observability [5]. The method requires the model to be linear regarding the inputs $u(t)$ and allows one to extend the analysis to the entire class of bounded and measurable input functions (*e.g.* the set of piecewise continuous functions [35]):

$$\begin{aligned} \frac{dx}{dt} &= f_0(x,\theta) + \sum_{i=1}^{m} u_i(t) f_i(x,\theta), \quad x(0) = x_0(\theta) \\ y(t,\theta) &= h(x,\theta) \end{aligned} \tag{7}$$

where x is the state vector, $u = \{u_1, u_2, ..., u_m\}$ the input vector, θ the parameter vector and y the output vector. The analysis relies on the model output functions $h(x,\theta)$ and their successive Lie derivatives, $L_{f_{j0}} ... L_{f_{jk}} h(x,\theta)$, as evaluated at a specific time where simple

enough mathematical expressions can obtained (usually $t=0$). The Lie derivative along a vector field f_i is equivalent to

$$L_{f_i} = \sum_{j=1}^{n} f_{j,i}\left(x,\theta\right)\frac{\partial}{\partial x_j} \tag{8}$$

where $f_{j,i}$ is the *j*-th component of f_i.

As in the power series expansion method, all Lie derivatives at a given time are assumed to be known to calculate the solution set for the parameters. Again, any conclusions about structural identifiability will rely on the number of solutions obtained [29,36]. The mathematical expressions obtained with the generating series method are usually simpler than those provided by previous method [37].

Like the generating series method, the method based on the local state isomorphism theorem, also known as the "similarity transformation" method, assumes that the entire class of bounded and measurable functions are available [30,38]. Also, the model should be locally reduced at $x_0\left(\theta\right)$ for almost any θ. This implies that the system must fulfil the Controllability Rank Condition (CRC) and Observability Rank Condition (ORC) [39,40] at $x_0\left(\theta\right)$. These conditions determine controllability and observability in non-linear systems and can be considered an extension of those for linear systems, which were established by Kalman [41]. If these conditions are verified, the local state isomorphism theorem establishes that, if two states $\bar{x}$ and $\tilde{x}$ corresponding to two different parameter sets $\bar{\theta}$ and $\tilde{\theta}$, respectively, are considered, then the corresponding models will have the same input/output behaviour for any input $u\left(t\right)$ and $t>0$ if and only if local state isomorphism $\lambda : V \to R^n,\ \bar{x} \to \tilde{x} = \lambda\left(\bar{x}\right)$ [where V is a neighbourhood of $\bar{x}\left(0\right)$] exists which fulfils the following conditions:

$$rank\left(\left.\frac{\partial\lambda}{\partial x^T}\right)\right|_{x=\bar{x}} = n \tag{9}$$

$$\lambda\left(\bar{x}(0)\right) = \tilde{x}(0) \tag{10}$$

$$f\left(\lambda(\bar{x}),\tilde{\theta}\right) = \left.\frac{\partial\lambda}{\partial x^T}\right|_{x=\bar{x}} \cdot f\left(\lambda(\bar{x}),\bar{\theta}\right) \tag{11}$$

$$g\left(\lambda(\bar{x}),\tilde{\theta}\right) = \left.\frac{\partial\lambda}{\partial x^T}\right|_{x=\bar{x}} \cdot g\left(\lambda(\bar{x}),\bar{\theta}\right) \tag{12}$$

$$h\left(\lambda(\bar{x}),\tilde{\theta}\right) = h\left(\bar{x},\bar{\theta}\right) \tag{13}$$

If all solutions of the system (9) to (13) for $\bar{\theta}$ and λ are obtained and, for almost any $\tilde{\theta}$, the only possible solution is $\bar{\theta} = \tilde{\theta}$, $\lambda(\bar{x}) = \tilde{x}$, then the model is structurally globally identifiable [30,42].

Overall, this method is only applicable to models of low complexity. Otherwise, the mathematical expressions required to verify conditions (9) to (13) are rather complex.

Differential algebra has also been used in this context. This method reformulates model equations as linear regressions regarding parameters by using traditional algebraic operations jointly with differentiation [32,43,44]. The method is based on the Ritt-Wu algorithm [45,46], which is only applicable to polynomial or rational models. Despite these restrictions, many models in biotechnology and other areas fulfil this condition. Additional restrictions require that the models have non-zero, differentiable inputs. The application of this method to structural identifiability analysis is described elsewhere [44]; also, a software implementation was developed in [47] and an improved version can be found in [48].

Notwithstanding its elegant formulation, experience has shown that differential algebra methods not only converge on the solution very slowly, but also fail with fairly complex models [49]. Also, they share some difficulties with the previous methods as regards mathematical handling of even low-order models [44].

Finally, some methods use a combination of differential algebra and either generating series or power series expansion methodology [16].

A different set of approaches uses reparametrization methods [15,23,29,50–53] to transform the original models into structurally identifiable forms. However, these methods are not systematic and the physical meaning of the ensuing parameters can be lost through reparametrization. Although some authors have tried to reduce heuristics at some stages of the process, some intuition and experience are still needed for proper implementation.

Some reported methods are only applicable under special circumstances. Thus, some necessary and sufficient conditions for verifying structural identifiability of uncontrolled systems with or without restrictions have been reported in [14] and [48,54–56], respectively. Finally, two methods for solving structural identifiability of high-order linear systems regarding states and non-linear regarding parameters are described in [57].

1.2.2. Numerical methods

In general, applying analytical methods to practical problems is not easy. This has fostered the development of numerical methods to assess theoretical local identifiability and, in some cases, global identifiability. Thus, a method for parameter estimation by use of a global optimization algorithm based on interval analysis was proposed in [58]. The method works on data previously generated by a known model around a specific point or region of the parametric space. This procedure allows checking whether the system has only one possible solution and is therefore globally identifiable. Its greatest drawback is the long computing time needed, which is only acceptable for models with up to three dimensions. This severely restricts its use with the models frequently encountered in practice. An improved alternative was developed in [59], which uses a forward–backward contractor to reduce the number of

bisections on interval analysis. This allows the global algorithm to be more efficient and to deal with more complex problems.

One other numerical approach [60] uses DAE (Differential-Algebraic Equations) solvers to analyse structural local identifiability. Models are solved in two different ways; one involves the use of preset values for the parameters and the other considers them as unknowns. Because the number of equations will usually differ from that of parameters, the system is expanded adding as many equations as necessary taking successive derivatives with respect to time. If the system has a unique, constant solution, then it can be deemed structurally locally identifiable; otherwise (*e.g.* if it has a non-constant solution or more than one solution), then either the model is unidentifiable or unsuitable derivatives were chosen.

Finally, one other numerical method for structural identifiability analysis, which relies on similarities between this concept and observability, is described in detail [61]. It considers parameters as constant state variables with null derivatives and examines their local observability together with the remainder "real" state variables. A probabilistic semi-numerical algorithm is used for computing, within a reasonable time, locally observable variables and the number of non-observable variables. The algorithm certifies that any variables deemed observable actually are. A method for transforming non-observable models into observable models with an identical input/output behaviour was also developed; the raised models are known as symmetric models.

2. Proposed model

Basically, acetic acid fermentation is an aerobic biological reaction effected by acetic acid bacteria; these cause partial oxidation of the medium, mainly by converting ethanol into acetic acid:

$$C_2H_5OH + O_2 \rightarrow CH_3COOH + H_2O \qquad \Delta H = -493\ kJ \tag{14}$$

Vinegar production relies heavily on this stage, bacterial activity on which is markedly dependent on the particular environmental conditions.

The proposed model is based on previously reported models [62–73], and also on extensive research work on process optimization.

If one assumes a thoroughly mixed liquid phase and isothermal conditions in a semi-continuous operation mode, then the following mass balances can be written:

$$V\frac{dX_v}{dt}+X_v\frac{dV}{dt}=V\left(r_{X_c}-r_{X_d}\right) \tag{15}$$

$$V\frac{dX_d}{dt}+X_d\frac{dV}{dt}=V\left(r_{X_d}-r_{lisis}\right) \tag{16}$$

$$V\frac{dE}{dt}+E\frac{dV}{dt}=F_i\cdot E_0-V\cdot r_E \tag{17}$$

$$V\frac{dA}{dt}+A\frac{dV}{dt}=V\cdot r_A \tag{18}$$

$$V\frac{dO}{dt}+O\frac{dV}{dt}=F_i\cdot O^0+V\left[\beta\left(O^0-O\right)-r_O\right] \tag{19}$$

$$\frac{dV}{dt}=F_i \tag{20}$$

where:

X_v : viable cell concentration ($g\cdot L^{-1}$).

X_d : non-viable cell concentration ($g\cdot L^{-1}$).

E : ethanol concentration ($g\cdot L^{-1}$).

A : acetic acid concentration ($g\cdot L^{-1}$).

O : dissolved oxygen concentration ($g\cdot L^{-1}$).

V : volume (L).

F_i : volumetric feed flow rate ($L\cdot h^{-1}$).

E_0 : ethanol concentration in raw material ($g\cdot L^{-1}$).

O^0 : dissolved oxygen concentration in equilibrium with air ($g \cdot L^{-1}$).

β : constant including K_La factor, aeration flow and volume (h^{-1}).

r_{X_c} : cell growth rate (g viable cells$\cdot L^{-1} \cdot h^{-1}$).

r_{X_d} : cell death rate (g non-viable cells$\cdot L^{-1} \cdot h^{-1}$).

r_{lisis} : cell lysis rate (g lysed cells$\cdot L^{-1} \cdot h^{-1}$).

r_E : ethanol uptake rate (g. ethanol$\cdot L^{-1} \cdot h^{-1}$).

r_A : acetic acid formation rate (g. acetic acid$\cdot L^{-1} \cdot h^{-1}$).

r_O : dissolved oxygen uptake rate (g. oxygen$\cdot L^{-1} \cdot h^{-1}$).

and the following kinetic expressions can be put forward for the growth rate:

$$r_{X_c} = \mu_c \cdot X_v \tag{21}$$

$$\mu_c = \mu_{\max} \cdot f_e \cdot f_a \cdot f_o \tag{22}$$

where:

μ_c : specific cell growth rate (h^{-1}).

$\mu_{\max}$: maximum specific cell growth rate (h^{-1}).

f_e, f_a and f_o : terms representing the influence of ethanol, acetic acid and dissolved oxygen on cell growth.

Also, the following expressions can be used to weight the maximum specific cell growth, which ultimately represent its kinetics:

$$f_e = \frac{E}{E + K_{SE} + \frac{E^2}{K_{IE}}} \tag{23}$$

$$f_a = \frac{1}{1 + \left(\frac{A}{K_{IA}} \right)^4} \tag{24}$$

$$f_o = \frac{O}{O + K_{SO}} \tag{25}$$

where:

K_{SE} : ethanol saturation constant (g ethanol·L^{-1}).

K_{IE} : ethanol inhibition constant (g ethanol·L^{-1}).

K_{IA} : acetic acid inhibition constant (g acetic acid·L^{-1}).

K_{SO} : dissolved oxygen saturation constant (g oxygen·L^{-1}).

The previous ethanol limitation and cell growth inhibition expression for f_e was proposed by Andrews [74]. f_a only considers growth inhibition by acetic acid. Finally, f_o uses a simple Monod structure to represent the influence of dissolved oxygen on cell growth.

In addition, the following expressions are proposed to represent the cell death rate:

$$r_{X_d} = \mu_d \cdot X_v \tag{26}$$

$$\mu_d = \mu_d^0 \cdot f_{dE} \cdot f_{dA} \tag{27}$$

$$f_{dE} = 1 + \left(\frac{E}{K_{mE}} \right)^4 \tag{28}$$

$$f_{dA} = 1 + \left(\frac{A}{K_{mA}} \right)^4 \tag{29}$$

where:

μ_d : specific cell death rate (h^{-1}).

μ_d^0 : minimum specific cell death rate (h^{-1}).

K_{mE} : ethanol cell death constant (g ethanol·L^{-1})

K_{mA} : acetic acid cell death constant (g acetic acid·L^{-1}).

The kinetic expression for μ_d assumes a minimum specific cell death rate (μ_d^0) due mainly to cell ageing. Such a rate is weighted by f_{dE} and f_{dA} terms, which consider the influence of the substrate and product.

A first-order kinetic law exclusively dependent on the non-viable cell concentration is used to represent the cell lysis rate:

$$r_{lisis} = \mu_{lisis}^0 \cdot X_d \tag{30}$$

where:

μ_{lisis}^0: specific lysis rate (h^{-1}).

The expression for the ethanol uptake rate, r_E, obtained from equation (17) is as follows:

$$r_E = a_{E/X} \cdot r_{X_c} \tag{31}$$

where:

$a_{E/X}$: a yield factor accounting for the amount of ethanol to be consumed in order to produce the amount of energy required for biomass growth (g ethanol·g^{-1} biomass). $a_{E/X}$ is approximately 116.96 g ethanol·g^{-1} cell.

The expression for the acetic acid formation rate is

$$r_A = \frac{r_E}{Y_{E/A}} \tag{32}$$

where:

$Y_{E/A}$: stoichiometric coefficient of ethanol consumption for acetic acid formation (0.767 g ethanol·g^{-1} acetic acid).

As regards the oxygen mass balance (19), coefficient β [75] allows one to use (O^0), which is constant and roughly equal to $7.5 \cdot 10^{-3}$ g O_2 L^{-1}.

$$\beta = \frac{K_L a}{1 + \frac{K_L a}{V \cdot V_m} \cdot \frac{RT}{H}} \tag{33}$$

$$V \cdot V_m = \frac{Q}{V} \tag{34}$$

where:

$K_L a$: global volumetric mass transfer coefficient for the liquid phase (h^{-1}).

$V \cdot V_m$: relation between the input air flow and volume (h^{-1}).

R : gas constant (0.082 $atm \cdot L \cdot K^{-1} \cdot mol^{-1}$).

T : temperature (K).

H : Henry constant.

Q : input air flow rate ($L \cdot h^{-1}$).

V : volume (L).

The dissolved oxygen uptake rate can be represented by

$$r_O = \frac{r_E}{Y_{E/O}} \tag{35}$$

where:

$Y_{E/O}$: ethanol/oxygen stoichiometric coefficient (1.44 g ethanol$\cdot g^{-1}$ oxygen).

In summary, if all previous expressions for the proposed model are considered, the final state–space form used for structural identifiability analysis is as follows:

$$\begin{aligned} \frac{dX_v}{dt} &= \mu_{\max} \cdot \frac{E}{E + K_{SE} + \frac{E^2}{K_{IE}}} \cdot \frac{1}{1 + \left(\frac{A}{K_{IA}}\right)^4} \cdot \frac{O}{O + K_{SO}} \cdot X_v - \\ &- \mu_d^0 \cdot \left[1 + \left(\frac{E}{K_{mE}}\right)^4\right] \cdot \left[1 + \left(\frac{A}{K_{mA}}\right)^4\right] \cdot X_v - \frac{1}{V} X_v \cdot F_i \end{aligned} \tag{36}$$

$$\frac{dX_d}{dt} = \mu_d^0 \cdot \left[1 + \left(\frac{E}{K_{mE}}\right)^4\right] \cdot \left[1 + \left(\frac{A}{K_{mA}}\right)^4\right] \cdot X_v - \mu_{lisis}^0 \cdot X_d - \frac{1}{V} X_d \cdot F_i \tag{37}$$

$$\frac{dE}{dt} = \frac{F_i}{V}\left(E_0 - E\right) - a_{E/X} \cdot \mu_{\max} \cdot \frac{E}{E + K_{SE} + \frac{E^2}{K_{IE}}} \cdot \frac{1}{1 + \left(\frac{A}{K_{IA}}\right)^4} \cdot \frac{O}{O + K_{SO}} \cdot X_v \tag{38}$$

$$\frac{dA}{dt}=\frac{1}{Y_{E/A}}\cdot\left[a_{E/X}\cdot\mu_{\max}\cdot\frac{E}{E+K_{SE}+\frac{E^2}{K_{IE}}}\cdot\frac{1}{1+\left(\frac{A}{K_{IA}}\right)^4}\cdot\frac{O}{O+K_{SO}}\cdot X_v\right]-\frac{1}{V}A\cdot F_i \quad (39)$$

$$\frac{dO}{dt}=\frac{F_i}{V}\left(O^0-O\right)+\frac{K_La}{1+\frac{K_La\cdot V}{Q}\cdot\frac{RT}{H}}\cdot\left(O^0-O\right)-$$

$$-\frac{1}{Y_{E/O}}\cdot\left[a_{E/X}\cdot\mu_{\max}\cdot\frac{E}{E+K_{SE}+\frac{E^2}{K_{IE}}}\cdot\frac{1}{1+\left(\frac{A}{K_{IA}}\right)^4}\cdot\frac{O}{O+K_{SO}}\cdot X_v\right] \quad (40)$$

$$\frac{dV}{dt}=F_i \quad (41)$$

$$y_1=X_v \quad (42)$$

$$y_2=X_v+X_d \quad (43)$$

$$y_3=E \quad (44)$$

The parameters introduced in the model are $\mu_{\max}$, K_{SE}, K_{IE}, K_{IA}, K_{SO}, μ_d^0, K_{mE}, K_{mA} and μ_{lisis}^0.

3. Structural identifiability analysis of the proposed model

3.1. Analytical methods

The proposed model is non-linear, with 6 state variables and 9 parameters, and has the volumetric feed flow (F_i) as its sole input. Also, it only uses the viable cell concentration, total cell concentration and ethanol concentration as outputs (measured variables); neither the acetic acid concentration —which is related, though variably, to the ethanol concentration—, nor the volume —which represents the balance between the feed and product flows— are employed. Also, it was impossible to obtain reliable dissolved oxygen concentration

measures owing to the characteristics of the culture and reactor; therefore, this variable was excluded as a measured output.

The mathematical structure of the proposed model is amenable to analysis with any type of method for verifying structural identifiability (Taylor series expansion, generating series, local state isomorphism and differential algebra). However, in order to preserve the physical meaning of the parameters —if any—, we chose to exclude analytical methods involving any model transformation.

The Taylor series expansion approach was implemented in several trial-and-error cycles that provided mathematical expressions for the initial coefficients of the outputs of the power series expansions until a system of 9 independent equations in 9 unknowns (parameters) was obtained. The aim was to find a solvable system. However, we encountered several problems in the process, namely:

- Possible combinations of the output power series expansion coefficients for the systems can grow exponentially; therefore, if no solvable system is found, it is impossible to ascertain whether the model is unidentifiable or a different combination of coefficients should be tested.
- The computation time needed to obtain each set of power series expansion coefficients —which sometimes entails calculating high-order derivatives— is substantial, even if specific software tools for symbolic computation are used. As a result, the computation time required to solve each equation system is very long.

We can therefore conclude that verifying structural identifiability in the proposed model with the Taylor series expansion method is very difficult. Indeed, we reached no reliable conclusion from its application to several equation systems. By way of example, let us illustrate the process for one of the many systems studied —an unsolvable system, obviously. The nine parameters involved would require at least nine independent equations established

by computing the power series expansion coefficients. Based on the above-described measured outputs,

$$\begin{aligned} y_1 &= X_v \\ y_2 &= X_v + X_d \\ y_3 &= E \end{aligned} \tag{45}$$

where y_i are the outputs. At $t = 0$,

$$\begin{aligned} y_1(0) &= X_v(0) \\ y_2(0) &= X_v(0) + X_d(0) \\ y_3(0) &= E(0) \end{aligned} \tag{46}$$

where $X_v(0)$, $X_d(0)$ and $E(0)$ are initial conditions of these three state variables, corresponding to the viable cell concentration, total cell concentration and ethanol concentration, respectively. For simplicity, state variables can be renamed x_i and parameters p_i:

$$x_1 = X_v;\ x_2 = X_d;\ x_3 = E;\ x_4 = A;\ x_5 = O;\ x_6 = V \tag{47}$$

$$\begin{gathered} p_1 = \mu_{\max};\ p_2 = K_{SE};\ p_3 = K_{IE};\ p_4 = K_{IA};\ p_5 = K_{SO}; \\ p_6 = \mu_d^0;\ p_7 = K_{ME};\ p_8 = K_{MA};\ p_9 = \mu_{lisis}^0 \end{gathered} \tag{48}$$

Because the initial conditions are known and constant, they are independent of the model parameters. Therefore, they cannot be used to solve for the parameters in equations (5) and (6). To simplify the notation in the equations, the following can be done:

$$c_1 = x_1(0);\ c_2 = x_2(0);\ c_3 = x_3(0);\ c_4 = x_4(0);\ c_5 = x_5(0);\ c_6 = x_6(0) \tag{49}$$

Therefore, appropriate expressions for $\frac{dy_i}{dt}(0)$ are needed which can be obtained by using Maple:

$$\frac{dy_1}{dt}(0) = \frac{p_1 c_1 c_3 c_5}{\left(c_3 + p_2 + \frac{c_3^2}{p_3}\right)\left(1 + \frac{c_4^4}{p_4^4}\right)(c_5 + p_5)} - p_6\left(1 + \frac{c_3^4}{p_7^4}\right)\left(1 + \frac{c_4^4}{p_8^4}\right)c_1 - \frac{c_1 \cdot F_i(0)}{c_6} \tag{50}$$

$$\frac{dy_2}{dt}(0)=\frac{dy_1}{dt}(0)+p_6\left(1+\frac{c_3^4}{p_7^4}\right)\left(1+\frac{c_4^4}{p_8^4}\right)c_1-p_9c_2-\frac{c_2\cdot F_i(0)}{c_6} \tag{51}$$

$$\frac{dy_3}{dt}(0)=-\frac{a_{E/X}\cdot p_1c_1c_3c_5}{\left(c_3+p_2+\frac{c_3^2}{p_3}\right)\left(1+\frac{c_4^4}{p_4^4}\right)(c_5+p_5)}+\frac{F_iE_0-c_3\cdot F_i(0)}{c_6} \tag{52}$$

Equations(50), (51) and (52) are only three of at least nine independent equations required to build a solvable system. Therefore, additional output power series expansion coefficients must be obtained. We computed $\frac{d^2y_i}{dt^2}(0)$ and $\frac{d^3y_i}{dt^3}(0)$ to this end, but their mathematical expressions were too complex to be shown here. Also, the ensuing system provided no solution with Maple software. The Maple code used is shown in Appendix A. These results led us to try an alternative method.

Such a method was the generating series method, which is very similar to the previous one but was expected to be less computationally demanding based on the ability to isolate output terms. The proposed model for the acetic acid fermentation process can be rewritten as follows in terms of equations (7):

$$\begin{gathered}
\frac{dX_v}{dt}=r_{Xc}-r_{Xd}-\frac{X_v}{V}F_i\\
\frac{dX_d}{dt}=r_{Xd}-r_{lisis}-\frac{X_d}{V}F_i\\
\frac{dE}{dt}=-r_{EX}+\frac{E_0-E}{V}F_i\\
\frac{dA}{dt}=r_A-\frac{A}{V}F_i\\
\frac{dO}{dt}=\beta\left(O^0-O\right)-r_{OE}+\frac{O^0-O}{V}F_i\\
\frac{dV}{dt}=F_i\\
y_1=X_v\\
y_1=X_v+X_d\\
y_1=E
\end{gathered} \tag{53}$$

Therefore,

$$f_0(x,\theta)=\begin{bmatrix} r_{Xc}-r_{Xd} \\ r_{Xd}-r_{lisis} \\ -r_{EX} \\ r_A \\ \beta\left(O^0-O\right)-r_{OE} \\ 0 \end{bmatrix} \qquad f_1(x,\theta)=\begin{bmatrix} -\dfrac{X_v}{V} \\ -\dfrac{X_d}{V} \\ \dfrac{E_0-E}{V} \\ -\dfrac{A}{V} \\ \dfrac{O^0-O}{V} \\ 1 \end{bmatrix} \tag{54}$$

$$h(x,\theta)=\begin{bmatrix} X_v \\ X_v+X_d \\ E \end{bmatrix}$$

This method exhibited the same problems as the power series expansion. Although Lie derivatives gave simpler mathematical expressions, the equation system was unsolvable, so no reliable conclusions about structural identifiability in the model could be drawn. By way of example, below is described one of many systems studied here. Because solving for the parameters required using at least nine equations, we chose to use the following Lie derivatives: (evaluated at $t=0$): $L_{f_0}h_1(0)$, $L_{f_0}h_2(0)$, $L_{f_0}h_3(0)$, $L_{f_1}L_{f_0}h_1(0)$, $L_{f_1}L_{f_0}h_2(0)$, $L_{f_1}L_{f_0}h_3(0)$, $L_{f_1}L_{f_1}L_{f_0}h_1(0)$, $L_{f_1}L_{f_1}L_{f_0}h_2(0)$ and $L_{f_1}L_{f_1}L_{f_0}h_3(0)$. We used the nomenclature introduced in equations (47), (48) and (49), and computed the first three derivatives with Maple:

$$L_{f_0}h_1(0)=\left(\frac{\partial h_1}{\partial x}\cdot f_0\right)(0)=\frac{p_1c_1c_3c_5}{\left(c_3+p_2+\dfrac{c_3^2}{p_3}\right)\left(1+\dfrac{c_4^4}{p_4^4}\right)(c_5+p_5)}-p_6\left(1+\frac{c_3^4}{p_7^4}\right)\left(1+\frac{c_4^4}{p_8^4}\right)c_1 \tag{55}$$

$$L_{f_0}h_2(0)=\left(\frac{\partial h_2}{\partial x}\cdot f_0\right)(0)=\frac{p_1c_1c_3c_5}{\left(c_3+p_2+\dfrac{c_3^2}{p_3}\right)\left(1+\dfrac{c_4^4}{p_4^4}\right)(c_5+p_5)}-p_9c_2 \tag{56}$$

$$L_{f_0} h_3(0) = \left(\frac{\partial h_3}{\partial x} \cdot f_0 \right)(0) = \frac{a_{E/X} p_1 c_1 c_3 c_5}{\left(c_3 + p_2 + \frac{c_3^2}{p_3} \right) \left(1 + \frac{c_4^4}{p_4^4} \right) (c_5 + p_5)} \tag{57}$$

The mathematical expressions for the other derivatives were exceedingly complex and are thus not shown here. Using Maple code provided no unique solution for the nine Lie derivative system.

Applying the local state isomorphism theorem method required initially verifying the controllability and observability rank conditions in the initial state. This entailed analysing the ranks of matrices built from Lie brackets and Lie derivatives, respectively.

Lie brackets were computed in this way:

$$\left[f_0(x,\theta), f_1(x,\theta) \right] = \frac{\partial f_1}{\partial x} f_0 - \frac{\partial f_0}{\partial x} f_1 \tag{58}$$

where $f_0(x,\theta)$ and $f_1(x,\theta)$ are vector fields.

Verifying the controllability rank condition required building an $n \times n$ matrix (n being the model order) by computing as many Lie brackets as needed to fill its columns. If the rank of the matrix in the initial state and for any value of each model parameter is n, then the controllability condition will be fulfilled. Otherwise, either the model does not satisfy this condition or a different matrix needs to be built from other Lie brackets.

Similarly, verifying the observability rank condition required building an $n \times n$ matrix, the rows of which would be filled with gradients of the Lie derivatives with respect to the state vector. Likewise, if the rank of the matrix in the initial state and for any value of each model parameter is n, then the observability condition will be fulfilled.

The difficulty involved in computing the Lie brackets and Lie derivatives for the proposed model, and their infinite potential combinations, made the problem unsolvable in practice.

In any case, we tested some controllability and observability matrices for the proposed model, but no solid evidence of their ranks being 6 ($=n$ from equation (53)) could be obtained.

By way of example of the verification of the controllability rank condition, we built the following controllability matrix from the model equations (54)

$$\left[f_0 \quad f_1 \quad [f_0, f_1] \quad \left[[f_0, f_1], f_0\right] \quad \left[[f_0, f_1], f_1\right] \quad \left[\left[[f_0, f_1], f_1\right], f_1\right]\right] \tag{59}$$

The mathematical expressions for these Lie brackets are too complex to be shown here. Their complexity precluded confirming whether the matrix rank for any parameter vector would be 6.

We used the following matrix, among others, to verify the observability rank condition:

$$\begin{bmatrix} \nabla_x h_1 \\ \nabla_x h_2 \\ \nabla_x h_3 \\ \nabla_x \left(L_{f_0} h_1\right) \\ \nabla_x \left(L_{f_0} h_2\right) \\ \nabla_x \left(L_{f_0} h_3\right) \end{bmatrix} \tag{60}$$

The only simple expressions in this case were those for $\nabla_x h_1$, $\nabla_x h_2$ and $\nabla_x h_3$:

$$\begin{aligned} \nabla_x h_1 &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \nabla_x h_2 &= \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \nabla_x h_3 &= \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \end{aligned} \tag{61}$$

Other gradients were mathematically more complex and are thus not shown. As with the controllability rank condition, we could not ascertain whether the rank of the observability matrix was 6. Therefore, this method also failed to afford clear-cut conclusions about structural identifiability in the proposed model.

Although the differential algebra method was applicable in theory, the complexity of the model required extremely elaborate algebraic and differentiation operations involving a vast

number of potential choices. Also, although several alternatives were used to obtain a set of linear equations for the parameters, no definite conclusions could be reached.

3.2. Numerical methods

Based on the results of the analytical methods, only numerical methods could provide an effective way of solving the problem. Although they only afford conclusions about structural local identifiability, numerical methods can be applied to much more complex models and with a higher computing efficiency.

The first method studied here is described in detail elsewhere [60]. We used Ecosimpro software (Empresarios Agrupados, S.A., Madrid, Spain) as the numerical solver for application to the proposed model. Although the computational cost was virtually negligible, non-linear algebraic loops and/or structural singularities often caused convergence problems. In any case, the main problem was to make an appropriate choice of equations in order to obtain a solvable system. If the system concerned has no solution, then alternative equations (*e.g.* expressions containing additional derivatives) must be tested. If no solution is eventually found, this does not necessarily mean that the model is structurally unidentifiable. Application of this method to the proposed model provided no solvable system. The initial system was of the following form [see nomenclature in equations (38) and (39)]:

$$\dot{x}_1 = \frac{\overline{p}_1 x_1 x_3 x_5}{\left(x_3 + \overline{p}_2 + \frac{x_3^2}{\overline{p}_3}\right)\left(1 + \frac{x_4^4}{\overline{p}_4^4}\right)\left(x_5 + \overline{p}_5\right)} - \overline{p}_6\left(1 + \frac{x_3^4}{\overline{p}_7^4}\right)\left(1 + \frac{x_4^4}{\overline{p}_8^4}\right)x_1 - \frac{x_1 F_i}{x_6} \tag{62}$$

$$\dot{x}_2 = \overline{p}_6\left(1 + \frac{x_3^4}{\overline{p}_7^4}\right)\left(1 + \frac{x_4^4}{\overline{p}_8^4}\right)x_1 - \overline{p}_9 x_2 - \frac{x_2 F_i}{x_6} \tag{63}$$

$$\dot{x}_3 = -\frac{a_{E/X}\overline{p}_1 x_1 x_3 x_5}{\left(x_3 + \overline{p}_2 + \frac{x_3^2}{\overline{p}_3}\right)\left(1 + \frac{x_4^4}{\overline{p}_4^4}\right)\left(x_5 + \overline{p}_5\right)} + \frac{F_i\left(E_0 - x_3\right)}{x_6} \tag{64}$$

$$\dot{x}_4 = \frac{a_{E/X}\overline{p}_1 x_1 x_3 x_5}{\left(x_3 + \overline{p}_2 + \frac{x_3^2}{\overline{p}_3}\right)\left(1 + \frac{x_4^4}{\overline{p}_4^4}\right)\left(x_5 + \overline{p}_5\right)Y_{E/A}} - \frac{x_4 F_i}{x_6} \tag{65}$$

$$\dot{x}_5 = \frac{K_L a\left(O^0 - x_5\right)}{1 + \frac{K_L a \cdot x_6}{Q}\cdot\frac{RT}{H}} \cdot \frac{a_{E/X}\overline{p}_1 x_1 x_3 x_5}{\left(x_3 + \overline{p}_2 + \frac{x_3^2}{\overline{p}_3}\right)\left(1 + \frac{x_4^4}{\overline{p}_4^4}\right)\left(x_5 + \overline{p}_5\right)Y_{E/O}} + \frac{F_i\left(O^0 - x_5\right)}{x_6} \tag{66}$$

$$\dot{x}_6 = F_i \tag{67}$$

$$y_1 = x_1 \tag{68}$$

$$y_2 = x_1 + x_2 \tag{69}$$

$$y_3 = x_3 \tag{70}$$

$$\dot{\tilde{x}}_1 = \frac{p_1 \tilde{x}_1 \tilde{x}_3 \tilde{x}_5}{\left(\tilde{x}_3 + p_2 + \frac{\tilde{x}_3^2}{p_3}\right)\left(1 + \frac{\tilde{x}_4^4}{p_4^4}\right)\left(\tilde{x}_5 + p_5\right)} - p_6\left(1 + \frac{\tilde{x}_3^4}{p_7^4}\right)\left(1 + \frac{\tilde{x}_4^4}{p_8^4}\right)\tilde{x}_1 - \frac{\tilde{x}_1 F_i}{\tilde{x}_6} \tag{71}$$

$$\dot{\tilde{x}}_2 = p_6\left(1 + \frac{\tilde{x}_3^4}{p_7^4}\right)\left(1 + \frac{\tilde{x}_4^4}{p_8^4}\right)\tilde{x}_1 - p_9\tilde{x}_2 - \frac{\tilde{x}_2 F_i}{\tilde{x}_6} \tag{72}$$

$$\dot{\tilde{x}}_3 = -\frac{a_{E/X} p_1 \tilde{x}_1 \tilde{x}_3 \tilde{x}_5}{\left(\tilde{x}_3 + p_2 + \frac{\tilde{x}_3^2}{p_3}\right)\left(1 + \frac{\tilde{x}_4^4}{p_4^4}\right)\left(\tilde{x}_5 + p_5\right)} + \frac{F_i\left(E_0 - \tilde{x}_3\right)}{\tilde{x}_6} \tag{73}$$

$$\dot{\tilde{x}}_4 = \frac{a_{E/X} p_1 \tilde{x}_1 \tilde{x}_3 \tilde{x}_5}{\left(\tilde{x}_3 + p_2 + \frac{\tilde{x}_3^2}{p_3}\right)\left(1 + \frac{\tilde{x}_4^4}{p_4^4}\right)\left(\tilde{x}_5 + p_5\right)Y_{E/A}} - \frac{\tilde{x}_4 F_i}{\tilde{x}_6} \tag{74}$$

$$\dot{\tilde{x}}_5 = \frac{K_L a\left(O^0 - \tilde{x}_5\right)}{1 + \frac{K_L a \cdot \tilde{x}_6}{Q}\cdot\frac{RT}{H}} \cdot \frac{a_{E/X} p_1 \tilde{x}_1 \tilde{x}_3 \tilde{x}_5}{\left(\tilde{x}_3 + p_2 + \frac{\tilde{x}_3^2}{p_3}\right)\left(1 + \frac{\tilde{x}_4^4}{p_4^4}\right)\left(\tilde{x}_5 + p_5\right)Y_{E/A}} + \frac{F_i\left(O^0 - \tilde{x}_5\right)}{\tilde{x}_6} \tag{75}$$

$$\dot{\tilde{x}}_6 = F_i \tag{76}$$

$$y_1 = \tilde{x}_1 \tag{77}$$

$$y_2 = \tilde{x}_1 + \tilde{x}_2 \tag{78}$$

$$y_3 = \tilde{x}_3 \tag{79}$$

where x and $\tilde{x}$ are state vectors, $\overline{p}$ is the known parameter vector, p the unknown parameter vector and y the output vector. Equations (62) to (79) contain 21 unknowns. Nevertheless, they can be reduced to 18 since, from equations (68) to (70) and (77) to (79), it is clear that $\tilde{x}_1 = x_1$, $\tilde{x}_2 = x_2$ and $\tilde{x}_3 = x_3$. It is therefore necessary to add six equations in order to obtain a solvable system. Such equations must be obtained by differentiating with respect to time some of the equations (71) to (75) on the assumption of constancy in the parameters. According to the proponent of this method [60], the equations to be differentiated should be chosen in accordance with two essential criteria: the equations should contain as many parameters as possible and excessively complex, high-order derivatives —numerical integration of which can be rather difficult— should be avoided. By way of example, we chose to derive equations (72) to (75) once and equation (71) twice —using Maple in both cases. The ensuing mathematical expressions were too complex to be shown here. We used the following initial conditions for integration:

$$x_1(0) = 0.15;\quad x_2(0) = 0.02;\quad x_3(0) = 11.6;\quad x_4(0) = 105;\quad x_5(0) = 3\cdot 10^{-3};\quad x_6(0) = 4 \tag{80}$$

and the following parameter values (falling in the preset range):

$$\overline{p}_1 = 0.6;\ \overline{p}_2 = 3.5;\ \overline{p}_3 = 3;\ \overline{p}_4 = 90;\ \overline{p}_5 = 3\cdot 10^{-4};\ \overline{p}_6 = 2.5\cdot 10^{-5};\ \overline{p}_7 = 12.5;\ \overline{p}_8 = 35;\ \overline{p}_9 = 0.5 \tag{81}$$

The solver exhibited convergence problems from the first integration step and provided no solution. Alternative equation systems, different initial conditions and trial parameter values were tested, but no unique solution for p_i could be found. Therefore, despite its simplicity and low computational cost, the method failed as a means for verifying structural identifiability in our model.

Finally, we used a software tool developed as a Maple package by its proponent (http://www2.lifl.fr/~sedoglav/Software/) to implement the method introduced in [61]. This

method, which features virtually negligible computation times, provided a positive answer as regards theoretical local identifiability in the proposed model. The software tool certified that all model parameters were observable.

4. Conclusions

The model proposed for the acetic acid fermentation process has been examined for structural identifiability by using various analytical and numerical methods. The former included Taylor series expansion, generating series, local state isomorphism and differential algebra, all of which exhibited problems such as difficult implementation and substantial computational costs for models of small-medium dimension and complexity frequently found in practice.

Numerical methods were easier to implement than the analytical methods, but only afforded verification of local structural identifiability.

The results of our study of structural identifiability in the proposed model for the acetic acid fermentation process allow us to draw the following conclusions:

- No clear-cut criteria exist for choosing the best potential method for structural identifiability verification of a dynamic model. This is consistent with the variety of methods used for this purpose to date.
- The substantial computational cost involved entails the use of symbolic software to analyse models for theoretical identifiability. In some cases, computational complexity is so high that the problem is intractable, even with powerful software.

Only one of the studied methods provided positive results affording the conclusion that the proposed model is structurally locally identifiable.

While theoretical identifiability analysis is important, model parameters can only be estimated uniquely after practical identifiability has been verified. If a model is practically

identifiable, then it should be structurally identifiable as well; on the other hand, a model may not be practically identifiable either because of the lack of appropriate available experimental data or the model structure itself. In this situation, it is important to also have information about the theoretical identifiability of the model.

Acknowledgements

The authors wish to acknowledge funding of this research (AGL2002-01712, PET2006_0827) by Spain's Ministry of Science and Technology (MCyT) and Ministry of Education and Science. Also, they wish to thank the firm Vinagres y Salsas, S.A., Grupo SOS (Córdoba, Spain), for valuable help and advice.

Appendix A

```
x1:=f1(t): x2:=f2(t): x3:=f3(t): x4:=f4(t): x5:=f5(t): x6:=f6(t):

fe:=x3/(x3+p2+x3^2/p3):

fa:=1/(1+(x4/p4)^4):

fo:=x5/(x5+p5):

rxc:=p1*fe*fa*fo*x1:

rxd:=p6*(1+(x3/p7)^4)* (1+(x4/p8)^4)*x1:

rlysis:=p9*x2:

re:=aex*rxc:

ra:=re/Yea:

ro:=re/Yeo:

bt:=Kla/(1+RTH*x6*Kla/Q):

x6p:=Fi:

x1p:=rxc-rxd-x1*x6p/x6:

x2p:=rxd-rlysis-x2*x6p/x6:

x3p:=-re+(Fi*E0-x3*x6p)/x6:

x4p:=ra-x4*x6p/x6:

x5p:=bt*(O0-x5)-ro+(Fi*O0-x5*x6p)/x6:

x1pp:=diff(x1p,t):

x2pp:=diff(x2p,t):

x3pp:=diff(x3p,t):

x1ppp:=diff(x1pp,t):

x2ppp:=diff(x2pp,t):

x3ppp:=diff(x3pp,t):

solve({x1p=c1,  x2p=c2,  x3p=c3,  x1pp=c4,  x2pp=c5,  x3pp=c6,  x1ppp=c7,

x2ppp=c8, x3ppp=c9}, [p1,p2,p3,p4,p5,p6,p7,p8,p9])
```

References

[1] J.R. Banga, E. Balsa-Canto, C.G. Moles, and A.A. Alonso, Improving food processing using modern optimization methods. Trends in Food Science & Technology 14 (2003) 131–144.

[2] C. Gutiérrez, Optimización global de procesos de la industria alimentaria y biotecnológica, Dept. Ingeniería de Sistemas y Automática, University of Vigo, Vigo, Spain, 2003.

[3] K. Schittkowski, Parameter estimation in systems of nonlinear equations. Numerische Mathematik 68 (1994) 129–142.

[4] E. Walter, Identifiability of Parametric Models, Pergamon, Oxford, 1987.

[5] D. Dochain, and P.A. Vanrolleghem, Dynamical modelling and estimation in wastewater treatment processes, IWA Publishing, London, 2001.

[6] D. Dochain, P.A. Vanrolleghem, and M. Vandaele, Structural identifiability of biokinetic models of activated-sludge respiration. Water Research 29 (1995) 2571–2578.

[7] B. Petersen, Calibration, identifiability and optimal experimental design of activated sludge models, University of Ghent, Ghent, Belgium, 2000.

[8] B. Petersen, K. Gernaey, and P.A. Vanrolleghem, Practical identifiability of model parameters by combined respirometric–titrimetric measurements. Water Science and Technology 43 (2001) 347–355.

[9] S.V. Bourrel, J.P. Babary, S. Julien, M.T. Nihtila, and D. Dochain, Modelling and identification of a fixed-bed denitrification bioreactor. Systems Analysis Modelling Simulation 30 (1998) 289–309.

[10] L. Chen, and G. Bastin, Structural identifiability of the yield coefficients in bioprocess models when the reaction rates are unknown. Mathematical Biosciences 132 (1996) 35–67.

[11] S. Julien, J.P. Babary, and P. Lessard, Identifiability and identification of an activated sludge process model. Systems Analysis Modelling Simulation 37 (2000) 481–499.

[12] K.J. Keesman, H. Spanjers, and G. van Straten, Analysis of endogenous process behaviour in activated sludge. Biotechnology and Bioengineering 57 (1998) 155–163.

[13] N. Noykova, T.G. Muller, M. Gyllenberg, and J. Timmer, Quantitative analyses of anaerobic wastewater treatment processes: Identifiability and parameter estimation. Biotechnology and Bioengineering 78 (2002) 89–103.

[14] L. Denis-Vidal, and G. Joly-Blanchard, An easy to check criterion for (un)identifiability of uncontrolled systems and its applications. IEEE Transactions on Automatic Control 45 (2000) 768–771.

[15] N.D. Evans, and M.J. Chappell, Extensions to a procedure for generating locally identifiable reparameterisations of unidentifiable systems. Mathematical Biosciences 168 (2000) 137–159.

[16] G. Margaria, E. Riccomagno, M.J. Chappell, and H.P. Wynn, Differential algebra methods for the study of the structural identifiability of rational function state–space models in the biosciences. Mathematical Biosciences 174 (2001) 1–26.

[17] G. Margaria, E. Riccomagno, and L.J. White, Structural identifiability analysis of some highly structured families of state space models using differential algebra. Journal of Mathematical Biology 49 (2004) 433–454.

[18] T.J. Debus, P.E. Dupont, and R.D. Howe, Distinguishability and identifiability testing of contact state models. Advanced Robotics 19 (2005) 545–566.

[19] R. Serban, and J.S. Freeman, Identification and identifiability of unknown parameters in multibody dynamic systems. Multibody System Dynamics 5 (2001) 335–350.

[20] F. Couenne, C. Jallut, and M. Tayakout-Fayolle, On minimal representation of heterogeneous mass transfer for simulation and parameter estimation: Application to

breakthrough curves exploitation. Computers & Chemical Engineering 30 (2005) 42–53.

[21] S.P. Asprey, and S. Macchietto, Statistical tools for optimal dynamic model building. Computers & Chemical Engineering 24 (2000) 1261–1267.

[22] R. Bellman, and K.J. Aström, On structural identifiability. Mathematical Biosciences 7 (1970) 329–339.

[23] K.R. Godfrey, and J.J. Distefano, Identifiability of model parameters. in: E. Walter, (Ed.), Identifiability of Parametric Models, Pergamon Press, Oxford, 1987, pp 1–19.

[24] C. Cobelli, A. Lepschy, and G.R. Jacur, Identifiability of compartmental systems and related structural properties. Mathematical Biosciences 44 (1979) 1–18.

[25] C. Cobelli, and G. Toffolo, Identifiability from parameter bounds – Structural and numerical aspects. Mathematical Biosciences 71 (1984) 237–243.

[26] A. Raksanyi, Y. Lecourtier, E. Walter, and A. Venot, Identifiability and distinguishability testing via computer algebra. Mathematical Biosciences 77 (1985) 245–266.

[27] H. Pohjanpalo, System Identifiability based on power-series expansion of solution. Mathematical Biosciences 41 (1978) 21–33.

[28] M.J. Chappell, K.R. Godfrey, and S. Vajda, Global identifiability of the parameters of non-linear systems with specified inputs – a comparison of methods. Mathematical Biosciences 102 (1990) 41–73.

[29] E. Walter, Identifiability of State Space Models, Springer, New York, 1982.

[30] S. Vajda, K.R. Godfrey, and H. Rabitz, Similarity transformation approach to identifiability analysis of nonlinear compartmental models. Mathematical Biosciences 93 (1989) 217–248.

[31] S. Vajda, Structural Identifiability of dynamical systems. International Journal of Systems Science 14 (1983) 1229–1247.

[32] S. Diop, and M. Fliess, On nonlinear observability. in: C. Commault, (Ed.), Proceedings of the First European Control Conference, Hermes, Grenoble, Switzerland, 1991, pp 152–157.

[33] S. Vajda, Deterministic identifiability and algebraic invariants for polynomial systems. IEEE Transactions on Automatic Control 32 (1987) 182–184.

[34] H. Pohjanpalo, Identifiability of deterministic differential models in state space, Technical Report 56, Technical Research Centre of Finland, Espoo, 1982.

[35] E. Walter and Y. Lecourtier, Global approaches to identifiability testing for linear and nonlinear state–space models. Mathematics and Computers in Simulation 24 (1982) 472–482.

[36] Y. Lecourtier, F. Lamnabhi-Lagarrigue, and E. Walter, Volterra and generating power series approaches to identifiability testing. in: E. Walter, (Ed.), Identifiability of Parametric Models, Pergamon Press, Oxford, 1987.

[37] E. Walter, and L. Pronzato, On the identifiability and distinguishability of nonlinear parametric models. Mathematics and Computers in Simulation 42 (1996) 125–134.

[38] E.T. Tunali, and T.J. Tarn, New results for identifiability of nonlinear systems. IEEE Transactions on Automatic Control 32 (1987) 146–154.

[39] R. Hermann, and A.J. Krener, Nonlinear controllability and observability. IEEE Transactions on Automatic Control 22 (1977) 728–740.

[40] H.J. Sussman, Existence and uniqueness of minimal realizations of nonlinear systems. Mathematical Systems Theory 10 (1977) 263–284.

[41] R.E. Kalman, A new approach to linear filtering and prediction problems. Transactions of the ASME – Journal of Basic Engineering 82 (1960) 35–45.

[42] S. Vajda, and H. Rabitz, State isomorphism approach to global identifiability of nonlinear systems. IEEE Transactions on Automatic Control 34 (1989) 220–223.

[43] M. Fliess, and S.T. Glad, An algebraic approach to linear and nonlinear control. in: H.L. Trentelman, and J.C. Willems, (Eds.), Essays on Control: Perspectives in the Theory and its Applications, Boston, MA, 1993, pp 223–265.

[44] L. Ljung, and T. Glad, On global identifiability for arbitrary model parametrizations. Automatica 30 (1994) 265–276.

[45] J.F. Ritt, Differential equations from the algebraic standpoint, American Mathematical Society Colloquium Publications, 1932.

[46] W.J. Wu, On zeros of algebraic equations – an application of Ritt principle. Kexue Tongbao 31 (1986) 1–5.

[47] D.M. Wang, An implementation of the characteristic set method in MAPLE. in: J. Pfalzgraf, and D. Wang, (Eds.), Automated Practical Reasoning: Algebraic Approaches, Springer, New York, 1995, pp 187–201.

[48] L. Denis-Vidal, G. Joly-Blanchard, and C. Noiret, Some effective approaches to check the identifiability of uncontrolled nonlinear systems. Mathematics and Computers in Simulation 57 (2001) 35–44.

[49] T.G. Müller, Modelling complex systems with differential equations, Albert Ludwigs Universität, Freiburg, 2002.

[50] M.J. Chappell, and R.N. Gunn, A procedure for generating locally identifiable reparameterisations of unidentifiable non-linear systems by the similarity transformation approach. Mathematical Biosciences 148 (1998) 21–41.

[51] J.J. Distefano, Complete parameter bounds and quasi-identifiability conditions for a class of unidentifiable linear systems. Mathematical Biosciences 65 (1983) 51–68.

[52] R.N. Gunn, M.J. Chappell, and V.J. Cunningham, Reparameterisation of unidentifiable systems using the Taylor series approach. in: D.A. Linkens, and E. Carson, (Eds.), Proceedings of the 3rd IFAC Symposium on Modelling and Control in Biomedical Systems, Pergamon Press, Oxford, 1997, pp 247–252.

[53] E. Walter, and Y. Lecourtier, Unidentifiable compartmental models – What to do. Mathematical Biosciences 56 (1981) 1–25.

[54] N.D. Evans, M.J. Chapman, M.J. Chappell, and K.R. Godfrey, Identifiability of uncontrolled nonlinear rational systems. Automatica 38 (2002) 1799–1805.

[55] N.D. Evans, R.J. Errington, M. Shelley, G.P. Feeney, M.J. Chapman, K.R. Godfrey, P.J. Smith and M.J. Chappell, A mathematical model for the in vitro kinetics of the anti-cancer drug topotecan. Mathematical Biosciences 189 (2004) 185–217.

[56] N.D. Evans, L.J. White, M.J. Chapman, M.J. Chappell, K.R. Godfrey, The structural identifiability of the susceptible infected recovered model with seasonal forcing. Mathematical Biosciences 194 (2005) 175–197.

[57] H.G.M. Dotsch, and P.M.J. VandenHof, Test for local structural identifiability of high-order non-linearly parametrized state space models. Automatica 32 (1996) 875–883.

[58] I. Braems, L. Jaulin, M. Kieffer, and E. Walter, Guaranteed numerical alternatives to structural identifiability testing, Proceedings of the 40th IEEE Conference on Decision and Control, Orlando, CA, 2001.

[59] E. Walter, I. Braems, L. Jaulin, and M. Kieffer, Guaranteed numerical computation as an alternative to computer algebra for testing models for identifiability. in: R. Alt, (Ed.), Numerical Software with Result Verification, Springer–Verlag, Berlin, 2004, pp 124–131.

[60] M. Gerdin, Using DAE solvers to examine local identifiability for linear and nonlinear systems, Proceedings of the 14th IFAC Symposium of System Identification (SYSID-2006), 2006.

[61] A. Sedoglavic, A probabilistic algorithm to test local algebraic observability in polynomial time. Journal of Symbolic Computation 33 (2002) 735–755.

[62] R. Bar, J.L. Gainer, and D.J. Kirwan, An unusual pattern of product inhibition – Batch acetic acid fermentation. Biotechnology and Bioengineering 29 (1987) 796–798.

[63] J.M. González-Saiz, C. Pizarro, and D. Garrido-Vidal, Evaluation of kinetic models for industrial acetic acid fermentation: proposal of a new model optimized by genetic algorithms. Biotechnology Progress 19 (2003) 599–611.

[64] T. Ito, H. Sota, H. Honda, K. Shimizu, and T. Kobayashi, Efficient acetic acid production by repeated fed-batch fermentation using 2 fermenters. Applied Microbiology and Biotechnology 36 (1991) 295–299.

[65] R.K. Kruppa, and D. Vortmeyer, Transient growth and product-formation kinetics of acetic acid bacteria. Bioprocess Engineering 20 (1999) 545–551.

[66] A. Nanba, A. Tamura, and S. Nagai, Synergistic effects of acetic acid and ethanol on the growth of Acetobacter sp. Journal of Fermentation Technology 62 (1984) 501–505.

[67] Y.S. Park, H. Ohtake, M. Fukaya, H. Okumura, Y. Kawamura, and K. Toda, Enhancement of acetic acid production in a high cell density culture of Acetobacter aceti. Journal of Fermentation and Bioengineering 68 (1989) 315–319.

[68] Y.S. Park, H. Ohtake, K. Toda, M. Fukaya, H. Okumura, and Y. Kawamura, Acetic acid production using a fermenter equipped with a hollow fiber filter module. Biotechnology and Bioengineering 33 (1989) 918–923.

[69] Y.S. Park, H. Ohtake, M. Fukaya, H. Okumura, Y. Kawamura, and K. Toda, Effects of dissolved oxygen and acetic acid concentrations on acetic acid production in

continuous culture of Acetobacter aceti. Journal of Fermentation and Bioengineering 68 (1989) 96–101.

[70] Y.S. Park, M. Fukaya, H. Okumura, Y. Kawamura, and K. Toda, Production of acetic acid by a repeated batch culture with cell recycle of Acetobacter aceti. Biotechnology Letters 13 (1991) 271–276.

[71] Y.S. Park, K. Toda, M. Fukaya, H. Okumura, and Y. Kawamura, Production of a high-concentration acetic-acid by Acetobacter aceti using a repeated fed-batch culture with cell recycling. Applied Microbiology and Biotechnology 35 (1991) 149–153.

[72] Y.S. Park, and K. Toda, Multistage biofilm reactor for acetic-acid production at high concentration. Biotechnology Letters 14 (1992) 609–612.

[73] L.E. Romero, J.M. Gomez, I. Caro, and D. Cantero, Kinetic model for growth of Acetobacter aceti in submerged culture. Chemical Engineering Journal and the Biochemical Engineering Journal 54 (1994) B15–B24.

[74] J.F. Andrews, A mathematical model for continuous culture of microorganisms utilizing inhibitory substrates. Biotechnology and Bioengineering 10 (1968) 707–723.

[75] F.J. Nieto, Algunos aspectos de la tecnología de la fermentación acética. in: C. Llaguno, and M.C. Polo, (Eds.), El Vinagre de Vino, CSIC, Madrid, Spain, 1991.