Preprints of the
9th International Symposium on Advanced Control of Chemical Processes
The International Federation of Automatic Control
June 7-10, 2015, Whistler, British Columbia, Canada

Parametric identifier of metabolic networks based on robust differentiation

A. M. Sepúlveda Gálvez * **J. A. Badillo-Corona** ** **I. Chairez****

* *SEPI-UPIBI, Instituto Politécnico Nacional, Mexico D.F. 07340, Mexico (e-mail: sega_asg@hotmail.com)*
** *Department of Bioprocesses-UPIBI, Instituto Politécnico Nacional , Mexico City 07340, Mexico (e-mail: jbadilloc@ipn.mx, isaac_chairez@yahoo.com)*

Abstract: This study proposes a new robust parametric identifier for systems that describe the dynamical behavior of metabolic networks. This identifier implements a robust parallel differentiator that recovers the time derivative for all the metabolites involved in the metabolic network. The differentiator is based on the well-known Super-Twisting Algorithm which is applied over the variation of each metabolites that is included within the metabolic network. The derivatives are fed into a parallel nonlinear least mean square scheme that is successful in recovering the parameters that characterizes the metabolic network. This identifier is applied to a simplified 22-reactions metabolic network of hydrogen production in *Escherichia coli* using glucose as substrate. The metabolic network is simulated with parameters obtained from previous studies and they are recovered using the parametric identifier proposed in this study. All the parameter are recovered with less than 5% error.

Keywords: Metabolic network, Parameter identification, Robust differentiation, Super-Twisting differentiator, Hydrogen production, *Escherichia coli*.

1. INTRODUCTION

Research on renewable fuels and environmentally-friendly fuels has substantially increased in the past few years due to a reduction on global oil reserves, the high dependence on fossil fuels and the strong public interest in protecting global climate (Ghosh et al., 2013; Mathews et al., 2010). Recently, hydrogen (H_2) has been projected as an energy carrier and substitute for fossil fuels because of its ease of production, its non-polluting nature and its large energy content per mass (142 $MJkg^{-1}$ H_2) compared to existing renewable energy sources such as ethanol and biodiesel (Manish et al., 2007; Patel and Kalia, 2013).

The production of H_2 by facultative anaerobic organisms such as *Escherichia coli* (*E. coli*) is a characteristic of mixed acid fermentation. Efforts to improve H_2 production have focused on identification and engineering of oxygen-tolerant hydrogenases, improvement in H_2 molar yields, development of efficient H_2 separation techniques from biorreactor head-space and metabolic pathway engineering (Fan et al., 2009). Metabolic pathways can be analyzed with a focus on their stoichiometry (time invariant) or their kinetics. Otherwise, dynamics of this kind of systems can be expressed as a set of Ordinary Differential Equations (ODEs), but one drawback of representing a genetic or a metabolic network with ODEs is the large number of parameters that are unknown and must be estimated (Chou and Voit, 2009).

Due to the lack of information regarding the molecular mechanisms that govern the system coupled with the large size of the metabolic or genetic networks that are modeled, the parametric estimation becomes an important issue for the modeling of biologic systems (Berthoumieux et al., 2011; Ashyraliyev et al., 2009). The adjustment of parameters in mathematical models has several difficulties. One of them is that some parameters may compensate the lack of sensitivity of other parameters resulting in arbitrariness in the specification of their values. Another difficulty is that different values of the parameters may be equally consistent with the experimental data. A third difficulty is that to find the optimal parameter values of the model the exploration of a large space of values may be required (Fernández Slezak et al., 2010).

The majority of current methods for parameter estimation, in principle, formulate the parameter estimation problem as a nonlinear optimization problem with differential-algebraic constraints that describe dynamics of biochemical networks. Some techniques as the Volterra-series approach and the method of maximum likelihood are already been investigated as solution to the problem of nonlinear estimation. Other approaches to approximation include extended Kalman Filter and sequential Monte Carlo methods (Sun et al., 2008), as well as the sliding mode algorithms.

Yan and Edwards (2007) establish that the sliding mode approach can be divided in two different steps: the first one is the design of a sliding surface in which the system have the desired performance if it is restricted to the surface; the second step consists in the design of a control law with variable structure which drives the system trajectories in finite time to the sliding surface.

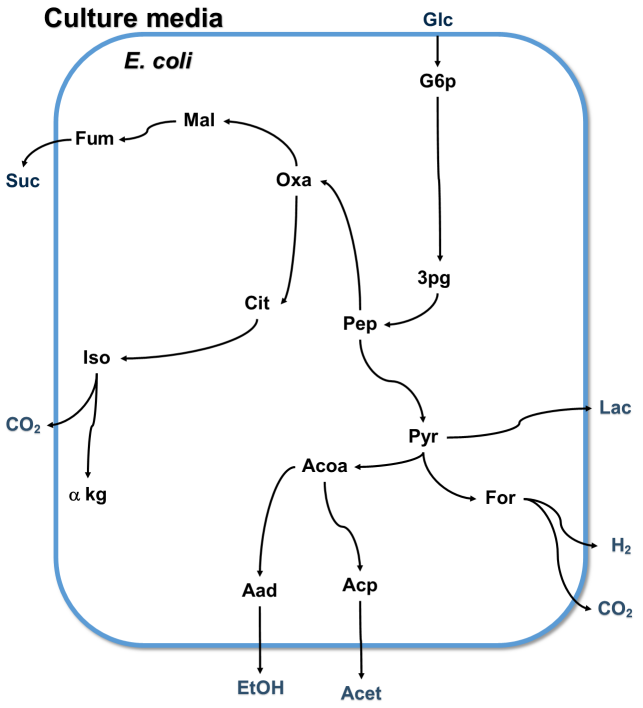

Fig. 1. Metabolic pathway of H_2 production in *E. coli*.

Other approach called higher order sliding modes, has been proposed in order to reduce the chattering phenomenon present in the conventional sliding mode. In this approach the sign function acts on its higher order time derivative and presents 190 a better accuracy with respect to discrete sampling time than conventional sliding modes (Defoort et al., 2009).

The aim of this work is to propose a parameter estimation algorithm based on high order sliding mode for an ordinary differential equations system representing hydrogen production in a strain of *E. coli*.

2. PROBLEM FORMULATION

According to databases that contain genetic information of *E. coli* (e.g. Kegg and EcoCyc), the metabolic pathway for producing H_2 in *E. coli* involves 23 reactions with glucose (Glc) as substrate and, succinate (Suc), ethanol (EtOH), lactate (Lac), carbon dioxide (CO_2),(H_2) and acetate (Acet) as extracellular products with the concomitant generation of biomass (X). A thermodynamic analysis, according to the methodology proposed by Henry et al. (2006), was made to reduce the number of reactions and to prove the feasibility of the proposed reactions. This analysis indicates that the minimal number of reactions to work with, are 18 with 22 metabolites involved (Figure 1).

One of the main challenges when ODEs are used for modeling metabolic networks is the parameter identification. Because the metabolic pathways of biological systems are described by a high number of reactions, the number of parameters to be estimated is also high. Therefore, the implementation of algorithms that allow the simultaneous identification from existing information is necessary.

Moreover, the complexity of experimental techniques allows to obtain only a small number of experimental data that may not contain all the dynamic information from the system. To overcome this obstacle, in this work it is proposed to use the Super-Twisting Algorithm (STA) as a differentiator to retrieve all the dynamic information of the system from experimental data and, from it, identify all the parameters. The proposed modeling strategy is tested over the mathematical model of the metabolic pathway for hydrogen production in *E. coli*.

$$
\left.\begin{aligned}
\frac{dX}{dt} &= \frac{\mu_{max} Glc}{K_s + Glc} X \\
\frac{dGlc}{dt} &= -q_S \frac{\mu_{max} Glc}{K_s + Glc} X \\
\frac{dG6p}{dt} &= \frac{\nu^{r1}_{max} Glc}{K^{r1}_M + Glc} - 2\frac{\nu^{r2}_{max} G6p}{K^{r2}_M + G6p} \\
\frac{dPdgp}{dt} &= 2\frac{\nu^{r2}_{max} G6p}{K^{r2}_M + G6p} - \frac{\nu^{r3}_{max} Pdgp}{K^{r2}_M + Pdgp} \\
\frac{dPep}{dt} &= \frac{\nu^{r3}_{max} Pdgp}{K^{r3}_M + Pdgp} - \frac{\nu^{r4}_{max}(0.909Pep)}{K^{r4}_M + (0.909Pep)} - \\
&\quad \cdot \frac{\nu^{r5}_{max}(0.032Pep)}{K^{r5}_M + (0.032Pep)} \\
\frac{dPyr}{dt} &= \frac{\nu^{r4}_{max}(0.909Pep)}{K^{r4}_M + (0.909Pep)} - \frac{\nu^{r13}_{max}(0.969Pyr)}{K^{r13}_M + (0.969Pyr)} - \\
&\quad \cdot \frac{\nu^{r12}_{max}(0.0303Pyr)}{K^{r12}_M + (0.0303Pyr)} \\
\frac{dOxa}{dt} &= \frac{\nu^{r5}_{max}(0.032Pep)}{K^{r5}_M + (0.032Pep)} - \frac{\nu^{r9}_{max}(0.949Oxa)}{K^{r9}_M + (0.949Oxa)} - \\
&\quad \cdot \frac{\nu^{r6}_{max}(0.025Oxa)}{K^{r6}_M + (0.025Oxa)} \\
\frac{dCit}{dt} &= \frac{\nu^{r6}_{max}(0.025Oxa)}{K^{r6}_M + (0.025Oxa)} - \frac{\nu^{r7}_{max} Cit}{K^{r7}_M + Cit} \\
\frac{dIso}{dt} &= \frac{\nu^{r7}_{max} Cit}{K^{r7}_M + Cit} - \frac{\nu^{r8}_{max} Iso}{K^{r8}_M + Iso} \\
\frac{d\alpha kg}{dt} &= \frac{\nu^{r8}_{max} Iso}{K^{r8}_M + Iso} \\
\frac{dMal}{dt} &= \frac{\nu^{r9}_{max}(0.949Oxa)}{K^{r9}_M + (0.949Oxa)} - \frac{\nu^{r10}_{max} Mal}{K^{r10}_M + Mal} \\
\frac{dFum}{dt} &= \frac{\nu^{r10}_{max} Mal}{K^{r10}_M + Mal} - \frac{\nu^{r11}_{max} Fum}{K^{r10}_M + Fum} \\
\frac{dAcoa}{dt} &= \frac{\nu^{r13}_{max}(0.969Pyr)}{K^{r13}_M + (0.969Pyr)} - \frac{\nu^{r15}_{max}(0.570Acoa)}{K^{r15}_M + (0.969Acoa)} \\
&\quad \cdot - \frac{\nu^{r14}_{max}(0.426Acoa)}{K^{r14}_M + (0.426Acoa} \\
\frac{dFor}{dt} &= \frac{\nu^{r13}_{max}(0.969Pyr)}{K^{r13}_M + (0.969Pyr)} - \frac{\nu^{r17}_{max} For}{K^{r17}_M + For} \\
\frac{dSuc}{dt} &= \frac{\nu^{r11}_{max} Fum}{K^{r11}_M + Fum} \\
\frac{dLac}{dt} &= \frac{\nu^{r12}_{max}(0.03035Pyr)}{K^{r12}_M + (0.03035Pyr)} \\
\frac{dAcp}{dt} &= \frac{\nu^{r14}_{max}(0.4262Acoa)}{K^{r14}_M + (0.4262Acoa)} - \frac{\nu^{r18}_{max} Acp}{K^{r18}_M + Acp} \\
\frac{dAad}{dt} &= \frac{\nu^{r15}_{max}(0.5709Acoa)}{K^{R15}_M + (0.9695Acoa)} - \frac{\nu^{r16}_{max} Aad}{K^{r16}_M + Aad} \\
\frac{dEtOH}{dt} &= \frac{\nu^{r16}_{max} Aad}{K^{r16}_M + Aad} \\
\frac{dAcet}{dt} &= \frac{\nu^{r18}_{max} Acp}{K^{r18}_M + Acp} \\
\frac{dH_2}{dt} &= \frac{\nu^{r17}_{max} For}{K^{r17}_M + For} \\
\frac{dCO_2}{dt} &= \frac{\nu^{r8}_{max} Iso}{K^{r8}_M + Iso} + \frac{\nu^{r17}_{max} For}{K^{r17}_M + For}
\end{aligned}\right\} \quad (1)
$$

3. MODEL DESCRIPTION

Consider that model (1) represents the metabolic network. This model, based on ODEs, was proposed yo represent the 18 enzymatic reactions of the metabolic pathway for (H_2) production in *E. coli*.

Model (1) was used to verify that the proposed methodology works for this particular biological system and contain 22 ODEs that describe the dynamics of the metabolites considered in the metabolic pathway under study. The model takes into account the stoichiometric balance of each reaction. The first equation in (1) represents the growth of the bacterial strain (X) and is modeled according to the Monod model taking Glc as substrate .

The equation has two parameters that correspond to the maximum specific growth rate of *E. coli* (μ_{max}) in h^{-1} and, K_s is the affinity constant in Mol. The second equation in (1) represents the time course of Glc consumption. In addition to the parameters μ_{max} and K_s, this equation has a parameter q_S that represents the consumption rate of Glc, hence the reason of μ_S represented with a negative sign. These two equations represent the part of the microbial kinetics. The remaining equations in (1) represent the reactions that occur inside the cell namely the metabolic pathway dynamical representation. The time course of all metabolite were modeled as the difference between the mass produced and the mass consumed in each enzymatic reaction. The rates of the enzymatic reactions are considered as a Michaelis-Menten dynamic, so every single equation has, at least, two parameters: the maximum velocity of reaction (ν_{max}) and Michaelis constant (K_M). All parameters for ν_{max} were taken from reaction rates obtained by Aispuro Castro (2011) and K_M values were taken from the EcoCyc database.
The mathematical model has a total number of 39 parameters that must be identified. In this study, it is assumed that every metabolite of this model can be measured, however, the information obtained from the model only contains specific discrete time experimental data. The following section proposes a methodology that allows to recover the entire dynamic of each metabolite from the model.

4. ROBUST PARAMETRIC IDENTIFIER

Let us assume that each differential equation of the model presented above satisfies:

$$\frac{dz_i(t)}{dt} = \frac{f^i(z(t),\alpha_1^i)}{g^i(z(t),\alpha_2^i)} + \zeta^i(z,t), \quad z_i \in \mathbb{R}$$

where $z_i \in \mathbb{R}$ is either variable in the same model and n is the number of variables. The vector $z \in \mathbb{R}^n$ represents all the variables involved in the metabolic network used in this study. Nonlinear functions $f^i : \mathbb{R}^{n+P_i^1} \longrightarrow \mathbb{R}$, $g^i : \mathbb{R}^{n+P_i^2} \longrightarrow \mathbb{R}$ satisfy:

$$|f^i(z_1,\cdot) - f^i(z_2,\cdot)| \leq L^i\|z_1 - z_2\| \ \ \forall z_1, z_2 \in \mathbb{R}^n$$
$$|g^i(z,\alpha)| \neq 0 \ \ \forall z \in \mathbb{R}^n, \ \ \forall \alpha \in \mathbb{R}^{P_i^2}$$

In this study, it is assumed that all the functions f^i and g^i are linearly parametrized by α_i^1 and α_i^2, respectively. Therefore, a couple of nonlinear functions ψ_1^i and ψ_2^i exists such that:

$$f^i(z) = \left[\alpha_1^i\right]^\top \psi_1^i(z) + \varphi_1^i(z)$$
$$g^i(z) = \left[\alpha_2^i\right]^\top \psi_2^i(z) + \varphi_2^i(z)$$

In the unrealistic case when the derivative of the whole vector z can be measured, the non-perturbed ($\zeta^i = 0$) ODE described in Moreno and Osorio (2012) can be alternatively represented as:

$$y^i(t) = \left[\theta^i\right]^\top x^i(t) + \zeta^i(z(t),t)$$

where

$$\theta^i = \begin{bmatrix} \alpha_1^i \\ \alpha_2^i \end{bmatrix} \quad x^i(t) = \begin{bmatrix} \psi_1^i(z) \\ \psi_2^i(z)\dfrac{dz_i(t)}{dt} \end{bmatrix}$$

$$y^i(t) = \left[\frac{dz_i(t)}{dt}\varphi_2^i(z) - \varphi_1^i(z) \right]$$

This form obeys the regular linear regression form. The solution for the parametric identification can be obtained as in Ljung (1999):

$$\bar{\theta^i}(t) = \left(\int_0^t x^i(\tau)\left[x^i(\tau)\right]^\top d\tau \right)^{-1} \left(\int_0^t x^i(\tau) y^i(\tau) d\tau \right)$$

However, the previous equation needs the time derivative of z_i. As it was stated, it is not usual to have this derivative on line and simultaneously for all the variables involved in the metabolic network. This is the main motivation to propose a robust (with respect to parametric uncertainties and external perturbations) differentiator to approximate $\frac{dz_i(t)}{dt}$. In this study the Super Twisting Algorithm (STA) is considered as a feasible method to solve this problem. STA as differentiator satisfies the structure:

$$\begin{aligned} \frac{d\hat{z}_a^i(t)}{dt} &= \hat{z}_b^i(t) + k_1^i|\delta^i(t)|^{\frac{1}{2}}\text{sign}(\delta^i(t)) \\ \frac{d\hat{z}_b^i(t)}{dt} &= k_2^i\text{sign}(\delta^i(t)) \end{aligned}$$

Here, the differentiator error is represented by the variable δ^i which is defined as $\delta^i = z^i - \hat{z}_a^i$. Indeed, this algorithm serves as differentiator of the uncertain signal $z^i(t)$. This fact can be supported considering that signal $z^i(t)$ can be represented as:

$$\begin{aligned} z^i(t) &= z_a^i(t); \ \ \frac{dz^i(t)}{dt} = z_b^i(t) \\ \frac{dz_a^i(t)}{dt} &= z_b^i(t) \\ \frac{dz_b^i(t)}{dt} &= \frac{d^2 z_a^i(t)}{dt^2}, \quad |\frac{d^2 z^i(t)}{dt^2}| < z^{+,i} \quad z^{+,i} \in \mathbb{R}^+ \end{aligned}$$

Then, the differentiation error $\delta_1^i = z_a^i - \hat{z}_a^i, \delta_2^i = z_b^i - \hat{z}_b^i$ satisfies:

$$\frac{d\delta_1^i(t)}{dt} = \delta_2^i(t) + k_1^i|\delta_1^i(t)|^{\frac{1}{2}} sign(\delta_1^i(t))$$
$$\frac{d\delta_2^i(t)}{dt} = k_2^i sign(\delta_1^i(t)) + \frac{d^2 z^i(t)}{dt^2}$$

For the purposes of this study, the sign function was used as follows:

$$sign(z) = \begin{cases} -1 & if \ z < 0 \\ \in [-1,+1] & if \ z = 0 \\ 1 & if \ z > 0 \end{cases}$$

Table 1. Real and estimated model parameters and relative error between them

Parameter	Real Value (h^{-1})	Estimated Value (h^{-1})	Relative error (%)	Parameter	Real Value (mM)	Estimated Value (mM)	Relative error (%)
μ_{max}	0.768*	0.760	1.002	K_S	1.65*	1.666	0.96
q_S	0.36*	0.03636	1.0				
Parameter	Real Value (mMh^{-1})	Estimated Value (mMh^{-1})	Absolute error	Parameter	Real Value (mM)	Estimated Value (mM)	Absolute error
ν_{max}^{R1}	3.14*	2.983	5.0	K_M^{R1}	0.78**	0.819	5.0
ν_{max}^{R2}	2.99*	2.841	4.98	K_M^{R2}	0.42**	0.4323	2.92
ν_{max}^{R3}	2.07*	1.967	4.97	K_M^{R3}	0.097**	0.09215	5.0
ν_{max}^{R4}	2.46*	2.337	5.0	K_M^{R4}	0.5**	0.5018	0.36
ν_{max}^{R5}	0.091*	0.08645	5.0	K_M^{R5}	0.07**	0.0735	5.0
ν_{max}^{R6}	0.00229*	0.00229	0	K_M^{R6}	0.07**	0.0735	5.0
ν_{max}^{R7}	0.00229*	0.00229	0	K_M^{R7}	11**	11	0
ν_{max}^{R8}	0.00229*	0.00229	0	K_M^{R8}	0.5**	0.4776	0
ν_{max}^{R9}	0.0864*	0.08208	5	K_M^{R9}	2.6**	2.73	5.0
ν_{max}^{R10}	0.0864*	0.08208	5.0	K_M^{R10}	1.1**	1.155	5
ν_{max}^{R11}	0.0864*	0.08208	5	K_M^{R11}	0.5**	0.525	5
ν_{max}^{R12}	0.0709*	0.0673	4.97	K_M^{R12}	26.5**	27.82	4.98
ν_{max}^{R13}	2.39*	2.271	4.97	K_M^{R13}	2.0**	2.028	1.4
ν_{max}^{R14}	1.02*	0.969	5.0	K_M^{R14}	0.0095**	0.00963	1.42
ν_{max}^{R15}	1.36*	1.295	4.77	K_M^{R15}	0.13**	0.133	2.31
ν_{max}^{R16}	1.36*	1.292	5.0	K_M^{R16}	0.5**	0.5092	1.84
ν_{max}^{R17}	2.39*	2.271	4.97	K_M^{R17}	26**	27.3	5.0
ν_{max}^{R18}	1.02*	0.9727	4.63	K_M^{R18}	1.5**	1.533	2.2

* Taken from (Aispuro Castro, 2011)
** Taken from EcoCyc database.

According to the result presented by Moreno and Osorio (2012), it can be proven that

$$|\delta_1^i| = 0, \quad |\delta_2^i| = 0 \quad \forall t \geqslant T^{*,i}$$

where

$$T^{*,i} = \sqrt{\frac{0.5V^i(0)}{2}}$$

The sense of V^i satisfies the expression $V^i(\xi^i) = \left[\xi^i\right]^\top P^i \xi^i$ with $\left[\xi^i\right]^\top = \left[|\delta_1^i|^{\frac{1}{2}} sign(\delta_1^i) \;\; \delta_2^i\right]$. Each matrix $P^i \in \mathbb{R}^{2\times 2}$ is a positive definite and symmetric matrix.

Finally if the variables $y^i(t)$ and $x^i(t)$ are changed by $\hat{x}^i$ and $\hat{y}^i$ the parametric identification can be solved as

$$\theta^{*,i}(t) = \left(\int_0^t \hat{x}^i(\tau)\left[\hat{x}^i(\tau)\right]^\top d\tau\right)^{-1}\left(\int_0^t \hat{x}^i(\tau)\hat{y}^i(\tau)d\tau\right) \tag{2}$$

5. NUMERICAL SIMULATIONS

This section presents the application of the STA as a differentiator to obtain the approximation of the time derivative signal of all states of the suggested model (1) and the subsequent parameter identification using the corresponding least squares algorithm.
Simulations of the model were made in Matlab/Simulink, using a fixed step numerical integration algorithm (ODE4-Runge-Kutta) with an integration step of 0.0001 hours. Initial conditions and reference parameters for X, Glc, H_2, Suc, EtOH, Lac, CO_2, Acet and values of μ_{max}, K_S, q_S and all ν_{max} parameters were taken from Aispuro Castro (2011). Initial conditions for the other compounds and K_M values were taken form the EcoCyc database.
The results of these simulations were taken as the real values of the dynamic for each compound. This strategy was used to test parametric identification algorithm proposed in this study. However the derivatives of all compounds were obtained using a perturbation signal produced by a random bounded value that tries to serve as an artificial representation of the experimentally error that is usual in this kind of systems.

Figure 2 shows the performance of the STA as differentiator when comparing the states of the three principal compounds and their derivatives. The convergence of the states are reached at finite time (less than 1.0 hour). This is a remarkable characteristic because for biological systems, the enzymatic reaction occurs in small periods of time.

Using the derivatives obtained with STA, a robust nonlinear least squares algorithm is used to obtain the parameter values presented in (4). The algorithm is based on the well-known Hammerstein-Wiener recursive parameter identification method. This algorithm was implemented in parallel and on-line to obtain the solution of the problem presented in equation 2.

Table 4 shows the whole set of identified parameters for all the reaction considered in the model (1). The third and sixth columns show the percentage error obtained as a result of applying the identification algorithm using the estimated derivative. None of these errors is over 5.0%. This is regularly considered as an acceptable result considering the relevance of the information obtained in the corresponding study.

Gains k_1 and k_2 of the STA were fixed individually for each compound looking for the minimal error between the real values and the STA trajectories. Although figure 3 shows

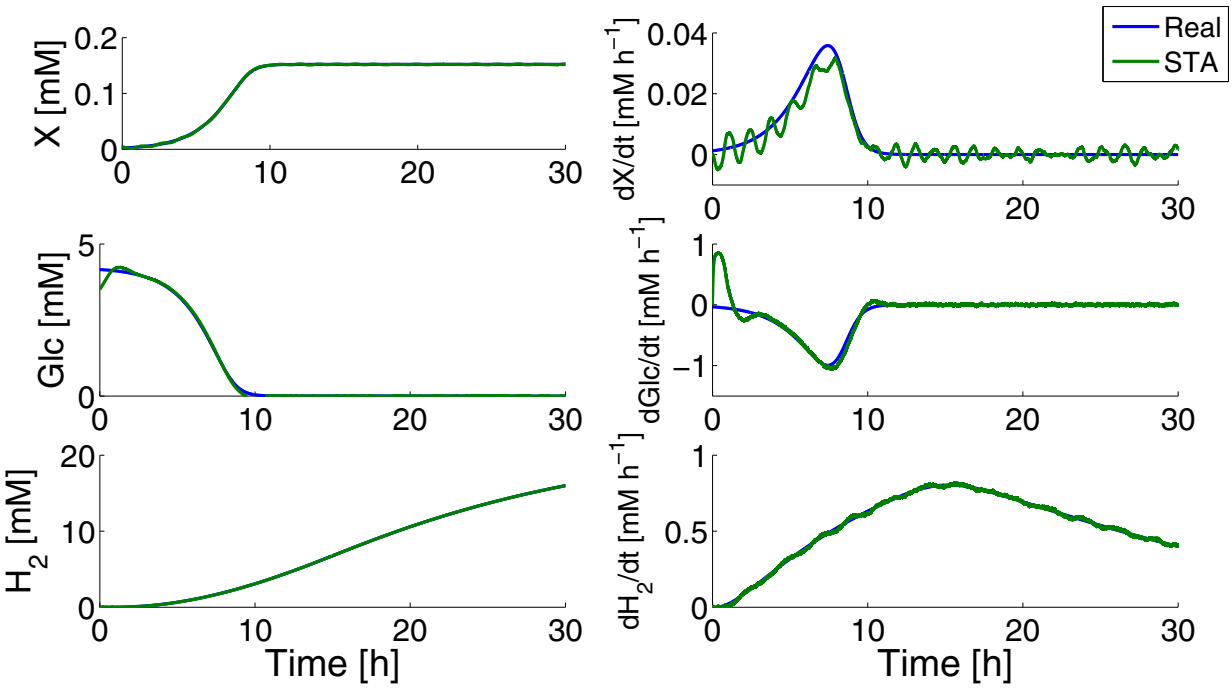

Fig. 2. Comparison between states and derivatives obtained with STA and real ones.

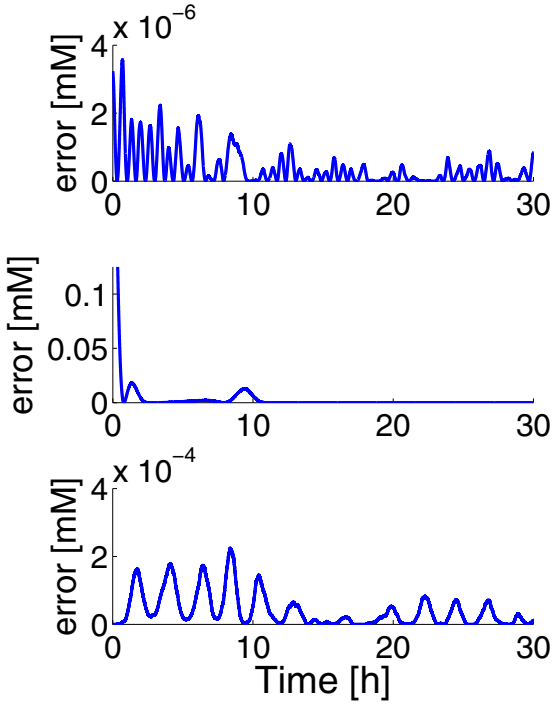

Fig. 3. Error values between states obtained with STA and real ones.

that error values are different from zero, this fact does not represent a problem because the parametric identification algorithm works in a more efficient way if a lot of dynamic information is present (persistent excitation condition).

The parameter values obtained are used in a new simulation where the system is tested with the estimated parameters. This strategy is used to obtain a validation of the technique proposed in this study. These results are compared with the values obtained in the simulation of the metabolic model. Figure 4 shows the comparison between the dynamic of the real states and derivatives and those obtained after the parameter identification (API). The difference between dynamics of states and derivatives can be considered acceptable for the type of biological systems analyzed in this study.

6. CONCLUSIONS

This study describes a methodology based on the application of the high order STA to recover an approximation of the time derivative information of variables involved in the model of a metabolic network. These approximate derivatives are used to solve a parallel parametric identification algorithm. This mixed scheme is used to recover the relevant parameters that characterize the time course evolution of the metabolic network. This methodology produces an efficient parameter identification considering the large quantity of elements that must be calculated to characterize the metabolic network. As one can understand, the application of STA yields to improve the reconstruction of experimental information based on the proposed model without losing the sense of biological character for each parameter. Knowing the unrealistic scenario of measure all metabolites of a metabolic pathway, this technique seems to be a simple but useful option, in the modeling of biological systems described by a large number of ODEs with many parameters involved, as a immediate step after the implementation of a state observer in order to know the time course of all metabolites from the experimental measurement of a few ones.

REFERENCES

Aispuro Castro, K.R. (2011). *Determinación de los flujos*

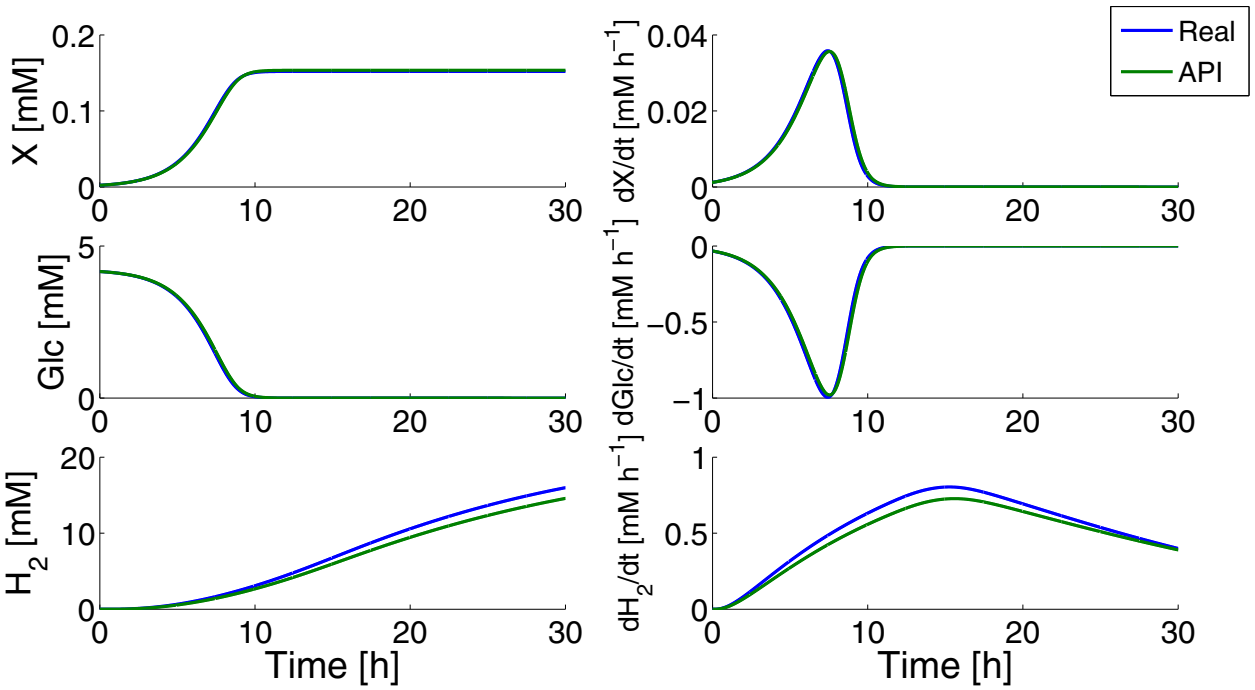

Fig. 4. Comparison between the time course of real states and derivatives and those after the parameter identification (API).

metabólicos en la producción de hidrógeno. Ph.D. thesis, Unidad Profesional Interdisciplinaria de Biotecnología-IPN.

Ashyraliyev, M., Fomekong-Nanfack, Y., Kaandorp, J.a., and Blom, J.G. (2009). Systems biology: parameter estimation for biochemical models. *The FEBS journal*, 276(4), 886–902. doi: 10.1111/j.1742-4658.2008.06844.x.

Berthoumieux, S., Brilli, M., de Jong, H., Kahn, D., and Cinquemani, E. (2011). Identification of metabolic network models from incomplete high-throughput datasets. *Bioinformatics*, 27(13), i186–95. doi:10.1093/bioinformatics/btr225.

Chou, I.C. and Voit, E.O. (2009). Recent developments in parameter estimation and structure identification of biochemical and genomic systems. *Mathematical biosciences*, 219(2), 57–83. doi: 10.1016/j.mbs.2009.03.002.

Defoort, M., Floquet, T., Kokosy, A., and Perruquetti, W. (2009). A novel higher order sliding mode control scheme. *Systems & Control Letters*, 58(2), 102–108. doi:10.1016/j.sysconle.2008.09.004.

Fan, Z., Yuan, L., and Chatterjee, R. (2009). Increased hydrogen production by genetic engineering of Escherichia coli. *PloS one*, 4(2), e4432. doi:10.1371/journal.pone.0004432.

Fernández Slezak, D., Suárez, C., Cecchi, G.a., Marshall, G., and Stolovitzky, G. (2010). When the optimal is not the best: parameter estimation in complex biological models. *PloS one*, 5(10), e13283. doi:10.1371/journal.pone.0013283.

Ghosh, D., Bisaillon, A., and Hallenbeck, P.C. (2013). Increasing the metabolic capacity of Escherichia coli for hydrogen production through heterologous expression of the Ralstonia eutropha SH operon. *Biotechnology for biofuels*, 6(1), 122. doi:10.1186/1754-6834-6-122.

Henry, C.S., Jankowski, M.D., Broadbelt, L.J., and Hatzimanikatis, V. (2006). Genome-Scale Thermodynamic Analysis of Escherichia coli Metabolism. *Biophysical Journal*, 90(4), 1453–1461. doi: 10.1529/biophysj.105.071720.

Ljung, L. (1999). *System Identification: Theory for the User (2nd Edition)*. Prentice Hall, 2 edition.

Manish, S., Venkatesh, K., and Banerjee, R. (2007). Metabolic flux analysis of biological hydrogen production by Escherichia coli. *International Journal of Hydrogen Energy*, 32(16), 3820–3830. doi:10.1016/j.ijhydene.2007.03.033.

Mathews, J., Li, Q., and Wang, G. (2010). Characterization of hydrogen production by engineered Escherichia coli strains using rich defined media. *Biotechnology and Bioprocess Engineering*, 15(4), 686–695. doi:10.1007/s12257-009-3139-4.

Moreno, J.A. and Osorio, M. (2012). Strict Lyapunov functions for the Super-Twisting algorithm. *IEEE transactions on automatic control*, 57(4), 1035–1040.

Patel, S.K.S. and Kalia, V.C. (2013). Integrative Biological Hydrogen Production: An Overview. *Indian journal of microbiology*, 53(1), 3–10. doi:10.1007/s12088-012-0287-6.

Sun, X., Jin, L., and Xiong, M. (2008). Extended kalman filter for estimation of parameters in nonlinear state-space models of biochemical networks. *PloS one*, 3(11), e3758. doi: 10.1371/journal.pone.0003758.

Yan, X.G. and Edwards, C. (2007). Nonlinear robust fault reconstruction and estimation using a sliding mode observer. *Automatica*, 43(9), 1605–1614. doi:10.1016/j.automatica.2007.02.008.