



The influence of experimental data quality on parametric identification accuracy: sulfate-reducing process

V. Peña-Caballero¹, P. A. López-Pérez^{2**}, M. I. Neria-González³, O. Madrigal-Serrano², R. Aguilar-López^{1*}

¹Department of Biotechnology & Bioengineering, CINVESTAV-IPN, Instituto Politécnico Nacional 2508, San Pedro Zacatenco, 07360 Mexico, D.F.

²Department of Energy, Universidad Politécnica Metropolitana de Hidalgo, Boulevard de acceso a Tolcayuca N° 1009 Exhacienda San Javier, Tolcayuca, Hgo. C.P. 43860.

³Chemical and Biochemical Engineering Division, Tecnológico de Estudios Superiores de Ecatepec, Tecnológico, 55210, Valle de Anáhuac, Ecatepec de Morelos, Estado de Mexico.

* E-mail.- raguilar@cinvestav.mx

** E-mail.- paperez@upmh.edu.mx

Resumen

Este documento se enfoca a la investigación del modelado de un proceso sulfato reductor utilizando a *Desulfovibrio alaskensis* 6SR la cual creció usando agua congénita y Postgate C como medio de cultivo; para los dos casos se implementaron modelos cinéticos no estructurados como el de Monod y Levenspiel para representar el crecimiento de *D. alaskensis* 6SR en sus respectivos medios. Se evaluó la influencia de la calidad de los datos experimentales sobre la exactitud de la identificación paramétrica del modelado. Después de la estimación de los parámetros, los intervalos de confianza se evaluaron a través de un método numérico basado en la matriz de información de Fisher (MIF). Con esta información de la MIF, se observó que la fiabilidad del valor de los parámetros estimados aumenta al disminuir el error de medición de los datos y al aumentar la frecuencia de muestreo, las ilustraciones numéricas consideradas en este trabajo permiten mostrar una mejora utilizando la MIF.

Abstract

In this paper a modeling approach of the sulfate-reducing process with sulfate reducing bacteria *Desulfovibrio alaskensis* 6SR, using congenial water and medium Postgate C as culture media was investigated; for this purpose, a typical unstructured kinetic models Monod and Levenspiel were used to modeling the kinetics of *D. alaskensis* 6SR growth, respectively. The influence of experimental data quality on parametric identification accuracy was evaluated on modeling approach. After estimating the parameters, the confidence intervals were assessed through a numerical method based on the Fisher Information Matrix. With this information from the FIM, it was observed that reliability of value of estimated parameters decreases with increasing measurement error of the data. Numerical illustrations considered in this work show considerable improvement of the new FIM estimator.

Keywords: sulfate-reducing process, modeling, *D. alaskensis* 6SR, Fisher Information Matrix

1. Introduction

Sulfate-reducing processes (SRP) have now been widely studied due to their multiple applications as bioremediation processes. For example its application in metal removal, recovery of valuable metals as metallic sulfide (López *et al.* 2013; Gallegos-Garcia *et al.* 2009), as well as in the removal and reduction of heavy metals (e.g., Immobilization of heavy metals through microbial mediated reduction and precipitation is now of considerable interest) (White *et al.* 2003).

Sulfate-reducing bacteria (SRB) of the genus *Desulfovibrio* are the most widely studied due to their potential application in the process of bioremediation, with *D. vulgaris* and *D. desulfuricans* species which have major applications in the bioremediation of soil and water with different metals (Acha *et al.* 2012). But in recent years, new species of this genus has been isolated and identified as species *D. alaskensis* (Feio *et al.* 2004) and *D. alaskensis* 6SR (Neria-González *et al.* 2006). that were isolated in pipelines carrying oil. *D. alaskensis* 6SR has shown high resistance to heavy metals like cadmium and chromium and high

concentrations of hydrogen sulfide (López *et al.* 2013). To exploit these advantages of bacteria *D. alaskensis* 6SR in bioremediation processes, an important aspect is to develop a mathematical model for this biological system for specific process in order to analyze the system behavior, in particular the response to external stimuli and perturbations; understanding biological processes; suggest new hypotheses and experiments to test them. However, modeling of these processes, i.e., developing a mathematical model, is directly related to the identification of parameters (or the selection) present in the equations that characterize the process dynamics. Model parameters are estimated through minimization algorithms with respect to experimental data and, afterwards, the calibration of model can be used for process improvement (e.g., in process design or process control). Therefore, the estimation of precise parameters values is a major issue in the construction of biological models, because the model behavior may be strongly dependent on the parameters (Van Riel *et al.* 2006), whereby each parameter has an uncertainty which needs to be considered

(Omlin *et al.* 2001; Marsili-Libelli *et al.* 2003).

The Fisher information matrix (FIM) is a critical quantity in several aspects of mathematical modeling, including input selection and confidence region calculation. Therefore, the name “information matrix” is used to indicate that a larger FIM (in the matrix sense (positive semi-definiteness as just mentioned) is associated with a smaller covariance matrix (i.e., more information), while a smaller FIM is associated with a larger covariance matrix (i.e., less information). Some important areas of applications of FIM include, to name a few, confidence interval computation of model parameter (Šimandl *et al.* 2001), configuration of experimental design involving linear (Chryssolouris *et al.* 1996) and nonlinear models (Spall *et al.* 2003).

Recently, the importance of assessing the precision and confidence interval of the parameter estimation from experimental data has been investigated by researchers in different systems (Chryssolouris *et al.* 1996; Spall *et al.* 2003; Insel *et al.* 2003), but confidence interval assessment is not a straightforward task since many different factors are involved such as the experimental data and of course the model structure. This enables identifying a set of experiments and model settings that deliver the most sensitive to the unknown parameters measurement data and thus avoiding a trial and error approach to seek for solutions.

The model itself relates the equations, input factors, parameters and variables characterizing the process with the input factors belonging to a set of probability distribution associated with a quite number of uncertainty sources including

errors arising from measurement, insufficient amount of information and poor or partial understanding of the mechanisms and driving forces. However, it is of paramount importance in the modeling practice, evaluate the confidence region of the model by way of assessing the uncertainties relating the model inputs with the output in any given situation (Saltelli *et al.* 2000). Hence a robust technique for model identification and validation should be used especially for complex systems described by differential-algebraic equations, DAEs (Donoso-Bravo *et al.* 2011).

In this paper, we model the SRP using SRB *D. alaskensis* 6SR, grown in two different media: congenial water (CW) and Postgate C medium (PCM), the importance of confidence interval of the parameter estimation from experimental data for SRP in each media was investigated for each model proposed. We consider two different kinetic models unstructured: Monod and Levenspiel to model the growth of bacteria, respectively using CW and PCM as medium of growth.

The parameter estimation was carried out with classical simplex Nelder and Mead minimization algorithm, using as a cost function the norm of the differences between the experimental data and the model. The confidence intervals were assessed through a numerical method based on the Fisher Information Matrix (FIM). The current work proposes an extension of the re-sampling algorithm in order to enhance the statistical qualities of the estimator of the FIM. This modified re-sampling algorithm is useful in those cases where the FIM has a structure with some elements being analytically known from prior information and the others being unknown.

2. Methodology

2.1 Experimental section

2.1.1 Microorganism: sulfate-reducing bacteria (SRB)

D. alaskensis 6SR was isolated from biofilm samples, it was identified by 16S rRNA gene sequencing and analysis (Neria-González *et al.* 2006). This strain was routinely cultivated and maintained on Postgate C medium according to Neria-González *et al.* (2006).

2.1.2 Batch cultures: congenital water

A sample of congenital water (CW) was obtained of an oil pipeline located in the Mexican Southeast region. Chemical determination of water: chlorides 64 000 g L⁻¹, sulfur 178 g L⁻¹, sulfate 350 to 400 g L⁻¹, pH 8.84. A 1000 ml aliquot of CW was saturated with N₂ by 1 hour and was enriched with sodium lactate 6 ml, yeast extract 0.5 g, and reducing solution 5 ml (acid ascorbic 1 g L⁻¹, and sodium thioglycolate, 1 g L⁻¹). Potential of hydrogen (pH) was adjusted to 7 with 1 N KOH. Then 90 ml CW medium was distributed in serum bottles of 160 ml using Hungate's technique¹⁰ and they were autoclaved at 120 °C for 15 min. The initial cultures of *D. alaskensis* 6SR in Postgate C medium¹¹ were used to inoculate 45 ml of CW medium. Culture was incubated at 20 days to 37°C under batch operation conditions.

Subsequently, 10 ml aliquot of this culture was used to inoculate several bottles with CW medium at different times under same conditions. Bacterial growth was followed through optical density (OD) measurements, sulfate consumption and sulfide production.

Samples from the cultures were taken anaerobically each hour. Sulfate in the medium was measured by the turbidimetric method based on the precipitation of barium. Also, the production of sulfide was measured by a colorimetric method. The OD reading for cell growth was transformed to dry weight (concentration) through a standard growth curve. Data were analyzed to determine the growth kinetics parameters according to the Monod model.

2.1.3 Batch cultures: Postgate C medium

Experiments were carried out using the modified nutrient Postgate C medium, which contains NaCl 30 g L⁻¹ and this does not contain Fe(II) (Postgate *et al.* 1984). The composition of the growth medium for *D. alaskensis* 6SR, consisted of 4.5 g/l sodium lactate, 1 g L⁻¹ sodium citrate, 1 g L⁻¹ yeast extract, 4.5 g L⁻¹ Na₂SO₄, 0.06 g L⁻¹ CaCl₂ • H₂O, 1.0 g L⁻¹ NH₄Cl, 0.5 g L⁻¹ KH₂PO₄, 2.0 g L⁻¹ MgSO₂ • 7H₂O, 30 g L⁻¹ NaCl and the volume was made up to 1 L with distilled water and pH of the solution was adjusted to 7 ± 0.2. Before autoclaving, the medium was flushed with nitrogen (100 ml/min – 30 min) to remove dissolved oxygen in medium and the head space.

The medium was adjusted to pH 7 and were placed 45 ml into serological bottles. These vessels were capped with crimped aluminum butyl rubber stoppers and sterilized by autoclaving at 120 °C for 15 min. Each bottle was inoculated with a 5 ml aliquot of a preculture of *D.*

alaskensis 6SR (OD_{580} between 0.35 and 0.4). All cultures were incubated at 37 °C for 8 days under batch operation conditions, in both cases using sulfate and sodium lactate as electron acceptor and donor respectively, using the same techniques previously described.

3. Numerical section

3.1 Mathematical model of the bioreactor

3.1.1 Case study: Monod as kinetic model for SRP

The following mathematical model is proposed, based on classical mass balances for biomass, sulfate (substrate) and sulfide (product) concentrations, considering batch operation:

Biomass (X) for growth of *D. alaskensis* 6SR in CW:

$$\frac{d}{dt}X \equiv \left(\frac{dX}{dt}\right)_{\text{growth}} = f_1(x, t, \theta)$$

(1)

Substrate (S):

$$\frac{d}{dt}S \equiv q_S \times X = -Y_S \times \left(\frac{dX}{dt}\right)_{\text{growth}} = f_2(x, t, \theta)$$

(2)

Sulfide (P):

$$\frac{d}{dt}P \equiv q_P X = Y_P \left(\frac{dX}{dt}\right)_{\text{growth}} = f_3(x, t, \theta)$$

(3)

3.1.2 Case study: Levenspiel as kinetic model for SRP

Biomass (X) for growth of *D. alaskensis* 6SR in Postgate C as medium:

$$\frac{d}{dt}X \equiv \left(\frac{dX}{dt}\right)_{\text{growth}} - k_d X = f_1(x, t, \theta)$$

(4)

Substrate (S):

$$\frac{d}{dt}S \equiv q_S X = -Y_S \left(\frac{dX}{dt}\right)_{\text{growth}} = f_2(x, t, \theta)$$

(5)

Sulfide (P):

$$\frac{d}{dt}P \equiv q_P X = Y_P \left(\frac{dX}{dt}\right)_{\text{growth}} = f_3(x, t, \theta)$$

(6)

In this study, increase in cell concentration ($\left(\frac{dX}{dt}\right)_{\text{growth}}$), for cultures

of SRP obeying the Monod *et al.* (1949) rate law (congenial water as medium) and Levenspiel *et al.* (1988) rate law (Postgate C as medium) for example, these term becomes:

$$\left(\frac{dX}{dt}\right)_{\text{growth}} = \mu_{\max} \frac{S}{k_S + S} X$$

(7)

$$\left(\frac{dX}{dt}\right)_{\text{growth}} = \mu_{\max} \frac{S}{k_S + S} \left[1 - \frac{P}{P^*}\right]^n X$$

(8)

In order to express the models (1)-(3) and (4)-(6) in the standard state-space form. The state vector is defined as

$$x = [X, S, P]^T; \quad y = [X, S, P]^T$$

(9)

Then, the models in eqs (1)-(3) and (4)-(6) can be represented as

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{t}, \boldsymbol{\theta}); \mathbf{x}(t_0) = \mathbf{x}(0)$$

(10a)

$$\mathbf{y} = \mathbf{h}(\mathbf{x})$$

(10b)

Therefore, the biological systems model discussed here (equations (1)-(6) and (4)-(6) with eqs (7) and (8), respectively) are system of ODE's that is dependent on a certain parameter set $\boldsymbol{\theta}$ and initial conditions $\mathbf{x}_i(t_0) = \mathbf{x}_i(0)$.

Where

\mathbf{X} , \mathbf{S} , and \mathbf{P} are respectively cell concentration, substrate, and product in the reaction mixed.

$\mathbf{y} = [\mathbf{X}, \mathbf{S}, \mathbf{P}]^T$ = measurement vector (ideal case)

$\mathbf{x} = [\mathbf{X}, \mathbf{S}, \mathbf{P}]^T$ = system states in the reactor

$\mathbf{f}(\mathbf{x}, \mathbf{t}, \boldsymbol{\theta})$ = reaction kinetics

$\boldsymbol{\theta}(\bullet)$ = parameter vector

q_p = specific product production rate

q_s = specific substrate consumption rate

μ_{max} = maximum specific growth rate

k_d = constant of cell death

k_s = substrate saturation constant

Y_s, Y_p = yields
 $\bar{x} \quad \bar{x}$

P^* = product inhibition constant

n = constant

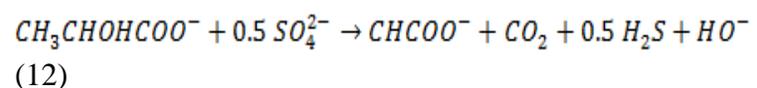
t = time

The output measurement considered in this study was the substrate (sulfate) uptake rate (SUR):

$$\left(\frac{ds}{dt}\right)_{\text{reaction}} \equiv \text{SUR}$$

(11)

Basic models for the SRP here (equations (1)-(6)) considered that sulfate-reducing microorganism utilize sulfate and the carbon source (lactate in this study) to produce carbon dioxide and hydrogen sulfide (Cao *et al.* 2012). Consequently, for sulfate reducing reactors, the rates of sulfate consumption as well as the rates of hydrogen production and sulfide are proportional to the rate of consumption of carbon source. Basis of proportionality is defined by the stoichiometry of the reaction (equation (12)). However, in this work we only considered in developing the basic models, compounds in the liquid phase (equations (1)-(6)).



3.2 Parameter estimation and confidence interval

Estimating the parameter ($\boldsymbol{\theta}$) related to the models proposed for the SRP can be estimated using a minimization algorithm (equation (13)), this algorithm considers the squared errors, J , in this algorithm considers the minimization of the squared errors between model outputs $\mathbf{y}(\mathbf{k}, \boldsymbol{\theta})$ and the measured outputs $\mathbf{y}_M(\mathbf{k})$, with \mathbf{k} as a certain sampling point. Q_k is defined as a

weighting matrix to balance the effect of each kind of measurement.

$$J = \sum_{k=1}^N [y(k, \theta) - y_M(k)]^T Q_k [y(k, \theta) - y_M(k)] \quad (13)$$

where N is the number of measurements and θ is the parameters set used to calculate the model outputs.

Confidence interval for the estimated parameters is based on the Fisher Information matrix (FIM) (Dochain *et al.* 2001).

This matrix relating the amount of information contained in the experimental data. The FIM can be calculated by linearizing the output signals of the system studied in the area around optimal parameter values θ_0 .

Linearization of the outputs for each parameter can be expressed as in equation (14). Mathematically, the sensitivity coefficients are the first-order derivatives of model outputs with respect to the model parameters

$$S_{x_i \theta} = \frac{\partial x_i}{\partial \theta}; \quad i = 1, \dots, n \quad (14)$$

Where x_i is the i th model output and θ is the model input parameter. These differential equations are differential output x_i with respect to the model input parameter θ over time (Khalil *et al.* 2002).

$$\frac{\partial S_{x_i \theta}}{\partial t} \equiv \frac{\partial}{\partial t} \left(\frac{\partial x_i}{\partial \theta} \right) = \frac{\partial f_{i(x,t,\theta)}}{\partial \theta} + \sum_{i=1}^n \frac{\partial f_{i(x,t,\theta)}}{\partial x_i} \times \frac{\partial x_i}{\partial \theta} \quad (15)$$

Thus, (15) can be rewritten as

$$\frac{\partial}{\partial t} S_{x_i \theta} = \begin{bmatrix} \frac{\partial f_1(x,t,\theta)}{\partial x_1} & \dots & \frac{\partial f_1(x,t,\theta)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(x,t,\theta)}{\partial x_1} & \dots & \frac{\partial f_n(x,t,\theta)}{\partial x_n} \end{bmatrix} \begin{bmatrix} S_{x_1 \theta} \\ \vdots \\ S_{x_n \theta} \end{bmatrix} + \begin{bmatrix} \frac{\partial f_1(x,t,\theta)}{\partial \theta} \\ \vdots \\ \frac{\partial f_n(x,t,\theta)}{\partial \theta} \end{bmatrix} \quad (16)$$

A procedure for calculating $S_{x_i \theta}$ is to append the variational equations (16) with the original state equation (11) to obtaining the $(n + np)$ augmented equation

$$\dot{x} = f(x, t, \theta), \quad x(t_0) = x(0) \quad (17a)$$

$$\dot{S} = f_x(x, t, \theta) \times S + f_\theta(x, t, \theta), \quad S(t_0) = S(0) \quad (17b)$$

The matrix f_θ , f_x , S , $x(0)$ and $S(0)$ have the following definition with $p = 1$ (i.e., only considered a parameter):

$$f_\theta \triangleq \frac{\partial f}{\partial \theta} = \begin{bmatrix} \frac{\partial f_1(x,t,\theta)}{\partial \theta} \\ \vdots \\ \frac{\partial f_n(x,t,\theta)}{\partial \theta} \end{bmatrix}, \quad f_x \triangleq \frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f_1(x,t,\theta)}{\partial x_1} & \dots & \frac{\partial f_1(x,t,\theta)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(x,t,\theta)}{\partial x_1} & \dots & \frac{\partial f_n(x,t,\theta)}{\partial x_n} \end{bmatrix}, \quad S = \begin{bmatrix} S_{x_1 \theta} \\ \vdots \\ S_{x_n \theta} \end{bmatrix} \\ , \quad x(0) = \begin{bmatrix} x_1(0) \\ \vdots \\ x_n(0) \end{bmatrix}, \quad S_0 = \begin{bmatrix} S_{x_1 \theta}(0) \\ \vdots \\ S_{x_n \theta}(0) \end{bmatrix} \quad (18)$$

The FIM is used to summarize information related to the uncertainties and dependencies between the parameter estimates (θ):

$$FIN = \sum_{k=1}^N S_{x\theta}^T(k) Q_k S_{x\theta}(k) \quad (19)$$

The FIM is a square matrix with same number of columns and rows as the number parameters to estimate. Q_k is also a square matrix with the same number of columns and rows as output measurements.

If we consider the first case (eqs (1)-(3) and (7)), the parametric sensitivity matrix S is given by eq (17b). This model contains four model parameters, μ_{max} , k_s , $\frac{Y_S}{X}$ and $\frac{Y_P}{X}$ but in our case we only considered two model parameters μ_{max} , k_s and the others were considered constant and according with these model parameters:

$$S \triangleq \begin{bmatrix} S_{x_1\theta_1} & \dots & S_{x_1\theta_j} \\ \vdots & \ddots & \vdots \\ S_{x_i\theta_1} & \dots & S_{x_i\theta_j} \end{bmatrix};$$

$i = 1, \dots, n; j = 1, \dots, p$

(20)

then considering the eqs (17a); (18) and (20), therefore, the eq (17a) is defined by the following matrices with $\theta \triangleq [\mu_{max}, k_s] = [\theta_1, \theta_2]$; $x \triangleq [x_1, x_2, x_3]^T = [X, S, P]^T$, therefore, the Jacobians f_θ and f_x of $f(x, t, \theta)$ can be obtained easily, then, sensitivity of the states on the variation of the model parameters (μ_{max} , k_s) for this study case is $\in \mathbb{R}^{n \times p}$:

$$\dot{S}_{y_1\theta_1} = \frac{\theta_1 S}{(\theta_2 + S)} S_{y_1\theta_1} + \frac{\theta_1 \theta_2}{(\theta_2 + S)^2} S_{y_2\theta_1} + \frac{SX}{(\theta_2 + S)}$$

(21)

$$\dot{S}_{y_1\theta_2} = \frac{\theta_1 S}{(\theta_2 + S)} S_{y_1\theta_2} + \frac{\theta_1 k_s}{(\theta_2 + S)^2} S_{y_2\theta_2} - \frac{\theta_1 SX}{(\theta_2 + S)^2}$$

(22)

$$\dot{S}_{y_2\theta_1} = -\frac{Y_S \theta_1 S}{(\theta_2 + S)} S_{y_1\theta_1} - \frac{Y_S \theta_1 \theta_2}{(\theta_2 + S)^2} S_{y_2\theta_1} - \frac{Y_S SX}{(\theta_2 + S)}$$

(23)

$$\dot{S}_{y_2\theta_2} = -\frac{Y_S \theta_1 S}{(\theta_2 + S)} S_{y_1\theta_2} - \frac{Y_S \theta_1 k_s}{(\theta_2 + S)^2} S_{y_2\theta_2} + \frac{Y_S \theta_1 SX}{(\theta_2 + S)^2}$$

(24)

$$\dot{S}_{y_3\theta_1} = \frac{Y_P \theta_1 S}{(\theta_2 + S)} S_{y_1\theta_1} + \frac{Y_P \theta_1 \theta_2}{(\theta_2 + S)^2} S_{y_2\theta_1} + \frac{Y_P SX}{(\theta_2 + S)}$$

(25)

$$\dot{S}_{y_3\theta_2} = \frac{Y_P \theta_1 S}{(\theta_2 + S)} S_{y_1\theta_2} + \frac{Y_P \theta_1 k_s}{(\theta_2 + S)^2} S_{y_2\theta_2} - \frac{Y_P \theta_1 SX}{(\theta_2 + S)^2}$$

(26)

Therefore, equations (21)-(26) are the sensitivity functions for system (1)-(3), where (21)-(22), (23)-(24) and (25)-(26) represent sensitivity to the dynamics of biomass, substrate and product respectively. According with equations (21)-(26) the order for system (17b) is 6 for the system (1)-(3) with Monod as rate law (i.e., kinetic model for *D. alaskensis* 6SR). A similar procedure can be used to obtain the differential equations for sensitivity functions to process using the Levenspiel model (equation (8)) in the equations (4)-(6) with $x \triangleq [x_1, x_2, x_3]^T = [X, S, P]^T$ and $\theta \triangleq [\theta_1, \theta_2, \theta_3, \theta_4] = (\mu_{max}, k_s, n, k_d)$, equations not shown here, in this case have the system (17b) is of order $n \times p = 12$ (i.e., 12 sensitivity functions) and the order for system (17a)-(17b) is $n + n \times p = 15$.

If we continue with the calculation of the FIM for this system of study, the following FIM is obtained from eq (19):

$$FIM = \sum_{k=1}^N \begin{bmatrix} a_{11}(k) & a_{12}(k) \\ a_{21}(k) & a_{22}(k) \end{bmatrix} = \sum_{k=1}^N S_{x\theta}^T(k) Q_k S_{x\theta}(k) \quad (27)$$

$$\begin{bmatrix} a_{11}(k) & a_{12}(k) \\ a_{21}(k) & a_{22}(k) \end{bmatrix} = \begin{bmatrix} S_{y_1\theta_1}(k, \theta_0) & S_{y_2\theta_1}(k, \theta_0) & S_{y_3\theta_1}(k, \theta_0) \\ S_{y_1\theta_2}(k, \theta_0) & S_{y_2\theta_2}(k, \theta_0) & S_{y_3\theta_2}(k, \theta_0) \end{bmatrix}^T \begin{bmatrix} Q_1(k) & 0 \\ 0 & Q_2(k) \end{bmatrix} \begin{bmatrix} S_{y_1\theta_1}(k, \theta_0) & S_{y_1\theta_2}(k, \theta_0) \\ S_{y_2\theta_1}(k, \theta_0) & S_{y_2\theta_2}(k, \theta_0) \\ S_{y_3\theta_1}(k, \theta_0) & S_{y_3\theta_2}(k, \theta_0) \end{bmatrix}$$

here the FIM elements for certain time (k) corresponds to $Q_i(k)$ corresponds to the inverse of the covariance of the measurement noise of the output variable y_i for each sampling point. If the error is considered constant throughout the experiment $Q_i(k)$ becomes a scalar instead of a vector.

The calculation of the FIM matrix can obtain the following information about the quality of experimental data: 1) summarizes the quantity and quality of information obtained in each experiment as it considers the output sensitivity functions (\dot{S}) and the measurement errors of the experimental data, 2) the inverse of the FIM provides the lower bound of the parameter estimation error covariance matrix, which can be used for assessing the estimation uncertainty of θ_0 (Peters & Hanzon *et al.* 1998). According with 1) and 2) the error covariance matrix is given by eq (28).

$$COV(\theta_0) \geq FIM^{-1} \quad (28)$$

Then, the square root of the diagonal elements of the FIM^{-1} can be used to approximate the standard errors for the

estimated parameters θ_0 as in shown in eq (29)

$$\sigma(\theta_j) = \sqrt{FIM^{-1}} \quad (29)$$

Hence, the higher the FIM values, the lower the standard errors estimated.

4. Results and analysis

4.1 Experimental results

Results showed that biomass, sulfate and dissolve sulfide in both cases were consumed (sulfate) and produced (biomass and sulfide) throughout the experiment at the first 50-100 hours in both experiments. Behavior of the SRB *D. alaskensis* 6SR in both study of cases was similar and marked by high levels sulfate reduction, and sulfide production. Results of sulfate reduction, sulfide production from batch cultures with CW and PCM are shown representatively in **Fig. 1** and **Fig 2** (experimental data with markers).

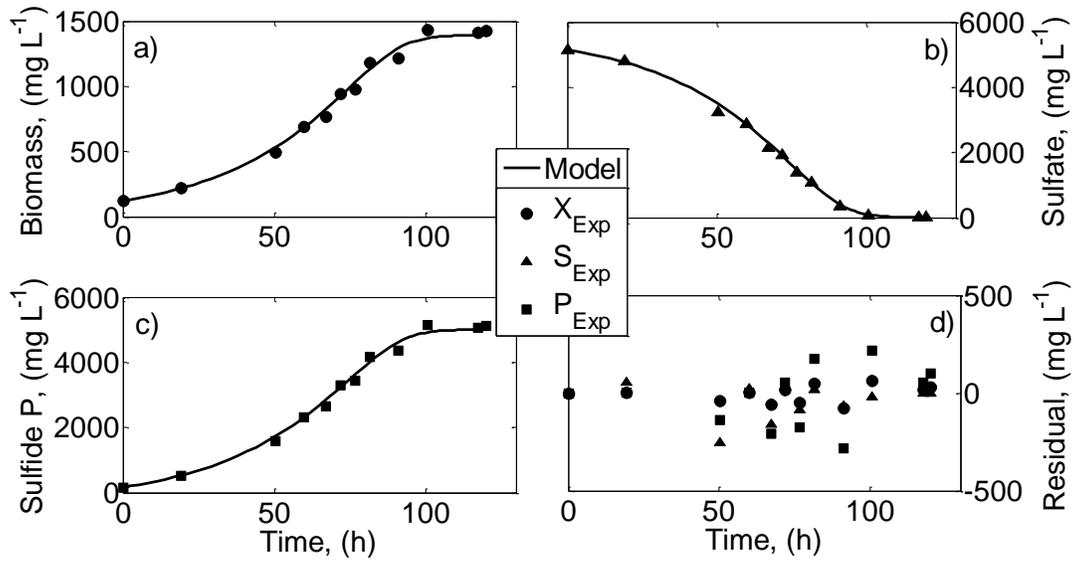


Fig. 1- Monod model validation with experimental data: (a) biomass, (b) sulfate and (c) sulfide, and (d) plot between residual and corresponding state variables. Data represents the mean of three replicates.

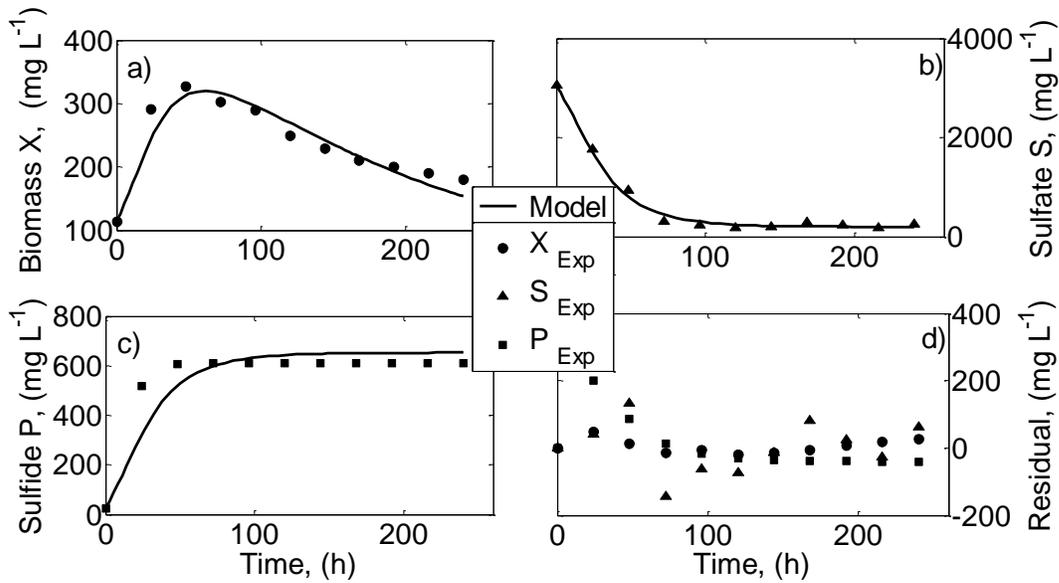


Fig. 2- Levenspiel model validation with experimental data: (a) biomass, (b) sulfate and (c) sulfide, and (d) Plot between residual and corresponding state variables. Data represents of the mean eight replicates.

However, the inhibition of the SRB began to be detected in PCM from hour 100 onwards (**Fig. 2 a**), this inhibition effect was attributed to high hydrogen sulfide concentration ($550 \pm 50 \text{ mg L}^{-1}$ as main in both experiments) as pointed out by Reis *et al.* (1992) in cultures using SRB. Final

biomass, sulfate and dissolve sulfide concentrations for experiments are summarized in **Table 1**. According with these data, 9.7-21.3 % sulfate was converted to dissolved sulfide ($45\text{-}56 \text{ mg L}^{-1}$).

Table 1- Biomass, sulfate and dissolve sulfide concentrations at the end of experiments

Medium	¹ CW (mg L)	² PCM (mg L)	The sulfate removal efficiency (%)	The sulfate consumption rate ($\text{mg L}^{-1}\text{h}^{-1}$)
Biomass	1400	320	¹ 97.5	¹ 73.92
Sulfate	$\cong 0$	< 500	² 83.6	² 72.02
Sulfide	500 (^a 9.7%)	650 (^a 21.31%)		
pH	8.0	7.6		

^aHydrogen sulfide in liquid phase with respect to the initial concentration of sulfate

4.2 Experimental procedures

First all, data were generated experimentally in batch cultures of SRB *D. alaskensis* 6SR with two different media (WC and PC). According to the analysis of experimental data, the growth kinetics of *D. alaskensis* 6SR was modeled with unstructured kinetic models. Bacteria growth in CW was modeled with Monod model (**Fig. 1** solid line) and with Levenspiel model when bacteria used Postgate C as medium (**Fig. 2** solid line), experimental data are shown with markers. In both cases of study, the response variables measured were biomass **X**, sulfate **S**, and hydrogen sulfide **P**. The parameter vector θ was estimate through the minimization function FIM in search. The cost function

was defined as the norm of vector resultant of the difference between experimental modeled data (eq (19)).

The **X**, **S** and **P** profiles obtained using the optimal parameters (solid line **Figs. 1** and **2**) were used for the calculation of sensitivity functions (eqs (17a)-(17b)) for each specific sampling time **k** for the two case studies (**Fig. 3**). In Fig. 3, shows only the behavior of the sensitivity function for the sulfate (eqs (23)-(24)), because from the control standpoint, this state has been considered sufficient to regulate and monitoring the PSR, therefore, the sensitivity of this state on the variation of the parameters in the process directly affect the rest of the states (Aguilar-Lopez *et al.* 2010). According with these results, for Monod

as kinetic model for SRP the maximum specific growth rate μ_{\max} , was the parameter for which the SRP was more sensitive compared with the parameter k_s (Fig. 3 a) and b) respectively, solid line).

But for the PSR described by Levenspiel as kinetic model were the parameters k_d and μ_{\max} (Fig. 3 .c) to f) respectively), also SRP was sensitive to the parameter n (Fig. 3.e).

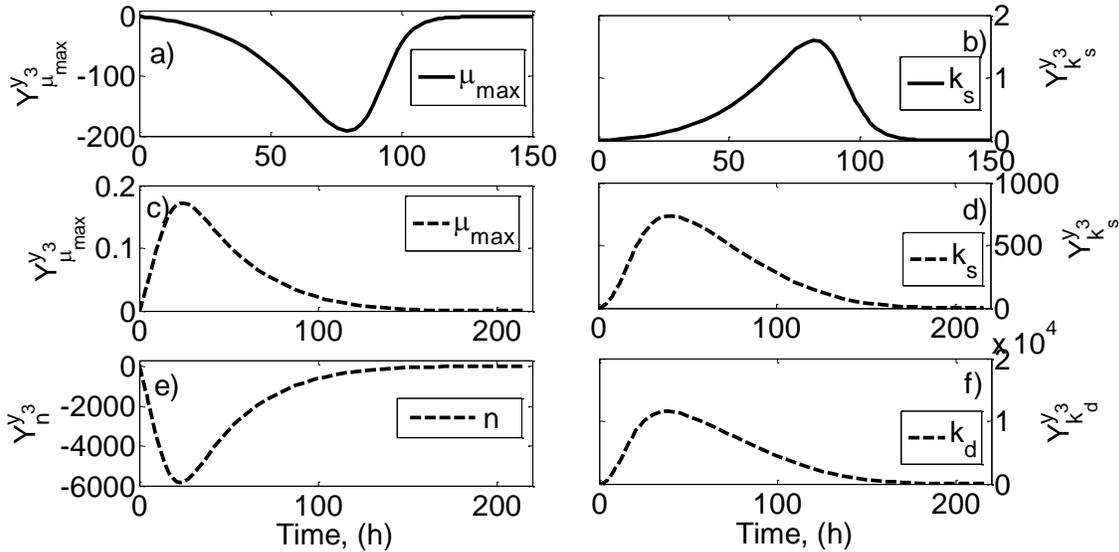


Fig. 3- Sensitivity functions of substrate S , with respect to μ_{\max} (a) and k_s (b) to Monod as kinetic model for SRP and μ_{\max} (c) and k_s (d), n (e) and k_d (f) to Levenspiel as kinetic model for SRP.

4.3 Influence of quality of experimental data

In order to infer the variability in the process of estimating the parameters in the SRP, the influence of the data quality was studied through variation of experimental data measurement error, according whit the model parameters and initial conditions in **Tables 1** and **2**. The results obtained using different measurement errors are depicted in **Fig. 4**. A value of 0.02 for the measurement of error (the optimal value of the parameters i.e., which minimize the cost function are shown in **Table 3**), and then we considered the variation of measurement error in the following

domain $[0.02 - 1.2]$. And as expected, an increase in the measurement error generates different values in the vector of estimated parameters (**Fig. 4 a**) and **c**), respectively Monod and Levenspiel) consequently, an increase in the confidence interval (**Fig. 4 b**) and **d**), respectively Monod and Levenspiel). Finally, in the SRP modeling in the two case studies, will assure the reliability of the values of the parameters estimated in two steps: parameter optimization and parameter error assessment.

Table 1- Biomass, sulfate and dissolve sulfide concentrations at the end of experiments

Medium	¹ CW (mg L)	² PCM (mg L)	The sulfate removal efficiency (%)	The sulfate consumption rate (mg L ⁻¹ h ⁻¹)
Biomass	1400	320	¹ 97.5	¹ 73.92
Sulfate	≅ 0	< 500		
Sulfide	500 (^a 9.7%)	650 (^a 21.31%)	² 83.6	² 72.02
pH	8.0	7.6		

^aHydrogen sulfide in liquid phase with respect to the initial concentration of sulfate

Table 2. Model Parameters

Known parameters	Rate law, $\left(\frac{dX}{dt}\right)_{\text{growth}}$		Unknown parameters
	¹ Monod	² Levenspiel	
Yield : $\frac{Y_S}{\bar{X}}$	4.05	8.8311	¹ $\theta = (\mu_{\max}, k_S)$
Yield: $\frac{Y_P}{\bar{X}}$	3.81	1.9449	² $\theta(\mu_{\max}, k_S, n, k_d)$
Measurement error	0.02	0.02	
Final time (batch culture)	150 h	216 h	

Table 3. Initial conditions

Initial condition for states	Rate law, $\left(\frac{dX}{dt}\right)_{\text{growth}}$		Equation
	¹ Monod	² Levenspiel	
	Values		
Initial biomass: $X(t_0) = X(0)$	123 (mg L ⁻¹)	110 (mg L ⁻¹)	(17a)
Initial substrate: $S(t_0) = S(0)$	5150 (mg L ⁻¹)	3050 (mg L ⁻¹)	
Initial product: $P(t_0) = P(0)$	164 (mg L ⁻¹)	30 (mg L ⁻¹)	

initial conditions sensitivity

(17b)

functions: $S_{np}(0)$ $S_{3 \times 2}(0)$ $S_{3 \times 4}(0)$

$$S \triangleq \begin{bmatrix} S_{11}(0) & \dots & S_{1j}(0) \\ \vdots & \ddots & \vdots \\ S_{i1}(0) & \dots & S_{np}(0) \end{bmatrix} \quad 0 \text{ (mg L}^{-1}\text{)} \quad 0 \text{ (mg L}^{-1}\text{)} \quad (17b)$$

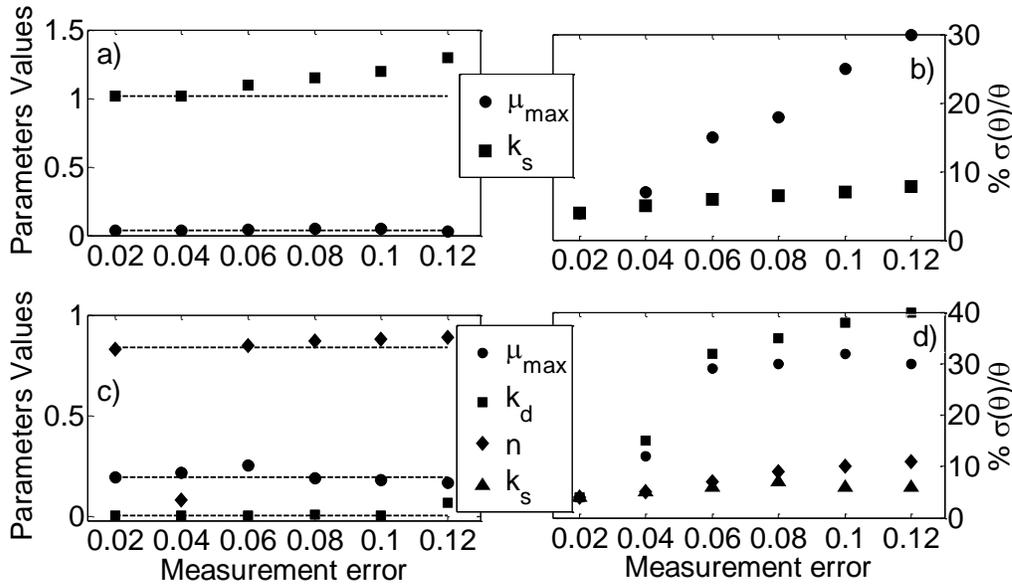


Fig. 4- The influence of the data quality: measurement error on the parameter estimation (left) above congenital water a) below Postgate C medium c) and on the confidence interval assessment (right) above congenital water b) below Postgate C medium d)

5. Conclusions

A simple procedure to model the sulfate-reducing process, using the bacteria *Desulfovibrio alaskensis* 6SR is presented. The procedure was applied to model the growth of bacteria in the two different culture media (congenital water and Postgate C). Growth kinetics was modeled by unstructured kinetic models (Monod and Levenspiel). The influence of the quality of experimental data on parametric identification process was

studied by Fisher information matrix (in terms of proximity of the estimated value to the real one and in terms of confidence interval assessment). Overall, the results showed that increasing the measurement error (quality) implies a lower accuracy in the estimation of parameters. At this point, it is helpful to recall that such models are useful in the sense that they will describe specific situation, but that the full range of their utility is unknown. Accurate estimation and model validation can be performed using this methodology approach. It also provides a solution for the correlated parameters and high

standard deviations which are known to be common problems for the estimation of parameters from batch experiments. As a result, the initial conditions of the experiments should be well defined, depending upon the selected model.

Acknowledgments

P.A.L.P and V.P.C would like to thank to Conacyt. Thanks are due to Elvira Leal, and Cirino Chávez by collaborative on Chemical Analysis. The technical assistance of Fernando Maldonado García (Department of Biotechnology & Bioengineering, CINVESTAV-IPN)

References

1. López-Pérez, P. A., Neria-González, M. I. & Aguilar-López, R. (2013). Cadmium concentration stabilization in a class of continuous sulfate reducing bioreactor via sulfide concentration control. *Chemical Papers*, 67 (3), 326–335.
2. Gallegos-Garcia, M., Celis, L. B., Rangel-Méndez, R. & Razo-Flores, E. (2009). Precipitation and recovery of metal sulfides from metal containing acidic wastewater in a sulfidogenic down-flow fluidized bed reactor. *Biotechnol Bioeng*, 102(1), 91–9.
3. White, C., Dennis, J. S., Gadd, G. M. A. (2003). Mathematical process model for cadmium precipitation by sulphate-reducing bacterial biofilms. *Biodegradation*, 14, 139-151.
4. Acha, D., Hintelmann, H. & Pabon, C. A. (2012). Sulfate-reducing Bacteria and Mercury Methylation in the Water Column of the Lake 658 of the Experimental Lake Area. *Geomicrobiol J*, 29(7), 667-674.
5. Feio, M. J., Zinkevich, V., Beech, I. B., Brossa, E. L., Eaton, P., Schmitt, J. & Guezennec, J. (2004). *Desulfovibrio alaskensis* sp. nov., a sulphate reducing bacterium from a soured oil reservoir. *Int J Syst Evol Micr*, 54, 1747–1752.
6. Neria-González, I., Wang, E. T., Ramírez Z, F., Romero, J. M. & Hernández-Rodríguez, C. (2006). Characterization of bacterial community associated to biofilms of corroded oil pipelines from the Southeast of Mexico. *Anaerobe*, 12, 122-133.
7. Van Riel, N. A. (2006). Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments. *Brief Bioinform*, 7(4), 364-374.
8. Omlin, M., Brun, R. & Reichert, P. (2001). Biogeochemical model of Lake Zürich: sensitivity, identifiability and uncertainty analysis. *Ecological Modelling*, 141, 105–123.
9. Marsili-Libelli, S., Guerrizio, S. & Checchi, N. (2003). Confidence regions of estimated parameters for ecological systems. *Ecol Model*, 165, 127–146.
10. Šimandl, M., Kráľovec, J. & Tichavský P., (2001). Filtering, predictive, and smoothing Cramer-Rao bounds for discrete-time nonlinear dynamic systems. *Automatica*. 37(11), 1703-1716.
11. Chryssolouris, G., Lee, M. & Ramsey, A. (1996). Confidence interval prediction for neural networks models. *IEEE T Neural Network*, 7(1), 229-232.
12. Spall, J. C. (2003). *Introduction to Stochastic Search and Optimization: Estimation, Simulation and Control*. Wiley-Interscience.
13. Insel, G., Orhon, D. & Vanrolleghem, P. A. (2003). Identification and modelling of aerobic hydrolysis - application of optimal experimental design. *J Chem Technol Biot*, 78(4), 437–445.
14. Saltelli, A. (2000). *What is Sensitivity Analysis?* John Wiley & sons Ltd. European Commission Joint Research Centre, Italy

15. Postgate, J. R. In the sulphate-reducing Bacteria (2nd edn). (1984). Cambridge University Press. 30-36.
16. Donoso-Bravo, A., Mailier, J., Martin, C., Rodríguez, J., Aceves-Lara, C. A. & Vande Wouwer, A. (2011). Model selection, identification and validation in anaerobic digestion: a review. *Water res*, 45(17), 5347–5364.
17. Monod, J. (1949). The growth of bacterial cultures. *Annu Rev Microbiol*, 3, 371-394
18. Han, K. & Levenspiel, O. (1988). Extended Monod kinetics for substrate, product and cell inhibition. *Biotechnol. Bioeng*, 32, 430-437.
19. Cao, J. Y., Zhang, G. J., Mao, Z. S., Li, Y. Y., Fang, Z. H. & Yang, C. (2012). Influence of electron donors on the growth and activity of sulfate-reducing bacteria. *Int. J. Miner. Process*, 106, 58-64.
20. Dochain D., Vanrolleghem, P.A. (2001). *Dynamical modelling and estimation in wastewater treatment processes*. IWA Publishing, London, UK.
21. Khalil, H. K. (2002). *Nonlinear Systems*. Prentice Hall, NJ, USA, third edition.
22. Peteers, R. L. & Hanzon, B. (1998). Symbolic computation of Fisher information matrices for parametrized state-space systems. *Automatica*, 35, 1059-1071.
23. Reis, M. A. M. (1992). Effect of Hydrogen-Sulfide on Growth of Sulfate Reducing Bacteria. *Biotechnol Bioeng*, 40(5), 593-600.
14. Aguilar-López, R., Pena-Caballero V. & Neria-Gonzalez M. I. (2010). Control of a Class of Sulfate Reducing Chemostat Via Feedback Polynomial Injection. *J Appl Res Technol*, 8(2), 274-287.

