Environmental Modelling & Software 85 (2016) 356–373

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Global parameter optimization for biokinetic modeling of simple batch experiments

Alma Mašić, Kai M. Udert, Kris Villez*

Eawag – Swiss Institute for Aquatic Science and Technology, Process Engineering, Dübendorf, Switzerland

ARTICLE INFO

Article history:
Received 9 January 2016
Received in revised form
19 May 2016
Accepted 22 June 2016

Keywords:
Biotechnology
Deterministic search
Global optimality
Kinetic modeling
Nitrification
Parameter estimation

ABSTRACT

Environmental process modeling is challenged by the lack of high quality data, stochastic variations, and nonlinear behavior. Conventionally, parameter optimization is based on stochastic sampling techniques to deal with the nonlinear behavior of the proposed models. Despite widespread use, such tools cannot guarantee globally optimal parameter estimates. It can be especially difficult in practice to differentiate between lack of algorithm convergence, convergence to a non-global local optimum, and model structure deficits. For this reason, we use a deterministic global optimization algorithm for kinetic model identification and demonstrate it with a model describing a typical batch experiment. A combination of interval arithmetic, reformulations, and relaxations allows globally optimal identification of all (six) model parameters. In addition, the results suggest that further improvements may be obtained by modification of the optimization problem or by proof of the hypothesized pseudo-convex nature of the problem suggested by our results.

1. Introduction

Despite abundant literature, model identification is a challenging task for environmental systems which keeps drawing considerable attention (Marsili-Libelli, 2010). In response, protocols have been developed to simplify the model identification task (Jakeman et al., 2006). One important aspect is that environmental process and system models are typically nonlinear in their parameters. Despite this problem, nonlinear parameter estimation is often solved with gradient-based optimization techniques that may not converge (Checchi et al. 2007) or which may converge to a local optimum (Jakeman et al., 2006; Rieger et al., 2012). Alternatively, stochastic optimization tools in combination with sensitivity-based parameter selection techniques (e.g., Benedetti et al., 2011; Sin et al., 2008) can ease this task. While fruitful in many cases, stochastic methods can still converge to a local optimum or may not converge at all. This is a significant drawback if the model structure itself is uncertain and subject to selection or modification. In other areas of engineering, deterministic optimization techniques are more popular. Whereas stochastic optimization methods *increase the chances* of finding global optima (in finite time), deterministic methods *find global optima without failure* (in finite time). Unfortunately, deterministic optimization still requires a deep understanding of the optimization problem and the most efficient algorithms tend to be tailored to a small set of optimization problems. However, with this work we show that deterministic optimization is at least applicable for modeling of simple batch respirometric experiments involving a single reaction. Since such experiments are typical for biological wastewater treatment process modeling, we argue that the provided parameter identification method is broadly applicable.

In addition to the nonlinear nature of the modeled processes, other factors complicating model identification include *(i)* the stochastic nature of their inputs, *(ii)* the lack of detailed understanding of metabolic pathways, and *(iii)* the large number of empirically determined parameters further leading to a lack of practical or even structural identifiability. While these issues are important, they are not addressed, diminished, or amplified by this work. Thus, we consider the experimental design and the produced experimental data as a given and focus on solving parameter estimation problems to global optimality.

Acronyms: AOB, ammonia oxidizing bacteria; DO, dissoved oxygen; NOB, nitrite oxidizing bacteria; ODE, ordinary differential equations; OUR, oxygen uptake rate; QP, quadratic program; TNN, total nitrite nitrogen; WLS, weighted least squares; WRMSR, weighted root mean squared residuals.

* Corresponding author. Eawag – Swiss Institute for Aquatic Science and Technology, Process Engineering, Überlandstrasse 133, P.O. Box 611, CH-8600 Dübendorf, Switzerland.
E-mail address: kris.villez@eawag.ch (K. Villez).

http://dx.doi.org/10.1016/j.envsoft.2016.06.015
1364-8152/

To showcase the real-world applicability of the developed optimization method, a data set collected for the purpose of kinetic model identification of a biological urine treatment process is used. Separate collection and treatment of urine is a new approach to optimize sanitation. Two possible applications are the recovery and recycling of nutrients to agriculture (Udert and Wächter, 2012) and the prevention of corrosion in sewers by nitrate dosage (Jiang et al., 2011; Oosterhuis and van Loosdrecht, 2009). Nitrification of urine is applied in both approaches, either to stabilize volatile ammonia or to produce the electron acceptor nitrate. Stable nitrification requires balanced activities of both bacterial groups involved in the process, ammonium oxidizing bacteria (AOB) and nitrite oxidizing bacteria (NOB). However, stable nitrification is challenging in urine due to the high pH value and the high concentrations of ammonia, organic substances, and salts. Three major process failures can occur (Fumasoli, 2016). First, both AOB and NOB are inhibited at high pH values due to high concentrations of free ammonia. Second, at intermediate pH values AOB grow too fast and produce large amounts of nitrite, which inhibit NOB. Third, acid-tolerant AOB grow in when the operational pH is low. In turn, the pH value can decrease even further leading to the chemical production of large amounts of volatile nitrogen compounds, especially nitric oxide (Fumasoli et al., 2015). The main operational parameter is the pH value. It directly influences the energy generation of the bacteria, but it also determines *(i)* limitation effects by free ammonia and carbonate and *(ii)* inhibition effects by free ammonia and nitrous acid (Fumasoli, 2016). Keeping these effects apart and determining the respective kinetic constants is challenging. Consequently, mechanistic computer models can be a helpful tool to include all effects and the necessary chemical and microbial processes (Fumasoli, 2016). Jubany et al. (2005) showed that consecutive dosage of nitrite and fitting the oxygen uptake rate can be used to determine the kinetics of NOB in high-strength ammonia wastewaters. This approach to experimental data collection is also applied in our study in order to demonstrate our optimization algorithm.

The next section describes the experimental data and the applied optimization algorithm. Afterwards, results are shown and discussed in separate sections. The major conclusions are summarized at the end.

2. Materials and methods

2.1. Notation and symbols

The notation conventions applied in this study are given in Table 1. All symbols used in this study are given in Table 2. In addition, inequalities of the form $\boldsymbol{x} \leq \boldsymbol{y}$ express that every element in $\boldsymbol{x}$ is smaller or equal to the corresponding element in $\boldsymbol{y}$, i.e. $\boldsymbol{x} \leq \boldsymbol{y} \Leftrightarrow \forall l : x_l = \boldsymbol{x}(l) \leq \boldsymbol{y}(l) = y_l$. Similarly, we write for matrices that $\boldsymbol{X} \leq \boldsymbol{Y} \Leftrightarrow \forall l, m : X_{l,m} = \boldsymbol{X}(l,m) \leq \boldsymbol{Y}(l,m) = Y_{l,m}$.

Table 1
Notation conventions.

Notation	Description
x, θ	Scalar
$\boldsymbol{x}$, $\boldsymbol{x}_m$, $\boldsymbol{\theta}$	Column vector
x_l, $\boldsymbol{x}(l)$	lth scalar element of vector $\boldsymbol{x}$
$X_{l,m}$, $\boldsymbol{X}(l,m)$	Scalar element of matrix $\boldsymbol{X}$ at row l and column m
$\boldsymbol{X}$	Matrix
$\boldsymbol{X}_{l,:}$	lth row from matrix $\boldsymbol{X}$
$\boldsymbol{X}_{:,m}$	mth column from matrix $\boldsymbol{X}$
$\tilde{x}$	Measurement
$\hat{x}$	Optimal estimate
$\breve{x}$, $\hat{x}$	Relaxed estimate or value
x^+	Positive part of x ($\max(x,0)$)
x^-	Negative part of x ($\min(x,0)$)
$\underline{x}$	Lower bound (h,h_j,h_1,h_2), lower interval limit (q,s,θ)
$\overline{x}$	Upper bound (h,h_j,h_1,h_2), upper interval limit (q,s,θ)
$\mathscr{X}$	Set of feasible solutions

2.2. Assumed model structure and general problem statement

The parameter optimization method as developed in this work applies to process models whose dynamics can be formulated as follows:

$$\dot{s}(t) = -q(s(t), \boldsymbol{\theta}), \quad s(0) = 1 \tag{1}$$

with $s(t)$ the single state variable and $q(s(t),\boldsymbol{\theta})$ a single rate expression. The state variable can only take on nonnegative values ($s(t) \geq 0$) and the rate expression $q(s(t),\boldsymbol{\theta})$ is nonnegative and nonincreasing in its parameters ($\boldsymbol{\theta}$, dimensions: $p \times 1$) over its whole domain:

$$\forall s \in \mathbb{R}_{\geq 0}, \forall \boldsymbol{\theta}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \mathbb{R}^p : \begin{cases} q(s,\boldsymbol{\theta}) \geq 0 \\ \boldsymbol{\theta}_1 \leq \boldsymbol{\theta}_2 \Leftrightarrow q(s,\boldsymbol{\theta}_1) \geq q(s,\boldsymbol{\theta}_2) \end{cases} \tag{2}$$

The process state and/or the rate of change ($s(t)$ and $q(t)$) are measured through equations of the following form:

$$\tilde{y}_{j,k_j} = y_{j,k_j} + e_{j,k_j}, \quad e_{j,} \sim \mathscr{N}\left(0, \sigma_{j,k_j}\right) \tag{3}$$

$$y_{j,k_j} = \beta_{j,0} + \beta_{j,1} s\left(t_{k_j}\right) + \beta_{j,2} q\left(s\left(t_{k_j}\right), \boldsymbol{\theta}\right) = \left[1 \;\; s\left(t_{k_j}\right) \;\; q\left(s\left(t_{k_j}\right), \boldsymbol{\theta}\right)\right] \boldsymbol{\beta}_j, \quad j = 1,\ldots,J, \; k_j = 1,\ldots,K_j \tag{4}$$

$$\boldsymbol{\beta}_j \in \Omega_j \subset \mathbb{R}^3_{\geq 0} \quad j = 1,\ldots,J \tag{5}$$

These measurement equations deliver K_j measurements $\tilde{y}_{j,k_j}$ of J measured variables y_{j,k_j} at sampling times t_{k_j}, where $k_j = 1,\ldots,K_j$ and $j = 1,\ldots J$. The measurement errors e_{j,k_j} are assumed to be sampled independently from zero mean normal distributions with standard deviations σ_{j,k_j}. These standard deviations are assumed known. In addition, the vectors $\boldsymbol{\beta}_j = \left[\beta_{j,0} \quad \beta_{j,1} \quad \beta_{j,2}\right]^{\mathrm{T}}$ ($j = 1,\ldots,J$) are bound to belong to a subset of the nonnegative real space, Ω_j. These subsets are assumed known and are required to be convex. The vectors $\boldsymbol{\theta}$ and $\boldsymbol{\beta}_j$ ($j = 1,\ldots J$) constitute the parameters of the model and are to be estimated.

2.3. Parameter estimation methods

2.3.1. Definition of optimality

We define optimal parameter estimation as maximum likelihood estimation, that is, we aim to find the values for the parameters which maximize the likelihood. Let $\boldsymbol{\gamma}$ denote the vector containing all parameters:

$$\boldsymbol{\gamma} = \left[\boldsymbol{\theta}^{\mathrm{T}} \quad \boldsymbol{\beta}_1^{\mathrm{T}} \quad \boldsymbol{\beta}_2^{\mathrm{T}} \quad \ldots \quad \boldsymbol{\beta}_j^{\mathrm{T}} \quad \ldots \quad \boldsymbol{\beta}_J^{\mathrm{T}}\right]^{\mathrm{T}} \tag{6}$$

and let $h(\boldsymbol{\gamma})$ be the negative log-likelihood function. The optimization problem is then:

$$\hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}} h(\boldsymbol{\gamma}) \tag{7}$$

Given assumptions and definitions discussed above, the negative log-likelihood corresponds to the following weighted least