Available at www.sciencedirect.com

ScienceDirect

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

Linear parameter estimation of rational biokinetic functions

*T.G. Doeswijk, K.J. Keesman**

Systems and Control Group, Wageningen University, P.O. Box 17, 6700 AA Wageningen, The Netherlands

ARTICLE INFO

Article history:
Received 22 May 2008
Received in revised form
18 September 2008
Accepted 7 October 2008
Published online 18 October 2008

Keywords:
Michaelis–Menten
Substrate inhibition
Activated sludge
Estimation
Total least squares
Bias compensation

ABSTRACT

For rational biokinetic functions such as the Michaelis–Menten equation, in general, a nonlinear least-squares method is a good estimator. However, a major drawback of a nonlinear least-squares estimator is that it can end up in a local minimum. Rearranging and linearizing rational biokinetic functions for parameter estimation is common practice (*e.g.* Lineweaver–Burk linearization). By rearranging, however, the error is distorted. In addition, the rearranged model frequently leads to a so-called 'errors-in-variables' estimation problem. Applying the ordinary least squares (OLS) method to the linearly reparameterized function ensures a global minimum, but its estimates become biased if the regression variables contain errors and thus bias compensation is needed. Therefore, in this paper, a bias compensated total least squares (CTLS) method, which as OLS is a direct method, is proposed to solve the estimation problem. The applicability of a general linear reparameterization procedure and the advances of CTLS over ordinary least squares and nonlinear least squares approaches are shown by two simulation examples. The examples contain Michaelis–Menten kinetics and enzyme kinetics with substrate inhibition. Furthermore, CTLS is demonstrated with real data of an activated sludge experiment. It is concluded that for rational biokinetic models CTLS is a powerful alternative to the existing least-squares methods.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Rational biokinetic functions frequently appear in studies with a biological component, as in *e.g.* activated sludge models and other biotechnological applications, in metabolic pathway modeling, gene engineering, and so on. A simple example of a rational biokinetic function is the well-known Michaelis–Menten equation:

$$v = \frac{v_{max} S}{K_m + S} \tag{1}$$

Parameters that appear nonlinearly in a rational mathematical model, as in this case v_{max} and K_m, can be estimated by linear techniques after rearranging and reparameterizing the equations. Several types of linearization like the so-called Lineweaver–Burk (L–B) and Eady–Hofstee (E–H) linearizations (Doran, 1995) have been proposed in the past and are widely used in practice. It is recognized, however, that by rearranging the model in a linear-in-the-parameters form the error becomes distorted (Cornish-Bowden, 2002). A nonlinear least squares (NLS) approach, using iterative search techniques, avoids this error distortion introduced by linearization. It, therefore, often results in a more facile determination of the kinetic constants (Garfinkel et al., 1977). In addition to this, accurate parameter estimates may result when applying an optimal input design (Keesman and Stigter, 2002; Stigter and Keesman, 2004). However, proper initial estimates are required as local minima do frequently occur (Yildirim et al., 2003; Doeswijk, 2007). As a consequence of this, wrong model predictions can easily occur. In some extreme cases, NLS can even result in a loop whenever two local minima are situated close to each other and when the corresponding sum of squares of the residuals are equal.

* *Corresponding author*. Tel.: +31 317 484952; fax: +31 317 484957.
E-mail addresses: timo.doeswijk@wur.nl (T.G. Doeswijk), karel.keesman@wur.nl (K.J. Keesman).
0043-1354/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.watres.2008.10.019

Nomenclature

C	dissolved oxygen concentration
e	measurement error
K_c	Monod half-saturation constant for dissolved oxygen
K_i	inhibition constant
K_m	Michaelis–Menten constant
K_s	Monod half-saturation constant for substrate
N	number of samples
$\mathcal{N}(\bullet)$	normal distribution
p	parameter vector
r	respiration rate
$\mathbb{R}$	set of real numbers
S	substrate concentration
t	time
v	substrate conversion rate
V_m	maximum substrate conversion rate
X	regression matrix
X_{BH}	biomass concentration
y	measurement
Y	yield coefficient
θ	reparameterized parameter vector
σ	standard deviation
μ_m	maximum specific growth rate
Γ	covariance matrix
Φ	data matrix

Subscripts

k	discrete-time index

Superscripts

0	nominal value
T	transpose of a vector or matrix
~	transformed variable
^	estimated value

A natural solution would be to use a multi-start procedure. However, even with multi-start procedures, a global minimum cannot be assured nor an approximation error in terms of the parameter estimates or the sum of squares can be given. Hence, the relevance of reparameterization methods that result in a linear estimation problem, allowing a direct solution with inherent unique global minimum.

Usually, without any special care, biased estimates result when regression variables contain errors (see *e.g.* Norton, 1986). Let us therefore have a closer look at Eq. (1). We note that usually the rate v in Eq. (1) is assumed to be the uncertain variable. But, as the rate is obtained from measured substrate concentrations, it is reasonable to expect that the substrate concentration S at the right-hand side of Eq. (1) is also uncertain. This assumption makes parameter estimation related to the Michaelis–Menten equation a so-called errors-in-variables problem. Consequently, after a linear reparameterization of Eq. (1), an ordinary least squares (OLS) solution to the resulting error-in-variables problem leads to biased estimates. An appropriate direct solution to these type of problems is given by total least squares (TLS) methods (Golub and Van Loan, 1980; Van Huffel and Vandewalle, 1991). However, nonlinearities in the uncertain data, in our case as a result of the linear reparameterization, may still lead to biased estimates (see Box, 1971). Hence, bias compensation must be introduced. As a consequence of all of this, a bias compensated total least squares (CTLS) estimator (Vandersteen, 1998), that is asymptotically unbiased for static polynomial functions and a direct method, is proposed in this paper and applied to some relevant biokinetic models.

The objective of the paper is to illustrate the wide applicability of our linearly reparameterizing method for rational (biokinetic) functions and to evaluate the performance of CTLS on these functions. As for NLS, we will not provide the technical details of CTLS. As yet, we believe that it suffices to show how the ingredients for CTLS, as data and covariance matrix, can be specified for some real cases. The outline of the paper is as follows. First, a general approach to linearly reparameterize rational functions is proposed and the handling of data errors is discussed in some more detail. Subsequently, the procedure, using CTLS for the estimation of the parameters, is firstly illustrated to a model with Michaelis–Menten kinetics. Furthermore, to show the wide applicability of our model reparameterization and CTLS, the parameters of enzyme kinetics with substrate inhibition are estimated as well. Finally, the practical usability is shown by estimating the biokinetic parameters in an activated sludge model from real respirometric data.

2. Preliminaries

2.1. Standard least-squares estimation

Consider a linear regression model, written in matrix notation as

$$y = X\theta + e \tag{2}$$

where for $t = 1, \ldots, N$ we define the measured output vector $y := [y(1), \ldots, y(N)]^T$, error vector $e := [e(1), \ldots, e(N)]^T$, which both are column vectors of length N, and parameter vector $\theta := [\theta_1, \ldots, \theta_p]^T$ with $N \geq p + 1$. Furthermore, X is the $N \times p$ data matrix, containing the regressor values at different time instances t. It is widely known that the *ordinary least-squares estimate* of θ is given by

$$\widehat{\theta} = (X^T X)^{-1} X^T y \tag{3}$$

under the assumption that the $p \times p$ matrix $X^T X$ is invertible.

For nonlinear-in-the-parameter models of the form

$$y = f(X, \theta) + e \tag{4}$$

with $f(X, \theta)$ a vector function relating the explanatory variables to the output, no direct solutions to the estimation problem exist. The estimates are found after minimizing a predefined objective function. Commonly, this objective function is chosen as

$$J(\theta) = \varepsilon(\theta)^T \varepsilon(\theta) = \sum_{t=1}^{N} \varepsilon^2(\theta; t) \tag{5}$$

where $\varepsilon(\theta) = y - f(X, \widehat{\theta})$ is an N-dimensional vector that contains the parameter dependent prediction errors. This function is also indicated as the sum of squares and is, in mathematical jargon, the squared 2-norm of the prediction

Download English Version:

https://daneshyari.com/en/article/4483273

Download Persian Version:

https://daneshyari.com/article/4483273

Daneshyari.com