Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Marine Systems 55 (2005) 205–221

JOURNAL OF MARINE SYSTEMS

www.elsevier.com/locate/jmarsys

Refined parameter and uncertainty estimation when both variables are subject to error. Case study: estimation of Si consumption and regeneration rates in a marine environment

Anouk de Brauwere[a,*], Fjo De Ridder[b], Marc Elskens[a], Johan Schoukens[b], Rik Pintelon[b], Willy Baeyens[a]

[a]*Department of Analytical and Environmental Chemistry, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium*
[b]*Department of Electricity and Instrumentation, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium*

Received 27 December 2003; accepted 21 September 2004
Available online 2 December 2004

Abstract

The problem of estimating parameters and their uncertainty from experimental measurements in marine ecosystems is a common task and often necessitates solving nonlinear equations. If the measurements are subject to individually varying errors (i.e., heteroscedastic data), the parameters are often estimated using a Weighted Least Squares (WLS) method. For estimating the parameter uncertainties, a linearized expression for the covariance matrix exists. Yet, both methods assume that the errors on the independent variable, also called "input", is negligible, which is often not true. For instance, in order to determine uptake and regeneration rates of silicic acid by phytoplankton, concentration and isotopic abundance measurements are performed at the beginning (input) and at the end (output) of an incubation experiment. Here, the so-called input and output are measurements of the same quantities, i.e., determined in exactly the same way, only differing by the time at which the measurements were performed. Clearly, there is no reason to assume that the input measurements are subject to less error than the output measurements. We propose a refinement of the two abovementioned estimation methods which enlarges their applicability to cases where input noise is not negligible. The refined methods are evaluated on the uptake and regeneration processes of silicic acid and compared to the original procedures using Monte-Carlo simulations. The results reveal a smaller bias for the refined WLS estimator compared with the original one. An additional advantage of using the refined WLS cost function is that its residual value can be interpreted as a sample from a χ^2 distribution. This property is especially useful because it enables an internal quality control of the results. In addition, the parameter uncertainty estimation is significantly improved. By neglecting the effect of the input noise, a (potentially) important origin of the parameter variation is simply ignored. Therefore, without the refinement, the parameter uncertainties are systematically

* Corresponding author. Tel.: +32 2 629 3264; fax: +32 2 629 3274.
E-mail address: adebrauw@vub.ac.be (A. de Brauwere).

doi:10.1016/j.jmarsys.2004.09.006

underestimated. Using the refined method, this systematic error disappears and on the whole, the parameter standard deviations are accurately estimated.

Keywords: Marine environment; Monte-Carlo simulations; Nonlinearity; Silicon cycling; Uncertainty estimation; Weighted Least Squares method

1. Introduction

Studying natural systems often involves the determination of physical parameters which are not experimentally accessible, e.g., rate constants. Generally, in order to quantify these parameters, some other, related system variables are measured instead. Then, using a chosen formal relation between the unknown parameters and the measurable quantities (a so-called model), the latter can finally be calculated.

This apparently straightforward procedure is often complicated by

(i) the presence of noise on all measurements and
(ii) nonlinear model structures,

which are often encountered simultaneously when studying natural phenomena. Yet, the well-established parameter and uncertainty estimation procedures are either designed for linear models or imply the absence of noise for one of the variables, usually called the independent or input variable. In this context, this article presents a method which deals with both aspects simultaneously. In fact, it corresponds to a refinement of existing nonlinear estimation methods, adapting them for those cases where input noise cannot be neglected.

In the literature, the case of nonlinear models is widely discussed. To put it simply, nonlinear models complicate the estimation of parameters and their uncertainty because analytic expressions for the parameters as a function of the measurements are not available anymore. Concerning the parameter estimation, effort is usually concentrated upon finding appropriate numerical optimization routines. That is to say, the objective is to find the optimal parameter values corresponding to the global minimum of the cost function, as fast as possible and without diverging or "falling" in local minima (Marsili-Libelli, 1992; Chan et al., 1993; Tang and Wang, 2002). Usually, it boils down to finding the best compromise between these criteria, which differs for each specific model structure considered. Besides the estimation of the parameters, attention has also been directed to the parameter uncertainty estimation, especially using linearized approximations (Donaldson and Schnabel, 1987; Marsili-Libelli et al., 2003). The accuracy of the uncertainty estimation is found to be related to "how nonlinear the model is" (Bates and Watts, 1980; Donaldson and Schnabel, 1987; Grimstad et al., 2001).

Nevertheless, to our knowledge, none of the nonlinear estimation procedures explicitly take into account the individual input noise influence. Instead, they either suppose all input and output variables to have the same noise variance, as in methods like Perpendicular Least Squares and Model II (also called "reduced major-axis") regression (Riggs et al., 1978; Jones, 1979; Valsami et al., 2000). Or they presume the input variations to be negligible compared to the output variations and hence simply ignore their effect, as in the traditional Weighted Least Squares (WLS) method (Valsami et al., 2000). However, in many cases, these assumptions of identical noise variances or negligible input noise are not valid. For instance, when analyzing chemical reaction data, it often happens that the so-called output is a measurement of a chemical concentration, whilst one of the input variables is the initial concentration of the same material, i.e., determined in exactly the same way. If in that case the absence of input noise is assumed, a potential bias is introduced in the estimation results. This will be illustrated in the current article. The individual effects of input and output noise should be taken into account during the parameter estimation and certainly during the parameter uncertainty estimation. Methods have been proposed using parameter transformations in order to make the negligible input noise assumption approximately fulfilled (Mandel, 1984) or to make the model as linear as possible

(Bates and Watts, 1981). These methods are actually trying to minimize the error made by neglecting the input noise contribution. To our knowledge, procedures that really take the input noise information into account have only been developed for linear models. Clutton-Brock (1967) and Chandler (1972) proposed a parameter estimation procedure that takes into account the linearized contribution of the input noise. In Cecchi (1991) and Moreno and Bruzzone (1993), the same reasoning is extended to the estimation of the parameter uncertainties, but again for linear models only. The fact that even these linear estimation procedures are not yet commonly used, is illustrated by their absence in statistical software.

To summarize, it seems that no method is readily available to estimate parameters and their uncertainty, taking into account all measurement variations when the model is nonlinear, although it is frequently encountered in studies of natural phenomena. This article is an attempt to fill this gap. That is to say, the existing nonlinear Weighted Least Squares (WLS) and the Linearized Covariance Matrix (LCM) procedures are refined, especially adapting them to take into account input variations.

The refined methods for estimating the parameters and their uncertainty are applied to a simple nonlinear compartmental model describing the consumption and regeneration of silicic acid in aquatic systems (see Section 5.1). A better understanding of silica recycling in pelagic waters is very important. Indeed, the role of diatoms in global marine primary production and in the export of organic matter has been confirmed (Nelson et al., 1995; Buesseler, 1998). Diatoms are major contributors to primary production in oceanic regions and in coastal upwelling systems (e.g., Dugdale and Wilkerson, 1998; Brzezinski et al., 2003).

The performance of the refined methods is demonstrated by comparing the results to Monte-Carlo simulations. In Section 2, the general problem statement is developed. This section defines the general framework and introduces the concepts of nonlinear models and noisy measurements. The traditional estimation methods are discussed in Section 3. The proposed refinement is then shortly presented in Section 4. The performance evaluation procedure is explained in Section 5. The results of this performance assessment are finally presented and discussed in Section 6.

2. Problem statement

The purpose of this section is to sketch a global view of all problems to be tackled. In an attempt to systematize the whole, the different steps leading to parameter and uncertainty estimation are discussed separately, considering the special assumptions made and the difficulties arising at each stage.

2.1. The system

According to Nihoul (1975), to model a marine system, it is first required to define the system without ambiguity, separate it from the outside world and identify the exchanges between the system and the exterior. In our case, the system consists of internal processes as well as interactions with the environment, and as a result, some measurable variables fluctuate dependently. This means that if one or several input variables, $\mathbf{x}_0$, are varied, one or several output variables, $\mathbf{y}_0$, seem to change as a result:

$$\mathbf{y}_0 = \text{system}(\mathbf{x}_0). \tag{1}$$

This situation is schematically represented in Fig. 1a. The question usually is to identify and quantify the processes that cause this relation.

2.2. The measurements

In order to characterize the system, it is necessary to collect information about it. From the previous paragraph, it follows logically that this is done by measuring the input and output variables. Due to all kinds of fluctuations, the measured values $\mathbf{x}$ and $\mathbf{y}$ generally are not equal to the true values $\mathbf{x}_0$ and $\mathbf{y}_0$. Consider the following additive noise model,

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{n}_\mathbf{x} \tag{2a}$$

and

$$\mathbf{y} = \mathbf{y}_0 + \mathbf{n}_\mathbf{y} \tag{2b}$$

where $\mathbf{n}_\mathbf{x}$ and $\mathbf{n}_\mathbf{y}$ represent the input and output noise vectors.

In many applications, it is explicitly assumed that $\mathbf{n}_\mathbf{x}$ is zero or small with respect to $\mathbf{n}_\mathbf{y}$ (see Mandel, 1984 for a meticulous formulation), and as a consequence, the input noise contribution can be

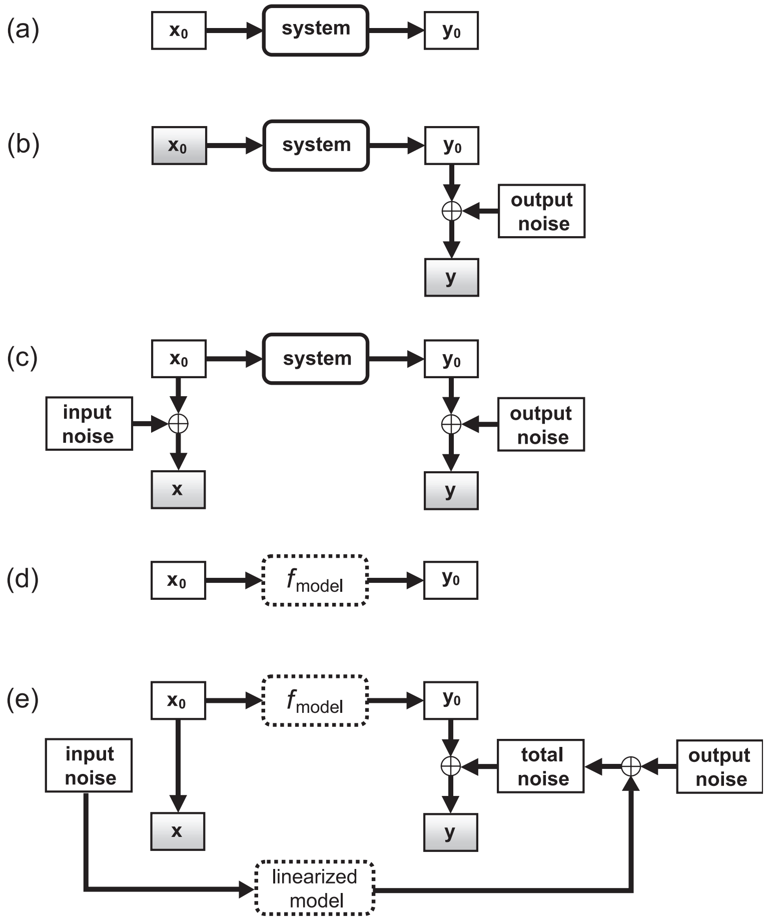

Fig. 1. Schematic representations of the system, model, measurements and noise concepts. Meaning of the symbols: the rounded boxes stand for transformations, the others are variables. The grey boxes are the measured variables. (a) Illustration of the system concept. (b) The measurements, in the situation of only output noise, or negligible input noise. (c) The measurements, in the situation of nonnegligible noise on input and output. (d) Illustration of the concept "model". The difference with (a) is that the model is a mathematical expression, whereas the system is not. (e) Scheme of the noise transformation underlying the refined total noise estimation. In fact, it is as if there were no input noise added to the input variables, instead the input noise is transformed into output noise and added to the original output noise. The resulting total noise is treated as if it were merely output noise, but actually it represents an estimation of the total variation.

neglected (Fig. 1b). In this article, the more general situation is considered: input noise cannot be neglected (Fig. 1c). The noise elements $n_{x,i}$ and $n_{y,i}$ are assumed to be independent $N(0, \sigma_{x,i})$ and $N(0, \sigma_{y,i})$ distributed random variables (i=1, ..., N). The variances $\sigma^2_{x,i}$ and $\sigma^2_{y,i}$ are known or estimated and are stored in the vectors $\sigma^2_{\mathbf{x}}$ and $\sigma^2_{\mathbf{y}}$.

2.3. The model

The next step in the identification procedure is to select a parametric model for the system. Whereas the system is merely a conceptual, qualitative idea, the model is a formal, mathematical description of the system processes under study (Fig. 1d). The model relates the true input and output measurements via the mathematical expression f_{model} and the true model parameters $\boldsymbol{\theta}_0$:

$$\mathbf{y}_0 = f_{\text{model}}(\mathbf{x}_0, \boldsymbol{\theta}_0). \tag{3}$$

From expression (3), it is clear that the model enables to predict the output measurements, given the input measurements and the model parameters.

Many types of models exist. The models considered in this article are nonlinear in the parameters. This means that they cannot be written as

$$f_{\mathrm{model}}(\mathbf{x}, \boldsymbol{\theta}) = K(\mathbf{x})\boldsymbol{\theta}. \quad (4)$$

For instance, a model $\boldsymbol{f}_{\mathrm{model}}(\boldsymbol{x},\ \boldsymbol{\theta})=\boldsymbol{\theta}_1\boldsymbol{x}+\boldsymbol{\theta}_2\boldsymbol{x}^2$ is linear in the parameters, whereas a model $\boldsymbol{f}_{\mathrm{model}}(\boldsymbol{x},\ \boldsymbol{\theta})=(\boldsymbol{\theta}_1+\boldsymbol{\theta}_2\boldsymbol{x})/(\boldsymbol{\theta}_3+\boldsymbol{\theta}_4\boldsymbol{x}^2)$ is not. Furthermore, it is assumed that no model errors are present. This means that $\boldsymbol{f}_{\mathrm{model}}$ is the true description of the system processes (in practice a sufficiently good description), and that the remaining difference between predicted model value and measurement is only of stochastic origin, i.e., due to random measurement variation (discussed in Section 2.2).

Once the model type, i.e., $\boldsymbol{f}_{\mathrm{model}}$, has been chosen, the problem of determining the system processes can be reformulated as quantifying the set of model parameters $\boldsymbol{\theta}$.

3. Absence of input noise: existing estimation methods

In this section, existing parameter and parameter uncertainty estimation methods are briefly discussed. These methods assume the absence of input noise. In Section 4, they will be adapted to cope with situations where input noise cannot be neglected. But first, the original methods are outlined.

3.1. The parameter estimator

From the previous paragraphs, it should be clear that after defining the system, carrying out the measurements, and choosing the model, the remaining step is to combine the model and the measurements in order to estimate the model parameters.

If input noise is negligible, the Weighted Least Squares (WLS) estimator is commonly used (Bard, 1974; Valsami et al., 2000). The WLS estimator of θ is the parameter value, denoted by $\hat{\boldsymbol{\theta}}_{\mathrm{WLS}}$, that minimizes the WLS cost function V_{WLS}:

$$\hat{\boldsymbol{\theta}}_{\mathrm{WLS}} = \arg\min_{\boldsymbol{\theta}} V_{\mathrm{WLS}} \quad (5)$$

with

$$V_{\mathrm{WLS}} = \frac{1}{2}\sum_{i=1}^{N}\frac{(y_i - f_{\mathrm{model}}(\mathbf{x},\boldsymbol{\theta}))^2}{\hat{\sigma}_{y,i}^2} = \frac{1}{2}\sum_{i=1}^{N}\frac{e_i^2}{\hat{\sigma}_{y,i}^2} = \frac{1}{2}\mathbf{e}^T\mathbf{W}\mathbf{e}. \quad (6)$$

N is the number of data, $\mathbf{e}$ stands for the difference, or "error", between the output measurements $\mathbf{y}$ and the model prediction $f_{\mathrm{model}}(\mathbf{x}, \boldsymbol{\theta})$. $\mathbf{W}$ represents a diagonal weighting matrix containing the inverse of the output variances $\hat{\sigma}_{y,i}^2$,

$$\mathbf{W} = \begin{pmatrix} \frac{1}{\hat{\sigma}_{y,1}^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\hat{\sigma}_{y,2}^2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \frac{1}{\hat{\sigma}_{y,N}^2} \end{pmatrix} \quad (7)$$

Note that in the WLS cost function (6), the squared error is "weighted" by the corresponding *output* variance. It allows to emphasize measurements subject to a small uncertainty and to suppress measurements with a large uncertainty (Massart et al., 1997; Pintelon and Schoukens, 2001).

However, by only considering the output variances, it is presumed that $\hat{\sigma}_{y,i}^2$ is a good estimate of the variance of the error, $y_i - f_{\mathrm{model}}(\mathbf{x},\boldsymbol{\theta})$, in Eq. (6) and hence that the input variations are negligible compared with the noise on the output. If this assumption is violated, this can lead to a bias in $\hat{\boldsymbol{\theta}}_{\mathrm{WLS}}$ (Powell and Macdonald, 1972). In Section 4, a refinement is suggested that better adapts the WLS estimator for those cases.

3.2. A note on the residual cost function value

In the previous paragraph, we saw that $\hat{\boldsymbol{\theta}}_{\mathrm{WLS}}$ corresponds to that parameter value for which the WLS cost function in Eq. (6) is minimal. Due to output noise, even for the optimal parameters there will be a residual difference between the measurements $\mathbf{y}$ and the predicted model values $f_{\mathrm{model}}(\mathbf{x},\ \hat{\boldsymbol{\theta}}_{\mathrm{WLS}})$. Therefore, the minimal value of the WLS cost function $V_{\mathrm{WLS}}(\hat{\boldsymbol{\theta}}_{\mathrm{WLS}})$ will not be zero, but equal to a residual value. Under the previously stated noise assumptions, this residual value has the interesting feature of being a sample from a χ^2 distribution with $N - n_\theta$

degrees of freedom (Box, 1970; Rod and Hancil, 1980), where N stands for the number of measurements and n_θ for the number of free parameters in the model. Formally,

$$V_{\mathrm{WLS}}(\hat{\boldsymbol{\theta}}_{\mathrm{WLS}}) \sim \chi^2_{N-n_\theta}. \tag{8}$$

Using this information, it is possible to assess the "probability" of a residual cost function value. If the residual value falls beyond a chosen confidence limit, the results should be rejected for being "unlikely". For instance, if the residual value is significantly higher than expected, the remaining difference between model and measurements is too high to be explained only by stochastic measurement noise. This significant difference between the expected and observed residual cost function is an indicator of systematic errors. These can be due to (i) model errors, (ii) artifacts in the measurements or (iii) ill estimation of the experimental uncertainties. In order to determine exactly which of these three is the cause of the unexpected residual cost function, it is necessary to analyze replicate data sets.

To summarize, using a WLS cost function has the additional benefit of providing an internal quality control of the results, via interpretation of its residual value. Yet, note that this property is only true as long as the assumption of negligible input noise holds.

3.3. The parameter uncertainty estimation

In order to interpret the estimated parameter values, it is necessary to assess the uncertainty associated with them. For nonlinear models, straightforward "error propagation" calculations are generally impossible since no explicit expression for the parameters as a function of the measurements is available.

Instead, Monte-Carlo, bootstrap or jackknife simulations are often used to estimate the parameter uncertainty (Meinrath et al., 2000 and references therein). These procedures consist of simulating a large number of data sets, based on the true measurements, but either with a different simulated random noise realization (Monte-Carlo) or with a different sampling from the original measurement record (bootstrap and jackknife). For each such "synthetic data set", the optimal parameters are estimated. This results in a distribution of best-fit parameter values from which the statistical properties can be examined.

Usually, these methods are said to deliver accurate results as long as enough simulations are performed (e.g., Massart et al., 1997). However, it is often not mentioned that an additional condition for reliable results is the use of an accurate parameter estimator. Indeed, these simulation methods use perturbations around the optimized parameter values. Therefore, it is implicitly assumed that these parameter values are unbiased. For instance, if input noise is negligible, the original WLS estimator (cf. Section 3.1) is unbiased and can be used. However, if input noise is not negligible, this is not true anymore and another, unbiased estimator should be chosen. In short, these simulation methods are useful tools to assess parameter uncertainties, as long as some conditions are met. Their main disadvantages are that they are (i) time-consuming, due to the large number of simulations to be processed, and (ii) difficult to incorporate in an automated routine calculation, because, e.g., optimization problems may occur for particular noise realizations or resamplings.

Alternatively, procedures based on an approximation of the parameter covariance matrix (Cov) often provide reasonable parameter uncertainty estimates. We consider the following first order approximation of Cov, assuming that the model is approximately linear with respect to the parameters near the optimal values:

$$\mathrm{Cov}(\hat{\boldsymbol{\theta}}) = \left(\mathbf{J}(\hat{\boldsymbol{\theta}})^T \mathbf{W}\mathbf{J}(\hat{\boldsymbol{\theta}})\right)^{-1}, \tag{9}$$

where $J(\hat{\boldsymbol{\theta}})$ stands for the Jacobian matrix of the model function $f_{\mathrm{model}}(\mathbf{x}, \boldsymbol{\theta})$ at $\hat{\boldsymbol{\theta}}$,

$$\mathbf{J} = \left.\frac{\partial f_{\mathrm{model}}(\mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}. \tag{10}$$

The diagonal elements of $\mathrm{Cov}(\hat{\boldsymbol{\theta}})$ are the parameter variances. Using the Linearized Covariance Matrix expression (9) for estimating the parameter uncertainty has the benefit of being very easy and fast to compute (Donaldson and Schnabel, 1987), e.g., it can usually be integrated in the parameter optimization algorithm as a final calculation. The main disadvantage is that in case of nonlinear models, the calculated parameter variances are only approximations. However, expression (9) has been proven to be asymptotically exact (e.g., Pintelon and Schoukens, 2001), i.e., for increasing number of measurements Eq. (9) approaches the true covariance matrix, even if the model is nonlinear. If the number of data is small, the accuracy of the approximation varies for each data set and each model, depending on the "degree of nonlinearity" and the signal-to-noise ratio of the model-measurement combination.

Finally, note that **W** in Eq. (9), as defined in Eq. (7), is composed only by the output variances $\hat{\sigma}_{y,i}^2$. Hence, it is again implicitly assumed that the input variations are negligible. Obviously, if this assumption is not valid, a significant origin of variation is ignored and the estimated parameter variances will underestimate the true variations. For these cases, expression (9) will be refined in Section 4 in order to include the input noise contribution to the parameter variation.

Other approximations for Cov exist, e.g., using the Hessian matrix instead of the Jacobian. However, it has been shown that these computationally more expensive and numerically less stable alternative procedures do not improve the uncertainty estimation significantly (Donaldson and Schnabel, 1987; Varah, 1990).

4. Presence of input noise: refinement of the existing methods

In Section 3, the Weighted Least Squares parameter estimator and the Linearized Covariance Matrix expression are introduced. Both are based on the assumption that input noise is negligible compared with output noise. This Section presents a refinement of the abovementioned methods in order to include the input noise effect as well. The proposed adjustment is based on a refinement of the weighting matrix **W** (Eq. (7)), which appears in both the expression of the WLS cost function (Eq. (6)) and of the parameter covariance matrix (Eq. (9)). **W** should express the total variance of the error ($\mathbf{y}-f_{\text{model}}(\mathbf{x}, \boldsymbol{\theta})$). Traditionally, as only the output is supposed to be noisy, $\boldsymbol{f}_{\text{model}}(\mathbf{x}, \boldsymbol{\theta})$ is noise-free and the total variation equals the variation on **y**. In that case, **W** is composed by the inverse variances of the individual output measurements $\hat{\boldsymbol{\sigma}}_{\boldsymbol{y},\boldsymbol{i}}^2$ (cf. Eq. (7)).

Now, the aim is to take into account the input variances as well. For that, their effect on the $\boldsymbol{f}_{\text{model}}(\mathbf{x}, \boldsymbol{\theta})$ term should be known or estimated. This is achieved by linearization of the input noise contribution, as was already described by Clutton-Brock (1967) in the framework of likelihood functions. Chandler (1972) has reformulated the problem in terms of Weighted Least Squares. Fig. 1e schematically represents this adjusted noise conceptualization. In fact, the input noise is transformed into output noise via a linearized approximation, and added to the original output noise, by that way delivering an improved estimation of the total variation. Now, the refined weighting matrix, $\mathbf{W}_{\text{r}}$, is still a diagonal matrix, but composed by inverse variances $\hat{\boldsymbol{\sigma}}_{\boldsymbol{i}}^2$, which are first-order corrections of $\hat{\boldsymbol{\sigma}}_{\boldsymbol{y},\boldsymbol{i}}^2$:

$$\hat{\sigma}_i^2 = \hat{\sigma}_{y,i}^2 + \sum_{j=1}^{N} \left| \frac{\partial y_i}{\partial x_j} \right|^2 \hat{\sigma}_{x,j}^2 = \left(\mathbf{W}_{\text{r}[i]}\right)^{-1}. \tag{11}$$

Note that in the special case where input variations appear to be negligible compared to output variations, the refined weighting matrix simplifies to the original **W**, as the correction term in Eq. (11) tends to zero. Therefore, Eq. (11) is generally applicable, and it is not necessary to make a priori checks about whether or not the input noise is negligible. In conclusion, $\mathbf{W}_{\text{r}}$ is a generalization of **W** enlarging its applicability from cases where only output noise is significant to cases where all variables are disturbed by noise.

Substituting **W** by this refined $\mathbf{W}_{\text{r}}$ in the expression of the original WLS cost function (Eq. (6)), creates a refined WLS cost function, which finally takes the input variations into account. The refined WLS estimator, $\hat{\boldsymbol{\theta}}_{\text{rWLS}}$, is defined in analogy with Eqs. (5) and (6) as

$$\hat{\boldsymbol{\theta}}_{\text{rWLS}} = \arg\min_{\boldsymbol{\theta}} V_{\text{rWLS}} \tag{12}$$

with

$$V_{\text{rWLS}} = \frac{1}{2} \sum_{i=1}^{N} \frac{(y_i - f_{\text{model}}(\mathbf{x}, \boldsymbol{\theta}))^2}{\hat{\sigma}_i^2} = \frac{1}{2} \sum_{i=1}^{N} \frac{e_i^2}{\hat{\sigma}_i^2} = \frac{1}{2} \mathbf{e}^T \mathbf{W}_{\text{r}} \mathbf{e}. \tag{13}$$

If $\hat{\boldsymbol{\sigma}}_{\boldsymbol{i}}^2$ (Eq. (11)) is a good approximation of the total variation, the residual V_{rWLS} value will again be a χ^2 distributed variable with $N-n_\theta$ degrees of freedom.

The refinement is expected to have most effect on the parameter uncertainty estimation. Indeed, using $\mathbf{W}_r$ in the expression for the parameter covariance matrix (Eq. (9)), gives

$$\mathrm{Cov}(\hat{\boldsymbol{\theta}}) = \left(\mathbf{J}(\hat{\boldsymbol{\theta}})^T \mathbf{W}_r \mathbf{J}(\hat{\boldsymbol{\theta}})\right)^{-1}, \tag{14}$$

and enables more reliable parameter uncertainty estimation, not ignoring any (known) origin of variation. Since $\mathbf{W}_r$ is a generalization of $\mathbf{W}$, both refined expressions are also valid in the special case where input noise is negligible compared with output noise.

The presented refinement of the weighting matrix (Eq. (11)) that includes the input variations has certainly the benefit of its simplicity, especially because the original expressions for the WLS cost function (Eq. (6)) and the LCM (Eq. (9)) can still be used. On the other hand, it should be kept in mind that for models not linear in the input variables, the refinement is still a first-order approximation. How well this approximation works is difficult to predict. Depending on the model, the values of the parameters and the measurements, the approximation will under- or overestimate to a different extent. In Section 6, the refined estimation methods are tested on real data and on simulations.

5. Assessment of the refinement

In the previous section, we have presented a refinement of (i) the WLS parameter estimation and (ii) the LCM for estimating the parameter uncertainties. The refinement is an attempt to enlarge their applicability to experiments in which all variables are disturbed by noise, by also including the contribution of input noise.

The remainder of the article is devoted to the empirical assessment of the proposed refinements. The results are shown and discussed in Section 6. But before that, the procedures used for that assessment are outlined in the following sections. First, the system chosen to apply the methods to is introduced (Section 5.1). Next, the assessment procedures of the parameter estimation (Section 5.2), the residual cost function distribution (Section 5.3) and the parameter uncertainty estimation (Section 5.4) are described.

5.1. Uptake and regeneration of silicic acid: model and experiment

The system investigated consists of the consumption and regeneration of (dissolved) silicic acid by phytoplankton. The question is to determine the rate at which these processes occur. For that, the samples are spiked with ^{30}Si-enriched silicic acid (Beucher et al., 2004). Then, the changes in concentration and isotopic abundance of silicic acid (D) and biosilica (P) are assessed by measuring them just after the spike (t=0) and after a certain incubation period (t=t_{inc}). With isotopic abundance, α, is meant $[^{30}\mathrm{Si}]/([^{28}\mathrm{Si}]+[^{29}\mathrm{Si}]+[^{30}\mathrm{Si}]) \cdot 100\%$.

In order to extract the unknown rates from these measurements, a two-compartmental model is used, which is depicted in Fig. 2.

The differential equations associated with this model are

$$\begin{cases} \dfrac{\partial \mathrm{D}}{\partial t} = R - U \\ \dfrac{\partial æ_{\mathrm{D}}}{\partial t} = -R\dfrac{æ_{\mathrm{D}}}{\mathrm{D}} \\ \dfrac{\partial \mathrm{P}}{\partial t} = U - R \\ \dfrac{\partial æ_{\mathrm{P}}}{\partial t} = R\dfrac{æ_{\mathrm{P}}}{\mathrm{P}} + U\dfrac{æ_{\mathrm{D}} - æ_{\mathrm{P}}}{\mathrm{P}} \end{cases}. \tag{15}$$

The symbols are defined in the Glossary. The unknown parameters are the zero order rate constants R (regeneration rate) and U (uptake rate). The input variables are the measurements at the beginning of the experiment (t=0): D(0), $æ_{\mathrm{D}}(0)$, P(0) and $æ_{\mathrm{P}}(0)$. Note that the isotopic enrichment is expressed using at.% excess, æ, instead of the abundance α, because this allows a simplification of the equations. The output variables are measurements of the same quantities but performed after the incubation (t=t_{inc}): D(t), $æ_{\mathrm{D}}(t)$, P(t) and $æ_{\mathrm{P}}(t)$. Clearly, input and output measurements are equivalent and thus it is not justified to assume that the

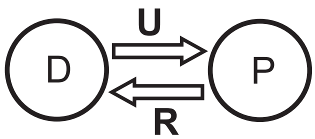

Fig. 2. The compartmental model used for estimating the uptake and regeneration rates of silicic acid. Compartments represent the pools of silicic acid (D) and biosilica (P). The arrows symbolize the rates of silicic acid uptake (U) and regeneration (R).

input variations are negligible. In addition, it should be noted that the input and the output variables, respectively, are composed of four different quantities, which are each measured in a different experiment. Consequently, the noise variance of one variable is, in general, different and independent from that of another variable (heteroscedastic data). In other words, each error term (i.e., the numerator in Eqs. (6) and (13)) should be weighed by a different variance and the weighting matrix is indeed diagonal (cf. Eq. (7)). With four output measurements and two parameters to be estimated, the system has two degrees of freedom.

When the equations (Eq. (15)) are integrated, the actual model equations, i.e., of the form $\mathbf{y}=f_{\text{model}}(\mathbf{x}, \boldsymbol{\theta})$ (cf. Eq. (3)), are obtained:

$$\begin{cases} \mathrm{D}(t) = \mathrm{D}(0) + (R-U)t \\ æ_{\mathrm{D}}(t) = (æ_{\mathrm{D}}(0) - æ_{\mathrm{P}}(0))\left(1 + \frac{R-U}{\mathrm{D}(0)}t\right)^{\frac{R}{U-R}} \\ \mathrm{P}(t) = \mathrm{P}(0) + (U-R)t \\ æ_{\mathrm{P}} = \dfrac{(æ_{\mathrm{D}}(0) - æ_{\mathrm{P}}(0))\mathrm{D}(0)}{\mathrm{P}(0) + (U-R)t}\left(1 - \left(1 + \dfrac{R-U}{\mathrm{D}(0)}t\right)^{\frac{R}{U-R}}\right) \end{cases} \tag{16}$$

Due to the second and the fourth equations, this model is nonlinear in the parameters R and U and in the input variables D(0), $æ_{\mathrm{D}}(0)$, P(0) and $æ_{\mathrm{P}}(0)$.

The basic assumptions underlying these equations are:

(1) All the Si that leaves the dissolved pool appears as particulate biosilica and Si recycling is regarded as a process that transfers Si from the particulate to the dissolved pool.
(2) Isotopic fractionation during uptake and regeneration is negligible.
(3) The regeneration of substrate consists of nonenriched compounds, i.e., the abundance of regenerated species equals the natural abundance (æ=0%).
(4) The flux rates R and U are constant during the experiment.

5.2. Assessment of the refined parameter estimation

In Section 4, a refinement of the WLS parameter estimation is proposed, which should improve the estimate in those cases where input noise may not be neglected. But how can this be verified? It is not possible to assess the accuracy of one estimation, since it is just a sample from a stochastic distribution. In other words, the optimal parameter values depend on the specific noise realization of the data; so for each noise realization, the estimator would have a different “accuracy”. Instead, the *consistency* of the estimation can be evaluated. This property of the estimation describes whether or not the optimal parameters converge in probability to the true values when the number of data tends to infinity. By taking into account an increasing number of data, the influence of noise should progressively be cancelled out. An estimator $\hat{\boldsymbol{\theta}}(N)$ is consistent if

$$\lim_{N\to\infty} P\left(|\hat{\boldsymbol{\theta}}(N) - \boldsymbol{\theta}_0| > \varepsilon > 0\right), \tag{17}$$

with ε an arbitrary small number. When both input and output variables are disturbed by noise, the original WLS estimator, $\hat{\boldsymbol{\theta}}_{\text{WLS}}$ as defined in Eq. (5) is inconsistent (e.g., Powell and Macdonald, 1972). Using simulations, the inconsistency of $\hat{\boldsymbol{\theta}}_{\text{WLS}}$ has been empirically confirmed, and a smaller bias has been observed for the refined estimator. The results are shown in Section 6.1.

5.3. Assessment of the residual cost function distribution

In the situation where input noise is not negligible, the residual V_{rWLS} value is expected to be a sample from a χ^2 distribution with $N-n_\theta$ degrees of freedom (cf. Eq. (8) and Section 3), while the residual V_{WLS} is expected not to be. These expectations are verified by Monte-Carlo simulations. One thousand different noise realizations were generated for the same data set. These 1000 artificial data sets were optimized minimizing (i) the original WLS cost function and (ii) the refined cost function. Next, the distributions of the resulting 1000 residual V_{WLS} and V_{rWLS} values were compared to the theoretically expected χ^2 distribution. The results are given in Section 6.2.

5.4. Assessment of the refined parameter uncertainty estimation

Two different tests were performed to assess the uncertainty estimation. First, a collection of 42 data sets performed by and described in Beucher et al., 2004 were considered. The dissolved Si concentration ranges from 0.85 to 10.5 μM, whereas the particulate

Si (biosilica) varies between 0.24 and 4.0 μM. The experimental relative standard deviations range from 2% to 20%. For each data set, the parameter standard deviations were calculated via the LCM, (i) using the output **W** and (ii) using the refined $\mathbf{W}_r$. Both uncertainty estimations are expected to give different results, namely, the first will give lower standard deviations than the latter. But so far, it is not yet possible to evaluate the accuracy of each of the standard deviation estimates. For that, a Monte-Carlo simulation with again 1000 different noise realizations was performed, for each data set. The parameter variation derived from the simulation is taken as the true standard deviation (with an uncertainty of ±4%). In Section 6.3, the comparison between (i) the true standard deviations, (ii) the estimated ones using the LCM expression with **W** and (iii) the estimated ones using the LCM expression with $\mathbf{W}_r$ is shown and discussed.

Secondly, the performance of both estimations was assessed for one experimental condition, but with increasing input to output noise ratio. It is expected that the LCM estimation using the output **W** will increasingly underestimate the parameter variation, as the noise ratio augments. But how well will the estimation using $\mathbf{W}_r$ cope with increasing input noise? This was checked, again using Monte-Carlo simulations to derive the true parameter standard deviations. The results are shown and discussed in Section 6.3.

6. Results and discussion

6.1. Consistency of parameter estimation

Fig. 3 shows the estimation results using the original and refined WLS cost functions for an increasing

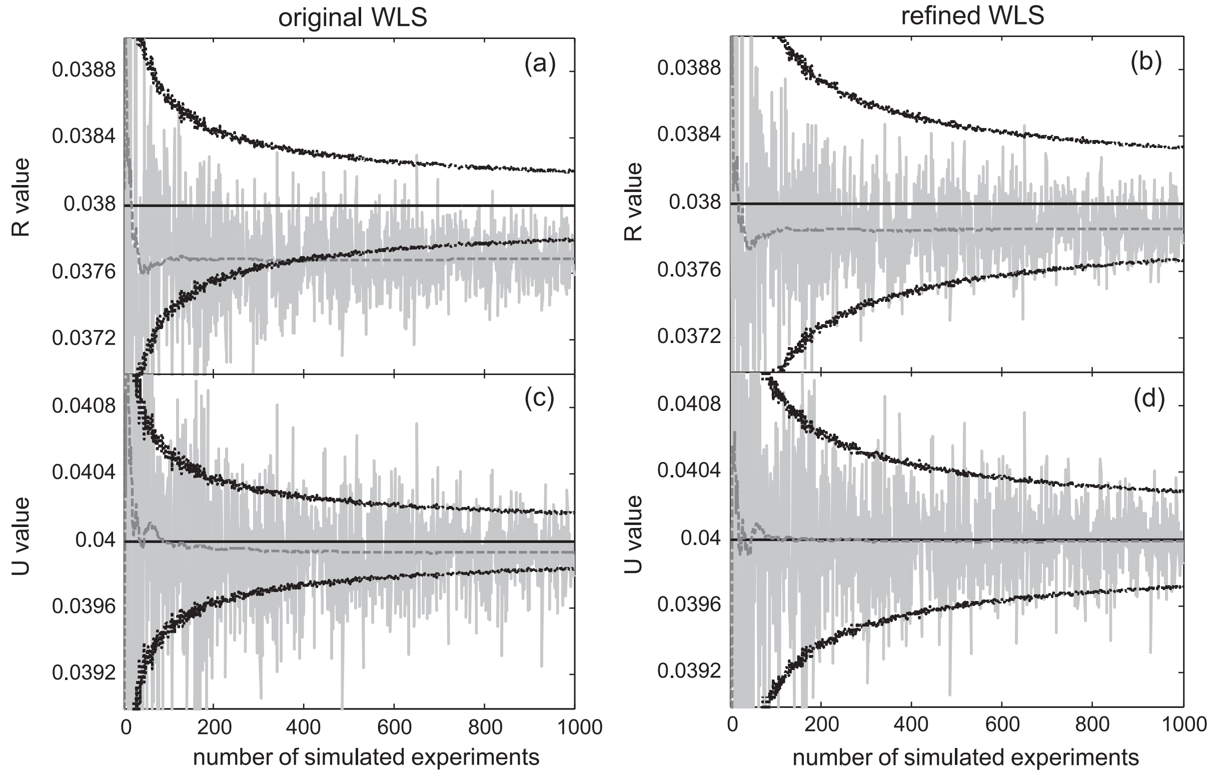

Fig. 3. Assessment of estimation consistency. The estimated parameter values (in μM/h) are shown as a function of the number of simulated experiments used for the estimation. Meaning of the symbols: the light grey solid line represents the estimated parameter value, the dashed grey line shows the mean of the estimated parameter values, the bold black line shows the true parameter value and the dotted lines indicate the estimated 95% confidence interval. (a and b) Consistency of original WLS estimation. The estimation shows a small bias: the estimation converges to a value lower than the true value. (c and d) Consistency of refined WLS estimation. The estimation shows a smaller bias.

number of data. It is expected that the original WLS procedure gives inconsistent estimations (cf. Section 5.2), and that the refinement decreases this inconsistence. The results of the simulation test illustrate these expectations. Indeed, the following observations can be made:

(i) The original WLS estimator gives inconsistent results for both parameters R and U: indeed, in Fig. 3a and c, the mean of estimated parameter values (grey dashed line) does not converge to the true value (solid black line), but to a lower value.
(ii) When the refined estimator is used, this inconsistency is decreased (Fig. 3b and d). This illustrates that the proposed refinement represents an improvement of the classical WLS procedure.
(iii) The refined R estimation still exhibits a certain bias (Fig. 3b). This is due to the fact that the proposed refinement takes the input variations into account via a linearization, which is only a first order approximation when the model is nonlinear in the input variables.
(iv) The observed inconsistency is (for both estimators) larger for parameter R than for U.

Summarizing, the simulation test illustrates that the original WLS estimator is inconsistent and that the refinement decreases the inconsistency.

Finally, it should be noted that this estimator property is mostly of theoretical importance, since in practice this asymptotical limit is never reached (see Eq. (17)). Nevertheless, we have included this section to show that the proposed refinement also influences the parameter estimation itself.

6.2. Distribution of the residual cost function values

Using a WLS estimator has the benefit of delivering an interpretable residual cost function in addition to the estimated optimal parameters. However, for this special feature to be true, it is necessary that the weighting matrix reflects the total variation of the error vector $\mathbf{e}$ (cf. Eq. (6)). Therefore, when input noise cannot be neglected, it is expected that the residual values of the original WLS cost function are not χ^2 distributed. Fig. 4a displays the results of the simulations performed to test this: the residual cost function values, plotted as a histogram, do not match the theoretical χ^2 function. Rather, there are too few low values and too many high values, which is the expected pattern for an underestimation of the total variation. Thus, even if, perhaps, the original WLS procedure produces quite reliable estimated parameter values (cf. Section 5.1), it is not possible to interpret its residual cost function in order to get information about the "acceptability" of the whole.

On the other hand, the residual values of the refined WLS cost function seem to fit very well the theoretical χ^2 curve (Fig. 4b). Consequently, using the refined WLS estimation offers the additional benefit of an interpretable residual cost function value. Moreover, these results confirm that the refined weighting matrix $\mathbf{W}_\mathrm{r}$ is an acceptable approximation of the total variation, which is clearly better than the original $\mathbf{W}$.

6.3. Recovery of uncertainty estimation

The LCM uncertainty estimations are evaluated by two ways. First, they are applied to real field data with fixed input and output noise variances. Secondly, synthetic data are used for which the input noise to output noise ratio is varied. The results of the first test are shown in Fig. 5. Forty-two data sets from Beucher et al. (2004) were analyzed as explained in Section 5.4. Because the standard deviations are of different magnitude, the results are shown as recoveries, i.e., the estimated standard deviation divided by the true one (derived from Monte-Carlo simulations).

Clearly and as expected, the uncertainty estimation using the original weighting matrix, which only considers the output variances, systematically underestimates the true parameter variation (the recoveries are systematically below 1). The average recoveries are 65% for parameter R and 62% for parameter U. On the contrary, the refined estimation is mostly a good approximation of the true variation. Sometimes, it under- or overestimates the true standard deviation, but there does not seem to be a systematic error. This is illustrated by the much better average recoveries: 94% and 92% for R and U, respectively. This confirms that the refined uncertainty estimation is, on the whole, a good approximation of the true parameter variation.

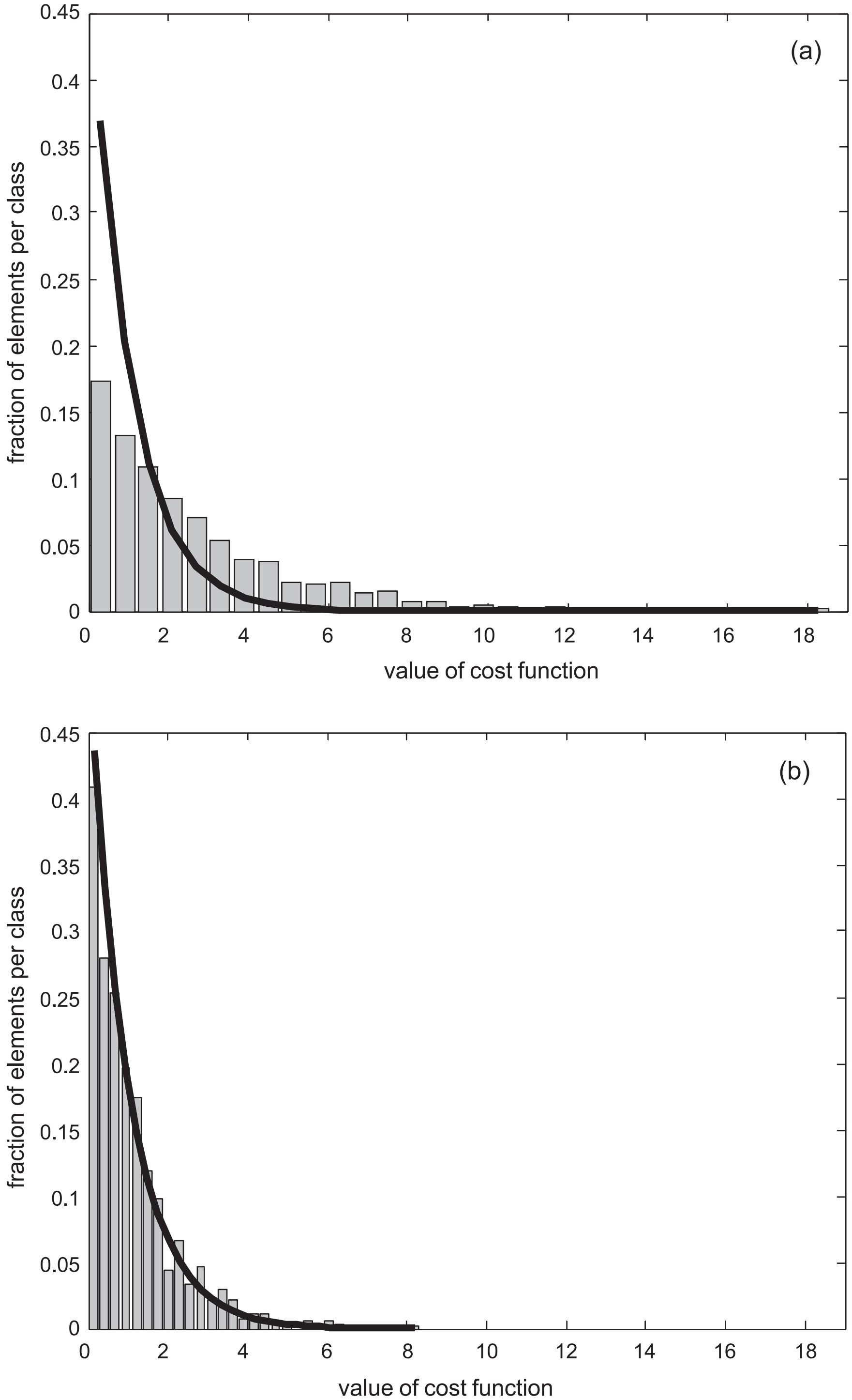

Fig. 4. Residual cost function values shown in a normalized histogram. The normalized histogram of simulated residual cost function values ( ) is compared to the expected probability density function with two degrees of freedom ( ). (a) Original WLS cost functions: the histogram does not correspond to the expected distribution. (b) Refined WLS cost functions: the histogram matches the expected function.

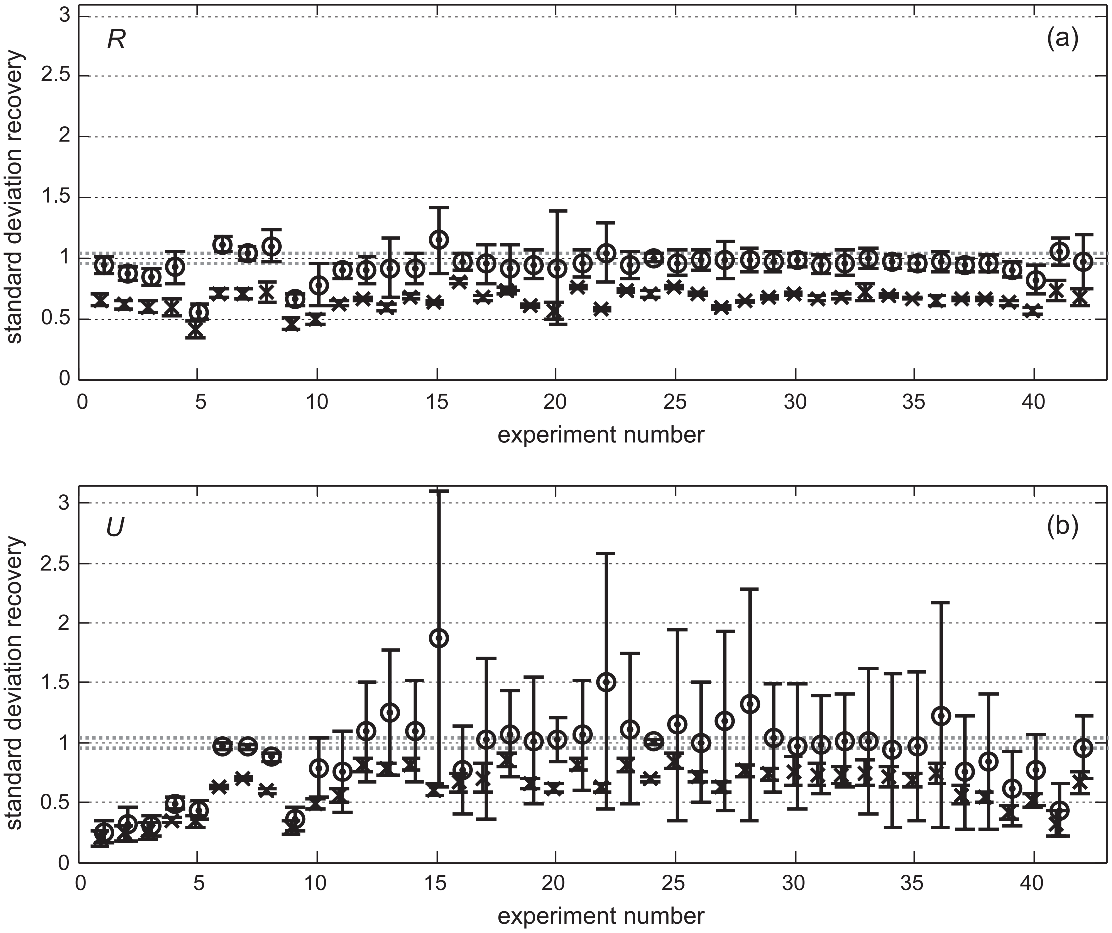

Fig. 5. Assessment of the uncertainty estimation applied to real data. The recovery achieved by the estimation using the original **W** (×) and using the refined $\mathbf{W}_r$ (○) are compared the ideal recovery of one. By recovery, it is meant: the estimated standard deviation divided by the true standard deviation. The grey dotted lines represent the approximate 95% confidence interval of the Monte-Carlo recovery. The error bars represent the approximate 95% confidence intervals of the estimated standard deviation recoveries. The recovery results are shown for the standard deviation estimation of (a) parameter R and (b) parameter U.

Nevertheless, it should be kept in mind that the LCM remains a linear approximation. Consequently, the recovery of the estimated standard deviations depends on the experimental conditions, but in an unpredictable way. Actually, the LCM expression with $\mathbf{W}_r$ (Eq. (14)) contains two linearizations. The first is with respect to the parameters, which is why it delivers only approximated parameter uncertainties when the model is nonlinear and the number of measurements is small. The other linearization, with respect to the input variables, comes from $\mathbf{W}_r$. If the model is linear in the input variables, this linearization is exact. Otherwise, as is the case with the Si cycling system under study, it is an approximation. The approximation could be improved by taking into account higher-order terms. But this would demand more computational time and could also increase the risk of instabilities. Another drawback is that higher-order moments should be accurately known, which is not obvious in general.

The results of the second test, in order to assess the performance of the uncertainty estimations for increasing input noise to output noise ratio, are displayed in Fig. 6. The output noise amplitude was always 5% of the output; this means that each noise realization was a sample from a $N(0,\ 0.05\cdot\mathbf{y})$

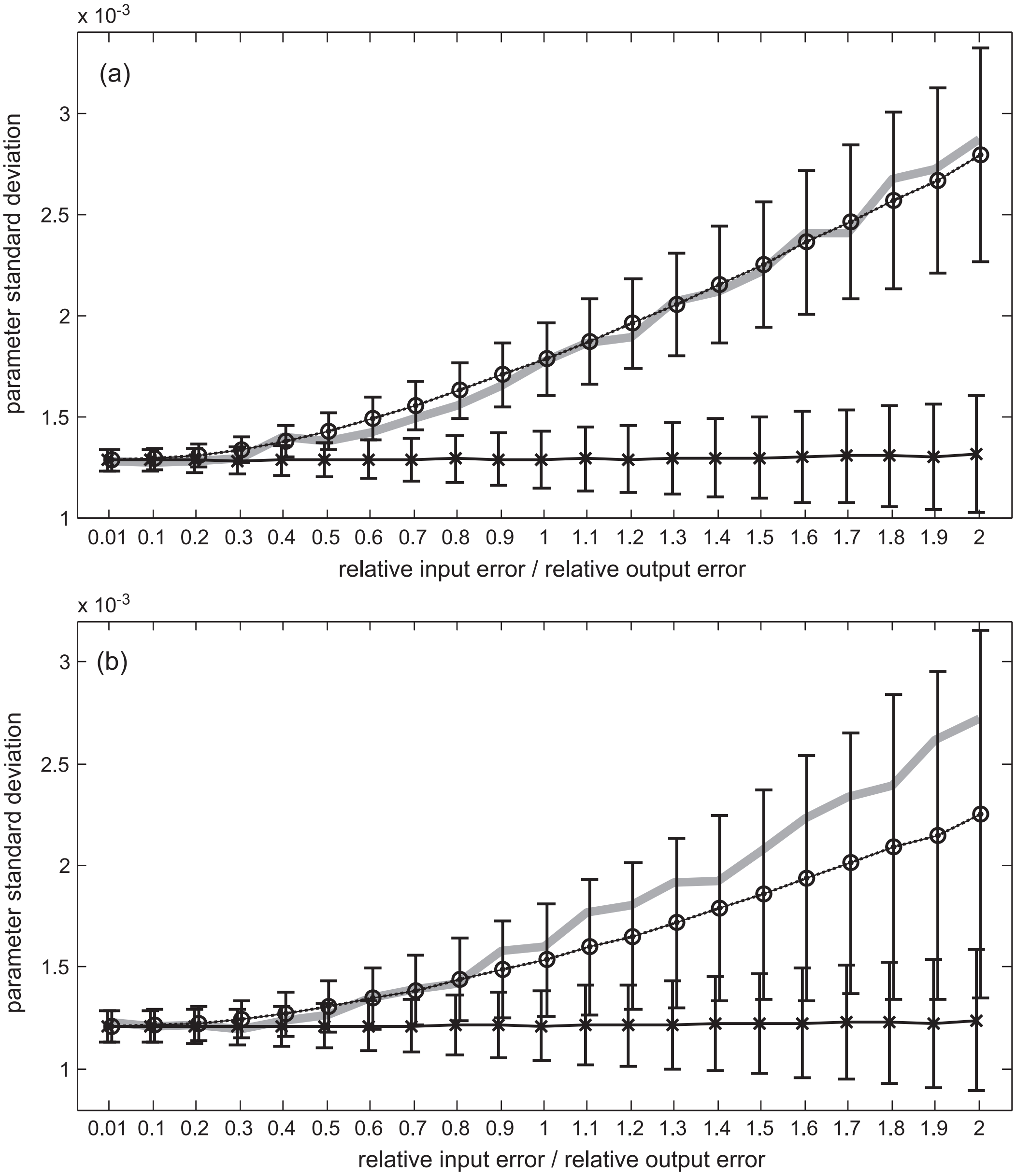

Fig. 6. Assessment of the uncertainty estimation for increasing input noise. The meaning of the symbols: The estimated standard deviation for U and R estimated using the original $\mathbf{W}$ (×) and using the refined $\mathbf{W}_r$ (○) are compared to the standard deviation derived from 1000 Monte-Carlo simulations (▬). The error bars represent the approximate 95% confidence intervals of the standard deviation estimates. The relative output noise amplitude is always 5% of the output. Thus, the relative input noise amplitude varies from 0.05% to 10% of the input. (a) The standard deviation of parameter R. (b) The standard deviation of parameter U.

distribution. While the output noise amplitude was fixed, the relative input noise amplitude was varied from 0.05% to 10% of the input. Fig. 6 nicely illustrates that for increasing relative input noise, the true parameter standard deviations increase too. This is understandable, because when the measurements used to estimate the parameters are increasingly variable, the parameter estimation itself will be subject to an enhanced uncertainty as well. However, this trend is not at all present in the LCM estimations with only output weighing. Rather, the estimated standard deviations remain approximately constant and thus are increasingly inaccurate. Since this estimation ignores an increasingly important origin of variation, namely, the input noise, and only considers the output noise which remains constant, this result is expected.

Finally and most importantly, the refined standard deviation estimations do show an increasing trend, which follows relatively well the true standard deviation. Apparently, the refined method produces satisfactory estimations of the parameter standard deviations, even when the input noise level is of the same or higher magnitude as the output noise. When the input noise becomes small with respect to the output noise, the figures show that the refined estimation converges to the original estimation. This is expected since the refined weighting matrix converges to the original one as the input noise variance tends to zero.

7. Conclusions

When a nonlinear model—in our case for the uptake and regeneration of Si in pelagic waters—is used for describing measurements which are all disturbed by noise, it is not so obvious to estimate the model parameters or, even less obvious, the parameter uncertainties. It is for these problems that we have formulated a refinement of the well-known Weighted Least Squares estimator and the Linearized Covariance Matrix expression, which originally neglect the effect of input noise. The refinement is based on a generalization of the weighting matrix by taking into account the input variations as well.

First, a comparison between the original and refined WLS estimator on a hypothetical data set revealed that the refined estimator exhibits a smaller bias.

Second, the analysis of the residual cost function values of both estimators showed that only the refined cost function is nicely χ^2 distributed, which confirms that the refined weighting matrix $\mathbf{W}_r$ is a good reflection of the total variation, and certainly better than the original **W**. Using this statistical feature of the refined WLS cost function, it is possible to identify systematic errors in individual measurement-model results. This property is important because it allows an internal quality control of the results and it can be used to select the most appropriate model from a set of models for a given data set (de Brauwere et al., in press).

Furthermore, the estimation of the parameter uncertainty using the Linearized Covariance Matrix expression was tested on field data of dissolved and particulate Si in the Bay of Brest (France) from Beucher et al. (2004). The estimated standard deviations were compared to values derived from Monte-Carlo simulations. Overall, the recovery was good for the refined method (average recoveries of 94% and 92% for R and U, respectively), whereas with the original **W**, the estimations systematically and significantly underestimated the parameter variations (average recoveries of 65% and 62%).

Altogether, the refinement of the weighting matrix corresponds to a better estimation of the total variation and this especially improves the estimation of the parameter standard deviations using the Linearized Covariance Matrix expression. Finally, it is worth emphasizing that although the methodology was developed for and exemplified by the silicic acid processes, it is applicable to all parameter estimation problems based on measurements that are subject to stochastic errors.

Glossary

Symbol	Description
$\hat{\boldsymbol{\theta}}_{\mathbf{(r)WLS}}$	Optimal values of parameters θ corresponding to the minimal (r)WLS cost function, as defined in Eqs. (6) and (12) ($n_\theta \times 1$ vector)
$\hat{\boldsymbol{\theta}}$	Optimal values for the parameters $\boldsymbol{\theta}$ ($n_\theta \times 1$ vector)
$\boldsymbol{\theta}$	Parameter vector ($n_\theta \times 1$)
$\boldsymbol{\theta}_0$	True parameter values ($n_\theta \times 1$ vector)
α_n	Natural abundance of Si=3.4%
α_D, α_P	Isotopic abundance of Si of silicic acid (D), or biosilica (P)=$[^{30}Si]_X/([^{28}Si]_X+[^{29}Si]_X+[^{30}Si]_X)\cdot 100\%$
$\boldsymbol{\sigma}_y^2$, $\hat{\boldsymbol{\sigma}}_y^2$	True or estimated values of output noise variances ($N\times 1$ vector composed by $\sigma_{y,i}^2$ or $\hat{\sigma}_{y,i}^2$)
$\boldsymbol{\sigma}_x^2$, $\hat{\boldsymbol{\sigma}}_x^2$	True or estimated values of input noise variances ($N\times 1$ vector composed by $\sigma_{x,i}^2$ or $\hat{\sigma}_{x,i}^2$)
æ$_D$, æ$_P$	Atom% excess of Si in silicic acid pool (D), or biosilica (P)=$\alpha-\alpha_n$

$\mathrm{Cov}(\theta)$ Covariance matrix associated with parameters θ ($n_\theta \times n_\theta$ matrix)
D Total concentration of (dissolved) silicic acid=$[Si(OH)_4]$
E Error vector ($N \times 1$ vector)
f_{model} Functional expression of model
$\mathbf{J}(\theta)$ Jacobian matrix, calculated for parameter values θ ($N \times n_\theta$ matrix)
LCM Linearized Covariance Matrix, see Eqs. (9) and (14)
N Number of output measurements
n_θ Number of parameters to estimate
$\mathbf{n_x}$ Input noise terms, as defined in Eq. (2a) ($N \times 1$ vector)
$\mathbf{n_y}$ Output noise terms, as defined in Eq. (2b) ($N \times 1$ vector)
P Total concentration of biosilica=$[BSiO_2]$
R Regeneration rate of Si, dissolution rate of biosilica
rWLS Refined Weighted Least Squares
U Uptake rate of Si, production rate of biosilica
$V_{(\mathrm{r})\mathrm{WLS}}$ (refined) WLS cost function, defined in Eqs. (7) and (13)
W Original weighting matrix, assuming only output noise ($N \times N$ matrix)
WLS Weighted Least Squares
$\mathbf{W}_\mathrm{r}$ Refined weighting matrix, including the effect of input noise ($N \times N$ matrix)
x Input variable or measured value of input variable ($N \times 1$ vector)
$\mathbf{x}_0$ True values of input variables ($N \times 1$ vector)
y Output variable or measured value of output variable ($N \times 1$ vector)
$\mathbf{y}_0$ True values of output variables ($N \times 1$ vector)

Acknowledgements

Anouk de Brauwere is a research assistant of the Flemish Fund for Scientific Research (FWO-Vlaanderen) and is grateful for their support. This work was also supported by the Belgian government (1UAP V/22) and the Flemish government in cooperation with the Vrije Universiteit Brussel (GOA221 DSWER4 and GOA-ILiNoS). Finally we wish to thank Charlotte Beucher for kindly letting us use her data.

References

Bard, Y., 1974. Nonlinear Parameter Estimation. Academic Press, New York.

Bates, D.M., Watts, D.G., 1980. Relative curvature measures of nonlinearity. J. R. Stat. Soc., Ser. B Methodol. 42 (1), 1–25.

Bates, D.M., Watts, D.G., 1981. Parameter transformations for improved approximate confidence-regions in non-linear least-squares. Ann. Stat. 9 (6), 1152–1167.

Beucher, C., Tréguer, P., Corvaisier, R., Hapette, A., Elskens, M., 2004. Production and dissolution of biosilica, and changing microphytoplankton dominance in the Bay of Brest (France). Mar. Ecol., Prog. Ser., 26757–26769.

Box, M.J., 1970. Improved parameter estimation. Technometrics 12 (2), 219–229.

Brzezinski, M.A., Jones, J.L., Bidle, K.D., Azam, F., 2003. The balance between silica production and silica dissolution in the sea: Insights from Monterey Bay, California, applied to the global data set. Limnol. Oceanogr. 48 (5), 1846–1854.

Buesseler, K.O., 1998. The decoupling of production and particulate export in the surface ocean. Glob. Biogeochem. Cycles 12 (2), 297–310.

Cecchi, G.C., 1991. Error analysis of the parameters of a least-squares determined curve when both variables have uncertainties. Meas. Sci. Technol. 2 (12), 1127–1128.

Chan, I.S., Goldstein, A.A., Bassingthwaighte, J.B., 1993. SENSOP—a derivative-free solver for nonlinear least-squares with sensitivity scaling. Ann. Biomed. Eng. 21 (6), 621–631.

Chandler, J.P., 1972. Iterative procedure for estimating functions when both variables are subject to error. Technometrics 14 (1), 71–76.

Clutton-Brock, M., 1967. Likelihood distributions for estimating functions when both variables are subject to error. Technometrics 9 (2), 261–269.

de Brauwere, A., De Ridder, F., Pintelon, R., Elskens, M., Schoukens, J., Baeyens, W. Model selection through a statistical analysis of the minimum of Weighted Least Squares cost function. Chemometr. Intell. Lab. Syst., in press.

Donaldson, J.R., Schnabel, R.B., 1987. Computational experience with confidence-regions and confidence-intervals for nonlinear least-squares. Technometrics 29 (1), 67–82.

Dugdale, R.C., Wilkerson, F.P., 1998. Silicate regulation of new production in the equatorial pacific upwelling. Nature 391 (6664), 270–273.

Grimstad, A.A., Kolltveit, K., Mannseth, T., Nordtvedt, J.E., 2001. Assessing the validity of a linearized accuracy measure for a nonlinear parameter estimation problem. Inverse Probl. 17 (5), 1373–1390.

Jones, T.A., 1979. Fitting straight lines when both variables are subject to error: 1. Maximum likelihood and least-squares estimation. Math. Geol. 11 (1), 1–25.

Mandel, J., 1984. Fitting straight-lines when both variables are subject to error. J. Qual. Technol. 16 (1), 1–14.

Marsili-Libelli, S., 1992. Parameter estimation of ecological models. Ecol. Model. 62 (4), 233–258.

Marsili-Libelli, S., Guerrizio, S., Checchi, N., 2003. Confidence regions of estimated parameters for ecological systems. Ecol. Model. 165 (2–3), 127–146.

Massart, D.L., Vandeginste, B.G., Buydens, L.M., De Jong, S., Lewi, P., Smeyers-Verbeke, 1997. Handbook of Chemometrics and Qualimetrics. Elsevier, Amsterdam.

Meinrath, G., Ekberg, C., Landgren, A., Liljenzin, J.O., 2000. Assessment of uncertainty in parameter evaluation and prediction. Talanta 51 (2), 231–246.

Moreno, C., Bruzzone, H., 1993. Parameters variances of a least-squares determined straight-line with errors in both coordinates. Meas. Sci. Technol. 4 (5), 635–636.

Nelson, D.M., Treguer, P., Brzezinski, M.A., Leynaert, A., Queguiner, B., 1995. Production and dissolution of biogenic silica in the ocean—revised global estimates, comparison with regional data and relationship to biogenic sedimentation. Glob. Biogeochem. Cycles 9 (3), 359–372.

Nihoul, J.C.J., 1975. Modelling of marine systems. Elsevier Oceanography Series, vol. 10. Elsevier, Amsterdam.

Pintelon, R., Schoukens, J., 2001. System Identification: A Frequency Domain Approach. IEEE, New York.

Powell, D.R., Macdonald, J.R., 1972. Rapidly convergent iterative method for solution of generalized nonlinear least-squares problem. Comput. J. 15 (2), 148–155.

Riggs, D.S., Guarnieri, J.A., Addelman, S., 1978. Fitting straight lines when both variables are subject to error. Life Sci. 22 (13-1), 1305–1360.

Rod, V., Hancil, V., 1980. Iterative estimation of model parameters when measurements of all variables are subject to error. Comput. Chem. Eng. 4 (1), 33–38.

Tang, S.Z., Wang, Y.H., 2002. A parameter estimation program for the error-in-variable model. Ecol. Model. 156 (2–3), 225–236.

Valsami, G., Iliadis, A., Macheras, P., 2000. Non-linear regression analysis with errors in both variables: estimation of co-operative binding parameters. Biopharm. Drug Dispos. 21 (1), 7–14.

Varah, J.M., 1990. Relative sizes of the Hessian terms in nonlinear parameter-estimation. SIAM J. Sci. Statist. Comput. 11 (1), 174–179.