

Using Statistical Bias Description for Multi-objective Calibration of a Lake Water Quality Model

Anne Dietzel and Peter Reichert

*Eawag, Swiss Federal Institute of Aquatic Science and Technology,
8600 Dübendorf, Switzerland, and
ETH Zurich, Institute of Biogeochemistry and Pollutant Dynamics,
8092 Zurich, Switzerland
Anne.Dietzel@eawag.ch, Peter.Reichert@eawag.ch*

Abstract: Models of environmental systems are simplified representations of reality. For this reason, their results are affected by systematic errors. This bias makes it difficult to get reliable uncertainty estimates of model parameters and predictions. We were faced with that difficulty when applying the lake model BELAMO to data from three Swiss lakes. This model combines the description of biogeochemical and ecological processes in lakes. Considering bias in the model output, lead to a description of system observations as the sum of the output of the deterministic model, bias and observation error. The identifiability problem between model output and bias was addressed by specifying informative priors for the standard deviations of the observation errors and choosing means of zero for the observation error and the bias. The resulting multi-objective calibration problem was solved by using the prior of the bias to specify how much model error we are willing to accept for which output variable. To avoid the very high computational demand of conventional Bayesian numerical techniques, the maximum of the posterior was calculated and a local Gaussian approximation was used to estimate parameter and model prediction uncertainty. Parameter estimations for 9 to 20 years until 1995 were conducted. The remaining 10 years of data were used for model validation and to compare with estimated prediction uncertainty. The results show the large influence of the bias for the model output. The results for the validation period indicate the large uncertainty in model prediction, but also the ability to estimate the role of model bias with the suggested technique.

Keywords: Lake water quality modelling; Multi-objective calibration; Bias; Uncertainty

1. INTRODUCTION

Models substantially contribute to formalizing and summarizing knowledge, analyzing observations and testing hypotheses about environmental systems. As all models are simplified representations of reality, results of environmental models are, in addition to input and parametric uncertainty, affected by systematic errors.

When trying to identify model parameters from data by using statistical inference, model bias leads to a violation of typical statistical assumptions and thus makes it difficult to get reliable uncertainty estimates. One option to derive uncertainty bounds of model predictions is to use a statistical description of bias in model output (Craig et al. [1996], Craig et al. [2001], Kennedy and O'Hagan [2001], Higdon et al. [2004], Bayarri et al. [2007]).

Another problem resulting from systematic deviations between model results and data is that frequently used model parameter estimation techniques often cause results that do not fulfil calibration objectives of the user. To address this problem, manual calibration criteria were included in multi-criteria optimization techniques (Yapo et al. [1996], Gupta et al. [1996], Madsen [2000], Madsen et al. [2002], Gupta et al. [2003], Boyle et al. [2003]). This procedure can lead to better model calibration, but does not provide probabilistic information required for estimating prediction uncertainty.

To assess the parameter, structural and prediction uncertainty under the problems mentioned above, a combination of statistical description of model bias and the ideas underlying multi-objective model calibration was used as described by Reichert and Schuwirth [2010]: The prior of a statistical description of bias was used to weigh between different calibration objectives. This multi-objective calibration technique was applied to the lake model BELAMO. Thereby, a calibration period of 20 years (9 years for Greifensee) was chosen (ending in 1995). Validation of the model for the following 10 years was done and uncertainty bounds of our knowledge about the true values of the model output variables were derived for the calibration as well as the validation period.

2. METHODS

We briefly summarize the technique suggested by Reichert and Schuwirth [2010], the used simplified approach to assess parameter and prediction uncertainty bounds and the numerical implementation of the proposed technique.

2.1 Review of inference in the presence of bias

When applying a model to a specific study area, one has to cope with the problem of the "right" parameter choice. As for most parameters direct measurements are missing, the user is faced with the difficulty of the choice of parameter values that are in a realistic order of magnitude and allow for model results that are in agreement with measurements for all state variables for the whole simulation period. This can be done by combining prior knowledge of parameter values with information gained from the observations of model output by Bayesian inference. However, in order not to violate the statistical assumptions of the model, it has to explicitly account for bias in model output. Following the literature on statistical bias description (Craig et al. [1996], Craig et al. [2001], Kennedy and O'Hagan [2001], Higdon et al. [2004], Bayarri et al. [2007]), the observations are described as the sum of the output of the deterministic model, $\mathbf{y}_M^L(\mathbf{x}, \boldsymbol{\theta})$, bias, $\mathbf{B}_M^L(\mathbf{x}, \boldsymbol{\xi})$, and observation error, $\mathbf{E}^L(\boldsymbol{\psi})$:

$$\mathbf{Y}_M^L(\mathbf{x}, \boldsymbol{\theta}, \boldsymbol{\psi}, \boldsymbol{\xi}) = \mathbf{y}_M^L(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{B}_M^L(\mathbf{x}, \boldsymbol{\xi}) + \mathbf{E}^L(\boldsymbol{\psi}), \quad (1)$$

where \mathbf{Y}_M^L is the vector of random variables representing the observations at the observation layout L . The observation layout, L , defines which variables to observe or evaluate at which points in time and space. \mathbf{Y}_M^L depends on external influence factors, \mathbf{x} , unknown model parameters, $\boldsymbol{\theta}$, and additional parameters, $\boldsymbol{\psi}$ and $\boldsymbol{\xi}$, of the error terms.

Equation (1) leads to a hierarchical model with the bias, \mathbf{B}_M^L , as an intermediate variable. Integrating out this intermediate variable, a likelihood function can be obtained as a function of the parameters, $\boldsymbol{\Gamma} = (\boldsymbol{\Theta}, \boldsymbol{\Psi}, \boldsymbol{\Xi})$, only:

$$f_{\mathbf{Y}_M^L | \boldsymbol{\Theta}, \boldsymbol{\Psi}, \boldsymbol{\Xi}}(\mathbf{y}^L | \boldsymbol{\theta}, \boldsymbol{\psi}, \boldsymbol{\xi}, \mathbf{x}) = \int f_{\mathbf{E}^L | \boldsymbol{\Psi}}(\mathbf{y}^L - \mathbf{y}_M^L(\mathbf{x}, \boldsymbol{\theta}) - \mathbf{b}^L | \boldsymbol{\psi}) \cdot f_{\mathbf{B}_M^L | \boldsymbol{\Xi}}(\mathbf{b}^L | \boldsymbol{\xi}, \mathbf{x}) d\mathbf{b}^L. \quad (2)$$

In this equation, $f_{\mathbf{E}^L | \boldsymbol{\Psi}}$ is the probability density of the observation errors and $f_{\mathbf{B}_M^L | \boldsymbol{\Xi}}$ is the prior density of the bias given the parameters $\boldsymbol{\xi}$ and external inputs \mathbf{x} .

We assume a normally distributed observation error with mean zero and covariance matrix $\Sigma_{\mathbf{E}^L}(\boldsymbol{\psi})$

$$f_{\mathbf{E}^L|\boldsymbol{\psi}}(\boldsymbol{\varepsilon}^L|\boldsymbol{\psi}) = \frac{1}{\sqrt{2\pi}^{n_L}} \frac{1}{\sqrt{\det(\Sigma_{\mathbf{E}^L}(\boldsymbol{\psi}))}} \exp\left(-\frac{1}{2}(\boldsymbol{\varepsilon}^L)^\top \Sigma_{\mathbf{E}^L}(\boldsymbol{\psi})^{-1} \boldsymbol{\varepsilon}^L\right) \quad (3)$$

and a Gaussian stochastic process with a correlation structure that decays continuously with the distance of the corresponding independent input variables, \mathbf{x} , to describe our prior knowledge of the bias

$$f_{\mathbf{B}_M^L|\boldsymbol{\xi}}(\mathbf{b}^L|\boldsymbol{\xi}, \mathbf{x}) = \frac{1}{\sqrt{2\pi}^{n_L}} \frac{1}{\sqrt{\det(\Sigma_{\mathbf{B}_M^L}(\boldsymbol{\xi}, \mathbf{x}))}} \exp\left(-\frac{1}{2}(\mathbf{b}^L)^\top \Sigma_{\mathbf{B}_M^L}(\boldsymbol{\xi}, \mathbf{x})^{-1} \mathbf{b}^L\right). \quad (4)$$

A simple form of the covariance matrix with matrix elements is used

$$\Sigma_{\mathbf{B}_M^L, i, j}(\boldsymbol{\xi}, \mathbf{x}) = \begin{cases} 0 & \text{if } y_i^L \text{ and } y_j^L \text{ are of different type} \\ \sigma_{\mathbf{B}_M^L, i}(\boldsymbol{\xi}) \sigma_{\mathbf{B}_M^L, j}(\boldsymbol{\xi}) \exp\left(-\sum_k \beta_k(\boldsymbol{\xi})(x_{i,k} - x_{j,k})^2\right) & \text{else} \end{cases} \quad (5)$$

With these assumptions the integration in equation (2) can be done analytically. This leads to the likelihood function

$$\begin{aligned} & f_{\mathbf{Y}_M^L|\boldsymbol{\theta}, \boldsymbol{\psi}, \boldsymbol{\xi}, \mathbf{x}}(\mathbf{y}^L|\boldsymbol{\theta}, \boldsymbol{\psi}, \boldsymbol{\xi}, \mathbf{x}) \\ &= \frac{1}{\sqrt{2\pi}^{n_L}} \frac{1}{\sqrt{\det(\Sigma_{\mathbf{E}^L})}} \frac{1}{\sqrt{\det(\Sigma_{\mathbf{B}_M^L})}} \frac{1}{\sqrt{\det(\Sigma_{\mathbf{E}^L}^{-1} + \Sigma_{\mathbf{B}_M^L}^{-1})}} \\ & \cdot \exp\left(-\frac{1}{2}[\mathbf{y}^L - \mathbf{y}_M^L(\mathbf{x}, \boldsymbol{\theta})]^\top (\Sigma_{\mathbf{E}^L}^{-1} - \Sigma_{\mathbf{E}^L}^{-1}(\Sigma_{\mathbf{E}^L}^{-1} + \Sigma_{\mathbf{B}_M^L}^{-1})^{-1} \Sigma_{\mathbf{E}^L}^{-1})[\mathbf{y}^L - \mathbf{y}_M^L(\mathbf{x}, \boldsymbol{\theta})]\right) \end{aligned} \quad (6)$$

(see Reichert and Schuwirth [2010] for more details). The identifiability problem between model and bias is solved by using means of zero for the bias and observation errors, an informative prior for the parameter(s) of the error model, $\boldsymbol{\xi}$, and by using the parameters of the bias, $\boldsymbol{\psi}$, for specifying how much bias one is willing to accept for each model variable.

2.2 Numerical implementation

Inference

An observation layout L_1 is used for inference. It represents points in time and space where observations are available for all considered variables. The mode of the posterior of the model parameters given the observations can be calculated by

$$\boldsymbol{\gamma}^0(\mathbf{y}^{L_1}) = \operatorname{argmax}_{\boldsymbol{\gamma}}(\log(f_{\mathbf{Y}_M^L|\Gamma}(\mathbf{y}^{L_1}, \boldsymbol{\gamma}, \mathbf{x}) f_{\text{pri}}(\boldsymbol{\gamma}))) \quad (7)$$

We then approximate the posterior by a normal distribution with the same mode and curvature at the mode, $\mathbf{N}(\boldsymbol{\gamma}^0, \boldsymbol{\Sigma})$. Here, the variance-covariance matrix $\boldsymbol{\Sigma}$ is given as

$$\Sigma_{\Gamma} \approx -\left(\frac{\partial^2 \log(f_{\text{post}}(\boldsymbol{\gamma}, \mathbf{y}^{L_1}, \mathbf{x}))}{\partial \boldsymbol{\gamma}^\top \partial \boldsymbol{\gamma}} \Bigg|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^0}\right)^{-1} = -\left(\frac{\partial^2 \log(f_{\mathbf{Y}_M^L|\Gamma}(\mathbf{y}^{L_1}, \boldsymbol{\gamma}, \mathbf{x}) f_{\text{pri}}(\boldsymbol{\gamma}))}{\partial \boldsymbol{\gamma}^\top \partial \boldsymbol{\gamma}} \Bigg|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^0}\right)^{-1} \quad (8)$$

where the second derivative of the log posterior can be approximated numerically as

$$\begin{aligned} \frac{\partial^2 \log(f_{\text{post}}(\boldsymbol{\gamma}, \mathbf{y}^{L_1}, \mathbf{x}))}{\partial \gamma_i \partial \gamma_j} \Big|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^0} &= \frac{\log(f_{\mathbf{Y}_M^{L_1} \Gamma}(\mathbf{y}^{L_1}, \boldsymbol{\gamma}^0 + \delta_i \mathbf{e}_i + \delta_j \mathbf{e}_j, \mathbf{x}) f_{\text{pri}}(\boldsymbol{\gamma}^0 + \delta_i \mathbf{e}_i + \delta_j \mathbf{e}_j))}{4 \delta_i \delta_j} \\ &\quad - \log(f_{\mathbf{Y}_M^{L_1} \Gamma}(\mathbf{y}^{L_1}, \boldsymbol{\gamma}^0 + \delta_i \mathbf{e}_i - \delta_j \mathbf{e}_j, \mathbf{x}) f_{\text{pri}}(\boldsymbol{\gamma}^0 + \delta_i \mathbf{e}_i - \delta_j \mathbf{e}_j)) \\ &\quad - \log(f_{\mathbf{Y}_M^{L_1} \Gamma}(\mathbf{y}^{L_1}, \boldsymbol{\gamma}^0 - \delta_i \mathbf{e}_i + \delta_j \mathbf{e}_j, \mathbf{x}) f_{\text{pri}}(\boldsymbol{\gamma}^0 - \delta_i \mathbf{e}_i + \delta_j \mathbf{e}_j)) \\ &\quad + \log(f_{\mathbf{Y}_M^{L_1} \Gamma}(\mathbf{y}^{L_1}, \boldsymbol{\gamma}^0 - \delta_i \mathbf{e}_i - \delta_j \mathbf{e}_j, \mathbf{x}) f_{\text{pri}}(\boldsymbol{\gamma}^0 - \delta_i \mathbf{e}_i - \delta_j \mathbf{e}_j)) \end{aligned} \quad (9)$$

(Gelman et al. [1995]). The approximate posterior distribution of the parameters is then given as $\Gamma | \mathbf{Y}_M^L : \mathbf{N}(\boldsymbol{\gamma}^0, \boldsymbol{\Sigma}_\Gamma)$.

Prediction

For the prediction, in contrast to Reichert and Schuwirth [2010], linearised error propagation is used to reduce computation time. The simplified prediction for layout L_2 is given by

$$\begin{aligned} \mathbf{y}_M^{L_2} + \mathbf{B}_M^{L_2} | \mathbf{Y}_M^{L_1} : \mathbf{N}(\mathbf{y}_M^{L_2}(\boldsymbol{\gamma}^0) + \mathbb{E}[\mathbf{B}_M^{L_2} | \mathbf{Y}_M^{L_1}, \Gamma](\mathbf{y}^{L_1}, \boldsymbol{\gamma}^0), \\ \mathbf{V}^{L_2} \boldsymbol{\Sigma}_\Gamma (\mathbf{V}^{L_2})^\top + \text{Var}[\mathbf{B}_M^{L_2} | \mathbf{Y}_M^{L_1}, \Gamma](\mathbf{y}^{L_1}, \boldsymbol{\gamma}^0)) \end{aligned} \quad (11)$$

with

$$\mathbf{V}^{L_2} = \frac{\partial(\mathbf{y}_M^{L_2} + \mathbb{E}[\mathbf{B}_M^{L_2} | \mathbf{Y}_M^{L_1}, \Gamma])}{\partial \boldsymbol{\gamma}^\top} \Big|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^0} \quad (12)$$

The conditional distribution of the bias becomes a multivariate normal distribution under our assumptions. $\mathbb{E}[\mathbf{B}_M^{L_2} | \mathbf{Y}_M^{L_1}, \Gamma]$ and $\text{Var}[\mathbf{B}_M^{L_2} | \mathbf{Y}_M^{L_1}, \Gamma]$ are the mean and covariance matrix of the bias. In our case, layout L_2 represents the validation period of 10 years.

Similarly, we can calculate the prediction for layout L_1 analogue to L_2 with mean

$$\mathbb{E}[\mathbf{B}_M^{L_1} | \mathbf{Y}_M^{L_1}, \Gamma] = \text{Var}[\mathbf{B}_M^{L_1} | \mathbf{Y}_M^{L_1}, \Gamma] \cdot \boldsymbol{\Sigma}_{\mathbf{E}^{L_1}}^{-1} \cdot (\mathbf{y}^{L_1} - \mathbf{y}_M^{L_1}(\mathbf{x}, \boldsymbol{\theta})) \quad (13)$$

and covariance matrix

$$\text{Var}[\mathbf{B}_M^{L_1} | \mathbf{Y}_M^{L_1}, \Gamma] = (\boldsymbol{\Sigma}_{\mathbf{E}^{L_1}}^{-1} + \boldsymbol{\Sigma}_{\mathbf{B}_M^{L_1}}^{-1})^{-1} \quad (14)$$

The marginals of the predictions (11) and (13) can be combined to marginal predictions of the complete layout $L_1 \cup L_2$.

3. DIDACTICAL EXAMPLE

As a simple example, we use a continuous-time, linear model with two output variables

$$\begin{aligned} g(t, a, b, c) &= a + b t \\ h(t, a, b, c) &= c(a + b t) \end{aligned} \quad (15)$$

as was used in Reichert and Schuwirth [2010]. The observation layout consists of observing both variables, g and h , at the time points $\{t_1, t_2, \dots, t_n\}$. This leads to the deterministic model function

$$\mathbf{y}_M^L(\boldsymbol{\theta}) = (g(t_1, \boldsymbol{\theta}), \dots, g(t_n, \boldsymbol{\theta}), h(t_1, \boldsymbol{\theta}), \dots, h(t_n, \boldsymbol{\theta}))^T \quad (16)$$

with the parameter vector $\boldsymbol{\theta} = (a, b, c)^T$. The error model consists of independent normal distributions with standard deviations σ_{E_g} and σ_{E_h} for the variables g and h at all points in time. This leads to the parameters $\boldsymbol{\psi} = (\sigma_{E_g}, \sigma_{E_h})^T$ of the error model. Finally, we assume a Gaussian stochastic process in time with standard deviations σ_{B_g} and σ_{B_h} for each of the model variables g and h , for describing the bias. Using the correlation time $t_{\text{corr}} = \beta^{-0.5}$ instead of β to parameterize equation (5), leads to the parameter vector $\boldsymbol{\xi} = (t_{\text{corr}}, \sigma_{B_g}, \sigma_{B_h})^T$ for the bias.

The model given by the equations (15) and (16) represents in a very simple way coupling of two output variables in a multivariate model (by the multiplicative factor c). The intercepts and slopes of the two model output variables are not independent. We will produce synthetic data for this model with two types of bias: The first type consists in choosing the intercepts and slopes of the two model output variables independently. Second, we add a "zig-zag" line of bias to the model outputs. As the model cannot fit both linear components equally well, the quality of fit of each component will depend on the choice of the prior of the bias. By specifying the prior of the bias, the user can choose how much bias is acceptable in each of the two variables. As this can hardly be done in absolute terms a priori, we choose an exponential prior for the standard deviations of bias in both output variables. We assume independent normal priors for a and b (with means of 1 and 0.2 and standard deviations of 0.5 and 0.2, respectively), lognormal priors for c , t_{corr} , σ_{E_g} and σ_{E_h} (means 2, 3, 0.2 and 0.2 and standard deviations 2, 0.3, 0.02 and 0.02, respectively), and exponential priors for σ_{B_g} and σ_{B_h} (with means of 0.2 and 0.5).

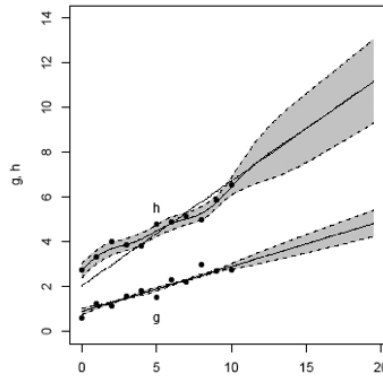


Figure 1. Data points (markers), median (solid) and 95% credibility interval (shaded area with dashed boundaries) of $\mathbf{y}_M^{L1,2} + \mathbf{B}_M^{L1,2} | \mathbf{Y}_M^{L1}$ and median of $\mathbf{y}_M^{L1,2}$.

Figure 1 shows the model predictions for $\mathbf{y}_M^{L1} + \mathbf{B}_M^{L1} | \mathbf{Y}_M^{L1}$ and $\mathbf{y}_M^{L2} + \mathbf{B}_M^{L2} | \mathbf{Y}_M^{L1}$ for one case of prior choice of bias. The results demonstrate that our posterior knowledge is much more

precise in time domains with data than in the extrapolation range. Further results of the residuals between model output and observations show that the suggested technique is able to divide the residuals probabilistically into bias and observation error. A test of a different prior choice of bias indicates that in case of conflicting objectives (good fit of variable g versus good fit of variable h) the prior standard deviation of the bias is a crucial value for trading-off one objective versus the other. The results of the linearized error propagation technique are shown in Fig. 1, and show slightly smaller uncertainty bounds than the results without the linearization shown in Reichert and Schuwirth [2010].

4. APPLICATION TO BIOGEOCHEMICAL AND ECOLOGICAL LAKE MODEL

In the following sections we go through the steps described in section 2 for the application case of a joint calibration of the lake model BELAMO applied to three lakes.

4.1 BELAMO: Model description

The calibration technique described in section 2 was applied to the **B**iogeochemical and **E**cological **L**Ake **M**odel (BELAMO). It aims at a joint calculation of mass balances of nutrients, oxygen, organic particles, phytoplankton and zooplankton. Its box version describes the lake as four boxes: epilimnion, hypolimnion and two sediment boxes. In these boxes, concentrations of ammonium, nitrate, phosphate, oxygen, degradable and inert dead organic particles and (in the aggregated version used for this paper) one group of phytoplankton and one group of zooplankton are modelled.

BELAMO was implemented in AQUASIM (version 2.1f), a computer program for the identification and simulation of aquatic systems (Reichert [1994], Reichert [1998]). For a detailed description of the box version of the model see Mieleitner and Reichert [2008]. For a detailed description of the changes made compared to the model used in Mieleitner and Reichert [2008] and illustrative figures of the model structure and processes accounted for in each of the model compartments, see Dietzel et al. [2010]. Manual calibrations showed the difficulty of calibrating the model evenly well for all output variables and all lakes.

4.2 Study area

The model was applied to long-term observations of the three Swiss lakes Greifensee, Lake Zurich and Walensee. As measured by prevailing phosphorus concentrations Greifensee is still eutrophic, Lake Zurich rather mesotrophic and Walensee is an oligotrophic lake. For a detailed description of the main lake attributes see Mieleitner et al. [2006].

4.3 Data

Monthly measured profiles of physical, chemical and biological variables for Lake Zurich and Walensee were obtained from 1976 to 2005. For Greifensee, monthly to daily measurements of physical, chemical and biological variables were obtained from 1987 – 2004. Information on inflows into the lakes (physical and chemical parameters) and meteorological data were received from federal and cantonal agencies.

4.4 Prior distributions

The parameters included in the parameter estimation procedure comprise a selection of influential and most uncertain model parameters. To be able to estimate the importance of bias in the calibration, fixed standard deviations were assumed for the observation error representing the parameter set Ψ . Finally, exponential priors were chosen for the standard deviations of the bias of all considered output variables, for which the probability density

increases with decreasing value of the standard deviation. This reflects the desire to avoid bias if possible. In general, the bias is assumed to be larger than the observation error.

For the parameter estimation a newer version of UNCSIM (Reichert [2005]), a program package for statistical inference, identifiability analysis and uncertainty analysis, was used by coupling with AQUASIM. Linear approximation of the posterior distribution and estimation of the prediction uncertainty was done with R.

4.5 Results

The calibration technique described above was applied to a calibration period of 9 (for Greifensee) and 20 (for Lake Zurich and Walensee) years. The parameter vector θ for the inference of this application case consists of maximum specific phyto- and zooplankton growth and death rates, the fraction of inert material in lake inputs as well as maximum aerobic and anaerobic mineralization rates. The prediction uncertainty was estimated for the validation period of the following 10 years. Due to the large computation time, for a first rough estimate of the results, the prediction uncertainty was only estimated with the maximum specific phytoplankton growth and death rates and the fraction of inert material in lake inputs. Here the deterministic model function $y_M^L(\theta)$ consists of discrete-time model outputs of the different variables oxygen, phosphate, nitrate, phytoplankton and zooplankton in the epilimnion and hypolimnion of all three lakes.

Fig. 2 and 3 show examples of simulation results of BELAMO for both the calibration and validation time. They represent the median of the distribution of model results ($y_M^{L,2}$) due to the posterior distribution of model parameters θ conditional on the model results of layout 1 without contribution of the bias. Furthermore, the 68% credibility intervals are shown. The results depict our knowledge about the true state of the system (without measurement errors). In the first part, observations were used to enhance our knowledge.

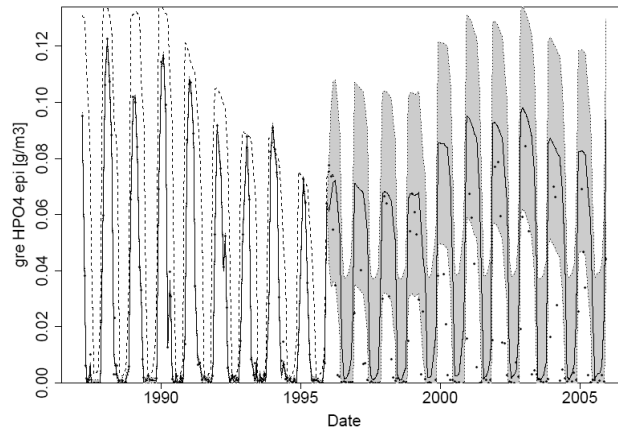


Figure 2. Phosphate concentration in the epilimnion of Greifensee. Data points (markers), median (solid) and 68% credibility interval (shaded area with dotted boundaries) of $y_M^{L,2} + B_M^{L,2} | Y_M^{L,1}$ (model results plus bias conditional on observations of layout 1) for both layouts and median of $y_M^{L,2}$ (results without bias) for the whole simulation time (dashed).

In the second part, observations were neglected, which can be seen as prediction of the “future” state. Hence, plotted data points were used for the calibration of the model only for the observation layout (layout 1). Observations within layout 2 are presented for validation of the predictive power of the model.

Depending on the correlation time (in our case around 2 months), the gained information about the bias for layout 1 gives information also for the beginning of the validation time.

After the correlation between the bias of layout 1 and 2 decreased to 0, the median of $\mathbf{y}_M^{L1,2} + \mathbf{B}_M^{L1,2} | \mathbf{Y}_M^{L1}$ (solid line in Fig. 2 and 3) equals the median of $\mathbf{y}_M^{L1,2}$ (dotted line in both figures).

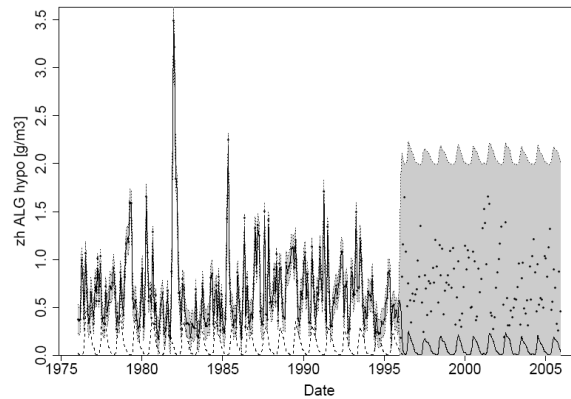


Figure 3. Phytoplankton concentration in the hypolimnion of Lake Zurich. See caption of Fig. 2 for description of the different symbols.

5. DISCUSSION

For both the didactical example and the application case, it becomes obvious that simulation results plus bias are mostly able to reproduce the data during the calibration time. This results from sufficient knowledge about the system in times where measurements are available. Furthermore, the observation error was assumed to be small compared to the contribution of the bias. Both aspects lead to small uncertainties during the calibration time, which is especially true for the shown oxygen results in Greifensee.

The observations of the calibration time help to get information about the model parameters, which improve the prediction. However, due to the deviations between model results and data, the uncertainty is estimated to be much larger for the validation period. As the model results are in less good agreement with the data for the application to BELAMO than for the didactical example, the difference in the uncertainties is larger in this case. This difference is even more extreme for the phytoplankton than for the oxygen results. Hence the estimated uncertainty for the validation period is very large. Those results are meaningful, as poor models should not be used for prediction purposes. In case they are, the user should be aware of the high uncertainty.

A comparison of the remaining results not shown in this paper indicates that annual patterns of both phyto- and zooplankton seem to be most difficult to be represented by the model. The deviations most likely result from simplifications in the spatial and functional aggregation in the model. Those simplifications are less realistic for biological than for chemical variables.

Furthermore, the good representation of data by the model results plus bias compared to the relatively poor pure model results demonstrates the large contribution of bias in model results. This is especially true for the rather complex model BELAMO. A more detailed insight into the importance of bias for the different output variables could be given by the comparison of prior and posterior marginal distributions of the model parameters and the standard deviations of the bias not presented in this paper. Those results indicate how much can be learned about the parameters due to the available data.

In general, the main aim should be to decrease bias by improving the model. But especially in complex models bias will be present nevertheless and a transparent way has to be found to deal with it. The proposed technique makes the estimation of uncertainties in model predictions possible while still fulfilling the statistical assumptions of the error model.

Although the technique requires subjective choice, i.e. choice of prior distributions of standard deviations of bias for each of the output variables, it is a transparent way of uncertainty analysis and makes explicit what often is done implicitly. This subjective choice influences the results of uncertainty estimates, but it is unavoidable for multi-objective calibrations, as there are no objective criteria to distribute the bias among different output variables. However, the decision is made by the best prior knowledge the modeller has about the system. In specific, the user gets information about the importance of the bias in model simulation. The results show the importance of using such techniques, on the one hand, and, on the other hand, the need to carefully use deterministic models as BELAMO for prediction purposes. The deviations between data and model results demonstrate the difficulty of a joint calibration of three different application objects over a long-term period.

REFERENCES

- Bayarri, M.J., J.O. Berger, R. Paulo, J. Sacks, J.A. Cafeo, J. Cavendish, C.-H. Lin, and J. Tu, A framework for validation of computer models, *Technometrics*, 49(2), 138-154, 2007.
- Boyle, D.P., H.V. Gupta, and S. Sorooshian, Multicriteria calibration of hydrologic models. In: Q. Duan, H.V. Gupta, S. Sorooshian, A.N. Rousseau, and R. Turcotte (Eds.), *Calibration of Watershed Models*, pp. 185-196, American Geophysical Union, Washington, DC, 2003.
- Craig, P.S., M. Goldstein, J.C. Rougier, and A.H. Seheult, Bayesian forecasting for complex systems using computer simulators, *Journal of the American Statistical Association*, 96, 717-729, 2001.
- Craig, P.S., M. Goldstein, A.H. Seheult, and J.A. Smith, Bayes linear strategies for matching hydrocarbon reservoir history. In: J.M. Bernardo, J.O. Berger, and A.P. Dawid (Eds.), *Bayesian Statistics 5*, pp. 69-95, Clarendon Press, Oxford, 1996.
- Dietzel, A., J. Mieleitner, S. Kardaetz, and P. Reichert. Effects of changing driving forces on water quality and plankton dynamics in three Swiss lakes - Long-term simulations with BELAMO, in preparation.
- Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin, *Bayesian Data Analysis*, Chapman & Hall, London, 1995.
- Gupta, H.V., L.A. Bastidas, J.A. Vrugt, and S. Sorooshian, Multiple criteria global optimization for watershed model calibration. In: Q. Duan, H.V. Gupta, S. Sorooshian, A.N. Rousseau, and R. Turcotte (Eds.), *Calibration of Watershed Models*, pp. 125-132., American Geophysical Union, Washington, DC, 2003.
- Gupta, H.V., S. Sorooshian, and P.O. Yapo, Toward improved calibration of hydrologic models: Multiple and non-commensurable measures of information, *Water Resources Research*, 34(4), 751-763, 1998.
- Higdon, D., M. Kennedy, J.C. Cavendish, J.A. Cafeo, and R.D. Ryne, Combining field data and computer simulations for calibration and prediction, *SIAM Journal on Scientific Computing*, 26(2), 448-466, 2004.
- Kennedy, M.C. and A. O'Hagan, Bayesian calibration of computer models, *J. R. Statist. Soc. B: Statistical Methodology*, 63(3), 425-464, 2001.
- Madsen, H., Automatic calibration of a conceptual rainfall-runoff model using multiple objectives, *Journal of Hydrology*, 235, 276-288, 2000.
- Madsen, H., G. Wilson, and H. C. Ammentorp, Comparison of different automated strategies for calibration of rainfall-runoff models, *Journal of Hydrology*, 261, 48-59, 2002.
- Mieleitner, J. and P. Reichert, Analysis of the transferability of a biogeochemical lake model to lakes of different trophic state, *Ecological Modelling*, 194, 49-61, 2006.
- Mieleitner, J. and P. Reichert, Modelling functional groups of phytoplankton in three lakes of different trophic state, *Ecological Modelling*, 211, 279-291, 2008.
- Reichert, P., AQUASIM - A tool for simulation and data analysis of aquatic systems, *Water Science and Technology*, 30(2), 21-30, 1994.
- Reichert, P., AQUASIM 2.0 - User Manual, Technical report, Swiss Federal Institute for Environmental Science and Technology (EAWAG), Dübendorf, Switzerland, 1998.

- Reichert, P., UNCSIM - a computer programme for statistical inference and sensitivity, identifiability, and uncertainty analysis. In: Proceedings of the 2005 European Simulation and Modelling Conference (ESM 2005), Porto, Portugal, pp. 51-55, EUROSIS-ETI, Oct. 24-26, 2005.
- Reichert, P. and N. Schuwirth, Linking statistical description of bias to multi-objective model calibration, in preparation.
- Yapo, P. O., H. V. Gupta, and S. Sorooshian, Multi-objective global optimization for hydrologic models, *Journal of Hydrology*, 204, 83-97, 1998.