NETLAKE toolbox for the analysis of high-frequency data from lakes

Factsheet #8

Bayesian calibration of mechanistic models of lake metabolism

Mark Honti

Objective

Resolve the identification issue (different pairs of production and respiration rates produce very similar dissolved oxygen time-series) that arises when complex mechanistic (process-based) metabolic models are calibrated against high-frequency dissolved oxygen (DO) measurements.

Specific application

Changes in DO are primarily related to net ecosystem production (NEP), and as such the time-dynamics of DO shows an aggregated picture on lake metabolism. Physical (e.g. gas exchange with the atmosphere, transport in the water) and chemical (e.g. many redox reactions) processes also contribute to these changes. Thus, it is difficult to disentangle major metabolic processes such as gross primary production (GPP) and ecosystem respiration (R). Several combinations of GPP and R result in very similar DO dynamics resulting in uncertain estimation of photosynthetic and respiration parameters. Instead of seeking for the parameter combination best fitting the data, Bayesian calibration narrows the domain of parameter combinations that yield similarly good fit on the basis of your prior expectations about parameter values. Sampling of posterior parameter distributions yield uncertainty distributions for each parameter.

Background

- Understanding lake metabolism.
- Experience in mechanistic modelling.
- Basic experience in programming.
- Understanding basic statistics (concepts of probabilities, probability distributions).

Type of data and requirements

For the most basic metabolic model, high frequency (30 min or less) records of DO, water temperature (vertical temperature profile), Photosynthetically Active Radiation (PAR), and wind velocity are needed. In shallow lakes, the coefficient of diffuse light attenuation (~turbidity) is used.

Extended metabolic models may use a set of additional data:

- Phytoplankton biomass (~chlorophyll fluorescence)
- Eddy diffusivity in stratified lakes
- Flow velocity and direction
- Wind direction
- pH, conductivity, alkalinity, CO_2 concentration

Basic procedures

Bayesian parameter inference is an advanced calibration technique, so it is assumed that a mechanistic metabolism model is already up and running.

The first step is to formulate expectations on the parameter values that is to set up the so-called prior distributions. This is usually done by explicitly listing the expected range and expected high probability region (if any) for each parameter based on literature values, expert opinion and domain of meaningfulness (e.g. values below or above a threshold are accepted or not). This information is then compiled into a proper statistical distribution for each parameter. The types and parameterisations (like: mean, standard deviation, etc.) of prior distributions express your subjective willingness to accept a certain value for the parameter in question.

Common prior distribution types are:

- uniform (there is a strictly defined meaningful domain, but there is no preferred choice within that domain),
- beta (the domain of meaningfulness is between 0 and 1 with a peak somewhere in between),
- normal (there is a preference for the mean value, there are no limits, deviations from the preference are accepted in both directions with the same decreasing probability),
- log-normal (negatives are not accepted, a certain deviation above the preference is accepted with higher probability than below it)

Besides these typical examples, any proper unimodal (=having a single peak) statistical distribution will do, if it properly expresses your subjective scientific expectations against the parameter.

The core of the procedure can either be done by modifying your present calibration routine or by plugging your model into a Bayesian calibration framework (e.g. JAGS or BUGS). The first option is discussed below.

Bayesian parameter inference requires the goodness-of-fit measure to be a proper statistical likelihood function. Therefore, if you previously used RMSE, Nash-Sutcliffe or similar informal measures, you have to modify the evaluation module of your script. For high-frequency DO data equidistantly sampled in time the best-suited formal statistical likelihood function is the first-order autoregressive error model. This has 2 parameters: the standard deviation of error innovations (e.g. the change of error from one timestep to the other) and the one-step autocorrelation coefficient. The log-likelihood (*log L*) of a certain parameter combination is calculated from the residual time-series (*E*) as follows:

$$\log L = -\frac{n}{2}\log(4\pi) - \frac{1}{2}\sum_i I_i^2$$

where *n* is the length of the residual series, and I_i are the scaled innovations of the residual series at each timestep except the very first one ($I_i = \frac{E_i - \rho \cdot E_{i-1}}{\sigma}$, where ρ is the autocorrelation between steps and σ is the standard deviation of error innovations). The likelihood is used in combination with the prior probability to evaluate model performance:

$$P_{\text{post}} \propto P_{\text{prior}} \cdot L$$

P_{post} is the posterior probability function that should be used as a new objective function in the calibration procedure. In practice, log posterior probability is used to prevent numerical underflows (when small numbers are accidentally rounded to 0) during computation:

$$\log P_{\text{post}} \propto \log P_{\text{prior}} + \log L$$

Using the autoregressive error model one arrives at the following equation for log posterior probability:

$$\log P_{\text{post}} = -\frac{n}{2}\log(4\pi) - \frac{1}{2}\sum_i I_i^2 + \sum_j \log P_{j,\text{prior}}$$

where *j* iterates over the model parameters. The log prior probability of individual parameters ($P_{j,prior}$) couldn't be expanded further in the above equation as it depends on the type of the prior distribution (e.g. normal, lognormal, uniform, etc.) assumed for the given parameter.

The optimal parameter combination will be a compromise between model fit and your subjective expectations. When parameter identification is poor, this compromise usually fits almost as well as unconstrained calibration. It is worth noting that unconstrained calibration does not deliver the objective truth, which may or may not be revealed by unrealistic parameter values.

The uncertainty of posterior parameters can be derived by producing a numerical sample from the posterior parameter distribution using Markov Chain Monte Carlo (MCMC) sampling. The core of this rejection sampling algorithm (Metropolis-Hastings sampler) is:

1. Start with any arbitrary parameter combination. For practical reasons, the combination that belongs to the maximum posterior probability is preferred, if available.
2. Create a new parameter combination from (1) by using a "jump" or "proposal" distribution: Generate a random normal number for each parameter with mean centred at the previous parameter value.
3. Evaluate the log posterior with the new parameter combination. If it is higher than the log posterior of the previous combination, accept the new parameter values and go back to 2. If the new posterior probability is lower, accept the new parameter combination with $P_{post,new} / P_{post,previous}$ probability (or $\exp\{\log P_{\text{post,new}} - \log P_{\text{post,old}}\}$ when log probabilities were used) and go back to step 2.

Repeating this cycle sufficient times (10^3 to 10^4 iterations), the set of parameter values that have been accepted at step 3 will converge to a proper sample from the posterior parameter distribution. The first portion of the sample is usually discarded because it is distorted by the stabilisation of the sample. The second part of the sample should look like thick noise bands in terms of both posterior probability and parameter values.

The posterior uncertainty of individual parameters can be visualized by extracting the posterior marginal distributions from the sample in plots of density functions of each parameter (Figure 1).

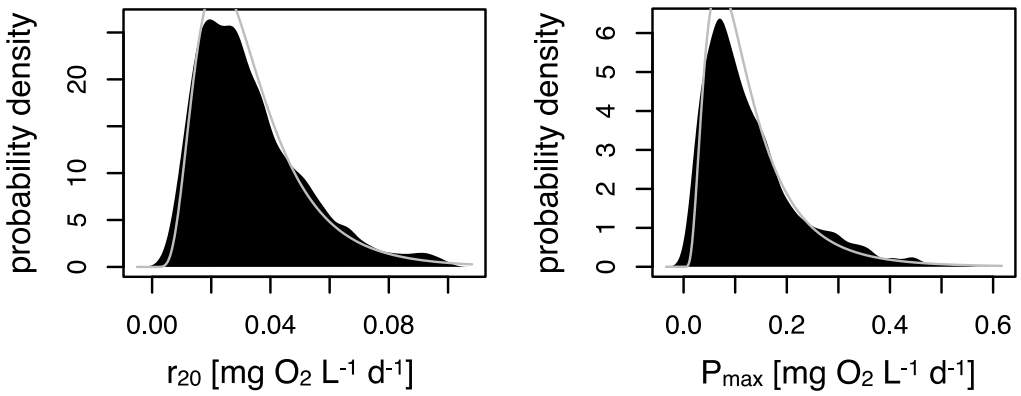

Figure 1. Selected posterior parameter marginals (black shading) from an MCMC sample. r_{20}: community respiration rate at 20˚C, P_{max}: maximal rate of gross primary production. The thin grey lines show a fitted lognormal distribution.

Although the principle of MCMC is simple and any implementation following the basic algorithm will work, there are several intricate tricks to make the sampler more efficient. These include a gradual fine-tuning of the proposal distribution to reflect the size and correlation structure of the posterior, thinning the sample to reduce serial correlation, and many others. Therefore, it is generally advisable to use the many existing MCMC implementations of 'R' or any other statistical environment.

Pitfalls and tips

Bayesian calibration does not resolve the identification problem of metabolic parameters in an objective way. As priors are subjective, posteriors represent a subjective compromise. Different priors would lead to different posteriors.

Bayesian calibration suffers from the general issue of parameter interpretability just like any other calibration method used for any type of mechanistic model: parameters are optimised during the calibration to compensate for structural deficiencies of the model. Therefore, parameters are biased and hence should be considered as abstract quantities with limited and uncertain physical, chemical or biological meaning. This limits the confidence in analysing calibrated parameter values.

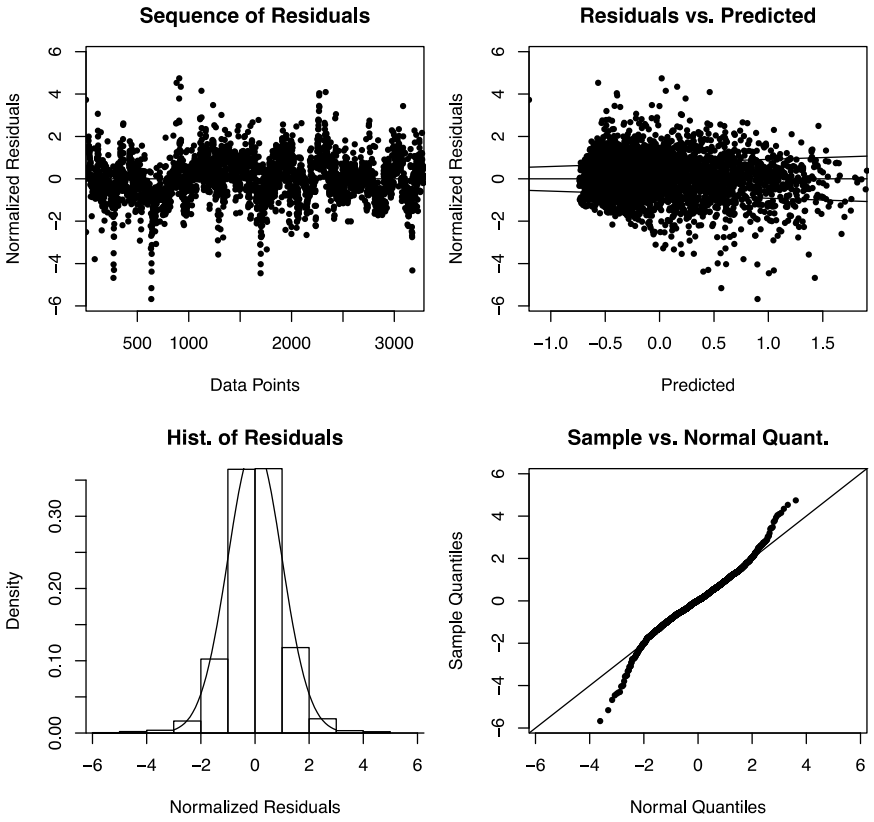

Figure 2. Residual diagnostic plots. Top left: sequence or trace plot; top right: residuals as function of the predicted (DO) value; bottom left: histogram of standardised residuals and a fitted normal distribution; bottom right: Q-Q plot.

Tips

- *Validating the error model.* In formal statistical approaches the likelihood function has to be validated against the posterior residuals to ensure that the statistical assumptions behind the error model are correct or at least not far from reality. This is usually done by testing each assumption on the residuals between observations and the maximum posterior probability solution. In the case of a metabolic model and autoregressive errors, this means testing if residuals have no significant autocorrelation beyond a 1-step lag (*acf* plot), and that innovations are normally distributed with a mean of zero (Q-Q plot). Figure 2 shows a thorough analysis for

independent, normally distributed residuals via plotting their sequence, their dependence on the predicted variable, their density function and a Q-Q plot (layout courtesy of Peter Reichert, EAWAG).

- *Checking MCMC progress.* To assure that MCMC converges successfully, it is common to launch parallel chains and observe whether they converge to the same region. Typical chain lengths are in the range of 2,000–100,000 iteration cycles. It can be shown that the proposal distribution is acceptably tuned if the mean acceptance probability is between 15 and 40 %.
- *Interpreting posteriors.* Posteriors may show two typical relations to priors. If they are very similar to priors, the calibration data did not contain any new and meaningful information about the parameters. This indicates weak identifiability. If posterior distributions are significantly narrower than prior ones, data contained useful information on parameters and hence, priors were suppressed to some degree. Nevertheless, posteriors still remain conditional on priors unless an infinitely long dataset is used for calibration.

Further reading

Key References:

As Bayesian statistics is a fully-fledged discipline within statistics, there are dozens of thick textbooks on the topic. A good example is:

Gelman A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B. 2013. Bayesian Data Analysis. 3rd edition. CRC Press.

Other useful references:

Application examples related to advanced calibration of metabolic models include

Use of first-order autoregressive error model in calibration:

Van de Bogert, M.C., Carpenter, S.R., Cole, J.J. Pace, M. 2007. Assessing pelagic and benthic metabolism using free water measurements. *Limnology and Oceanography: Methods* 5: 145-155.

Use of first-order autoregressive error model in calibration, parameter uncertainty assessed with bootstrapping:

Solomon, C.T., Bruesewitz, D.A., Richardson, D., Rose, K., Van de Bogert, M., Hanson, P., Kratz, T., Larget, B., Adrian, R., Babin, B.L., Chiu, C.Y., Hamilton, D.P., Gaiser, E., Hendricks, S., Istvánovics, V., Laas, A., O'Donnell, D.M., Pace, M., Ryder, E., Staehr, P.A., Torgersen, T., Vanni, M.J., Weathers, K., Zhu., G. 2013. Ecosystem respiration: Drivers of daily variability and background respiration in lakes around the globe. *Limnology and Oceanography* 58: 849-866.

Use of first-order autoregressive error model in calibration, parameter uncertainty assessed with PEST (informal likelihood procedure with Monte Carlo):

Hanson, P.C., Carpenter, S.R., Kimura, N., Wu, C., Cornelius, S.P., Kratz, T.K. 2008. Evaluation of metabolism models for free-water dissolved oxygen methods in lakes. *Limnology and Oceanography: Methods* 6: 454-465.

Use of Kalman filter (an example of linearised Bayesian updater) with independent, identically distributed error:
Batt, R.D., Carpenter, S.R. 2012. Free-water lake metabolism: Addressing noisy time series with a Kalman filter. *Limnology and Oceanography: Methods 10:* 20-30.

BaMM - Proper Bayesian inference with independent, identically distributed error and simple multi-objective calibration:
Holtgrieve, G.W., Schindler, D.E., Branch, T.A., A'mar, Z. 2010. Simultaneous quantification of aquatic ecosystem metabolism and reaeration using a Bayesian statistical model of oxygen dynamics. *Limnology and Oceanography* 55: 1047–1063.

A complex Bayesian error model demo on DO data from a Swiss river:
Reichert, P., Schuwirth, N. 2012. Linking statistical description of bias to multi-objective model calibration. *Water Resources Research* 48: W09543.

Code

Due to the task-specific requirements there aren't any ready solutions that would meet all limnological needs, but there are solid frameworks which help to carry out the basic steps of Bayesian parameter inference and uncertainty analysis. It is advised to start with the examples attached to these frameworks and develop your own likelihood function, etc.

Rpackages for Bayesian inference can be downloaded from CRAN (by the 'install.packages' command): *mcmc*, *rjags*.

The LakeMetabolizer Rpackage can help you to assemble your metabolic model.

Contact details

Mark Honti. Budapest University of Technology and Economics, Hungary.
honti@vit.bme.hu

Suggested citation

Honti, M. 2016. Bayesian calibration of mechanistic models of lake metabolism. In Obrador, B., Jones, I.D. and Jennings, E. (Eds.) *NETLAKE toolbox for the analysis of high-frequency data from lakes* (Factsheet 8). Technical report. NETLAKE COST Action ES1201. pp. 40-46. http://eprints.dkit.ie/id/eprint/539.

Acknowledgement

This factsheet is based upon work from the NETLAKE COST Action ES1201, supported by COST (European Cooperation in Science and Technology).