USER GUIDE FOR RUNNING BASE v2

User guide correspondence: Darren Giling, darren.giling@idiv.de

Supplementary Material for:

Grace, M.R., Giling, D.P., Hladyz, S., Caron, V., Thompson, R.M. and Mac Nally, R. (2015). Fast processing of diel oxygen curves: estimating stream metabolism with BASE (BAyesian Single-station Estimation). *Limnology & Oceanography: Methods*, 13, 103–114.

http://onlinelibrary.wiley.com/doi/10.1002/lom3.10011/abstract

Contents

1. *Introduction*

This is a guide to estimate single-station whole-stream metabolic rates from diel dissolved oxygen (DO) curves using the BASE program described by Grace *et al.* (2015). Please refer to and cite Grace *et al.* (2015) for the program, and note the updates detailed in Section 2.

Grace *et al.* (2015) provides a brief overview of whole-stream metabolism methods and a description of the Bayesian estimation model. Here, we describe how to structure the input data and to run the model to calculate metabolic rates from diel DO curves. Processing can be done in batch mode. Running the model is straightforward, but requires familiarity with R (R Development Core Team 2011). The parameter estimation is performed with the JAGS software (Plummer 2003). The model does not require experience with JAGS, but to gain a better understanding of the methods and outputs (e.g. for checking model convergence), we recommend consulting introductory texts on Bayesian methods (e.g. Kéry 2010; McCarthy 2007).

2. *Update history*

- June 2014: BASE v1 released with Grace *et al.* (2015)
- July 2016: BASE v2 released, major changes:
    - Changes to the diel DO model structure following Song *et al.* (2016). Results published by Song *et al.* (2016) showed that BASE v1 was underestimating metabolic rates in some systems due to two differences in the model formulation compared to other aquatic metabolic models (e.g. Hall and Tank 2005; Hanson et al. 2008; Holtgrieve et al. 2010; Van de Bogert et al. 2007):
        1. BASE v1 used a 'stepwise' approach to model changes in DO concentration (Δ[DO]) between successive measurements rather than DO concentration ([DO]) directly.

2. BASE v1 used the measured DO concentration ($[DO]_{measured}$) to estimate oxygen deficiency for reaeration rates instead of the modelled DO concentration ($[DO]_{measured}$).

In light of Song *et al.* (2016)'s findings, both these inconsistencies in the BASE model were rectified:

BASE v1:

$$\Delta[DO]_t/\Delta t = AI_t^p - R\left(\theta^{(T_t-\bar{T})}\right) + K_{DO}.\left(1.0241^{(T_t-\bar{T})}\right).([DO]_{sat,t} - [DO]_{meas,t})$$

BASE v2:

$$[DO]_{t+1} = [DO]_t + AI_t^p - R\left(\theta^{(T_t-\bar{T})}\right) + K_{DO}.\left(1.0241^{(T_t-\bar{T})}\right).([DO]_{sat,t} - [DO]_{modelled,t})$$

Where *t* indicates the timestep, *A* is a constant, *p* is an exponent describing incident light use, Θ describes temperature dependence of respiration, *T* is water temperature and *sat*, *meas* and *mod* indicate [DO] at saturation, observed concentration and modelled concentration. Please refer to Song *et al.* (2016) for a description of the assumptions underlying the alternate models. The updated model of BASE v2 improved fit (visual and R^2) of modelled DO to observed values and reduced uncertainty in parameter estimates of the validation dataset of Grace *et al.* (2015). There was greater congruence between BASE v2 estimates and BaMM (Holtgrieve et al. 2010) estimates, the 'accurate' method of Song *et al.* (2016), across a wide range of stream characteristics and range of metabolic rates (Figure 1). Some remaining differences may be due to differences in the photosynthesis-irradiance (PI) curve used in each model.

- Change in software for MCMC algorithm from OpenBUGS (Lunn et al. 2000) to JAGS (Plummer 2003) to allow for parallel core processing of the three chains, greatly increasing computational speed. Fastest results will therefore be achieved on a processor with at least 3 cores.

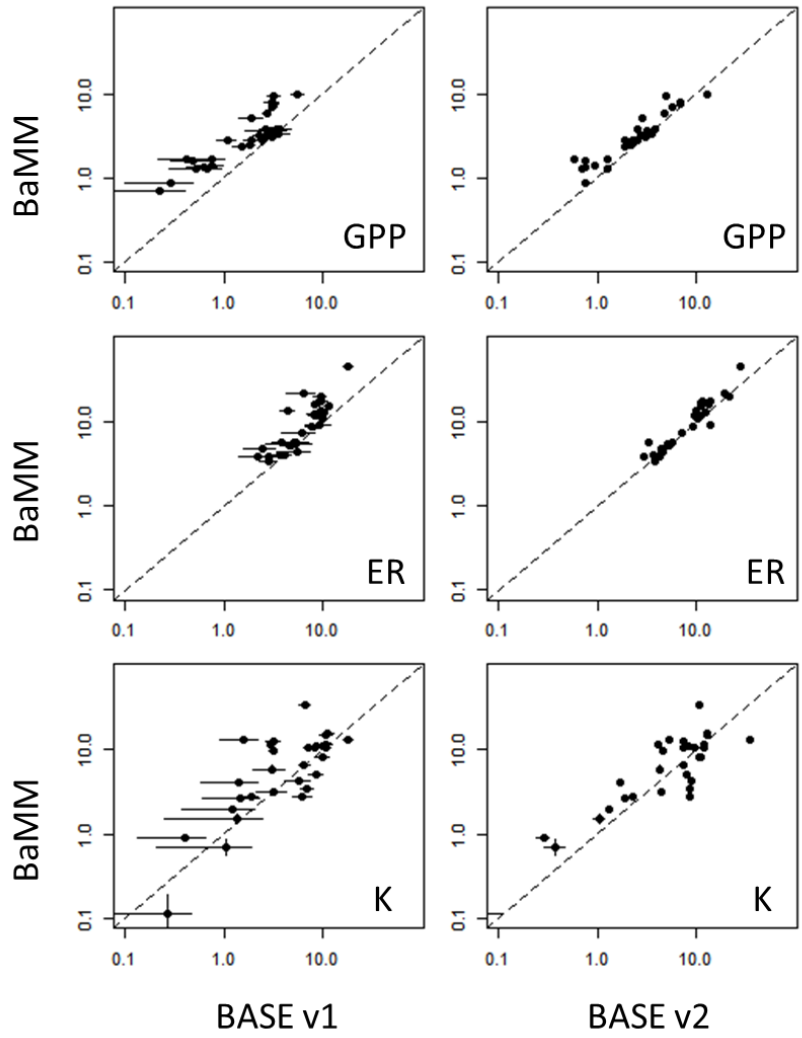

Figure 1. Comparisons between BASE (3-parameter model) and BaMM (Holtgrieve et al. 2010) for converged estimates of log-transformed GPP (mg O_2 L^{-1} day^{-1}), ER (mg O_2 L^{-1} day^{-1}) and K (day^{-1}). Data are those validation streams from Grace *et al.* (2016) with a measurement interval of 5 or 10 minutes. Note reduced uncertainty in BASE v2 estimates compared to BASE v1.

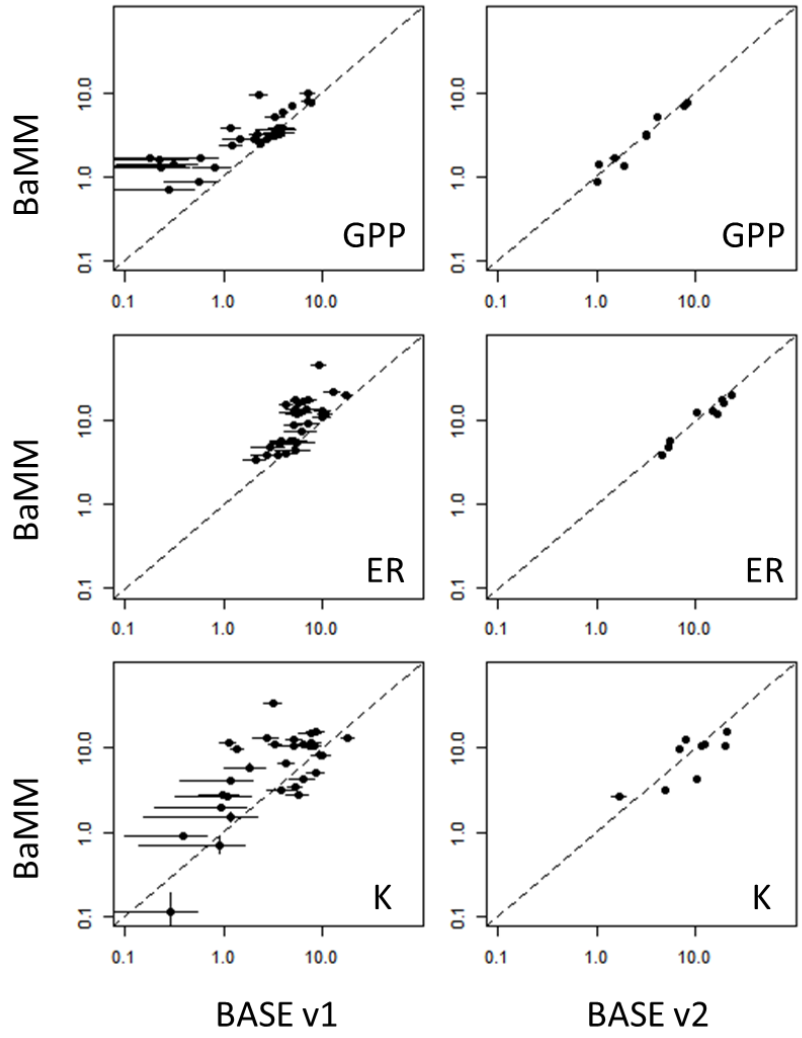

Figure 2. Comparisons between BASE (3-parameter model) and BaMM (Holtgrieve et al. 2010) for converged estimates of log-transformed GPP (mg O_2 L^{-1} day^{-1}), ER (mg O_2 L^{-1} day^{-1}) and K (day^{-1}). Data are those validation streams from Grace *et al.* (2016) with a measurement interval of 5 or 10 minutes. Note reduced uncertainty in BASE v2 estimates compared to BASE v1. Convergence rates for the 5-parameter model were lower for BASE v2.

3. *Required software and model files*

Download and install R (http://www.r-project.org/) and JAGS (http://mcmc-jags.sourceforge.net/). Update software if necessary.

Step 1

Extract the zipped 'BASE_v2' folder to a location on your hard drive. This folder contains the R and JAGS code, and the subfolders 'input' (data) and 'output' (results). You may create new folders, but do not alter the folder structure inside the 'BASE' folder.

4. *Code description*

There are two files (located in the 'BASE' folder) required for the model:

Script 1: *Call_BASE_v2.R*

Script 2: *BASE_metab_model_v2.txt*

Script 1 is an R script used to define the diel data vectors and Bayesian model parameters within the R environment. The Bayesian model is then called from R to run in JAGS (occurs in the background), looping iteratively through each data file (diel time-series) defined in the 'input' folder (see Section 5). Results are written to file after each data-file iteration in the 'output' folder (see Section 7).

Script 2 (the model file) is the model that takes the data and parameters packaged by Script 1 to run the Bayesian model. Temperature and salinity corrections are made, and the daytime regression model is fitted to measured DO data; key outputs are estimates of A (constant used to calculate GPP), R (instantaneous respiration rates) and K (the reaeration coefficient). See Grace *et al.* (2015) for a full description of the daytime regression model.

5. *Input file location and format*

An example data set of diel curves is provided with the download in the 'input' folder. We recommend familiarizing yourself with the model by first using this data series, and subsequently using one of the comma-separated values (.csv) files as a template to input your own data. This will ensure correct formatting.

Rates can be estimated for multiple diel time-series in one model run. Each DO diel time-series is provided in a separate csv file within the 'input' folder. Each data file must be a 24-hour time series of DO measurements (5 or 10 minute data intervals are commonly used). *Dissolved oxygen concentration should be corrected for probe drift prior to running the metabolic model (Grace and Imberger 2006).* The first data point should be midnight (24h00) of the day of interest and the last data point should be midnight of the following day. See Section 8 for description of code to split long time series into separate files.

Example input:

Date	Time	I	tempC	DO.meas	atmo.pressure	salinity
2012-04-05	0:00:00	0	16.32	8.11	0.976886	0.071737
2012-04-05	0:05:00	0	16.29	8.112	0.976886	0.071737
2012-04-05	0:10:00	0	16.27	8.107	0.976886	0.071737
...	...	...	...	...	...	...
2012-04-05	23:55:00	0	16.68	8.077	0.977577	0.071737
2012-04-06	0:00:00	0	16.67	8.064	0.977281	0.071737

Where:

I	Photosynthetic active radiation (PAR; in μ mol m^{-2} s^{-1}).
tempC	Stream water temperature (in degrees Celsius).
DO.meas	Measured dissolved oxygen concentration (in mg L^{-1}).
atmo.pressure	Measured atmospheric pressure in atmospheres. Can be constant (i.e. fill every time interval with same value) and inferred from stream altitude if barometric data is unavailable. A default of 1 can be used if pressure and altitude are unknown.

salinity	Water salinity (in ppt). Can be constant (i.e. fill every time interval with same value) or a time-series. Salinity does not play a large role in determining DO saturation in freshwaters; a default of 0 can be used in salinity is low and unknown.

IMPORTANT:

- Correct column headings are vital (they are case-sensitive).
- All columns must contain data for each time interval.
- For the 'Separate files to days' code, ensure correct date and time formatting. Date must be separated by dashes (-), and not by slashes. Time should be formatted to hh:mm:ss.

6. Running the model

6.1 Default model

Step 2

Open the R script *Call_BASE_v2.R* by selecting 'File', then 'Open script...'

The first time you run the model you will need to install several R packages. They can be installed using the following code:

```
install.packages("coda")
install.packages("R2jags")
install.packages("zoo")
```

Sections of the code in one or both scripts must be updated to adjust the model for your system and data. This section describes the minimum lines of code that must be specified to run the default (3-parameter) model (these lines are all within Script 1). The model can be customized by including prior information for K or changing how some model parameters are estimated, if desired (described in Section 6.2).

The capital letters listed below (A, B and C) are paired with corresponding letters in the code of Script 1 where the code is to be updated.

Step 3: Amend code at lines A, B and C in Script 1

(A) Define the location of 'BASE' folder

Tell R where to find the unzipped folder. Replace "[your directory]" in the code at (A) with the location of the 'BASE_v2' folder on your disc. Use forward slashes to indicate folder levels. Windows explorer uses back slashes, so you will have to change these. For example:

```
folder.location <- "C:/Desktop/Analysis"
```

(B) Define the measurement interval

Define the measurement interval of your DO time-series (in seconds).

(C) Define the number of model iterations

Define the total number of Bayesian model iterations and number of burn-in ('settling') iterations. By default this is set to 20000 iterations with 10000 burn-in, which should be sufficient in most cases for the 3-parameter model. This can be reduced (e.g. to 200/100) to quickly test if the model is functioning properly before proceeding with the full analysis. The number of required iterations can be assessed visually and by inspecting the convergence statistics (see section 7.2).

Step 4: Run the model

Once the code at A, B and C in Script 1 has been amended to your requirements, start the model by running the entire Script 1 within R. Select the 'Edit' menu, then 'Run all'.

This will call the JAGS program, which will operate in the background, and may take up to several minutes for each diel cycle (i.e. file in the input folder). The script will loop through each file in the input folder without further user input. You do not need to run Script 2 manually. Pressing ESC in R will stop the model from looping to the next file. Results are saved after each file is completed. The model outputs are described in Section 7. The user can check the progress of the model by seeing how many fitting plots have been written (see description Section 7). Do not open the results csv file while the model is running because further results cannot then be written.

6.2 Optional model customization

By default, BASE is set to estimate parameters *A*, *R* and *K* simultaneously (i.e. a 3-parameter model) and the parameters Θ (theta) and *p* have fixed values.

There is an option to change these defaults:

- Priors for K: estimated or measured K

  K can be estimated from the model and data with uninformative priors ($K \geq 0$), or you can inform the priors with mean and uncertainty of a measured K (e.g. based on stream morphology or if measured using SF_6 injections).

- Fixed or estimated theta or p

  The constants for temperature dependence and light saturation (Θ and p) can be estimated from the model and data (within narrow, physically realistic bounds), which may enhance model fit. Theta or p are estimated along with GPP, ER and K, making a 4-parameter (theta or p is estimated) model or 5-parameter (theta and p are estimated) model. Selecting the most appropriate model is described in section 7.3 – Model selection.

To alter the defaults, adjust the lines of code (described below) in Script 2 (Y and Z). This is performed by commenting (i.e. adding "#") or un-commenting (removing "#") the appropriate lines of code for the model you are running. Script 2 should be opened in Notepad (or similar),

ensuring that you save any changes before running the model. Note that the alternative lines must have one commented and one uncommented line; the model will produce an error if both are commented or both are uncommented. Start the model by running all of Script 1 in R (Script 2 is not run manually).

(Y) Priors for K

Inform JAGS if you are using informative or uninformative priors here.

(Y1) Uninformative priors

(Y2) Informative priors

(Z) Treatment of theta and p

Inform JAGS if theta and p should be treated as fixed or estimated.

(Z1) theta and p fixed

(Z2) theta and p estimated

NOTE: You can choose to treat either p or theta as estimated and the other as fixed by selecting the appropriate combination of code from lines Z1 and Z2.

7. *Model outputs*

7.1 Results table (BASE_results.csv)

The results table ('BASE_results.csv', located in the 'output' folder) provides the means and standard deviations for the metabolic rates and other parameters estimated by the model. Each row of the csv file is the result for one input file. Rates of 'GPP' (daily gross primary production) and 'ER' (daily ecosystem respiration; calculated from the instantaneous rates R) are expressed in mg O_2 L^{-1} day^{-1}. The reaeration coefficient ('K') is expressed in day^{-1}.

IMPORTANT: Rename or move the results file and plots if you wish to keep the results because they will be overwritten the next time you run the model.

7.2 Model validation (quick guide pg 10): assessing convergence and fit

It is vital to ensure that the MCMC chains of model parameters have adequately converged to a stationary distribution or the model may be inappropriate. The 'R-hat' statistic gives an indication of convergence. Values close to 1 indicate good convergence, while values >1.1 indicate poor mixing. R-hat for model parameters (*A*, *R*, *K*, *theta*, *p* and *GPP*) are included in the results table. Poor mixing can (but not always) be improved by increasing the number of iterations. The output csv contains a column titled 'convergence.check', which tests whether all the R-hat values are < 1.1 and can be used to quickly assess convergence (returns 'fine' when all R-hats < 1.1).

Convergence also can be visually assessed by examining that the distributions are stationary and chains are well mixed on the MCMC trace plots. The model prints a jpg multi-panel plot for each diel model in the 'validation plots' folder of the 'output' folder, including MCMC trace plots for *A*, *p*, *R*, *K* and *Θ*. The three chains should be converged (overlapping) and stationary (centred) (Fig. 3). When set as fixed values the trace plots for *p* and *Θ* show a horizontal lines with no variation.

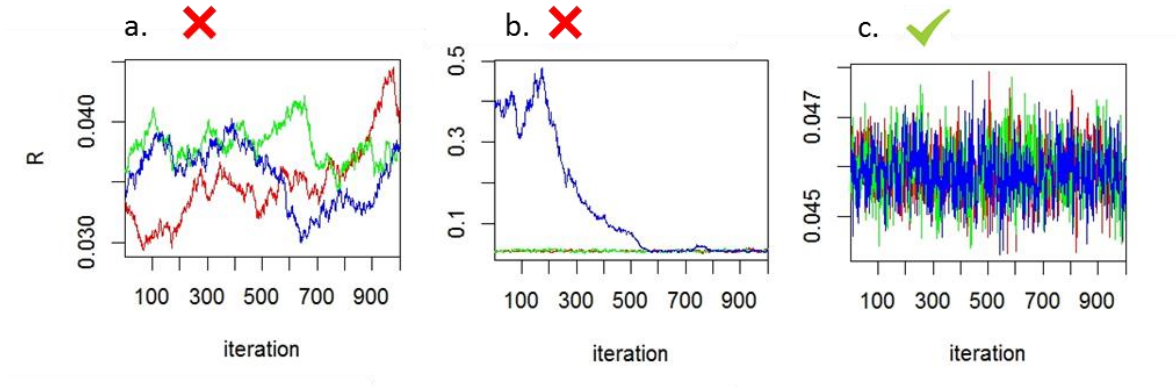

Figure 3. Example trace plots showing (a) three chains with poor mixing, (b) an example where a longer burn-in period is required, and (c) converged and stationary chains.

The model fit also can be confirmed visually using the validation plots, which show the measured (empty circles) and predicted (black line) DO curve for each diel period, as well as the raw temperature and PAR data. These plots can be used to visually confirm curve fits and quickly identify

any discrepancies in the data or model. It may be possible to observe if there are any inconsistencies in the data that may indicate a violation in the assumption of the free-water DO method that reaeration, GPP and ER are the only processes contributing to change in DO. For example, a sharp increase or decrease in temperature or DO may indicate another source of water started to enter the system, or a lack of increase in DO percent saturation during daylight hours may indicate low biological activity compared to reaeration.

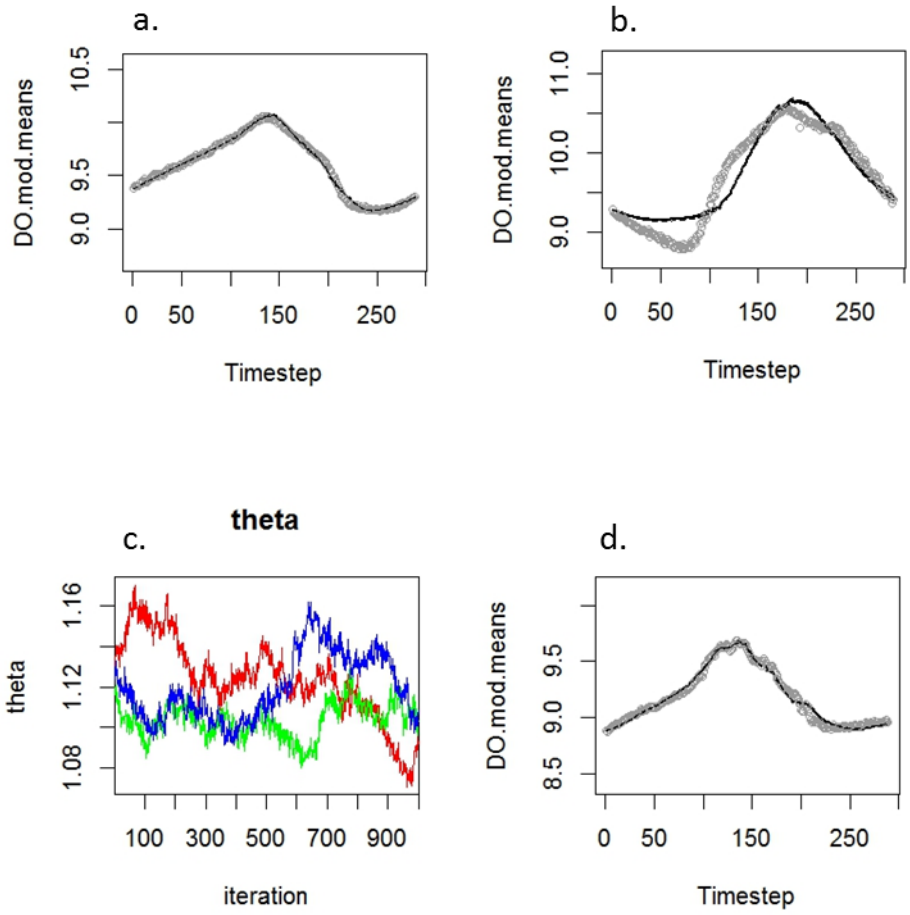

Figure 4. Examples of models with (a) good fit and (b) poor fit. Note that even if model fit is visually adequate (d), the parameter estimates are unreliable if chains are unconverged (c).

The results table also returns the effective number of parameters ('pD'). This value normally should be positive. Negative pD may indicate the posterior mean is not a good measure of the posterior distribution, and there is likely an issue with the model.

Assessing model fit quantitatively

There are three quantitative measures of model fit included in the results table: (1) the posterior predictive p-value (PPP), the (2) R^2 value, and (3) the residual mean square error (rmse). The PPP compares lack of fit of the model to the actual data against lack to fit to a distribution of possible model discrepancies by using data simulated from the parameterized model (Gelman et al. 1996). A PPP value (the PPfit.mean column in the results table) of close to 0.5 indicates a very plausible model, while values <0.1 or >0.9 indicate that the parameterized model is not a plausible explanation of the observed data. The correlation (R^2) between the observed and modelled DO data is reported in the results table. Due to the temporal and correlated nature of the DO time-series, the R^2 may be high when model estimates are consistently above or below measured values. In these cases, the poor fit may be indicated by the residual mean square error (rmse) and maximum run length fraction (mrl.fraction). The rmse is specific to the magnitude of the dataset and should be assessed against models from days at the same site. The rmse is expressed relative to the point-to-point variation in the dataset by the output column 'rmse.relative'. The maximum run length fraction (mrl.fraction) is the proportion of time occupied by the longest run of values for which the estimated DO is below or above the measured DO. A high maximum run length proportion may indicate consistent over- or under-estimation of DO and plots should be inspected.

Quick guide for model validation

To check the model is reliable:

- Ensure all parameters have converged (R-hats < 1.1; convergence.check = 'fine')
- Check PPfit is between 0.1 and 0.9 (closer to 0.5 is best)
- Check pD is positive
- Visually confirm the model is appropriate

7.3 Model selection

The results table returns the Deviance Information Criterion (DIC), an assessment of how well the model will predict a replicate dataset. DIC takes into account the complexity of the model and can be used for model selection (for example from the 3- or 5-parameter model). Lower DIC is desirable, and DIC may be negative.

There is no hard rule, but we recommend a difference in DIC of ≥ 5 between models is evidence that the model with lower DIC best predicts the data because values exceeding 5 correspond to 10-fold or greater support for the model with the lower DIC.

For example:

3-parameter (default) model: DIC = -1797	*alternative, simpler model*
5-parameter (customized) model: DIC -1807	*'best', more complex model*

In this case, the difference in DIC between the alternative and 'best' model is 10 (-1797 − -1807). We would rule the 5-parameter model (with lower DIC) provides a substantially better prediction of the data than the 3-parameter model, despite its additional complexity (two additional parameters estimated).

8. Additional code

8.1 'Separate files to days' code

The code and folders inside the 'Separate files to days' folder can be used to quickly convert a single csv file with long-term sonde data (i.e. many days or months) into separate csv files for each date, which are required by the metabolism model. The input structure for the 'Separate files to days' code is the same as described in Section 4 and an example file is included. The code will split the data according to the 'Date' column and include one extra row from the next date, with the assumption this is the midnight data point. The user needs only to open the script in R, update the input and output directory lines of code, and then run the entire script.

References

Gelman, A., X.-L. Meng, and H. Stern. 1996. Posterior predictive assessment of model fitness via realized discrepancies. Statistica Sinica **6:** 733-807.

Grace, M. R., D. P. Giling, S. Hladyz, V. Caron, R. M. Thompson, and R. Mac Nally. 2015. Fast processing of diel oxygen curves: estimating stream metabolism with BASE (BAyesian Single-station Estimation). Limnol. Oceanogr. Methods **13:** 103-114.

Grace, M. R., and S. J. Imberger. 2006. Stream Metabolism: Performing & Interpreting Measurements, p. 204. Water Studies Centre Monash University, Murray Darling Basin Commission and New South Wales Department of Environment and Climate Change. Accessed at http://www.sci.monash.edu.au/wsc/docs/tech-manual-v3.pdf.

Hall, R. O., and J. L. Tank. 2005. Correcting whole-stream estimates of metabolism for groundwater input. Limnology and Oceanography-Methods **3:** 222-229.

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

Holtgrieve, G. W., D. E. Schindler, T. A. Branch, and Z. T. A'mar. 2010. Simultaneous quantification of aquatic ecosystem metabolism and reaeration using a Bayesian statistical model of oxygen dynamics. Limnol Oceanogr **55:** 1047-1063.

Kéry, M. 2010. Introduction to WinBUGS for Ecologists: A Bayesian Approach to Regression, ANOVA and Related Analyses. Academic Press.

Lunn, D. J., A. Thomas, N. Best, and D. Spiegelhalter. 2000. WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing **10:** 325-337.

Mccarthy, M. A. 2007. Bayesian Methods for Ecology. Cambridge University Press.

Plummer, M. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003).

R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Song, C., W. K. Dodds, M. T. Trentman, J. Rüegg, and F. Ballantyne. 2016. Methods of approximation influence aquatic ecosystem metabolism estimates. Limnol. Oceanogr. Methods**:** n/a-n/a.

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