Article

Robust estimation of lake metabolism by coupling high frequency dissolved oxygen and chlorophyll fluorescence data in a Bayesian framework

Mark Honti,[1] Vera Istvánovics,[1]* Peter A. Staehr,[2] Ludmila S. Brighenti,[3] Mengyuan Zhu,[4] and Guangwei Zhu[5]

[1] *MTA-BME Water Research Group, Hungarian Academy of Sciences, Budapest, Hungary*
[2] *Department of Bioscience, Aarhus University, Roskilde, Denmark*
[3] *Post-Graduation Program in Ecology, Conservation and Management of Wildlife, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil*
[4] *State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, Jiangsu, China*
[5] *Taihu Laboratory for Lake Ecosystem Research, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China*
* Corresponding author: istvanovics.vera@gmail.com

Abstract

Gross primary production (GPP) and community respiration (R) are increasingly calculated from high-frequency measurements of dissolved oxygen (DO) by fitting dynamic metabolic models to the observed DO time series. Because different combinations of metabolic components result in nearly the same DO time series, theoretical problems burden this inverse modeling approach. Bayesian parameter inference could improve identification of processes by including independent knowledge in the estimation procedure. This method, however, requires model development because parameters of existing metabolic models are too abstract to achieve a significant improvement. Because algal biomass is a key determinant of GPP and R, and high-frequency data on phytoplankton biomass are increasingly available, coupling DO and biomass time series within a Bayesian framework has a high potential to support identification of individual metabolic components. We demonstrate this potential in 3 lakes. Phytoplankton data were simulated via a sequential Bayesian learning procedure coupled with an error model that accounted for systematic errors caused by structural deficiencies of the metabolic model. This method provided ecologically coherent, and therefore presumably robust, estimates for biomass-specific metabolic rates and contributes to a better understanding of metabolic responses to natural and anthropogenic disturbances.

Key words: Bayesian parameter inference, dynamic model, net primary production, photosynthesis, respiration, sequential learning

Introduction

Several recent studies have concluded that lakes and reservoirs actively contribute to the global carbon cycle, and their carbon budgets should be taken into account at both regional and global scales (Einsele et al. 2001, Cole et al. 2007, Tranvik et al. 2009). Despite its importance, representative rates of carbon metabolism are extremely difficult to obtain at the ecosystem scale. Dissolved carbon dioxide is difficult to measure and interpret because of its reactivity in water (Hanson et al. 2003). Diurnal dynamics of dissolved oxygen (DO) concentration have been considered the most useful and easily measurable proxy of aquatic ecosystem productivity since the pioneering work of Odum (1956). The increasing use of automated high-frequency optical DO sensors during the last decade has made it easy to obtain long,

high-resolution, precise time series of DO. This development has been perceived as an opportunity to conduct research into lake metabolism by applying previously developed data analysis techniques on high-frequency DO data (Staehr et al. 2010).

Diurnal changes of DO in the water column are influenced by many processes, such as gross primary production (GPP), community respiration (R), atmospheric gas exchange (X), horizontal and vertical transport fluxes (T), and others (O) usually assumed to be less important (e.g., nitrification, precipitation, groundwater). Consequently, the change in DO can be a weak proxy for metabolism; the biological oxygen balance (net ecosystem production [NEP] = GPP − R) covers only a part of the diel DO variability (Rose et al. 2014). The task of deriving metabolic rates from DO data is therefore challenging; the rates should be extracted from an aggregate of weakly known components. This task is thus a classical mathematical identification problem with 5 unknowns in a single equation, solved in practice by reducing the degrees of freedom at the cost of simplifying assumptions. Thus, T and O are typically neglected, X is calculated by empirical formulae, and R is assumed to be constant during a day and is estimated from night data when GPP is zero. These assumptions are, however, crude approximations and not always supported by other studies. For example, respiration rates can vary significantly over the day due to photorespiration of autotrophs (García-Camacho et al. 2012), changing physiological status of organisms (Markager et al. 1992), and changing biomass of both autotrophic and heterotrophic organisms (Solomon et al. 2013, Sadro et al. 2014). Empirical models of gas transfer coefficients are strongly site- and case-specific (Cole and Caraco

1998, Crusius and Wanninkhof 2003, MacIntyre et al. 2010) and hence induce major uncertainty when applied elsewhere (see also, Dugan et al. 2016). Horizontal and/or vertical transport would be negligible only in homogeneous, completely mixed waterbodies, but ecosystems exhibit horizontal and vertical density gradients.

High-frequency DO time-series reveal that DO dynamics in lakes seldom follow a smooth, almost sinusoidal path predicted by simple metabolic models (Hanson et al. 2008). The interplay of ecological and physical processes typically produces a periodic yet rough curve, with irregular oscillations on the top of the daily period (e.g., figures 4 and 7 in Hanson et al. 2008; Fig. 1). Some of the systematic deviations between the smooth modeled curves and reality can be associated with non-modeled phenomena, whereas others remain unexplained even when all available supplementary data are considered (Rose et al. 2014).

Systematic deviations of observed data from existing models violate the statistical assumption of independent model residuals, the basis of traditional measures of error such as sum of squared deviations or root mean square error. To avoid introducing a strong bias to the calibrated parameter values and simultaneously underestimating parameter uncertainty (Reichert and Schuwirth 2012), one has to account for the strong autocorrelation between model residuals (Van de Bogert et al. 2007, Hanson et al. 2008, Solomon et al. 2013). Autoregressive error models can be thought of as Bayesian descriptions of model structural uncertainty (Bayarri et al. 2007, Reichert and Schuwirth 2012). They eliminate erroneous under-estimation of the parameter and prediction uncertainties that common fit measures would commit in the
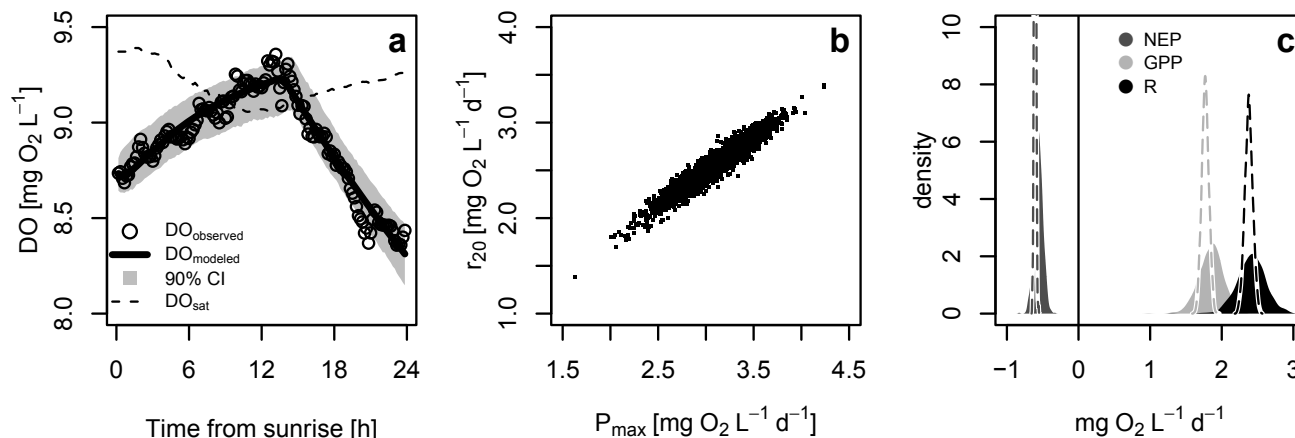
Fig. 1. The parameter identification problem and the underestimated uncertainty using the nonlinear metabolic model of Holtgrieve et al. (2010) on a daily DO curve from Lake Balaton, Hungary (28 Apr 2013). (a) Observed and modeled DO; (b) correlation between high-probability values of maximal daily rate of photosynthesis (P_{max}) and temperature-corrected daily heterotrophic respiration rate (r_{20}); (c) estimated daily sum of NEP, GPP, and R with independent normal (dashed lines) and autoregressive error models (shaded areas).

presence of systematic deviations. Consequently, confidence intervals for parameters will widen when structural uncertainty is handled properly (Fig. 1c). The elevated uncertainty reveals that even the simplest linear metabolic models can suffer from serious parameter identification issues (Hanson et al. 2008). Although several studies have made progress reducing the uncertainty of ecosystem metabolism by constraining and improving estimates of X (e.g., Cole et al. 2010, Dugan et al. 2016), rates of GPP and R are uncertain and strongly correlated in both model calculations (Fig. 1b) and reality. This issue makes it difficult to identify parameters of photosynthesis and respiration when the ratio of these 2 processes varies.

Parameter identification can be improved using Bayesian calibration (parameter inference), during which independent external knowledge and/or data are introduced via prior parameter distributions. This procedure restricts parameter variability and thus reduces correlations between parameters. The main output from Bayesian calibration is the posterior parameter distribution, a formal statistical compromise between prior knowledge or expectations about parameter values and the fit to observed data, the latter being expressed by the likelihood. The price of this compromise in terms of fit quality is usually minor because difficult parameter identification means the model has enough degrees of freedom to adapt to prior expectations without essentially changing the simulated DO dynamics.

Bayesian calibration has been tested on DO data with both independent (Holtgrieve et al. 2010) and autoregressive (Reichert and Schuwirth 2012) error models in flowing waters. Each of these proof-of-concept studies covered only a few consecutive days of data, however, and therefore the models were not confronted with significant changes in the metabolic parameters over time. Here, we extend this modeling framework to lakes and apply the framework to estimate ecosystem metabolism in 3 lakes over the period of several months of measurements.

The benefit of a Bayesian parameter inference depends strongly on the information content of the priors. The posterior distribution is the product of the data-independent prior distribution and the data-dependent likelihood. Narrow prior distributions reflect confident knowledge about parameters, and, being dominant over the likelihood, they effectively concentrate the posterior, hence reducing posterior uncertainty. Vague or high entropy priors, however, do not impose much restriction on parameter values, and the procedure converges to the classical statistical parameter inference because prior probability becomes

almost invariant to the parameters, and therefore the likelihood will determine the shape of the posterior. A fundamental obstacle to obtaining informative priors for metabolic parameters is that parameters of most models are so abstract that estimating them from measurements is impossible. For example, the commonly used efficiency of photosynthesis in terms of surface photosynthetically active radiation (α in Holtgrieve et al. 2010; IP in Hanson et al. 2008; P in Reichert and Schuwirth 2012; ι in Solomon et al. 2013) depends on the composition and biomass of phytoplankton as well as on the vertical diffuse light attenuation coefficient. Consequently, this parameter can vary several orders of magnitude in and between systems. This shortcoming has limited the use of Bayesian parameter inference in studies of lake metabolism.

In addition to the increasing use of high frequency DO sensors, technologies for monitoring phytoplankton have rapidly improved. Various fluorescence-based sensors produce estimates of phytoplankton biomass, and some can also measure the photosynthetic properties of algae. Although fluorescence sensors are widespread in automatic lake monitoring, to our knowledge their data have never been used in connection with metabolic studies. The lack of a bridge between metabolic studies and data on primary producers represents a key missing link between studies of ecosystem metabolism and phytoplankton because GPP and NEP should be closely related to changes in autotrophic biomass. Based on the link between fluorescence and metabolism, fluorescence data could either be used to validate estimates of GPP and NEP or to strengthen prior knowledge about primary producers during parameter inference.

The aim of this study was to (1) develop an improved Bayesian parameter inference procedure that utilizes high frequency chlorophyll fluorescence data to obtain estimates of lake metabolism that are coherent with the observed dynamics of phytoplankton as estimated from high frequency chlorophyll fluorescence; (2) compare the results of this novel approach to metabolic components calculated by the commonly used classical statistical approach that does not use fluorescence data; and (3) test whether the new approach is applicable in various types of lakes with different high frequency monitoring configurations.

Table. 1. Process rates, stoichiometric factors, and affected state variables of the metabolic model.

Process	State variables			Rate
	DO (O_2)	B_a (Chl)	P^B_{max}	
	[mg L^{-1}]	[μg L^{-1}]	[mg O_2 (μg Chl d)$^{-1}$]	
atmospheric gas exchange	1			k_X/z_{mix} ($DO_{sat} - DO$)
photosynthesis	1	$1/c_{ob}$	c_{po} /B_a	$B_a$tanh($\alpha^B I$ / P^B_{max})
autotrophic respiration	−1	$−1/c_{ob}$		0.05 B_a P^B_{max} $\theta^{(T-20)}$
heterotrophic respiration	1			r_{20} $\theta^{(T-20)}$
decline of productivity			−1	k_P ($P^B_{max} - P^B_{max,base}$)

Table. 2. Stoichiometric constants in the metabolic model.

Name	Unit	Converts between
c_{ob}	mg O_2 L^{-1} (μg Chl L^{-1})$^{-1}$	oxygen production and biomass
c_{of}[†]	mg O_2 L^{-1} fluorescence^{-1}	oxygen production and fluorescence
c_{po}	d^{-1}	increase in productivity and photosynthesis

[†] c_{of} is used to convert prior estimates for P^B_{max} and α^B from their original fluorescence units to oxygen

Table. 3. Parameters of the metabolic model.

Symbol	Unit	Description
k_X[†]	m d^{-1}	O_2 exchange coefficient between water and atmosphere
z_{mix}[†]	m	mixing depth
DO_{sat}[†]	mg O_2 L^{-1}	saturating O_2 concentration
α^B	mg O_2 ([μmol m^{-2} s^{-1}] [μg Chl]) d)$^{-1}$	Biomass-specific light utilization efficiency
$P^B_{max,base}$	mg O_2 (μg Chl d)$^{-1}$	Base value for biomass-specific maximum rate of photosynthesis (P^B_{max})
θ	–	Temperature correction factor for respiration (set to 1.07)
r_{20}	mg O_2 L^{-1} d^{-1}	Heterotrophic respiration rate at 20°C
k_P	d^{-1}	first-order decline rate of biomass-specific maximum productivity (P^B_{max}) toward its base value

[†] calculated instead of calibrated

Methods

Model structure

We present the structure of the metabolic model in the matrix format proposed by Rauch et al. (1998; Table 1–3). The same model equations are presented in a full textual form in the Supplemental Material. The backbone of the model is the nonlinear photosynthesis–light relationship of Jassby and Platt (1976). Respiration is split into autotrophic and heterotrophic components (R_a and R_h, respectively), both of which are adjusted to water temperature using the van't Hoff–Arrhenius equation following Holtgrieve et al.

(2010) and Obrador et al. (2014). To link metabolism to autotrophic biomass (B_a) as closely as possible, we include feedback between metabolism and biomass. Photosynthesis and autotrophic respiration are both products of biomass and biomass-specific daily metabolic rates; they increase and decrease B_a, respectively (Table 1). Unlike primary production, daily heterotrophic respiration could not be made biomass-specific because of the lack of information on the biomass of major groups of heterotrophs.

Maximal rate of biomass-specific photosynthesis (P^B_{max}) is a dynamic state variable in our model to mimic the well-known diurnal changes in photosynthetic properties of algae (Henley 1993, Istvánovics et al. 2005, Honti and

Istvánovics 2011). The dynamic photosynthesis module allows algae to attain the daily maximum rate of photosynthesis gradually; P^B_{max} can increase from its base value ($P^B_{max,base}$) by photosynthesis and decay back to $P^B_{max,base}$ during darkness, proportionally to the distance from $P^B_{max,base}$.

We used established empirical formulae (Schmidt number: Wanninkhof 1992; piston velocity: Cole and Caraco 1998; exchange coefficient: Jähne et al. 1987) to estimate the physical parameters of gas exchange (Supplemental Material).

Calibration and uncertainty assessment

We applied a sequential Bayesian learning procedure to estimate parameters within a several months-long measurement season. We defined a model day to last from one sunrise to the next. Calibration progressed in units defined by a sliding window covering 3 model days. For each calibration unit, a Bayesian parameter inference and uncertainty analysis was performed. Structural errors of the metabolic model were described in a Bayesian manner using a first-order autoregressive error model for DO (Reichert and Schuwirth 2012). In parallel with fitting the model to DO measurements, we calibrated daily mean biomass of algae by incorporating both of these components into a single likelihood function.

Prior parameter distributions were obtained in 2 ways. For parameters that could directly be derived from high frequency data (daily mean $P^B_{max} \approx P^B_{max,base}$, daily mean α^B), measured daily mean values with a prescribed coefficient of variation of 20% were taken as priors in each calibration unit. For nonmeasured parameters (r_{20} and k_P), posterior distributions obtained in the preceding day represented prior distributions in the actual day. This procedure is the essence of sequential learning that can also be thought of as a prior expectation about gradually changing matter and energy fluxes at an ecosystem level. Sequential learning does not eliminate all sudden shifts in lake metabolism parameters, however; optimal posterior parameter values can be far from the high prior probability regions when data provide strong evidence about the inappropriateness of the prior expectations.

The posterior distribution of parameters was sampled for each calibration unit with Markov chain Monte Carlo sampling using the traditional Metropolis algorithm (Gamerman 1997). The covariance matrix of the jump distribution was tuned during the burn-in period so that the average acceptance rate was within the optimal range between 15 and 40% afterward (Gelman et al. 1996). The differential equations of the metabolic model were solved with the LSODA solver (Hindmarsh 1983, Petzold 1983), which automatically selected between stiff and nonstiff solution methods according to the behavior of the system.

To evaluate the importance of priors and sequential learning, we also performed the calibration and uncertainty assessment without parameter priors except for restricting some parameters (such as $r_{20} > 0$, $B_a > 0$) to the positive domain. This procedure was equivalent to a classical statistical calibration approach. In the absence of priors, calibration units (3 model days in a sliding window) were completely independent of each other.

Study sites

The model was tested on high frequency data from 3 lakes: Balaton (Hungary), Buresø (Lake Bure, Denmark), and Taihu (Lake Tai, China). These lakes are substantially different in their hydromorphological properties and trophic status (Table 4). All sites are members of the Global Lake Ecological Observatory Network (GLEON; gleon.org).

At Lake Balaton, depth varied between 1.4 and 1.7 m at the monitoring site. There, wind, global radiation, water temperature, DO, and turbidity were recorded every minute using a meteorology station (Mettech, Hungary), 2 LDO optical sensors (Hach, USA), 5 thermometers, and 5 light scattering sensors (WetLabs, USA). A delayed fluorescence spectroscope (TETT Ltd., Hungary) measured the biomass and the photosynthesis–irradiance relationship of phytoplankton every 20 min in the middle of the water column. Delayed fluorescence (DF) was converted to chlorophyll with a conversion factor derived from regressing weekly fluorometric chlorophyll data (Turner Designs TD-700, Canada; EPA Method 445.0) with the corresponding delayed fluorescence signal (Istvánovics et al. 2005; regression for 2013 was Chl [μg L^{-1}] = 1.06 DF – 1.221; $n = 25$; $R^2 = 0.86$). In deep Buresø, a winch moved a YSI 6600 v2 multisonde (Yellow Springs Instruments, USA) through the 9 m-deep water column twice per hour. The multisonde measured DO, temperature, and direct chlorophyll fluorescence. The buoy was also equipped with a self-cleaning rack of light sensors. In shallow Taihu, the buoy measured DO (D-Opto optical sensor, Zebra-tech, New Zealand), water temperature (TempHion T-2 sensor, Geotech, USA), turbidity, and direct chlorophyll fluorescence (Seapoint Sensors Inc., USA) every 10 min, 1 m above the sediment–water interface in Meiliang Bay, the most hypertrophic area of the lake. The Taihu station was also equipped with a WXT510 weather station (Vaisala, Finland).

In Balaton, water quality variables were averaged over the entire shallow water column, whereas in Buresø averages of the mixed layer were calculated. Mixing depth was determined by the Lake Analyzer (Read et al. 2011). In Balaton, 10 min averages of vertically averaged variables were taken (Table 1) because of the high computing requirement of calibrating the model with 1 min data. The metabolic model was fitted to the averaged data.

Table. 4. Lakes and data used in this study.

	Balaton	Buresø	Taihu
Lake shape and the location of the monitoring buoy/ station (grey dot)			
Area [km²]	596	0.761	2238
Mean depth [m]	3.2	6.7	1.9
Mixing regime	polymictic	seasonally stratified	polymictic
Catchment area [km²]	5181	2.6	36 900
WRT[†] [years]	5	11	<1
Altitude [m a.s.l.] [‡]	104	26	3
Latitude	46°	55°50′	31°
Trophic state	meso-eutrophic	mesotrophic	hyper-eutrophic
Dominant summer algae	*Ceratium hirundinella*, *Cylindrospermopsis raciborskii*, diatoms	*Ceratium* spp., *Chrysochromulina* spp., *Rhodomonas* spp., *Volvocales*	*Microcystis* spp.
Fetch at station [km]	0.1–25	0.2–0.9	0.1–55
Data timespan	23 Apr–23 Oct 2013	30 Mar–2 Dec 2012	18 May–4 July 2014
Number of modelled days	114	240	43
Algal parameters	B_a, P^B_{max}, α^B	B_a	B_a
Time resolution* [min]	10	30	10

* Time resolution of the DO series used for modeling; [†]WRT: water residence time; [‡]m.a.s.l.: metres above sea level

Results

The metabolism model frequently fit the DO data well with both the classical statistical and the Bayesian calibration approaches, albeit with important similarities and differences in the outcomes. The correlation coefficient between observed and simulated DO curves showed a considerable overlap using the 2 calibration methods. The classical approach produced the best possible fit, with a mean correlation coefficient of 0.88 (range 0.48–0.99); the values were 0.83 (0.24–0.99) for Bayesian parameter inference. Thus, fit to the observed DO time-series was usually not compromised noticeably by including prior knowledge. A further similarity was that both approaches benefited from restricting the domain of parameter values. Because of the parameter identification problems, the classical statistical procedure without any restrictions would have resulted in values obviously in conflict with the meaning attributed to the parameters (such as $r_{20} < 0$, $B_a < 0$) despite a good fit to the DO data.

These parameters were therefore constrained to the positive domain. Finally, the magnitude of GPP, NEP, and R did not vary much between the classical and Bayesian calibration (Fig. 2) because of a stronger connection of GPP, NEP, and R to DO data relative to the indirect connection of DO to single metabolic parameters. Data were sometimes poorly representative of the true, high phytoplankton biomass in Taihu because of the difficulties of sampling large, heterogeneously distributed *Microcystis* colonies. In this case, the model fit was often poor, and hence the correspondence between the 2 calibration approaches became weak for derived metabolic components.

The essential difference between the classical statistical and Bayesian approaches was observed in the internal coherence of metabolic parameters. The Bayesian procedure delivered gradually changing metabolic parameters (exemplified by the daily temperature-corrected rate of heterotrophic respiration in Fig. 3) and reasonably followed the observed dynamics of phytoplankton

biomass (Fig. 4). This finding suggested that the Bayesian solution was more credible from an ecological perspective and therefore might yield more robust predictions under changing boundary conditions.

Despite the generally good model fit, there were cases in each lake when the structure of the applied metabolic model was obviously inappropriate to describe the observed data (Fig. 5). Our model was defined to simulate DO dynamics in a closed, fully mixed mass of water, and it was unable to treat cases when non-simulated processes significantly influenced DO concentrations.

As expected, a strong linear correlation existed between daily GPP and R (Fig. 6) in our lakes. The intercept constant of the linear models roughly equaled the mean annual value of r_{20} in each lake. NEP was nearly zero over the measuring period in both Balaton

and Buresø and decreased steadily in Taihu because of the high heterotrophic respiration. In Balaton, the ratio of heterotrophic to autotrophic respiration fluctuated ~50–60% during the first half of the measurement period, followed by a gradual increase after the collapse of the late summer phytoplankton bloom (Fig. 7). Both the decreasing autotrophic respiration and enhanced heterotrophic respiration contributed to this increase. Although R_h/R_a ratios were as low in Taihu as in Balaton, they never dropped below 100% in Buresø despite a steady decrease throughout the season. In Buresø, heterotrophic respiration decreased sharply after the autumn overturn.
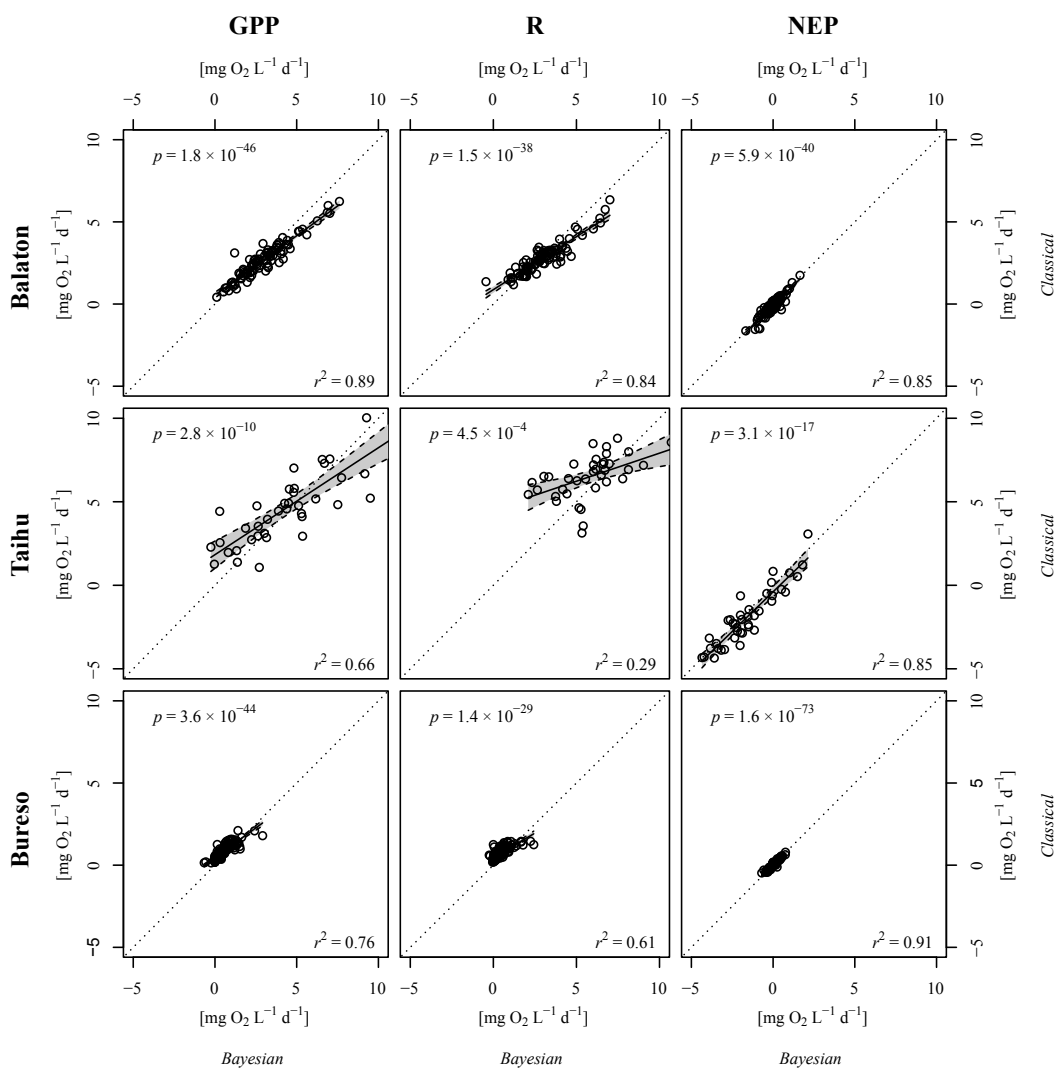

Fig. 2. Daily rates of total GPP, NEP, and R in lakes Balaton, Taihu, and Buresø with the classical statistical model fitting procedure (Classical) vs. the proposed Bayesian approach using priors and sequential learning (Bayesian). Open circles: data pairs; line: linear regression on data pairs; shading: 95% confidence interval of regression.
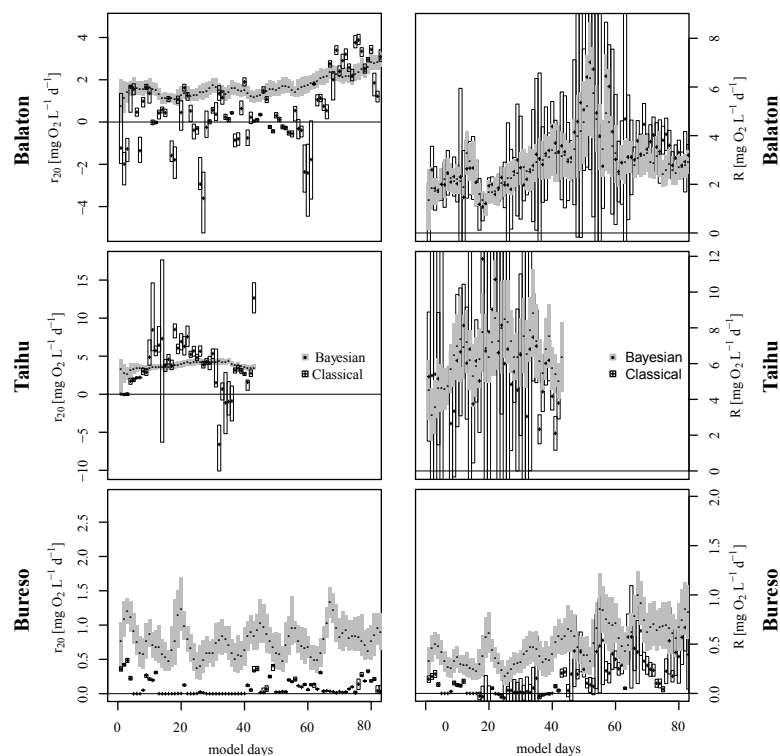
Fig. 3. Daily posterior means and 95% confidence intervals of heterotrophic respiration rate (r_{20}, left panels) and total daily respiration (R, right panels) with the classical statistical model fitting procedure (Classical - mean: cross; confidence interval: open box) vs. the proposed Bayesian approach using both priors and sequential learning (Bayesian - mean: dot; confidence interval: gray box).
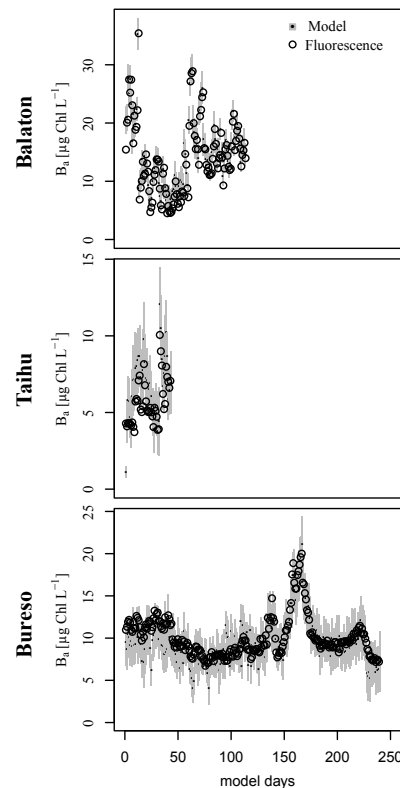
Fig. 4. Modeled posterior distributions of daily mean autotrophic biomass vs. observed chlorophyll fluorescence used to set up prior distributions for B_a in lakes Balaton, Taihu, and Buresø.
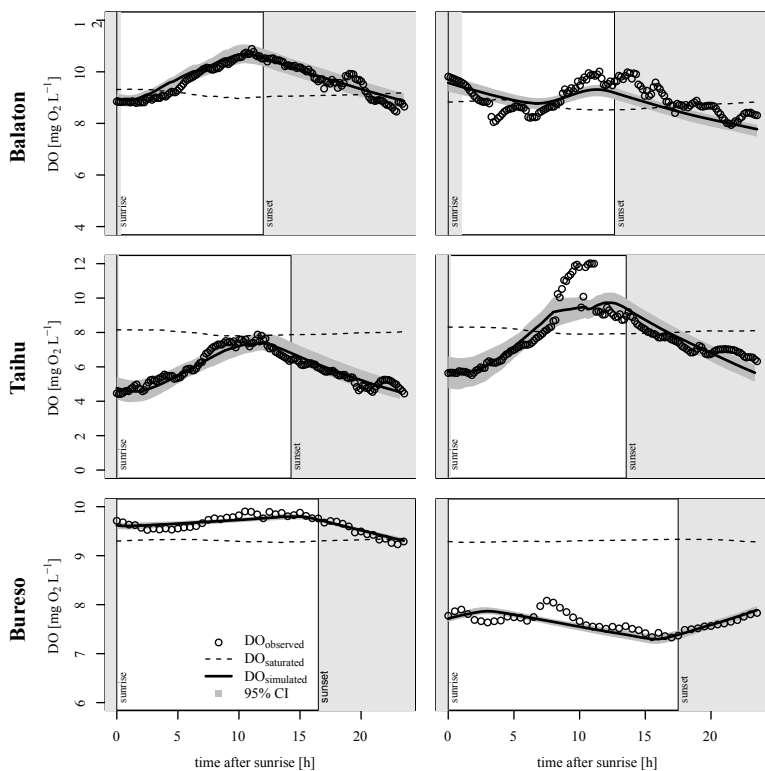
Fig. 5. Examples of days with good (left) and poor (right) model fit for lakes Balaton, Taihu, and Buresø.

Discussion

Goodness of fit

In general, our metabolic model could reasonably be fitted to the observed DO series (Fig. 5). Bayesian parameter inference decreased the goodness of fit slightly compared to the best possible fit of our model to the data obtained by the classical statistical calibration. Despite the slightly worse fit to the data, parameters estimated in the Bayesian approach displayed features more likely to be ecologically realistic. For example, the slow and consistent change in parameter values over the course of several days is likely to be more realistic than the sudden and large changes observed using a classical approach (e.g., Solomon et al. 2013). Therefore, despite the tradeoffs in model fit, using a Bayesian approach represents a means to derive an ecologically more coherent set of metabolic parameters (Fig. 3 and 4).

Because our model was defined to simulate DO dynamics in a closed, fully mixed mass of water, model fit was poor on certain days (Fig. 5). Considering that transport may significantly influence metabolism estimates in some systems (Antenucci et al. 2013), plausible explanations could be temporary stratification of shallow water columns in Balaton and Taihu, mixing with DO-poor hypolimnetic water in Buresø, and lateral transport of water masses containing different concentrations of DO in each of the 3 lakes. Our recent online velocity measurements at the monitoring site in Balaton (Honti and Istvánovics, unpubl. data), however, cast doubt on advection having a predominant role in modifying DO series measured at a single site. Horizontal flow velocities are typically in the range of a few mm s^{-1} and increase to 1.5–5 cm s^{-1} only during large storms. These values indicate that the radius of the area from where water can travel to our site in a day is typically from 200 m to a maximum of 1–4 km. Low wind speeds promote spatial differentiation of water masses and biota, but the action radius of advection seems to be too low to transport water with substantially different DO content to the monitoring site. By contrast, high winds efficiently homogenize water masses and planktonic communities.

The empirical functions used to describe atmospheric gas transfer might also represent a major source of model structural error and large systematic deviations of modeled DO from measurements, just by the sheer volume of gas exchange in shallow waters. The present consensus seems to be that oxygen exchange rates obtained with various empirical functions were unequivocally too low to signifi-
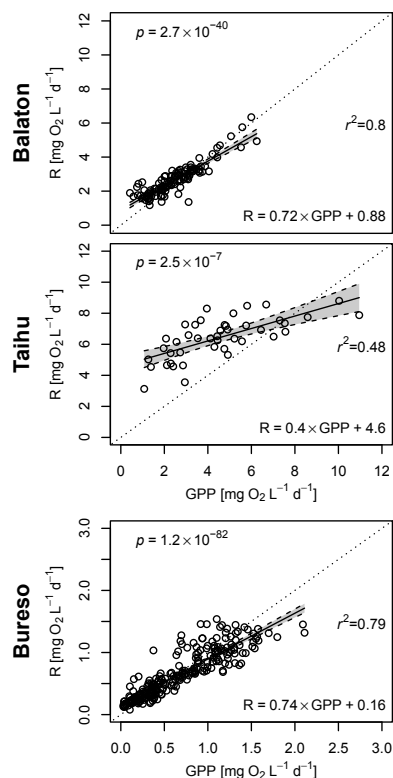
Fig. 6. Modeled daily total R as a function of daily total GPP in lakes Balaton, Taihu, and Buresø. Circles: data pairs; line: linear regression on data pairs; shading: 95% confidence interval of regression.
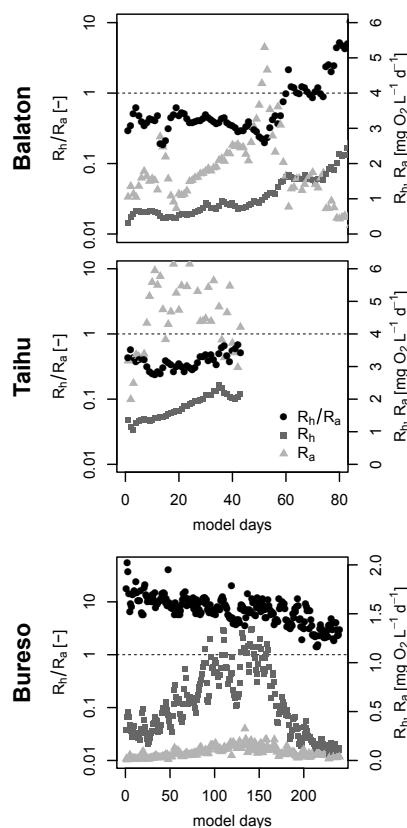
Fig. 7. Time-series of R_a, R_h, and their ratio in lakes Balaton, Taihu, and Buresø.

cantly influence simulated GPP and R (Antenucci et al. 2013, Rose et al. 2014). Other studies, however, noted distinct phase transitions in gas exchange with increasing wave height, and breaking waves may result in sudden multiplication of gas transfer rate in large lakes (Livingstone and Imboden 1993).

Linking metabolism to biomass

Although NEP should be generally related to the changes in autotrophic biomass, most metabolic models do not use autotrophic or heterotrophic biomass and instead employ abstract parameters of daily photosynthesis and community respiration rates. Reluctance to include biomass into metabolic models is a pragmatic modeling solution to the problem that changes in biomass may be decoupled from the actual balance of photosynthesis and respiration at the timescale of days. Abrupt changes in algal biomass may be triggered by factors outside the scope of metabolic models, such as the sudden collapse of an algal bloom due to viral infection (Bratbak et al. 1993), a large storm that causes damage to filamentous cyanobacteria (Padisák et al. 1990), termination of a diatom bloom by silica limitation, and/or onset of stable thermal stratification in late spring/early summer (Huber et al. 2008). Under these types of transitions, metabolic models cannot provide a precise estimate of metabolizing biomass.

A novelty of our approach is the decomposition of photosynthesis and autotrophic respiration into products of a changing algal biomass and biomass-specific rates such that ecological coherence of calibrated parameters could be tested. Although this decomposition makes our metabolic model more complicated compared to some published models (e.g., Hanson et al. 2008, Solomon et al. 2013), it is a prerequisite to increase the predictive power of modeling. When GPP, NEP, and R are adjustable model parameters, metabolic fluxes will be system- and case-specific. They cannot be used for making predictions under different boundary conditions (changing boundary conditions being the essence of modeling carbon cycling of lakes under a changing climate), metabolic shifts upon successful biotic invasions, or altered nutrient loads. Although these fluxes tended to be similar using the classical statistical and the Bayesian calibration (Fig. 2), a simple model fitting procedure would not suffice for predictive purposes. Biomass-specific metabolic rates are more general than fluxes; they are more likely to be applicable under different boundary conditions.

It seems unlikely that biomass-specific rates could be estimated in an ecologically coherent way from DO time series alone. Therefore, the decomposition is achieved by using high frequency chlorophyll fluorescence data in the calibration of a metabolic model, which has a dynamic biomass compartment directly linked to metabolism. Linking the metabolic model to high frequency fluorescence data required several technical improvements. First, it was necessary to establish feedback between metabolism and biomass. GPP increased with B_a and decreased with R_a (Table 1–3; Supplemental Material), complementing the wide-open material balance of common metabolic models. Second, linking the metabolic model to high frequency chlorophyll fluorescence required a specialized solution. Fluorescence is a weak proxy for active algal biomass because both direct and delayed fluorescence depend on diurnally changing photosynthetic properties of cells in addition to biomass (Falkowski and Raven 1997, Istvánovics et al. 2005). Therefore, algal biomass was estimated from daily mean fluorescence data, and 80% of the fluorescence-based estimate of B_a was applied as prior distribution of the initial biomass in the beginning of each calibration unit. Third, biomass-specific parameters allowed us to tune GPP and R_a in accordance with the intra-day changes of biomass; however, simulating diurnal biomass changes in single-day units caused simulated B_a to shift dramatically on the boundaries between subsequent model days. These shifts were most probably caused by the strong internal memory of the metabolic model, due to which calibration did not try to eliminate errors in the final few hours of the calibration window as opposed to initial or midday errors. Because errors in simulated DO typically had a few hours collateral forward effect, errors committed in the terminal stage were statistically preferable. To avoid jumps of simulated B_a between calibration windows, we used a 3-day moving window to estimate parameters assigned to the middle day, which helped filter out parameter combinations that resulted in a close fit during the middle day but performed poorly in the neighboring days.

Multiday calibration windows seem to be inevitable for ensuring the continuity of B_a, but the span of such windows is strongly limited by the nonpredictability of biomass dynamics on the longer run. Notably, most metabolic models estimate model parameters for short periods (1–4 days; Hanson et al. 2008, Staehr et al. 2012, Solomon et al. 2013, Obrador et al. 2014), indicating that the properties of biotic communities may evolve over time, and assuming parameter stability for longer intervals would be incorrect. Splitting the measurement period in parts, however, disentangles links between neighboring calibration units, and thus no kind of connection is guaranteed between the parameter values of subsequent calibration windows. The moving window and sequential learning applied in our study solved this discrepancy.

Because total heterotrophic biomass was not measured in any of our test systems, the daily rate of heterotrophic respiration could not be decomposed into biomass and a biomass-specific rate. The sequential Bayesian learning procedure resulted in a slow gradual change in this parameter and a less smooth but still gradual change in total respiration (Fig. 3), in sharp contrast with the results of other studies that demonstrated substantial day-to-day variation in respiration (Cole et al. 2000, Staehr and Sand-Jensen 2007, Coloso et al. 2011). Solomon et al. (2013) fitted a simple metabolic model to high frequency DO series and used bootstrap analysis to estimate uncertainty of GPP and R. They concluded that a significant part of daily variability in R was due to real ecological variability with somewhat higher impact of model uncertainty in spatially more heterogeneous large than in small lakes; however, neither Solomon et al. (2013) nor other studies discussed therein could provide a plausible and generally valid explanation for large daily fluctuations in R. As we demonstrated, this variability was most probably dependent on the technique of model calibration (Fig. 3). We argue that confined day-to-day variability in both r_{20} and R is ecologically more credible than large fluctuations. Smooth seasonal variation reflects resilience of the ecosystem, with sudden shifts in lake metabolism being rare events (Batt et al. 2013). Moreover, the application of sequential learning does not exclude the occurrence of abrupt changes in the time-series of calibrated parameters. When measured data provide strong evidence about the inappropriateness of the prior parameters for a given calibration unit, the likelihood term will dominate the posterior probability, and significant departure from the prior distributions can take place. In the absence of such strong indication, however, gradual adjustments are preferred in line with the characteristics of a resilient system. We found only a single case (after model day 90 in Buresø; Fig. 3) when the continuity of r_{20} had to be interrupted to maintain model fit.

Evaluation of the error model

Formal statistical calibration methods require that the statistical assumptions about model residuals (in our case the differences between modeled and observed DO) are validated. If the assumptions are violated, simulation results should be rejected because the model parameters may be heavily biased and model uncertainty may be significantly overestimated or underestimated. Because each metabolic model is an extremely simplified representation of the ecosystem, structural uncertainty of the model would obviously dominate measurement uncertainty of DO. The first-order autoregressive error model (Van de Bogert et al. 2007, Hanson et al. 2008)

could be considered a proper Bayesian representation of structural errors (Bayarri et al. 2007, Reichert and Schuwirth 2012) and was therefore used during calibration. The calibrated error parameters underpinned our assumption about the strong autocorrelation of errors. The correlation half-life of errors, that is the time required to reduce the correlation to 0.5, ranged between 1 and 9 h, suggesting that studies using a noncorrelated error model underestimated the uncertainty of both predictions and parameters and therefore produced statistically unacceptable results.

Interpretation of model parameters

Although Bayesian calibration and uncertainty analysis helps to deal with models that have more degrees of freedom than the actual information content of the dataset used for calibration, it is important to make a clear distinction between the uncertainties of the modeled time series and those of the parameters obtained during such a calibration or uncertainty analysis.

Bayesian parameter inference guarantees proper uncertainty intervals for the predicted time series (in this case DO) when statistical assumptions behind the applied error model are not violated by model residuals (Mantovan and Todini 2006). The properness of these uncertainty intervals is meant in a statistical sense (e.g., the number of outliers, correlations between realizations). The autoregressive error model ensured that the essential statistical properties of the model residuals were properly reflected in the error model, and therefore the estimated uncertainty intervals of DO, and most likely those of underlying processes (GPP, R), could be considered robust.

Model parameters, however, are not in the same quality category as uncertainty intervals of modeled time series. In a mathematical sense, the only task of parameters is to tune the model's outcome to match observations as closely as possible; their attributed meaning does not play any role during calibration (Mantovan and Todini 2006). Parameters will become inevitably biased to a certain degree because they compensate for deficiencies of the model structure, for errors in input data, and for errors in the observed DO time series. As noted earlier, in our case, structural model uncertainty was the dominant source of error. Examples of this included boosting respiration rates and slowing down production to describe the mixing of DO-depleted hypolimnetic water into the epilimnion in Buresø and adjusting production rates to follow oscillations caused by temporary stratification and mixing events in Balaton. As a consequence, calibrated parameters partly or fully lose their original meaning as physical/chemical/ecological characteristics (Doherty and Christensen 2011), a drawback that applies to any

calibration procedure, including non-Bayesian approaches. In addition to this dependence on data, model structure, and the error model, posterior parameter values are also conditional on priors, the essential subjective component of Bayesian methods. Improperly defined priors (for example based on nonrepresentative data, as in Taihu) may prevent statistically optimal model fits and hence introduce a strong bias in posterior parameters. These issues prevent the in-depth interpretation of calibrated parameter values and also require serious precaution if a model is used outside the boundary conditions for which it was calibrated. Calibrated metabolic rates can best be used for cross-system comparison or regional modeling when calibration was successful in terms of a good fit to observations and no obvious structural deficiencies could compromise the theoretical meaning of the calibrated parameters.

Comparison of metabolism in the test lakes

Daily R was tightly coupled to daily GPP in mesotrophic Balaton and Buresø, whereas the coupling was weak in hypertrophic Taihu (Fig. 6), a pattern in agreement with that observed in a wide selection of lakes (Solomon et al. 2013). The intercept of linear relationships of GPP versus R roughly equaled mean heterotrophic respiration. R was an order of magnitude higher in Taihu than in our mesotrophic lakes. Mean GPP that compensated for mean total R and ensured an NEP of zero increased with the trophic status of the 3 lakes, with O_2 values of 0.6 mg L^{-1} d^{-1} in Buresø, 3.1 mg L^{-1} d^{-1} in Balaton, and 7.7 mg L^{-1} d^{-1} in Taihu. Thus, likely because it receives substantial amounts of allochthonous organic matter from both agriculture and aquaculture (Jinglu et al. 2007, Zeng et al. 2007), Taihu was frequently net heterotrophic, whereas Buresø was nearly always autotrophic. The slope of the GPP versus R relation was somewhat lower in both Balaton and Taihu (0.72 and 0.40, respectively) than that obtained by Solomon et al. (2013) in these 2 lakes (0.85 and 0.55, respectively).

NEP cumulated over the season was small compared to the magnitude of seasonal GPP and R in each lake. This difference explains the well-known difficulties estimating GPP and R and suggests that the sign of the overall seasonal metabolic balance of a lake may easily change under slightly different boundary conditions. Of the 3 lakes, only Buresø closed the measurement period with a slightly autotrophic overall balance. Balaton was slightly heterotrophic at the end, probably because the measurements did not cover the highly productive early spring period. The similar magnitude of GPP and R does not apply for Taihu, where R surpassed GPP during the entire period covered by data, and heterotrophy at the measurement site was obvious.

The ratio of heterotrophic to autotrophic respiration was surprisingly different in our lakes (Fig. 7). In Buresø, R_h significantly exceeded R_a during the stratified period, and R_h/R_a ratios approached 1 only after complete autumn mixing. By contrast, the R_h/R_a ratios were low in the 2 shallow lakes. This pattern did not match the trophic character of the test lakes; both Balaton and Buresø were nearly balanced in terms of autotrophy and heterotrophy, whereas Taihu was net heterotrophic. We propose that R_h/R_a ratios might reflect overall ecological efficiency in lakes. Low ratios in Balaton were in line with the low efficiency of energy transfer from primary producers to fish (0.04–0.1%; Bíró and Vörös 1990). Enhanced heterotrophic respiration after the collapse of the summer phytoplankton bloom (Fig. 7) might be attributed to rapid growth of chironomid larvae that reach their biomass maxima in late autumn–early spring (Specziár and Vörös 2001). Although the limnology of Buresø is poorly known, its trophic status and dominant algal species (Table 4) make a high overall ecological efficiency likely. In Taihu, low R_h/R_a ratios might arise from a particularly low energy transfer efficiency between trophic levels (Li et al. 2009) caused by *Microcystis* dominance.

Conclusions

Because of the complexity of metabolic models and the associated parameter identification problems, we recommend a Bayesian approach for the calibration of these models so that additional knowledge about parameters can be incorporated. Expressing metabolic rates in a biomass-specific form may enhance the credibility of model parameters. Linking metabolism and biomass increases the ecological usefulness of these models. Although decomposition into biomass-specific rates and biomass further increases model complexity and worsens the parameter identification problem, the potential gain in ecological coherence compensates for these adverse effects.

Ecologically coherent, and therefore presumably robust, estimates for biomass-specific metabolic rates will increase the predictive power of metabolic models under different boundary conditions. Ecological knowledge suggests that abrupt shifts in biomass and metabolic rates occur rarely. To ensure the general time-continuity of these model parameters, we propose a sequential Bayesian parameter learning procedure.

The sometimes obvious structural deficiencies of our model suggest that an enhanced parameter estimation technique alone is not sufficient to derive generally valid metabolic rates from high frequency DO data. There is still substantial potential in measuring underlying processes of DO metabolism and improving metabolic models with these observations.

Acknowledgements

References

Antenucci JP, Tan KM, Eikaas HS, Imberger J. 2013. The importance of transport processes and spatial gradients on *in situ* estimates of lake metabolism. Hydrobiologia. 700:9–21.

Batt RD, Carpenter SR, Cole JJ, Pace ML, Johnson RA. 2013. Changes in ecosystem resilience detected in automated measures of ecosystem metabolism during a whole-lake manipulation. P Natl Acad Sci USA. 110:17298–17403.

Bayarri MJ, Berger JO, Paulo R, Sacks J, Cafeo JA, Cavendish J, Lin CH, Tu J. 2007. A framework for validation of computer models. Technometrics. 49:138–154.

Bíró P, Vörös L. 1990. Trophic relationships between primary producers and fish yields in Lake Balaton. Hydrobiologia. 191:213–221.

Bratbak GE, Egge JK, Heldal M. 1993. Viral mortality of the marine alga *Emiliania huxleyi* (Haptophyceae) and termination of algal bloom. Mar Ecol Prog Ser. 93:39–48.

Cole JJ, Bade DL, Bastviken D, Pace ML, van de Bogert M. 2010. Multiple approaches to estimating air-water gas exchange in small lakes. Limnol Oceanogr Meth. 8:285–293.

Cole JJ, Caraco NF. 1998. Atmospheric exchange of carbon dioxide in a low-wind oligotrophic lake measured by the addition of SF6. Limnol Oceanogr. 43:647–656.

Cole JJ, Pace ML, Carpenter SL, Kitchell JF. 2000. Persistence of net heterotrophy in lakes during nutrient addition and food web manipulations. Limnol-Oceanogr. 45:1718–1730.

Cole JJ, Prairie YT, Caraco NF, McDowell WH, Tranvik LJ, Striegl RG, Duarte CM, Kortelainen P, Downing JA, Middelburg JJ, Melack J. 2007. Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. Ecosystems. 10:171–184.

Coloso JJ, Cole JJ, Pace ML. 2011. Short-term variation in thermal stratification complicates estimation of lake metabolism. Aquat Sci. 73:305–315.

Crusius J, Wanninkhof R. 2003. Gas transfer velocities measured at low wind speed over a lake. Limnol Oceanogr. 48:1010–1017.

Doherty J, Christensen S. 2011. Use of paired simple and complex models to reduce predictive bias and quantify uncertainty. Water Resour Res. 47:W12534. doi:10.1029/2011WR010763

Einsele G, Jianping Y, Hinderer M. 2001. Atmospharic carbon burial in modern lake basins and its significance for the global carbon budget. Glob Planet Change. 30:167–195.

Falkowski PG, Raven JA. 1997. Aquatic photosynthesis. Malden (MA): Blackwell Science.

Gamerman D. 1997. Markov Chain Monte Carlo - Stochastic simulation for Bayesian inference. London (UK): Chapman and Hall.

García-Camacho F, Sánchez-Mirón A, Molina-Grima E, Camacho-Rubio F, Merchuck JC. 2012. A mechanistic model of photosynthesis in microalgae including photoacclimation dynamics. J Theor Biol. 304:1–15.

Gelman A, Roberts GO, Gilks WR. 1996. Efficient Metropolis jumping rules. In: Bernardo JM, Berger JO, Dawid AP, Smith FM, editors. Bayesian statistics. Oxford University Press. 5:599–607.

Hanson PC, Bade DL, Carpenter SR, Kratz TK. 2003. Lake metabolism: relationships with dissolved organic carbon and phosphorus. Limnol Oceanogr. 48:1112–1119.

Hanson PC, Carpenter SR, Kimura N, Wu C, Cornelius SP, Kratz TK. 2008. Evaluation of metabolism models for free-water dissolved oxygen methods in lakes. Limnol Oceanogr-Meth. 6:454–465.

Henley WJ. 1993. Measurement and interpretation of photosynthetic light-response curves in algae in the context of photoinhibition and diel changes. J Phycol. 29:729–739.

Hindmarsh AC. 1983. ODEPACK, a systematized collection of ODE solvers. IMACS T Sci Comput. 1:55–64.

Holtgrieve GW, Schindler DE, Branch TA, A'Mar ZT. 2010. Simultaneous quantification of aquatic ecosystem metabolism and reaeration using a Bayesian statistical model of oxygen dynamics. Limnol Oceanogr. 55:1047–1063.

Honti M, Istvánovics V. 2011. Online monitoring of phytoplankton light response curves using a novel delayed fluorescence device. Lake Reserv Res Manage. 16:153–158.

Huber V, Adrian R, Gerten D. 2008. Phytoplankton response to climate warming modified by trophic state. Limnol Oceanogr. 53:1–13.

Istvánovics V, Honti M, Osztoics A, Shafik HM, Padisák J, Yacobi Y, Eckert W. 2005. Continuous monitoring of phytoplankton dynamics in Lake Balaton (Hungary) using on-line delayed fluorescence excitation spectroscopy. Freshwater Biol. 50:1950–1970.

Jähne B, Münnich KO, Bösinger R, Dutzi A, Huber W, Libner P. 1987. On the parameters influencing air-water gas exchange. J Geophys Res-Oceans. 92(C2):1937–1949.

Jassby AD, Platt T. 1976. Mathematical formulation of the relationship between photosynthesis and light for phytoplankton. Limnol Oceanogr. 21:540–547.

Jinglu W, Chengmin H, Haiao Z, Schleser GH, Battarbee R. 2007. Sedimentary evidence for recent eutrophication in the northern basin of Lake Taihu, China: human impacts on a large shallow lake. J Paleolimnol. 38:13–23.

Li YK, Chen Y, Song B, Olson D, Yu N, Chen LQ. 2009. Ecosystem structure and functioning of Lake Taihu (China) and the impacts of fishing. Fish Res. 95:309–324.

Livingstone D, Imboden DM. 1993. The non-linear influence of wind-speed variability on gas transfer in lakes. Tellus. 45:275–295.

MacIntyre S, Jonsson A, Jansson M, Aberg J, Turney DE, Miller SD. 2010. Buoyancy flux, turbulence, and the gas transfer coefficient in a stratified lake. Geophys Res Lett. 37:L24604.

Mantovan P, Todini E. 2006. Hydrological forecasting uncertainty assessment: incoherence of the GLUE methodology. J Hydrol. 330:368–381.

Markager S, Jespersen A-M, Madsen T, Berdalet E, Weisburd R. 1992. Diel changes in dark respiration in a plankton community. Hydrobiologia. 238:119–130.

Obrador B, Staehr PA, Christensen JPC. 2014. Vertical patterns of metabolism in three contrasting stratified lakes. Limnol Oceanogr. 59:1228–1240.

Odum HT. 1956. Primary production in flowing waters. Limnol Oceanogr. 1:102–117.

Odum HT, Odum EP. 1955. Trophic structure and productivity of a windward coral reef community on Eniwetok Atoll. Ecol Monogr. 25:291–320.

Padisák J, G.-Tóth L, Rajczy M. 1990. Stir-up effect of wind on a more-or-less stratified shallow lake phytoplankton community, Lake Balaton, Hungary. Hydrobiologia. 191:249–254.

Petzold L. 1983. Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations. SIAM J Sci Stat Computing. 4:136–148.

Rauch W, Henze M, Koncsos L, Reichert P, Shanahan P, Somlyódy L, Vanrolleghem P. 1998. River water quality modelling: I. State of the art. Water Sci Technol. 38:237–244.

Read JS, Hamilton DP, Jones ID, Muraoka K, Winslow LA, Kroiss R, Wu CH, Gaiser E. 2011. Derivation of lake mixing and stratification indices from high-resolution lake buoy data. Environ Model Softw. 26:1325–1336.

Reichert P, Schuwirth N. 2012. Linking statistical bias description to multiobjective model calibration. Water Resour Res. 48:W09543.

Rose KC, Winslow LA, Read JS, Read EK, Solomon CT, Adrian R, Hanson PC. 2014. Improving the precision of lake ecosystem metabolism estimates by identifying predictors of model uncertainty. Limnol Oceanogr-Meth. 12:303–312.

Sadro S, Holtgrieve GW, Solomon CT, Koch GR. 2014. Widespread variability in overnight patterns of ecosystem respiration linked to gradients in dissolved organic matter, residence time, and productivity in a global set of lakes. Limnol Oceanogr. 59:1666–1678.

Solomon CT, Bruesewitz DA, Richardson DC, Rose KC, Van de Bogert MC, Hanson PC, Kratz TK, Larget B, Adrian R, Babin BL, et al. 2013. Ecosystem respiration: drivers of daily variability and background respiration in lakes around the globe. Limnol Oceanogr. 58:849–866.

Specziár A, Vörös L. 2001. Long term dynamics of Lake Balaton's chironomid fauna and its dependence on the phytoplankton production. Arch Hydrobiol. 152:119–142.

Staehr PA, Bade D, Van de Bogert MC, Koch GR, Williamson C, Hanson P, Cole JJ, Kratz T. 2010. Lake metabolism and the diel oxygen technique: state of the science. Limnol Oceanogr-Meth. 8:628–644.

Staehr PA, Christensen JPA, Batt RD, Read JS. 2012. Ecosystem metabolism in a stratified lake. Limnol Oceanogr. 57:1317–1330.

Staehr PA, Sand-Jensen K. 2007. Temporal dynamics and regulation of lake metabolism. Limnol Oceanogr. 52:108–120.

Tranvik LJ, Downing JA, Cotner JB, Loiselle SA, Striegl RG, Ballatore TJ, Dillon P, Finlay K, Fortino K, Knoll LB, et al. 2009. Lakes and reservoirs as regulators of carbon cycling and climate. Limnol Oceanogr. 54:2298–2314.

Van de Bogert MC, Carpenter SR, Cole JJ, Pace ML. 2007. Assessing pelagic and benthic metabolism using free water measurements. Limnol Oceanogr-Meth. 5:145–155.

Wanninkhof R. 1992. Relationship between wind speed and gas exchange over the ocean. J Geophys Res: Oceans. 97(C5):7373–7382.

Zeng QF, Kong FX, Zhang EL, Tan X. 2007. Effects of anthropogenic organic matter inputs on stable carbon and nitrogen isotopes in organisms from microbial food chain in Taihu Lake. Huanjing Kexue 28:1670–1674. Chinese.

Supplementary information

A detailed description of the metabolic model is available for download via the Inland Waters website, https://www.fba.org.uk/journals/index.php/IW/article/viewFile/877/638