LIMNOLOGY AND OCEANOGRAPHY Methods
ASLO Association for the Sciences of Limnology and Oceanography

Fitting metabolic models to dissolved oxygen data: The Estuarine BAyesian Single-station Estimation method (EBASE)

Journal:	*Limnology and Oceanography: Methods*
Manuscript ID	LOM-23-11-0095
Wiley - Manuscript type:	New Methods
Date Submitted by the Author:	20-Nov-2023
Complete List of Authors:	Beck, Marcus; Tampa Bay Estuary Program Arriola, Jill; The Pennsylvania State University Herrmann, Maria; The Pennsylvania State University Najjar, Raymond ; The Pennsylvania State University
Keywords:	Bayesian, dissolved oxygen, ecosystem metabolism, estuary, gas exchange, open-source

Abstract:	Continuous measurements of dissolved oxygen (DO) are useful for quantifying ecosystem metabolism, which is critical for understanding estuarine biogeochemistry and ecology, but current methods applied to these data may lead to unphysical results and poorly constrained errors. Here we present a new approach for estimating estuarine metabolism: EBASE (Estuarine BAyesian Single-station Estimation). EBASE applies a Bayesian framework to a simple process-based model and DO observations, allowing the estimation of critical model parameters, specifically light efficiency and respiration, as informed by a set of prior distributions. EBASE improves upon the stream-based model from which it was derived by accommodating missing DO data and allowing the user to set the time period over which parameters are estimated. We demonstrate that EBASE can recover known metabolic parameters from a synthetic time series, even in the presence of noise (e.g., due to tidal advection) and when prior distributions are uninformed. Optimization periods of 7 days and 30 days are more preferable than 1 day. A comparison with the more-conventional method of Odum reveals the ability of EBASE to avoid unphysical results (such as negative photosynthesis and respiration) and improves when the DO data are detided. EBASE is available using open-source software and can be readily applied to multiple years of long-term monitoring data that are available in many estuaries. Overall, EBASE provides an accessible method to parameterize a simple metabolic model appropriate for estuarine systems and will provide additional understanding of processes that influence ecosystem status and condition.

Fitting metabolic models to dissolved oxygen data: The Estuarine BAyesian Single-station Estimation method (EBASE)

Marcus W Beck[1,✉], Jill M Arriola[2], Maria Herrmann[2], and Raymond G Najjar[2]

[1] Tampa Bay Estuary Program, St. Petersburg, Florida 33701 USA

[2] The Pennsylvania State University, University Park, Pennsylvania 16802 USA

✉ Correspondence: Marcus W Beck <mbeck@tbep.org>

Running head: EBASE metabolic model

Key words: Bayesian, dissolved oxygen, ecosystem metabolism, estuary, gas exchange, open-source

Abstract

Continuous measurements of dissolved oxygen (DO) are useful for quantifying ecosystem metabolism, which is critical for understanding estuarine biogeochemistry and ecology, but current methods applied to these data may lead to unphysical results and poorly constrained errors. Here we present a new approach for estimating estuarine metabolism: EBASE (Estuarine BAyesian Single-station Estimation). EBASE applies a Bayesian framework to a simple process-based model and DO observations, allowing the estimation of critical model parameters, specifically light efficiency and respiration, as informed by a set of prior distributions. EBASE improves upon the stream-based model from which it was derived by accommodating missing DO data and allowing the user to set the time period over which parameters are estimated. We demonstrate that EBASE can recover known metabolic parameters from a synthetic time series, even in the presence of noise (e.g., due to tidal advection) and when prior distributions are

uninformed. Optimization periods of 7 days and 30 days are more preferable than 1 day. A comparison with the more-conventional method of Odum reveals the ability of EBASE to avoid unphysical results (such as negative photosynthesis and respiration) and improves when the DO data are detided. EBASE is available using open-source software and can be readily applied to multiple years of long-term monitoring data that are available in many estuaries. Overall, EBASE provides an accessible method to parameterize a simple metabolic model appropriate for estuarine systems and will provide additional understanding of processes that influence ecosystem status and condition.

Introduction

Estuaries play a key role in the global cycling of elements because of high rates of biogeochemical activity at the interface between land, ocean, and atmosphere. As materials are transported from land to ocean, estuaries profoundly transform or filter these materials through various processes, chief among which are primary production and respiration, collectively referred to as metabolism (Schubel and Kennedy 1984). For example, estuarine respiration of organic carbon delivered from land produces carbon dioxide, which may be lost to the atmosphere via outgassing (e.g, Cai 2011). Similarly, estuarine primary production consumes dissolved inorganic carbon delivered from land, forming particulate organic carbon that may be buried in estuarine sediments (e.g., Hu et al. 2006). Estuarine metabolic processes are fundamentally important, as illustrated by the dissolved inorganic carbon budget of eastern North American estuaries (Najjar et al. 2018); net primary production and heterotrophic respiration far exceed inputs from upland sources and outputs to the atmosphere and ocean. As such, small changes in either net primary production or heterotrophic respiration can have a large impact on the difference between the two, net ecosystem metabolism (NEM, sometimes also referred to as

net ecosystem production), which ultimately governs the form and quantity of estuarine materials that are lost to the atmosphere, the sediments, and the ocean.

Methods for estimating metabolism in aquatic environments are numerous and include open-water techniques, bottle-based incubations, ecosystem budgets, oxygen isotopes, and use of inert gases (Kemp and Testa 2011; Staehr et al. 2011). Based on the variety of assumptions and limitations of each method, comparisons of metabolic rates within and between ecosystems have been challenging. For example, an assessment in Randers Fjord (Denmark) showed agreement in both the sign and magnitude among four different approaches (Gazeau et al. 2005). However, gross primary production and ecosystem respiration estimates from incubations were consistently lower than those from open-water methods, consistent with the estuarine respiration synthesis of Hopkinson and Smith (2005). Even within one technique, there are numerous choices to be made that alter the outcome. For example, in the commonly used ^{14}C technique for measuring primary production, Cloern et al. (2014) noted that investigators have made divergent choices with regard to filtering mesozooplankton grazers, measuring ^{14}C in dissolved organic carbon, changing spectral light quality with depth, the euphotic zone depth, the incubation time, and the incubation location (*in situ* vs. on-deck). These challenges are further compounded in estuaries, where high temporal and spatial variability have been noted in global syntheses of photosynthesis and respiration. Cloern et al. (2014) found that annual phytoplankton primary production can vary 10-fold within an estuary and up to 5-fold from year to year. Hopkinson and Smith (2005) focused on cross-system variations and found benthic and pelagic respiration rates varying by more than factors of 30 and 40, respectively, ranges that are broadly consistent with whole-system respiration rates based on the open-water method (Hoellein et al. 2013).

Advances in sensor technology for continuous collection of dissolved oxygen time series suggest that the open-water method (Odum 1956) could be leveraged for applications across multiple locations. This method exploits the diel cycle of dissolved oxygen and allows the simultaneous determination of whole-system gross primary production (P = net primary production + autotrophic respiration), ecosystem respiration (R = heterotrophic + autotrophic respiration), and NEM (P − R). A global synthesis of warm-season, open-water metabolism studies (Hoellein et al. 2013) was based on a total of 47 estimates, almost all of which were derived from a single study of U.S. estuaries (Caffrey 2004). A similar synthesis was done by Yvon-Durocher et al. (2012). These studies evaluated the open-water method as applied to continuous monitoring data available for several decades across the National Estuarine Research Reserve System (NERRS), a network of 30 estuaries across the U.S. Thus, the approach can be broadly applied where data are available, although it is not without limitations. In particular, tidal advection can violate assumptions of the method that the measurements are from the same parcel of water with a continuous metabolic history (Kemp and Boynton 1980). Striking evidence of this limitation is the large number of physically impossible negative P and R estimates in the open-water analysis of 28 3-year time series by Caffrey (2003). The computed rates were negative 23% of the time, on average, and as high as 69% for individual sites.

Another limitation of the open-water method is the estimation of the air–water oxygen flux, which is an important term in the oxygen budget. Most approaches estimate this flux using an empirically derived gas transfer velocity that is dependent on wind speed and temperature or held at a constant value (e.g., Caffrey 2004; Russell and Montagna 2007; Nagel et al. 2009; Beck et al. 2015; Murrell et al. 2018). Particularly challenging is that the exchange velocity in estuaries also depends on water depth and tides (Ho et al. 2016) and possibly additional factors, such as

fetch and the concentration of total suspended solids (Borges and Abril 2011). The importance of the air–water gas flux in the oxygen budget opens up the possibility that the oxygen data themselves may be used to determine the gas transfer velocity, an approach that has been applied in streams (e.g., Riley and Dodds 2013).

A unified model to explain the large variability of metabolic rates within and across estuaries is clearly lacking as a critical tool for understanding metabolic processes. Bayesian techniques have only been explored in limited estuarine applications despite their potential to address the above (Ciavatta et al. 2008; Tassone and Bukaveckas 2019). Bayesian techniques allow for incorporation of prior knowledge about model parameters and robust estimation of parameter probability distributions (Holtgrieve et al. 2010) using highly efficient computational algorithms (Grace et al. 2015; Winslow et al. 2016). These techniques are very powerful because they not only estimate P and R as standard open-water methods do, but they retrieve estimates and their uncertainties of parameters related to gas exchange, P, and R (e.g., the initial slope of the photosynthesis–irradiance curve), which reveals quantitative information about the mechanisms of metabolism and gas exchange. A potentially useful application developed for streams is the BAyesian Single-station Estimation (BASE) method (Grace et al. 2015). BASE has been demonstrated to accurately determine metabolic parameters and their uncertainty using continuous monitoring data that are similar to those available at many estuarine locations. As such, modification of BASE to include model characterizations more appropriate for estuaries, while maintaining the fundamental Bayesian approach, could prove useful.

The goal of this paper is to describe the development and application of a new Bayesian method for simultaneously determining gross primary production and ecosystem respiration in estuaries from high-resolution dissolved oxygen time series. We call this method EBASE (Estuarine

BAyesian Single-station Estimation), which builds and improves on the BASE method (Grace et al. 2015) using an approach more appropriate for estuarine time series. The rationale and changes made to BASE are described in detail, followed by a demonstration of how EBASE can estimate known metabolic parameters with reasonable certainty from a synthetic time series. Lastly, metabolic estimates from EBASE are compared with those of the Odum open-water method to identify key differences related to both the theoretical and statistical differences associated with each method. These comparisons combine results using real observations and those based on detided inputs to demonstrate effects of tidal advection on metabolic estimates. We also attempt to constrain the gas transfer velocity using this approach but are unable, at least with the data sets used in this study. EBASE is available as an open-source software package created with the R statistical programming language (Beck et al. 2023; R Core Team 2023) and the results herein provide practical suggestions for use of the software on novel datasets.

Materials and Procedures

EBASE uses a mass balance equation for dissolved oxygen, assuming a well-mixed water column of depth H and no lateral transport processes:

$$H\frac{dC}{dt} = P - R - D \qquad (1)$$

where the terms on the right side of the equation are gross primary production (P), ecosystem respiration (R), and the net upward diffusive gas flux at the air–water interface (D), respectively, processes that result in the change in dissolved oxygen concentration (dC) per unit time (dt). The equation has dimensions of moles per unit area and time. P is modeled as $aPAR$, where PAR is the surface photosynthetically active radiation and a is the light efficiency; R is constant; and D is modeled in the standard manner (e.g, Sarmiento and Gruber 2013) as $k_w(C_{sat} - C)$, where k_w

136 is the gas transfer velocity and C_{sat} is the saturation (or equilibrium) concentration, calculated
137 from salinity and temperature (García and Gordon 1992). The gas transfer velocity is modeled
138 using the formulation in Wanninkhof (2014):

139 $$k_w = bU_{10}^2\left(\frac{Sc}{600}\right)^{-0.5} \qquad (2)$$

140 where b is constant, U_{10} is the wind speed at 10 m above the water, and Sc is the Schmidt
141 number, which is the ratio of the kinematic viscosity of water to the molecular diffusivity of
142 oxygen in water (Sc is computed from water temperature and salinity using the polynomial fit in
143 Wanninkhof 2014). The b parameter is fairly well known for open-ocean conditions, but may be
144 different in estuaries, where the transfer velocity is influenced by factors other than wind speed
145 and Sc. As such, the three parameters estimated by EBASE are a, R, and b using the required
146 input data that include time series of dissolved oxygen, water temperature, salinity, PAR, wind
147 speed, and water column depth.

148 ## Changes to BASE

149 Equation 1 is similar to the oxygen mass balance equation in BASE (Grace et al. 2015), with the
150 main difference being the treatment of gas exchange. The gas transfer velocity, k_w (Equation 2),
151 increases as the water column in the vicinity of the air–water interface becomes more turbulent.
152 As such, k_w is usually modeled as a function of measurable variables that can potentially predict
153 turbulent mixing of the surface water, such as wind speed. k_w also increases with temperature
154 because the film of water at the air–water interface becomes thinner as the viscosity decreases
155 and the random motion of gas molecules increases. This dependence of k_w on temperature is

156 usually modeled with the Schmidt number as in Equation 1, which combines information on the
157 viscosity of the fluid and the diffusivity of the gas.

158 Grace et al. (2015) used the following for the gas transfer in BASE:

159 $$\frac{k_w}{H} = K \cdot 1.024^{(T - \bar{T})} \qquad (3)$$

160 where T is the instantaneous water temperature, $\bar{T}$ is the average water temperature over each 24-
161 hour period in the oxygen time series (both in °C), and K is the reaeration coefficient estimated
162 from the data. Equation 3 indicates that k_w increases with temperature and depth. The
163 temperature dependence is consistent with the known decrease of k_w with the Schmidt number,
164 but the depth dependence is difficult to justify. An expectation is that k_w would increase as water
165 depth decreases for a given current speed because turbulence generated at the sediment–water
166 interface by currents would more easily reach the surface (Ho et al. 2016). Based on these
167 differences and the lack of availability of other possible drivers of the transfer velocity in coastal
168 systems, the well-established wind- and temperature-based parameterization for gas exchange of
169 Wanninkhof (2014) was used for EBASE, but allowing b, which was estimated by Wanninkhof
170 (2014) to be 0.251 $(\mathrm{cm/hr})/(\mathrm{m^2/s^2})$, to be a free parameter that may reflect spatial and temporal
171 variability in factors other than wind and temperature that may influence transfer velocity.

172 BASE was also modified by removing the temperature dependence of respiration and the
173 potential for P to depend nonlinearly on PAR. While there is evidence for both dependencies in
174 aquatic systems, we found, in preliminary model runs, that model stability was enhanced when
175 these dependencies were removed. It would be straightforward to include such dependencies in

EBASE as our understanding of them improve and as data become available that will allow them to be robustly constrained.

For ease of application, an R package was developed to implement EBASE (Beck et al. 2023; R Core Team 2023), similar to the R package available for BASE (https://github.com/dgiling/BASEmetab). EBASE differs from BASE in model implementation in two ways. First, users can specify the model optimization period, where the time period can vary from a minimum of 1 day to the maximum of the entire length of the time series. This differs from the BASE approach where the optimization period is set at 1 day. Sensitivity of the EBASE method to different optimization periods is described below. Second, the EBASE software can accommodate missing observations in the input data by interpolating gaps prior to estimating the metabolic parameters. Results obtained from interpolated data are automatically removed from the output, eliminating the need to pre- and post-process the data when using EBASE. Again, this feature was not included in the original BASE method and a complete time series was required for use. A detailed web page (https://fawda123.github.io/EBASE/) describes how to use the package, with explanations of the functions for viewing model results, including a plot of the modeled dissolved oxygen with the observed, a time series plot of the oxygen budget terms (P, R, and D), and a time series plot of the credible intervals for the a, R, and b parameters. All results provided herein were generated using the EBASE R package.

EBASE model estimation

The parameters a, R, and b are estimated by likelihood given the observed data and prior distributions for the parameters. The "Just Another Gibbs Sampler" (JAGS) software (as for BASE, Plummer et al. 2003) is used with the EBASE R package (described above) to estimate probability distributions of the unknown parameters using Markov Chain Monte Carlo (MCMC)

simulations in a Bayesian framework applied to a finite-difference form of Equation 1. A JAGS model file is included with the package that implements Equation 1, with options to supply a custom model file as needed.

The metabolism estimates and their parameters returned by the Bayesian routine implemented in JAGS are affected by the prior distributions assigned to each. As for BASE, relatively uninformed prior distributions following a normal Gaussian distribution are used by default, although the priors can be changed based on previous knowledge of parameters specific to an ecosystem or as informed by other metabolic modeling approaches. Reasonable uninformed prior distributions for EBASE were chosen with mean values using approximate estimates from the literature (Caffrey 2004; Wanninkhof 2014) and standard deviations that were sufficiently large to allow the Bayesian routine to search an unconstrained parameter space.

Specifically, the prior mean for R was chosen as 300 $\mathrm{mmol/m^2/d}$ (~10 g $\mathrm{O_2/m^2/d}$) using a synthesis based on NERRS data (Table 2 in Caffrey 2004). Although this synthesis focused on small, shallow estuaries that are part of the NERRS, this dataset is the largest, longest, and most consistently collected in the US and it is reasonable to expect others may apply EBASE to these data. The prior mean for a was chosen assuming that P and R are comparable at long time scales, such that $aPAR = R$, or $a = (300\ \mathrm{mmol/m^2/d})/PAR$. Using the global mean surface PAR of 80 W/m^2, which is based on a global mean surface total shortwave radiation of 200 $\mathrm{W/m^2}$ (Kiehl and Trenberth 1997) and a PAR fraction of total shortwave radiation of 0.4 (Jacovides et al. 2004), the prior mean for a is reasonably set at 4 $(\mathrm{mmol/m^2/d})/(\mathrm{W/m^2})$. Lastly, the prior mean for b was chosen as 0.251 $(\mathrm{cm/hr})/(\mathrm{m^2/s^2})$ following Wanninkhof (2014).

The prior standard deviation for R was based on the standard deviation of the mean values of R at the NERRS sites in Caffrey (2004). The standard deviation of R was 171 $\mathrm{mmol/m^2/d}$, which was rounded to 50% of the mean, or 150 $\mathrm{mmol/m^2/d}$. The standard deviation for a was set proportionally to 2 $(\mathrm{mmol/m^2/d})/(\mathrm{W/m^2})$. For the prior standard deviation of b, we considered using the 20% error estimated for b by Wanninkhof (2014). However, we chose a larger value, 50%, because the wind speed-based gas transfer velocity formulation of Wanninkhof (2014) was mainly developed and calibrated for open-ocean applications, where the complications of surfactants (Frew 1997) and tidal currents (Ho et al. 2016), as well as fetch and suspended matter (Borges and Abril 2011), are expected to be minimal.

Together, the default prior distributions for a, R, and b (Figure 1) are:

$$a \sim N(4,2)I(0,\infty) \quad (\mathrm{mmol/m^2/d})/(\mathrm{W/m^2}) \qquad (4)$$

$$R \sim N(300,150)I(0,\infty) \quad \mathrm{mmol/m^2/d} \qquad (5)$$

$$b \sim N(0.251,0.125)I(0,0.502) \quad (\mathrm{cm/hr})/(\mathrm{m^2/s^2}) \qquad (6)$$

where $N(\alpha,\beta)$ indicates a normal distribution with mean α and standard deviation β and $I(\gamma,\theta)$ is 1 between γ and θ and 0 elsewhere; I has the effect of truncating the normal distribution to minimum and maximum values. All prior distributions were constrained to positive values based on known physical constraints and requirements of the model formula in Equation 1. Additionally, the prior distribution for b was constrained to an upper limit of 0.502 $(\mathrm{cm/hr})/(\mathrm{m^2/s^2})$ (~twice the default mean). Initial development of EBASE showed that using an undefined upper limit of b led to unstable and unreasonable estimates for all parameters. Functions in the EBASE R package allow alternative prior distributions from those above.

Assessment

EBASE comparison with known results

A preliminary assessment of the ability of EBASE to produce reasonable parameter estimates for a, R, and b (Equation 1) was conducted using a synthetic time series of dissolved oxygen concentration created from observed water temperature, salinity, air temperature, PAR, and wind speed measured every 15 minutes. This assessment evaluated if (1) EBASE estimates similar values for the known parameters used to create the synthetic time series, and (2) how the estimated values change with different model configurations. The synthetic time series used inputs from one year (2021) of continuous data from the NERRS reserve in Apalachicola Bay, Florida. Water temperature and salinity were from the Cat Point water quality monitoring station (29.7021° N, 84.8802° W) and the meteorological observations were from the nearby East Bay station (29.7694° N, 84.8815° W, ~10 km north of Cat Point). Missing observations in the water and meteorological data (5.1% and 3.4% of all observations in each dataset, respectively) were filled using autoregressive modelling (Akaike 1969) of the actual data versus time to create a complete dataset from January 1st to December 31st.

The synthetic oxygen time series was generated using Equation 1 as a forward calculating model with specified daily values for a, R, and b. The b parameter was fixed at 0.251 $(\mathrm{cm/hr})/(\mathrm{m^2/s^2})$ (Wanninkhof 2014). Daily values of a and R were estimatd by (1) applying the Odum open-water method as implemented in the WtRegDO R package (Beck 2023), (2) computing monthly averages of the resulting daily estimates of P and R, (3) computing the monthly average of the parameter a by dividing the monthly average of P by the monthly average of PAR, and (4) by fitting a spline function to the monthly averages of R and a.

The resulting synthetic time series of light efficiency, dissolved oxygen budget terms (P, D, and R), and dissolved oxygen concentration are shown in Figure 2. The light efficiency has a complex structure, with maxima in the summer, fall, and winter, and a strong minimum in spring (Figure 2a). The oxygen budget terms (Figure 2b) reveal that (1) photosynthesis has a structure similar to that of light efficiency, but noisier, reflecting the synoptic-scale variability of PAR; (2) respiration has the expected structure of being greatest during the summer (Caffrey 2004), stimulated by greater autochthonous organic matter variability and high temperature; and (3) the air–water gas flux mimic productivity on short time scales, with outgassing during high productivity and ingassing during low productivity. The synthetic dissolved oxygen time series shows high variability at multiple time scales (Figure 2c). There is a prominent diurnal cycle that is superimposed on subseasonal variability, presumably driven by radiation and winds.

Recovery of model parameters

The synthetic time series was used as input to EBASE to predict metabolic parameters with the default prior distributions above. Two optimization periods of 7 and 30 days were evaluated. This analysis provided a demonstration that the default priors are not entirely unreasonable based on the ability to reproduce the known parameters in the synthetic time series (Figure 3). Overall, EBASE produced reasonable results for a and R, with higher accuracy using the 30 day optimization period (Figure 3b) compared to the 7 day optimization period (Figure 3a). However, the model was unable to return the b parameter with acceptable accuracy, suggesting the data were insufficient to constrain this parameter. Further, many of the 95% credible intervals from the posterior distributions of each parameter did not overlap the known values from the synthetic time series. Estimates for each optimization period may be inaccurate despite overall similarities comparing all estimates to the known values for the whole year.

Sensitivity to priors

Several configurations of the EBASE model can influence how the parameters and metabolic rates are estimated. Two types of sensitivity analyses were conducted to evaluate changes in the results returned by EBASE, where in all cases, the results were compared to the known values in the synthetic time series. First, results were compared to the known values after changing characteristics of their prior distributions. Second, results were compared using a model period of optimization for 1, 7, and 30 days to assess how results varied based on the amount of data used for the Bayesian estimation. Details of these analyses are as follows.

The prior distributions for the EBASE parameters follow truncated normal distributions with defined means (μ) and standard deviations (σ) (Equations 4, 5, and 6). The sensitivity analysis evaluated changes in the distributions by varying the means and standard deviations from small to large to test the effect of changes in the central tendencies and more or less constrained ranges, respectively, on the results. Low and high values for the means and standard deviations for R were assessed following reasonable ranges from Caffrey (2004) using logic similar to that for choosing the default priors. The minimum and maximum mean values for R were evaluated as 138 and 1009 $(\mathrm{mmol/m^2/d})/(\mathrm{W/m^2})$, which we rounded to $1/2$ and 3 times the mean (150 and 900 $\mathrm{mmol/m^2/d}$) and proportionally for a as 2.0 and 12 $(\mathrm{mmol/m^2/d})/(\mathrm{W/m^2})$. The minimum and maximum values for the standard deviations for R and a were chosen as 10% and 1000% of the default mean values to test a range from high precision to completely uninformed values. The mean and standard deviation for the b parameter were not tested and the prior value was set as the fixed constant used in the synthetic time series, 0.251 $(\mathrm{cm/hr})/(\mathrm{m^2/s^2})$. There is minimal prior knowledge on expectations for the b parameter and initial assessments with the synthetic time series demonstrated that the parameter was poorly constrained, primarily because the fixed

parameter could not be estimated using observed data (i.e., wind, water temperature) that affected the actual realization of b. As such, each unique combination of low and high values for the means and standard deviations for a and R were evaluated, creating $2^4 = 16$ different combinations of prior distributions that were used with the synthetic time series in EBASE.

The optimization period used for a time series determines how much data are used to estimate the a and R parameters. Unlike BASE, the EBASE software allows flexibility in the period of time used for optimization, with a minimum being 1 day and a maximum being the length of the time series. For a chosen optimization period, single estimates for a and R are returned by the model (b is also returned, but was fixed for the sensitivity analysis). A short period of optimization (e.g., 1 day) may return unstable parameter estimates if forcing factors (e.g., wind, temperature) that affect estuarine biogeochemical rates occur at time scales longer the chosen period of optimization for EBASE. As such, potentially longer optimization periods may return more stable and robust estimates given that more data are used for parameter estimation, although at the risk of overgeneralizing metabolic rates if biogeochemical variation occurs at time scales less than the optimization period. Results from optimization periods of 1, 7, and 30 days were compared for the synthetic one-year time series. For the models with 7 and 30 day optimization periods, the final period in the annual time series was excluded from the results due to incomplete number of observations compared to earlier periods (i.e., the total number of days in the time series was not evenly divided by 7 or 30, resulting in a remainder for the last period). Each optimization period was also evaluated for every unique combination of prior distributions, totaling 48 (16 times three optimization periods) unique analyses for the assessment of changing priors of all parameters.

Changing the prior distributions affected the ability of EBASE to return the known parameters from the synthetic time series, with results varying by optimization period. Figure 4 shows a summary of the comparisons of the synthetic data with EBASE for each unique combination of prior distributions and the three optimization periods of 1, 7, and 30 days. The results are shown using Nash-Sutcliffe Efficiency (NSE) values (Nash and Sutcliffe 1970; Moriasi et al. 2007):

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(Y_i^{Syn} - Y_i^{EBASE}\right)^2}{\sum_{i=1}^{n}\left(Y_i^{Syn} - \bar{Y}^{Syn}\right)^2} \qquad (7)$$

where Y_i is the estimate from EBASE or the synthetic times series for the optimization period i, evaluated for each of five model outputs (D, DO, R, a, and P), each of 16 unique combinations of priors, and each of 3 optimization periods. The value n is the number of optimization periods in the one-year time series, i.e., $n = 365$ for comparability among the optimization periods. The NSE value is conceptually similar to the coefficient of determination, but varies from $-\infty$ to 1. Values in the positive range from 0 to 1 are desirable with 1 being a perfect fit, whereas values in the negative range indicate the mean of the synthetic data is a better predictor than EBASE.

Several conclusions can be made from the results in Figure 4 following the patterns of the NSE values. Overall, most models regardless of optimization period or priors were able to show some skill at reproducing the dissolved oxygen time series, which had a median NSE across all optimization periods and priors of 0.91. The only models where the NSE values were less than zero for dissolved oxygen were those where σ of the a and R parameters was low and μ for one of the a and R parameters was high. However, substantial variation in NSE values was observed among all other model outputs, with some very low values of NSE occurring while the DO simulation was fairly good, showing the potential for the model to produce the right answer for

the wrong reasons. After dissolved oxygen, gas exchange (D) was estimated with the highest median NSE of 0.73. This result makes sense, given that differences in D across all optimization periods and priors are due solely to difference in DO. Conversely, the a parameter was reproduced the most poorly, with a median NSE of -0.75.

By optimization period, median NSE values across all model outputs increased as the number of days used for model optimization increased (1 day = -2.42, 7 days = 0.26, 30 days = 0.6). The top three ranked models were identical for the 1 and 7 day optimization period, whereas the top three models for the 30 day optimization appeared less dependent on the choice of priors. Median NSE values across model outputs for the top models increased with optimization period (1 day = 0.09, 7 days = 0.44, 30 days = 0.64). The NSE values for all model outputs were greater than zero for the top three models in the 30 day optimization period, whereas the a parameter had NSE less than zero for the top model in the 7 day optimization period and R and a had negative NSE values for the top model in the 1 day optimization period. Less than half of the model outputs regardless of optimization period or priors had NSE < 0 (43%). As such, the prior combinations that produced the highest NSE values across all parameters were those where all prior distributions had low μ and σ values for both a and R for the 1 and 7 day optimization periods, whereas the combination of priors did not have a notable effect on the output for the 30 day optimization period. Notably, many of the top models for the shorter optimization periods had at least one high value for σ for a or R, suggesting that models with sufficient flexibility to search a larger parameter space can produce metabolic estimates most similar to actual values.

The actual model results for the best and worst performing combinations of prior distributions for the three optimization periods are shown in Figure 5 (all results are averaged at the time-step of the optimization period for visual comparability). The best performing models were those with

375 the highest median NSE values across all model outputs in Figure 4. The best performing models
376 for each optimization period (Figure 5 a, c, e) showed similar results where the model outputs
377 were similar to the synthetic time series, although more variation in the EBASE results were of
378 course observed with the 1 day optimization period. The worst performing models for the 1 and
379 7 day optimization periods (Figure 5 b, d) demonstrate the negative NSE values that were
380 obtained (i.e., EBASE values much lower or higher than the synthetic values). The worst
381 performing models were caused by constraining EBASE to a small parameter space, such as a
382 high μ and low σ for a as in the bottom row for subplots (b) and (d). In such cases, the DO
383 simulation (fourth row) can still be quite good as a result of compensating errors in a and R. The
384 EBASE-derived rates of photosynthesis and respiration are both much higher than in the
385 synthetic data for subplots (b) and (d), errors that compensate well enough to produce fairly
386 accurate recoveries of the DO time series. Further, resutls for the worst model for the 30 day
387 optimization period in subplot (f) are visually indistinguishable from those of the best model in
388 subplot (e), suggesting that poor estimates from inadequate priors can in part be mitigated using
389 longer optimization periods.

390 **Impact of noise on parameter recovery**

391 As a final evaluation of EBASE to return known metabolic parameters, a second synthetic time
392 series was created that added estimated effects of noise to the original synthetic time series. The
393 objective of the analysis is to develop an expectation of how metabolic estimates may change
394 using *in situ* data with noise that may substantially mask a biological signal in dissolved oxygen
395 measurements. The noise estimates were created using a weighted regression analysis (Beck et
396 al. 2015; Beck 2023) on observed dissolved oxygen time series for Apalachicola Bay in 2021.
397 This approach models dissolved oxygen as a function of time (day), hour, and water depth using

a moving window approach with weights of the independent variables appropriate for each time period within the windows. The results of this analysis return a predicted and detided dissolved oxygen time series. Two separate time series of noise, representing residual (any other random process not related to metabolism or tidal advection) or tidal noise, were derived from the observed, predicted and detided time series. The residual noise was estimated as the difference between the observed and predicted time series and the tidal noise was estimated as the difference between the predicted and detided time series (Figure 6a). Both the residual and tidal noise were added to the original synthetic time series in Figure 2c and are shown in Figure 6b. EBASE was then used on the synthetic time series with noise, then compared with results from the original synthetic time series. The default priors (Equations 4 and 5) were used with a 7 day optimization period. The prior for the b parameter was fixed at 0.251 $(\mathrm{cm/hr})/(\mathrm{m^2/s^2})$ following similar logic used for the sensitivity analysis described above.

Figure 7 shows the comparison of the results recovered from EBASE for the synthetic time series with noise (residual and tidal) and for the synthetic time series without noise. EBASE was able to recover reasonable estimates for both a and R, although the results from the noisy time series were slightly biased towards higher values as shown by the blue regression lines on the right side of Figure 7. The higher values for a will produce higher estimates for P, and combined with R, will produce inflated values for metabolic estimates compared to the results for the synthetic time series without noise. Although the results for the latter were not biased towards higher values, both estimates provided by EBASE had a relatively large range of variablity around the known values. As noted above, a longer optimization period could reduce this variability.

EBASE and Odum comparison with real observations

Metabolic estimates from EBASE were further compared to those from existing methods to benchmark and evaluate potential differences from alternative model formulations used by each approach. The observed one-year continuous time series described above for water quality and weather data from Apalachicola Bay was used to compare metabolic estimates from EBASE and the Odum open-water method (Odum 1956). The Odum method used an approach described in Murrell et al. (2018) created for estuarine application that uses a gas-exchange approach from Thébault et al. (2008). For a better comparison with the EBASE methods, the gas exchange parameterization in the Odum method was modified to use the Wanninkhof (2014) approach and the *b* parameter was fixed at 0.251 $(\mathrm{cm/hr})/(\mathrm{m^2/s^2})$ (as available in the WtRegDO R package, Beck 2023). An additional and fundamental difference between EBASE and the Odum method is that the former is likelihood-based with metabolic rates estimated by fitting the model to the data, whereas the latter is strictly arithmetic-based primarily on integration of the diel dissolved oxygen curve. Thus, potentially different results are expected based on the alternative statistical approaches to estimate the metabolic results. Each comparison also evaluated differences in the results using the observed dissolved oxygen time series and a detided time series using weighted regression (Beck et al. 2015; Beck 2023). The latter comparison provided an assessment of metabolic rates that are expected to be minimally influenced by tidal advection, as a fundamental assumption of metabolic models using *in situ* dissolved oxygen time series from fixed monitoring stations. The prior distributions for EBASE used the default values (Equations 4, 5, and 6). Similar to the previous analyses, the prior distribution for the *b* parameter was fixed given the difficulty in recovery. A 1 day model optimization period was used for EBASE for comparability with the Odum method that returns daily estimates.

Pairwise comparisons of the daily estimates of NEM, P, and R returned by each method were evaluated using simple summary statistics, including Pearson correlation coefficients (ρ) and root mean square differences (RMSD) of the linear regression fit comparing the same flux type between methods. Figure 8 and Table 1 provide a comparison of the metabolic estimates from the EBASE and Odum methods. Correlations for all metabolic estimates, regardless of the input time series, were positive and significantly correlated, as expected. However, correlations were generally stronger (and RMSDs smaller) for the metabolic results based on the detided dissolved oxygen time series. The observed time series clearly produced a much larger range of values for results from both the EBASE and Odum methods.

The comparisons in Figure 8 and Table 1 suggest that tidal effects on continuous monitoring data can severely violate assumptions required for methods of estimating ecosystem metabolism (as described in detail in Beck et al. 2015). These effects are easily identified by stronger correlations and lower RMSD using the detided dissolved oxygen. Further, effects of tidal advection can be seen with estimates from the Odum method as negative values for P and R, i.e., "anomalous" values. The effects of tidal advection on estimates from EBASE were more subtle as anomalous values cannot be returned based on constraints of the priors (i.e., none can be negative), although the range of values for the metabolic estimates were much larger with observed dissolved oxygen. A similar result (larger estimates for P and R) was observed using the synthetic time series with instrument and tidal noise (Figure 7). Overall, the results suggest that detiding the dissolved oxygen time series prior to estimating metabolic rates reduces tidal noise and produces more stable estimates with smaller ranges, regardless of the method used.

Discussion

The above analyses demonstrated that EBASE can successfully recover known metabolic parameters from a synthetic time series and provides robust estimates on real data covering a year of observations. The success of these results was predicated on the requirements of EBASE to address several critical needs that existing methods for estimating metabolism do not address for coastal applications. These needs were addressed primarily using a Bayesian framework that estimates the best fit of the model to the data with posterior probabilities describing the likelihood of the fit. This framework enables hypothesis testing of how a model can be fit to data by using prior knowledge in the model fitting process, which results in posterior probabilities describing the certainty of the parameter estimates (Hilborn and Mangel 2013). The advantages of the Bayesian approach have led to widespread adoption over frequentist approaches in many scientific applications, especially in the coastal and marine community (e.g., Myers et al. 2001; Borsuk et al. 2001; Brown et al. 2017), yet these methods have not been rigorously explored for estimating metabolic rates in estuaries. Our results demonstrated that even using uninformed priors, EBASE can reproduce known metabolic parameters and rates with reasonable certainty. However, initial analyses suggested an inability of the model to estimate stable values for the b parameter that influence gas exchange. This parameter was fixed for the sensitivity analyses and we were unable to evaluate the ability of EBASE to recover it from the synthetic time series. Future work should focus on understanding the sensitivity of EBASE to estimating b where the dominance of the relative processes that manifest signals in the dissolved oxygen time series could vary. For example, estuarine dissolved oxygen at Apalachicola Bay may be dominated by production and respiration, and less so from gas exchange, such that the available data may provide insufficient information for EBASE to estimate b. This aligns with others that have

487 suggested gas exchange may be a relatively small component of the metabolic signal in shallow
488 estuaries of the Gulf of Mexico (Murrell et al. 2018). In such cases, more constrained prior
489 distributions for b could be appropriate or EBASE could be tested in settings with expected
490 larger contributions of gas exchange expressed by dissolved oxygen.

491 Our results also confirmed that the effects of tidal advection on metabolic estimates can lead to
492 biases in understanding gross primary production and ecosystem respiration. This result has
493 previously been shown using the Odum open water method (Kemp and Boynton 1980; Caffrey
494 2003; Beck et al. 2015) and confirmed herein, whereby results using observed time series can
495 produce over- and under-estimates of metabolic rates with more variation than those without
496 tidal influence (Figures 7 and 8). As a result, detiding dissolved oxygen data is recommended
497 prior to estimating metabolism, regardless of whether EBASE or an alternative method is used.
498 In addition to inaccurate and imprecise metabolic estimates, tidal influences are also most likely
499 observed as negative (or anomalous) values in the production and respiration rates using the
500 Odum method (Caffrey 2003). Tidal effects may be less obvious using EBASE since negative
501 values cannot be obtained based on constraints of the prior distributions, although as noted
502 above, more variable estimates are returned. A potential future approach for developing EBASE
503 could be an explicit formulation of tidal advection in the model using a synoptic tidal height time
504 series in the Bayesian framework. Such an approach would have the advantage of simultaneously
505 detiding and estimating metabolic rates using the same model, whereas using weighted
506 regression prior to EBASE, as used herein, requires additional analysis and assumptions.
507 However, field measurements (e.g., control volume approaches, Falter et al. 2008; Nidzieko et al.
508 2014) may be required to provide reasonable constraints on the prior distributions for any
509 parameters that quantify advective influences on dissolved oxygen.

As described in the materials and procedures section, EBASE is an estuarine adaptation of BASE (Grace et al. 2015), which was developed for stream applications and some assumptions of the underlying model of the latter were inappropriate for the former. The most critical change made to BASE was adopting the gas transfer velocity parameterization in Wanninkhof (2014), which is primarily wind-based and more appropriate than the depth-dependent formulation in BASE. EBASE also removed the temperature dependency of respiration, consistent with applications of the Odum method to other shallow estuaries (Beck et al. 2015; Murrell et al. 2018). Overall, these changes also reduced the required inputs and many monitoring locations nationwide include the requisite data for applying EBASE (i.e., dissolved oxygen, water temperature, salinity, PAR, and wind speed, as collected by the NERRS).

EBASE also provides several computational advantages over BASE implemented in the R package that was created for the purposes herein (Beck et al. 2023). Most importantly, EBASE was developed to allow different model optimization periods, whereas BASE was developed to estimate results for a single day. Allowing EBASE to incorporate more than a day of observations in the optimization can likely produce more accurate and precise estimates for the metabolic parameters. Physical and biological characteristics of an ecosystem that affect the metabolic rates are likely acting at time scales longer than a day, e.g., ecological and biogeochemical characteristics influencing the a and R parameters (such as community composition and nutrient availability) may persist for several days. Exposing the model to additional observations may produce more stable results as the parameters for the best fit are estimated by the Bayesian routine as a function of the data that are evaluated each optimization period. However, we cannot provide a precise recommendation on the number of days to use for the optimization period as the exact length of time that ecosystem characteristics can affect

metabolic rates may not be known and likely varies by location. Our example using the synthetic time series from Apalachicola Bay suggested that the most precise results were obtained using the 30 day optimization period, although results are likely to vary across systems. For most applications, very short (e.g., 1 day) or very long (e.g., > 30 days) are not recommended, such that the former may produce unstable results and the latter may miss important events in the time series producing parameters that are overly generalized. However, an advantage of the Bayesian approach is the ability to evaluate different optimization periods and, therefore, the amount of data on which to estimate key parameters. Hypotheses can be generated on the amount of time over which critical metabolic parameters may vary across systems.

The EBASE R package can also accommodate missing observations in the input data. This feature allows the estimation of metabolic rates for time series that cover relatively long periods of time (e.g., more than year), when data gaps are likely to occur for several reasons (e.g., equipment malfunction and routine maintenance). The R package accommodates data gaps using linear interpolation for any required input used by EBASE, allowing the Bayesian routine to provide continuous estimates across the time series. The results are then automatically post-processed to remove any metabolic estimates that exceed a threshold for the number of interpolated observations. The analyst can determine the appropriate length of time to use for excluding results, although the default setting for EBASE is 12 hours, where results for an entire optimization period are removed if any interpolated period exceeded 12 hours within that period. Although this does not apply to the synthetic time series, some of the estimates using the 2021 observed data for Apalachicola Bay were excluded from the analysis. This feature allows EBASE to be applied more easily to other data sets because an analyst does not need to manually screen data that include missing values prior to analysis.

While the EBASE theory and software provide advantages for metabolic estimates, the approach is not without limitations. Like most models, explicit choices are needed to apply EBASE to time series data. First, the use of unconstrained prior distributions was shown to reproduce known metabolic parameters with reasonable certainty, although initial assessments suggested the b parameter was relatively unstable and unconstrained priors may not provide useful estimates. Parameters that are poorly estimated may require constrained priors, although doing so may produce biased results in the absence of additional data to justify a constraint. Field-based measurements or alternative metabolic models could be used for informing priors. As an example, initial application of BASE to estuarine data prior to developing EBASE demonstrated that unstable and unrealistic metabolic rates were returned with uninformed priors. The Odum method was used on the same data to develop an expectation of the range of values for the reaeration coefficient, K, which was then used to create an informed prior distribution for the same parameter using the BASE method. More stable and realistic metabolic estimates were then returned by BASE, demonstrating the value of the Bayesian approach that can incorporate prior knowledge. A similar approach using alternative metabolic models for informing priors could be applied to create more accurate estimates from EBASE. An additional choice for using EBASE is the model optimization period, although some guidance is provided above.

A final limitation of EBASE is the relatively long processing time when applying the method to a time series longer than a few days. For example, the execution time applying EBASE to a year of data on a conventional laptop computer was approximately 40 minutes, depending on the model settings. This time is compared to the Odum method applied to the same dataset, where results are obtained in a matter of seconds. Given the Bayesian framework used by EBASE, not considering the additional statistical advantages, it is not reasonable to expect comparable

processing times between the methods. Regardless, computation times could be reduced with additional hardware improvements (e.g., accessing multiple processing cores) or changing the model estimation methods (e.g., reducing the number of MCMC simulations used by JAGS), although the latter may negatively affect model accuracy.

Comments and Recommendations

The development of EBASE represents a new approach for estimating metabolic rates in estuarine settings that leverages existing methods in a Bayesian framework. Our results have demonstrated that unknown metabolic parameters can be estimated with reasonable certainty and EBASE could be a valuable approach to improve the understanding of critical ecosystem processes as the method can be readily applied to other settings. A notable limitation of EBASE was an inability to constrain the gas transfer velocity from a synthetic time series, although this result may be explained by a minimal contribution of gas exchange as expressed in the dissolved oxygen time series. Effects of tidal advection on dissolved oxygen measurements were also observed, demonstrating a violation of the assumption that the measured data represent a water sample having a continuous metabolic history. Further development of EBASE to accommodate a tidal signal, or exploring alternative methods for detiding dissolved oxygen data, are promising avenues of future research to more accurately quantify metabolic parameters in estuaries. Overall, EBASE represents a potentially powerful tool that removes many of the assumptions of existing methods by allowing unknown metabolic parameters to be estimated by the data and can also incorporate existing knowledge for further exploration of estuarine ecosystem properties.

References

Akaike, H. 1969. Fitting autoregressive models for prediction. Annals of the Institute of Statistical Mathematics **21**: 243–247. doi:10.1007/bf02532251

Beck, M. W. 2023. WtRegDO: Implement weighted regression on dissolved oxygen time series, R package version 1.0.1. https://github.com/fawda123/WtRegDO

Beck, M. W., J. D. Hagy III, and M. C. Murrell. 2015. Improving estimates of ecosystem metabolism by reducing effects of tidal advection on dissolved oxygen time series. Limnology and Oceanography: Methods **13**: 731–745. doi:10.1002/lom3.10062

Beck, M. W., M. Herrmann, J. Arriola, R. Najjar, and W. McGillis. 2023. EBASE: Estuarine Bayesian Single-Station Estimation Method for Ecosystem Metabolism, R package version 0.0.0.9015. https://fawda123.github.io/EBASE

Borges, A. V., and G. Abril. 2011. Carbon dioxide and methane dynamics in estuaries, p. 119–161. *In* E. Wolanski and D.S. McLusky [eds.], Treatise on estuarine and coastal science, volume 5. Academic Press, Waltham.

Borsuk, M. E., D. Higdon, C. A. Stow, and K. H. Reckhow. 2001. A Bayesian hierarchical model to predict benthic oxygen demand from organic matter loading in estuaries and coastal zones. Ecological Modelling **143**: 165–181. doi:10.1016/s0304-3800(01)00328-3

Brown, C. J., S. D. Jupiter, S. Albert, and others. 2017. Tracing the influence of land-use change on water quality and coral reefs using a Bayesian model. Scientific Reports **7**. doi:10.1038/s41598-017-05031-7

Caffrey, J. M. 2003. Production, respiration and net ecosystem metabolism in U.S. estuaries, p. 207–219. *In* B.D. Melzian, V. Engle, M. McAlister, S. Sandhu, and L.K. Eads [eds.], Coastal monitoring through partnerships. Springer Netherlands.

Caffrey, J. M. 2004. Factors controlling net ecosystem metabolism in U.S. estuaries. Estuaries **27**: 90–101. doi:10.1007/bf02803563

Cai, W.-J. 2011. Estuarine and coastal ocean carbon paradox: CO_2 sinks or sites of terrestrial carbon incineration? Annual Review of Marine Science **3**: 123–145. doi:10.1146/annurev-marine-120709-142723

Ciavatta, S., R. Pastres, C. Badetti, G. Ferrari, and M. B. Beck. 2008. Estimation of phytoplanktonic production and system respiration from data collected by a real-time monitoring network in the Lagoon of Venice. Ecological Modelling **212**: 28–36. doi:10.1016/j.ecolmodel.2007.10.025

Cloern, J. E., S. Foster, and A. Kleckner. 2014. Phytoplankton primary production in the world's estuarine-coastal ecosystems. Biogeosciences **11**: 2477–2501. doi:10.5194/bg-11-2477-2014

Falter, J. L., R. J. Lowe, M. J. Atkinson, S. G. Monismith, and D. W. Schar. 2008. Continuous measurements of net production over a shallow reef community using a modified Eulerian approach. Journal of Geophysical Research **113**. doi:10.1029/2007jc004663

Frew, N. M. 1997. The role of organic films in air-sea gas exchange, p. 121–172. *In* P.S. Liss and R.A. Duce [eds.], The sea surface and global change. Cambridge University Press.

García, H. E., and L. I. Gordon. 1992. Oxygen solubility in seawater: Better fitting equations. Limnology and oceanography **37**: 1307–1312. doi:10.4319/lo.1992.37.6.1307

Gazeau, F., A. Borges, C. Barrón, and others. 2005. Net ecosystem metabolism in a micro-tidal estuary (Randers Fjord, Denmark): evaluation of methods. Marine Ecology Progress Series **301**: 23–41. doi:10.3354/meps301023

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). Limnology and Oceanography: Methods **13**: 103–114. doi:10.1002/lom3.10011

Hilborn, R., and M. Mangel. 2013. The ecological detective: Confronting models with data, Princeton University Press.

Ho, D. T., N. Coffineau, B. Hickman, N. Chow, T. Koffman, and P. Schlosser. 2016. Influence of current velocity and wind speed on air-water gas exchange in a mangrove estuary. Geophysical Research Letters **43**: 3813–3821. doi:10.1002/2016GL068727

Hoellein, T. J., D. A. Bruesewitz, and D. C. Richardson. 2013. Revisiting Odum (1956): A synthesis of aquatic ecosystem metabolism. Limnology and Oceanography **58**: 2089–2100. doi:10.4319/lo.2013.58.6.2089

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. Limnology and Oceanography **55**: 1047–1063. doi:10.4319/lo.2010.55.3.1047

Hopkinson, C. S., and E. M. Smith. 2005. Estuarine respiration: An overview of benthic, pelagic, and whole system respiration, p. 122–146. *In* P. Del Giorgio and P. Williams [eds.], Respiration in aquatic ecosystems. Oxford University Press.

Hu, J., P. Peng, G. Jia, B. Mai, and G. Zhang. 2006. Distribution and sources of organic carbon, nitrogen and their isotopes in sediments of the subtropical Pearl River estuary and adjacent shelf, Southern China. Marine Chemistry **98**: 274–285. doi:10.1016/j.marchem.2005.03.008

Jacovides, C. P., F. S. Timvios, G. Papaioannou, D. N. Asimakopoulos, and C. M. Theofilou. 2004. Ratio of PAR to broadband solar radiation measured in Cyprus. Agricultural and Forest Meteorology **121**: 135–140. doi:10.1016/j.agrformet.2003.10.001

Kemp, W. M., and W. R. Boynton. 1980. Influence of biological and physical processes on dissolved oxygen dynamics in an estuarine system: Implications for measurement of community metabolism. Estuarine and Coastal Marine Science **11**: 407–431. doi:10.1016/s0302-3524(80)80065-x

Kemp, W. M., and J. M. Testa. 2011. Metabolic balance between ecosystem production and consumption, p. 83–118. *In* E. Wolanski and D. McLusky [eds.], Treatise on estuarine and coastal science. Elsevier.

675 Kiehl, J. T., and K. E. Trenberth. 1997. Earth's annual global mean energy budget. Bulletin of
676 the American Meteorological Society **78**: 197–208. doi:10.1175/1520-
677 0477(1997)078<0197:eagmeb>2.0.co;2

678 Moriasi, D. N., J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith.
679 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed
680 simulations. Transactions of the American Society of Agricultural and Biological Engineers **50**:
681 885–900. doi:10.13031/2013.23153

682 Murrell, M. C., J. M. Caffrey, D. T. Marcovich, M. W. Beck, B. M. Jarvis, and J. D. Hagy. 2018.
683 Seasonal oxygen dynamics in a warm temperate estuary: Effects of hydrologic variability on
684 measurements of primary production, respiration, and net metabolism. Estuaries and Coasts **41**:
685 690–707. doi:10.1007/s12237-017-0328-9

686 Myers, R. A., B. R. MacKenzie, K. G. Bowen, and N. J. Barrowman. 2001. What is the carrying
687 capacity for fish in the ocean? A meta-analysis of population dynamics of North Atlantic cod.
688 Canadian Journal of Fisheries and Aquatic Sciences **58**: 1464–1476. doi:10.1139/f01-082

689 Nagel, J. L., W. M. Kemp, J. C. Cornwell, M. S. Owens, D. Hinkle, and C. J. Madden. 2009.
690 Seasonal and regional variations in net ecosystem production in *Thalassia testudinum*
691 communities throughout Florida Bay. Contributions in Marine Science **38**: 91–108.

692 Najjar, R. G., M. Herrmann, R. Alexander, and others. 2018. Carbon budget of tidal wetlands,
693 estuaries, and shelf waters of eastern North America. Global Biogeochemical Cycles **32**: 389–
694 416. doi:10.1002/2017gb005790

695 Nash, J. E., and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models part i—a
696 discussion of principles. Journal of hydrology **10**: 282–290. doi:10.1016/0022-1694(70)90255-6

697 Nidzieko, N. J., J. A. Needoba, S. G. Monismith, and K. S. Johnson. 2014. Fortnightly tidal
698 modulations affect net community production in a mesotidal estuary. Estuaries and Coasts **37**:
699 91–110. doi:10.1007/s12237-013-9765-2

700 Odum, H. T. 1956. Primary production in flowing waters. Limnology and Oceanography **1**: 102–
701 117.

702 Plummer, M. and others. 2003. JAGS: A program for analysis of Bayesian graphical models
703 using Gibbs sampling. *Proceedings of the 3rd international workshop on distributed statistical*
704 *computing*. Vienna, Austria. 1–10.

705 R Core Team. 2023. R: A language and environment for statistical computing, v4.3.1, R
706 Foundation for Statistical Computing.

707 Riley, A. J., and W. K. Dodds. 2013. Whole-stream metabolism: strategies for measuring and
708 modeling diel trends of dissolved oxygen. Freshwater Science **32**: 56–69. doi:10.1899/12-058.1

709 Russell, M. J., and P. A. Montagna. 2007. Spatial and temporal variability and drivers of net
710 ecosystem metabolism in western Gulf of Mexico estuaries. Estuaries and Coasts **30**: 137–153.
711 doi:10.1007/bf02782974

712 Sarmiento, J. L., and N. Gruber. 2013. Oceanic carbon cycle, atmospheric CO2, and climate, *In*
713 J.L. Sarmiento [ed.], Ocean biogeochemical dynamics. Princeton University Press.

714 Schubel, J. R., and V. S. Kennedy. 1984. The estuary as a filter: An introduction, p. 1–11. *In*
715 V.S. Kennedy [ed.], The estuary as a filter. Elsevier.

716 Staehr, P. A., J. M. Testa, W. M. Kemp, J. J. Cole, K. Sand-Jensen, and S. V. Smith. 2011. The
717 metabolism of aquatic ecosystems: history, applications, and future challenges. Aquatic Sciences
718 **74**: 15–29. doi:10.1007/s00027-011-0199-2

719 Tassone, S. J., and P. A. Bukaveckas. 2019. Seasonal, interannual, and longitudinal patterns in
720 estuarine metabolism derived from diel oxygen data using multiple computational approaches.
721 Estuaries and Coasts **42**: 1032–1051. doi:10.1007/s12237-019-00526-0

722 Thébault, J., T. S. Schraga, J. E. Cloern, and E. G. Dunlavey. 2008. Primary production and
723 carrying capacity of former salt ponds after reconnection to San Francisco Bay. Wetlands **28**:
724 841–851. doi:10.1672/07-190.1

725 Wanninkhof, R. 2014. Relationship between wind speed and gas exchange over the ocean
726 revisited. Limnology and Oceanography: Methods **12**: 351–362. doi:10.4319/lom.2014.12.351

727 Winslow, L. A., J. A. Zwart, R. D. Batt, H. A. Dugan, R. I. Woolway, J. R. Corman, P. C.
728 Hanson, and J. S. Read. 2016. LakeMetabolizer: an R package for estimating lake metabolism
729 from free-water oxygen using diverse statistical models. Inland Waters **6**: 622–636.
730 doi:10.1080/iw-6.4.883

731 Yvon-Durocher, G., J. M. Caffrey, A. Cescatti, and others. 2012. Reconciling the temperature
732 dependence of respiration across timescales and ecosystem types. Nature **487**: 472–476.
733 doi:10.1038/nature11205

734 Acknowledgments

735 We thank Jason Garwood and Ethan Borque of the Apalachicola National Estuarine Research

736 Reserve for assisting with the deployment and processing of data used herein. We also thank

737 Wade McGillis for assisting in this effort. We thank Darren Giling for initial conversations on

738 applying BASE to estuarine data. This research was funded by award 1924559 from the National

739 Science Foundation to co-PIs Raymond Najjar, Maria Herrmann, and Kathleen Hill.

740

741 # Figure Legends

Figure 1: The default prior distributions for a, R, and b used in EBASE. All priors are normal distributions and truncated to positive values. b is also truncated to a maximum of 0.502 $(cm/hr)/(m^2/s^2)$*. Probabilities have been normalized to the maximum probability.*

Figure 2: Synthetic time series for one year of continuous data at Apalachicola Bay. Synthetic data were created using the (a) a parameter to estimate (b) metabolic estimates for gross primary production (P), ecosystem respiration (R), and gas exchange (D) and the (c) dissolved oxygen time series (including saturation, C_{sat}*).*

742

Figure 3: Comparison of EBASE results using default prior distributions to a synthetic time series created with known model parameters (a, R, and b). Results are compared for EBASE parameters estimated with a (a) 7 day and (b) 30 day optimization period. 95% credible intervals from the posterior distributions of the parameter estimates from EBASE are also shown. The results are averaged at the time step of the model optimization period.

743

Figure 4: Comparison of EBASE results to a synthetic time series created with known model parameters (a, R, and b). Results are compared for unique combinations of prior distributions and optimization period. The means (μ) and standard deviations (σ) are evaluated as low (L) or high (H) values (see text for details) and the optimization period for EBASE is 1 day, 7 days, or 30 days. The model predictions of dissolved oxygen (DO), gross primary production (P), ecosystem respiration (R), air–water oxygen flux (D), and light efficiency (a) are evaluated based on Nash-Sutcliffe Efficiency (NSE) values between the EBASE results and synthetic values. The b parameter is a fixed constant and was not evaluated. Rankings of each combination as the median NSE value across all five model outputs in a row are shown on the right, with the top three as black/bold and bottom three as black/italic.

744

Figure 5: Comparison of EBASE results for the best and worst performing models to a synthetic time series created with known metabolic parameters. The best and worst performing models for the 1 day (a, b), 7 day (c, d), and 30 day (e, f) optimization periods are based on the EBASE results with prior distributions that produced the highest and lowest median Nash-Sutcliffe Efficiency values across all parameter estimates (Figure 4). The results are averaged at the time step of the model optimization period.

745

Figure 6: A subset of the synthetic time series (Figure 2c) and the synthetic time series with noise for evaluating the ability of EBASE to recover known metabolic parameters. The (a)

estimated residual and tidal noise added to the (b) synthetic time series are shown.

746

Figure 7: Comparisons of EBASE to recover parameters a (top) and R (bottom) from a synthetic time series and a synthetic time series with residual and tidal noise included. The left plots show the time series and the right plots show a 1:1 comparison for each result. See Figure 6 for subsets of both dissolved oxygen time series.

747

Figure 8: Comparisons of daily metabolic estimates from the EBASE and Odum methods for one year of continuous monitoring data at Apalachicola Bay. Results for (a) gross primary production (P), (b) ecosystem respiration (R), and (c) net ecosystem metabolism (NEM) are compared using the observed and detided dissolved oxygen time series as input to each model. Black lines show the 1:1 comparison and the red lines show the linear regression fit.

748

Tables

Table 1: Summary statistics of daily metabolic estimates comparing the EBASE and Odum methods for gross primary production (P), ecosystem respiration (R), and net ecosystem metabolism (NEM) using the observed (left values) and detided (right values) dissolved oxygen time series as input to each model. Summary statistics include Pearson correlation coefficients and root mean square difference (RMSD) between the metabolic estimates obtained from each method. P-values for correlation coefficients all less than 0.005.

Parameter	ρ	RMSD (mmol $O^2/m^2/d$)
P	0.7 / 0.71	76.17 / 14.9
R	0.51 / 0.49	98.76 / 27.29
NEM	0.36 / 0.52	42.28 / 17.98

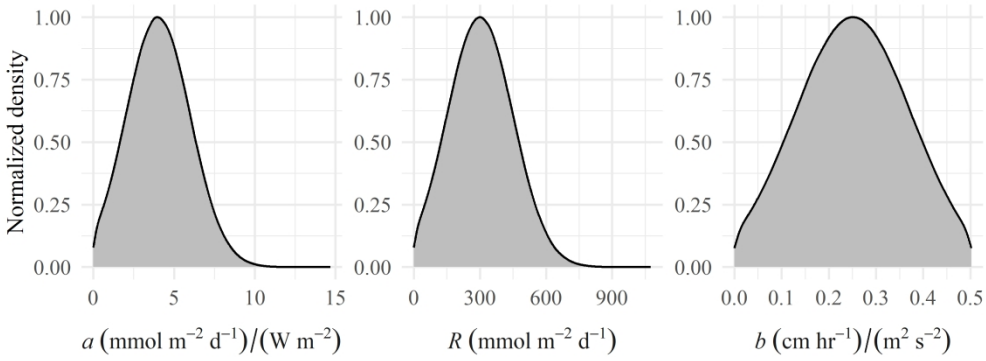

Figure 1

419x161mm (197 x 197 DPI)

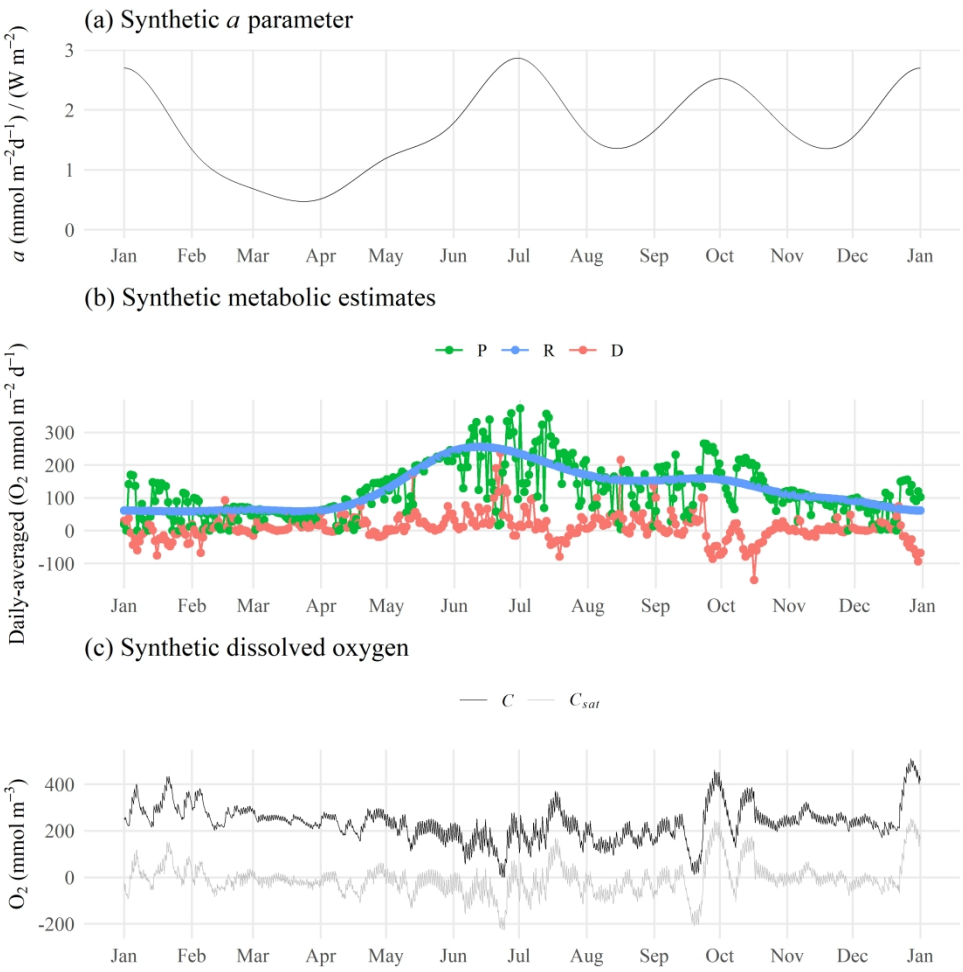

Figure 2

451x451mm (197 x 197 DPI)

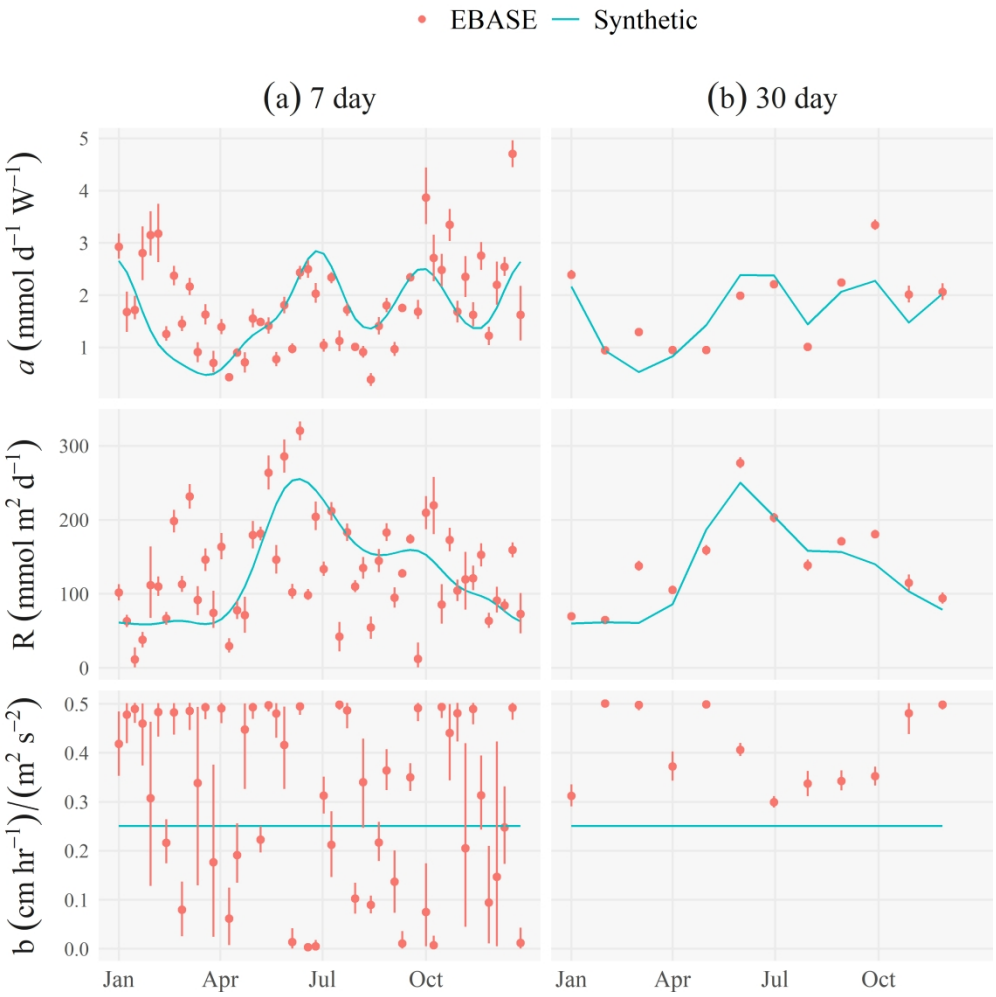

Figure 3

451x451mm (197 x 197 DPI)

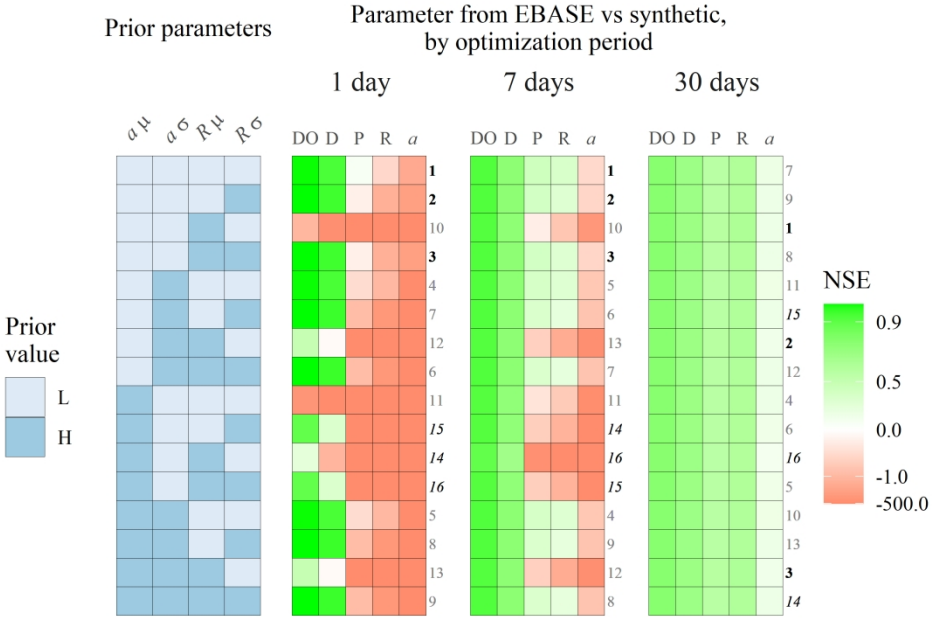

Figure 4

386x257mm (197 x 197 DPI)

Figure 5

580x515mm (197 x 197 DPI)

Figure 6

515x386mm (197 x 197 DPI)

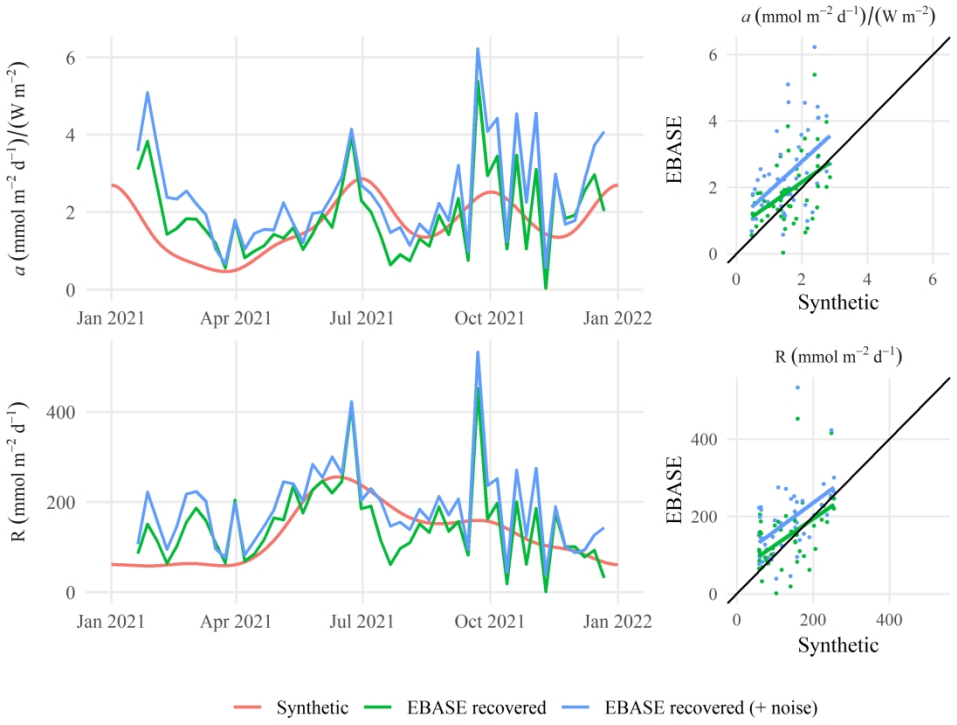

Figure 7

451x354mm (197 x 197 DPI)

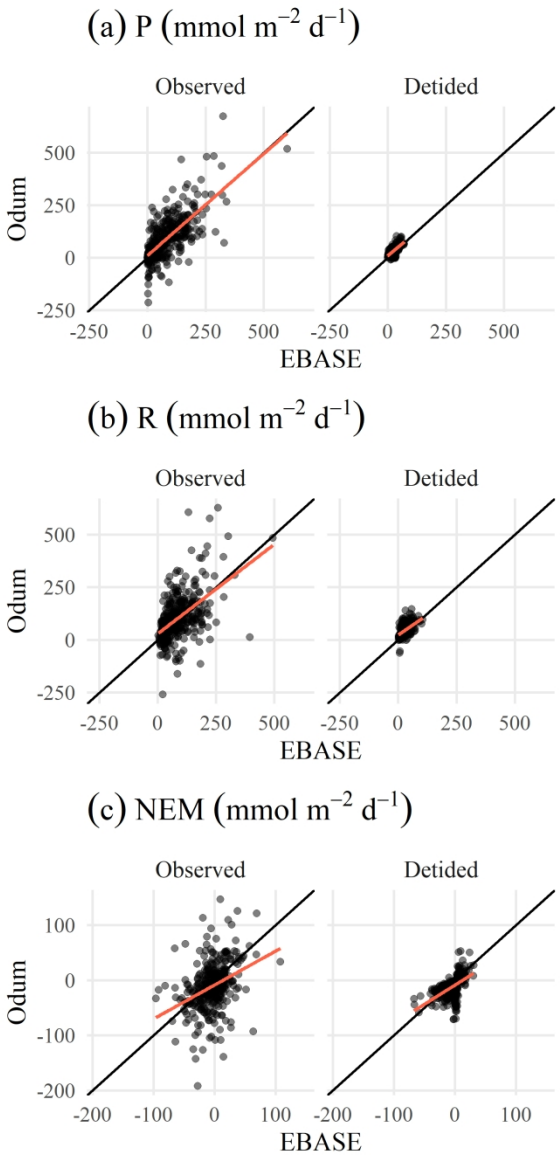

Figure 8

225x451mm (197 x 197 DPI)