**Improving estimates of ecosystem metabolism by reducing effects of tidal advection on dissolved oxygen time series**

**Marcus W. Beck1, James D. Hagy III2, Michael C. Murrell3**

**1***ORISE Research Participation Program*

*USEPA National Health and Environmental Effects Research Laboratory*

*Gulf Ecology Division, 1 Sabine Island Drive, Gulf Breeze, FL 32561*

*Phone: 850-934-2480, Fax: 850-934-2401, Email:* [*beck.marcus@epa.gov*](mailto:beck.marcus@epa.gov)

**2***USEPA National Health and Environmental Effects Research Laboratory*

*Gulf Ecology Division, 1 Sabine Island Drive, Gulf Breeze, FL 32561*

*Phone: 850-934-2455, Fax: 850-934-2401, Email:* [*hagy.jim@epa.gov*](mailto:hagy.jim@epa.gov)

**3***USEPA National Health and Environmental Effects Research Laboratory*

*Gulf Ecology Division, 1 Sabine Island Drive, Gulf Breeze, FL 32561*

*Phone: 850-934-2433, Fax: 850-934-2401, Email:* [*murrell.michael@epa.gov*](mailto:murrell.michael@epa.gov)

*Key words: Dissolved Oxygen, Ecosystem Metabolism, Open Water Method, Tides, Time Series, Weighted Regression*

*Running head: Improving Estimates of Estuary Metabolism*

***Abstract***

In aquatic systems, time series of dissolved oxygen (DO) have been used to compute estimates of ecosystem metabolism. Central to this open water method is the assumption that the dissolved oxygen time series is a Lagrangian specification of the flow field. However, most DO time series are collected at fixed locations, such that changes in dissolved oxygen are assumed to reflect metabolism and that effects of advection or mixing are negligible. A weighted regression model was applied to remove variability in DO time series from tides, thereby helping to partially relax this assumption and improve metabolism estimates. The method offers a distinct advantage over traditional deconvulution methods by targeting the periodicity of the tidal component while preserving the true biological signal. The model was first applied to simulated DO time series with specified biological and physical characteristics, and then applied to one year of continuous monitoring data from four stations within the National Estuarine Research Reserve System. The correlation of DO and metabolism estimates with tides was greatly reduced after using weighted regression. The model was especially effective when the magnitude of tidal influence was high and correlations between tidal change and the solar cycle were low at the time scales of interest. The model was less robust when tides and the solar cycle were correlated for protracted periods. By reducing the effects of physical transport on metabolism estimates, there may be increased potential to empirically relate metabolic rates to causal factors on timescales of several days to several weeks.

***Introduction***

Ecosystem metabolism in aquatic environments describes the balance between production and respiration processes that create and consume organic matter (Kemp and Testa 2012; Needoba et al. 2012). Experimental light and dark exposure experiments of water samples collected at discrete locations and times have traditionally been used to measure pelagic metabolic activity (Gaarder and Gran 1927). Although controlled and precise, bottle-based methods are labor-intensive and rarely sufficient to describe entire ecosystem rates. In contrast, open-water techniques have the potential to estimate whole system metabolism by use of continuous time series of dissolved oxygen (Odum 1956; D’Avanzo et al. 1996). Open-water estimates also provide a practical approach to capturing events and ecosystem changes over time because DO monitors may be deployed continuously and for sustained periods. While bottle-based measurements offer advantages in precision, it is not possible to sustain bottle measurements day after day, making it unlikely that sampling will occur during an unanticipated short-term event. Moreover, because more variability is captured in continuous data, statistical power for resolving long-term changes is increased. Although metabolic rates vary naturally at different spatial and temporal scales (Ziegler and Benner 1998; Caffrey 2004; Russell and Montagna 2007), anthropogenic nutrient sources are often contributing factors that increase rates of production (Nixon 1995; NRC 2000). Reliable estimates of whole ecosystem metabolism are critical for measuring both background rates of production and respiration, and to assess the potential effects of human activities on ecosystem function.

The accuracy of ecosystem metabolism measurements using the open-water method depends on the degree to which assumptions of the theory are satisfied (Staehr et al. 2010; Kemp and Testa 2012). A fundamental assumption is that the time series of dissolved oxygen (DO) represents a Lagrangian specification of the flow field (Needoba et al. 2012). The Lagrangian specification characterizes an individual fluid parcel through time, such as a parcel of water moving with the tide. However, most DO time series are an Eulerian specification of the flow field given that they are collected at fixed locations such as a mooring or dock. Time series at fixed locations may characterize water masses with different metabolic histories as water is advected past the sensor, leading to errors in metabolism estimates (Kemp and Boynton 1980; Russell and Montagna 2007). Despite this problem, the open-water method has been used with varying success in estuaries influenced by tidal transport and mixing (Caffrey 2004; Russell and Montagna 2007; Caffrey et al. 2014). The method has been more successfully applied to water quality time series in lakes, although stratification may limit estimates to the epilimnion (Staehr et al. 2010; Coloso et al. 2011; Batt and Carpenter 2012). Successful application of open water methods requires careful consideration of the placement of monitoring sondes, sampling frequency, deployment and duration, and reliability of extrapolating ecosystem rates from one or few fixed locations (Russell and Montagna 2007; Staehr et al. 2010). Sites located near bay inlets or along major tidal axes may produce DO time series that are strongly influenced by advection, leading to relatively large errors when using the open-water method.

Quantitative approaches have not been sufficiently developed to reduce the potential bias or noise in metabolism estimates resulting from advection. Despite this limitation, the open-water method has been applied in estuaries that are strongly influenced by tidal transport and mixing (Caffrey 2003; Russell and Montagna 2007; Caffrey et al. 2014; Nidzieko et al. 2014). For example, an extensive analysis by Caffrey (2003) applied the open-water method to 28 continuous monitoring stations at 14 US estuaries. A significant portion of the production and respiration estimates were negative (3 - 69% depending on site), suggesting that advection of water masses with different DO histories may be influencing the observed time series. Similarly, Collins et al. (2013) attributed nighttime DO increases in Long Island Sound to tidal advection. Further, Nidzieko et al. (2014) evaluated the effects of tidal advection on metabolism estimates in a mesotidal estuary using multiple sensors and a well-defined control volume model. They observed that metabolism estimates were strongly correlated with the spring-neap cycle such that net heterotrophy was more common during spring tides and balanced metabolism occurred during neap tides. This example illustrates that tides can modulate true metabolism in addition to causing variability in the DO time series that is unrelated to metabolism. Although useful as an in situ, site-specific approach, more broadly applicable statistical methods are needed given the increasing availability of continuous monitoring data. As one example, Batt and Carpenter (2012) explored use of a Kalman filter (Harvey 1989) to remove process and observation uncertainty from DO time series in lakes. Similar approaches have not been developed for estuaries and none specifically address the effects of tides.

This article describes the development and application of a method for improving estimates of ecosystem metabolism computed by reducing the effect of tidal advection on continuous DO observations. We used a weighted regression approach originally developed to resolve trends in pollutant concentrations in streams and rivers (Hirsch et al. 2010). The weighted regression approach creates dynamic predictions of DO as a function of time and tidal height, which are then used to filter, or detide, the DO signal. The model is based on the recognition that daily fluctuations in DO caused by metabolism are associated with the solar cycle, whereas other fluctuations in estuaries are likely associated with cyclical water level changes that do not have the same period or phase as the solar cycle. The weighted regression model was applied, rather than methods commonly used for detiding in physical oceanography, to allow for the complex and dynamic patterns of DO changes relative to advection. The method targets the tidal component as an explicit variable using dynamic model parameters that change through time. This approach makes it possible to estimate and reduce the effect of advection on the DO time series, even if this effect varies through time due to changes in metabolic rates. Local fitting provides an advantage relative to traditional deconvulution methods. The method is most useful in estuaries where tidal effects interact with horizontal differences in metabolism and associated DO concentrations. Although the method could be applied in open ocean environments, it may be less useful because horizontal DO gradients may be less in such systems, such that it could be sufficient to simply neglect advection as has commonly been done in the past. During initial method development, we used simulated DO time series with known characteristics to evaluate the ability of the weighted regression to remove the simulated effects of a tidally-advected DO gradient. Subsequently, the simulation results informed further development of the method as applied to four case studies chosen from the National Estuarine Research Reserve System (NERRS, Wenner et al. 2004). In all examples, tidal height is used as a proxy for DO changes resulting from advection because it was recorded with the NERRS data.

***Materials and Procedures***

***Weighted regression for modelling and filtering DO time series***

We adapted a weighted regression model to filter DO time series for apparent tidal effects. This model adapted the weighted regression on time, discharge, and season (WRTDS) method, which was developed for estimating pollutant concentrations in streams and rivers (Hirsch et al. 2010). As applied in this study, the functional form of the model is:

 (1)

where *DOobs* is a linear function of time *t* and tidal height *H*. Time is a continuous variable for the day and time of each observation as a proportion of the total number of observations. The beginning of each day was defined as the nearest thirty minute observation to sunrise for the location. Our model differed from the original WRTDS method, which included parameters to estimate variation of the response variable on a sinusoidal period. DO variation was not modeled using this approach because the expected metabolically-driven pattern is not strictly sinusoidal.

Tidal height is used as a proxy for advection because this measurement is widely available, easy to measure, and meets the model requirements, which is a variable that is correlated locally in time with DO variations attributable to advection. Importantly, tide height need not predict the actual magnitude of advection. This is important because, across estuarine locations, the relationship between tide height and tidal currents can be variable in both magnitude and phase, depending on local characteristics. The weighted regression model is well-suited for the situation in estuaries because it empirically fits the local relationship between tide height and DO, such that no a priori knowledge is needed to define the relationship between tide height and either the magnitude of advection or the horizontal gradient in DO.

Weighted regression was implemented as a moving window that allowed for estimation of DO throughout the time series by adapting to variation through time as a function of tide. Regression models were estimated sequentially for each observation in the time series using dynamic weight vectors that change with the center of the window. Weight vectors quantified the relevance of observations to the center of the window in respect to time, hour of the day, and tidal height. Weights were assigned to each variable using a tri-cube weighting function (Tukey 1977; Hirsch et al. 2010):

 (2)

where the weight *w* of each observation is inversely proportional to the distance *d* from the center of the window such that observations more similar to the point of reference are given higher weight in the regression. Observations at distances further than the maximum width of the window *h* have a weight of zero. The tri-cube weighting function is similar to a Gaussian distribution such that weights decrease gradually from the center until the maximum window width is reached. Regressions that use simpler windows (e.g., boxcar approach) are more sensitive to influential observations as they enter or leave the window, whereas the tri-cube function minimizes their effect through gradual weighting of observations from the center (Hirsch et al. 2010). The final weight vector for each observation is the product of the three separate weight vectors for time (day), hour, and tidal height. Windows for time (continuous throughout the time series) and hour (within each day) are used to weight observations based on distance (time) from the center of the window. The window for tidal height is used to weight observations based on the difference from mean tide as a proportion of tide range. For example, a half-window width of 0.5 means that observations are weighted proportionately within ±50% of the tide range referenced to the tidal height in the center of the window. A low weight is given to an observation if any of the three weighting values were not similar to the center of the window since the final weight vector is the product of the weight vectors for each variable (see the link in the multimedia section for graphical display of different weights).

The choice of window widths for weight vectors affects the model results; therefore, weights were selected via a rational and replicable method. Window widths that minimize prediction error are typically smaller than widths that should be used to minimize tidal effects. Similarly, window widths that more effectively filter the DO signal may produce less precise predictions for the observed data. Evaluations of the weighted regression method with simulated DO time series, described below, used multiple window widths to evaluate the ability of the model to filter the DO signal. The ability to predict observed DO was not the primary objective. Rather the window widths were evaluated based on their ability to remove tidal variation from the DO time series.

The approach to minimize advection effects in the observed DO time series differs slightly from Hirsch et al. (2010), who used a two-dimensional grid predicted for stream pollutant concentrations across the time series and the range of discharge values observed in the study system. Normalized or discharge-independent values for pollutant concentration were obtained by averaging grid predictions across the discharge values that were likely to occur on a given day. Rather than creating a two-dimensional grid of DO related to time and tidal height change, the normalized time series herein were the model predictions conditional on time and constant tidal height set to the mean:

 (3)

such that the normalized time series represents DO variation related to biological processes. The term ‘filter’ is used in reference to the removal of a specific variance component from the time series, while maintaining the structure of the biological component. Although the approach shares similarities with common filtering techniques, a distinction is noted such that weighted regression has a specific purpose that differs from those of filters (e.g., moving window averages or local smoothers, Shumway and Stoffer 2011).

***Assessment***

***Simulation of DO time series***

To test the ability of the weighted regression to filter the DO signal for apparent tide effects, multiple time series with known characteristics were simulated and filtered. A simulation approach was used prior to application with real data because the true biological signal can be specified as a known component for comparison with the filtered results from weighted regression. The following describes the theoretical basis for developing the simulated time series. Observed DO time series (*DOobs*) were simulated as the sum of variation from biological processes (*DObio*) and physical effects related to tidal advection (*DOadv*):

 (4)

Each component of *DOobs* was simulated separately. The biological DO signal (*DObio*) was the sum of diel variation (*DOdie*) plus uncertainty or noise (*DOunc*):

 (5)

 (6)

Biological DO signals are inherently noisy (Batt and Carpenter 2012) such that uncertainty in the biological DO signal was described as variation from observation and process uncertainty (*εobs* and *εpro*, Hilborn and Mangel 1997). Observation uncertainty is variation related to imprecision in sampling processes or methods, whereas process uncertainty is caused by unknown parameters that contribute to variance in a time-dependent fashion. For example, wind speed affects air-sea gas exchange (Ziegler and Benner 1998; Caffrey et al. 2014) such that periods of very high or low wind speed are a source of increased process uncertainty. Although this was not an explicit focus of the simulation analyses, wind effects could be considered an implicit component of process uncertainty in addition to the effects of other unmeasured or latent variables that influence DO in a time-dependent manner. Multiple 30 day time series were created with 30 minute time steps by varying the relative magnitude of each component of observed DO in equations (4) to (6) and used to test the effectiveness of weighted regression under different scenarios. Accordingly, observed DO was generalized as the sum of four separate time series (Fig. 1):

 (7)

Each component of the simulated time series was created as follows. The diel component, *DOdie*, was estimated as (Cryer and Chan 2008):

 (8)

such that the mean DO (*α*) was 8, amplitude (*β*) was 1, *f* was 1/48 to represent 30 minute intervals, *t* was the time series vector and Φ was the x-axis origin set for an arbitrary sunrise at 0630. The diel signal was increasing during the day and decreasing during the night for each 24 hour period and ranged from 7 to 9 mg L−1 (as a rough approximation from the case studies below). Although the representation of *DOdie* as a simple sine wave ignores more complex dynamics associated with the photosynthesis-irradiance relationship as applied to an ecosystem, it is sufficient to address the analytical objectives of the simulations. Uncertainty was added to the diel DO signal as the sum of observation and process uncertainty:

 (9)

where observation and process uncertainty (*εobs*, *εpro*) were simulated as normally distributed random variables with mean zero and standard deviation varying from zero to an upper limit, described below. Process uncertainty was estimated as a serially correlated variable using the cumulative sum of *n* observations plus random variation added at each time step for *t* = 1,…,*n*. The total uncertainty, *DOunc*, was added to the diel DO time series to create the biological DO time series (eq. (5) and Fig. 1).

A tidal series was simulated based on the principal lunar semidiurnal (M2) tide with a period of 12.42 hours (Foreman and Henry 1989). The amplitude was set to 1 meter and centered at 4 meters. The tidal time series simulated DO changes with advection, *DOadv* (eq. (7) and Fig. 1). Conceptually, this vector reflects both water movement and the horizontal DO gradient, which can be expressed as:

 (10)

where  is the DO gradient in horizontal space and  is horizontal displacement of water over time. The simulated tidal time series was used to create *DOadv* as a simple sine wave as in eq. (8). This time series was centered at zero to simulate increasing or decreasing DO with tidal height change. The final time series for observed DO was the sum of biological DO and advection DO (eq. (4) and Fig. 1). For time series from the NERRS monitoring sites, variations in DO associated with advection were modeled as proportional to the difference between observed and mean tide height as indicated by the parameter *β2* in eq. (1). Thus, we can express the tidal component of DO variations (*DOadv*) simply as:

 (11)

where *H* is the difference between observed and mean tide height.

***Evaluation of weighted regression with simulated DO time series***

Multiple time series were simulated by varying the conditions in eq. (7) (Fig. 2) to evaluate weighted regression under a range of conditions. Specifically, the simulated data varied in the relative amount of noise in the measurement (*εobs*, *εpro*), relative amplitude of the diel DO component (*DOdie*), and degree of association of the tide with the DO signal (*DOadv*). Three levels were evaluated for each variable: relative noise as 0, 1, and 2 standard deviations for both process and observation uncertainty, amplitude of diel biological DO as 0, 1, and 2 mg L−1, and DO change from tidal advection as 0, 1, and 2 mg L−1. A total of 81 time series were created based on the unique combinations of parameters (Fig. 2). Half-window widths (day, hour of day, and tide height) for the weighted regressions were evaluated for each time series: time as 1, 3, and 6 days, time of day as 1, 3, and 6 hours, and tidal height as 0.25, 0.5, and 1 as a proportion of the total range given the height at the center of the window. In total, 27 window width combinations were evaluated for each of 81 simulated time series, producing results for 2,187 weighted regressions.

The filtered DO time series were compared to the simulated data to evaluate the ability of weighted regression to quantify and remove the variability due to advection (*DOadv*), leaving the biological response (*DObio*) (eq. 4); Fig 3). Comparisons were made using Pearson correlation coefficients and the root mean square error (RMSE). The ability of weighted regression to isolate *DObio*, indicated by high correlation and low RMSE, was not reduced by the introduction of advection into the simulated time series (i.e., *DOobs*) because weighted regression was able to quantify and remove this variability. On the other hand, introduction of observation uncertainty (*obs*) decreased the ability of the model to quantify and remove advection. Perhaps counterintuitively, increased magnitude of advection improved the model performance because a larger advection signal can be fitted empirically despite the presence of random error. When the advection signal was small, it became difficult or impossible to quantify unless both process uncertainty and observation error were absent. This suggests that if the advection signal is in fact very small, it would be preferable not to quantify it and instead neglect it, as has been done previously. Model performance was not substantially affected by variation in half window widths relative to characteristics of the DO time series (Fig. 4).

***Validation of weighted regression with case studies***

Results from the simulated time series were used to inform the validation of weighted regression with real data, especially with respect to choosing half-window widths described below. Continuous monitoring data from NERRS was used to validate the weighted regression model by evaluating estimates of ecosystem metabolism obtained from observed and filtered DO time series. NERRS is a network of 28 shallow, productive estuary reserves established for long-term research, water-quality monitoring, education, and coastal stewardship (Wenner et al. 2004). Continuous water quality data have been collected at NERRS sites since 1994 through the System Wide Monitoring Program (SWMP, CDMO website, <http://cdmo.baruch.sc.edu>, accessed January, 2014). In addition to providing a basis for trend evaluation, data from SWMP provides an ideal opportunity to evaluate long-term variation in water quality parameters from biological and physical processes. Continuous SWMP data can be used to describe DO variation at sites with different characteristics, including variation from differences in tidal regime (Sanger et al. 2002) and rates of ecosystem production (Caffrey 2003, 2004). We selected sites from the SWMP database that had desirable characteristics for validating weighted regression. Four macrotidal sites were chosen based on apparent relationships between DO and tidal changes (Fig. 5 and Table 1). Many daily metabolism estimates from these sites were not interpretable (i.e., ‘anomalous’) when using standard open water methods. In other words, the advection effects on DO time series at these sites are very large.

The case study sites included Vierra Mouth station at Elkhorn Slough (ELKVM, California, 36.81◦ N, 121.78◦ W), Bayview Channel at Padilla Bay (PDBBY, Washington, 48.50◦ N, 122.50◦ W), Middle Blackwater River station at Rookery Bay (RKBMB, Florida, 25.93◦ N, 81.60◦ W), and Dean Creek station at Sapelo Island (SAPDC, Georgia, 31.39◦ N, 81.28◦ W). The ELKVM station is located at the mouth of the Elkhorn Slough estuary in approximately four meters of water (MLW, tidal range 2.0). Sediments are compacted mud and sand due to large tidal currents. The PDBBY station is located along the Bayview Channel, which is a major tributary of Padilla Bay that drains vegetated intertidal flats characterized by eelgrass and macroalgae. The site is in 1.5 meters of water (MLLW, tidal range 2.9) with sediments of fine silt and clay overlying sand. The RKBMB station is located at the mouth of the Middle Blackwater River in approximately 2 meters of water (MHW, tidal range 1.3). Salinity is oligohaline (2.3 ppt) to polyhaline (38.6 ppt). Sediments are a mixture of sand, silt, oyster shell, and organic matter. Finally, the SAPDC station is located in Dean Creek, a small tidal basin that is fed from the Doboy Sound on the south end of Sapelo Island. Mean low water depth is approximately one meter (tidal range ~3 m) and sediments consist of sand and mud with occasional oyster reefs.

The weighted regression model was applied to continuous DO time series and water level measurements from January 1 to December 31, 2012 at the four sites. NERRS sites are typically shallow and vertically-mixed such that one water quality monitor adequately characterizes DO for the entire water column. Tide predictions were obtained for each site using harmonic regression applied to the sonde depth data (oce package in R, Foreman and Henry 1989; Kelley and Richards 2014; RDCT 2014). The predicted time series of tidal height from the regressions were used for all models to avoid issues with missing values in the observed data. Tides at the study sites were semidiurnal or mixed semidiurnal. Metabolism was net heterotrophic on an annual basis (Table 1).

***Estimates of ecosystem metabolism before and after filtering***

The weighted regression method was applied to the annual data for each station to obtain a filtered DO time series for estimating metabolism. Ecosystem metabolism was estimated using the open-water technique (Odum 1956) as described in Caffrey et al. (2014). The method is used to infer net ecosystem metabolism using the mass balance equation:

 (14)

where the change in DO concentration (*DO*, g O2 m−3) over time (*t*, hours) is equal to photosynthetic rate (*P*, g O2 m−3 hr−1) minus respiration rate (*R*, g O2 m−3 hr−1), corrected for air-sea gas exchange (*D*, g O2 m−3 hr−1) (Caffrey et al. 2014). D is estimated as the difference between the DO saturation concentration and observed DO concentration, multiplied by a volumetric reaeration coefficient, ka (Thébault et al. 2008). This coefficient accounts for the influence of meteorological effects, such as wind, on air-sea gas exchange at the surface. The diffusion-corrected DO flux estimates as hourly rates of DO change were first averaged during day and night periods for each 24 hour ‘metabolic day’ in the time series. The ‘metabolic day’ was considered the period between sunsets on two adjacent calendar days. Respiration was hourly DO flux during night hours and net production was hourly DO flux during day hours. Total respiration (Rt) rates were assumed constant during day and night such that daily rates were calculated as the average DO flux during night hours multiplied by 24. Daily gross production (Pg) was the average DO flux during day hours minus the average respiration, multiplied by total sunlight time. Net ecosystem metabolism was gross production (positive) plus total respiration (negative). Finally, volumetric rates were converted to depth-integrated (areal) estimates (g O2 m−2 d−1, Table 1) by multiplying by the mean water-column depth at the site. Site depth was estimated from the pressure sensor on the water quality sonde, with a half-meter added to account for placement of the sonde just above the bottom.

The selection of half-window widths for filtering the DO time series was based on an evaluation of results using four performance metrics. The ‘optimal’ window widths for each case study were those that provided the best performance on all four metrics. First, the regression results were evaluated using correlations of DO and metabolism estimates with tidal height before and after application of the model. The ideal window widths are those that provided the maximum reduction in correlation relative to the observed data. In these analyses, daily metabolism estimates before and after filtering were compared with the mean rate of tidal height change (i.e., first derivative of the predicted tidal height) during day and night. Production rates were compared to mean tidal height change during the day and respiration rates were compared to mean change at night. The second and third performance metrics evaluated changes in annual mean and standard deviation of metabolism estimates before and after filtering. We assumed that annual mean values would not change after filtering because long-term averages of metabolism are not likely to be affected by advection. We evaluated this assumption by computing metabolism for a 12 month simulated DO time series with and without advection. These simulations showed that the means converge within the annual time scale, whereas variance remains substantially larger for the time series with advection even at the annual time scale. Optimum window widths in this context were considered those that maintained the mean values while reducing standard deviation relative to the observed data. Finally, results were evaluated based on the occurrence of ‘anomalous’ daily production or respiration estimates, where anomalous was defined as negative gross production estimates or positive respiration estimates. Anomalous values have been previously attributed to the effects of physical processes on DO time series (Caffrey 2003). Optimum window widths for this metric were those that provided the maximum reduction in anomalous values. Although anomalies could be caused by processes other than tidal advection, e.g., abiotic dark oxygen production (Pamatmat 1997), we assumed that physical processes were the dominant cause of these values given the tidal characteristics at each site.

Multiple combinations of half-window widths were evaluated for the case studies. These included half-window widths of 1, 3, 6, 9, and 12 days, 1, 3, 6, 9, and 12 hours, and 0.2, 0.4, 0.6, 0.8, and 1 for tidal height proportion, producing a total of 125 unique combinations. Half-window widths that maximized the four performance metrics for each case study were chosen as the ‘optimal’ values. Accordingly, the optimal half-window widths were 12, 6, and 0.8 (days, hours, tidal height) for Elkhorn Slough, 3, 6, and 0.6 for Padilla Bay, 3, 1, and 0.6 for Rookery Bay, and 3, 1, and 0.6 for Sapelo Island. In the fitted models, DO was positively correlated with predicted tidal height at all sites except Padilla Bay, where increases in water level were associated with decreases in DO concentration. Filtering nearly eliminated the correlations between water level changes and the DO time series and, similarly, dramatically reduced correlations between metabolism estimates and tidal height changes (Table 2, see the link in the multimedia section for graphical results of each case study).

The percentage of daily integrated metabolism estimates that were anomalous (negative production, positive respiration) were reduced for all sites after filtering (Table 3), perhaps indicating the relative effects of water movement. Before filtering, anomalous values ranged from 9.15 (as a percentage of the total estimates, Rookery Bay) to 21.80 (Padilla Bay) for production and 7.57 (Rookery Bay) to 20.68 (Elkhorn Slough) for respiration. Anomalous values were noticeably reduced for all case studies and were virtually eliminated in two of the four case studies. Metabolism estimates using filtered DO time series had decreased mean production (-63.1 % change from the annual mean) and respiration (-62.6 %) for Elkhorn Slough, increased mean production (17.8 %) and respiration (18.8 %) for Padilla Bay, and generally unchanged mean production and respiration for Rookery Bay and Sapelo Island (Table 3). Changes in mean estimates based on filtered DO time series suggests that the weighted regression removed variation attributed to both biological and physical processes. The implications of these undesirable results are described below. Decreases in the standard deviation for all metabolism estimates were observed for all cases after filtering. However, reduction in variation may also occur if the biological signal is reduced individually or in addition to physical variation. The extent to which this reduction is related to the former should be minimal, provided that the two are statistically distinguishable. Conversely, when the tidal and solar cycles are correlated at time scales of interest (e.g., weeks to months), it may not be possible to separate tidal effects using this model. The observed reduction in annual mean values, especially at Elkhorn Slough, suggests that in that case, biological effects were attributed to tides, causing a downward bias in estimated metabolism.

An example from Sapelo Island illustrates the effects of weighted regression on DO and metabolism estimates (Figs. 6 to 8). During a two-week period in February the tide changed from neap tides with relatively pronounced differences between the first and second high and low tides to spring tides with a larger tidal amplitude and smaller differences between first and second tides. During Feb 4-7, tide levels at sunset were lower than at sunrise, causing an apparent net decrease in DO during the day. Tide levels at sunrise were higher than at sunset, causing an apparent net increase in DO at night (Fig 6). These differences led to large errors in metabolism (Fig. 8). In the subsequent week, tides also had a dramatic effect on observed DO, but a much smaller effect on estimates of metabolism, which were slightly inflated by the tidal effect. In both weeks, weighted regression was highly effective for removing the advection signal from the DO time series (Figs. 6 and 7). Weighted regression also corrected the anomalous metabolism estimates in the latter portion of the first week, suggesting relatively stable rates through the two week period.

***Accuracy of results and effects of time averaging***

One question of interest is the period of observation within which observed DO is affected by tidal height changes and the extent to which this affects the interpretation of ecosystem metabolism. The effects of tidal variation on daily metabolism estimates may not be of concern if seasonal or annual means are of interest and averaging alone removes the noise that occurs on shorter time scales. The example from Sapelo Island highlights this point given that despite short term errors, mean production and respiration estimates before and after filtering were largely unchanged for the two-week period. Variance was substantially reduced, however. Alternatively, annual averages of production and respiration estimates were significantly different for Elkhorn Slough and Padilla Bay (Table 3). Therefore, evaluating the effects of applying weighted regression for different time scales of concern is critical to evaluating its potential use. Does the period of observation affect the ability of weighted regression to remove physical variation in the time series? When should filtering be applied if time averaging of observed data on longer time periods removes effects of advection? The first question is addressed by evaluating collinearity between tidal change and solar periods. The second question is addressed by comparing observed and filtered estimates that are averaged over different periods of observation.

Collinearity between tidal height change and the solar cycle affected the ability of weighted regression to model tidal advection effects at some locations. Nidzieko et al. (2014) found this condition in Elkhorn Slough during summer when high tides always occurred at night. To evaluate when this condition would present a problem for weighted regression, we computed for each case study the correlation between sun angle and tidal change (measured as an angular rate) using a moving window (Fig 9). Weighted regression can be expected to effectively separate biological and physical variation during periods when the correlation between tidal height change and the solar cycle is low within the window. Collinearity between sun angle and tidal height change exceeds ±0.2 on a seasonal basis in Elkhorn Slough and Padilla Bay, whereas correlations were much smaller at all times in Rookery Bay and Sapelo Island. Given the change in annual mean metabolism using the filtered DO time series and the relatively high collinearity between tidal change and solar cycling, the estimates of metabolism using weighted regression for Elkhorn Slough and Padilla Bay may have reduced variability due to tides, but still be in error. Results for these estuaries may only be interpretable for periods of the year when the correlation is low (e.g., April and October for Elkhorn Slough).

The observed and filtered daily estimates were averaged by month for each site to evaluate the effect of filtering on mean production and respiration at the monthly time scale (Table 3 and Fig. 10). Filtered production and respiration estimates for Padilla Bay and Rookery Bay exhibited monthly variation that was more characteristic of expected trends during warmer months than in the unfiltered estimates. Results for Sapelo Island suggested that time averaged estimates, either as monthly or annual means, were comparatively unchanged by filtering, although there were some more subtle differences seasonally and variance was reduced. Summer production and respiration for Elkhorn Slough were significantly reduced by filtering. It is likely that both the filtered and unfiltered estimates are erroneous due to collinearity between the tidal and solar cycle. Overall, these trends emphasize the importance of considering different averaging periods for interpreting metabolism estimates. Each case study showed differences in observed and filtered values for monthly means, suggesting tidal variation may influence metabolism estimates at relatively long time scales.

***Discussion***

The weighted regression approach was developed to improve estimates of ecosystem metabolism by removing variation associated with tides from DO time series. Analysis of simulated DO time series with known characteristics and subsequent extension to continuous monitoring data from selected NERRS sites suggested that the approach can in many circumstances successfully isolate and remove much of the tidal variation. In application to DO time series from NERRS sites, the method reduced apparent day-to-day variation in metabolic rates, suggesting that this variation could result substantially from artifacts of advection, rather than real drivers of metabolic processes. This hypothesis could be investigated further by applying weighted regression to DO time series while collecting contemporaneous measurements of metabolism and causal factors at a relatively high frequency. The regression method and open-water technique may have broad appeal for application in estuaries that are shallow and vertically well-mixed, such as those within the NERRS. Extension to stratified systems may be possible with additional data and model development.

Comparisons between filtered and biological DO time series from the simulations indicated that weighted regression can reduce the effects of tidal variation for a range of characteristics of DO time series. An examination of scenarios that produced abnormal results can provide additional insight into factors that affect the performance of weighted regression. For example, poor performance was observed when the observation uncertainty (*εobs*) was high and both process uncertainty (*εpro*) and tidal advection (*DOadv*) were low. These examples represent time series with excessive random variation (e.g., poor measurement quality), no auto-correlation, and no tidal influence. Poor performance is expected because the weighted regression is trying to model a minimal tidal signal in a very noisy DO time series. These results were observed even for time series with a large diel component of the biological DO signal, suggesting that the model may not be effective for microtidal systems with high noise and little serial correlation. From a practical perspective, weighted regression should not be applied to noisy time series if there is not sufficient evidence to suggest the variation is related to tides. Alternative approaches, such as the Kalman filter (Harvey 1989; Batt and Carpenter 2012), may be more appropriate if random variation is the primary source of uncertainty. Similarly, results with perfect or near-perfect correlations between filtered and biological DO time series were observed when observation uncertainty and tidal effects were not components of the simulated time series. Although there is no need to apply weighted regression to time series with no apparent tidal influences, the results will not be incorrect. Weighted regression will be most useful when applied to time series that meet certain conditions, as described below.

Weighted regression was successful in each of the case studies in reducing variation in the DO time series that was correlated with tides. Filtering reduced the frequency of anomalous metabolism estimates, such as negative production and positive respiration (Needoba et al. 2012). Filtering also reduced apparent errors in ‘normal’ estimates (i.e., positive production and negative respiration) that nonetheless may still include significant bias due to physical advection. For example, a large decrease in DO at night may appear ‘normal’ even if inflated by DO changes resulting from advection. In some cases, tidal processes may have real effects on ecosystem metabolism. For example, Nidzieko et al. (2014) observed that net metabolism at Elkhorn Slough was strongly heterotrophic when spring tides regularly flooded salt marshes on summer nights. Resolving these patterns may simply require better data. Similarly, Sasaki et al. (2009) showed that increased concentrations of organic substrates on falling tides could increase metabolic rates. Although one might hope that an improved ability to resolve temporal patterns in metabolism would make it possible to resolve such changes using only a DO time series, we suggest that this is still not possible. The open water metabolism method has generally been applied at the time scale of one day, rather than hourly. Weighted regression may improve estimates of daily metabolism, but will not make it possible to resolve intraday variations in metabolism. As long as there are not sustained periods of collinearity between solar and tidal cycles, weighted regression can be expected to improve daily estimates of average metabolism within a tidal excursion from the sensor location.

The weighted regression approach is effective and broadly applicable in part because it makes no assumptions as to the direction, magnitude, or stability of the relationship between DO and tidal variation over time. Although the functional form of the model is a simple linear regression with only two explanatory variables (eq. (1)), the moving window approach combined with the adaptive weighting scheme allows for quantification of complex tidal effects that may not be possible using alternative approaches. A similer approach by Batt and Carpenter (2012) uses a Kalman filter to improve estimates of ecosystem metabolism in lakes. The approach minimizes uncertainty in observed DO using a filter that combines information about the data generation process and the manner in which the data are observed (Harvey 1989). Although a similar approach could be used for estuaries, it may not be effective given that tidal advection may not be related to process or observation uncertainty as they are defined. Additionally, results from the case studies illustrated the ability of the weighted regression approach to model changes over time in the relationships between tidal change and DO. Results for Padilla Bay and Rookery Bay suggested that filtering had the largest effect during the summer, whereas the results for cooler months were not significantly different from the observed. The weighted regression method produced filtered time series that accommodated seasonal variation in DO conditional on tidal height change, whereas moving window filters or standard regression techniques would likely not have characterized these dynamic relationships.

***Comments and recommendations***

Results from the analysis of simulated data and case studies suggest that weighted regression can improve estimates of ecosystem metabolism by reducing the effects of tidal advection on DO time series. The weighted regression method dramatically reduced advection effects for some of the case studies, generating useful estimates of metabolism where otherwise the open water metabolism method would be subject to significant errors. Further analysis of the method illustrated the attributes of sites where it should be most useful and conversely, attributes of sites where problems may be encountered using the method. The weighted regression method will work best at sites that have a medium or large effect of advection on DO, but tidal changes are not correlated with the solar cycle for more than a few days. At sites with a dominant semidiurnal tide, such as at Sapelo Island, Georgia, the method can be expected to dramatically improve analysis using open water metabolism. Since this tidal regime occurs throughout the Atlantic coast of North America, Europe and Africa, as well as many other areas of the global coastline, this method should be broadly applicable, even if not universally applicable. Where tidal effects on DO time series are small, the method may reduce advection effects further, decreasing variance and increasing the potential to resolve drivers and changes in ecosystem metabolism. Where tides are large and strongly correlated with the solar cycle on seasonal time scales, the traditional open water metabolism method may be subject to very large errors. In these cases, however, the weighted regression attributed most DO variation to tides, resulting in large underestimates of metabolism. Thus, the computed metabolism estimates after weighted regression are different from the original metabolism computations, but equally incorrect.

Tidal currents in estuaries can move water across a mosaic of habitat types that likely have different biological rates. For example, the magnitudes of production and respiration, and the net metabolism may differ across salinity gradients, between open water and shallow water habitats, and inside and outside of tidal marshes and seagrass habitats. These complex estuarine seascapes are especially prevalent at sites like those in the NERRS network. The interaction between spatial differences in metabolic processes and tidal advection moving water masses among habitats contributes to tidal artifacts in open water metabolism calculations. Thus, the weighted regression method is ideally applied in estuarine environments similar to NERRS sites. Advection in an open coastal environment, which may simply move water within a single homogeneous habitat, might generate relatively little distortion of the metabolic signal within the DO record. In these cases, weighted regression could be used but may have little benefit. As a final note, it is worth mentioning that these computations are complex relative to the basic open water metabolism calculations. However, because we implemented the analysis in R, the freely-available programming environment for statistical analysis, scientific computing and graphics (RDCT 2014), many users may be able to replicate the approach using a package provided in the supporting information. As the traditional open-water metabolism computations have already been implemented in R and shared widely, the additional development of the method presented here could also be applied by using our code.

***References***

Batt, R. D., and S. R. Carpenter. 2012. Free-water lake metabolism: Addressing noisy time series with a Kalman filter. Limnology and Oceanography: Methods **10**: 20–30.

Caffrey, J. M. 2003. Production, respiration and net ecosystem metabolism in U.S. estuaries. Environmental Monitoring and Assessment **81**: 207–219.

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

Caffrey, J. M., M. C. Murrell, K. S. Amacker, J. Harper, S. Phipps, and M. Woodrey. 2014. Seasonal and inter-annual patterns in primary production, respiration and net ecosystem metabolism in 3 estuaries in the northeast Gulf of Mexico. Estuaries and Coasts **37**: 222–241.

Collins, J. R., P. A. Raymond, W. F. Bohlen, and M. M. Howard-Strobel. 2013. Estimates of new and total productivity in Central Long Island Sound from in situ measurements of nitrate and dissolved oxygen. Estuaries and Coasts **36**: 74–97.

Coloso, J. J., J. J. Cole, and M. L. Pace. 2011. Difficulty in discerning drivers of lake ecosystem metabolism with high-frequency data. Ecosystems **14**: 935–948.

Cryer, J. D., and K. S. Chan. 2008. Time series analysis with applications in r, second. Springer.

D’Avanzo, C., J. N. Kremer, and S. C. Wainwright. 1996. Ecosystem production and respiration in response to eutrophication in shallow temperate estuaries. Marine Ecology Progress Series **141**: 263–274.

Foreman, M. G. G., and R. F. Henry. 1989. The harmonic analysis of tidal model time series. Advances in Water Resources **12**: 109–120.

Gaarder, T., and H. H. Gran. 1927. Investigations of the production of plankton in the oslo fjord. Rapports et Procés-Verbaux des Réunions du Conseil Permanent International pour l’Exploration de le Mer **42**: 1–48.

Harvey, A. C. 1989. Forecasting, structural time series models and the kalman filter, Cambridge University Press.

Hilborn, R., and M. Mangel. 1997. The ecological detective: Confronting models with data, Princeton University Press.

Hirsch, R. M., D. L. Moyer, and S. A. Archfield. 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. Journal of the American Water Resources Association **46**: 857–880.

Kelley, D., and C. Richards. 2014. oce: Analysis of oceanographic data, R package version 0.9-15.

Kemp, W. M., and W. R. Boynton. 1980. Influence of biological and physical processes on dissolved oxygen dynamics in an estuarine system: Implications for the measurement of community metabolism. Estuarine and Coastal Marine Science **11**: 407–431.

Kemp, W. M., and J. M. Testa. 2012. Metabolic balance between ecosystem production and consumption, p. 83–118. *In* E. Wolanski and D.S. McLusky [eds.], Treatise on estuarine and coastal science. Academic Press.

Needoba, J. A., T. D. Peterson, and K. S. Johnson. 2012. Method for the quantification of aquatic primary production and net ecosystem metabolism using in situ dissolved oxygen sensors, p. 73–101. *In* S.M. Tiquia-Arashiro [ed.], Molecular biological technologies for ocean sensing. Springer.

Nidzieko, N. J., J. A. Needoba, S. G. Monismith, and K. S. Johnson. 2014. Fortnightly tidal modulations affect net community production in a mesotidal estuary. Estuaries and Coasts **37**: 91–110.

Nixon, S. W. 1995. Coastal marine eutrophication: A definition, social causes, and future concerns. Ophelia **41**: 199–219.

NRC (National Research Council). 2000. Clean Coastal Waters: Understanding and Reducing the Effects of Nutrient Pollution. National Academy Press, Washington, DC.

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

Pamatmat, M. M. 1997. Non-photosynthetic oxygen production and non-respiratory oxygen uptake in the dark: A theory of oxygen dynamics in plankton communities. Marine Biology **129**: 735–746.

RDCT (R Development Core Team). 2014. R: A language and environment for statistical computing, v3.1.0. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.

Russell, M. J., and P. A. Montagna. 2007. Spatial and tepmoral variability and drivers of net ecosystem metabolism in western Gulf of Mexico estuaries. Estuaries and Coasts **30**: 137–153.

Sanger, D. M., M. D. Arendt, Y. Chen, E. L. Wenner, A. F. Holland, D. Edwards, and J. Caffrey. 2002. A synthesis of water quality data: National estuarine research reserve system-wide monitoring program (1995-2000). National Estuarine Research Reserve Technical Report Series 2002:3. South Carolina Department of Natural Resources, Marine Resources Division Contribution No. 500.

Sasaki, A., Y. Hagimori, T. Nakatsubo, and A. Hoshika. 2009. Tidal effects on the organic carbon mineralization rate under aerobic conditions in sediments of an intertidal estuary. Ecological Research **24**: 723–729.

Shumway, R. H., and D. S. Stoffer. 2011. Time series analysis and its applications: With r examples, 3rd ed. Springer.

Staehr, P. A., D. Bade, M. C. V. de Bogert, G. R. Koch, C. Williamson, P. Hanson, J. J. Cole, and T. Kratz. 2010. Lake metabolism and the diel oxygen technique: State of the science. Limnology and Oceanography: Methods **8**: 628–644.

Thébault, J., T. S. Schraga, J. E. Cloern, and E. G. Dunlavey. 2008. Primary production and carrying capacity of former salt ponds after reconnection to San Francisco Bay. Wetlands **28**: 841–851.

Tukey, J. W. 1977. Exploratory data analysis, Addison-Wesley.

Wenner, E., D. Sanger, M. Arendt, A. F. Holland, and Y. Chen. 2004. Variability in dissolved oxygen and other water-quality variables within the National Estuarine Research Reserve System. Journal of Coastal Research **45**: 17–38.

Ziegler, S., and R. Benner. 1998. Ecosystem metabolism in a subtropical, seagrass-dominated lagoon. Marine Ecology Progress Series **173**: 1–12.

***Acknowledgments***

We acknowledge the significant efforts of research staff and field crews from the System Wide Monitoring Program of the National Estuarine Research Reserve System for providing access to high quality data sets. We thank Dr. Jane Caffrey and Dr. Erik Smith for reviewing early drafts of the manuscript. The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.

***Multimedia***

A simple R package with a sample dataset and code to implement weighted regression is available on GitHub, including functions to estimate ecosystem metabolism. See the README file on the web page for download instructions and examples (<https://github.com/fawda123/WtRegDO>). Interactive applications are also available that illustrate the weighting scheme described in the material and procedures section ([https://beckmw.shinyapps.io/weights widget),](https://beckmw.shinyapps.io/weights%20widget),%20) results for each simulation ([https://beckmw.shinyapps.io/detiding sims/),](https://beckmw.shinyapps.io/detiding%20sims/),%20) and results for each case study ([https://beckmw.shinyapps.io/detiding cases/](https://beckmw.shinyapps.io/detiding%20cases/)). Each link is a graphical summary of data based on interactive inputs to support the results in the manuscript.

***Figure legends***

Fig. 1: Example of each component of a simulated DO time series for testing weighted regression.

The time series were created using eqs. (4) to (13). Yellow indicates a twelve hour daylight period beginning at 630 each day.

Fig. 2: Representative examples of simulated time series of observed DO (*DOobs*, blue lines) and biological DO (*DObio*, as a component of observed, red lines) created by varying each of four parameters: strength of tidal association with DO signal (*DOadv*), amount of process uncertainty (*εpro*), amount of observation uncertainty (*εobs*), and strength of diel DO component (*DOdie*). Parameter values represent the minimum and maximum used in the simulations as mg L−1 of DO.

Fig. 3: Heat maps of correlations and errors (RMSE) for filtered DO time series (*DOdtd*) from weighted regression with ‘true’ biological DO (*DObio*) for varying simulation parameters: strength of tidal association with DO signal (*DOadv*), amount of process uncertainty (*εpro*), amount of observation uncertainty (*εobs*), and strength of diel DO component (*DOdie*). Each tile represents the correlation or error from results for a given combination of simulation parameters averaged for all window widths (as in Fig. 4).

Fig. 4: Heat maps of correlations and errors (RMSE) for filtered DO time series (*DOdtd*) from weighted regression with ‘true’ biological DO (*DObio*) for varying half window widths: days, hour of day, and proportion of tidal range. Each tile represents the correlation or error from results for a given combination of window widths averaged for all simulation parameters (as in Fig. 3).

Fig. 5: Locations of NERRS sites used as case studies to validate weighted regression. Stations at each reserve are ELKVM (Vierra Mouth at Elkhorn Slough), PDBBY (Bayview Channel at Padilla

Bay), RKBMB (Middle Blackwater River at Rookery Bay), and SAPDC (Dean Creek at Sapelo

Island).

Fig. 6: Continuous DO time series before (observed) and after (filtered) filtering with weighted regression (top) and tidal height (m) colored by total photosynthetically active radiation (bottom,

mmol m−2). Results are for the Sapelo Island station for a seven day period characterized by a neap tide and relatively unequal first and second high and low tides. These result in large errors in applying the open water method to unfiltered data, and improvements after removing tidal effects (Fig. 8). Weighted regression is applied with half window widths of six days, one hour within each day, and tidal height proportion of one half.

Fig. 7: Continuous DO time series before (observed) and after (filtered) filtering with weighted regression (top) and tidal height (m) colored by total photosynthetically active radiation (bottom,

mmol m−2). Results are for the Sapelo Island station for a seven day period characterized by spring tides and relatively small differences between first and second high and low tides. Artifacts due to advection are smaller than in the previous week because tide levels are relatively similar at the beginning and end of day and night periods. Weighted regression is applied with half window widths of six days, one hour within each day, and tidal height proportion of one half.

Fig. 8: Example of daily mean metabolism (net ecosystem metabolism, gross production, and total respiration) before (observed) and after (filtered) filtering with weighted regression. Results are for the Sapelo Island station for a two week period in February 2012 characterized by a transition from unequal neap tides (Fig 6) to relatively equal spring tides (Fig 7).

Fig. 9: Correlations of sun angle with tidal change (as an angular rate) using a half-window width of 12 days. Correlations larger or smaller than zero are periods when weighted regression may not effectively quantify variation from biological and physical sources in DO time series due to collinearity.

Fig. 10: Means and standard errors of daily metabolism estimates (gross production, total respiration) averaged by month. Averaged results are shown for observed and filtered DO time series. May was removed from Rookery Bay because of incomplete data.

***Tables***

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1: Summary statistics of tidal component amplitudes (m), selected water quality parameters (DO mg L−1, chlorophyll-a *μ*g L−1, salinity psu, water temperature ◦C) and metabolism estimates (gross production, respiration, and net ecosystem metabolism as g m−2 d−1 based on observed data) for each case study. Tidal components are principal lunar semidiurnal (O1, frequency 25.82 hours), solar diurnal (P1, 24.07 hours), lunar semidiurnal (M2, 12.42 hours), and solar semidiurnal (S2, 12 hours) estimated from harmonic regressions of tidal height (oce package in R, Foreman and Henry 1989, RDCT 2014). Water quality data are averages for the entire period of record for each site. Metabolism estimates are means of daily integrated values. | | | | | | | | | | | | | |
| **Site** | **Tidal amplitude** | | | |  | **Water quality** | | | |  | **Metabolism*a*** | | |
|  | O1 | P1 | M2 | S2 |  | DO | Chl | Sal | Temp |  | Pg | Rt | NEM |
| ELKVM | 0.24 | 0.12 | 0.48 | 0.13 |  | 7.87 | 3.87 | 32.43 | 13.78 |  | 8.14 | -8.19 | -0.05 |
| PDBBY | 0.46 | 0.23 | 0.63 | 0.15 |  | 8.97 | 2.24 | 29.17 | 10.44 |  | 5.95 | -5.90 | 0.05 |
| RKBMB | 0.13 | 0.04 | 0.36 | 0.10 |  | 4.48 | 4.50 | 30.53 | 25.85 |  | 3.02 | -3.62 | -0.60 |
| SAPDC | 0.10 | 0.02 | 0.54 | 0.07 |  | 4.96 | 5.98 | 27.30 | 21.77 |  | 4.89 | -6.04 | -1.16 |
| *a* Pg: gross production, Rt: respiration, NEM: net ecosystem metabolism, estimated using methods described herein without removing advection effects. | | | | | | | | | | | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| Table 2: Correlations of tidal changes at each site with continuous DO observations and metabolism estimates (gross production, respiration) before (observed) and after (ﬁltered) ﬁltering with weighted regression. Values are averages of monthly correlations. DO values are correlated with predicted tidal height at each observation, whereas metabolism estimates are correlated with mean tidal height change between observations during day or night periods for production and respiration, respectively. | | | |
| Site | DO | Pg*a* | Rt |
| **ELKVM** |  |  |  |
| Observed | 0.44 | 0.43 | 0.43 |
| Filtered | -0.04 | 0.04 | -0.01 |
| **PDBBY** |  |  |  |
| Observed | -0.49 | -0.11 | -0.29 |
| Filtered | 0.01 | -0.05 | 0.00 |
| **RKBMB** |  |  |  |
| Observed | 0.45 | 0.26 | 0.34 |
| Filtered | 0.02 | -0.04 | 0.03 |
| **SAPDC** |  |  |  |
| Observed | 0.62 | 0.47 | 0.64 |
| Filtered | 0.00 | -0.04 | 0.07 |
| *a*Pg: gross production, Rt: respiration, NEM: net ecosystem metabolism | | | |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Table 3: Summary of metabolism estimates (gross production, respiration) for case studies using DO time series before (observed) and after (ﬁltered) ﬁltering with weighted regression. Means and standard deviation are based on daily integrated metabolism estimates. Anomalous values are the percentage of metabolism estimates that were negative for gross production and positive for respiration. | | | | | | | |
| **Site** | **Pg*a*** | | |  | **Rt** | | |
|  | Mean | SD | Anom |  | Mean | SD | Anom |
| **ELKVM** |  |  |  |  |  |  |  |
| Observed | 8.14 | 11.44 | 18.64 |  | -8.19 | 11.82 | 20.68 |
| Filtered | 3.00 | 2.26 | 6.78 |  | -3.06 | 2.34 | 7.12 |
| **PDBBY** |  |  |  |  |  |  |  |
| Observed | 5.95 | 11.69 | 21.80 |  | -5.90 | 12.60 | 19.03 |
| Filtered | 7.01 | 6.85 | 4.84 |  | -7.01 | 6.81 | 5.54 |
| **RKBMB** |  |  |  |  |  |  |  |
| Observed | 3.02 | 2.55 | 9.15 |  | -3.62 | 2.65 | 7.57 |
| Filtered | 3.46 | 1.79 | 0.63 |  | -4.07 | 1.88 | 0.63 |
| **SAPDC** |  |  |  |  |  |  |  |
| Observed | 4.89 | 4.42 | 13.39 |  | -6.04 | 4.78 | 10.93 |
| Filtered | 4.94 | 1.74 | 0.00 |  | -6.13 | 1.99 | 0.00 |
| *a*Pg: gross production, Rt: respiration | | | | | | | |