Limnol. Oceanogr., 57(6), 2012, 1689–1700

doi:10.4319/lo.2012.57.6.1689

Spatial heterogeneity strongly affects estimates of ecosystem metabolism in two north temperate lakes

Matthew C. Van de Bogert,[a,*] Darren L. Bade,[b] Stephen R. Carpenter,[a] Jonathan J. Cole,[c] Michael L. Pace,[d] Paul C. Hanson,[a] and Owen C. Langman [a]

[a] Center for Limnology, University of Wisconsin, Madison, Wisconsin
[b] Department of Biological Sciences, Kent State University, Kent, Ohio
[c] Cary Institute of Ecosystem Studies, Millbrook, New York
[d] Department of Environmental Sciences, University of Virginia, Charlottesville, Virginia

Abstract

To characterize the spatial variability of metabolism estimates (gross primary production [GPP], respiration [R], and net ecosystem production [NEP]) in two Northern Wisconsin lakes, we collected data from 27 and 35 dissolved oxygen sensors placed in a two-dimensional array throughout the upper mixed layers over a period of 10 d per lake in midsummer. Averaged over the deployment, aerial metabolism estimates among sensor locations varied 1–2 orders of magnitude and were largely unrelated to physical habitat within the lake. For all sites and days, 76–90% of the explainable variance in GPP and R was attributable to location in the lake rather than day of the deployment. NEP, on the other hand, was less affected by location, with 79–93% of the explained variance attributable to the day of the deployment. Single-location estimates can yield errors of more than an order of magnitude in estimates of daily GPP and R and can mischaracterize the trophic status of the lake. Using a rarefaction approach, we found that using four randomly placed sensors increased the precision of the resulting daily metabolism estimates fourfold over single-location measures in both lakes.

Ecosystem metabolism is fundamentally important to the understanding of ecosystem function and has been a topic of research for decades (Odum 1957*a*; Fisher and Likens 1973; Hanson et al. 2003). Interest in aquatic metabolism has further increased in recent years with the recognition that aquatic ecosystems, despite their small footprint, can act as hot spots of carbon cycling on the landscape, both as storage reservoirs and rapid processors of terrestrially fixed carbon (Cole et al. 2007). Measurements of metabolism have been made using a variety of approaches, each with its own advantages and difficulties. Early labor-intensive free-water methods (Odum 1956, 1957*b*) were eclipsed by ^{14}C methods for primary production with a promise of higher sensitivity and a more direct link to the underlying physiological mechanisms. Over the past 15 yr, technological improvements in dissolved oxygen sensors have led to a renewed interest in free-water methods for estimating gross primary production (GPP), respiration (R), and net ecosystem production (NEP).

Despite the current widespread use of free-water approaches, questions related to the interpretation of free-water metabolism estimates remain (Van de Bogert et al. 2007; Staehr et al. 2010). Following a long tradition of sampling lakes at their deepest point, free-water studies of metabolism often use a single oxygen sensor. This approach has the implicit assumption that metabolism estimates from a central location will be representative of at least the pelagic zone, if not the entire epilimnion. Despite the recognition of biological and physical heterogeneity between and within pelagic and littoral habitats, the cost of obtaining and maintaining automated sampling devices often precludes sampling across known and unknown gradients of heterogeneity. For example, many studies have identified differences between pelagic and littoral habitats with respect to both primary and secondary production (Lindeman 1942; Wetzel 1995; Vadeboncoeur et al. 2002). Within littoral zones, patchiness of benthic substrate types may drive spatial heterogeneity in rates of metabolism (Vadeboncoeur and Lodge 2000; Lauster et al. 2006; Vadeboncoeur et al. 2006). In pelagic zones, studies have shown that even moderate physical forcing can create patches of differing nutrient concentrations and abundances of producers and consumers (Caron et al. 2008; Blukacz et al. 2010; Mackay et al. 2011). Despite recognition of this underlying heterogeneity, the use of single sensors has dominated free-water metabolism studies because of their high cost, the time investment necessary for calibration and maintenance, and uncertainty in the added value of sampling more broadly. Still, studies using free-water measurements at multiple sites are beginning to demonstrate marked spatial differences in metabolism estimates. These studies have largely focused on differences between littoral and pelagic zones but have also noted 2–3-fold variation among estimates from different sites within a given habitat (Van de Bogert et al. 2007; Sadro et al. 2011).

In addition to spatial patterns in lake metabolism, recent studies have drawn attention to the often quite large, and unexplained, day-to-day variability in metabolism estimates from measurements made at a single location (Staehr and Sand-Jensen 2007; Staehr et al. 2010; Coloso et al. 2011). Both Staehr and Sand-Jensen (2007) and Coloso et al. (2011) used multiple regression models to predict summertime GPP, R, and NEP, using several biological and physical predictors (e.g., chlorophyll *a*, zooplankton

* Corresponding author: matthew.vandebogert@gmail.com

biomass, wind speed, water column stability, and others) but found that in most cases their models left > 50% of the variance unexplained. While some of the day-to-day variation could be attributed to unmeasured drivers of metabolism (e.g., nutrient availability), it is also possible that some of the variation reflects within-lake heterogeneity due to measuring transient "hot spots" or "cold spots" of metabolism, or to physical processes causing the sensor to measure a parcel of water with differing metabolic and physical histories for some portion of the day (Serra et al. 2007; Mackay et al. 2011).

Differences in metabolism among sites may be explained by both underlying differences in ecosystem structure and the degree to which physical processes isolate or homogenize those differences. Under conditions of isolation, measurements should accurately characterize the metabolism at the specific sensor location; in this case, variation should be clearly attributable to mechanistic drivers. Van de Bogert et al. (2007), for example, demonstrated that when wind speed remained low, metabolism along a transect from littoral to pelagic habitats nicely fit a mixing model of water column and benthic metabolism, with sensors in the middle of the lake measuring water column metabolism and those nearshore capturing water column plus benthic metabolism. Sadro et al. (2011) similarly found the greatest differences between littoral and pelagic metabolism in an oligotrophic alpine lake when surface-water advection was low and Schmidt stability was high. When physical processes act towards homogenization of the system, however (e.g., horizontal advection due to temperature gradients or wind forcing), two things may happen. In the best case, metabolic signals among different locations would be averaged: a sensor in the middle of the lake might account for metabolism both in the water column and from benthic sources nearshore. Alternatively, the physical movement of water could jumble several heterogeneous signals: a sensor might measure pelagic metabolism for a few hours and then transition to measuring metabolism in a water mass that originated nearshore a few hours ago but has now moved offshore, thus displacing the water mass that the sensor was originally sampling. In either case, it is clear that single-sensor estimates of metabolism may miss important components of ecosystem heterogeneity.

Characterizing within-lake variability in metabolism estimates and developing methods for capturing this variation are important for advancing our understanding of ecosystem metabolism and its drivers. In this study we use an unprecedented number of sensors placed in a two-dimensional array throughout the epilimnia of two strongly stratified lakes of contrasting sizes to characterize the spatial (horizontal) pattern of metabolism estimates in both littoral and pelagic regions. We then contrast the spatial variation with temporal variation over a 10-d period as a benchmark for identifying the density of sensors necessary to obtain reasonably precise estimates of whole-ecosystem metabolism.

Methods

Study site and data collection—Data for this study were collected from two low-productivity seepage lakes, Sparkling Lake (46.008 N, 89.701 W) and Peter Lake (46.253 N, 89.504 W), located in the Northern Highlands Lake District of Wisconsin and Michigan's Upper Peninsula over a period of 10 d for each lake in July and August of 2007 (Table 1). Both lakes are strongly stratified during the summer season, with mean mixed-layer depths of 6.5 m (Sparkling) and 3 m (Peter) during our study period. The lakes provide a contrast in surface area and the proportion of the lake in littoral vs. pelagic zones (Fig. 1; Table 1). Sparkling Lake is 25 times larger than Peter Lake (by area) and has 27% of its area in the littoral zone (upper mixed-depth reaches to the sediment), whereas Peter Lake is 40% littoral by area. Macrophytes were sparse in both lakes over the deployment period.

Table 1. Characteristics of the study lakes over the period of sensor deployment.

Lake characteristic	Sparkling Lake	Peter Lake
Surface area (m^2)	640,000	24,000
Maximum depth (m)	20.0	19.6
Mixed-layer depth (m)	6.5	3.0
Thermocline depth (m)	8.5	4.5
Total phosphorus ($\mu g\ L^{-1}$)	6.0	12.4
Total nitrogen ($\mu g\ L^{-1}$)	239	363
Chlorophyll *a* ($\mu g\ L^{-1}$)	1.1	5.4
Dissolved organic carbon ($mg\ L^{-1}$)	3.3	5.8
Euphotic depth (m)	14.7	6.6
Mixed-layer water temperature, mean (°C)	23.3	25.5
Air temperature, mean (°C)	20.7	21.5
Daily air temperature range, mean (°C)	16.6–25.6	17.4–26.5
Wind speed, mean±SD ($m\ s^{-1}$)	2.0±1.6	1.0±0.85

Measurements of dissolved oxygen and temperature were made every 10 min at a depth of 0.7 m using multiple sondes dispersed horizontally throughout the mixed layer. We used 35 sondes in Sparkling Lake and 27 in Peter Lake (Fig. 1). The 35 sondes in Sparkling Lake included 31 from YSI Incorporated (15 of model 600XLM, 14 of model 6920, and 2 of model 6600) and four D-Opto sensors (Zebra-Tech). In Peter Lake we used 14 YSI model 6920 and 13 YSI model 600XLM sondes. Six of the YSI sondes used in both lakes were fitted with YSI ROX™ 6150 optical dissolved oxygen sensors and the remaining YSI sondes were fitted with model 6562 Rapid Pulse™ Clark-electrode dissolved oxygen sensors. The four D-Opto sensors used optical sensor technology. Before each deployment, all Clark-type sensors were reconditioned if necessary and had new membranes applied. Proper probe response was tested using the manufacturers' suggested diagnostics. Logging was initialized and all sensors were placed in water-saturated air in the laboratory for a minimum of 3–4 h prior to calibration. After calibration, sensors were allowed to continue logging in water-saturated air for at least 3 more hours prior to deployment to verify the calibration. After each deployment, sensors were moved to a common location and allowed to log for several hours to identify sensor-specific drift, and linear adjustments were made accordingly. Further details related to calibration are in Van de Bogert et al. (2007).

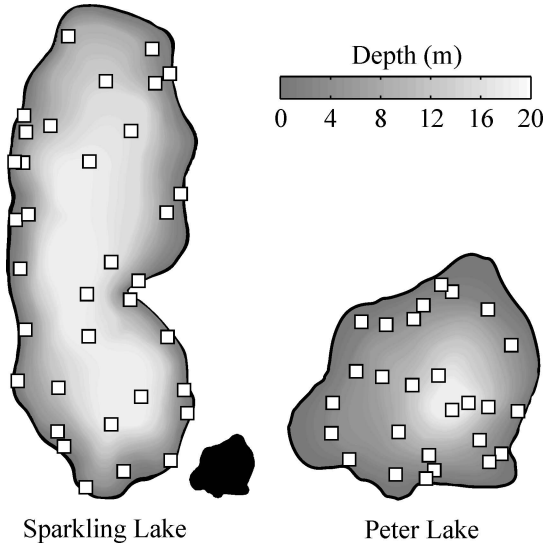

Fig. 1. Bathymetry and location of the sensors in each of our study lakes. Peter Lake is ~ 200 m across and the small silhouette next to the Sparkling Lake map shows its size relative to Sparkling Lake. The horizontal bar relates the map shading to depth (m).

Because previous studies have shown high variation in metabolism among shallow, littoral sites (Caraco and Cole 2002; Lauster et al. 2006) and found that metabolism varies as a function of distance from shore (Van de Bogert et al. 2007), we chose to use a stratified-random placement of sensors to be sure to adequately sample the littoral zone despite its smaller surface area and volume than the pelagic. Each lake was stratified into two littoral regions and two or three pelagic regions with a predetermined number of sensors allotted to each (Table 2; Fig. 1). Sensors were randomly assigned locations within each of the zones using rasterized bathymetric maps of the lakes and a random number generator in Matlab (The MathWorks, release 14). Within each lake, one pelagic sensor was placed over the deepest location of the lake, which is the location used for routine long-term sampling. In Sparkling Lake, this site corresponds to the location of a long-term monitoring buoy. After locations were determined, individual sensors were randomly assigned to each location, with the exception of the four D-Opto sensors in Sparkling Lake. These sensors are part of larger monitoring buoys used in the North Temperate Lakes Long Term Ecological Research program. One of these was located near the deep hole of the lake, whereas the other three were assigned to random locations along the north shore, south shore, and pelagic regions of the lake.

We calculated daily metabolism values (GPP, R, and NEP) based on the continuous dissolved oxygen (DO) and temperature data collected from each individual sensor according to the methods of Cole et al. (2000) following the approach of Odum (1956). We used the underlying model:

$$\frac{d\mathrm{DO}}{dt} = \mathrm{GPP} - \mathrm{R} + D \quad (1)$$

where $d\mathrm{DO}/dt$ is the change in DO over each time interval (converted to areal units by dividing by the depth of the mixed layer at the sensor location), and D is the diffusive exchange of oxygen with the atmosphere and is calculated as

$$D_t = k(\mathrm{DO}_{\mathrm{SAT},t} - \mathrm{DO}_t) \quad (2)$$

where k is the piston velocity, $\mathrm{DO_{SAT}}$ is the equilibrium DO concentration, and DO_t is the actual DO concentration at the sensor. The piston velocity, k, was calculated for each time interval according to the method of Jahne et al. (1987), using an oxygen- and temperature-specific Schmidt number (Wanninkhof 1992) and k_{600} (piston velocity for a gas with a Schmidt number of 600) estimated based on the low-wind equations of Cole and Caraco (1998). Wind speed was measured at 2 m above the water surface using an R.M. Young anemometer and average values were recorded every 10 min. We used the empirical equation of Smith (1985) to estimate the wind speed at 10 m above the water surface, following the approach of Cole et al. (2000). For converting from volumetric to areal units, we calculated the daily depth of the epilimnion (z_{mix}) as the first depth at which rate of temperature change is $> 1^{\circ}\mathrm{C\ m^{-1}}$ at sunrise. Over the course of each of the 10-d deployments, there were no episodic thermocline deepening events and the depth of the diel mixed layer did not change. For littoral sensor locations where z_{mix} was greater than the depth of the water column, z_{mix} was assigned the water depth for that site.

Table 2. Number of sensors (n) randomly placed in each of two littoral and two (Sparkling Lake) or three (Peter Lake) pelagic zones.

Zone	Sparkling Lake		Peter Lake	
	Depth (m)	n	Depth (m)	n
Littoral, nearshore	0–2	11	0–2	6
Littoral, offshore	2–6.5	12	2–3	5
Pelagic 1	6.5–15	6	3–7	7
Pelagic 2	>15	6	7–12	5
Pelagic 3			>12	4

Heterogeneity in metabolism estimates—In order to test the hypothesis that variation in metabolism estimates among sites and days is statistically significant, we conducted two-way fixed-effects analyses of variance (ANOVAs) for each metabolism parameter (GPP, R, NEP) in each lake. We then partitioned the variance attributable to day of deployment (n = 10 for GPP and NEP, n = 11 for R) and location within the lake (n = 35 for Sparkling Lake, n = 27 for Peter Lake) using the approach outlined in Gotelli and Ellison (2004). We used Fisher's protected least significant difference (LSD) to identify the smallest interval around each site's mean within which means of other sites are not considered significantly different.

Significant differences in metabolism estimates between pelagic and littoral habitats were identified using a two-way fixed-effects ANOVA with habitat (pelagic or littoral) and day of the deployment as factors (Matlab, Statistics Toolbox release 5.01). We did this for each metabolism parameter and for volumetric and areal units of metabolism. We used the

Tukey–Kramer honestly significant difference criteria for unequal sample sizes to identify the 95% confidence interval for the difference in means between the habitats. Variations in water depth over the littoral area lead to different patterns in metabolism when viewed as areal (mmol m^{-2} d^{-1}) compared to volumetric (mmol m^{-3} d^{-1}) estimates. For the littoral–pelagic comparisons, we provide results using both units. For all other analyses, we report only areal values.

We tested for spatial autocorrelation in our daily metabolism estimates by first estimating variograms for the observed lag distances (distances between sensors) for each lake and metabolism parameter using the geoR package in R (R Development Core Team 2011). We compared these variograms to simulation envelopes representing the range of values expected if there is no spatial dependence (Diggle and Ribeiro 2007). These envelopes were created empirically by randomly reassigning observed metabolism estimates to sensor locations and estimating variograms based on this now spatially independent sample. This was repeated 100 times, and the confidence envelopes were created using the minimum and maximum values of the simulated variograms at each lag distance.

Rarefaction of whole-lake metabolism estimate—Lake-wide epilimnetic metabolism was calculated from individual sites using a habitat-weighted average of littoral and pelagic metabolism where the mean of littoral sites was applied to the surface area of the lake's littoral zone and the pelagic mean was applied to the remaining area of the lake. We used a rarefaction procedure (analogous to Simberloff 1972; Heck et al. 1975) to identify how combining data from multiple sensors influences the precision of the lake-wide metabolism estimates. We identified the marginal increase in precision per additional sensor by randomly resampling metabolism estimates from the pool of sensors, combining estimates from multiple sensors into a single habitat-weighted average value for the lake. We then calculated the standard deviation of lake-wide metabolism estimates obtained from 1, 2, …, $n - 1$, n sample locations, where n is the total number of sensors deployed. The details of the procedure are provided below. Adding additional sensors will always yield a more precise estimate of metabolism; to provide one potential benchmark for identifying when the gain from an additional sensor is warranted, we compared the standard deviation from the rarefaction procedure to the lake-wide standard deviation attributable solely to the day-to-day factor from the two-way day vs. site ANOVA discussed above.

Pseudocode for rarefaction of whole-lake metabolism estimate—

(1) Calculate metabolism (GPP, R, NEP) for each site and day.
(2) For $k = 1$ to n, where n equals the total number of sites where sensors were deployed:
    (a) Randomly choose k sites from the pool of n sites.
    (b) Calculate average metabolism values (GPP, R, NEP) using the subset of k sensors.
    (c) Repeat 2a and 2b using another random subset of sites; continue repeating the procedure until either 1000 unique subsets have been sampled or the maximum number of subsets (given by the binomial coefficient ${}_nC_k$) has been reached, whichever is less.
    (d) Calculate the mean and standard deviation of the repeated measures of the mean metabolism values using k sites.
    (e) Repeat 2a through 2d for the next value of k.

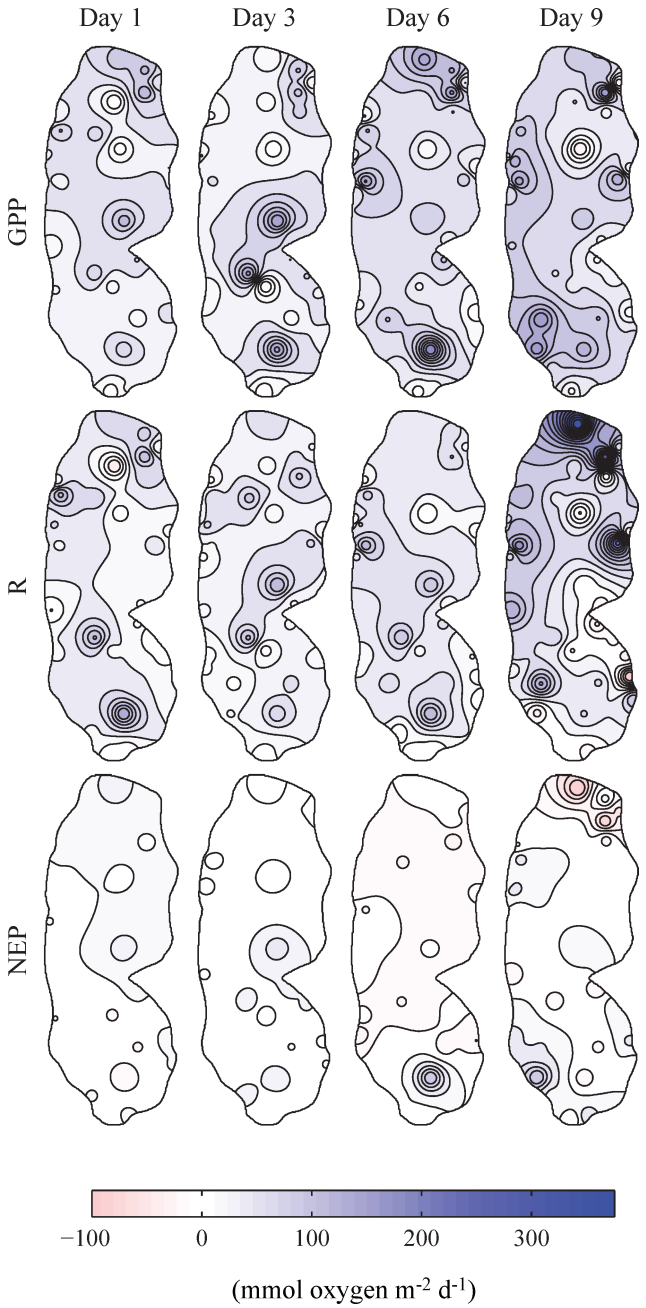

Fig. 2. Spatial pattern of daily GPP, R, and NEP estimates (mmol m^{-2} d^{-1}) for four representative days on Sparkling Lake. The color scale is the same for all three metabolism measures. The contour interval is 20 mmol m^{-2} d^{-1} for GPP and R and 10 mmol m^{-2} d^{-1} for NEP.

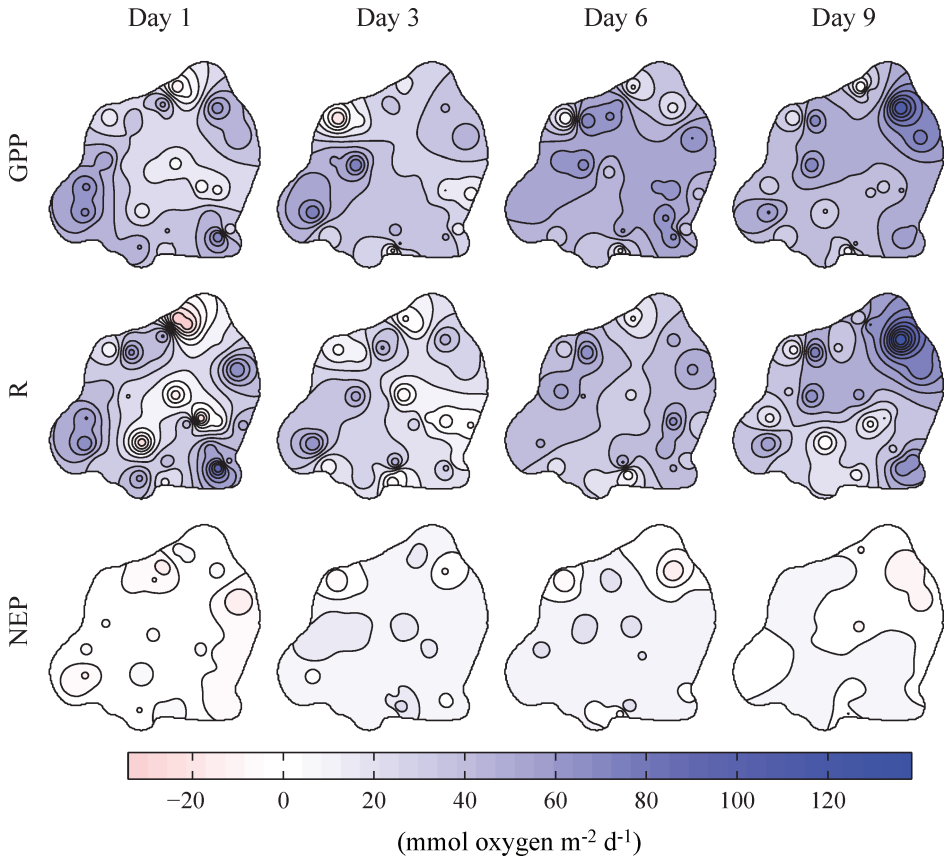

Fig. 3. Spatial pattern of daily GPP, R, and NEP estimates (mmol m^{-2} d^{-1}) for four representative days on Peter Lake. The color scale is the same for all three metabolism measures. The contour interval is 10 mmol m^{-2} d^{-1} for GPP and R and 5 mmol m^{-2} d^{-1} for NEP.

Targeted rarefaction procedure—In order to determine if targeting the placement of sensors in specific habitats (as opposed to a random sampling of our sites) could result in more precise metabolism estimates with fewer sensors, we conducted the rarefaction procedure above for $k = 2$ to 14 sensors, across all possible combinations of 1 to $k - 1$ sensors in each of the two habitats (littoral vs. pelagic). For example, when $k = 3$, we conducted the rarefaction analysis two times ($k - 1$), once using one pelagic and two littoral sensors, and again using two pelagic and one littoral sensor. Individual sensors within each habitat were still selected randomly. We compared the standard deviation of metabolism estimates from each of these habitat-targeted combinations to that of our random selection procedure and the benchmark of standard deviation attributable the day-to-day factor in the ANOVA described above.

Results

Variability of metabolism estimates among sites and days—Estimates of GPP, R, and NEP over all sites and days were highly variable in both lakes (Figs. 2, 3; Table 3). Over all sensor locations and days in Sparkling Lake, average habitat-weighted estimates of GPP, R, and NEP were 67, 57, and 11 mmol m^{-2} d^{-1}, with standard deviations of 37, 50, and 21 mmol m^{-2} d^{-1}, respectively. For Peter Lake, average GPP, R, and NEP were 44, 36, and 8.0 mmol m^{-2} d^{-1} with standard deviations of 15, 18, and 6.9 mmol m^{-2} d^{-1}, respectively. Maps of metabolism did not reveal consistent lake-wide spatial patterns in metabolism estimates; many sites with relatively high metabolism one day had low metabolism on another day of the deployment (Figs. 2, 3). Tests for spatial autocorrelation using empirical variograms revealed no spatial dependence for any lake or metabolism parameter over the distances we sampled; variograms were always within the confidence envelopes, signifying a lack of spatial dependence (data not shown). The ranges of metabolism estimates for individual locations were reduced by 60–85% when averaged over the deployment period (Fig. 4).

Partitioning variance among locations and days—We partitioned the variance in areal metabolism estimates among locations and days of the deployment using a two-way fixed-effects ANOVA for each lake and metabolism parameter combination (total = 6). We found significant differences among sites and days for all metabolism parameters in both lakes ($p < 0.001$) except for NEP in Sparkling Lake, which did not vary significantly among

Table 3. The grand mean (coefficient of variation), minimum (min.), and maximum (max.) values over all sites and days for GPP, R, and NEP in both lakes (mmol m^{-2} d^{-1}). *n* is the number of sensor-days included for each metabolism parameter.

	Sparkling Lake				Peter Lake			
	Mean, CV (%)	Min.	Max.	*n*	Mean, CV (%)	Min.	Max.	*n*
GPP	67(55)	−132	250	350	44(35)	−45	116	270
R	57(87)	−131	401	385	36(51)	−61	138	297
NEP	11(190)	−109	108	350	8.0(86)	−16	33	270

sites. Together, sensor location and day of deployment account for between 25% and 63% of the observed variance in metabolism rates (Fig. 5), with location in the lake accounting for the majority of explainable variance for GPP and R in both lakes. Sensor location accounts for 90% and 81% of the explained variance for GPP and R, respectively, in Sparkling Lake and 84% and 76% of the explained variance in Peter Lake GPP and R, respectively. Day of the deployment, on the other hand, contributed the largest portion of the explained variance in NEP for both lakes (93% and 79%, for Sparkling Lake and Peter Lake, respectively).

Using Fisher's LSD approach, we determined the smallest difference between any two sites that can be considered statistically significant ($\alpha = 0.05$). Comparing each site's deployment mean to the lake-wide deployment mean using the LSD as a yardstick for significance revealed only two sites that had NEP values significantly different from the lake-wide means in Peter Lake. Sparkling Lake did not have any significant differences for NEP. For GPP and R, across both lakes, between 14% and 37% of the sites had estimates significantly different from the corresponding lake-wide mean value (Fig. 6).

Averaging metabolism across sites resulted in daily metabolism values that were much less variable over the deployment than observed from single sensors alone (Table 4). Daily metabolism estimates based on habitat-weighted averages of all site-specific values had ranges that were 20–50% lower than the average ranges across all sensors and were 63–83% smaller than the maximum ranges seen over all sites. Estimates across days were much less variable than estimates across sites, with only 4–5% of the variance in GPP and R in Sparkling Lake and 10–12% of the variance in Peter Lake GPP and R being attributable to day of deployment.

Littoral–pelagic differences—For both lakes, areal GPP and R were greater at pelagic locations than littoral locations based on Tukey's honestly significant difference ($p < 0.01$). Areal NEP was also greater for pelagic locations in Peter Lake ($p < 0.05$), but was not significantly different between habitats in Sparkling Lake. When significant, areal metabolism was 25–47% higher in pelagic zones than in littoral zones (Table 5). Volumetric metabolism rates (GPP, R, and NEP) were greater for littoral locations than pelagic locations in Sparkling Lake (by 43–76%, $p < 0.0001$), but there were no statistically significant

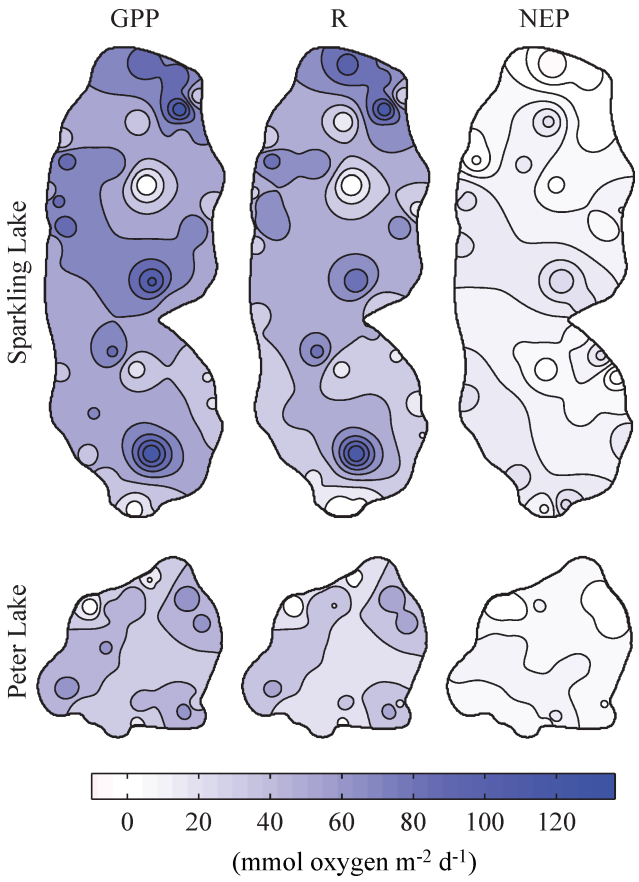

Fig. 4. Spatial pattern of mean metabolic rates (GPP, R, and NEP, mmol m^{-2} d^{-1}) averaged over all days for each site in each lake. The color scale is the same for both lakes and all components of metabolism. The contour interval is 20 mmol m^{-2} d^{-1} for GPP and R and 5 mmol m^{-2} d^{-1} for NEP for both lakes.

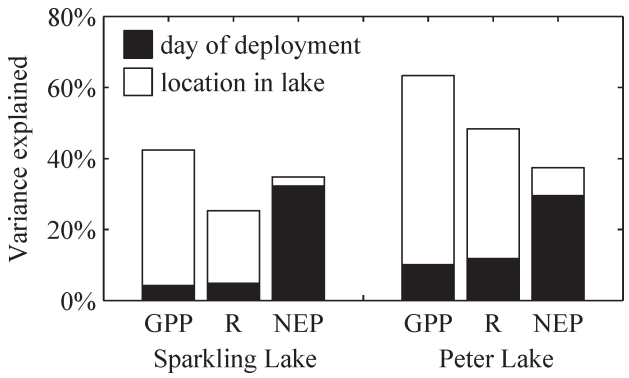

Fig. 5. Proportion of the variance in metabolism estimates attributable to location (open) or day of the deployment (shaded). The balance of the variance is attributable to residuals.

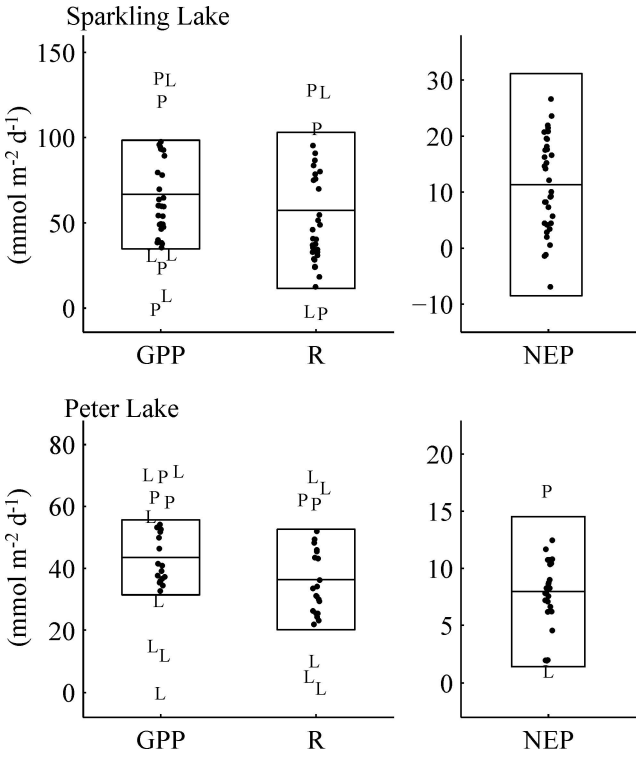

Fig. 6. Deployment-averaged mean GPP, R, and NEP for both lakes. Each point represents the 10- or 11-d average value at one location in the lake. The center of each box shows the lake- and deployment-wide mean, while the box boundaries show Fisher's LSD around the mean. Sites outside the box have statistically different means over the deployment from the lake-wide mean ($p < 0.05$) and are identified as P or L depending on whether they were located in the pelagic or littoral zones, respectively. Points are jittered about the x-axis to aid in distinguishing individual data points.

differences between pelagic and littoral volumetric rates in Peter Lake (Table 5).

Rarefaction—Increasing the number of sensors (k) used in estimating whole-lake epilimnetic metabolism reduces the standard deviation of multiple estimates using different subsets of k sensors, for all values of $k = 1$ to n (Fig. 7). In all cases (for GPP, R, and NEP in both lakes), the rarefaction analysis indicates that increasing the number of sensors from one to four reduced the standard deviation in the resulting estimate by half. Because precision is the reciprocal of variance, this means that four randomly selected sensors result in a metabolism estimate that is four times more precise than that based on a single sensor alone. Using two or three sensors resulted in estimates just over two and three times more precise, respectively, than single-site estimates alone. To reduce the standard deviation to just 25% of its single-site value (and thus increase precision 16-fold), 10–12 sensors are needed for our study lakes.

In order to obtain a standard deviation comparable to the benchmark of the standard deviation that can be attributed to the ANOVA-partitioned day-to-day variance over a 10-d period, seven and six sensors are needed in Peter Lake for GPP and R, respectively, while only two are needed for NEP. In Sparkling Lake, 14 and 12 sensors are needed for GPP and R to meet this benchmark, but only two are needed for NEP.

The standard deviation of metabolism estimates using k sensors is sensitive to the proportion of sensors coming from pelagic vs. littoral habitats. In the random sampling procedure, we ended up with a distribution of sensors taken from pelagic and littoral regions in roughly equal proportion to the numbers of sensors we placed in each habitat. Conducting the rarefaction procedure using randomly selected sensors with specified littoral–pelagic distributions revealed a strong dependence of the rarefaction curve on the distribution of those sensors within the lake. Sampling sensors systematically from the pool of available sensor locations revealed that placing sensors such that they disproportionately sample one habitat over the other (by a large margin) results in poorer precision of metabolism estimates (Fig. 8, darkest colored circles). Near-optimum precision was obtained when the sensors were drawn from a distribution where the density of sensors in each habitat was equal (white circles) and precision was maximized (standard deviation minimized) when the distribution of sensors in each habitat was nearest to the proportion of the area-weighted standard deviation attributable to each habitat. In Sparkling Lake, standard deviation was slightly larger in the pelagic habitat (Table 5); accordingly, the minimum points along the rarefaction curve (Fig. 8, light blue circles) correspond to the sensor subsets that include slightly more pelagic sensors relative to the proportion of the lake that is in the pelagic zone. The opposite was true of Peter Lake, where littoral standard deviation was slightly higher than pelagic and so the sensor subsets that minimize the standard deviation of the lake-wide estimate are those that have slightly more sensors in the littoral zone (light red circles).

Discussion

Spatial variation: Expectations vs. findings—We expected that the spatially intensive sampling of lake metabolism would reveal spatial patterns in estimates of GPP and R related to the littoral–pelagic gradient. For example, because of contact between the mixed layer and sediment, we expected metabolism nearshore to be higher (at least on a volumetric basis) than that in the pelagic zone due to the addition of benthic metabolism to the water column metabolism. Further, because previous studies found variation among littoral sites related to substrate type (Lauster et al. 2006), we expected that heterogeneity in substrate types (e.g., cobble, sand, muck, macrophytes) and variable water depth might lead to higher variability of metabolism estimates among littoral sites than among the more uniform pelagic sites. For this reason, we sampled the littoral zones with a higher density of sensors, especially in Sparkling Lake, which has a wider variety of substrate types. For pelagic sites, we expected to have less variation among the sites than we had in the littoral zone.

Contrary to our expectations, we generally found high variation among sites within both the littoral and pelagic regions of the lakes; we also found a large overlap in the ranges of metabolism estimates from littoral and pelagic

Table 4. Average areal metabolism (mmol m^{-2} d^{-1}) and coefficient of variation (in parentheses, %) across all sensors for each day of the deployment. $n = 35$ sensors per day in Sparkling Lake and 27 sensors per day in Peter Lake. No estimates are provided for GPP and NEP on the 11th day (—) because sensors were removed midday. SD is the standard deviation of daily estimates and SD_{day} is the standard deviation attributable solely to the day of deployment factor in a two-way site vs. day ANOVA.

	Sparkling Lake			Peter Lake		
Day	GPP	R	NEP	GPP	R	NEP
1	52(49)	48(65)	14(89)	36(43)	33(79)	2.7(222)
2	82(37)	28(90)	44(25)	47(33)	33(53)	17(31)
3	54(73)	49(52)	17(70)	36(39)	26(60)	13(41)
4	60(58)	24(164)	19(68)	41(42)	21(76)	13(21)
5	72(36)	58(83)	16(111)	36(32)	35(55)	−0.75(661)
6	70(50)	52(43)	4.3(500)	50(21)	39(29)	12(39)
7	48(82)	78(54)	−20(109)	36(31)	35(31)	4.0(113)
8	54(73)	58(98)	−13(155)	44(38)	29(48)	7.5(90)
9	77(49)	76(96)	13(201)	49(30)	45(46)	3.6(125)
10	98(41)	53(103)	19(68)	60(26)	46(39)	8.1(102)
11	—	106(60)	—	—	58(30)	—
Mean	67	57	11	44	36	8.0
SD	16	23	18	8	10	5.6
SD_{day}	9.8	14	16	7.2	9.2	5.1

sensors. We did see statistically significant differences between littoral and pelagic metabolism in Sparkling Lake for volumetric (GPP, R, and NEP) and areal metabolism (GPP and R only) and for Peter Lake areal metabolism (GPP, R, and NEP) (ANOVA, $p < 0.05$; Table 5). The differences in areal metabolism rates, however, were small, and likely detectable due to the statistical power of having a large number of samples. In comparison, the mean differences between littoral and pelagic GPP and R were smaller than the LSD calculated for comparing individual sites to the lake-wide mean; the average difference between littoral and pelagic sites was smaller than the difference necessary to consider estimates from two sites significantly different for the deployment.

Not only did metabolism estimates not follow the expected littoral to pelagic gradient, they seemed to not follow a spatial pattern at all. Variograms for each metabolism parameter revealed that there was no spatial autocorrelation in estimates at the scales we sampled for either lake. This is apparent from looking at the maps of metabolism: adjacent sites can differ greatly or be quite similar (Figs. 2, 3). For example, in Fig. 2, there are two pelagic sensors about one-third of the way up from the southern shore of Sparkling Lake that consistently yield estimates of metabolism on opposite sides of the daily mean value for both GPP and R. The lack of autocorrelation and spatial pattern does not preclude metabolism estimates being spatially correlated at shorter distances than we observed. Indeed, at some shorter distance, estimates must become correlated.

While previous studies have shown variation among littoral sites (Lauster et al. 2006) and differences between littoral and pelagic sites (Van de Bogert et al. 2007), few have investigated variability of metabolism estimates among pelagic sites using free-water oxygen measurements (but *see* Sadro et al. 2011). Researchers often consider the pelagic zone of lakes of this size well mixed and able to be characterized by measurements taken from a single

Table 5. Mean volumetric (e.g., GPP_v) and areal (e.g., GPP_a) metabolism estimates (with coefficient of variation in parentheses, over sites and days) by lake and habitat for the 10-d deployment. The 95% confidence interval for the mean difference between pelagic ($\bar{y}_P$) and littoral ($\bar{y}_L$) metabolism is based on Tukey–Kramer method for unequal sample sizes. ns indicates comparisons for which the difference in means was not significant (i.e., $p > 0.05$). n_L and n_P are the number of sensor-days included in the means for littoral and pelagic habitats, respectively.

	Sparkling Lake					Peter Lake				
	Littoral, $\bar{y}_L$ (%)	Pelagic, $\bar{y}_P$ (%)	$\bar{y}_P - \bar{y}_L$ ±95% CI	n_L	n_P	Littoral, $\bar{y}_L$ (%)	Pelagic, $\bar{y}_P$ (%)	$\bar{y}_P - \bar{y}_L$ ±95% CI	n_L	n_P
mmol oxygen m^{-3} d^{-1}										
GPP_v	25.1(69)	11.2(74)	−14±3.3**	230	120	17.0(65)	16.0(37)	ns	110	160
R_v	17.7(101)	10.1(112)	−7.6±3.3**	253	132	14.4(106)	13.2(56)	ns	121	176
NEP_v	7.2(185)	1.7(281)	−5.5±2.2**	230	120	3.6(134)	3.0(110)	ns	110	160
mmol oxygen m^{-2} d^{-1}										
GPP_a	57.5(76)	72.7(74)	15±10*	230	120	36.5(74)	48.0(37)	12±5.1*	110	160
R_a	44.6(125)	65.6(112)	21±13*	253	132	31.6(102)	39.5(56)	7.9±5.8*	121	176
NEP_a	12.1(220)	10.8(281)	ns	230	120	6.6(132)	8.9(110)	2.3±1.9	110	160

* $p<0.01$; ** $p<0.0001$.

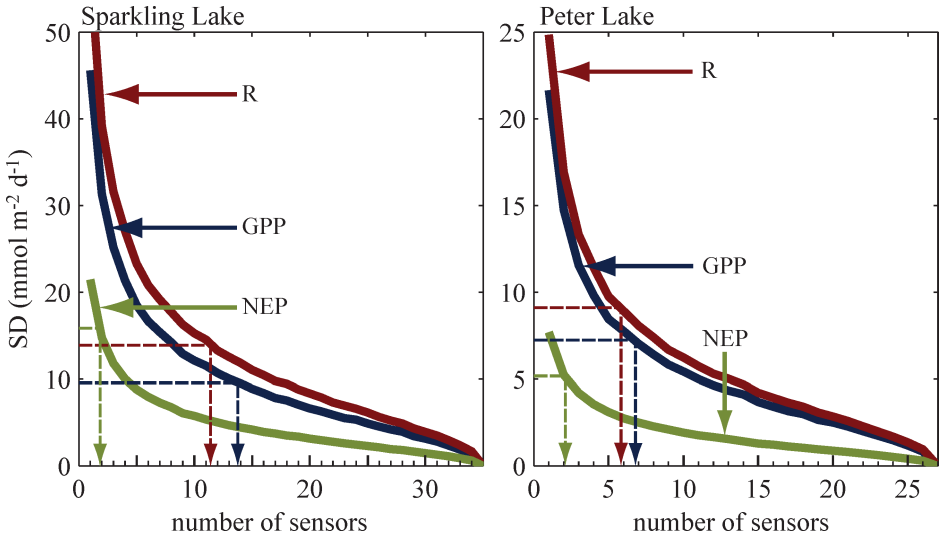

Fig. 7. Mean standard deviation (SD) of the estimate of whole-lake epilimnetic metabolism components (GPP, R, NEP) over all days of the deployment (y-axis; SD, mmol m^{-2} d^{-1}) vs. number of sensors used for the estimate (x-axis; k). For each x-axis value of k sensors, the bold line shows the standard deviation for up to 1000 estimates of metabolism obtained by a random selection of k sensors drawn without replacement from the pool of $n = 35$ sensors in Sparkling Lake or $n = 27$ sensors in Peter Lake (*see* text). The dashed lines identify the standard deviation attributable to day-to-day variation alone (over the 10-d deployment, two-way fixed-effects ANOVA) and the number of sensors necessary to reach that benchmark.

location. While this may be true for some limnological parameters (especially those with drivers acting at scales greater than the diel scale), this study clearly shows it is not the case for dissolved oxygen dynamics.

Variability in pelagic metabolism may be tied to both biological and physical processes. For instance, differential heating and cooling of water in the littoral zone may set up density-driven currents that can provide a potential explanation for some of the variability we see in metabolism estimates as well as why parameters with drivers acting at scales > 24 h may be more uniform in the lake. Water temperature anomalies (difference from lake-wide mean) are greatest by late afternoon and have more or less disappeared by just after sunrise the following day. These temperature differences can lead to transitional density-driven flows (Monismith et al. 1990; MacIntyre and Melack 1995). Unfortunately, we do not have sufficient physical data for our deployments to quantify the magnitudes of these flows nor the effect they may have on oxygen dynamics at a given site. Combining large arrays of sondes with equally large arrays of flow and temperature measurements seems like a productive avenue for future research.

The flows described above, if large enough, could create transient variability in drivers of metabolism and thus lead to spatially heterogeneous metabolism estimates. Indeed, pelagic zones have been shown to exhibit spatial heterogeneity in plankton populations, chlorophyll *a*, and nutrient concentrations (Pinel-Alloul and Pont 1991; Kamarainen et al. 2009; Mackay et al. 2011), all of which could translate into heterogeneity in metabolism estimates. However, sites with occasional GPP or R estimates that do not make sense biologically (e.g., "negative" respiration, or primary production in the dark) point to the possibility that at least some of the variation is driven by factors that are not biological, including the likely possibility of advection of water masses past the sensor.

High site-to-site variability in metabolism estimates (Figs. 2, 3; Table 3) calls into question the ability of any one site to represent the lake as a whole. Cases where long time series of daily metabolism estimates have been calculated often show large day-to-day variation that is only partially explained by driving variables (Staehr and Sand-Jensen 2007; Coloso et al. 2011), leaving much variability unexplained. The present study suggests that a large portion of this variation is not due to day-to-day differences at all, but rather is the result of variation within the lake itself on any given day (Fig. 5). In fact, only 4% (Sparkling Lake) to 12% (Peter Lake) of the total variance seen among locations over the 10-d deployment was attributable to day-to-day differences, while 20–50% of the variation seen was due to within-lake heterogeneity. Thus, lake-wide metabolism is much less variable, day to day, than the value observed at any one site. Further, much of the variance observed from single-sensor metabolism estimates is likely not daily variance, but the result of only sampling a small portion of the lake ecosystem. The implications of this result are striking for the interpretation of metabolism studies and lake monitoring.

Rarefaction: How many sensors are needed for daily estimates?—The cost of sensors and effort for upkeep make extensive spatial sampling prohibitive in most cases. Fortunately, an examination of the tradeoffs between sample size and confidence in metabolism estimates indicates that substantial gains can be made with just a

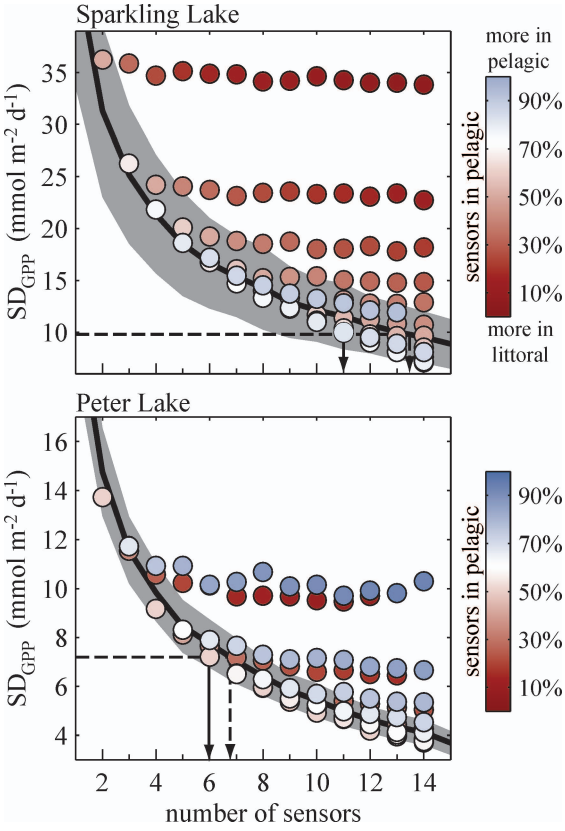

Fig. 8. Habitat-targeted rarefaction using GPP as a representative example in each lake. The dark black line is the rarefaction curve from a random selection of sensors from the pool. The shaded area is ± standard deviation (SD) from the mean rarefaction curve. For each x-value, 2 to 14, there are $x - 1$ points on the graph each of which shows the standard deviation of up to 1000 estimates using x sensors. The points represent all possible combinations of littoral and pelagic sensors for the given value of x, where there is at least one sensor in each habitat, e.g., for $x = 5$, there are four points on the graph corresponding to ratios of pelagic to littoral sensors of 1 : 4, 2 : 3, 3 : 2, and 4 : 1. The color of the points identifies the proportion of x sensors that are in the pelagic zone. The color bar is scaled such that the ratio corresponding to an equal density of sensors in each habitat is white and color intensifies as the sensors are concentrated in littoral (red) or pelagic (blue) habitats. The dashed line shows the benchmark of day-to-day standard deviation and the associated number of sensors needed to meet that benchmark (arrowhead). The vertical black line shows the updated number of sensors needed using the targeted approach.

few additional sensors. Confidence in metabolism estimates goes up with each additional sensor used, but with diminishing returns. While there is no clear-cut answer to the question of how many locations need to be sampled, we have examined some benchmarks for comparison. Using the partitioned day-to-day variance as a benchmark, we evaluated the number of sensors needed to match the precision of the daily metabolism estimate to the level of precision obtained given the temporal variation over the 10-d deployment. Obtaining estimates that reduce the uncertainty attributable to spatial heterogeneity to the level of uncertainty for a week's worth of values bring the result into a range that many ecologists could accept. At between 10 and 14 sensors for Sparkling Lake GPP and R, respectively, this target is daunting and remains beyond the reach of most research programs (including our own). However, between five and seven sensors are needed to meet this target for Peter Lake GPP and R which, while challenging, may be an attainable number, especially as prices for dissolved oxygen sensors decline. Because estimates of GPP and R are correlated, NEP (GPP − R) is somewhat decoupled from the spatial variation. Indeed, for estimates of NEP, just two randomly placed sensors achieve a level of precision on par with 10-d temporal precision for both lakes. Finally, because of the initial steepness of the rarefaction curve, the use of just four sensors increases the precision of the estimate substantially—in our case fourfold. Although more multiple-sonde studies are needed, these results suggest large gains can be made with relatively few additional sensors.

Because we did not find a clear and strong pattern between habitat and metabolism values, we initially conducted the rarefaction procedure by randomly selecting sensors from the pool of locations where sensors were deployed. This pool had nearly twice the density of sensors in the littoral zone than pelagic for Sparkling Lake, but was much closer to equal densities between habitats in Peter Lake. Targeting our selection of sensors to specific habitats showed a small potential for improved estimates over our a priori selection of sites (Fig. 8). We show results only for GPP, but the patterns and conclusions are the same for R and NEP as well. The lowest standard deviation was achieved when sensors were allocated to habitats in proportion to their contribution to the area-weighted littoral–pelagic standard deviation over all sites and days. For Sparkling Lake, a slightly higher density of sensors in the pelagic was needed to achieve the lowest standard deviation possible, reflecting the slightly higher variance among pelagic estimates in Sparkling Lake relative to littoral estimates. In Peter Lake, the reverse was true. With only a few sensors, however, it may not be possible to achieve this precise density (e.g., with four sensors, it is only possible to have 25%, 50%, or 75% of the sensors in either habitat). In this case, getting close to an even density between habitats is the best goal. Furthermore, without doing a study of the spatial variability in a particular lake, judging this optimum allocation is not possible. Getting close, with an approximate distribution of sensors equal to the relative proportion of the lake in each habitat is much better than over-sampling one or the other habitat.

Substituting time for spatial coverage—Although our data demonstrate that a single-sensor location chosen at random is not likely to represent the lake-wide mean determined by multiple sensor locations for a single day, many researchers are interested in time frames longer than a single day. It is likely that deployments for a week, month, or entire season at a single site may come closer to the lake-wide mean. Of the three metabolic results, 10-d mean NEP had the least spatial variability in both lakes. Site-to-site differences in NEP were not significantly different from zero in Sparkling Lake (ANOVA, $p >$

0.05). An ANOVA revealed that not all sites' NEP estimates were equal for Peter Lake ($p < 0.01$), but Fisher's LSD post hoc test identified only two sites (one pelagic and one littoral) whose mean NEP over the deployment was significantly different from the lake-wide mean. For R, 86% and 74% of the sites were within the LSD for Sparkling and Peter Lakes, respectively, and for GPP, 77% and 63% of sites were within the LSD. Thus, there is some indication that longer time averaging at a single site may provide an adequate weekly or monthly value of lake-wide metabolism. We caution, however, that some sites in both lakes have 10-d means that are either significantly higher or lower than the 10-d mean across all sites. These results suggest that some exploration of spatial variation is essential, but that fewer sensors might be adequate if the primary interest is in time frames longer than a single day.

While estimates of metabolism can vary greatly from one location within a lake to the next over periods of days, the challenges are not insurmountable. It can no longer be assumed that a single location is an accurate representation of a lake at the daily time scale. Fortunately, 20 sensors are not necessary; using even four would provide considerably more confidence in the daily lake-wide estimates and longer-term averages may be adequate.

Acknowledgments

The deployment of this number of sensors was possible because of the support of many projects and people. Sensors came from our own projects in addition to the North Temperate Lakes Long Term Ecological Research (NTL-LTER) project, Jim Hodgson of Saint Norbert College, and John Walker of the United States Geological Survey. We thank Jim Coloso, Ryan Kroiss, Luke Winslow, and Tim Meinke for help in deploying sondes and sensor platforms. We are grateful for the use of facilities at the University of Wisconsin–Trout Lake field station and the University of Notre Dame Environmental Research Center, and we thank Tim Kratz and Gary Belovsky for making these available to us. We thank Emily Stanley, Monica Turner, Steve Loheide, and two anonymous reviewers for comments which helped to improve this manuscript. This work was supported by the National Science Foundation's Division of Environmental Biology, the NTL-LTER program, the Mellon Foundation, and the Dorothy Powers Grant and Eugene Lodewick Grant memorial scholarship fund.

References

Blukacz, E. A., W. G. Sprules, B. J. Shuter, and J. P. Richards. 2010. Evaluating the effect of wind-driven patchiness on trophic interactions between zooplankton and phytoplankton. Limnol. Oceanogr. **55:** 1590–1600, doi:10.4319/lo.2010.55.4.1590

Caraco, N. F., and J. J. Cole. 2002. Contrasting impacts of a native and alien macrophyte on dissolved oxygen in a large river. Ecol. Appl. **12:** 1496–1509, doi:10.1890/1051-0761(2002)012[1496:CIOANA]2.0.CO;2

Caron, D. A., and others. 2008. Macro- to fine-scale spatial and temporal distributions and dynamics of phytoplankton and their environmental driving forces in a small montane lake in southern California, USA. Limnol. Oceanogr. **53:** 2333–2349, doi:10.4319/lo.2008.53.5_part_2.2333

Cole, J. J., and N. F. Caraco. 1998. Atmospheric exchange of carbon dioxide in a low-wind oligotrophic lake measured by the addition of SF_6. Limnol. Oceanogr. **43:** 647–656, doi:10.4319/lo.1998.43.4.0647

———, M. L. Pace, S. R. Carpenter, and J. F. Kitchell. 2000. Persistence of net heterotrophy in lakes during nutrient addition and food web manipulations. Limnol. Oceanogr. **45:** 1718–1730, doi:10.4319/lo.2000.45.8.1718

———, and others. 2007. Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon budget. Ecosystems **10:** 171–184, doi:10.1007/s10021-006-9013-8

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, doi:10.1007/s10021-011-9455-5

Diggle, P., and P. J. Ribeiro. 2007. Model-based geostatistics. Springer.

Fisher, S. G., and G. E. Likens. 1973. Energy flow in Bear Brook, New Hampshire: An integrative approach to stream ecosystem metabolism. Ecol. Monogr. **43:** 421–439, doi:10.2307/1942301

Gotelli, N. J., and A. M. Ellison. 2004. A primer of ecological statistics. Sinauer Associates.

Hanson, P. C., D. L. Bade, S. R. Carpenter, and T. K. Kratz. 2003. Lake metabolism: Relationships with dissolved organic carbon and phosphorus. Limnol. Oceanogr. **48:** 1112–1119, doi:10.4319/lo.2003.48.3.1112

Heck, K. L., G. Vanbelle, and D. Simberloff. 1975. Explicit calculation of rarefaction diversity measurement and determination of sufficient sample size. Ecology **56:** 1459–1461, doi:10.2307/1934716

Jahne, B., K. O. Munnich, R. Bosinger, A. Dutzi, W. Huber, and P. Libner. 1987. On the parameters influencing air-water gas-exchange. J. Geophys. Res. (C Oceans) **92:** 1937–1949, doi:10.1029/JC092iC02p01937

Kamarainen, A. M., H. L. Yuan, C. H. Wu, and S. R. Carpenter. 2009. Estimates of phosphorus entrainment in Lake Mendota: A comparison of one-dimensional and three-dimensional approaches. Limnol. Oceanogr.: Methods **7:** 553–567, doi:10.4319/lom.2009.7.553

Lauster, G. H., P. C. Hanson, and T. K. Kratz. 2006. Gross primary production and respiration differences among littoral and pelagic habitats in northern Wisconsin lakes. Can. J. Fish. Aquat. Sci. **63:** 1130–1141, doi:10.1139/f06-018

Lindeman, R. L. 1942. The trophic dynamic aspect of ecology. Ecology **23:** 399–418, doi:10.2307/1930126

MacIntyre, S., and J. M. Melack. 1995. Vertical and horizontal transport in lakes: Linking littoral, benthic, and pelagic habitats. J. N. Am. Benthol. Soc. **14:** 599–615, doi:10.2307/1467544

Mackay, E. B., I. D. Jones, S. J. Thackeray, and A. M. Folkard. 2011. Spatial heterogeneity in a small, temperate lake during archetypal weak forcing conditions. Fundam. Appl. Limnol. **179:** 27–40, doi:10.1127/1863-9135/2011/0179-0027

Monismith, S. G., J. Imberger, and M. L. Morison. 1990. Convective motions in the sidearm of a small reservoir. Limnol. Oceanogr. **35:** 1676–1702, doi:10.4319/lo.1990.35.8.1676

Odum, H. T. 1956. Primary production in flowing waters. Limnol. Oceanogr. **1:** 102–117, doi:10.4319/lo.1956.1.2.0102

———. 1957*a*. Primary production measurements in eleven Florida springs and a marine turtle-grass community. Limnol. Oceanogr. **2:** 85–97, doi:10.4319/lo.1957.2.2.0085

———. 1957*b*. Trophic structure and productivity of Silver Springs, Florida. Ecol. Monogr. **27:** 55–112, doi:10.2307/1948571

Pinel-Alloul, B., and D. Pont. 1991. Spatial-distribution patterns in fresh-water macrozooplankton—variation with scale. Can. J. Zool. **69:** 1557–1570, doi:10.1139/z91-219

R Development Core Team. 2011. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing [accessed 2011 December 1], Available from http://www.R-project.org/

Sadro, S., J. M. Melack, and S. MacIntyre. 2011. Spatial and temporal variability in the ecosystem metabolism of a high-elevation lake: Integrating benthic and pelagic habitats. Ecosystems **14:** 1123–1140, doi:10.1007/s10021-011-9471-5

Serra, T., J. Vidal, X. Casamitjana, M. Soler, and J. Colomer. 2007. The role of surface vertical mixing in phytoplankton distribution in a stratified reservoir. Limnol. Oceanogr. **52:** 620–634, doi:10.4319/lo.2007.52.2.0620

Simberloff, D. 1972. Properties of rarefaction diversity measurement. Am. Nat. **106:** 414–418, doi:10.1086/282781

Smith, S. V. 1985. Physical, chemical and biological characteristics of CO_2 gas flux across the air water interface. Plant Cell Environ. **8:** 387–398, doi:10.1111/j.1365-3040.1985.tb01674.x

Staehr, P. A., and others. 2010. Lake metabolism and the diel oxygen technique: State of the science. Limnol. Oceanogr.: Methods **8:** 628–644, doi:10.4319/lom.2010.8.0628

———, and K. Sand-Jensen. 2007. Temporal dynamics and regulation of lake metabolism. Limnol. Oceanogr. **52:** 108–120, doi:10.4319/lo.2007.52.1.0108

Vadeboncoeur, Y., J. Kalff, K. Christoffersen, and E. Jeppesen. 2006. Substratum as a driver of variation in periphyton chlorophyll and productivity in lakes. J. N. Am. Benthol. Soc. **25:** 379–392, doi:10.1899/0887-3593(2006)25[379:SAADOV]2.0.CO;2

———, and D. M. Lodge. 2000. Periphyton production on wood and sediment: Substratum-specific response to laboratory and whole-lake nutrient manipulations. J. N. Am. Benthol. Soc. **19:** 68–81, doi:10.2307/1468282

———, M. J. Vander Zanden, and D. M. Lodge. 2002. Putting the lake back together: Reintegrating benthic pathways into lake food web models. BioScience **52:** 44–54, doi:10.1641/0006-3568(2002)052[0044:PTLBTR]2.0.CO;2

Van de Bogert, M. C., S. R. Carpenter, J. J. Cole, and M. L. Pace. 2007. Assessing pelagic and benthic metabolism using free water measurements. Limnol. Oceanogr.: Methods **5:** 145–155, doi:10.4319/lom.2007.5.145

Wanninkhof, R. 1992. Relationship between wind-speed and gas-exchange over the ocean. J. Geophys. Res. (C, Oceans) **97:** 7373–7382, doi:10.1029/92JC00188

Wetzel, R. G. 1995. Death, detritus and energy flow in aquatic ecosystems. Freshw. Biol. **33:** 83–89, doi:10.1111/j.1365-2427.1995.tb00388.x

Associate editor: David A. Caron

Received: 05 December 2011
Accepted: 25 July 2012
Amended: 05 August 2012