AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2014GL061824

Key Points:

- Data aggregation biases estimates of regional and global lake CO_2 emissions
- The direction of bias depends on lake size
- Covariance-based correction factors adjust for aggregation bias

Supporting Information:

- Readme
- Text S1, Figure S1 and Tables S1 and S2

Correspondence to:
D. A. Seekell,
david.seekell@emg.umu.se

Citation:
Seekell, D. A., J. A. Carr, C. Gudasz, and J. Karlsson (2014), Upscaling carbon dioxide emissions from lakes, *Geophys. Res. Lett.*, *41*, 7555–7559, doi:10.1002/2014GL061824.

Received 12 SEP 2014
Accepted 14 OCT 2014
Accepted article online 16 OCT 2014
Published online 6 NOV 2014

Upscaling carbon dioxide emissions from lakes

David A. Seekell[1,2], Joel A. Carr[2], Cristian Gudasz[3], and Jan Karlsson[1]

[1]Department of Ecology and Environmental Science, Umeå University, Umeå, Sweden, [2]Department of Environmental Sciences, University of Virginia, Charlottesville, Virginia, USA, [3]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey, USA

Abstract Quantifying CO_2 fluxes from lakes to the atmosphere is important for balancing regional and global-scale carbon budgets. CO_2 emissions are estimated through statistical upscaling procedures that aggregate data from a large number of lakes. However, aggregation can bias flux estimates if the physical and chemical factors determining CO_2 exchange between water and the atmosphere are not independent. We evaluated the magnitude of aggregation biases with moment expansions and pCO_2 data from 5140 Swedish lakes. The direction of the aggregation bias depends on lake size; mean flux was overestimated by 4% for small lakes (0.01–0.1 km^2) but underestimated by 13% for large lakes (100–1000 km^2). Simple covariance-based correction factors were generated to adjust for upscaling biases. These correction factors represent an easily interpretable and implemented approach to improving the accuracy of regional and global estimates of lake CO_2 emissions.

1. Introduction

Quantifying CO_2 fluxes from lakes to the atmosphere is important for balancing regional and global carbon budgets [e.g., *Buffam et al.*, 2011; *Raymond et al.*, 2013; *IPCC*, 2013]. Estimates of global lake CO_2 emissions range from 0.07 to 0.86 Pg C yr^{-1}, and are significant relative to estimates of terrestrial and marine CO_2 sequestration [*Cole et al.*, 2007; *Battin et al.*, 2009; *Marotta et al.*, 2009; *Raymond et al.*, 2013]. Therefore, it is not surprising that constraining estimates of lake CO_2 emissions to better understand inland waters in a global context is a priority for limnological research [*Cole et al.*, 2007; *Downing*, 2009; *Tranvik et al.*, 2009].

Most lakes are small and their CO_2 fluxes cannot be resolved within the spatial resolution of top-down approaches like eddy covariance [*Battin et al.*, 2009; *Downing*, 2009]. Consequently, statistical upscaling procedures are used to estimate lake CO_2 emissions [*Battin et al.*, 2009; *Downing*, 2009]. Upscaling is the application of physical models of diffusive gas transfer from one system, to data aggregated from many lakes [*Downing*, 2009; *Marotta et al.*, 2009; *Raymond et al.*, 2013]. This approach is necessary because it is logistically infeasible to sample all lakes in a lake-rich region. However, applying physical models for fine-scale dynamics of individual systems to aggregated data creates inherent biases in estimates because of correlations between and variations among the aggregated components [*Rastetter et al.*, 1992; *Lim and Roderick*, 2014]. Specifically, upscaling can produce a biased estimate of mean CO_2 flux if the relationships between the physical and chemical factors that determine the diffusive flux are nonlinear; aggregating by taking a mean represents a first-order, linear approximation [*Rastetter et al.*, 1992; *Lim and Roderick*, 2014]. The magnitude of this bias increases with the strength of correlation between factors and with the variance of individual factors [*Rastetter et al.*, 1992; *Lim and Roderick*, 2014]. Such biases may be present in upscaling lake CO_2 emissions, but to our knowledge the magnitude of these potential aggregation errors has not been assessed.

To date, global limnology has mostly focused on revising estimates of global lake CO_2 emissions by including additional samples in data aggregations [*Tranvik et al.*, 2009; *Marotta et al.*, 2009; *Raymond et al.*, 2013]. However, adding data to aggregations does not necessarily eliminate biases, and in some cases it creates new errors [cf. *Stumpf and Porter*, 2012; *Seekell et al.*, 2013]. Despite this, studies evaluating methodological approaches for global limnology are scarce [e.g., *Lewis*, 2011; *Seekell and Pace*, 2011; *Seekell et al.*, 2013]. Here we evaluate the accuracy of statistical upscaling of lake CO_2 fluxes using a moment expansion. We use data from 5140 Swedish lakes to show bias in upscaling due to the correlations between and variation within the physical and chemical factors that determine the exchange of CO_2 between water and the atmosphere. In order to contribute to improving the accuracy of CO_2 emission estimates from lakes, we provide covariance-based correction factors to reduce this bias in upscaling analyses.

2. Upscaling Theory

The CO_2 flux (FCO_2, mol d^{-1}) from an individual lake is estimated based on Fick's law for diffusive flux across a concentration gradient [e.g., *Marotta et al*., 2009; *Raymond et al*., 2013]. Specifically, flux is the product of the gradient between the concentrations of CO_2 in the water and in the air (denoted here as ΔCO_2, mol m^{-3}), a gas transfer coefficient that represents the physical processes influencing the rate of gas transfer from water to air (k, m d^{-1}), a chemical enhancement factor that accounts for increased flux due to buffering reactions along the within-lake CO_2 gradient in alkaline conditions (α, dimensionless), and the surface area (A, m^2) of the lake [*Kuss and Schneider*, 2004].

$$F_{CO_2} = \alpha\ k\ \Delta CO_2 A$$

To estimate the average CO_2 flux of all lakes within a region, the factors comprising flux are measured for a subset of lakes and averaged. Fick's law is applied to the aggregated values of each parameter [e.g., *Marotta et al*., 2009].

$$\overline{F_{CO_2}} = \left(\overline{\alpha}\,\overline{k}\,\overline{\Delta CO_2}\,\overline{A}\right)$$

This average flux is multiplied by the total number of lakes in the region to estimate the overall flux. The total number of lakes, assumed or known exactly, is obtained from either high-resolution remotely sensed imagery, compilations of maps, or from statistical extrapolations [e.g., *Tranvik et al*., 2009; *Verpoorter et al*., 2012, 2014]. Other approaches to upscaling are possible (i.e., average the fluxes of individual lakes), but we analyze this product-of-means approach specifically because it has been applied in at least two important global-scale analyses of CO_2 emissions from lakes [*Marotta et al*., 2009; *Aufdenkampe et al*., 2011]. Additionally, a third global-scale study, *Raymond et al*. [2013], used a similar aggregation-based approach, although these authors applied different measures of central tendency for different parameters.

Bias in the upscaling equation can be evaluated by calculating covariance-based correction factors derived from a moment expansion (Taylor series expansions truncated to low orders) around the means of the factors determining flux $\left(\overline{\alpha}, \overline{k}, \overline{\Delta CO_2}, \overline{A}\right)$ [*Rastetter et al*., 1992; *Lim and Roderick*, 2014]. After, simplification, the second order expansion is

$$\overline{F_{CO_2}}^* = \overline{\alpha}\,\overline{k}\,\overline{\Delta CO_2}\,\overline{A}\left(1 + \chi_{\alpha,k} + \chi_{\alpha,\Delta CO_2} + \chi_{\alpha,A} + \chi_{k,\Delta CO_2} + \chi_{kA} + \chi_{\Delta CO_2,A}\right)$$

where χ is a bias correction coefficient calculated for each paired comparison of factors that is the product of the correlation between 2 factors and each of the factors' coefficient of variation [*Lim and Roderick*, 2014]. These bias correction coefficients can be summed to provide an overall bias correction term or intercompared in order to evaluate the relative contributions of different factors to the overall bias. Multiplying the bias correction term by the upscaled mean flux estimate provides an improved mean flux estimate, which can then be applied given estimates of lake abundance to generate overall CO_2 flux estimates.

3. Empirical Analysis

To evaluate the magnitude of aggregation biases in upscaling, we calculated covariance-based correction factors based on physical and chemical data for 5140 Swedish lakes. These lakes were sampled as part of national lake chemistry monitoring survey conducted between 1990 and 2012 (chemical data freely available from the Swedish University of Agricultural Sciences: http://www.slu.se/vatten-miljo; surface areas for the sampled lakes from *Nisell et al*. [2007]). The lakes, selected by a stratified (by size) random sample, were spread across the entire country. The samples were collected from surface waters in the autumn when the lakes are not stratified. We excluded records where there were no water temperature or total organic carbon measurements, where water temperature was <3°C at the time of sampling, where alkalinity was <0.005 meq L^{-1}, or where pH <5.2.

For each lake, α and ΔCO_2 were derived from surface water chemistry. Specifically, we estimated α from a previously published empirical relationship with pH, and CO_2 concentrations in lake water were calculated from pH, dissolved inorganic carbon, and ionization constants adjusted for water temperature [*Kuss and Schneider*, 2004; *Weyhenmeyer et al*., 2012]. The gas transfer velocity k for each lake was estimated based on a previously published empirical relationship with lake area [*Read et al*., 2012]. When a lake was sampled more

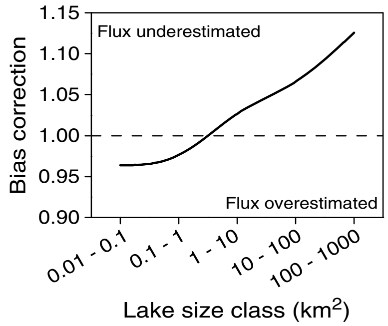

Figure 1. Bias correction terms for five logarithmic size classes of Swedish lakes. The bias correction term is multiplied by the mean flux estimate to create an unbiased estimate. Hence, a value of one indicates no bias. Values greater than one indicate that flux is underestimated and values less than one indicate that flux is overestimated. A spline is used to emphasize the pattern across size classes. Specific correction factors for each size class are reported in Table S1 in the supporting information.

than once, we averaged the calculated values for α, k, and ΔCO_2 prior to our analysis. Detailed formulae for our calculations are provided in the supporting information for the paper.

The covariance-based correction factors allow us to evaluate the magnitude of aggregation bias, rank the relative importance of different factors in creating this bias, and to correct for these biases in upscaling analyses [*Lim and Roderick*, 2014]. For our assessment, we first partitioned lakes into logarithmic size classes and then calculated correction factors for each size class. Partitioning lakes into logarithmic size classes is a common approach to minimizing bias in upscaling by reducing variance [e.g., *Rastetter et al.*, 1992; *Raymond et al.*, 2013]. We do not consider lakes $<0.01\ km^2$ in size because lakes below this size class are not accurately recorded on maps [*Verpoorter et al.*, 2012]. Additionally, we do not consider the largest lakes in Sweden ($>1000\ km^2$) in our analysis because their abundance is too low to calculate correction factors.

Covariance based correction factors evaluate bias in the mean flux, but overall flux is strongly influenced by the asymmetry in abundance across the lake size spectrum. We evaluated the magnitude of aggregation bias relative to overall lake CO_2 emissions by upscaling CO_2 fluxes for Swedish lakes ($0.01 - 1000\ km^2$) based on surface area and abundance data from high-resolution lake census [*Verpoorter et al.*, 2012]. The total flux (mol d^{-1}) is calculated by summing across lake size classes before and after applying size-class specific correction factors.

4. Results

$\overline{F_{CO_2}}$ is biased due to aggregation, but the direction and magnitude of bias depends on lake size (Figure 1). For instance, $\overline{F_{CO_2}}$ is overestimated by 3.6% for small lakes (i.e., $<0.1\ km^2$), but underestimated by about 12.6% for large lakes ($>100\ km^2$) (Figure 1). This bias is due to both the relationship between gas transfer and lake area, and the relationship between CO_2 and lake area (Figure 2). Gas transfer has a positive relationship with surface area across the full spectrum of lake sizes, although with considerable residual variability [*Read et al.*, 2012; *Vachon and Prairie*, 2013]. CO_2 has a negative relationship with surface area for small and medium lakes (e.g., $<10\ km^2$) but a positive relationship for large lakes (e.g., $>100\ km^2$). Hence, biases due to k and CO_2 act to mitigate each other for small lakes. For large lakes, the biases are in the same direction, additively increasing the overall bias (Figure 2). Biases related to other factors are small and similar across the lake size spectrum (Figure S1 in the supporting information).

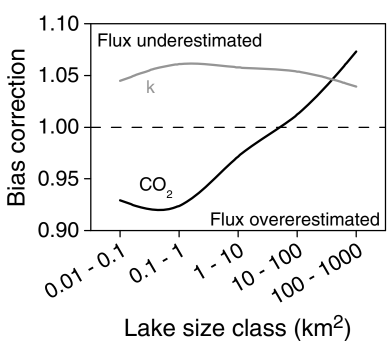

Figure 2. Bias correction terms due to covariance between k and A, and CO_2 and A, for five logarithmic size classes of Swedish lakes. For small lakes, bias due to aggregation offsets itself, mitigating the overall effect of bias on mean CO_2 flux. A spline is used to emphasize the pattern across size classes. Specific correction factors for each size class are reported in Table S2 in the supporting information.

After balancing the asymmetry in lake abundance and size through upscaling, aggregation bias resulted in an underestimate of overall emissions by about 4%. Specifically, CO_2 fluxes from lakes $0.01–1000\ km^2$ are 2.062×10^9 mol/d without correcting for aggregation bias. After bias correction, the emissions estimate is 2.143×10^9 mol/d.

5. Discussion

This study shows that upscaling CO_2 emissions from lakes at the regional scale (Sweden) is biased due to correlations between and variability within the physical and chemical factors comprising Fick's law. This bias is inherent to upscaling and mostly originates from variability within and relationships between lake area, k, and

ΔCO_2. Size-class specific covariance-based correction factors correct for biases, improving the accuracy of upscaling analyses [*Lim and Roderick*, 2014]. Aggregation bias is a general feature of regional and global upscaling analyses and cannot be overcome directly with additional sampling.

Upscaling without bias is possible with the statistical expectation function if the probability density functions are accurately described for process of interest [*Rastetter et al*., 1992]. This is not possible for upscaling CO_2 emissions from lakes because the distributions of relevant physical and chemical parameters are either poorly described or change geographically [e.g., *Seekell and Pace*, 2011; *Seekell et al*., 2013]. The bias correction coefficients presented here have no assumptions about the underlying distributions of the variables and hence are particularly useful for application to upscaling until theories or comprehensive descriptions of the statistical distributions of limnological variables are better understood [*Rastetter et al*., 1992].

The full extent of the impact of aggregation bias on regional and global upscaling estimates will depend on the specific conditions of each analysis. Because the direction of bias changes across size classes, it is possible that positive and negative biases cancel out and result in an overall estimate of CO_2 flux that is minimally or unbiased. Whether or not this happens will depend on regional characteristics, including the magnitude of bias within each size class and the relative contribution of small versus large lakes to the total surface area of lakes, which varies regionally such that the overall bias may be larger or smaller in other regions [*Downing et al*., 2006; *Verpoorter et al*., 2012]. For regions where overall estimates are not biased, the aggregations biases could still have important implications for understanding the contributions of small versus large lakes to landscape carbon cycling [*Downing*, 2010]. For example, the results herein indicate that small lakes contribute less in terms of CO_2 flux relative to large lakes than would be suggested by an uncorrected upscaling analysis.

We partitioned lakes in logarithmic size classes prior to analysis, but the covariance-based correction factors could also be calculated without partitioning. Analyses that do not partition in size classes will have much higher magnitudes of bias because variation in lake surface area will increase by orders of magnitude relative to the within-partition variation in lake size. Although partitioning samples by logarithmic lake size-class prior to upscaling does not eliminate aggregation bias, the combination of partitioning by lake size and bias correction coefficients should minimize most aggregation biases in $\overline{F_{CO_2}}$.

Partitioning by lake size may also minimize biases due to unequal probability sampling—when measurements from large lakes are overrepresented in databases relative to small lakes. This sampling bias is common to the type of environmental monitoring data that often form the basis for upscaling and can cause errors in upscaled estimates [*Wagner et al*., 2008]. Unequal probability sampling bias does not influence our analysis because lakes were sampled randomly within logarithmic lake size classes. The covariance based correction factors presented in this analysis will not correct for unequal probability sampling, but partitioning by lake size may reduce both unequal probability sampling bias (lakes within size-class partitions should have a more similar probability of being sampled than lakes between size-class partitions) and aggregation bias due to variability in lake area.

There are at least five other sources of potential errors in upscaling that we do not address in our analysis: (1) accuracy of the estimates of lake abundance, (2) geographic biases in sampling, (3) difficulties constraining the gas transfer velocity within and among lakes, (4) uncertainty in the accuracy of calculated compared to directly measured pCO_2 values, and (5) and the use of one or a few spot samples to extrapolate to annual fluxes [*Marotta et al*., 2009; *Tranvik et al*., 2009; *Raymond et al*., 2013; *Schilder et al*., 2013; *Morales-Pineda et al*., 2014]. These sources of error are not inherent to upscaling, and rather are due to the logistic challenges of data collection and the subsequent use of broad-scale data sets not originally intended and not optimally suited for upscaling. All of these sources of errors can be overcome by additional data collection. For instance, new remote sensing techniques are eliminating errors due to inaccuracy in lake abundance and size distributions [e.g., *Seekell and Pace*, 2011; *Verpoorter et al*., 2012, 2014], and constraining values of k is currently an area of active research [e.g., *Read et al*., 2012; *Vachon and Prairie*, 2013; *Schilder et al*., 2013]. However, some other errors may be difficult to overcome due to practical limitations in data collection. For instance, year-round measurements of pCO_2 are scarce and it may be logistically infeasible to collect both year-round measurements and sample the large number of lakes typically needed for estimating regional- or global-scale fluxes. Our analysis is not immune from these types of uncertainties, which may be substantial. For instance, *Marotta et al*. [2009] found that global estimates of CO_2 flux from lakes could double by using less conservative estimates of gas transfer coefficients (but still within the range of field observations).

Raymond et al. [2013] found that the range of uncertainty in a global estimate of CO_2 flux from lakes was twice the magnitude of an average estimate, based on an error analysis that resampled from the distributions of flux component across lakes. While intensive sampling could overcome these errors, it will not resolve the underlying aggregation bias that we describe here which is inherent to upscaling itself.

The results of this study are important for upscaling CO_2 emissions from inland waters. We show that aggregation biases exist and may influence both upscaled estimates of CO_2 flux and understanding of the relative contributions of small versus large lakes to CO_2 emissions. We emphasize that aggregation bias is a systematic error in upscaling itself and is not due to sampling logistics. Hence, this study emphasizes that aggregation bias that should be considered in future upscaling analyses, even if it is smaller than other upscaling uncertainties, because it is a known and easily corrected error [cf. *Sprugel*, 1983]. Despite advancements in measurements techniques and improved spatial and temporal resolution in available data on CO_2 fluxes, developing methods for upscaling will remain as an important task for accurate assessments of the role of inland waters in regional and global C cycles.

Acknowledgments
The chemical data for this paper are available from the Swedish University of Agricultural Sciences at http://www.slu.se/en/departments/aquatic-sciences-assessment/. The lake size data for this paper are from *Verpoorter et al*. [2012] and *Nisell et al*. [2007] and are available through Uppsala University.This paper is based on research supported by the U.S. National Science Foundation (Graduate Research Fellowship), the Virginia Space Grant Consortium (Graduate STEM Research Fellowship), the Carl Trygger Foundation for Scientific Research, Umeå University's Department of Ecology and Environmental Science, and the University of Virginia's Department of Environmental Sciences.

Wolfgang Knorr thanks two anonymous reviewers for their assistance in evaluating this paper.

References

Aufdenkampe, A. K., E. Mayorga, P. A. Raymond, J. M. Melack, S. C. Doney, S. R. Alin, R. E. Aalto, and K. Yoo (2011), Riverine coupling of biogeochemical cycles between land, oceans, and atmosphere, *Front. Ecol. Environ.*, *9*, 53–60, doi:10.1890/100014.

Battin, T. J., S. Luyssaert, L. A. Kaplan, A. K. Aufdenkampe, A. Richter, and L. J. Tranvik (2009), The boundless carbon cycle, *Nat. Geosci.*, *2*, 598–600, doi:10.1038/ngeo618.

Buffam, I., M. G. Turner, A. R. Desai, P. C. Hanson, J. A. Rusak, N. R. Lottig, E. H. Stanley, and S. R. Carpenter (2011), Integrating aquatic and terrestrial components to construct a complete carbon budget for a north temperate lake district, *Global Change Biol.*, *17*, 1193–1211, doi:10.1111/j.1365-2486.2010.02313.x.

Cole, J. J., et al. (2007), Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon budget, *Ecosystems*, *10*, 172–185, doi:10.1007/s10021-006-9013-8.

Downing, J. A. (2009), Global limnology: Up-scaling aquatic services and processes to planet, *Earth. Verh. Int. Ver. Limnol.*, *30*, 1149–1166.

Downing, J. A. (2010), Emerging global role of small lakes and ponds: Little things mean a lot, *Limnetica*, *29*, 9–24.

Downing, J. A., et al. (2006), The global abundance and size distribution of lakes, ponds, and impoundments, *Limnol. Oceanogr.*, *51*, 2388–2397, doi:10.4319/lo.2006.51.5.2388.

IPCC (2013), *Climate Change 2013: The Physical Science Basis*, Cambridge Univ. Press, Cambridge, U. K.

Kuss, J., and B. Schneider (2004), Chemical enhancement of the CO_2 gas exchange at a smooth seawater surface, *Mar. Chem.*, *91*, 165–174, doi:10.1016/j.marchem.2004.06.007.

Lewis, W. M. (2011), Global primary production of lakes: 19th Baldi Memorial Lecture, *Inland Waters*, *1*, 1–28, doi:10.5268/IW-1.1.384.

Lim, W. H., and M. L. Roderick (2014), Up-scaling short-term process-level understanding to longer timescales using a covariance-based approach, *Hydrol. Earth Syst. Sci.*, *18*, 31–45, doi:10.5194/hess-18-31-2014.

Marotta, H., C. M. Duarte, S. Sobek, and A. Enrich-Prast (2009), Large CO_2 disequilibria in tropical lakes, *Global Biogeochem. Cy.*, *23*, GB4022, doi:10.1029/2008GB003434.

Morales-Pineda, M., A. Cózar, I. Laiz, B. Úbeda, and J. Á. Gálvez (2014), Daily, biweekly, and seasonal temporal scales of pCO2 variability in two stratified Mediterranean reservoirs, *J. Geophys. Res. Biogeosci.*, *119*, 509–520, doi:10.1002/2013JG002317.

Nisell, J., A. Lindsjö, and J. Temnerud (2007), Rikstäckande virtuellt vattendrags nätverk för flödesbaserad modellering VIVAN, [In Swedish], Rapport 2007:17, Institutionen för miljöanalys, SLU.

Rastetter, E. B., A. W. King, B. J. Cosby, G. M. Hornberger, R. V. O'Neil, and J. E. Hobbie (1992), Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems, *Ecol. Appl.*, *2*, 55–70, doi:10.2307/1941889.

Raymond, P. A., et al. (2013), Global carbon dioxide emissions from inland waters, *Nature*, *502*, 355–359, doi:10.1038/nature12760.

Read, J. S., et al. (2012), Lake-size dependency of wind shear and convection as controls on gas exchange, *Geophys. Res. Lett.*, *39*, L09405, doi:10.1029/2012GL051886.

Schilder, J., D. Bastviken, M. van Hardenbroek, P. Kankaala, P. Rinta, T. Stötter, and O. Heiri (2013), Spatial heterogeneity and lake morphology affect diffusive greenhouse gas emission estimates of lakes, *Geophys. Res. Lett.*, *40*, 5752–5756, doi:10.1002/2013GL057669.

Seekell, D. A., and M. L. Pace (2011), Does the Pareto distribution adequately describe the size-distribution of lakes?, *Limnol. Oceanogr.*, *56*, 350–356, doi:10.4319/lo.2011.56.1.0350.

Seekell, D. A., M. L. Pace, L. J. Tranvik, and C. Verpoorter (2013), A fractal-based approach to lake size-distributions, *Geophys. Res. Lett.*, *40*, 517–521, doi:10.1002/grl.50139.

Sprugel, D. G. (1983), Correcting for bias in log-transformed allometric equations, *Ecology*, *64*, 209–210.

Stumpf, M. P. H., and M. A. Porter (2012), Critical truths about power laws, *Science*, *335*, 665–666, doi:10.1126/science.1216142.

Tranvik, L. J., et al. (2009), Lakes and reservoirs as regulators of carbon cycling and climate, *Limnol. Oceanogr.*, *54*, 2298–2314, doi:10.4319/lo.2009.54.6_part_2.2298.

Vachon, D., and Y. T. Prairie (2013), The ecosystem size and shape dependence of gas transfer velocity versus wind speed relationships in lakes, *Can. J. Fish. Aquat. Sci.*, *70*, 1757–1764, doi:10.1139/cjfas-2013-0241.

Verpoorter, C., T. Kutser, and L. Tranvik (2012), Automated mapping of water bodies using Landsat multispectral data, *Limnol. Oceanogr. Methods*, *10*, 1037–1059, doi:10.4319/lom.2012.10.1037.

Verpoorter, C., T. Kutser, D. A. Seekell, and L. J. Tranvik (2014), A global inventory of lakes based on high-resolution satellite imagery, *Geophys. Res. Lett.*, *41*, 6396–6402, doi:10.1002/2014GL060641.

Wagner, T., P. A. Soranno, K. S. Cheruvelil, W. H. Renwick, K. E. Webster, P. Vaux, and R. J. Abbitt (2008), Quantifying sample biases of inland lake sampling programs in relation to lake surface area and land use/cover, *Environ. Monit. Assess.*, *141*, 131–147, doi:10.1007/s10661-007-9883-z.

Weyhenmeyer, G. A., P. Kortelainen, S. Sobek, R. Müller, and M. Rantakari (2012), Carbon dioxide in boreal surface waters: A comparison of lakes and streams, *Ecosystems*, *15*, 1295–1307, doi:10.1007/s10021-012-9585-4.