Project Description

Background: Lakes and reservoirs are hot spots for biogeochemical processing, but are almost never considered in global scale analyses of biogeochemical cycles (Cole et al. 2007, Downing 2009). Recently, limnologists have become interested in developing more accurate estimates of the global distribution of inland aquatic systems (Downing et al. 2006, Seekell and Pace 2011). This interest is driven partly by the need to assess the importance of inland waters in processes such as the global carbon cycle, partly by the need to more fully assess the numbers and sizes of aquatic systems given global human pressures on inland waters, and partly by a desire to make limnology relevant to the public during a period of intense focus on global scale environmental issues (Downing 2009, cf. Jumars 1990). This work has led to new findings, including that inland lakes and reservoirs cover a much greater portion of the earth's land surface (~3%) and that inland waters process substantial amounts of organic carbon, relative to previous estimates (Downing et al. 2006, Tranvik et al. 2009, Butman and Raymond 2011). At the global scale, Tranvik et al. (2009) estimated that land exports of carbon to inland waters are twice as high as land exports of carbon to the ocean. Most of this carbon is subsequently exported to oceans (0.9 Pg yr^{-1}), is buried (0.6 Pg yr^{-1}), or is oxidized and evades into the atmosphere as carbon dioxide (at least 1.4 Pg yr^{-1}) (Tranvik et al 2009). Lake sediments may contain as much as 820 Pg of carbon (Cole et al. 2007). Globally, lakes are important sources of methane, with greater emissions to the atmosphere than the world's oceans (Bastviken et al. 2004). Conservative estimates of carbon dioxide and methane emissions from lakes suggest that lakes offset at least 79% of the terrestrial green house gas sink (Bastviken et al. 2011). Collectively, these results represent a new sub-discipline known as global limnology.

Despite an abundance of new empirical findings, detailed above, there has been limited conceptual development in global limnology. Consequently, there is a critical lack of theory from which to derive testable hypotheses and to guide new studies. There is also a lack of new methodologies suitable for global scale limnological analyses (Downing 2009). This is partly because limnology has traditionally focused on studying one or a few lakes intensively, and partly because the lack of theory that might foster new methods (Downing et al. 2009). There is a significant opportunity to impact global limnology through the development of new theories and methodologies. The purpose of the proposed work is to develop theoretical and methodological advancements for global limnology.

Research plan: This project will build upon the Graduate Fellow Applicant's (GFA) previous research on regional and global lake size-distributions (see Seekell and Pace 2011). This work challenged previous findings that the size distribution of lakes follows a power-law distribution and that small lakes dominate the total surface area of lakes. The proposed project will comprise two specific components related to lake-size distributions. The first component will focus on theoretical lake size-distributions and their variability with regional hypsography (surface area-elevation relationships). This component will use equations and simulations to explain why some hypsographic regions have power-law lake size-distributions and why some do not (Downing et al. 2006, Seekell et al. 2011). The second project component will evaluate variability in empirical biogeochemical scaling relationships due to regional hypsography with the purpose of improving accuracy in global up-scaling of the role of lakes in the global biogeochemical cycles, specifically the global carbon cycle. Advantages of the proposed collaborations with Nordic partners are presented after briefly outlining the two projects.

1) Derivation of a theoretical lake size-distribution.

Background and purpose.—The size-distribution of lakes is a fundamental but unresolved concern of limnology (Mandelbrot 1983, Tranvik et al. 2009, Seekell and Pace 2011). A size-distribution is an equation that describes the abundance of lakes of any given surface area (Downing et al. 2006, Seekell and Pace 2011). If the size distribution of lakes is known *a priori*, the total abundance of lakes in a region can be estimated simply by enumerating the surface area of the largest lakes on maps and extrapolating based on the size-distribution equation (Downing et al. 2006, Downing 2009). This procedure is important because small lakes have high rates of biogeochemical processing and storage, but these small lakes are generally not enumerated on maps (Downing et al. 2006). Hence, an accurate description of the size-distribution of lakes is critical for quantifying the role of lakes in global biogeochemical cycles (Tranvik et al. 2009).

A number of studies have used the Pareto (power-law) distribution to estimate the global or regional abundances and surface areas of lakes (e.g. Downing et al. 2006, Tranvik et al. 2009). This approach fits a log-abundance log-size regression based on the largest lakes recorded on maps. Abundance is the number of lakes greater than or equal to a given size (Downing et al 2006). Parameters from the regression are then used to estimate the number of small, unobserved lakes using the Pareto distribution probability density function (Downing et al. 2006). However, the GFA has previously shown that small lakes often depart from the power-law in highly resolved datasets. This is likely because many size-distributions mimic the Pareto distribution when small values are excluded, such as small lakes are on most maps. This means that good fit of large lakes to the Pareto distribution does not necessarily mean that all lakes fit a Pareto distribution (Seekell and Pace 2011). For instance, Figure 1 contains data simulated from a lognormal distribution where the number of small lakes in the lower tail of the distribution is variably truncated. The solid circles have 0% truncation, the open circles have 90% truncation, the solid squares have 99% truncation, and the open squares have 99.9% truncation (see Seekell and Pace 2011 for details). As degree of truncation increases the distribution appears more like a straight line, which is characteristic of a Pareto distribution. Hence, current empirical methods leave the important issue of lake-size distributions uncertain. Are the distributions more like the Pareto (power law) pattern or more like the lognormal pattern represented by solid circles in Figure 1? The purpose of this project component is to derive a theoretical lake size distribution to improve the accuracy of estimates of the role of lakes in regional and global biogeochemical cycles.

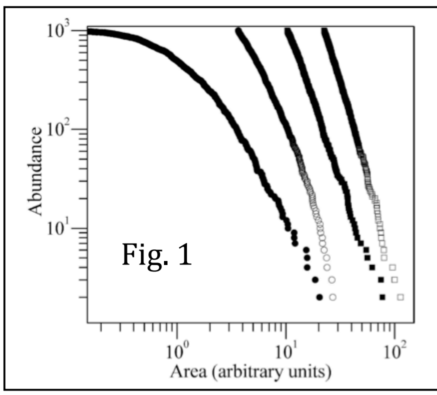
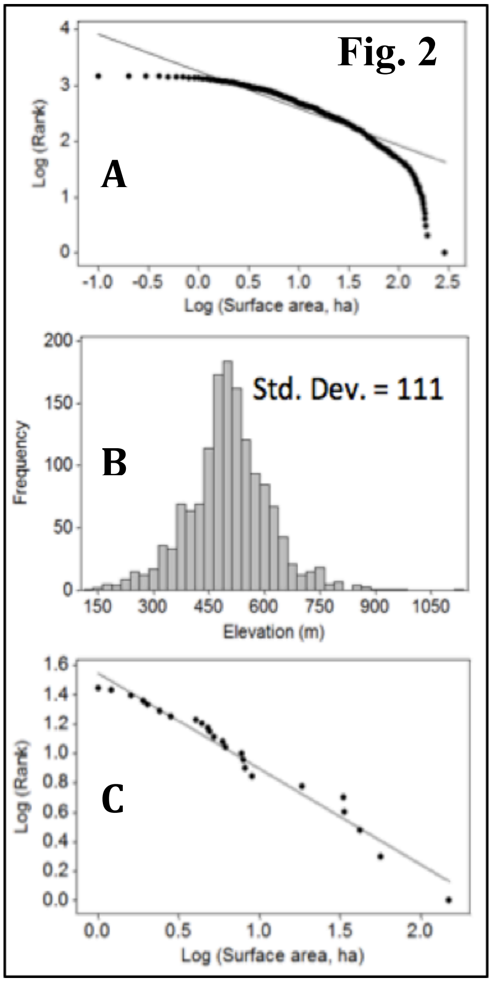


Fig. 1

Methods and preliminary results.—Lakes are fractals (Mandelbrot 1983, Kent and Wong 1982, Russ 1994). Theoretically their surface area can be known exactly but their shore length cannot be measured (Russ 1994). Based on fractal geometric theory, if individual lakes are fractals, the size-distribution of lakes will also be fractal (cf. Matshushita et al. 1991, Russ 1994). One property of fractal size-distributions is that they follow the Pareto distribution (Matshushita et al. 1991, Russ 1994). This property has been cited as justification for a Pareto distribution of lake sizes (e.g. Goodchild 1988, Hamilton et al. 1992). *The caveat in this theory is that the Pareto distribution only holds for lakes at the mean elevation* (cf. Matshushita 1991, Russ 1994). When lakes are not at the mean elevation, the size distribution is curved and the Pareto distribution overestimates the number of small lakes (cf. Ding and Yang 1995). Hence, I hypothesize the shape of the size-distribution of lakes should depend on the vertical relief of landscapes. Regions where lakes all occur near the mean elevation should follow Pareto distribution while mountainous regions where may lakes occur away from the mean elevation should depart from the Pareto distribution. A preliminary result supports this idea (Fig. 2). The size-distribution of lakes in the Adirondack Mountains of New York (data from Seekell and Pace 2011) deviates significantly from a power-law distribution, represented by a straight line (Fig. 2A). These lakes occur over a wide range of elevations (Fig. 2B). Interestingly, at the mean elevation these lakes do not deviate significantly from a power-law (Fig. 2C). Further the distribution in Fig. 2C has a fractal dimension that is almost exactly the expected value based upon theory and an independent estimate of the fractal dimension of this landscape by Turcotte (1995).



This project will formally derive the theoretical lake size-distribution within the framework of fractal geometry. Simulations of fractal landscapes and highly resolved lake size datasets (e.g., Figure 2) will be used to evaluate the derivations. The GFA will derive testable hypotheses based on these size-distributions about the dominant processes in different hypsographic regions and at different elevations within regions. Specifically, 1) flat regions should have a Pareto lake size-distributions while mountainous regions should not have Pareto distributions, 2) small lakes cannot theoretically dominate the global lake surface area and empirical results suggesting this are due to sampling artifacts, 3) The scaling relationships of carbon processing vary regionally with lake size distribution. This third hypothesis will be evaluated partly based on the results of this project component and partly based on an analysis of Swedish national lake survey data (see below).

2) Regional variability in lake biogeochemistry

Background and purpose.—The biogeochemical characteristics of lakes generally covary with ecosystem characteristics, particularly lake surface area (Downing 2009). Hence, simply multiplying mean areal rates or quantities by the number of lakes in the world may not be an adequate approach for up-scaling the role of lakes in the global carbon cycle (Downing 2009). Generally, global level predictive relationships are used to relate biogeochemical properties to lake surface area (e.g., Sobek et al. 2007). However, these relationships might vary geographically (cf. Weyhenmeyer and Karlsson 2009). This is likely due to a variety of factors including climate, land use history, and underlying geology. The purpose of this project component is to use novel statistical methods to identify regions with lakes characterized by similar biogeochemical scaling properties.

Methods and preliminary results.—Geographically weighted regression (GWR) analysis will be applied to Swedish national lake survey data (n > 2000 lakes) to create empirical models of biogeochemical scaling relationships (Brunsdon et al. 1998). GWR is a formal statistical method that tests for and models structured geographic variation in regression coefficients (Brunsdon et al. 1998). The GWR procedure results in a map of 'local' regression coefficients that describes scaling relationships for each lake. Each local regression coefficient is calibrated using a moving window and distance decay weighting function such that nearby lakes have more weight in calibrating the local regression coefficient than lakes further away (cf. Brunsdon et al. 1998). Cluster analysis is subsequently applied to the local scaling relationships to identify regions with similar scaling relationships between lake area and biogeochemical parameters (cf. Wimberly et al. 2008). This approach is distinct from the cluster analyses generally employed in limnology in that lakes are not being clustered based on the rate and quantities of biogeochemical parameters (e.g. Shannon and Brezonik 1972), but rather based on the similarity of scaling relationships. The delineated scaling regions will be tested for correspondence to hypsographic regions. Correspondence between these regions would allow improved accuracy in up-scaling because more accurate lake-size distributions and scaling relationships could be applied to each region, and regional estimates of green house gas emissions and carbon storage can subsequently be summed into continental scale estimates.

Sweden is an ideal country to apply these methods because of its long north-south climate gradient, its east-west altitudinal gradient, and its abundance of existing national lake survey data (cf. Göransson et al. 2004, Göransson et al. 2006). Initial analyses have found statistically significant geographic variability in relationships in Swedish lake survey data and that GWR models considerably improve the amount of explained variance over statistical models that do not account for geographically variable relationships. The methodology will also be applied to national lake survey data from Finland. Dr. Pirkko Kortelainen of the Finnish Environment Institute has made these existing data available for study.

Nordic partnership objectives and rational for proposed collaboration: The proposed Nordic Host, Dr. Lars J. Tranvik at Uppsala University, is an international leader in the study of the role of inland waters in the global carbon cycle. Tranvik has authored or co-authored most of major papers on global limnology (e.g. Downing et al. 2006, Cole et al. 2007) including major empirical analyses published in *Science* and *Nature Geoscience* (see Bastviken et al. 2011 and Battin et al. 2009, respectively). Tranvik has also authored an important review paper on global limnology (Tranvik et al. 2009).

The GFA and Nordic Host previously participated in a special workshop in San Juan, Puerto Rico (February 2011) on *The Role of Aquatic Networks in the Boreal Carbon Cycle* sponsored by the American Society of Limnology and Oceanography and the International Federation of Boreal Aquatic Research. A number of Swedish limnologists engaged in global limnological research attended the workshop. Tranvik actively collaborates with many of these limnologists and the Nordic Research Opportunity will provide the GFA an avenue for continued contact with these researchers. Specifically, the GFA has been invited to visit Umeå University's Abisko field station by Erik Lundin, a doctoral student working on catchment scale aquatic carbon cycling in the boreal region. Collaboration with Tranvik represents the most advantageous possible partnership both within Nordic countries, and globally because of his expertise in aquatic carbon cycling and global limnology and connection to his network of global limnologist collaborators.

The principal product of the proposed Nordic collaboration for this project is the development of testable hypotheses based on theoretical lake size-distribution derivations. The GFA is well versed in the theoretical aspects (e.g. Seekell and Pace 2011) of this project and the proposed host is an expert in aquatic carbon cycling (e.g. Tranvik et al. 2009). Collaboration will likely lead to additional insights and more refined hypotheses. Collaboration with Tranvik will allow the GFA access to extensive datasets from Swedish lake surveys that may be used to evaluate the newly developed hypotheses. These datasets will also serve as the basis for the second project component. A post-doctoral fellow in Tranvik's lab is currently using remote sensing techniques to gather data on the size-distribution of lakes in Sweden. Collaboration with Tranvik and his post-doc could lead to the development of data suitable for evaluating the theoretical derivations detailed in the proposed methods.

The proposed project, which utilizes existing data sources, will be completed in a sixteen-week period (September 1–December 22). This duration is a sufficient duration to also participate in seminars, discussion groups, and networking opportunities within Sweden.

Intellectual merit: The GFA's research seeks to better understand the role of lakes in global biogeochemical cycles. This important for understanding the global carbon cycle because inland waters may emit greenhouse gasses sufficient to largely offset the terrestrial carbon sinks (Richey et al. 2002, Bastviken et al. 2011). Further, inland waters may emit more methane and store more carbon in sediments than the world's oceans (Cole et al. 2007). Hence improved understanding the role of lakes, particularly in the global carbon cycle, stands to improve overall understanding (Cole et al. 2007, Downing 2009).

The first component of this project addresses one of the most fundamental aspects of limnology. Derivation of a lake size-distribution would be a major theoretical advancement. Lake chemistry is known to have regional dynamics (e.g. Pace and Cole 2002), but little is known about what contributes to the delineation of these regions. If these scaling regions delineated in the second component of this study correspond with the hypsographic regions of lake size-distributions, regional models could greatly increase the accuracy of up-scaling procedures. This study will provide new insights into the geographic sales of variability of aquatic biogeochemical processes while also informing sampling for future up-scaling studies.

Broader impacts: Theoretical and empirical results from this collaboration will be presented at the American Society of Limnology and Oceanography's 2013 Aquatic Sciences Meeting February 17–22 in New Orleans, Louisiana. This is the foremost conference in the field and is held every two years. The GFA will also present these results at the Fall 2013 meeting of the American Geophysical Union in San Francisco, California. As a result of this work at least two manuscripts should be co-authored by the GFA, the GFA's academic advisor, the Nordic host, and other possible collaborators. The theoretical derivations and hypotheses developed in the first component of the collaboration will be detailed in a manuscript appropriate for submission to *Limnology & Oceanography* (the flagship journal in aquatic sciences) or a similar journal. The methodological developments from the second component of the collaboration will be detailed in a manuscript for submission suitable for submission to a journal such as *Limnology & Oceanography* or *Journal of Geophysical Research-Biogeosciences*. Within six months after completion of the visit to Sweden, the GFA will submit a report to the Graduate Research Fellowship Director reviewing the experience and accomplishments, in addition to any reporting requirements from the Swedish Research Council.

The GFA will receive mentoring on limnological research methods from the Nordic host and attend lectures and seminars in the host's academic department. This participation will improve GFA's understanding of fundamental limnology, particularly aquatic biogeochemistry. The GFA's home department provides only minimal opportunity for interaction with other limnologists (the GFA's advisor is the only limnologist and only one introductory limnology course is offered by the University). The international collaboration will provide the GFA with an opportunity to develop a network from which to seek future research opportunities including possible postdoctoral positions.

References:

Bastviken D, Cole JJ, Pace ML, Tranvik LJ (2004) Methane emissions from lakes: Dependence of lake characteristics, two regional assessments, and a global estimate. *Global Biogeochemical Cycles* 18:GB4009.

Bastviken D, Tranvik LJ, Downing DA, Crill PM, Enrich-Prast A (2011) Freshwater methane emissions offset the continental carbon sink. *Science* 331:50.

Battin TJ, Luyssaert S, Kaplan LA, Aufdenkampe AK, Richter A, Tranvik LJ (2009) The boundless carbon cycle. *Nature Geoscience* 2:598–600.

Brunsdon C, Fotheringham S, Charlton M (1998) Geographically weighted regression-modelling spatial non-stationarity. *Journal of the Royal Statistical Society D* 47:431–443.

Butman D, Raymond PA (2011) Significant efflux of carbon dioxide from streams and rivers in the United States. *Nature Geoscience* 4:839–842.

Cole JJ, Prairie YT, Caraco NF, McDowell WH, Tranvik LJ, Striegl RG, Duarte CM, Kortelainen P, Downing JA, Middelburg J, Melack JM (2007) Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon cycles. *Ecosystems* 10:172–185.

Ding M, Yang W (1995) Distribution of the first return time in fractional Brownian motion and its application to the study of on-off intermittency. *Physical Review E* 52:207–213.

Downing JA, Prairie YT, Cole JJ, Duarte CM, Tranvik LJ, Striegl RG, McDowell WH, Kortelainen P, Caraco NF, Melack JM, Middelburg J (2006) The global abundance and size distribution of lakes, ponds, and impoundments. *Limnology & Oceanography.* 51:2388–2397.

Downing JA (2009) Global limnology: Up-scaling aquatic services and processes to planet Earth. *Verhandlungen Internationale Vereinigung für theoretische und angewandte Limnologie* 30:1149–1166.

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

Goodchild MF (1988) Lakes on fractal surfaces: a null hypothesis for lake-rich landscapes. *Mathematical Geology* 20:615–630.

Göransson E, Johnson RK, Wilander A (2004) Representativity of a mid-lake surface water chemistry sample. *Environmental Monitoring and Assessment* 95:221–238.

Göransson E, Bringmark E, Rapp L, Wilander A (2006) Modeling the effect of liming on calcium concentration in Swedish lakes. *Environmental Monitoring and Assessment* 119:331–348.

Jumars PA (1990) W(h)ither limnology? *Limnology & Oceanography* 35:1216–1218.

Hamilton SK, Melack JM, Goodchild MF, Lewis WM (1991) p. 145–163 in Carling PA, Petts GE (eds.) *Lowland Floodplain Rivers: Geomorphological Perspectives*. Wiley.

Kent C, Wong J (1982) An index of littoral zone complexity and its measurement. *Canadian Journal of Fisheries and Aquatic Sciences* 39:847–853.

Mandelbrots BB (1983) *The Fractal Geometry of Nature*. Freeman.

Matsushita M, Ouchi S, Honda K (1991) On the fractal structure and statistics of contour lines on a self-affine surface. *Journal of the Physical Society of Japan* 60:2109–2112.

Pace ML, Cole JJ (2002) Synchronous variation of dissolved organic carbon and color in lakes. *Limnology & Oceanography* 47:333–342.

Richey JE, Melack JM, Aufdenkampe AK, Ballester VM, Hess LL (2002) Outgassing from Amazonian rivers and wetlands as a large tropical source of atmospheric CO2. *Nature* 416:617–620.
Russ JC (1994) *Fractal Surfaces*. Plenum.
Seekell DA, Pace ML (2011) Does the Pareto distribution adequately describe the size-distribution of lakes? *Limnology & Oceanography.* 56:350–356.
Shannon EE, Brezonik PL (1972) Limnological characteristics of north and central Florida lakes. *Limnology & Oceanography* 17:97–110.
Sobek S, Tranvik LJ, Prairie YT, Kortelainen P, Cole JJ (2007) Patterns and regulation of dissolved organic carbon: an analysis of 7,500 widely distributed lakes. *Limnology & Oceanography* 52:1208–1219.
Tranvik LJ, Downing JA, Cotner JB, Loiselle SA, Striegl RG, Ballatore TJ, Dillon P, Finlay K, Fortino K, Knoll LB, Kortelainen PL, Kutser T, Larsen S, Laurion I, Leech DM, McCallister SL, McKnight DM, Melack JM, Overholt E, Porter JA, Prairie Y, Renwick WH, Roland F, Sherman BS, Schindler DW, Sobek S, Tremblay A, Vanni MJ, Verschoor AM, von Wachenfeldt E, Weyhenmeyer GA (2009) Lakes and reservoirs as regulators of carbon cycling and climate. *Limnology & Oceanography* 54:2298–2314.
Turcotte DL (1995) Scaling in geology: landforms and earthquakes. *Proceedings of the National Academy of Science of the United States of America* 92:6697–6704.
Weyhenmeyer GA, Karlsson J (2009) Nonlinear response of dissolved organic carbon concentration in boreal lakes to increasing temperatures. *Limnology & Oceanography* 54:2513–2519.
Wimberly MC, Yabsley MJ, Baer AD, Dugan VG, Davidson WR (2008) Spatial heterogeneity of climate and land-cover constraints on distributions of tick-borne pathogens. *Global Ecology and Biogeography* 17:189–202.