Eutrophication drives extreme seasonal CO_2 flux in lake ecosystems

Running head: Eutrophication shapes lake CO_2 flux

Ana M. Morales-Williams[1,2], Alan D. Wanamaker, Jr.[3], Clayton J. Williams[2], and John A. Downing[1,4]

[1]Iowa State University, Department of Ecology, Evolution, and Organismal Biology, 251 Bessey Hall, Ames, IA, 50014, USA

[2]University of Vermont, Rubenstein School of Environment and Natural Resources, Aiken Center, 81 Carrigan Dr., Burlington, VT, 05405

[3]Department of Geological and Atmospheric Science, Iowa State University, 12 Science 1, Ames, IA, 50011, USA

[4]Minnesota Sea Grant, University of Minnesota-Duluth, 31 West College St., Duluth, MN, 55812, USA

Corresponding Author: A.M. Morales-Williams, ana.morales@uvm.edu

Author contributions: JAD and AMM conceived and designed the study. AMM performed research, analyzed data, and wrote the paper. ADW contributed methods and provided feedback on manuscript drafts. CJW analyzed data and provided feedback on manuscript drafts.

Abstract

Lakes process a disproportionately large fraction of carbon relative to their size and spatial extent, representing an important component of the global carbon cycle. Alterations of ecosystem function via eutrophication change the balance of greenhouse gas (GHG) flux in these systems. Without eutrophication, lakes are net sources of CO_2 to the atmosphere, but in eutrophic lakes this function may be amplified or reversed due to cycling of abundant autochthonous carbon. Using a combination of high-frequency and discrete sensor measurements, we calculated continuous CO_2 flux during the ice-free season in 15 eutrophic lakes. We found net CO_2 influx over our sampling period in 5 lakes (-47 to -1865 mmol m^{-2}) and net efflux in 10 lakes (328 to 11,755 mmol m^{-2}). Across sites, predictive models indicated that the highest efflux rates were driven by nitrogen enrichment, and influx was best predicted by chlorophyll *a* concentration. Regardless of whether CO_2 flux was positive or negative, stable isotope analyses indicated that the dissolved inorganic carbon (DIC) pool was not derived from heterotrophic degradation of terrestrial organic carbon, but from degradation of autochthonous organic carbon, mineral dissolution, and atmospheric uptake. Optical characterization of dissolved organic matter (DOM) revealed an autochthonous organic matter pool. CO_2 influx was correlated with autochthony, while efflux was correlated with total nitrogen and watershed wetland cover. Our findings suggest that CO_2 uptake by primary producers during blooms can contribute to continuous CO_2 influx for days to months. Conversely, eutrophic lakes in our study that were net sources of CO_2 to the atmosphere showed among the highest rates reported in literature. These findings suggest that anthropogenic eutrophication has substantially altered biogeochemical processing of carbon on Earth.

Keywords: CO_2 flux, eutrophication, DOM, nitrogen

Manuscript highlights

- Five of 15 eutrophic lakes in this study were net CO_2 sinks, and influx was driven by indicators of autochthony, including chlorophyll a concentration and autochthonous dissolved organic matter.
- Nitrogen concentration and percent watershed wetland cover best predicted CO_2 efflux.
- Lakes that were net CO_2 sources reported here have substantially higher efflux rates than oligotrophic or mesotrophic lakes previously reported in the literature.

Introduction

Anthropogenic eutrophication is changing the role of lakes in the global carbon cycle. Intensification of industrial agriculture has resulted in massive increases in fertilizer use and the extent of irrigated cropland (Foley et al. 2005). Extensive cultivation alters watershed horizontal permeability, and thus the rate, timing, concentration, and quality of inorganic nutrients and dissolved organic matter (DOM) exported to downstream aquatic ecosystems (Foley et al. 2005; Petrone et al. 2011; Williams et al. 2015). Collectively, these processes contribute to degradation of water quality, hypoxia, and harmful cyanobacteria blooms (Heisler et al. 2008; Brooks et al. 2016). In the absence of eutrophication, inputs of terrestrial DOM to lakes fuel heterotrophic respiration in excess of primary production (Pace and Prairie 2005; Duarte and Prairie 2005). Coupled with watershed inputs of inorganic carbon, this often results in positive net CO_2 efflux from surface waters (Marcé *et al.*, 2015; Weyhenmeyer *et al.*, 2015; Wilkinson *et al.*, 2016). Because a disproportionate amount of lake carbon research has been conducted in northern temperate forested lakes (Sobek et al. 2005; Balmer and Downing 2011) relative to eutrophic, agriculturally impacted systems, the generalization is sometimes made that lakes function as sources of CO_2 to the atmosphere, that these rates are moderate (i.e., <50 mmol m^{-1}

day^{-1}), and that daytime influx is balanced or exceeded by diel respiratory flux (Kosten et al. 2010; López et al. 2011). This may not be true, however, of anthropogenically impacted aquatic ecosystems.

Anthropogenically eutrophic freshwater ecosystems differ from less impacted lakes in watershed cultivation and development (Heathcote and Downing 2011), nutrient concentrations, primary productivity (Heisler et al. 2008; Pacheco et al. 2014), and DOM quality (Williams et al. 2015). These differences substantially alter how lakes process, store, and export carbon (Heathcote and Downing 2011; Pacheco et al. 2014; Nõges et al. 2016; Wilkinson et al. 2016). Lakes with agricultural and urban catchments have higher microbial processing rates of organic matter than those with forested watersheds, and a greater contribution of microbial-derived, protein-like compounds (Williams et al. 2010, 2015; Petrone et al. 2011) which tend to persist longer than DOM derived from higher plants (Kellerman et al. 2015). Coupled with elevated nutrient concentrations, this can correspond with inorganic C uptake by primary producers exceeding that produced via heterotrophic respiration resulting in sustained depletion of water column CO_2 (Morales-Williams et al. 2017). In the absence of large inputs of humic, terrestrial DOM of higher plant origin, it is unclear if exogenous CO_2 inputs and mineral dissolution can support net CO_2 efflux from surface waters when primary production is very high, as is expected in eutrophic and hypereutrophic lake ecosystems.

Dissolved inorganic carbon (DIC) in lake surface waters is primarily derived from mineral dissolution and the bicarbonate buffering system, but is further mediated by a diversity of sources including equilibration with the atmosphere, heterotrophic respiration, and watershed inputs (Bade et al. 2004). The balance between CO_2 produced via these mechanisms and that fixed by primary production affects the potential flux of CO_2 between the lake surface and the

atmosphere, though net flux is ultimately controlled by turbulence at the air-water interface (Kling *et al.*, 1992; Del Giorgio *et al.*, 2009) and physical mixing events (i.e., gas release at fall turnover). Thus, while high rates of primary production fix large quantities of inorganic carbon during bloom events, the combined effects of processes that facilitate net efflux (heterotrophy, mineral dissolution, physical mixing) may prevent eutrophic and hypereutrophic lakes from acting as net CO_2 sinks. Alternately, if inorganic carbon contributions from watershed sources and heterotrophy are small relative to autochthonous carbon from primary production, eutrophic lakes would be expected to maintain continuous negative flux (CO_2 flux into the lake) during periods of stable stratification.

The purpose of this study was to investigate the variability in magnitude and duration of CO_2 flux in eutrophic and hypereutrophic lake ecosystems, and to assess the relative influence of biological and physical parameters on CO_2 flux direction and rate. We assessed the source of inorganic carbon pools and quality of dissolved organic matter across 15 eutrophic lakes using stable isotopic and optical methods. Using high frequency pH and temperature measurements, we calculated continuous CO_2 flux over one ice free season in these systems, and partitioned variability in flux attributable to endogenous (i.e., primary production and autochthonous organic matter) or exogenous (watershed inputs) sources. We hypothesized that periods of net CO_2 influx would be correlated with variables associated with endogenous biological mechanisms and that net efflux would correlate with physical mixing and DIC sourced from mineral dissolution rather than the degradation of terrestrial organic matter.

Materials and methods

Site selection and sampling design

Fifteen eutrophic lakes were chosen along an orthogonal gradient of watershed cultivation and interannual variability in Cyanobacteria dominance (Table 1, S1). These sites were selected based on long-term survey data from 132 lakes monitored by the Iowa State Limnology Laboratory between 2000 and 2010 (Ambient Lake Monitoring Program: https://programs.iowadnr.gov/aquia/Programs/Lakes). All lakes in this study are relatively shallow (<7 m max depth), and 13 of 15 are man-made. They are all algal-dominated systems and do not have productive macrophyte communities. Eight lakes are classified as dimictic (Arrowhead, Badger, Beeds, East Osceola, George Wyth, Keomah, Silver -Dickinson, and Springbrook); seven are polymictic (Blackhawk, Center, Five Island, Orient, Lower Gar, Silver-Palo Alto, and Rock Creek), though Silver-Dickinson did not stratify during our sampling season (Figure 1). Lakes were sampled for standard biological, chemical, and physical parameters during the ice-free season of 2012 once at ice out, twice per week in May and June, once per week in July and August, and once per month until the onset of ice cover. Samples for DOM characterization and stable isotope analysis of dissolved inorganic carbon ($\delta^{13}C_{DIC}$) were collected once in April, at every second sampling event in May, June, July, and August, and at every sampling event in September and November.

Water quality measurement and analysis

Lakes were sampled at the historic deep point (Table S1), which is the deepest point in each lake based on historical bathymetry. These sites have been sampled regularly by state monitoring programs for >15 years, so have been used here for consistency. If a thermocline was present (based on visual inspection of plotted thermal profile data at the time of sampling, YSI multiparameter sonde), integrated upper mixed zone water column samples were collected above the thermocline to a maximum 2 m depth using a vertical column sampler. If no thermocline was

present, 2 m integrated column samples were collected unless the lake was < 2 m deep, in which case samples were collected 0.5 m above maximum depth so as not to disturb the sediment. Samples were stored in coolers on ice until delivery to the laboratory within 24 hours of collection, then kept at 4°C and processed to a stable state or analyzed fully within 36 hours of collection. Total phosphorus (TP), dissolved organic carbon (DOC), and alkalinity (as mg L^{-1} $CaCO_3$) were analyzed using standard APHA methods (2012). Chlorophyll *a* samples were filtered onto GF/C filters, frozen, then sonicated and extracted in cold acetone under red light and analyzed fluorometrically (Arar and Collins 1997; Jeffrey et al. 1997). Total nitrogen (TN) and nitrate (NO_3^{2-}) were analyzed using the second derivative method (Crumpton et al. 1989). TN was analyzed as NO_3^{2-} after autoclave digestion with sodium hydroxide and persulfate. Vertical profiles of dissolved oxygen (DO), specific conductivity, temperature, and pH were measured with a YSI multi-parameter sonde.

High frequency pH and temperature sensors were deployed between 1.5 and 2 m depth at the deep point of each lake (TempHion pH/ISE/redox sensor probes; accuracy: ± 0.2 °C; 0.2 pH units; 0.1% mV). Measurements of pH and temperature were recorded every 15 minutes during the ice free season (early April through late November 2012) in order to calculate continuous CO_2 flux. For model calibration, discrete measurements of CO_2 were made at each sampling event using a Vaisala GMT220 atmospheric probe modified for aquatic measurements (Johnson et al. 2009). Using methods described and field tested in Johnson et al. (2009), we fitted the atmospheric sensor with a custom made gas-permeable membrane sealed at the sensor base with plasti-dip. Continuous sensors were calibrated monthly and cleaned weekly at each sampling event to remove any biofouling. Minimal biofouling did not affect sensor precision or rate of drift. The calibration re-set sensor drift and it's linearity in response to pH change. Continuous

sensor reported pH values were comparable and reliable within each lake and precise with measurements made 15 minutes apart. After calibration, continuous sensor reported pH often differed from that measured discretely with the YSI sonde. Here, we consider the YSI pH estimate the true value and continuous sensor reported pH values were linearly adjusted to better match YSI measurements (see data correction below and Plates S1-S15).

High frequency data correction and CO_2 modeling

Continuous, high frequency pH and temperature measurements were averaged by the hour for the full measurement period (Plates S1-S15). Hourly averaged temperature values were used in the pCO_2 calculation without further manipulation. Hourly averaged pH values were corrected for systematic error and drift by adjusting the pH values based on their deviation from the discrete pH measurements. The difference between hourly average and measured pH at each discrete sampling event was then linearly interpolated at an hourly interval between sampling events. The hourly interpolated pH difference was added to the hourly averaged, high frequency pH measurement. Next, for each lake, 90% confidence intervals were calculated for the adjusted pH values and pH values that fell outside the 90% confidence interval were removed. Finally, adjust pH values within the 90% confidence interval for each lake were visually checked by plotting the time series and adjusted pH values were manually removed if they deviated more than 1 pH unit from discrete value or they were noisy (rapid hour by hour bi-directional pH change). The final hourly averaged, adjusted pH values were used to calculate pCO_2. The impact of these adjustments is displayed in the supplemental information (Plates S1-S15). All pH cleaning and adjustment steps took place in R using base packages (R Core Team 2015).

Continuous aqueous pCO_2 was calculated based on carbonate equilibria using corrected hourly, adjusted pH and temperature data (Plates S1-S15), and linearly interpolated discrete

measurements of alkalinity and conductivity (Stumm & Morgan, 1996). Calculated pCO_2 generally overestimate measured pCO_2 in Iowa lakes (Plates S1-S15), which was also found in Wisconsin lakes (Golub et al. 2017). To correct for this over estimation, calculated pCO_2 was modeled with measured pCO_2 for each lake and this linear fit was used to predict pCO_2 concentration. For each model, the slope, intercept, coefficient of variation, root mean squared error (RSME) and relative squared error (RSE) of the mean and median predicted value were calculated (Table S2). RSE of the median varied widely across lakes, ranging from 7% for Silver Lake (Dickinson) to 80% for Springbrook Lake. For all but three lakes, R^2 was > 0.60 and RSE of the median was < 38%. RSME ranged from 76 ppm for Center Lake to 267 ppm for Badger Lake. Random error in the influence of high pH values on pCO_2 was not quantified in this study, but the RSE of each predicted model tended to be higher than the combined systematic and random error of 7.7% reported in Golub et al. (2017). This suggests that random measurement error was likely not the main driver of error between measured and calculated pCO_2. Given the uncertainty around calculated pCO_2 concentrations, linear model predicted pCO_2 concentrations were considered the best approximation of direct pCO_2 measurement (Plates S1-S15) and predicted pCO_2 was used in this study to estimate flux.

Hourly flux was calculated as described in Balmer & Downing (2011) and Wilkinson *et al.* (2016) using the equation

$$F_{CO_{2(t)}} = (CO_{2(t)} - CO_{2(eq)}) * k_H) * k_{CO_{2(t)}} \quad (1)$$

where $CO_{2\,(t)}$ is the concentration of CO_2 in surface water at time t, CO_{2eq} is the average atmospheric equilibrium concentration at time of sampling in 2012 (393 ppm, NOAA Earth System Research Laboratory, http://www.esrl.noaa.gov/), k_H is the Henry's Law constant for CO_2 at time t, and $k_{CO2(t)}$ is the piston velocity. k_H was calculated using the equation

$$e^{-58.0931 + 90.5069 \times \frac{100}{Temp_{h2o(t)}} + 22.294 \times \log\left(\frac{Temp_{h2o(t)}}{100}\right)}$$

where $Temp_{h20}$ is water temperature (°K) at time *t*.

$k_{CO2(t)}$ was calculated using the equation

$$k_{CO_{2(t)}} = \frac{2.07 + 0.215 \times wind(t)^{1.7}}{100} \times \left(\frac{1911 - 118.11 \times Temp_{h2o(t)} + 3.452 \times Temp_{h2o(t)}{}^{2} - 0.04132 \times Temp_{h2o(t)}{}^{3}}{600} \right)^{-0.5}$$

where wind(*t*) is the wind speed (m/s) measured at a height of 10 m at a frequency of 1 to 10 minutes and average at an hourly time interval (*t*) (Wanninkhof 1992; Cole and Caraco 1998; Wilkinson et al. 2016). Wind data was downloaded from the MESONET network (http://mesonet.agron.iastate.edu/request/download.phtml?network=AWOS) for the nearest Iowa Automated Weather Observation System (IA-AWOS) to each lake. Because an atmospheric average was used (393 ppm, NOAA Earth System Research Laboratory, http://www.esrl.noaa.gov/), rather than direct, on-site measurements of atmospheric CO_2, additional uncertainty exists for flux values close to atmospheric equilibrium.

Stable isotope analysis

To characterize the source of the inorganic carbon pool, $\delta^{13}C_{DIC}$ samples were filtered in the field to 0.2 µm and injected into helium gas-flushed septa-capped vials pre-charged with H_3PO_4 to cease biological activity and to sparge CO_2 (Raymond and Bauer 2001; Beirne et al. 2012). Samples were measured via a Finnigan MAT Delta Plus XL mass spectrometer in continuous flow mode connected to a Gas Bench with a CombiPAL autosampler. Reference

standards (NBS-19, NBS-18, and LSVEC) were used for isotopic corrections, and to assign the data to the appropriate isotopic scale. Average analytical uncertainty (analytical uncertainty and average correction factor) was ±0.06 ‰. Samples were analyzed by standard isotope ratio mass spectrometry methods (IRMS), and reported relative to the Vienna Pee Dee Belemnite in ‰ (Equation 1).

$$\delta^{13}C_{Sample} = [(^{13}C/^{12}C)_{sample}/(^{13}C/^{12}C)_{VPDB} - 1] \times 1000 \quad \text{Eq. 1}$$

DOM characterization

To assess the source and quality of DOM and concentration of DOC, water samples were syringe filtered in the field using 0.2 µm pore size polycarbonate membrane filters (Millipore) paired with combusted GF/F pre-filters. A small volume of sample was passed through filters to rinse prior to collecting samples. Samples were then collected in acid-washed and combusted amber glass bottles, and stored on ice until returning to the lab, then stored at 4°C until analysis. Samples were optically characterized by generating absorbance spectra and excitation-emission matrices (EEMs, Horiba Aqualog UV-Vis benchtop fluorometer/spectrophotometer). Absorbance scans (240 to 600 nm, 3 nm interval) and fluorescence EEMs (excitation: 240 to 600 nm, 3 nm interval; emission 213.7 to 620.5 nm, 3.28 nm interval) were run simultaneously at a fixed 5 nm bandpass. A Milli-Q water blank was run daily. Sample EEMs were corrected for inner filter effects and instrument bias and then blank subtracted (Cory et al. 2010; Williams et al. 2010; Murphy et al. 2010). EEMs were standardized to Raman Units using the area under the Raman peak from the daily Milli-Q blank scan.

Optical indices were calculated to evaluate DOM source (fluorescence index, FI), level of degradation (β:α ratio), and humification (humification index, HIX). FI, modified from McKnight *et al.* (2001) was calculated as the ratio of emission at 470 nm to emission at 520 nm

at an excitation of 370 nm, and is an indicator of DOM source material (terrestrial or microbial). The β:α ratio, an indicator of DOM degradation, was calculated using the excitation wavelength at 310 nm as the emission intensity at 380 nm divided by the emission intensity maximum between 420 and 435 nm (Parlanti *et al.* 2000;Wilson & Xenopoulos 2009). HIX was calculated as the ratio of peak area under emissions 434-480 nm and 300-346 nm at 255 nm excitation (Zsolnay et al. 1999), with corrections described in Ohno (2002).

Statistical analysis

We used regression tree and random forest analysis to identify environmental predictor variables for CO_2 flux. This approach was chosen because it is robust to outliers, does not assume data independence or normality, and handles missing data well. First, we generated a regression tree model using the rpart package (Therneau et al. 2017) to predict the magnitude of instantaneous influx or efflux from discrete variables, including chl *a*, $\delta^{13}C_{DIC,}$ FI, β:α, HIX, surface water temperature, wind gust speed, average wind speed, precipitation (daily average), sampling site depth, epilimnetic DO, TP, TN, DOC, thermocline depth, and Schmidt stability from all lakes. Thermocline depth and Schmidt stability were calculated using rLakeAnalyzer (Winslow et al. 2017). The time step of CO_2 flux used in this analysis was matched to that of discrete predictor variables. We used default settings and pruned the tree by setting the maximum tree depth to 4. Second, we generated a random forest model using the R randomForest package (Liaw and Wiener 2002). Using this approach, we generated ensembles of regression trees based on 500 randomized bootstrap samples of training data, where the number of variables tried at each node split were informed by RMSE.

To identify predictors of net CO_2 flux (sum of calculated continuous flux over ice-free sampling season, n=15), we first built a classification tree model using static predictor variables,

including maximum depth (Z_{max}), watershed to lake area ratio (WA:LA), and four land use categories (percent wetland, percent water, percent pasture, and percent row crop agriculture). We limited land use to these four categories to avoid overfitting the model, and used a minimum of 3 observations in a node for a split to be attempted, with a requirement that each split decrease overall lack of fit by a factor of 0.0001. Second, we generated a classification random forest model using the same variables and 500 randomized bootstrap samples of training data limited to two variables tried at each split. Terminal nodes were averaged across all trees to generate the relative importance of each predictor variable.

Results

Water chemistry and meteorology

Water quality data are summarized in Table 1 and Supplementary Table 1. Across the study period, TN ranged from 0.2 mg L^{-1} in Lake Orient in July, to 17.1 mg L^{-1} in Badger Lake in May. With the exception of Arrowhead, Badger, and Springbrook Lakes (7 to 33 µg TP L^{-1}, 7 to 88 µg L^{-1}, and 10 to 89 µg L^{-1}, respectively), all sampling sites had eutrophic or hypereutrophic TP concentrations across sampling events (TP > 25 µg L^{-1}), ranging from 27 µg L^{-1} in Beeds Lake to 885 µg L^{-1} in Lake Orient. The highest DOC concentrations were measured in October in Blackhawk Lake (14.6 mg L^{-1}), and the lowest in April in East Lake Osceola (2.7 mg L^{-1}). 2012 was a severe drought year in the Midwestern U.S., so average rainfall across sites was minimal (4.82 ± 2.15 mm). Average lake depth was 4.0 ± 1.55 m, and thermocline depth was 0.85 ± 1.26 m (Table S1).

CO_2 flux

The magnitude of net CO_2 flux for the ice-free season was negative for 5 lakes and positive for 10 lakes (Table S4). The largest net efflux for the ice-free season was observed in Badger Lake (11,755 mmol m^{-2}), and the largest net influx in Lake Orient (-1865 mmol m^{-2}). Average daily flux across sites and sampling events ranged from -45.4 to 757.3 mmol m^{-2} d^{-1} (Figure 1). Across lakes, the largest efflux events were observed during spring or fall mixing (Figure 1), while minimum values of both influx and efflux were observed during periods of stratification (Figure 1). The longest period of calculated continuous influx (negative flux, day and night) was 73 days in Lake Orient, while the longest period of calculated net efflux (positive flux, day and night) was 96 days in Lake Arrowhead out of 193 days of continuous sampling (Figure 1).

Organic and inorganic carbon sources

Across seasons and sites, the DOM pool was dominated by autochthonous, degraded organic matter, not of higher plant origin (Table S3; Figure 2; β:α: 0.77 ± 0.05; FI: 1.6 ± 0.06; HIX_{Ohno}: 0.82 ± 0.06). β:α, an indicator of the level of degradation of the DOM pool, ranged from 0.65 (degraded) in Springbrook Lake in April to 0.91 (newly produced) in Arrowhead Lake in July. FI values ranged from 1.5 to 1.8, and did not show substantial variation across sites or seasons. Values approaching 1.8 indicate microbial and algal leachate; lower values approaching 1.2 indicate terrestrial organic matter of higher plant origin or soil organic matter. HIX, indicative of DOM humic content, was between 0.54 (not humic) and 0.92 (humic), with the lowest values in Blackhawk Lake in August, and highest in Five Island Lake in July. Mean $\delta^{13}C$ signatures of the ambient DIC pool were -1.16 ± 3.40 ‰, with a range of -12.57 ‰ to 5.78 ‰ (Figure 3). The highest $\delta^{13}C_{DIC}$ values were measured in Center Lake in September, and the lowest in Lake Orient in July.

Predictors of discrete CO_2 flux

Discrete CO_2 flux was best predicted by chl *a* and TN concentration. Regression tree models revealed that discrete CO_2 influx across eutrophic lakes in this study was best predicted by chl *a* concentration greater than or equal to 24 μg L^{-1} , while efflux was predicted by TN greater than 12 mg L^{-1} and chl *a* less than 24 μg L^{-1} (Figure 4). At high concentrations of chl *a*, instantaneous wind speed equal to or exceeding 13 m s^{-1} predicted with the highest rate of influx, followed by HIX < 0.82. When chl *a* concentration was less than 24 μg L^{-1}, efflux was predicted by DO less than 9.3 mg L^{-1} and β:α less than 0.76 (Figure 4). The random forest model explained 31.25% of variation in the discrete dataset, and indicated that the most important predictor of CO_2 flux is chl *a*, followed by TN and HIX (Table 2).

Predictors of net CO_2 flux

Using static predictors, we found that net CO_2 influx during the ice-free season was best predicted by lack of watershed wetland cover (< 0.5%) by both classification tree and random forest models (Figure 5, Table 3). When percent wetland exceeded 0.5, the classification tree model predicted that lakes are rendered CO_2 sinks only if their WA:LA is less than 14, and Z_{max} is less than 5.3. When percent wetland cover exceeds 0.5, lakes are predicted to be CO_2 sources if WA:LA is greater than 14, or if WA:LA is less than 14 but Z_{max} is greater than 5.3. The random forest classification model ranked the relative importance of each static predictor as follows: % wetland, % pasture, % water, Z_{max}, WA:LA, and agriculture (Table 4).

Discussion

Our results indicate that anthropogenically eutrophic and hypereutrophic lakes exhibit extreme rates of CO_2 flux associated with autochthony and land-use. Five of the 15 lakes in this study maintained CO_2 influx, day and night, for days to months at a time (Fig. 1, Supplemental

plates S1-S15). In these lakes, atmospheric influx was predicted by indicators of autochthonous primary production (chlorophyll *a*, dissolved oxygen, newly produced dissolved organic matter) and small watershed to lake area ratios (Fig. 4 and 5). Lakes that were net CO_2 emitters over the sampling period had rates at the high end of literature reported values and were best predicted by nitrogen enrichment and wetland cover. This reflects spring post-drought release of nitrogen from agricultural soils in lakes with large watershed to lake area ratios (Howarth et al. 2012; Al-Kaisi et al. 2013; Loecke et al. 2017). Previously reported net flux rates for oligotrophic and mesotrophic lakes have ranged from 0.58 to 4.08 mol m^{-2} sampling period $^{-1}$ (Fig. 6; Table S4; Stets et al. 2009; Jones et al. 2016) compared to net efflux rates in eutrophic lakes reported here ranging from 0.33 to 11.76 mol m^{-2} sampling period $^{-1}$ (Fig. 6; Table S4). Because stable isotope analyses did not show evidence of degradation of terrestrial organic matter, these large flux events may be a result of nitrate and nitrite photodegradation (Brezonik and Fulkerson-Brekken 1998a; Schwarzenbach et al. 2003) primed by accumulated autochthonous organic matter in eutrophic and hypereutrophic lakes. Nitrate and nitrite are important intermediates in photochemistry of freshwater lakes, generating hydroxyl radicals (*OH) that are rapidly scavenged by all types of DOM (allochthonous and autochthonous) at approximately equal rates (Brezonik and Fulkerson-Brekken 1998a). This suggests that eutrophication processes resulting from both nitrogen and phosphorus loading can fundamentally alter gas flux and the contribution of inland waters to the global carbon budget, but that eutrophic lakes vary substantially in response due to regional variability in land use and land cover characteristics (Balmer and Downing 2011; Jones et al. 2016; Ouyang et al. 2017).

Across sites and seasons, we found that the organic matter pool was primarily autochthonous in our study lakes. Stable isotope analysis indicated that DIC was derived from

the atmosphere and mineral dissolution, rarely from heterotrophic degradation of terrestrial organic matter (Fig. 3). These findings demonstrate that human activity and eutrophication have not only degraded water quality and altered organic matter composition in lakes (Foley et al. 2005; Li et al. 2008; Williams et al. 2015), but may have much farther reaching effects on CO_2 flux and the role of lakes in the global carbon cycle. With increased land use alteration for agriculture and urban centers, more freshwater ecosystems will be subject to these pressures and may shift to eutrophic and hypereutrophic states (Foley et al. 2005; Heisler et al. 2008). This will depend on local conservation laws, as eutrophication in some parts of the world has leveled off or is declining due to fertilizer use restrictions (Bennett et al. 2001). In the agricultural Midwest U.S., however, where industrial row-crop agriculture dominates the landscape, reductions in non-point source nutrient pollution remain voluntary and are not enforced. Recent changes to the U.S. Clean Water Rule further reduce regulatory oversight of headwater streams, wetlands, and groundwater, which will have cascading impacts on lakes. Our results indicate that land-use alterations and eutrophication can substantially influence CO_2 flux rates, both as sustained influx resulting in net CO_2 sinks or as very large seasonal CO_2 efflux.

In these lakes, flux was negatively correlated with variables associated with eutrophication and primary productivity. Previous work in experimentally eutrophied ecosystems has suggested that the magnitude of the inorganic carbon demand of autochthonous primary producers will be less than the combined contributions of exogenous watershed CO_2 inputs and heterotrophic degradation of terrestrial organic matter (Wilkinson et al. 2016). While this may be accurate in northern temperate lakes having high contributions of terrestrial plant-derived, humic, and aromatic organic carbon (Sobek et al. 2005; Kothawala et al. 2014), it is not the case in lakes with agricultural watersheds and autochthonous carbon pools. This is evidenced by

previous studies (Balmer and Downing 2011; Pacheco et al. 2014) and our 5 lakes having net CO_2 influx during the open water season. Optical characterization of DOM in these 5 lakes indicated that their organic matter pools were dominated by compounds resembling newly produced bacterial and algal leachate with low humic content (Table S3). Correspondingly, our model indicated that non-humic DOM (HIX < 0.82) was an important predictor of CO_2 influx at high chl *a* concentration.

While 5 of the lakes in this study exhibited net CO_2 influx, 10 showed the opposite trend and were net emitters of CO_2. High efflux was best predicted by TN and watershed wetland cover. 2012 was a severe drought year in the Midwestern U.S., resulting in high nitrate accumulation in agricultural soils (Al-Kaisi et al. 2013), which has been shown in many studies to increase nitrate export during rain events (Watmough et al. 2004; Mosley 2015). Badger Lake, which had more than double the net efflux than any other lake in this study (11,755 mmol m^{-2}), also had record high nitrate levels in 2012, ranging from 0.4 to 16.8 mg NO_3-N L^{-1} with a mean value of 9.4 ± 5.8 mg NO_3-N L^{-1}. One possible explanation for the co-occurrence of elevated nitrate concentrations and flux rates is the photodegradation of nitrate and nitrite in surface waters. This process would generate hydroxyl radicals which mineralize organic carbon, increasing CO_2 efflux (Brezonik and Fulkerson-Brekken 1998b; Molot et al. 2003; Filstrup and Downing 2017). It is also possible that DIC export during rain events could result in outgassing of terrestrially-derived CO_2, though TN was consistently the best predictor of efflux in our regression tree and random forest models.

With the exception of Badger Lake, rates of net efflux during the 2012 ice free seasons in these lakes (April 1 to mid-November) ranged from 327 to 5474 mmol m^{-2}. These values are in a comparable range of previous studies in temperate lakes and reservoirs (Kosten et al. 2010;

Barros et al. 2011; Pacheco et al. 2014; Jones et al. 2016), but on average 3 to 4 times higher than previous studies in the same lakes (Pacheco et al. 2014) and more recent studies in artificially fertilized northern temperate lakes (Wilkinson et al. 2016) (Figure 6, Supplemental Table S4). Calculated flux in Blackhawk Lake reported in Pacheco *et al.* (2014) was determined based on monthly discrete measurements, compared with high frequency measurements in this study. Discrete daytime measurements do not capture nighttime respiratory flux that high frequency sensors do, and may miss large transient fluxes associated with mixing events. The largest periods of efflux in our study occurred during spring and fall mixing (Figure 1). These periods were not captured in Wilkinson *et al.* (2016), which calculated flux for 100 days between late May and late August, and is one possible explanation for the large difference in net efflux between the two studies of lakes with comparable trophic status. Similar seasonal trends were found by Jones *et al.* (2016) across a gradient of oligotrophic to eutrophic lakes, where the majority of annual efflux occurred during spring and fall mixing events, and by Trolle *et al.* 2012 where calculated CO_2 influx and efflux was lowest in summer months across 151 Danish lakes.

Stable isotopic analysis of DIC pools across all lakes in this study ranged from -12.57 to 5.78 $\delta^{13}C_{DIC}$ ‰, indicating they were derived primarily from mineral dissolution and atmospheric sources, but not from heterotrophic degradation of terrestrial organic matter (Figure 3). While $\delta^{13}C_{DIC}$ was not an important predictor of CO_2 flux in our models, our measured range demonstrates that the inorganic carbon in these lakes is not sourced from degradation of terrestrial organic matter, which is an important determinant of efflux in northern temperate and boreal lakes. Published values of $\delta^{13}C_{DIC}$ in lake surface waters generally range from -29.6 ‰ to +2.6 ‰, where the lowest values indicate heterotrophic degradation of terrestrial organic matter (Bade et al. 2004). Depending on proximity to industry and urban areas, atmospheric sources

range from -7.5 ‰ to -12 ‰, though the global atmosphere is fairly well mixed and has a nominal value of around -8.5 ‰ (Mook 1986; Boutton 1991). At high pH values, chemically enhanced diffusion can result in fractionation that would decrease $\delta^{13}C_{DIC}$ values (Bade and Cole 2006). If bloom-forming phytoplankton are taking up mineral bicarbonate rather than CO_2, heterotrophic degradation of autochthonous material should result in $\delta^{13}C_{DIC}$ values between -15 and -10 ‰ (Morales-Williams et al. 2017). Values associated with carbonate dissolution typically span from -15‰ to 0‰, however such values (and higher) can also be attributable to sediment methanogenic fermentation in shallow systems (Boutton 1991). Due to high rates of water column primary productivity, hypolimnetic hypoxia and anoxia, and sediment organic carbon accumulation in these systems (Heathcote and Downing 2011), methanogenic fermentation is a plausible explanation for elevated $\delta^{13}C_{DIC}$ values measured in this study.

Non-humic DOM was a predictor of CO_2 influx at high chl *a* concentration in hypereutrophic lakes, demonstrating that autochthonous carbon is an important driver of CO_2 dynamics in these systems. Optical characterization of DOM indicated that across lakes, the organic matter pool was composed primarily of endogenous material, both fresh and degraded, with moderate humic content (Table S3, Figure 2). Average FI values, indicating DOM source material, were 1.6 ± 0.06. Values approaching 1.2 would indicate terrestrial organic matter of higher plant origin, while values approaching 1.8 are reflective of algal and microbial leachate (McKnight et al. 2001). The average value of β:α across sites and sampling events was 0.77 ± 0.05, suggesting mixed contributions of fresh and degraded material in these lakes. Lower values of β:α indicate DOM is highly degraded (~0.5), while higher values indicate the DOM pool is fresh, or recently produced (~1.0) (Parlanti et al. 2000; Wilson and Xenopoulos 2008). Similarly, HIX values in our study suggested contributions of humic organic matter in the DOM pool.

Because FI values indicate the DOM pool is of microbial and algal origin, and $\delta^{13}C_{DIC}$ values do not indicate degradation of terrestrial organic matter is occurring in these systems, these HIX values are likely attributable to microbial humics and reflect rapid processing of endogenous material. These patterns are supported by the β:α ratio, which suggests a large portion of the DOM pool in these lakes has been processed and degraded.

Our results demonstrate that anthropogenically eutrophic lakes can function as significant sources and sinks of CO_2. While previous work with some exceptions has demonstrated that lakes generally act as net sources of CO_2 to the atmosphere (Cole et al. 1994; Sobek et al. 2005; Wilkinson et al. 2016), we show that inorganic carbon uptake by primary producers can far exceed contributions from heterotrophy and mineral dissolution. While two-thirds of our study sites were net CO_2 emitters, we show that hypereutrophic lakes with <0.5% watershed wetland cover maintained negative flux (i.e., continuous CO_2 uptake) for months at a time, meaning high rates of primary production in these impacted ecosystems was not balanced or exceeded by respiration or exogenous DIC inputs (Wilkinson et al. 2016). Lakes that were net sources of CO_2 had substantially higher flux rates than oligotrophic or mesotrophic lakes previously reported in the literature, and these trends were best predicted by nitrogen enrichment. Our findings further indicate that the carbon supplies of these lake food webs are supported by autochthonous sources, have minimal contributions of terrestrial organic matter, and are cycled by autochthonous processes, as evidenced by both optical characterization of DOM and stable isotope analyses. Taken together, these findings further demonstrate that anthropogenic eutrophication has fundamentally changed lake biogeochemistry, gas flux, and their role in the global carbon cycle. As global land use changes to accommodate a large and growing human population, it is likely that more freshwater ecosystems will shift to eutrophic and hypereutrophic

states (Cole et al. 2007; Tranvik et al. 2009) in regions where non-point source pollution remains unregulated. The impacts of these processes on lake carbon cycles will depend on the extent of eutrophication, and regional scale watershed characteristics. In our models, nitrogen loading was the strongest predictor of extreme efflux rates in eutrophic lakes, and net influx was best predicted by chlorophyll and lack of watershed wetland cover. Extensive watershed cultivation without wetland buffers would be expected to drive lakes toward net CO_2 sinks, while post-drought nitrogen release from agricultural watersheds is expected to result in high rates of CO_2 efflux. It will, therefore, be critical to integrate eutrophic and hypereutrophic systems into the global carbon budget and evaluate the effects of these changes at global scales.

Acknowledgements

We thank Amber Erickson, Lisa Whitehouse, and Suzanne Ankerstjerne for chemical and analytical assistance, and Adam Heathcote for his contributions to site selection and sampling design. Lakes in this study occupy indigenous lands of the Meskwaki, Sauk, Ho-Chunk, Ioway, and Dakota, forcibly ceded in multiple treaties between 1824 and 1853 (http://www.iowahild.com/index.html). The Meskwaki Nation currently resides on 7,000 acres in Meskwakenuk in Tama County, IA. This study was funded by a grant from the National Science Foundation to John A. Downing, DEB-1021525.

References

Al-Kaisi, M. M., R. W. Elmore, J. G. Guzman, and others. 2013. Drought impact on crop production and the soil environment: 2012 experiences from Iowa. J. Soil Water Conserv. **68**: 19A-24A. doi:10.2489/jswc.68.1.19A

Arar, E. J., and G. B. Collins. 1997. Method 445.0: In vitro determination of chlorophyll a and pheophyton a in marine and freshwater algae by fluorescence: Revision 1.2. EPA 22.

Bade, D. L., S. R. Carpenter, J. J. Cole, P. C. Hanson, and R. H. Hesslein. 2004. Controls of delta 13 C-DIC in lakes : Geochemistry , lake metabolism , and morphometry. Limnol. Oceanogr. **49**: 1160–1172.

Bade, D. L., and J. J. Cole. 2006. Impact of chemically enhanced diffusion on dissolved inorganic carbon stable isotopes in a fertilized lake. J. Geophys. Res. Ocean. **111**: 1–10. doi:10.1029/2004JC002684

Balmer, M. B., and J. A. Downing. 2011. Carbon dioxide concentrations in eutrophic lakes : undersaturation implies atmospheric uptake. Inl. Waters **1**: 125–132. doi:10.5268/IW-1.2.366

Barros, N., J. J. Cole, L. J. Tranvik, Y. T. Prairie, D. Bastviken, V. L. M. Huszar, P. del Giorgio, and F. Roland. 2011. Carbon emission from hydroelectric reservoirs linked to reservoir age and latitude. Nat. Geosci. **4**: 593–596. doi:10.1038/ngeo1211

Beirne, E. C., A. D. Wanamaker, and S. C. Feindel. 2012. Experimental validation of environmental controls on the δ13C of Arctica islandica (ocean quahog) shell carbonate. Geochim. Cosmochim. Acta **84**: 395–409. doi:10.1016/j.gca.2012.01.021

Bennett, E. M. ., S. R. Carpenter, and N. F. Caraco. 2001. Human Impact on Erodable Phosphorus and Eutrophication: A Global Perspective. Bioscience **51**: 227. doi:10.1641/0006-3568(2001)051[0227:hioepa]2.0.co;2

Boutton, T. W. 1991. Stable carbon isotope ratios of natural materials: Atmospheric, terrestrial, marine, and freshwater environments, p. 173–183. *In* D.C. Coleman and B. Fry [eds.], Carbon Isotope Techniques.

Brezonik, P. L., and J. Fulkerson-Brekken. 1998a. Nitrate-induced photolysis in natural waters: Controls on concentrations of hydroxyl radical photo-intermediates by natural scavenging agents. Environ. Sci. Technol. **32**: 3004–3010. doi:10.1021/es9802908

Brezonik, P. L., and J. Fulkerson-Brekken. 1998b. Nitrate-Induced Photolysis in Natural Waters: Controls on Concentrations of Hydroxyl Radical Photo-Intermediates by Natural Scavenging Agents. Environ. Sci. Technol. **32**: 3004–3010. doi:10.1021/es9802908

Brooks, B. W., J. M. Lazorchak, M. D. A. Howard, and others. 2016. Are harmful algal blooms becoming the greatest inland water quality threat to public health and aquatic ecosystems? Environ. Toxicol. Chem. **35**: 6–13. doi:10.1002/etc.3220

Cole, J. J., and N. F. Caraco. 1998. Atmospheric exchange of carbon dioxide in a low-wind oligotrophic lake measured by the addition of SF6. Limnol. Oceanogr. **43**: 647–656. doi:10.4319/lo.1998.43.4.0647

Cole, J. J., N. F. Caraco, G. W. Kling, and T. K. Kratz. 1994. Carbon dioxide supersaturation in the surface waters of lakes. Science **265**: 1568–70. doi:10.1126/science.265.5178.1568

Cole, J. J., Y. T. Prairie, N. F. Caraco, 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

Cory, R., M. Miller, D. M. McKnight, J. J. Guerard, and P. L. Miller. 2010. Effect of instrument-specific response on the analysis of fulvic acid fluorescence spectra. Limnol. Oceanogr. Methods **8**: 67–78. doi:10.4319/lom.2010.8.67

Crumpton, W. D., T. M. Isenhart, and P. D. Mitchell. 1989. Nitrate and organic N analyses with second-derivative spectroscopy. Limnol. Oceanogr. **37**: 907–913.

Duarte, C. M., and Y. T. Prairie. 2005. Prevalence of Heterotrophy and Atmospheric CO2

Emissions from Aquatic Ecosystems. Ecosystems **8**: 862–870. doi:10.1007/s10021-005-0177-4

Filstrup, C. T., and J. A. Downing. 2017. Relationship of chlorophyll to phosphorus and nitrogen in nutrient-rich lakes. Inl. Waters **7**: 385–400. doi:10.1080/20442041.2017.1375176

Finlay, K., P. R. Leavitt, a. Patoine, and B. Wissel. 2010. Magnitudes and controls of organic and inorganic carbon flux through a chain of hard-water lakes on the northern Great Plains. Limnol. Oceanogr. **55**: 1551–1564. doi:10.4319/lo.2010.55.4.1551

Finlay, K., P. R. Leavitt, B. Wissel, and Y. T. Prairie. 2009. Regulation of spatial and temporal variability of carbon flux in six hard-water lakes of the northern Great Plains. **54**: 2553–2564.

Foley, J. a, R. Defries, G. P. Asner, and others. 2005. Global consequences of land use. Science **309**: 570–4. doi:10.1126/science.1111772

Del Giorgio, P. A., J. J. Cole, N. F. Caraco, and R. H. Peters. 2009. Linking Planktonic Biomass and Metabolism to Net Gas Fluxes in Northern Temperate Lakes. Ecology **80**: 1422–1431.

Golub, M., A. R. Desai, G. A. McKinley, C. K. Remucal, and E. H. Stanley. 2017. Large Uncertainty in Estimating pCO2 From Carbonate Equilibria in Lakes. J. Geophys. Res. Biogeosciences **122**: 2909–2924. doi:10.1002/2017JG003794

Heathcote, A. J., and J. a. Downing. 2011. Impacts of Eutrophication on Carbon Burial in Freshwater Lakes in an Intensively Agricultural Landscape. Ecosystems **15**: 60–70. doi:10.1007/s10021-011-9488-9

Heisler, J., P. M. Glibert, J. M. Burkholder, and others. 2008. Eutrophication and harmful algal blooms: A scientific consensus. Harmful Algae **8**: 3–13. doi:10.1016/j.hal.2008.08.006

Howarth, R., D. Swaney, G. Billen, and others. 2012. Nitrogen fluxes from the landscape are

controlled by net anthropogenic nitrogen inputs and by climate. Front. Ecol. Environ. **10**: 37–43. doi:10.1038/news050808-1

Jeffrey, S. W., R. F. C. Mantoura, and S.W. Wright. 1997. Phytoplankton Pigments in Oceanography.

Johnson, M., M. Billett, K. Dinsmore, M. Wallin, K. E. Dyson, and R. S. Jassal. 2009. Direct and continuous measurement of dissolved carbon dioxide in freshwater aquatic systems—method and applications. Ecohydrology. doi:10.1002/eco

Jones, J. R., D. V Obrecht, J. L. Graham, M. B. Balmer, C. T. Filstrup, and J. A. Downing. 2016. Seasonal patterns in carbon dioxide in 15 mid-continent (USA) reservoirs. Inl. Waters 265–272. doi:10.5268/IW-6.2.982

Jones, S. E., T. K. Kratz, C.-Y. Chiu, and K. D. McMAHON. 2009. Influence of typhoons on annual CO 2 flux from a subtropical, humic lake. Glob. Chang. Biol. **15**: 243–254. doi:10.1111/j.1365-2486.2008.01723.x

Kellerman, A. M., D. N. Kothawala, T. Dittmar, and L. J. Tranvik. 2015. Persistence of dissolved organic matter in lakes related to its molecular characteristics. Nat. Geosci. **8**: 454–459. doi:10.1038/NGEO2440

Kling, G., G. Kipphut, and M. Miller. 1992. The flux of CO2 and CH4 from lakes and rivers in arctic Alaska. Hydrobiologia **240**: 23–36.

Kosten, S., F. Roland, D. M. L. Da Motta Marques, E. H. Van Nes, N. Mazzeo, L. D. S. L. Sternberg, M. Scheffer, and J. J. Cole. 2010. Climate-dependent CO 2 emissions from lakes. Global Biogeochem. Cycles **24**: 1–7. doi:10.1029/2009GB003618

Kothawala, D. N., C. a Stedmon, R. a Müller, G. a Weyhenmeyer, S. J. Köhler, and L. J. Tranvik. 2014. Controls of dissolved organic matter quality: evidence from a large-scale

boreal lake survey. Glob. Chang. Biol. **20**: 1101–14. doi:10.1111/gcb.12488

Li, L., Z. Yu, R. E. Moeller, and G. E. Bebout. 2008. Complex trajectories of aquatic and terrestrial ecosystem shifts caused by multiple human-induced environmental stresses. Geochim. Cosmochim. Acta **72**: 4338–4351. doi:10.1016/j.gca.2008.06.026

Liaw, A., and M. Wiener. 2002. Classification and regression by randomForest. R news.

Loecke, T. D., A. J. Burgin, D. A. Riveros-Iregui, A. S. Ward, S. A. Thomas, C. A. Davis, and M. A. S. Clair. 2017. Weather whiplash in agricultural regions drives deterioration of water quality. Biogeochemistry **133**: 7–15. doi:10.1007/s10533-017-0315-z

López, P., R. Marcé, and J. Armengol. 2011. Net heterotrophy and CO 2 evasion from a productive calcareous reservoir: Adding complexity to the metabolism-CO 2 evasion issue. J. Geophys. Res. **116**: 1–14. doi:10.1029/2010JG001614

Marcé, R., B. Obrador, Josep-Anton Morguí, J. L. R. P. López, and A. Joan. 2015. Carbonate weathering as a driver of CO 2 supersaturation in lakes. Nat. Geosci. **8**: 1–5. doi:10.1038/NGEO2341

McKnight, D., E. Boyer, P. Westerhoff, P. Doran, and T. Kulbe. 2001. Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. Limnol. Oceanogr. **46**: 38–48.

Molot, L. A., S. A. Miller, P. J. Dillon, and C. G. Trick. 2003. A simple method for assaying extracellular hydroxyl radical activity and its application to natural and synthetic waters. **213**: 203–213. doi:10.1139/F03-014

Mook, W. G. 1986. 13C in Atmospheric CO2. Netherlands J. Sea Res. **20**: 211–223.

Morales-Williams, A. M., A. D. Wanamaker, and J. A. Downing. 2017. Cyanobacterial carbon concentrating mechanisms facilitate sustained CO2 depletion in eutrophic lakes.

Biogeosciences **14**: 2865–2875. doi:10.5194/bg-14-2865-2017

Mosley, L. M. 2015. Drought impacts on the water quality of freshwater systems; review and integration. Earth-Science Rev. **140**: 203–214. doi:10.1016/j.earscirev.2014.11.010

Murphy, K. R., K. D. Butler, R. G. M. Spencer, C. a Stedmon, J. R. Boehme, and G. R. Aiken. 2010. Measurement of dissolved organic matter fluorescence in aquatic environments: an interlaboratory comparison. Environ. Sci. Technol. **44**: 9405–12. doi:10.1021/es102362t

Nõges, P., F. Cremona, A. Laas, and others. 2016. Role of a productive lake in carbon sequestration within a calcareous catchment. Sci. Total Environ. **550**: 225–230. doi:10.1016/j.scitotenv.2016.01.088

Ohno, T. 2002. Fluorescence Inner-Filtering Correction for Determining the Humification Index of Dissolved Organic Matter. Environ. Sci. Technol. **36**: 742–746. doi:10.1021/es0155276

Ouyang, Z., C. Shao, H. Chu, R. Becker, T. Bridgeman, C. A. Stepien, R. John, and J. Chen. 2017. The effect of algal blooms on carbon emissions in western lake erie: An integration of remote sensing and eddy covariance measurements. Remote Sens. **9**: 1–19. doi:10.3390/rs9010044

Pace, M. L., and Y. T. Prairie. 2005. Respiration in Lakes in Respiration in Aquatic Ecosystems, P. Del Giorgio and P. Williams [eds.]. Oxford University Press.

Pacheco, F., F. Roland, and J. Downing. 2014. Eutrophication reverses whole-lake carbon budgets. Inl. Waters **4**: 41–48. doi:10.5268/IW-4.1.614

Parlanti, E., K. Wo, L. Geo, and M. Lamotte. 2000. Dissolved organic matter fluorescence spectroscopy as a tool to estimate biological activity in a coastal zone submitted to anthropogenic inputs. Org. Geochem. **31**: 1765–1781.

Petrone, K. C., J. B. Fellman, E. Hood, M. J. Donn, and P. F. Grierson. 2011. The origin and

function of dissolved organic matter in agro-urban coastal streams. J. Geophys. Res. **116**. doi:10.1029/2010JG001537

R Core Team. 2015. R Core Team (2015). R: A language and environment for statistical computing.

Raymond, P. A., and J. E. Bauer. 2001. DOC cycling in a temperate estuary : A mass balance approach using natural 14 C and C isotopes. Limnol. Oceanogr. **46**: 655–667.

Schwarzenbach, R., P. Gschwend, and D. Imbiden. 2003. Indirect photolysis: Reactions with photooxidants in natural waters and in the atmosphere. In Environmental Organic Chemistry, 2nd ed. Wiley.

Sobek, S., L. Tranvik, and J. Cole. 2005. Temperature independence of carbon dioxide supersaturation in global lakes. Global Biogeochem. Cycles **19**: 1–10. doi:10.1029/2004GB002264

Stets, E. G., R. G. Striegl, G. R. Aiken, D. O. Rosenberry, and T. C. Winter. 2009. Hydrologic support of carbon dioxide flux revealed by whole-lake carbon budgets. J. Geophys. Res. **114**: 1–14. doi:10.1029/2008JG000783

Stumm, W., and J. Morgan. 1996. Aquatic chemistry: chemical equilibria and rates in natural waters, 3rd ed. Wiley-Interscience.

Therneau, T., B. Atkinson, and B. Ripley. 2017. rpart: Recursive Partitioning and Regression Trees. 2017.

Tranvik, L. J., J. A. Downing, J. B. Cotner, and others. 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

Trolle, D., P. A. Staehr, T. A. Davidson, R. Bjerring, T. L. Lauridsen, M. Søndergaard, and E.

Jeppesen. 2012. Seasonal Dynamics of CO 2 Flux Across the Surface of Shallow Temperate Lakes. Ecosystems **15**: 336–347. doi:10.1007/s10021-011-9513-z

Wanninkhof, R. 1992. Relationship Between Wind Speed and Gas Exchange. J. Geophys. Res. **97**: 7373–7382.

Watmough, S. A., M. C. Eimers, J. Aherne, and P. J. Dillon. 2004. Climate Effects on Stream Nitrate Concentrations at 16 Forested Catchments in South Central Ontario. Environ. Sci. Technol. **38**: 2383–2388. doi:10.1021/es035126l

Weyhenmeyer, G. A., S. Kosten, M. B. Wallin, L. J. Tranvik, E. Jeppesen, and F. Roland. 2015. Significant fraction of CO2 emissions from boreal lakes derived from hydrologic inorganic carbon inputs. Nat. Geosci. **8**: 933–936. doi:10.1038/NGEO2582

Wilkinson, G. M., C. D. Buelo, J. J. Cole, and M. L. Pace. 2016. Exogenously produced CO 2 doubles the CO 2 efflux from three north temperate lakes. Geophys. Res. Lett. **43**. doi:10.1002/2016GL067732

Williams, C. J., P. C. Frost, A. M. Morales-Williams, J. H. Larson, W. B. Richardson, A. S. Chiandet, and M. A. Xenopoulos. 2015. Human activities cause distinct dissolved organic matter composition across freshwater ecosystems. Glob. Chang. Biol. 613–626. doi:10.1111/gcb.13094

Williams, C., Y. Yamashita, H. Wilson, R. Jaffé, and M. Xenopoulos. 2010. Unraveling the role of land use and microbial activity in shaping dissolved organic matter characteristics in stream ecosystems. Limnol. Oceanogr. **55**: 1159–1171. doi:10.4319/lo.2010.55.3.1159

Wilson, H. F., and M. A. Xenopoulos. 2008. Effects of agricultural land use on the composition of fluvial dissolved organic matter. Nat. Geosci. **2**: 37–41. doi:10.1038/ngeo391

Winslow, L., J. Read, R. Woolway, J. Brentrup, T. Leach, and J. Zwart. 2017. rLakeAnalyzer:

Lake Physics Tools. R package version 1.11.0. 2017.

Zsolnay, A., E. Baigar, M. Jimenez, B. Steinweg, and F. Saccomandi. 1999. Differentiating with fluorescence spectroscopy the sources of dissolved organic matter in soils subjected to drying. Chemosphere **38**: 45–50. doi:10.1016/S0045-6535(98)00166-0

2012. APHA Standard Methods for the examination of waste and wastewater, 22nd ed. American Public Health Association.

Table Legend

Table 1. Discrete summary data for lakes included in this study measured between April (ice-off) and November 2012. Total phosphorus (TP), total nitrogen (TN), chlorophyll *a* (Chl *a*), DOC, pH, and alkalinity are reported as average values of n=23 sampling events ± standard deviation.

Table 2. Relative importance of predictor variables of discrete CO_2 flux in random forest regression model. Importance is estimated by IncNodePurity which indicates the total decrease in node impurities from splitting on the variable averaged over all attempted trees and measured by residual sum of squares.

Table 3. Relative importance of categorical predictor variables of net CO2 flux in random forest classification model. Importance is estimated by MeanDecreaseGini, which informs splits based on contribution of the variable to model accuracy and degree of misclassification. Higher values indicate greater predictive importance.

704 **Table 1**

Lake	*Latitude*	*Longitude*	*TP ($\mu g\ L^{-1}$)*	*TN ($mg\ L^{-1}$)*	*Chl a ($\mu g\ L^{-1}$)*	*DOC ($mg\ L^{-1}$)*	*pH*	*Alkalinity ($mg\ CaCO_3\ L^{-1}$)*
Arrowhead	42.297218	-95.051228	26 ± 9	0.9 ± 0.2	11 ± 6	5.9 ± 0.50	8.4 ± 0.1	189.7 ± 8.4
Badger	42.586161	-94.192562	58 ± 35	9.4 ± 5.8	34 ± 35	4.4 ± 1.8	8.3 ± 0.3	169.8 ± 33.4
Beeds	42.770320	-93.236436	76 ± 49	7.5 ± 4.6	48 ± 40	4.0 ± 0.8	8.4 ± 0.3	189.8 ± 37.0
Black Hawk	42.296334	-95.029191	226 ± 119	2.4 ± 0.6	78 ± 35	9.0 ± 1.3	8.7 ± 0.3	187.3 ± 12.1
Center	43.412607	-95.136293	104 ± 50	1.9 ± 0.3	42 ± 36	10.5 ± 0.6	8.5 ± 0.2	160.5 ± 4.2
East Osceola	41.032548	-93.742649	196 ± 78	1.9 ± 0.5	80 ± 48	8.8 ± 1.2	8.9 ± 0.6	111.7 ± 27.0
Five Island	43.145274	-94.658204	106 ± 51	2.1 ± 0.4	67 ± 38	7.7 ± 1.8	8.4 ± 0.2	163.5 ± 8.1
George Wyth	42.534834	-92.400362	62 ± 22	1.0 ± 0.2	26 ± 7	4.4 ± 0.3	8.4 ± 0.2	142.0 ± 28.0
Keomah	41.295123	-92.537482	107 ± 105	1.4 ± 0.7	45 ± 52	6.3 ± 1.2	8.7 ± 0.4	116.3 ± 16.6
Orient	41.196669	-94.436084	398 ± 286	2.3 ± 1.3	144 ± 105	7.4 ± 1.3	9.4 ± 0.6	94.4 ± 24.7
Lower Gar	43.352299	-95.120186	96 ± 35	1.7 ± 0.3	51 ± 23	7.9 ±1.2	8.6 ± 0.1	186.3 ± 13.6
Rock Creek	41.736936	-92.851859	115 ± 45	1.7 ± 0.5	53 ± 50	4.9 ± 1.2	8.5 ± 0.2	148.5 ± 7.5
Silver (Dickinson)	43.439162	-95.336799	161 ± 85	2.2 ± 0.9	35 ± 59	6.9 ± 0.9	8.3 ± 0.2	170.4 ± 13.5
Silver (Palo Alto)	43.030775	-94.883701	340 ± 206	2.6 ± 0.6	118 ± 60	8.2 ± 2.1	8.8 ± 0.3	174.9 ± 28.0
Springbrook	41.775930	-94.466736	38 ± 26	1.8 ± 0.9	18 ± 14	3.5 ± 0.4	8.4 ± 0.3	183.5 ± 22.0

Table 2.

Predictor	**Importance (IncNodePurity)**
Chl *a*	6311
TN	5800
HIX	4101
DO	3354
Precipitation	3243
$\delta^{13}C_{DIC}$	2672
BA	2448
TP	2363
Wind gust speed	2281
DOC	2204
Wind speed	2166
Schmidt stability	1487
Epilimnion temperature	1438
Site depth	1032
FI	990
Thermocline depth	576

Table 3.

Predictor	**Importance (MeanDecreaseGini)**
% Wetland	1.46
% Pasture	1.34
% Water	0.99
Z_{max}	0.89
Watershed to lake area	0.84
% Agriculture	0.74

Figure captions

Figure 1. Time series of average daily CO_2 flux (mmol C m^{-2} d^{-1}) calculated for lakes in this study and corresponding thermal heatmap visualizing seasonal stratification patterns. Color legend units are °C. Black lines on flux plots are modeled flux; grey lines are 95% confidence intervals.

Figure 2. Distribution of DOM quality indices measured in this study. (a) Fluorescence index (FI). Indicator of DOM source material. (b) β:α ratio. Index of DOM degradation (c) Humification index (HIX).

Figure 3. Distribution of isotopic composition of dissolved organic carbon (δ^{13}DIC) across lakes and sampling events. Values between -25 to -30 ‰ indicate heterotrophic degradation of terrestrial organic matter. -15 to -10 ‰ reflect degradation of bloom organic matter when primary producers are taking up mineral bicarbonate (~ -10 ‰) rather than CO_2. -8.5 ‰ indicates atmospheric CO_2. Few observations of δ^{13}DIC of atmospheric origin is likely attributable to rapid fractionation by surface blooms at the air-water interface (e.g., Morales-Williams et al. 2017). Values between -10 and 0 ‰ and higher reflect mineral dissolution and methanogenic fermentation.

Figure 4. Regression tree model visualizing predictors of discrete CO_2 flux. Terminal node values represent instantaneous flux rates where influx appears on the left, and efflux on the right. Abbreviations are as follows: TN indicates total nitrogen (mg L^{-1}), chl.a indicates chlorophyll a

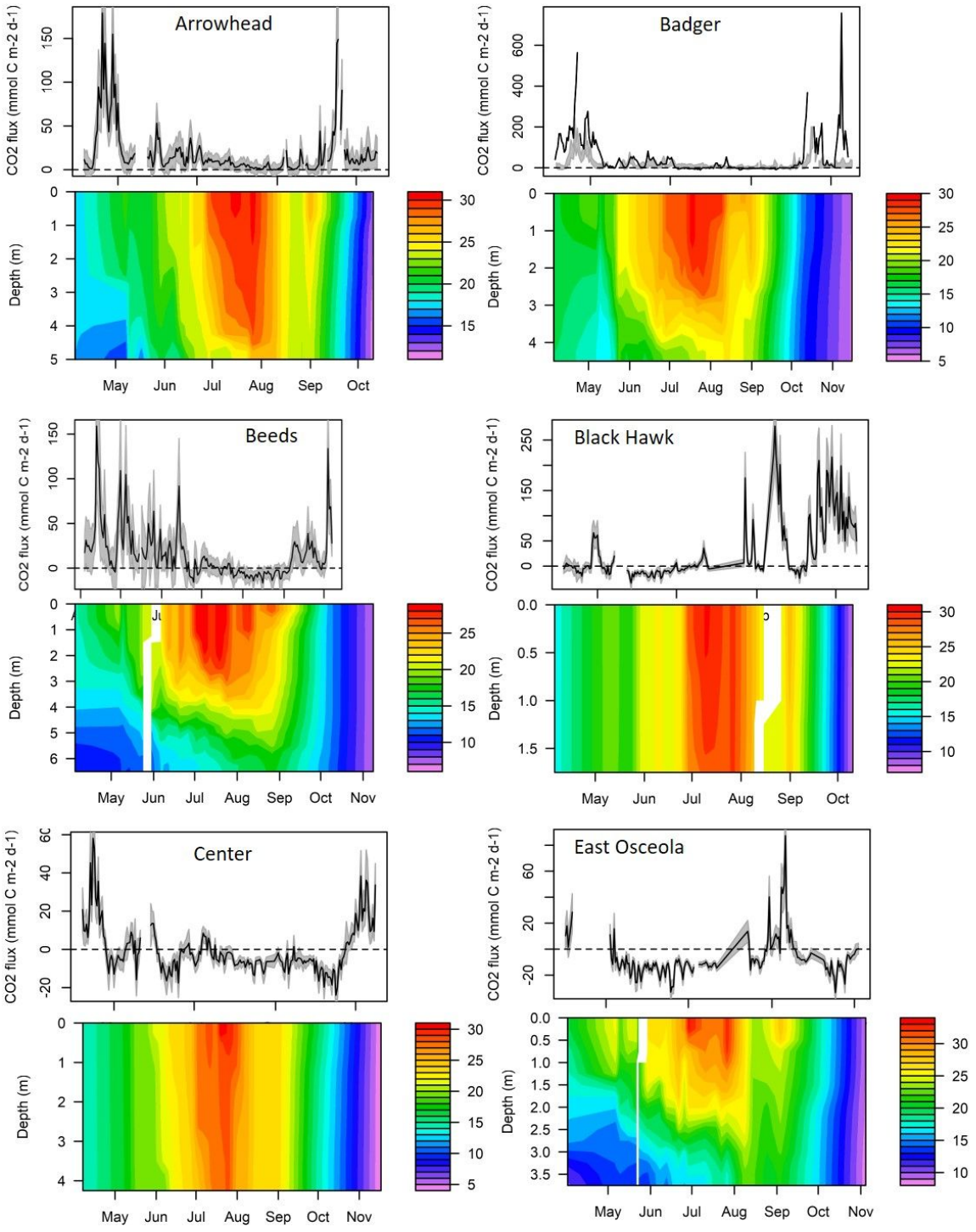
(ug L^{-1}), DO indicates epilimnetic or surface (integrated 2 m) dissolved oxygen (mg L^{-1}), HIX indicates the humification index, and BA indicates the beta:alpha index.

Figure 5. Classification tree model visualizing static predictors of net CO_2 flux, indicated as a net source or sink. Abbreviations are as follows: WtoL indicates watershed to lake area ratio, and Z_{max} indicates maximum lake depth.

Figure 6. Range of flux values previously reported in literature across trophic state. Data sources and sampling periods can be found in supplemental table S4.

757

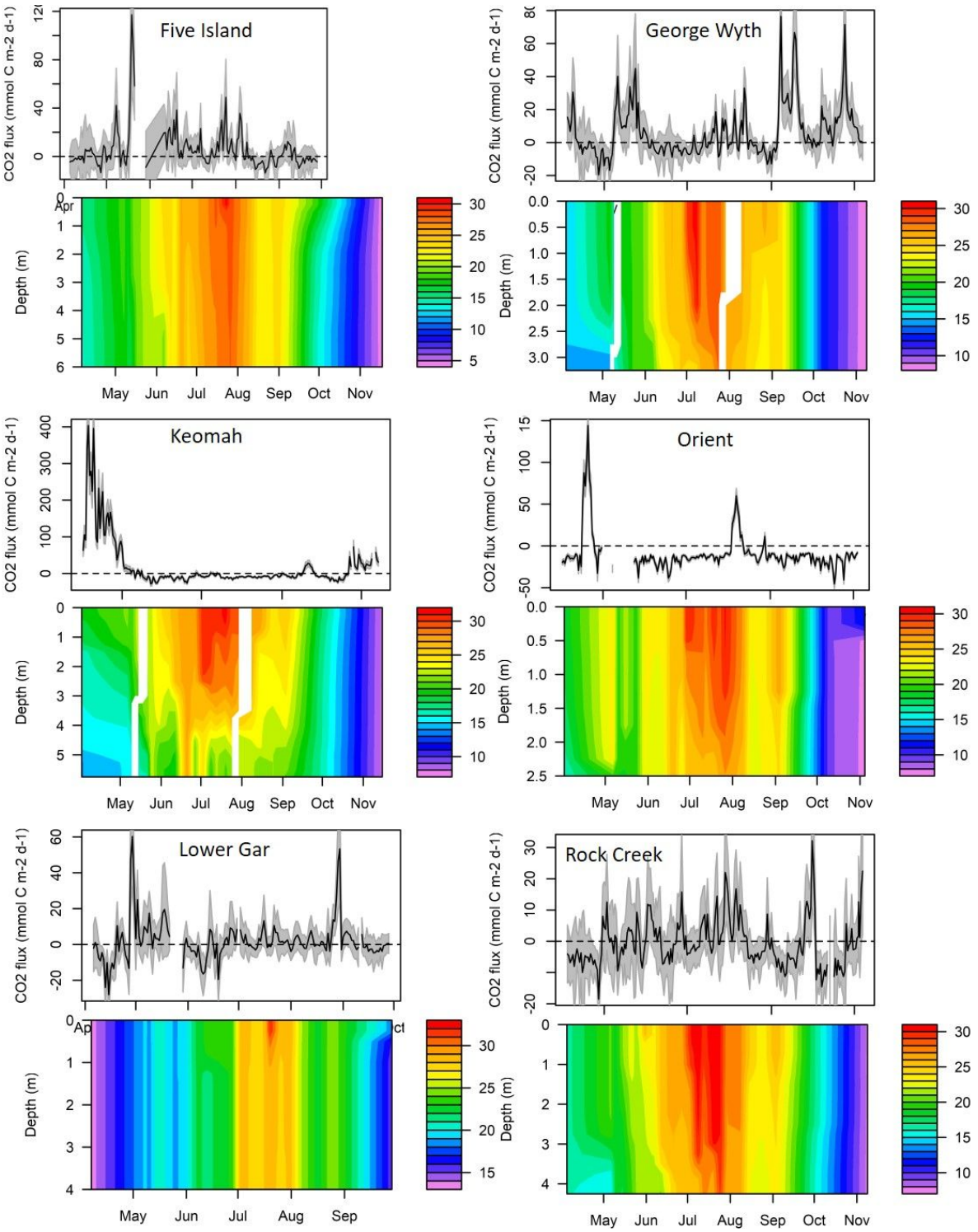
758 Figure 1



759

760

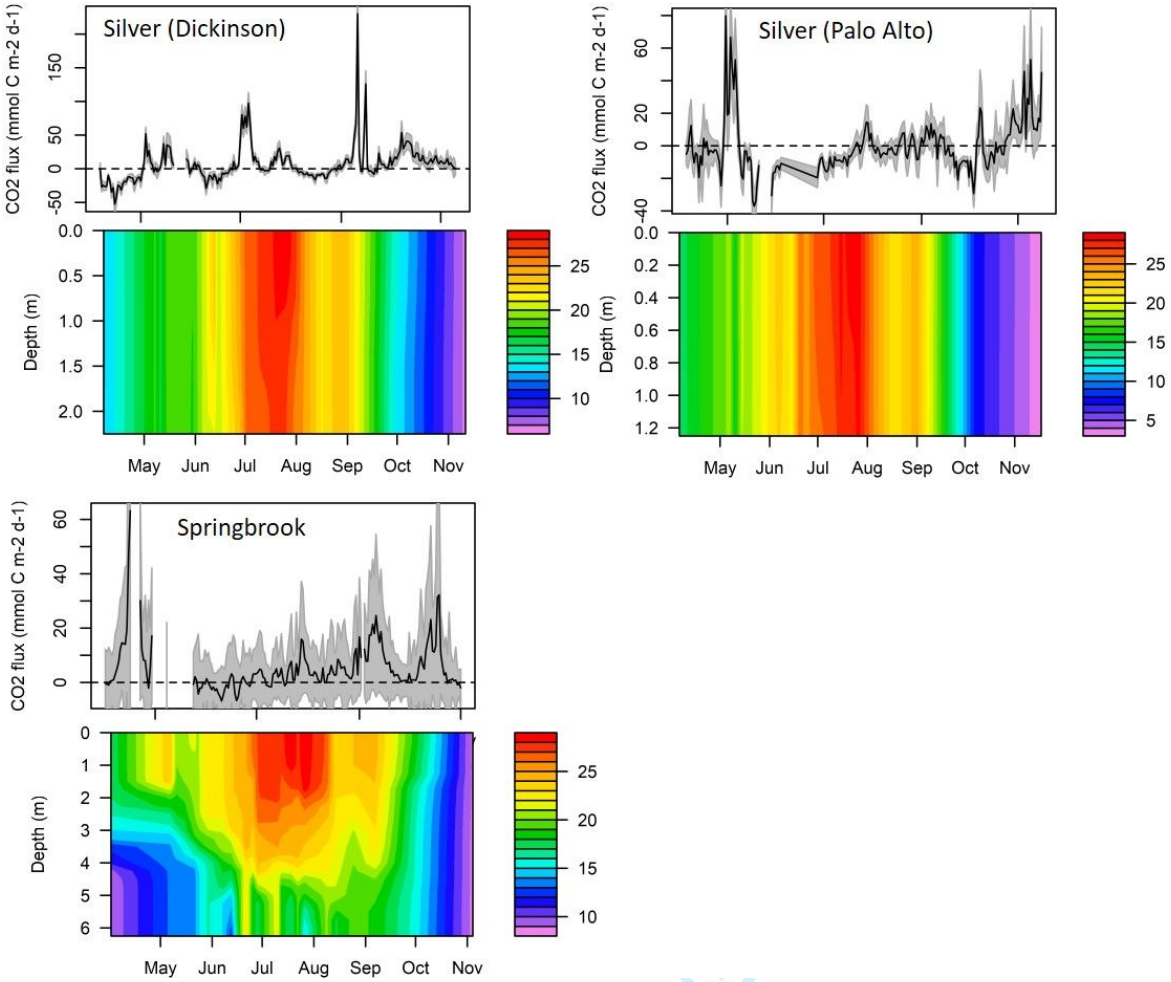
761



762

763

764



765

766

767

768

769

770

771

772

773

774

775

Figure 2

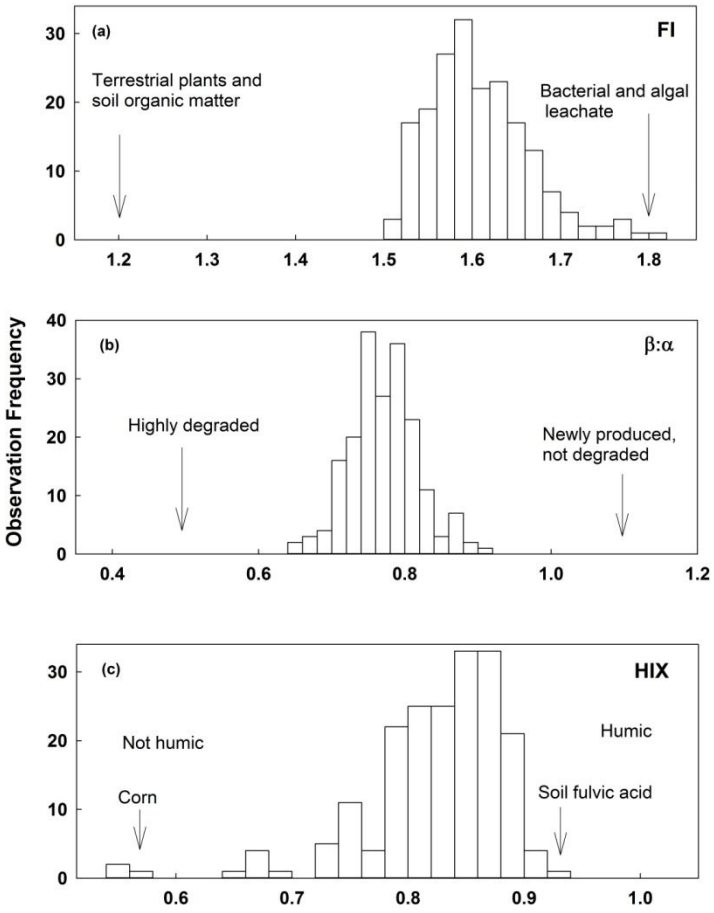
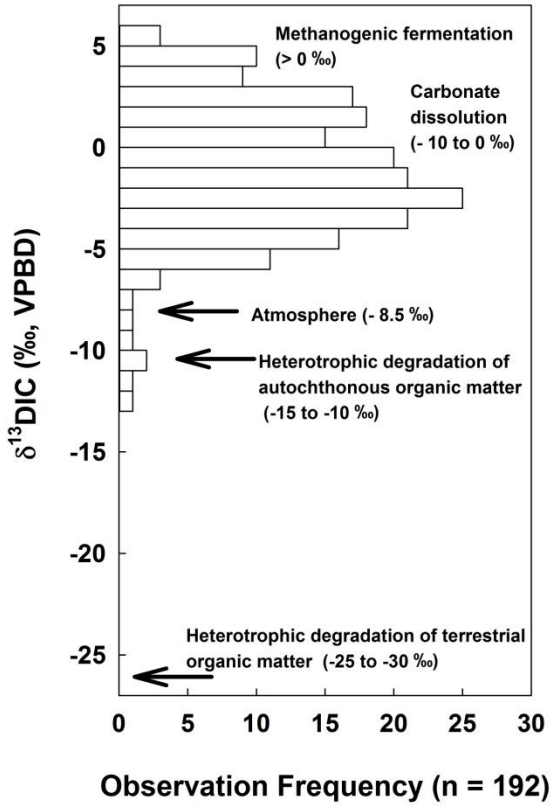
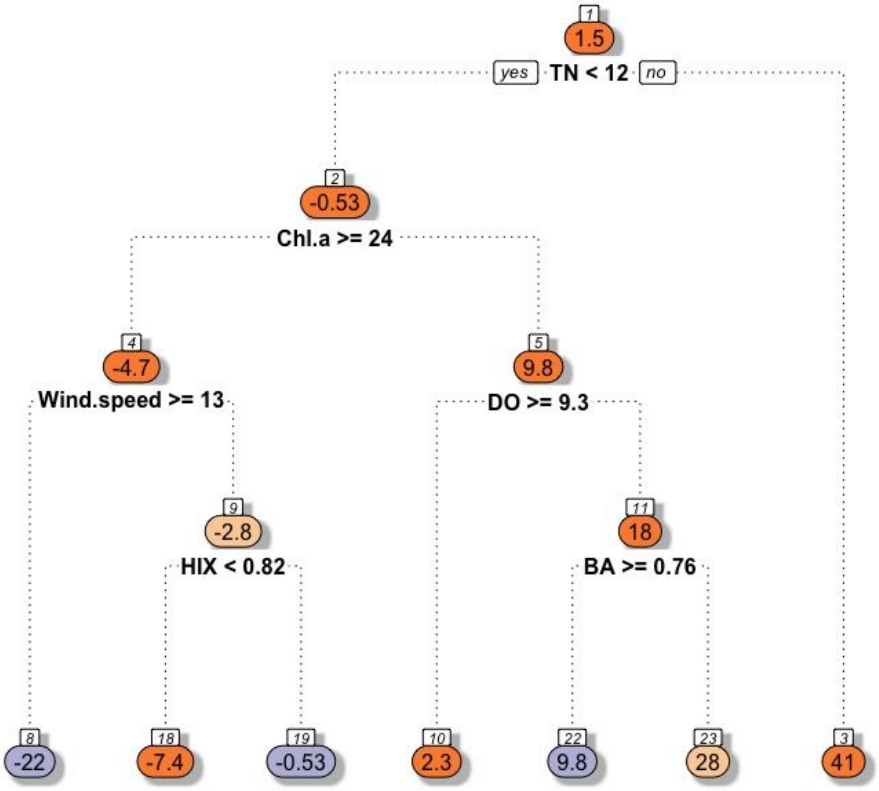

(a)
FI
Terrestrial plants and soil organic matter
Bacterial and algal leachate
30
20
10
0
1.2
1.3
1.4
1.5
1.6
1.7
1.8
(b)
β:α
Highly degraded
Newly produced, not degraded
Observation Frequency
40
30
20
10
0
0.4
0.6
0.8
1.0
1.2
(c)
HIX
Not humic
Humic
Corn
Soil fulvic acid
30
20
10
0
0.6
0.7
0.8
0.9
1.0

Figure 3



822 Figure 4

823



824

825

826

827

828

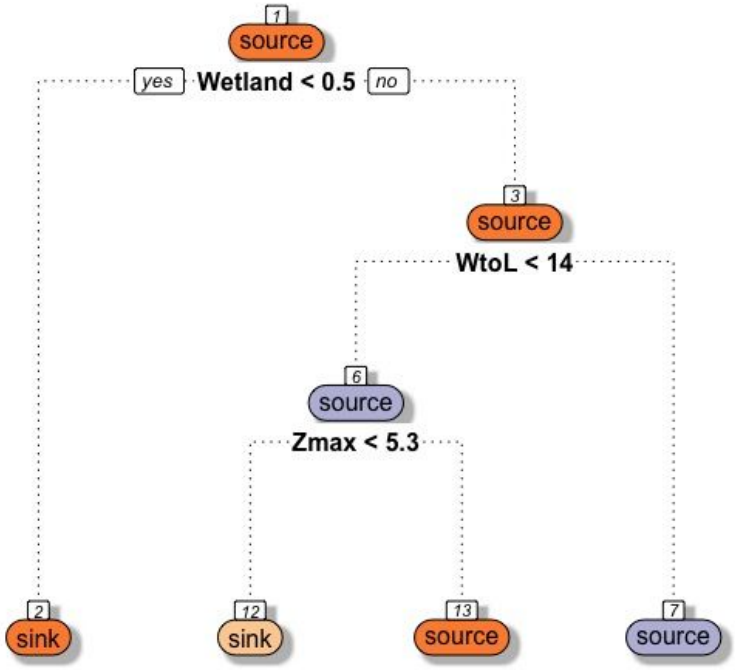
829

830

831

832

833 Figure 5

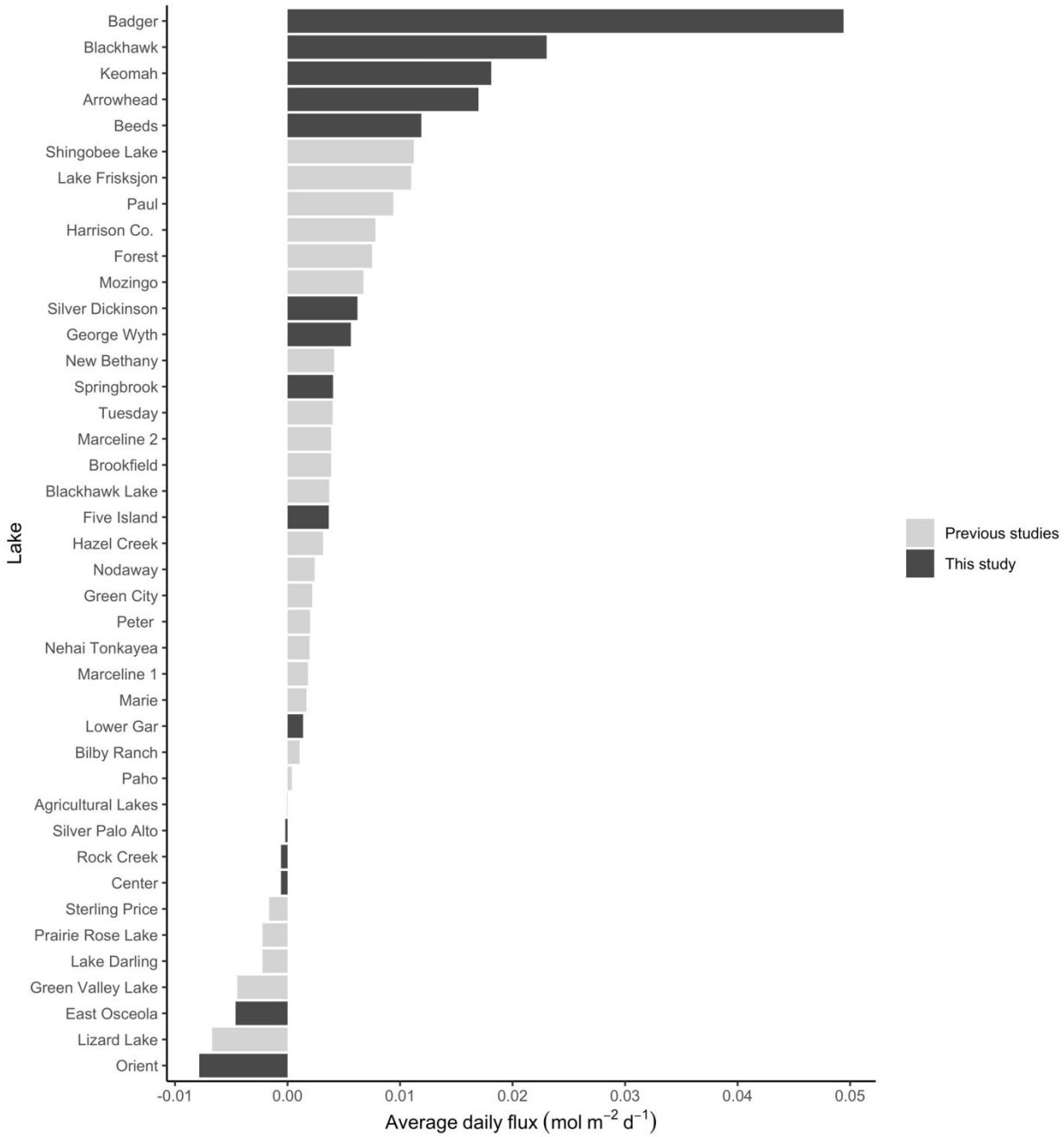


834

835

836
837
838
839
840
841
842
843
844
845
846
847

848 Figure 6.
849



850

Lake Orient, IA, 28 June 2012
photo credit: Ana Morales-Williams

1151x863mm (72 x 72 DPI)

Lake Keomah, Iowa boat launch, 14 August 2012
photo credit: Ana Morales-Williams

1151x863mm (72 x 72 DPI)