Contribution of gas concentration and transfer velocity to CO_2 flux variability in northern lakes

David Rudberg [1]*, **Jonathan Schenk** [1], **Gustav Pajala** [1], **Henrique Sawakuchi** [1], **Anna Sieczko** [1],
Ingrid Sundgren,[1] **Nguyen Thanh Duc,**[1] **Jan Karlsson** [2], **Sally MacIntyre,**[3,4] **John Melack** [3,4],
David Bastviken [1]

[1]Department of Thematic Studies—Environmental Change, Linköping University, Linköping, Sweden
[2]Climate Impacts Research Centre, Department of Ecology and Environmental Sciences, Umeå University, Umeå, Sweden
[3]Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, California, USA
[4]Earth Research Institute, University of California, Santa Barbara, California, USA

Abstract

The CO_2 flux (F_{CO_2}) from lakes to the atmosphere is a large component of the global carbon cycle and depends on the air–water CO_2 concentration gradient (ΔCO_2) and the gas transfer velocity (k). Both ΔCO_2 and k can vary on multiple timescales and understanding their contributions to F_{CO_2} is important for explaining variability in fluxes and developing optimal sampling designs. We measured F_{CO_2} and ΔCO_2 and derived k for one full ice-free period in 18 lakes using floating chambers and estimated the contributions of ΔCO_2 and k to F_{CO_2} variability. Generally, k contributed more than ΔCO_2 to short-term (1–9 d) F_{CO_2} variability. With increased temporal period, the contribution of k to F_{CO_2} variability decreased, and in some lakes resulted in ΔCO_2 contributing more than k to F_{CO_2} variability over the full ice-free period. Increased contribution of ΔCO_2 to F_{CO_2} variability over time occurred across all lakes but was most apparent in large-volume southern-boreal lakes and in deeper (> 2 m) parts of lakes, whereas k was linked to F_{CO_2} variability in shallow waters. Accordingly, knowing the variability of both k and ΔCO_2 over time and space is needed for accurate modeling of F_{CO_2} from these variables. We conclude that priority in F_{CO_2} assessments should be given to direct measurements of F_{CO_2} at multiple sites when possible, or otherwise from spatially distributed measurements of ΔCO_2 combined with k-models that incorporate spatial variability of lake thermal structure and meteorology.

Lakes occupy only a small portion of the Earth's surface but are estimated to emit large amounts of carbon dioxide (CO_2) to the atmosphere with recent emission estimates ranging from 0.3 to 0.6 Pg carbon (C) per year (Holgerson and Raymond 2016; Raymond et al. 2013). The CO_2 emissions (F_{CO_2}) from lakes depend on the gas transfer velocity (k) and differences in concentration of CO_2 in the surface water and its theoretical concentration when in equilibrium with the atmosphere (ΔCO_2):

$$F_{CO_2} = k \cdot \Delta CO_2 \qquad (1)$$

The variable k in Eq. (1) parameterizes the physical gas transport across the surface boundary layer and is modified by physical factors that influence turbulence, such as wind speed, direction and fetch, heating and heat loss, and currents (Banerjee and MacIntyre 2004; Read et al. 2012; Tedford et al. 2014; MacIntyre et al. 2021). In contrast, ΔCO_2 represents the balance between lake input of C, in-lake factors, and export via F_{CO_2}. Lake input of C comes from the surrounding catchment (Duarte and Prairie 2005; Maberly et al. 2013; Marcé et al. 2015) and is modified by catchment size and biogeohydrological properties and the magnitude and intensity of rainfall. In-lake factors constitute, for example, uptake, processing, and respiration of C by aquatic organisms (Lauster et al. 2006; Cole et al. 2007; Reis and Barbosa 2014) and are modified by light intensity, nutrient concentrations, and the standing stock of organisms.

*Correspondence: david.rudberg@liu.se

Additional Supporting Information may be found in the online version of this article.

The many factors involved makes it challenging to assess the relative influence of k and ΔCO_2 on F_{CO_2}, especially because the effect of such factors is manifested at different temporal scales ranging from hours or days (e.g., diurnal changes in winds, temperature, primary production, and respiration) to seasons or years (e.g., seasonal and annual variability in temperature, nutrient availability, and C input). Besides, k and ΔCO_2 may interact in contrasting ways, making it difficult to interpret their individual effect on F_{CO_2}. For example, during periods of stratification CO_2 may accumulate in the hypolimnion and the rate at which CO_2 can be transported to the surface depends on how effective winds and cooling are at deepening the upper mixed layer. These mixing-effects can range from being transient and associated with day–night patterns of wind and temperature (Jammet et al. 2017; Czikowsky et al. 2018) to being seasonal and linked to lake mixing in spring or autumn (López Bellido et al. 2009; Ducharme-Riel et al. 2015; Jansen et al. 2019).

In addition, temporal patterns in ΔCO_2 and k can vary from one location to another in a lake (Natchimuthu et al. 2017). For example, at shallower depths the time for CO_2 from the sediment to reach the surface can be shorter, and the sediment is more sensitive to resuspension, which can increase near-shore ΔCO_2 (Tonetta et al. 2016). In contrast, there can be lower ΔCO_2 at nearshore locations in lakes with uptake of CO_2 by primary production in the littoral zone (Schilder et al. 2013; Peixoto et al. 2016). Winds can also produce local patterns in turbulence and k (Read et al. 2012; Schilder et al. 2013; Vachon et al. 2013) and may tilt the thermocline, causing upwelling of CO_2 resulting in local increases in F_{CO_2} in both small lakes (Natchimuthu et al. 2017; MacIntyre et al. 2021; Rudberg et al. 2021) and larger ones (Heiskanen et al. 2014; Czikowsky et al. 2018). Alongside within-lake variability in ΔCO_2 and k, variability in lake depth, size and morphology, catchment hydrology (discharge and precipitation), and temperature and wind patterns can lead to between-lake variability of k and ΔCO_2.

Interaction effects between k and ΔCO_2 have been identified in river systems (Rocher-Ros et al. 2019), where turbulence is generally high. There, ΔCO_2 was regulated by emission rate and thus covaried with k in the sense that ΔCO_2 was low when k was high and vice versa. Although multiple studies have investigated the variability of either ΔCO_2 or k in lakes (e.g., Natchimuthu et al. 2017; Jansen et al. 2020; MacIntyre et al. 2021), studies comparing the relative influence of ΔCO_2 or k on F_{CO_2} variability at different temporal scales and across multiple lakes are rare. Yet, such knowledge is needed for fundamental understanding of gas fluxes in lakes and for improving process-based modeling of F_{CO_2} (Tan et al. 2017). It is also key for assessing when and where ΔCO_2 and k need to be measured to reliably assess F_{CO_2}.

Here we used floating chambers and surface water samples to measure F_{CO_2} and ΔCO_2, respectively, at deep and shallow locations for short-term (1–9 d) and monthly intervals in 18 lakes, in hemi-boreal, boreal, and sub-arctic regions in Sweden. We combined these measurements to derive spatially resolved values of k and identified: (i) changes in variability of F_{CO_2}, ΔCO_2, and k over time, (ii) the relative contributions of ΔCO_2 and k to F_{CO_2} variability for different time scales, and (iii) how differences in the relative contributions of ΔCO_2 and k to F_{CO_2} variability at both short-term (1–9 d) and seasonal time scales can be related to the characteristics of lakes and other environmental variables that affect processes within lakes. Based on our findings, we also discuss how to effectively assess the variability in F_{CO_2} across spatiotemporal scales in lakes.

Methods

Study lakes and sampling design

Eighteen lakes were studied between 2018 and 2020 (six lakes per year) (Supporting Information Table S1; Fig. 1). The lakes are distributed within six regions spanning sub-arctic to hemi-boreal environments in Sweden. Lakes were chosen to cover landscape characteristics and carbon and nutrient conditions representative of ranges observed in Swedish lakes (Supporting Information Fig. S1). Ranges of dissolved organic carbon (DOC), absorbance at 420 nm and total phosphorous (TP) were: 2–22 mg L^{-1}, 0.006–0.11 cm^{-1}, and 4–280 μg L^{-1}.

All lakes were visited monthly for week-long sampling campaigns during the full ice-free period for 1 yr. Sampling was conducted on 2–5 occasions during each week-long visit, with 24–48 h between each sampling. Samples of pCO$_{2aq}$ and pCO$_{2air}$ were collected to derive ΔCO_2 and combined with F_{CO_2} measurements to estimate k along three horizontal transects from shore toward the lake's center. Each transect consisted of 2–4 sampling locations that represented both shallow ($\sim$0–2 m) and deep ($>$2 m) parts of the respective lakes (Fig. 1). Typically, shallow and deep locations constituted near-shore and offshore locations, respectively, but for logistical reasons the most central part of the lakes were not always covered. In three of the shallowest lakes, Dammsjön (DAM), Lammen (LAM), and Grinnsjön (GRI), and in BD6, where sampling was limited to one of two bays, deep sampling was partly made at depths less than 2 m and reflects longer distance to shore rather than a great depth difference.

Floating chambers were deployed at each sampling location and were used to measure both F_{CO_2} and pCO$_{2aq}$. The chamber design was similar to that used in Natchimuthu et al. (2017) and each chamber consisted of an inverted plastic bucket (volume: 5.9–8.6 liters, diameter: 0.3 m), covered with aluminum tape to minimize internal heating by sunlight. The chamber had floats attached to its sides allowing the chamber walls to immerse approximately 3 cm into the water. The chambers were attached to another small float with a $\sim$1-m line, which in turn was connected to weights; this mooring allowed the chambers to move up and down with waves and water level changes. This type of chamber design yield similar flux estimates as other methods that do not interfere with the water surface (Cole et al. 2010; Gålfalk et al. 2013).
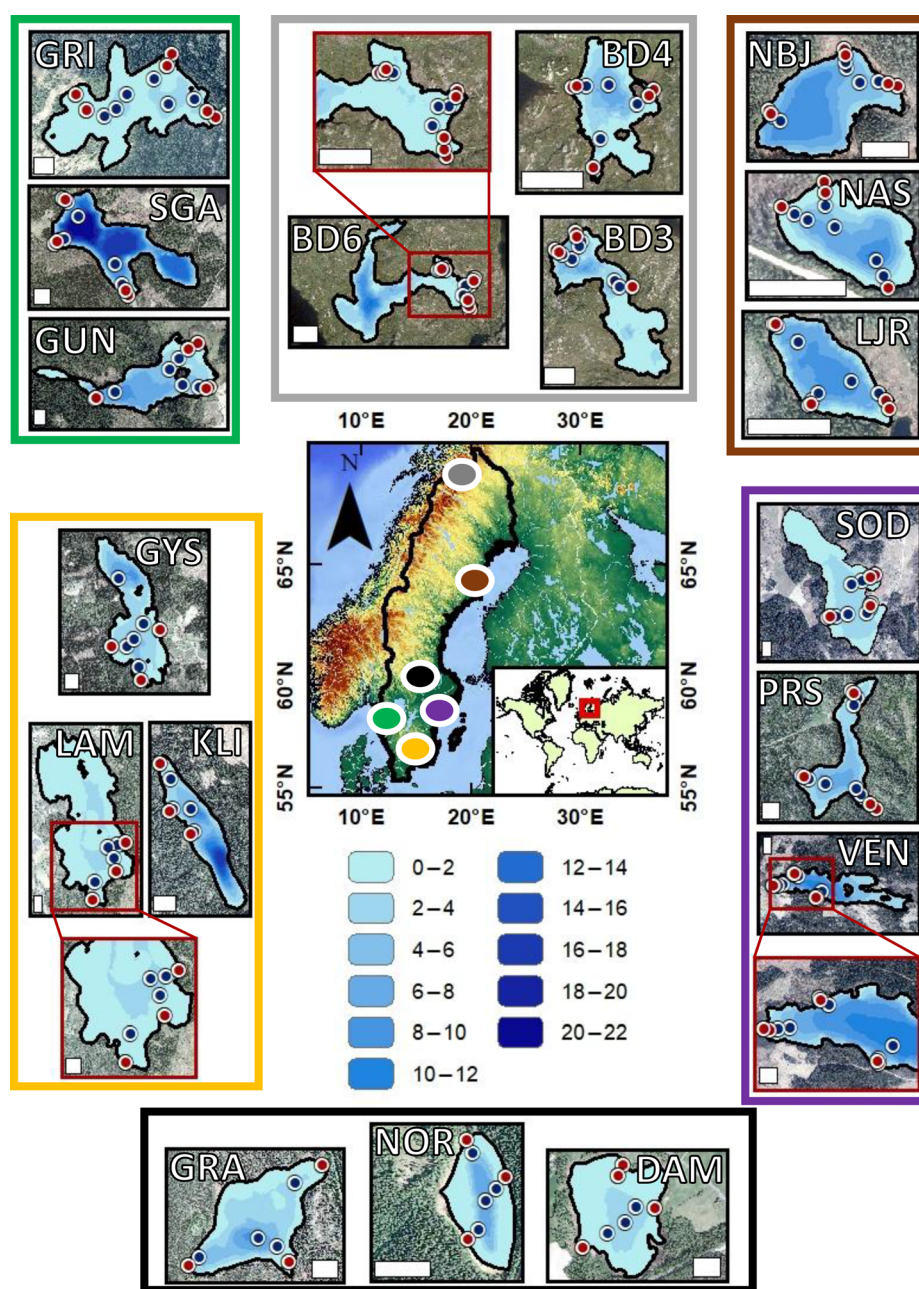
Fig. 1. Location of lakes and sampling locations of F_{CO_2} and pCO_{2aq} within the study lakes, with red and blue markers denoting shallow and deep sampling locations, respectively. Blue color toning represents depth intervals of 2 m, calculated in ArcGIS (3D Analyst, Natural Neighbor). The colors in the main map correspond to lakes within frames of similar color. Horizontal white scale bars have length of 100 m. Orthophotos were retrieved from Lantmäteriet, Sweden.

A 0.2- to 0.3-m polyurethane tube (inner and outer diameter of 3 and 5 mm, respectively) connected to a three-way Luer-lock valve was attached through the top of the chamber for manual collection of CO_2 from inside the chamber.

We retrieved air temperature and atmospheric pressure at 2 m height, wind speed at 10 m height, and precipitation from the meteorological reanalysis model MESAN developed by the Swedish Meteorological and Hydrological Institute, which interpolates measurements from nearby weather stations combined with a meteorological model to determine hourly means of meteorological variables on a 2.5 × 2.5 km grid (Häggmark et al. 2000). Although the MESAN model provides estimates for wind speed at 10 m height, it gives consistent ($R^2 = 0.78$–0.87) values of wind speed and wind gust speed compared to values measured at lake level (Rudberg et al. 2021).

Field measurements

Measurements of pCO$_{2aq}$ and pCO$_{2air}$

pCO$_{2aq}$ was measured with two separate methods; a primary method combining headspace- and dissolved inorganic carbon (DIC) samples (pCO$_{2HS}$) and dissociation constants from Wallin et al. ([2010](#)), used in 95% of cases and a secondary method involving measurements from equilibrated floating chambers (pCO$_{2FC}$), used in the remaining cases. Sampling of pCO$_{2aq}$ and pCO$_{2air}$ is only briefly described below and details are found in the Supporting Information (Supporting Information Text S2). Sampling and measuring DIC are described in Rudberg et al. ([2021](#)).

The pCO$_{2HS}$-method involved collecting water samples at ~ 0.1 m depth and then extracting them into the headspace of a syringe. We used a 140-mL syringe with 35 mL of air and equilibrated the sample for 2–3 min. Values of pCO$_{2aq}$ were converted from ppm (obtained from the gas analysis) to μatm, by accounting for changes in the carbonate system equilibrium in the vial headspace during sampling (Koschorreck et al. [2021](#)), as described in Rudberg et al. ([2021](#)). In cases where headspace equilibration samples were not collected at the specific sampling location but elsewhere in the sampling transect (21% of pCO$_{2HS}$ values), we instead calculated pCO$_{2aq}$ using headspace samples from the closest sampling location.

In cases where no headspace measurements were collected on the sampling transect, we estimated pCO$_{2aq}$ using the pCO$_{2FC}$-method. The pCO$_{2FC}$-method involved collecting gas samples from the headspace of floating chambers after deployment times of ~ 24 or 48 h, which is sufficient time for the chambers to fully equilibrate with CO$_{2aq}$ (Bastviken et al. [2015](#)). The chamber headspace was sampled with three 60-mL syringes and the gas was used to flush a 20-mL vial followed by injecting sufficient gas in the vial to create an overpressure. The sample integrity was confirmed when the overpressure in the vial was released prior to analysis in the lab (*see* below). Values of pCO$_{2aq}$ were converted from ppm to μatm by multiplying with the pressure during sampling. There is a k-dependent delay in capturing the real pCO$_{2aq}$ with floating chambers, which has been estimated to 1 and 3 h for wind speeds > 6 and < 1 m s^{-1}, respectively (Bastviken et al. [2015](#)). Such delay times imply a smoothing of short-term temporal variability of pCO$_{2aq}$. The bias in pCO$_{2FC}$ caused by this delay was investigated by comparing pCO$_{2HS}$ and pCO$_{2FC}$-samples that were collected within a few minutes of each other. A standardized major axis (SMA) linear regression between the two methods showed no consistent bias (slope: 1.04, $R^2 = 0.8$; Supporting Information Fig. S2), after removing outliers (7 of 790 values).

Sampling of pCO$_{2air}$ was made the same way as for the pCO$_{2FC}$-method by collecting air directly above the water surface. Assuming that variability of pCO$_{2air}$ was small within a transect, it was only sampled from the location nearest to the shore at each transect.

Values of pCO$_{2air}$ and pCO$_{2aq}$ were converted from ppm to μatm by multiplying with the pressure during sampling. ΔCO$_2$ (mmol m^{-3}) was derived by first subtracting transect location pCO$_{2air}$ from sampling location pCO$_{2aq}$ and then multiplying with Henry's constant (M atm^{-1}), calculated from surface water temperatures at the time of sampling according to Weiss ([1974](#)).

Measurements of F_{CO_2}

Each floating chamber was equipped with a CO$_2$ sensor (K33 ELG, SenseAir AB) to measure CO$_2$ concentration in the chamber headspace at 5- to 6-min intervals. The sensors also logged temperature and relative humidity. These CO$_2$ sensors have proved reliable under non-condensing conditions (Bastviken et al. [2015](#)). The CO$_2$ sensor was placed inside the chamber within a plastic box attached to the chamber top, to protect from splashes. The box was drilled with multiple 0.5- to 1-cm-diameter holes to facilitate gas transport. To protect the sensors against humidity, critical areas on the circuit boards were coated with polyurethane resin. Before each field measurement the sensors were cross calibrated for approximately 3 h by comparison to reference CO$_2$ measurements made with an Ultraportable Greenhouse Gas Analyzer (Los Gatos Research). Sensors whose measurements deviated from the reference were not used in field measurements. Following retrieval of data from CO$_2$ sensors, the data were filtered to account for humidity peaks, and the CO$_2$ concentration was corrected for temperature and humidity effects, as described in Bastviken et al. ([2015](#)). In nine cases, we identified nonlinear patterns in our sensor measurements associated with short-term instabilities and that sensor CO$_2$ increase at the time near deployment was not reflecting the long-term trend. Those cases were excluded from further analysis.

F_{CO_2} was calculated through an iterative procedure, accounting for measurements within a time window of 6–45 min from the chamber deployment. Measurements made within the first 6 min after deployment were discarded to minimize bias from incomplete mixing inside the chamber or possible turbulence caused by deploying the chamber. A slope representing CO$_2$ change over time was estimated for the first three measurements within the time window, and additional slopes were calculated after including one measurement point at a time for the whole time window. Among these slopes, the one with highest R^2 was used for calculation of F_{CO_2} according to Eq. ([3](#)). Cases when the highest R^2 of the slopes was < 0.9 were only considered if the root mean square error was below 10 ppm to not discriminate against small fluxes, as R^2 generally decreases when the surface water CO$_2$ concentration is close to equilibrium with the atmosphere. For the calculation of F_{CO_2} in Eq. ([2](#)), $\Delta f / \Delta t$ represents the change in mole fraction over time (ppm s^{-1}), P is atmospheric pressure (atm) at the time of sampling, V and A correspond to the volume (m^3) and area (m^2) of the chamber headspace, T is the air temperature (K) during sampling, derived from the chamber CO$_2$

sensor, R is the ideal gas constant ($\mathrm{L\,atm\,K^{-1}\,mol^{-1}}$) and $86.4 \cdot 10^3$ is the number of seconds per day, used to convert units from $\mathrm{s^{-1}}$ to $\mathrm{d^{-1}}$.

$$F_{\mathrm{CO_2}} = \frac{1}{10^6} \cdot \frac{\Delta f}{\Delta t} \cdot \frac{PV}{RTA} \cdot 86.4 \cdot 10^3 \qquad (2)$$

Values of $p\mathrm{CO_{2aq}}$, $p\mathrm{CO_{2air}}$ and associated chamber CO_2 sensor values allowed further assessment on data quality. We checked whether the concentration change measured by the CO_2 sensor (i.e., increase or decrease) over the first hour after deployment fully equilibrated with water ($p\mathrm{CO_{2aq}}$ minus $p\mathrm{CO_{2air}}$), which was seen as unrealistic as it normally takes $> 1\,\mathrm{h}$ for that to occur (Bastviken et al. 2015). To limit the effect of outliers, we discarded measurements where the above criterion was violated for at least three consecutive sensor measurements. Six hundred twenty of 1810 (34%) measurements did not pass this quality assessment and were excluded from analysis. Our motivation for this filtering approach and potential impacts on results are discussed in detail in Supporting Information Text S3. Briefly, 32% of the discarded measurements were made at such near-saturation conditions ($|p\mathrm{CO_{2aq}} - p\mathrm{CO_{2air}}| < 100\,\mu\mathrm{atm}$), when the uncertainty of k estimates is high (Supporting Information Fig. S3). An additional 14% of the discarded CO_2 sensor measurements used for estimating fluxes were made during reducing air temperatures when relative humidity neared 100%, conditions where condensation inside the chamber headspace can potentially bias measurements (but such conditions did not always lead to data discards). One possible explanation for the remaining ($\sim 50\%$) discards is that $p\mathrm{CO_{2aq}}$ derived from syringe sampling at $\sim 0.1\,\mathrm{m}$ depth ($p\mathrm{CO_{2HS}}$ method) was lower than the $p\mathrm{CO_{2aq}}$ in the surface water boundary layer, potentially due to microbial processes or photochemical reactions in the surface microlayer, enhancing CO_2 production (Conrad and Seiler 1988; Zhou and Mopper 1997; Cunliffe et al. 2011), which would contribute to underestimating $p\mathrm{CO_{2aq}}$ and result in overestimating k (Pajala et al. 2023), and in turn more rejected data.

Estimation of k from $F_{\mathrm{CO_2}}$ and ΔCO_2

Values of k were derived from sampling location and time-specific $F_{\mathrm{CO_2}}$ and ΔCO_2, by solving Eq. (1) for k. The direction of $F_{\mathrm{CO_2}}$ is out of the lake (i.e., positive) when ΔCO_2 is positive, and likewise $F_{\mathrm{CO_2}}$ is into the lake (i.e., negative) when ΔCO_2 is negative. Hence, k can only be positive. Nevertheless, 7% of our k-estimates were negative, which can happen near equilibrium (i.e., when ΔCO_2 is low) due to uncertainties in the measurements. These values were excluded from the analysis.

Additional lake chemistry and environmental variables

Values of pH in surface water were either derived indirectly by solving the carbonic acid equilibria equation (Stumm and Morgan 1996), using monthly measurements of DIC and $p\mathrm{CO_{2aq}}$ at $\sim 0.1\,\mathrm{m}$ depth at each chamber position as described in Rudberg et al. (2021), or measured directly at $\sim 0.1\,\mathrm{m}$ depth at the lake center with a pH electrode (HACH IntelliCAL PHC201). Measurements of surface water $p\mathrm{CO_{2aq}}$ is explained above. Surface water DIC was sampled with a syringe according to Rudberg et al. (2021). In short, 5 mL of water was injected to a 22-mL N_2-filled vial holding 100 μL concentrated phosphoric acid. The vial was kept at 20°C before the headspace gas of the vial was analyzed with gas chromatography (*see* below). DIC concentration in the water phase of the vial was derived by applying Henry's law to the concentration in the vial headspace. The in situ DIC concentration was calculated as the combined molar DIC amount in the water phase and headspace of the vial divided by the volume of the vial water phase.

Near-surface water (0.5–1 m depth) measurements of chlorophyll a (Chl a), total nitrogen (TN), TP, total organic carbon (TOC), DOC, and absorbance were made by sampling water with a Ruttner sampler at the deepest location of the lake. Part of this water was used to fill 125-mL acid-washed high-density polyethylene bottles that were transferred to a freezer upon arrival to the lab and kept at −18°C before being analyzed for TP. The remaining water was stored in a 4-L cubitainer until arriving at the lab the same day. At the lab, part of this water was transferred to 50-mL polypropylene tubes and stored at 4°C, to be analyzed for duplicate measurements of TN and TOC. Another part of the water was filtered using 47-mm Whatman GF/F glass fiber filters (nominal pore size 0.7 μm) with 100 mL of the filtrate transferred to two 50-mL polypropylene tubes and stored at 4°C to be analyzed for duplicate measurements of DOC and the remaining filtrate used for measurements of absorbance. The remaining water in the cubitainer was used for filtering with GF/C filters, upon which the filters were folded and wrapped in aluminum foil and stored at −18°C to be analyzed for Chl a. TP was analyzed with a segmented flow analyzer (AutoAnalyzer II, Seal Analytical) according to Murphy and Riley (1962). TN, TOC, and DOC were analyzed with a Total Organic Carbon Analyzer (TOC-V CSH, Shimadzu with a TN analysis unit), and certified standards containing 1, 2, and 5 $\mathrm{mg\,L^{-1}}$ TN and 10 $\mathrm{mg\,L^{-1}}$ TOC were used for calibration. Samples that we assumed had TOC concentrations higher than 20 $\mathrm{mg\,L^{-1}}$ were diluted with deionized water prior to analysis of TOC and DOC. Light absorbance at a 420 nm wavelength was measured with a spectrophotometer (Aqualog, Horiba). Chl a was extracted with pure methanol according to Swedish standard procedures (SS 28170) with a method based on Richards and Thompson (1952) and was analyzed with spectrophotometry.

Depth gradient measurements of temperature and light intensity in the photosynthetically active wavelength range (400–700 nm) were made at monthly intervals during the ice-free period at the lake center. Temperature measurements were made vertically at 0.5- to 2-m intervals with a multi-parameter probe (AP-5000 Aquaprobe or LD0101 IntelliCAL HACH

probe). Light intensity was measured with an underwater quantum sensor (LI-192SA LI-COR) every 10 cm down to 1 m depth and thereafter in 50-cm intervals. The slope of the regression between the logarithm of light intensity vs. depth was used to calculate light attenuation, according to Beer-Lamberts Law.

Dissolved oxygen (DO) saturation was measured at 0.1–0.25 m depth throughout the ice-free period using one centrally located PME miniDOT logger. Measurements were made at 10-min intervals or were recalculated to 10-min averages if higher measurement frequencies were used, in accordance with the sensor response time. Water temperature was obtained from the miniDOT loggers at 10-min intervals and at second intervals from thermistors (RBRsolo², RBR) located within each sampling transect at 0.1 m depth.

Gas analysis

Analysis of gas samples and of headspace gas in the DIC samples were made with an Agilent 7890A gas chromatograph (Agilent Technologies, equipped with a 1.8 m × 3.175 mm Porapak Q 80/100 column from Supelco, and a thermal conductivity detector) through automatic injection, with a 7697A headspace sampler (Agilent Technologies) attached. After confirming a linear and stable response by analyzing serially diluted samples from a certified high concentration standard (50,000 ± 1000 ppm) upon GC system evaluation, an independent certified standard of 1985 ± 40 ppm was used for calibration for each batch of samples being analyzed. Prior to analysis, vial overpressure was released by inserting a 0.5-mm hypodermic needle holding a water droplet making outflow of excess gas visible and forming a barrier to gas inflow. This setup made it possible to verify sample integrity during storage. Stopper CO_2 permeability was characterized by independent experiments and sample analyses were corrected to sample storage time.

Data analysis
Temporal variability in F_{CO_2}, ΔCO_2, and k

Interquartile range (IQR) was used for assessing temporal variability in F_{CO_2}, ΔCO_2, and k. By calculating IQR on datasets that represented different temporal scales in the respective lakes, we derived measures on lake-wise temporal variability. The temporal scales ranged from one monthly sampling campaign (1–9 d) and consecutive monthly intervals (i.e., multiple monthly sampling campaigns) to the full ice-free period. For each lake and temporal scale, IQR was calculated using all possible monthly combinations of data belonging to that time scale, to not favor certain monthly combinations over others. The combined mean IQR of these individual estimates was calculated and correspond to temporal variability for that lake and scale.

Variation in F_{CO_2} explained by ΔCO_2 or k

We used dominance analysis to estimate the relative influence of ΔCO_2 and k on F_{CO_2} variability (Budescu 1993; Azen and Budescu 2003). In dominance analysis, relative influence of predictor variables is derived by first calculating the additional contributions that each predictor variable provides to the model R^2 across all possible linear subset models (in our case two subset models, one with only ΔCO_2 included and one with only k included as the predictor variable). The additional contribution of each predictor is calculated as the increase in R^2 that results from adding that predictor to the regression model, and results in an output between 0 and 1, where higher output indicates greater relative influence of that predictor variable to the response variable.

We based our model used for dominance analysis on the logarithm of Eq. (1), satisfying the requirement of linear relationships between predictor variables, $\log(\Delta CO_2)$ and $\log(k)$, and the dependent variable, $\log(F_{CO_2})$. We used absolute values of F_{CO_2}, ΔCO_2, and k in dominance analysis because logarithms cannot be calculated on negative values. It was also preferrable to use absolute values to not bias results in cases of negative F_{CO_2} (otherwise, k will be positive when F_{CO_2} is negative, and the relationship between F_{CO_2} and k will be non-linear if positive and negative F_{CO_2} are used in the same regression; Supporting Information Fig. S4). The model is displayed in Eq. (3).

$$\log|F_{CO_2}| = \log|k| + \log|\Delta CO_2| \qquad (3)$$

Due to the small number of measurements with negative F_{CO_2}, we did not conduct separate dominance analyses for sets of values associated with negative and positive F_{CO_2}, respectively. Instead, all values were combined in the same analysis. By doing so, we presume that k and ΔCO_2 influence F_{CO_2} in similar ways when the flux is directed into (4% of cases) and out of (96% of cases) the lake, respectively.

Influence indexes were derived at various temporal scales by applying dominance analysis on separate subsets of data that represented measurements of F_{CO_2}, ΔCO_2, and k from different temporal scales. The method was designed as an iterative Monte-Carlo procedure. The temporal scales were the same as those used for estimating temporal variability in our data (explained above). Similar to that method, individual estimates of influence indexes were calculated for each time scale, consisting of different monthly combinations of data to not favor certain monthly combinations over others (i.e., for a temporal scale of 2 months, one set included data from March and April, another April and May, and so forth). The method followed the general procedure: (1) The first month in the dataset was set as the starting month. (2) Data belonging to the starting month was added to data for consecutive month(s), where the total of months compiled depended on the monthly scale considered. (3) We identified the smallest number (n_{min}) of observations for any of the months included

in the data. (4) We extracted a random subset of n_{min} unique samples from each month of data, to ensure similar weights for months with unequal sample sizes. (5) We calculated influence indexes of ΔCO_2 and k to F_{CO_2} variability for the subset of samples with dominance analysis; 6. Steps 4 and 5 were repeated 100 times. (7) We calculated the respective mean relative influence of ΔCO_2 and k to F_{CO_2} variability from the iterations. (8) Steps 1–7 were repeated but using different starting months as reference. (9) After separate mean relative influences had been estimated accordingly for all monthly combinations that represented the same time scale, the collective mean of these values was calculated and correspond to *lake mean relative influence* (here denoted I) for that specific time scale.

In the text, we refer to *lake mean relative influence* of ΔCO_2 and k on F_{CO_2} variability at a given temporal scale as I_{CO_2} and I_k, respectively. The sum of I_{CO_2} and I_k is always 1, which means that $I_{CO_2} > 0.5$ indicates greatest influence of ΔCO_2 on F_{CO_2}. Likewise, $I_{CO_2} < 0.5$ indicates greatest influence of k on F_{CO_2}. We refer to *short-term relative influence* as the shortest temporal scale included here (one monthly sampling campaign; 1–9 sampling days) and refer to *long-term relative influence* as covering the full ice-free period in our lakes.

We calculated I_{CO_2} and I_k for all sampling depths combined in each lake. We also investigated spatial within-lake variability by comparing results for shallow and deep sampling locations, according to definitions of "shallow" and "deep" mentioned earlier and displayed in Fig. 1. Due to limited sampling at deep locations in BD3 and BD6 (Supporting Information Fig. S5), those lakes were excluded from analyses concerning I_{CO_2} and I_k at deep sampling locations.

Comparison with lake characteristics

To explore how much short-term and long-term I_{CO_2} could be linked to lake characteristics, univariable linear, logarithmic, and exponential ordinary least squares (OLS) regressions as well as multivariable linear OLS regressions of I_{CO_2} (one value per lake) were made with a range of variables (Table 1). In multivariable regressions, the potential predictor variables listed in Table 1 were used in combinations of two at a time. Shallow and deep location I_{CO_2} were statistically different ($p < 0.05$, t-test) at long-term scales among the lakes, hence regressions were conducted on shallow and deep locations, respectively. At short-term scales, shallow and deep location I_{CO_2} were not statistically different ($p = 0.62$, Kruskal–Wallis test) and were thus combined in the regression analysis.

All lake water and weather variables were compiled by first averaging values by sampling occasion and then calculating a total average for the respective lakes. Values of wind speed, averaged by both sampling occasion and the full year, were analyzed. Lake water and meteorological weather variables thus represent mean values for the whole lake in all regressions (i.e., in regressions with whole-lake, shallow, and deep

Table 1. Variables used in OLS regression analysis.

Lake and catchment properties[a]	Lake water and weather variables
Catchment area, mean catchment elevation slope, lake area, lake to catchment area ratio, lake volume, lake water residence time, latitude, altitude, mean and maximum lake depth, percent catchment coverage of forest, farmland, clearcut and wetland, respectively, duration of ice-free period, fraction of ice-free period that lakes were stratified	ΔCO_2, light absorbance (420 nm), light attenuation (400–700 nm), Chl a, pH, DOC, TOC, DIC, TN, TP, DO saturation, water temperature, air temperature, precipitation, wind speed

[a]See Supporting Information Text S1 for more information on how variables were calculated.

I_{CO_2}). Variable values for the respective lakes are listed in Supporting Information Table S2.

Model stability was evaluated by repeating the regression analysis but leaving one lake out each time. In cases where model significance was determined by the inclusion of one lake alone, the model was deemed unstable and was excluded from analysis. Furthermore, predictor variables included in multiple regressions were examined for collinearity by calculating variance inflation factors (VIF), which is a measure of the maximum correlation between the respective predictor variables. Significant regressions were only accepted if VIF was < 5 for the predictor variables (Sheather 2009). Finally, models were ranked based on adjusted R^2 (calculated according to Eq. 4).

Statistical analysis

Probability values below 0.05 have been considered statistically significant to reject null-hypotheses in statistical tests and regressions. Reported R^2 values are adjusted to the number of observations (n) and predictors (p) in the model (adjusted R^2):

$$\text{Adjusted } R^2 = 1 - (1 - R^2)\frac{n-1}{n-p} \quad (4)$$

In boxplots, boxes indicate the distribution of data, where the box itself represents the range of the middle 50% of data (IQR), the horizontal lines in each box show the median value, the whiskers above and below boxes show the first and last quartile of data within $\pm 1.5 \times IQR$, and points represent outliers ($> |1.5 \times IQR|$).

SMA regression was done in R Studio v.1.1.456, using package *smatr* (Warton et al. 2012). All other statistical analyses, data handling, and plotting were made in Python v.3.7, using

packages *scipy*, version 1.9.1 (Virtanen et al. 2020), *statsmodels*, version 0.13.2 (Seabold and Perktold 2010) and *dominance_analysis*, version 1.1.8 (Azen and Budescu 2003; https://dominance-analysis.github.io/dominance-analysis). Normality tests for residuals were made with *scipy.stats.shapiro*. Pearson (for normally distributed residuals) and Spearman (for non-normally distributed residuals) correlation tests were made with *scipy.stats.pearsonr* and *scipy.stats.spearmanr*, respectively. Tests for differences in means between groups with normal distribution were made with *t*-tests, using *scipy.stats.ttest_ind*, and applying Welch's correction in cases where variances were unequal. Tests for equal variance between groups were made with Bartlett's test, using *scipy.stats.bartlett*. For cases where normality could not be achieved with or without log-transformation, differences in means between groups were assessed with the Kruskal–Wallis test, using *scipy.stats.kruskal*. OLS regressions were calculated with *statsmodels.api.formula.ols*. VIF were calculated with *statsmodels.stats.outliers_influence.variance_inflation_factor*. Dominance analysis was made using *dominance_analysis.Dominance*.

Results

Magnitude and variability of F_{CO_2}, ΔCO_2, and *k*

Mean ice-free F_{CO_2}, ΔCO_2, and *k* ranged between 2 and 73 mmol m^{-2} d^{-1}, 4 and 78 mmol m^{-3}, and 0.8 and 3.2 m d^{-1}, respectively, for the different lakes (Table 2; Fig. 2). F_{CO_2} was predominantly positive in our lakes (i.e., flux out of the lake) but was negative (i.e., flux into the lake) for 3%, 6%, 16%, and 38% of measurements in oligotrophic BD6 and our three eutrophic lakes, GUN, VEN, and SOD. Mean F_{CO_2} and ΔCO_2 corresponded well ($r = 0.81$) and were linked to light absorbance at 420 nm ($r = 0.74$ and 0.79) which in turn was linked to DOC ($r = 0.9$). No clear trend was identified between F_{CO_2} and *k* (Fig. 2). Variability in ΔCO_2 increased from temporal scales of 1–9 d up to the full ice-free period (10 months or less). In contrast, variability in *k* was reduced or leveled out at scales > 5 months in most lakes (Fig. 3). Variability in F_{CO_2} differed between lakes. In lakes VEN, GUN, LAM, and GRI, F_{CO_2} variability increased at lower time scales but was reduced at scales > 6–8 months. In most other lakes, F_{CO_2} variability either increased or was leveled out over time.

Influence of ΔCO_2 and *k* on F_{CO_2} variability

The lake mean relative influence of ΔCO_2 on F_{CO_2} variability was generally below 0.5 at temporal scales between 1–9 d and 6 months, meaning that *k* had most influence on F_{CO_2} variability at those temporal scales. However, I_{CO_2} in the lakes studied here increased with greater temporal scale and at scales longer than 7 months average lake I_{CO_2} was instead slightly above 0.5 (i.e., ΔCO_2 had most influence on F_{CO_2}). Despite this general trend there was large between-lake

Table 2. General information including duration of sampling campaigns, total number of sampling campaigns ($n_{campaigns}$) and observations (n_{obs}) and mean, standard deviation (SD), and minimum–maximum range of F_{CO_2}, ΔCO_2, and *k* for the respective lakes. Statistics are calculated as collective sampling month means (i.e., by first averaging for each monthly sampling campaign and then calculating a combined average of those). For calculation of SD, only sampling months with > 3 measurements have been considered.

Lake	Sampling period	$n_{campaigns}$	n_{obs}	F_{CO_2} (mmol m^{-2} d^{-1})			ΔCO_2 (mmol m^{-3})			*k* (m d^{-1})		
				Mean	SD	Range	Mean	SD	Range	Mean	SD	Range
BD3	17 Jul to 10 Sep	3	27	18.6	9.8	1–49	6.0	2.4	2–19	3.21	1.41	0.2–7.3
BD4	25 Jun to 11 Sep	4	50	27.8	19.0	0.1–90	18.5	10.9	3–48	1.82	1.25	0.01–5.5
BD6	25 Jun to 10 Sep	4	30	19.1	6.2	-2–62	8.7	5.1	-5–21	2.09	0.75	0.3–4.4
NAS	07 Jun to 30 Sep	5	72	34.1	24.9	2–151	26.2	4.1	10–61	1.41	0.89	0.1–5.5
NBJ	07 Jun to 24 Oct	6	102	57.8	21.9	4–141	78.2	17.4	31–143	0.78	0.32	0–1.9
LJR	05 Jun to 30 Sep	5	73	31.1	15.4	6–117	24.9	2	11–73	1.55	0.75	0.3–5.1
NOR	04 Apr to 04 Nov	8	62	52.3	21.5	2–150	37.8	5	2–84	1.71	0.72	0.2–8.5
GRA	06 Apr to 04 Nov	8	44	49	15.5	10–131	26.0	4.9	12–57	2.04	0.60	0.3–3.9
DAM	03 Apr to 04 Nov	8	64	72.1	31.1	3–315	43.7	9.7	8–82	1.67	0.67	0.2–3.9
VEN	06 May to 03 Dec	10	103	32.4	26.6	−27–181	29.3	14	−12–107	1.5	1.1	0.1–10.5
SOD	07 May to 07 Nov	8	52	1.8	8.7	−26–45	4.4	8.2	−15–50	1.44	0.62	0.1–3.7
PRS	06 May to 07 Nov	9	125	23.3	14.1	3–128	26.1	10.5	3–85	1.21	0.77	0.1–5.0
GUN	13 Apr to 29 Nov	9	81	50.7	23.7	−15–200	37.7	6.7	−10–86	1.51	0.67	0.2–4.2
GRI	13 Apr to 29 Nov	9	101	39.4	23.9	3–171	41.0	9.6	6–138	1.27	0.78	0.1–4.7
SGA	13 Apr to 29 Nov	9	69	28.2	16.7	1–139	20.8	9.3	1–108	1.75	0.88	0.2–5.3
LAM	20 Apr to 18 Nov	8	39	46.1	21.1	4–125	24.6	12.4	4–84	2.32	1.26	0.5–8.2
KLI	20 Apr to 18 Nov	6	35	30.8	14.9	2–98	15.7	5	3–66	2.23	1.08	0.3–5.7
GYS	20 Apr to 18 Nov	8	53	38.7	24.1	3–144	21.2	6.1	2–72	1.95	1.01	0.3–5.8
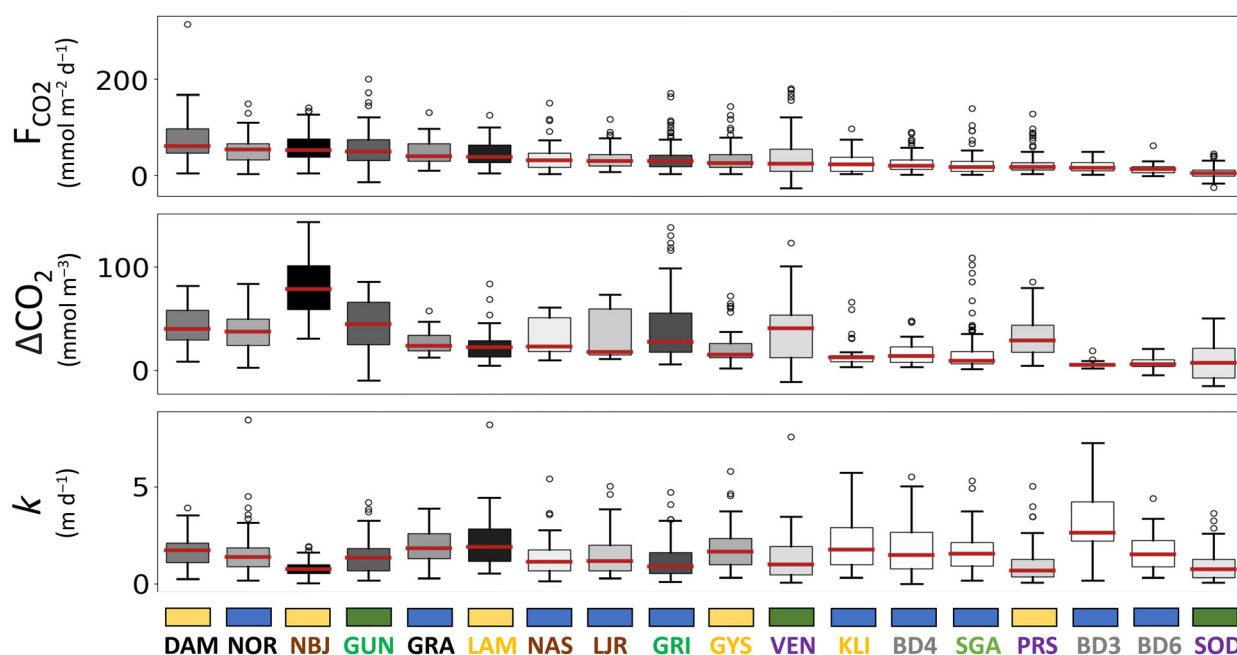
Fig. 2. Magnitude and variability of F_{CO_2}, ΔCO_2, and k for all study lakes. Lakes are sorted by median F_{CO_2}, displayed with red lines. Darker shading of boxes in the boxplot represents greater light absorbance measured at 420 nm. Colors of lake names in the horizontal axis are connected to similarly colored lake regions in Fig. 1. Colors above lake names denote trophic state (blue: oligotrophic, yellow: mesotrophic, green: eutrophic) based on TP, according to Carlson (1977).
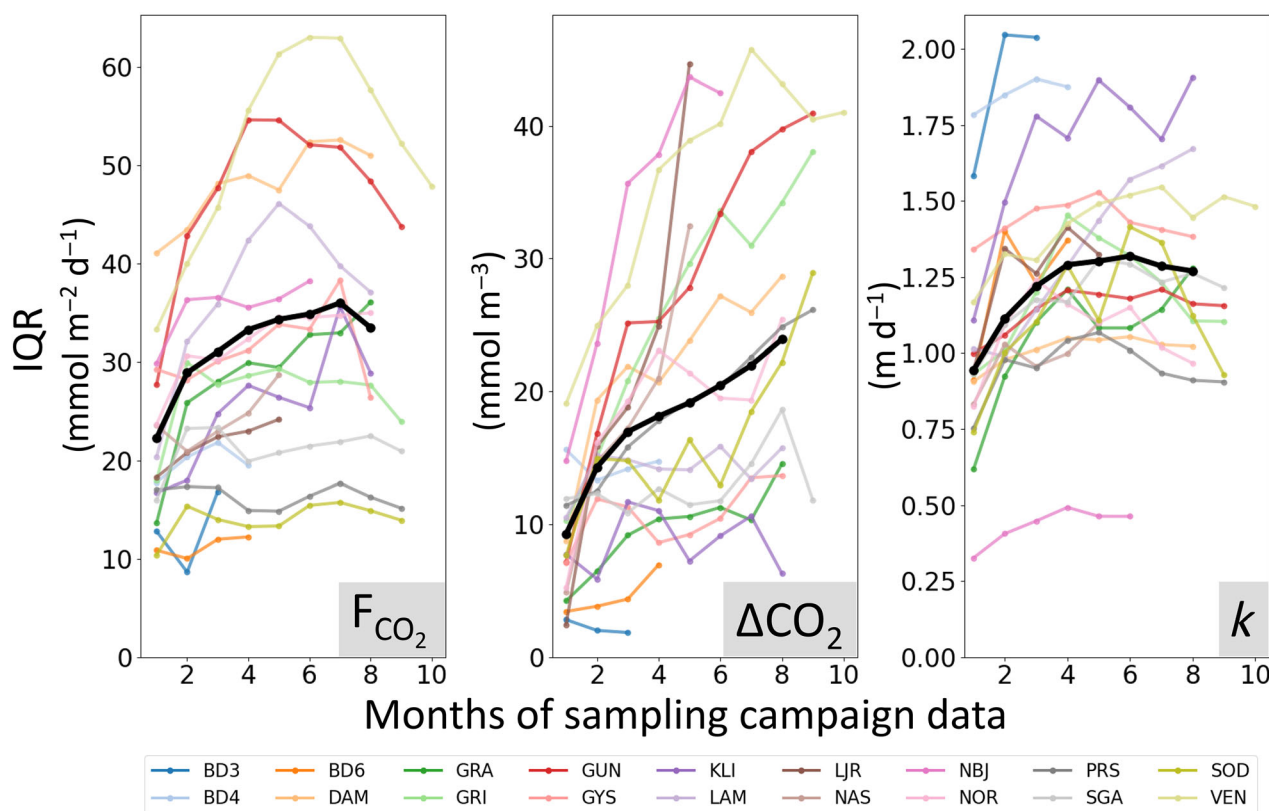
Fig. 3. Temporal variability of F_{CO_2}, ΔCO_2, and k, represented by the IQR of measured values aggregated by periods of 1–10 monthly sampling campaigns. The black lines show the average across lakes with at least eight consecutive months of data.

variability, with differences in I ranging between 0.3 and 0.5 depending on temporal scale.

The change in I with temporal scale (increase for I_{CO_2} and decrease for I_k) was generally greatest at deeper sampling locations (Fig. 4) and when only these deep sampling locations were considered, I_{CO_2} had a clear positive trend with increased timescale in all lakes except in BD3, BD6, SOD, and GRA (Supporting Information Fig. S5). Accordingly, I_{CO_2} at deep sampling locations was significantly greater than at shallow locations when data for the full open water period in the lakes were used ($p < 0.05$).

Short-term regulation of ΔCO_2 and k on F_{CO_2} variability

For short-time periods (1–9 d) catchment slope, catchment fraction of open field, and lake area were positively related to relative influence of ΔCO_2 on F_{CO_2} variability and combinations of those variables produced the best models ($R^2 = 0.54$–0.73; Fig. 5a,b; all equations and statistics are shown in Supporting Information Table S3). Of these three variables, catchment slope and fraction of open field were correlated to wind speed ($r = 0.72$ and 0.84).

Other significant models ($R^2 = 0.37$–0.48) combined either catchment slope, catchment fraction of open field, or lake area, with lake physiochemical variables (TOC, absorbance, ΔCO_2, Chl a, DIC, DOC, TP, and light attenuation). Similar degrees of explanation were identified when catchment fraction of open field was combined with either latitude or air temperature ($R^2 = 0.4$–0.42), and when lake area was combined with wind speed ($R^2 = 0.42$ and 0.41 using mean values

from sampling occasions and the full year, respectively; Fig. 5c).

Ice-free period regulation of ΔCO_2 and k on F_{CO_2} variability

No significant relations between long-term I_{CO_2} and driver variables were found for shallow sampling locations. When deep sampling locations were considered, yearly precipitation, lake volume, and ice-free period duration had positive relationships with long-term I_{CO_2} on their own ($R^2 = 0.39$–0.5; all equations and statistics are shown in Supporting Information Table S4). The degree of explanation increased when two of these variables were combined together ($R^2 = 0.61$–0.72; Fig. 6a,c), when precipitation was combined with catchment area or lake-to-catchment area ratio ($R^2 = 0.7$–0.72) or when ice-free period duration was combined with mean depth ($R^2 = 0.64$, Fig. 6b). Lake volume was also significant in regressions with DIC, TP or Chl a, but those multiple regressions mostly explained deviation in lake SOD, which had much higher values of DIC, TP, and Chl a than other lakes.

Discussion

Our study of variability in F_{CO_2} shows that variable influence changes depending on time scale. Over shorter time scales, k generally had greater influence than ΔCO_2 on F_{CO_2} variability. On longer time scales, the relative influence of ΔCO_2 for F_{CO_2} variability increased and became the dominant factor for F_{CO_2} variability over deep locations, which was in line with the seasonal pattern of ΔCO_2 observed in most of
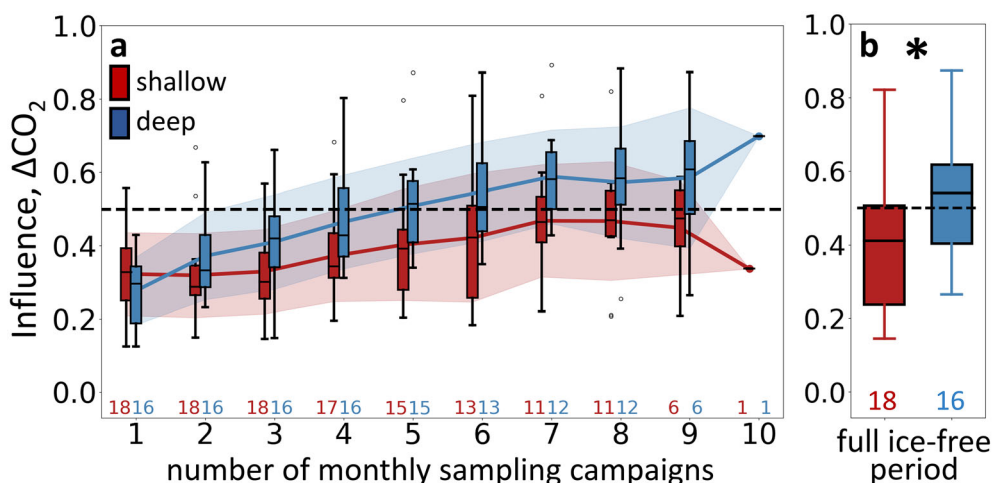
Fig. 4. Distribution of lake mean relative influence of ΔCO_2 on F_{CO_2} variability (I_{CO_2}) at different temporal scales—(**a**) 1–10 sampling campaigns and (**b**) full ice-free periods—for shallow (red) and deep (blue) sampling locations in our study lakes. Lakes BD3 and BD6 are excluded in the analysis for deep locations due to limited sampling there. Boxes show the distribution of I_{CO_2} values from the different lakes (each box with its horizontal lines, whiskers, and points represents maximum one I_{CO_2} value per lake) and colored areas cover values within one standard deviation from the mean. Colored numbers below each boxplot pair denote the number of lakes included in the assessments. The dashed horizontal line indicates the separation between greater influence of ΔCO_2 on F_{CO_2} variability (above the line; influence of $\Delta CO_2 > 0.5$) and greater influence of k on F_{CO_2} variability (below the line; influence of $\Delta CO_2 < 0.5$). Individual estimates of relative influence used for shallow and deep sampling locations in the lakes are displayed in Supporting Information Fig. S5. Asterisk in (**b**) imply significant differences for the relative influence of ΔCO_2 on F_{CO_2} variability (and likewise significant differences for the relative influence of k on F_{CO_2}) at shallow and deep sampling locations.
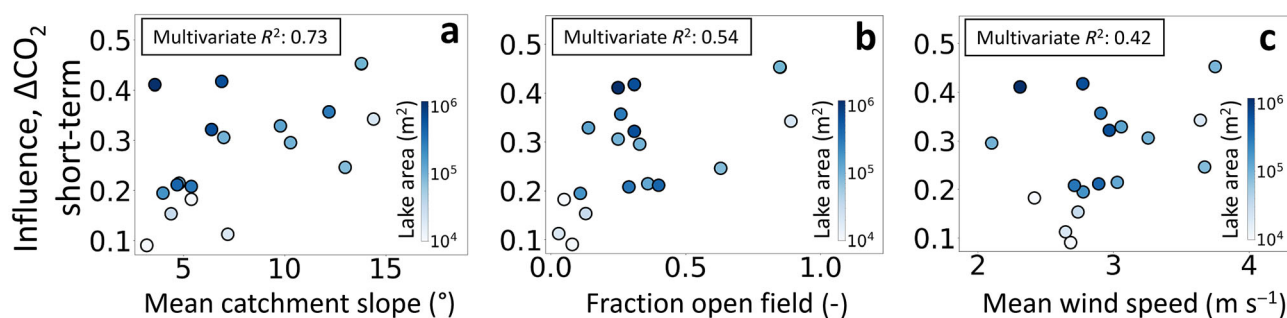
Fig. 5. Multivariate regressions for short-term (1–9 d) lake mean relative influence of ΔCO_2 on F_{CO_2} variability (I_{CO_2}), using data from all sampling depths, together with lake area and (**a**) mean catchment slope, (**b**) catchment fraction of open field, and (**c**) sampling mean wind speed. In each regression, the variable lake area is displayed in colors ranging from low (white) to high (blue) values on a logarithmic scale.

our lakes (Supporting Information Fig. S7) and the greater variability observed over longer time scales for ΔCO_2 compared to *k* (Fig. 3). Nevertheless, *k* was still the dominant factor for F_{CO_2} variability at shallow near-shore locations even on longer time scales. These time-scale differences can be generalized to ΔCO_2 contributing more than *k* to F_{CO_2} variability on longer time scales over deep locations, and *k* contributing more than ΔCO_2 to F_{CO_2} variability at shallow near-shore locations regardless of temporal scale (Fig. 7). Natchimuthu et al. (2017) found that ΔCO_2 contributed more than *k* to F_{CO_2} variability over weekly scales, whereas *k* contributed more at hourly scales. However, such findings were reported for one lake only and the influence of ΔCO_2 on F_{CO_2} variability exceeded influence of *k* on F_{CO_2} variability in only some of the 18 lakes we studied. Thus, the relative importance of *k* and ΔCO_2 for F_{CO_2} differ over time and space in lakes and among lakes.

Short-term influence on F_{CO_2} variability across lakes

For short-time periods (1–9 d) catchment slope, catchment fraction of open field, and lake area were most relevant for relative influence of ΔCO_2 on F_{CO_2} variability. Of these three variables, catchment slope and fraction of open field were correlated to wind speed. The third variable, lake area, is in
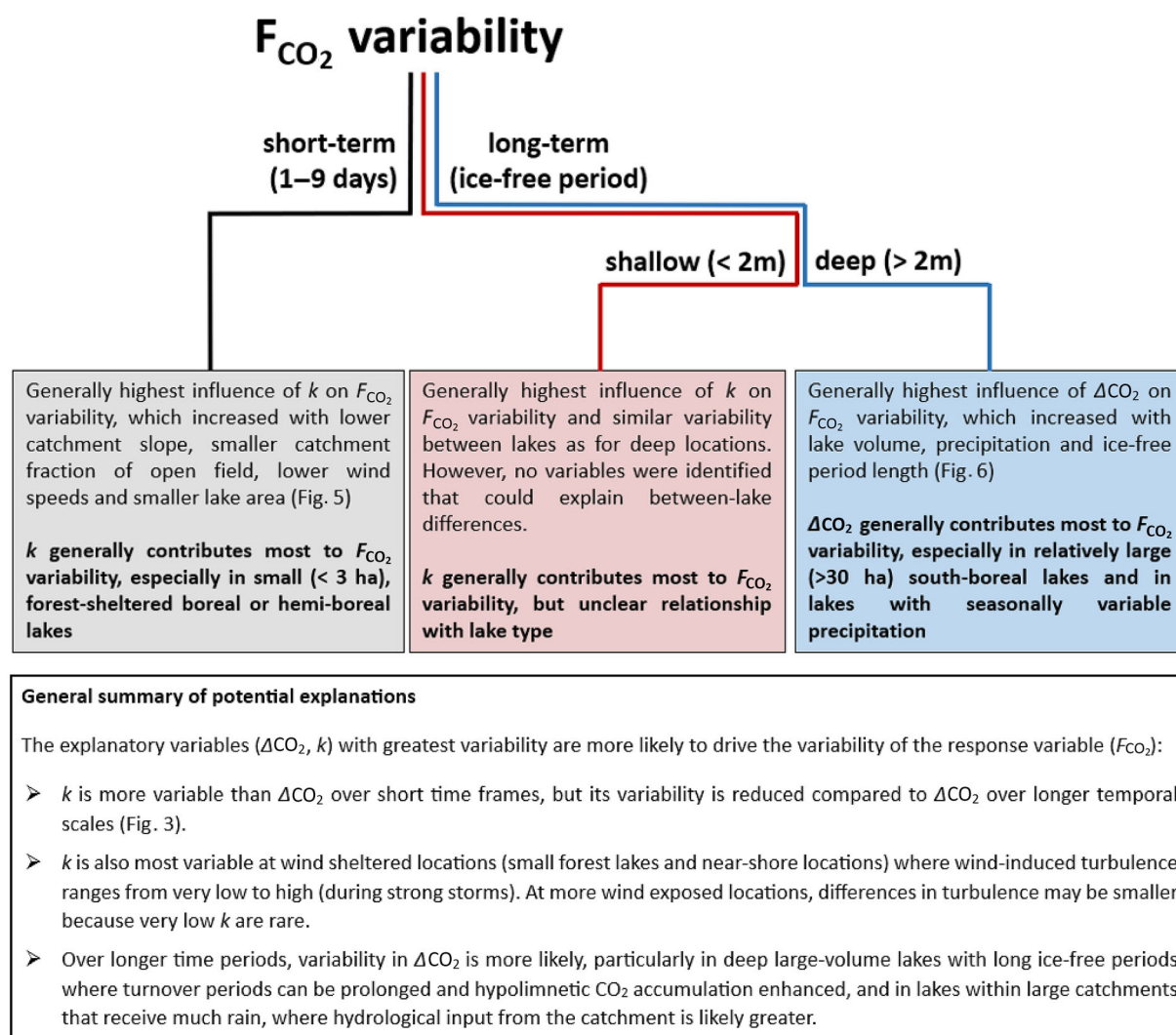
turn important for wind fetch (Read et al. 2012; Vachon et al. 2013; Klaus and Vachon 2020). The above suggests high influence of ΔCO_2 on F_{CO_2} variability on a short-term basis in lakes with characteristics favoring high winds (Fig. 5c). A lower relative influence of *k* on F_{CO_2} variability in such lakes, which usually have high values of *k* (Schilder et al. 2013), may seem counterintuitive. However, the magnitude of *k* does not reflect variability in *k*, and in lakes where *k* is commonly stable from constant wind exposure, *k* also varies less and is thereby less important for F_{CO_2} variability. Similarly, the short-term relative influence of *k* on F_{CO_2} variability was high in small lakes sheltered by forests (LJR, NOR, NBJ, and NAS; 1–3 ha and 69–89% forest coverage), where wind speeds were below the lake average and ΔCO_2 was generally high (Supporting Information Fig. S7). There, variability in *k* between calm (most of the time) and higher wind speed (occasionally) can be greater than in lakes with consistently high wind speeds. In a short-term study of a small boreal lake located directly beside NBJ, MacIntyre et al. (2021) found that *k* was primarily wind driven yet highly variable even at low wind speeds, which supports the possibility of variable *k* in small wind sheltered lakes.

Short-term relative influence of ΔCO_2 on F_{CO_2} variability was higher in lakes with low ΔCO_2. In our data, the
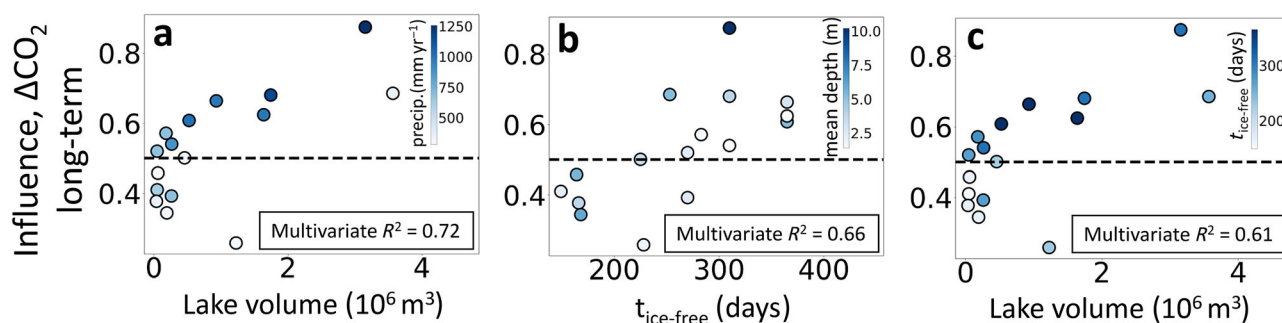
Fig. 6. Multivariate regressions for differences in long-term (ice-free period) lake mean relative influence of ΔCO_2 on F_{CO_2} variability (I_{CO_2}) at deep sampling locations, including the variables lake volume, duration of ice-free period ($t_{ice-free}$), yearly precipitation and mean lake depth. In each regression, one of the variables, (**a**) precipitation, (**b**) mean depth, and (**c**) $t_{ice-free}$, is displayed in colors ranging from low (white) to high (blue) values. BD3 and BD6 are excluded in the regressions due to limited sampling at deep locations. Dashed horizontal lines in the panels represent the divide between higher influence of $\Delta CO2$ on F_{CO_2} variability (above line) and higher influence of *k* on F_{CO_2} variability (below line).

F_{CO_2} variability

short-term (1–9 days)

long-term (ice-free period)

shallow (< 2m)   **deep (> 2m)**

Generally highest influence of *k* on F_{CO_2} variability, which increased with lower catchment slope, smaller catchment fraction of open field, lower wind speeds and smaller lake area (Fig. 5) ***k* generally contributes most to F_{CO_2} variability, especially in small (< 3 ha), forest-sheltered boreal or hemi-boreal lakes**	Generally highest influence of *k* on F_{CO_2} variability and similar variability between lakes as for deep locations. However, no variables were identified that could explain between-lake differences. ***k* generally contributes most to F_{CO_2} variability, but unclear relationship with lake type**	Generally highest influence of ΔCO_2 on F_{CO_2} variability, which increased with lake volume, precipitation and ice-free period length (Fig. 6) **ΔCO_2 generally contributes most to F_{CO_2} variability, especially in relatively large (>30 ha) south-boreal lakes and in lakes with seasonally variable precipitation**

General summary of potential explanations

The explanatory variables (ΔCO_2, *k*) with greatest variability are more likely to drive the variability of the response variable (F_{CO_2}):

➤ *k* is more variable than ΔCO_2 over short time frames, but its variability is reduced compared to ΔCO_2 over longer temporal scales (Fig. 3).

➤ *k* is also most variable at wind sheltered locations (small forest lakes and near-shore locations) where wind-induced turbulence ranges from very low to high (during strong storms). At more wind exposed locations, differences in turbulence may be smaller because very low *k* are rare.

➤ Over longer time periods, variability in ΔCO_2 is more likely, particularly in deep large-volume lakes with long ice-free periods where turnover periods can be prolonged and hypolimnetic CO_2 accumulation enhanced, and in lakes within large catchments that receive much rain, where hydrological input from the catchment is likely greater.

Fig. 7. Conceptual flow-chart for lake mean relative influence of ΔCO_2 and *k* on F_{CO_2} variability. Red and blue lines and boxes denote shallow and deep locations, respectively. The gray box denotes shallow and deep locations combined. Text in colored boxes separate between results (normal font) and implications (bold font). The bottom box summarizes potential explanations for results.

unproductive clearwater lakes (BD3, BD4, BD6, SGA) and the most eutrophic lake (SOD) have low levels of ΔCO_2. In these low-ΔCO_2 lakes, even small alterations in ΔCO_2 can be of high relative importance for F_{CO_2}. In such situations, the interactions between ΔCO_2 and *k* can be extra important for F_{CO_2} variability. For example, if minor relative changes in *k* are still large enough to induce sufficient transport of CO_2 from, for example, epilimnetic sediments to surface waters to generate greater relative changes in ΔCO_2, this would imply a greater influence of ΔCO_2 than *k* on F_{CO_2} variability.

Lower levels of ΔCO_2 in more eutrophic lakes may instead be caused by higher rates of primary productivity than respiration during daytime (when most of our samples were collected), where primary production produces O_2 and consumes CO_2 and respiration produces CO_2 and consumes O_2.

The effect of primary productivity (and thus its consumption of CO_2) depends on factors that vary on daily and diel timescales, such as temperature and light availability (Gomez-Gener et al. 2021). Thus, higher primary productivity can positively affect the relative influence of ΔCO_2 on short-term variability of F_{CO_2}. In lakes where respiration and primary productivity is important for surface water CO_2, alterations in DO saturation can reflect surface water CO_2 concentrations (Cole et al. 2000; Hanson et al. 2003; Hanson et al. 2006). We found strong correlations between DO saturation and ΔCO_2 ($r = -0.88$) and they also had similar importance in regressions (Supporting Information Table S3). Therefore, we believe that primary productivity can control short-term relative influence of ΔCO_2 on F_{CO_2} variability in eutrophic lakes.

Long-term influence on F_{CO_2} variability across lakes

Duration of the ice-free period, lake volume and yearly precipitation dominated among variables being most relevant for long-term relative influence of ΔCO_2 on F_{CO_2} variability at deep lake locations and explained up to 74% of variability when used in multiple regressions (Supporting Information Table S4). Deep large-volume lakes are likely to have more gradual and prolonged seasonal mixing periods. The effect of such mixing periods on F_{CO_2} variability may be accentuated in lower-latitude lakes with longer ice-free periods, where a warmer climate may enhance hypolimnetic CO_2 production and accumulation. Prolonged seasonal mixing periods and enhanced accumulation can in turn generate greater seasonal variability in F_{CO_2} and ΔCO_2 when deep-water mixes to the surface. The precipitation effect on seasonal F_{CO_2} variability can be related to transport of CO_2 from the catchment to the lake with surface water or via soil and ground water. If greater mean yearly precipitation implies greater variability in precipitation and CO_2 transport to the lake, this will also trigger greater variability in F_{CO_2} and ΔCO_2.

Factors controlling the influence of ΔCO_2 vs. k on F_{CO_2} within lakes

Short-term relative influence of ΔCO_2 on F_{CO_2} variability at shallow and deep locations was similar among our lakes, with differences in I_{CO_2} between these two depth categories averaging at 10% for our lakes. For longer time periods, differences were greater (19% on average) and whereas we found variables that could explain a substantial portion ($R^2 \le 0.71$) of between-lake variability in I_{CO_2} at deep sampling locations (Fig. 6), we found none that could explain such patterns at shallow locations. This could be due to complex interactions between ΔCO_2 and k. For example, at shallow locations a greater k will enhance not only gas exchange rates to the atmosphere, but also the gas exchange with shallow sediments, sustaining a steeper gas concentration gradient (del Giorgio and Williams. 2005; Bergström et al. 2010; Bartosiewicz et al. 2015). Hence, k variability can have a dual effect on F_{CO_2} and thereby emerge as the most important factor for F_{CO_2} variability from shallow waters. In effect, ΔCO_2 variability at shallow locations may, similarly to k, be more linked to short-term patterns in wind speed and wind direction and less linked to seasonal patterns.

In addition, the use of lake mean values in regressions also means that some variables, such as wind speed, may not adequately represent near-shore conditions or may represent it differently in different lakes. This may have resulted in a lack of models for relative influence on F_{CO_2} at shallow locations. Furthermore, other drivers for long-term relative influence on F_{CO_2} at shallow locations than those considered here may be important, such as groundwater or riverine input (Striegl and Michmerhuizen 1998; Humborg et al. 2010; Maberly et al. 2013), catchment geology and productivity (Maberly et al. 2013), sediment characteristics (Bergström et al. 2010; Chmiel et al. 2016), internal waves and seiches (Heiskanen

et al. 2014; MacIntyre et al. 2021), and macrophyte density (Peixoto et al. 2016). These factors need to be addressed in other studies. Overall, we here show that the consideration of temporal and spatial scales to disentangle processes important for spatial and temporal variability is important when comparing patterns in F_{CO_2} within and among lakes.

Implications for assessing F_{CO_2}

Clearly, both k and ΔCO_2 control F_{CO_2} variability, but with a relative importance that differs between temporal scales. For time periods ranging from individual monthly sampling campaigns (1–9 d) to several months, k is a key variable although ΔCO_2 sets base-line conditions for fluxes. Over seasonal time periods ΔCO_2 is a main driver in many lakes, especially in lakes with large volumes and long ice-free periods, allowing for prolonged seasonal mixing periods and greater storage of DIC, or in lakes within large and precipitation-prone catchments, allowing for greater input of carbon through the catchment (Fig. 7).

The contribution from shallow and deep areas, respectively, to total lake F_{CO_2} corresponded to the lake area that they encompassed (shallow areas encompassed 10%–70% of lake area; Supporting Information Fig. S6), showing that both shallow and deep areas are equally important for whole lake F_{CO_2}. Sampling designs that combine spatially distributed measurements of monthly pCO$_{2aq}$ with single-location estimates from k models that account for turbulence both within and above the lake (Tedford et al. 2014; Czikowsky et al. 2018; MacIntyre et al. 2021), may provide reliable estimates of ice-free F_{CO_2} at deep central parts of lakes but are likely to bias F_{CO_2} at shallow near-shore parts of lakes. To limit bias in such sampling designs, additional measurements at shallow locations as input to long-term modeling using k would be beneficial. However, and importantly, the large within- and among-lake differences in dominance of ΔCO_2 vs. k on F_{CO_2} combined with previous concerns regarding lack of reliability in commonly used wind-based k models (Schilder et al. 2013; Klaus and Vachon 2020) lead to the general advice to prioritize direct spatio-temporally distributed measurements of F_{CO_2} whenever possible or alternatively to assess k from models and with sampling designs that incorporate spatial variability of lake thermal structure and meteorology (e.g., MacIntyre et al. 2021).

Constraining the current uncertainty in F_{CO_2} is critical for accurate regional and global budgets. Our work here shows that variability over the short term depends on the gas transfer velocity but over the long term depends on variability in the concentration gradient of CO_2 at the air–water interface. We have shown that the latter depends on lake volume, precipitation, and duration of the ice-free period in boreal and sub-arctic lakes, implying that duration of seasonal mixing periods, hypolimnetic storage of DIC and input of carbon from the catchment are key factors. While modeling of emissions from lakes and reservoirs on global scales uses typical values of k and of ΔCO_2 over broad spatial areas

(Raymond et al. 2013), our method based on lake-attributes and landscape characteristics provides an approach for further refinement. Further improvement will come from studies such as this one which identify key drivers and with development of more effective low-cost sensor measurements for CO_2 assessments, facilitating estimates of CO_2 emissions from diverse environments.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Azen, R., and D. V. Budescu. 2003. The dominance analysis approach for comparing predictors in multiple regression. Psychol. Methods **8**: 129–148.

Banerjee, S., and S. MacIntyre. 2004. The air–water interface: Turbulence and scalar exchange, p. 181–237. *In* G. E. Likens [ed.], PIV and water waves. World Scientific.

Bartosiewicz, M., I. Laurion, and S. MacIntyre. 2015. Greenhouse gas emission and storage in a small shallow lake. Hydrobiologia **757**: 101–115.

Bastviken, D., I. Sundgren, S. Natchimuthu, H. Reyier, and M. Gålfalk. 2015. Technical note: Cost-efficient approaches to measure carbon dioxide (CO_2) fluxes and concentrations in terrestrial and aquatic environments using mini loggers. Biogeosciences **12**: 3849–3859.

Bergström, I., P. Kortelainen, J. Sarvala, and K. Salonen. 2010. Effects of temperature and sediment properties on benthic CO_2 production in an oligotrophic boreal lake. Freshw. Biol. **55**: 1747–1757.

Budescu, D. V. 1993. Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. Psychol. Bull. **114**: 542–551.

Carlson, R. E. 1977. A trophic state index for lakes. Limnol. Oceanogr. **22**: 361–369.

Chmiel, H. E., and others. 2016. The role of sediments in the carbon budget of a small boreal lake. Limnol. Oceanogr. **61**: 1814–1825.

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

Cole, J. J., and others. 2007. Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon budget. Ecosystems **10**: 172–185.

Cole, J. J., D. L. Bade, D. Bastviken, M. L. Pace, and M. Van de Bogert. 2010. Multiple approaches to estimating air–water gas exchange in small lakes. Limnol. Oceanogr. Methods **8**: 285–293.

Conrad, R., and W. Seiler. 1988. Influence of the surface microlayer on the flux of nonconservative trace gases (CO, H_2, CH_4, N_2O) across the ocean-atmosphere interface. J. Atmos. Chem. **6**: 83–94.

Cunliffe, M., R. C. Upstill-Goddard, and J. C. Murrell. 2011. Microbiology of aquatic surface microlayers. FEMS Microbiol. Rev. **35**: 233–246.

Czikowsky, M. J., S. MacIntyre, E. W. Tedford, J. Vidal, and S. D. Miller. 2018. Effects of wind and buoyancy on carbon dioxide distribution and air–water flux of a stratified temperate lake. J. Geophys. Res. Biogeo. **123**: 2305–2322.

del Giorgio, P. A., and P. L. B. Williams. 2005. The global significance of respiration in aquatic ecosystems: From single cells to the biosphere. Respir. Aquat. Ecosyst. **1**: 267–303.

Duarte, C. M., and Y. T. Prairie. 2005. Prevalence of heterotrophy and atmospheric CO_2 emissions from aquatic ecosystems. Ecosystems **8**: 862–870.

Ducharme-Riel, V., D. Vachon, P. del Giorgio, and Y. Prairie. 2015. The relative contribution of winter under-ice and summer hypolimnetic CO_2 accumulation to the annual CO_2 emissions from northern lakes. Ecosystems **18**: 547–559.

Gålfalk, M., D. Bastviken, S. Fredriksson, and L. Arneborg. 2013. Determination of the piston velocity for water-air interfaces using flux chambers, acoustic Doppler velocimetry, and IR imaging of the water surface. J. Geophys. Res. Biogeo. **118**: 770–782.

Gómez-Gener, L., and others. 2021. Global carbon dioxide efflux from rivers enhanced by high nocturnal emissions. Nat. Geosci. **14**: 289–294.

Häggmark, L., K.-I. Ivarsson, S. Gollvik, and P.-O. Olofsson. 2000. Mesan, an operational mesoscale analysis system. Tellus A Dyn. Meteorol. Oceanogr. **52**: 2–20.

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

Hanson, P. C., S. R. Carpenter, D. E. Armstrong, E. H. Stanley, and T. K. Kratz. 2006. Lake dissolved inorganic carbon and dissolved oxygen: Changing drivers from days to decades. Ecological monographs **76**: 343–363.

Heiskanen, J. J., and others. 2014. Effects of cooling and internal wave motions on gas transfer coefficients in a boreal lake. Tellus B Chem. Phys. Meteorol. **66**: 22827.

Holgerson, M. A., and P. A. Raymond. 2016. Large contribution to inland water CO_2 and CH_4 emissions from very small ponds. Nat. Geosci. **9**: 222–226.

Humborg, C., and others. 2010. CO_2 supersaturation along the aquatic conduit in Swedish watersheds as constrained by terrestrial respiration, aquatic respiration and weathering. Glob. Chang. Biol. **16**: 1966–1978.

Jammet, M., and others. 2017. Year-round CH_4 and CO_2 flux dynamics in two contrasting freshwater ecosystems of the subarctic. Biogeosciences **14**: 5189–5216.

Jansen, J., and others. 2019. Climate-sensitive controls on large spring emissions of CH_4 and CO_2 from northern lakes. J. Geophys. Res. Biogeo. **124**: 2379–2399.

Jansen, J., and others. 2020. Drivers of diffusive CH$_4$ emissions from shallow subarctic lakes on daily to multi-year timescales. Biogeosciences **17**: 1911–1932.

Klaus, M., and D. Vachon. 2020. Challenges of predicting gas transfer velocity from wind measurements over global lakes. Aquat. Sci. **82**: 53.

Koschorreck, M., Y. T. Prairie, J. Kim, and R. Marcé. 2021. Technical note: CO$_2$ is not like CH$_4$—Limits of and corrections to the headspace method to analyse pCO$_2$ in fresh water. Biogeosciences **18**: 1619–1627.

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

López Bellido, J., T. Tulonen, P. Kankaala, and A. Ojala. 2009. CO$_2$ and CH$_4$ fluxes during spring and autumn mixing periods in a boreal lake (Pääjärvi, southern Finland). J. Geophys. Res. **114**: G04007.

Maberly, S., P. Barker, A. Stott, and M. De Ville. 2013. Catchment productivity controls CO$_2$ emissions from lakes. Nat. Clim. Change **3**: 391–394.

MacIntyre, S., and others. 2021. Turbulence in a small boreal lake: Consequences for air–water gas exchange. Limnol. Oceanogr. **66**: 827–854.

Marcé, R., B. Obrador, J.-A. Morguí, J. L. Riera, P. López, and J. Armengol. 2015. Carbonate weathering as a driver of CO$_2$ supersaturation in lakes. Nat. Geosci. **8**: 107–111.

Murphy, J., and J. P. Riley. 1962. A modified single solution method for the determination of phosphate in natural waters. Anal. Chim. Acta **27**: 31–36.

Natchimuthu, S., I. Sundgren, M. Gålfalk, L. Klemedtsson, and D. Bastviken. 2017. Spatiotemporal variability of lake pCO$_2$ and CO$_2$ fluxes in a hemiboreal catchment. J. Geophys. Res. Biogeosci. **122**: 30–49.

Pajala, G., and others. 2023. Higher apparent gas transfer velocities for CO$_2$ compared to CH$_4$ in small lakes. Environ. Sci. Technol. **57**: 8578–8587.

Peixoto, R., H. Marotta, D. Bastviken, and A. Enrich-Prast. 2016. Floating aquatic macrophytes can substantially offset open water CO$_2$ emissions from tropical floodplain lake ecosystems. Ecosystems **19**: 724–736.

Raymond, P. A., and others. 2013. Global carbon dioxide emissions from inland waters. Nature **503**: 355–359.

Read, J. S., and others. 2012. Lake-size dependency of wind shear and convection as controls on gas exchange. Geophys. Res. Lett. **39**: 1–5.

Reis, P. C., and F. A. Barbosa. 2014. Diurnal sampling reveals significant variation in CO$_2$ emission from a tropical productive lake. Braz. J. Biol. **74**: S113–S119.

Richards, F. A., and T. G. Thompson. 1952. The estimation and characterization of plankton populations by pigment analyses. J. Mar. Res. **11**: 156.

Rocher-Ros, G., R. A. Sponseller, W. Lidberg, C. M. Mörth, and R. Giesler. 2019. Landscape process domains drive patterns of CO$_2$ evasion from river networks. Limnol. Oceanogr. Lett. **4**: 87–95.

Rudberg, D., and others. 2021. Diel variability of CO$_2$ emissions from northern lakes. J. Geophys. Res. Biogeosci. **126**: e2021JG006246.

Schilder, J., and others. 2013. Spatial heterogeneity and lake morphology affect diffusive greenhouse gas emission estimates of lakes. Geophys. Res. Lett. **40**: 5752–5756.

Seabold, S., and J. Perktold. 2010. Statsmodels: Econometric and statistical modeling with python. Proceedings of the 9th Python in Science Conference **57**: 92–96.

Sheather, S. 2009. A modern approach to regression with R. Springer Science & Business Media.

Striegl, R. G., and C. M. Michmerhuizen. 1998. Hydrologic influence on methane and carbon dioxide dynamics at two north-central Minnesota lakes. Limnol. Oceanogr. **43**: 1519–1529.

Stumm, W., and J. J. Morgan. 1996. Aquatic chemistry: Chemical equilibria and rates in natural waters, 3rd ed. John Wiley & Sons.

Tan, Z., and others. 2017. Modeling CO$_2$ emissions from Arctic lakes: Model development and site-level study. J. Adv. Model. Earth Syst. **9**: 2190–2213.

Tedford, E. W., S. MacIntyre, S. D. Miller, and M. J. Czikowsky. 2014. Similarity scaling of turbulence in a temperate lake during fall cooling. J. Geophys. Res. Oceans **119**: 4689–4713.

Tonetta, D., P. A. Staehr, R. Schmitt, and M. M. Petrucio. 2016. Physical conditions driving the spatial and temporal variability in aquatic metabolism of a subtropical coastal lake. Limnologica **58**: 30–40.

Vachon, D., Y. T. Prairie, and R. Smith. 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.

Virtanen, P., and others. 2020. SciPy 1.0: Fundamental algorithms for scientific computing in Python. Nat. Methods **17**: 261–272.

Wallin, M., I. Buffam, M. Öquist, H. Laudon, and K. Bishop. 2010. Temporal and spatial variability of dissolved inorganic carbon in a boreal stream network: Concentrations and downstream fluxes. J. Geophys. Res. Biogeo. **115**: G02014.

Warton, D. I., R. A. Duursma, D. S. Falster, and S. Taskinen. 2012. Smatr 3– An R package for estimation and inference about allometric lines. Methods Ecol. Evol. **3**: 257–259.

Weiss, R. F. 1974. Carbon dioxide in water and seawater: The solubility of a non-ideal gas. Mar. Chem. **2**: 203–215.

Zhou, X., and K. Mopper. 1997. Photochemical production of low-molecular-weight carbonyl compounds in seawater and surface microlayer and their air–sea exchange. Mar. Chem. **56**: 201–213.

Conflict of Interest

None declared.

Submitted 03 June 2022
Revised 20 December 2023
Accepted 16 January 2024

Associate editor: Robert O. Hall, Jr