Received: 3 June 2017 | Accepted: 24 August 2017

DOI: 10.1111/gcb.13902

PRIMARY RESEARCH ARTICLE

WILEY Global Change Biology

CO_2 evasion from boreal lakes: Revised estimate, drivers of spatial variability, and future projections

Adam Hastie[1] | Ronny Lauerwald[1,2] | Gesa Weyhenmeyer[3] | Sebastian Sobek[3] | Charles Verpoorter[3,4] | Pierre Regnier[1]

[1]Biogeochemistry and Earth System Modelling, Department of Geoscience, Environment and Society, Université Libre de Bruxelles, Bruxelles, Belgium

[2]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK

[3]Department of Ecology and Genetics/Limnology, Uppsala University (UU), Uppsala, Sweden

[4]Laboratoire d'Océanologie et de Géosciences (LOG), Université du Littoral Côte d'Opale, CNRS, UMR 8187, University of Lille, Wimereux, France

Correspondence
Adam Hastie, Biogeochemistry and Earth System Modelling, Department of Geoscience, Environment and Society, Université Libre de Bruxelles, Bruxelles, Belgium.
Email: Adam.Hastie@ulb.ac.be

Funding information
Horizon 2020 Framework Programme, Grant/Award Number: 643052; Marie Sklodowska-Curie, Grant/Award Number: 643052; European Union's Horizon 2020 research and innovation program, Grant/Award Number: 703813; Swedish Research Council, Grant/Award Number: 2016-04153; Knut and Alice Wallenberg Foundation; European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC, Grant/Award Number: 336642

Abstract

Lakes (including reservoirs) are an important component of the global carbon (C) cycle, as acknowledged by the fifth assessment report of the IPCC. In the context of lakes, the boreal region is disproportionately important contributing to 27% of the worldwide lake area, despite representing just 14% of global land surface area. In this study, we used a statistical approach to derive a prediction equation for the partial pressure of CO_2 (pCO_2) in lakes as a function of lake area, terrestrial net primary productivity (NPP), and precipitation (r^2 = .56), and to create the first high-resolution, circumboreal map (0.5°) of lake pCO_2. The map of pCO_2 was combined with lake area from the recently published GLOWABO database and three different estimates of the gas transfer velocity k to produce a resulting map of CO_2 evasion (FCO_2). For the boreal region, we estimate an average, lake area weighted, pCO_2 of 966 (678–1,325) μatm and a total FCO_2 of 189 (74–347) Tg C $year^{-1}$, and evaluate the corresponding uncertainties based on Monte Carlo simulation. Our estimate of FCO_2 is approximately twofold greater than previous estimates, as a result of methodological and data source differences. We use our results along with published estimates of the other C fluxes through inland waters to derive a C budget for the boreal region, and find that FCO_2 from lakes is the most significant flux of the land-ocean aquatic continuum, and of a similar magnitude as emissions from forest fires. Using the model and applying it to spatially resolved projections of terrestrial NPP and precipitation while keeping everything else constant, we predict a 107% increase in boreal lake FCO_2 under emission scenario RCP8.5 by 2100. Our projections are largely driven by increases in terrestrial NPP over the same period, showing the very close connection between the terrestrial and aquatic C cycle.

KEYWORDS
boreal, carbon budget, climate change, CO_2, future projections, lake, precipitation, terrestrial NPP

1 | INTRODUCTION

Lakes (including reservoirs) are an important component of the global carbon (C) cycle, as acknowledged by the fifth assessment report of the IPCC (Ciais et al., 2013). Global, regional, and local studies commonly report carbon dioxide (CO_2) supersaturation (e.g., Lapierre & del Giorgio, 2012; Raymond et al., 2013; Sobek, Tranvik, & Cole, 2005; Weyhenmeyer, Kortelainen, Sobek, Muller, & Rantakari, 2012), resulting in an evasive flux of CO_2 (FCO_2 in this paper), which on a global scale, equates to approximately 0.32–0.64 Pg C $year^{-1}$

 © 2017 John Wiley & Sons Ltd

(Aufdenkampe et al., 2011; Cole, Caraco, Kling, & Kratz, 1994; Cole et al., 2007; Holgerson & Raymond, 2016; Raymond et al., 2013; Tranvik et al., 2009). This CO_2 outgassing corresponds to roughly 12%–25% of the total carbon flux mobilized from soils and the bedrock into aquatic systems (Regnier et al., 2013).

In the context of lakes, the boreal forest region (BF, as defined in Potapov, Hansen, Stehman, Loveland, & Pittman, 2008) is disproportionately important. According to the satellite-based GLOWABO product (Verpoorter, Kutser, Seekell, & Tranvik, 2014), 1.35×10^6 km^2 out of 5×10^6 km^2 lakes globally are located in this region, contributing to 27% of the worldwide lake area, despite the BF representing just 14% of global land surface area. Boreal waters are also predicted to be very sensitive to future climate change; in particular with regard to increasing temperature and terrestrial net primary productivity (NPP), as well as mobilization of C from thawing permafrost soils (Gauthier, Bernier, Kuuluvainen, Shvidenko, & Schepaschenko, 2015; Intergovernmental Panel on Climate Change (IPCC), Climate Change, 2013; Koven, 2013; Price et al., 2011). It is thus important to understand the consequences of future climate change for lake C cycling and lake FCO_2. In order to project future boreal lake CO_2 partial pressure (pCO_2) and FCO_2, we first need robust and spatially resolved estimates for the present day, and we then need to identify the key environmental drivers of lake FCO_2.

The spatial heterogeneity in lake FCO_2 has so far only been assessed using an approach averaging observed local lake values of pCO_2 across a region of interest, from which an evasion rate was calculated and applied to the entire lake surface area of the region. At the national scale for the USA, McDonald, Stets, Striegl, and Butman (2013) and Butman et al. (2016) resolved pCO_2 and FCO_2 using 16 and 19 regions, respectively. Similarly, Humborg et al. (2010) calculated average pCO_2 values and FCO_2 rates for five lake size classes to estimate the total FCO_2 from Swedish lakes, although only a single region was used here. Raymond et al. (2013) presented the first global map of lake pCO_2 and FCO_2 for different lake size classes at the resolution of the so-called COSCAT segmentation of Meybeck, Dürr, and Vörösmarty (2006), which subdivides the world's catchments into 231 units. While such regionalization provides robust estimates where data availability is high, they are of less use where observations are scarce or missing. In the case of boreal lakes, large datasets are available for Sweden, Finland, and the southeastern part of Canada. In contrast, data availability is scarce across very large portions of NW America and the Asian continent. Thus, for the regions that lack empirical data, pCO_2 and FCO_2 need to be modeled.

In order to extrapolate pCO_2 to those locations where data are scarce or absent, globally available environmental drivers need to be identified. In previous studies, the spatial variability of lake pCO_2 in high latitude regions has been linked to a wide range of variables including dissolved or total organic carbon (DOC/TOC) concentration in lake water (e.g., Humborg et al., 2010; Lapierre & del Giorgio, 2012; Roehm, Prairie, & del Giorgio, 2009; Weyhenmeyer et al., 2012 for most recent contributions), lake area and depth (Humborg et al., 2010; Kortelainen et al., 2006; Raymond et al., 2013; Roehm et al., 2009; Sobek, Algesten, Bergström, Jansson, & Tranvik, 2003), dissolved inorganic carbon (DIC) input from the catchment (Maberly, Barker, Stott, & De Ville, 2013; Perga et al., 2016; Weyhenmeyer et al., 2015), chlorophyll-a concentration (Kortelainen et al., 2006, 2013; Maberly et al., 2013; Perga et al., 2016; Roehm et al., 2009), and precipitation (Rantakari & Kortelainen, 2005; Sobek et al., 2003). While hydrochemical and physical variables observed in the field give valuable insights into in-lake processes controlling pCO_2 and its short-term variability, they are of limited use for extrapolations, as lake pCO_2 can only be predicted where these variables have been locally observed. Considering the vast number of lakes in the boreal zone, the proportion that is covered by sampling programs is small. Therefore, geodata sets of potential environmental controls related to climate, terrain, geology, and vegetation, which cover the global landmass in a consistent way, are better alternatives for large-scale assessments.

Lauerwald (2013), Hartmann, Moosdorf, Kempe, and Raymond showed that a significant proportion of the spatial variability in river pCO_2 in North America could be explained by catchment variables derived from geodatabases. Using a multiple regression analysis, they found that 43% ($r^2 = .43$) of the river pCO_2 variability across North America is related to annual mean precipitation, annual mean air temperature, and mean catchment slope gradient. A similar approach was later applied to derive global maps of river pCO_2 and FCO_2 (Lauerwald, Laruelle, Hartmann, Ciais, & Regnier, 2015). A multiple regression using terrestrial NPP, population density, mean catchment slope gradient, as well as mean air temperature at the sampling location was able to explain 47% ($r^2 = .47$) of the spatial variability in global river pCO_2. The map of predicted pCO_2 was then combined with estimates of stream surface area and gas exchange velocity k to derive a map of global FCO_2 from rivers at 0.5° resolution.

In this study, we used globally available environmental drivers derived from geodatabases to predict a 0.5° map of lake pCO_2 for the entire BF biome from a limited, spatially lumped set of sampling data. The map of pCO_2 was then combined with lake area from the GLOWABO database (Verpoorter et al., 2014) and estimates of gas transfer velocity k to produce the first high-resolution map of boreal lake FCO_2. We then merged our new map with the spatially resolved river FCO_2 from Lauerwald et al. (2015), lateral land-ocean C exports (GlobalNews, Mayorga et al., 2010), and lake C burial (Heathcote, Anderson, Prairie, Engstrom, & del Giorgio, 2015), to derive a full aquatic carbon budget for the circumboreal region. Finally, based on spatially resolved projections of terrestrial NPP and precipitation, we used the model to predict changes in lake pCO_2 and FCO_2 over the twenty-first century.

2 | MATERIALS AND METHODS

2.1 | Lake pCO_2 data

Lake chemistry data were collated from a number of different databases pertaining to distinct regions (Figure 1): (i) Swedish lake pCO_2

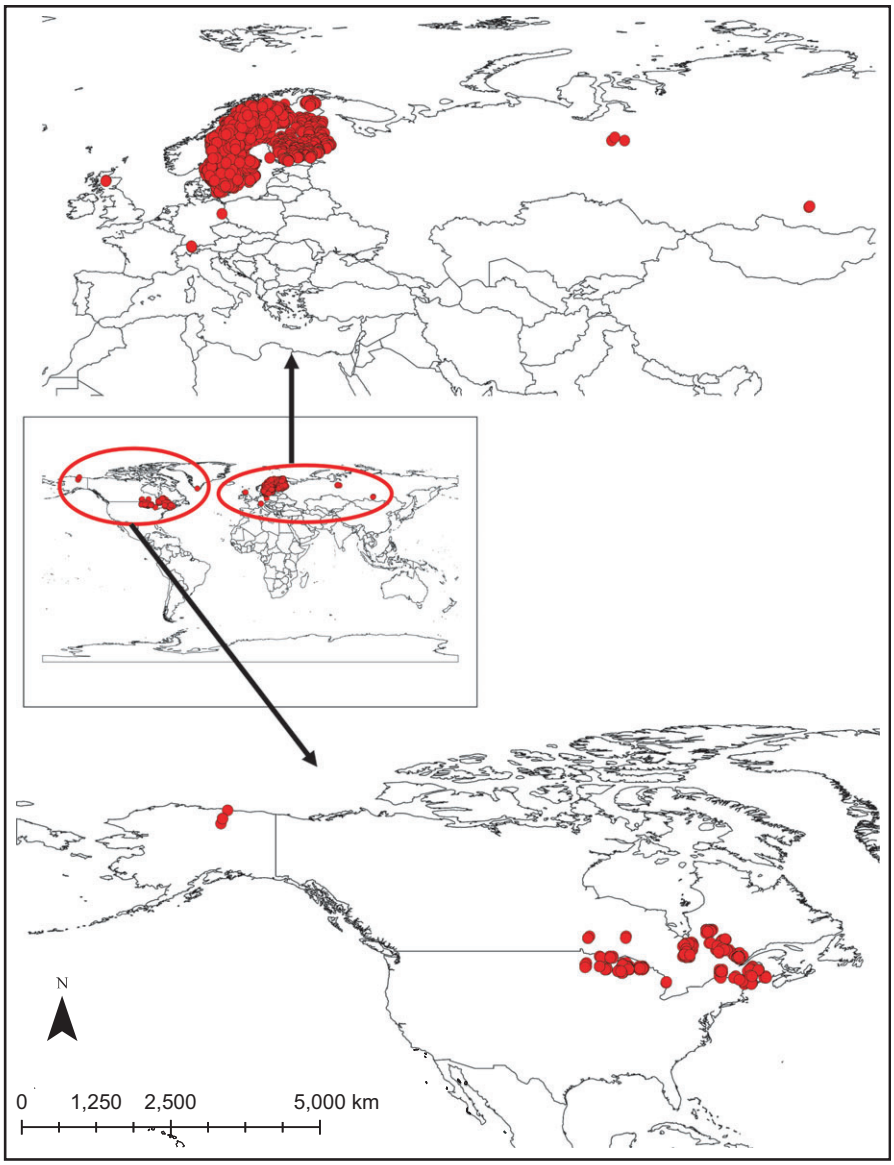

FIGURE 1 Geographical distribution of measured lake pCO_2 data (n = ~27,000 samples)

data were taken from the Swedish lake chemistry database published in Weyhenmeyer et al. (2012), (ii) pCO_2 data from a number of other boreal regions (in particular, the whole of Scandinavia and North America) were taken from a global lake pCO_2 database published in Sobek et al. (2005), (iii) Canadian lake pCO_2 data were taken from Lapierre and del Giorgio (2012), and (iv) Siberian data were taken from Shirokova et al. (2013). The Siberian data were reported in μmol/L and converted into partial pressure using Henry's constant, adjusted for water temperature, following Telmer and Veizer (1999).

Approximately 99% of the pCO_2 values were calculated from measured alkalinity, pH, and water temperature. Samples with a pH ≤5.4 were discarded because calculating pCO_2 from alkalinity, pH, and water temperature is highly error prone in low-pH and low-alkalinity waters (Abril et al., 2015; Raymond et al., 2013). A recent study in Sweden (Wallin, Löfgren, Erlandsson, & Bishop, 2014) found a slight overestimation of the average pCO_2 in samples calculated with the alkalinity-based method. The majority of these samples had low alkalinity (<0.07 meq/L). As such, we compared the median and average pCO_2 values in our dataset above and below this alkalinity threshold. In fact, in our data the samples with an alkalinity >0.07 meq/L had a marginally higher median/average pCO_2 than those ≤0.07 meq/L. We further investigated the effect of alkalinity on the indirect pCO_2 data using simple linear regression. A significant but very weak correlation ($r = -.0308$, $p < .05$, $N = 26{,}530$) was found between the indirectly observed pCO_2 data and alkalinity (after discarding samples with pH ≤ 5.4). Therefore, we conclude that using a pH threshold of 5.4 is a sufficient criterion for filtering out unreliable data. Moreover, another study in Sweden and Finland (Denfeld, Kortelainen, Rantakari, Sobek, & Weyhenmeyer, 2016) did not find a statistically significant difference between directly and indirectly measured pCO_2 data (Wilcoxon each pair test: $p > .05$). The remaining 1% of the samples (which includes 25% of North American data) corresponded to direct pCO_2 measurements (see Sobek et al., 2005 and Lapierre & del Giorgio, 2012 for more details).

In order to remove pCO_2 data from lakes potentially covered by ice, where possible, only data measured at a water temperature

greater than 4°C and between the months of April and November were retained in the analysis. For a small number of samples (<1%), water temperature was not reported and in these cases, all data from April to November were retained. After the data selection procedure described above, approximately 27,000 pCO_2 samples were retained. See discussion for further deliberation on the issue of ice cover.

2.2 | Data aggregation

Based on coordinates and lake name, median pCO_2 was calculated per sampling location. Then, using the median pCO_2 for all sampling locations within each 0.5° grid cell, a spatially aggregated mean pCO_2 was calculated at this resolution. This produced a single "open water" lake pCO_2 value for 584 grid cells located within the circumboreal region. In order to prevent a geographical bias and to have a similar number of pCO_2 grids representing Scandinavia and North America, only Scandinavian grids with a mean pCO_2 calculated from at least 22 median pCO_2 values (in other words, 22 sampled lakes) were retained in our training data. This rule did not apply to the grids containing the very largest lakes in the Scandinavian dataset, which were included regardless of the number of samples used in data aggregation to ensure that they are represented. Indeed, the largest lakes in the boreal region are larger than the area of a 0.5° grid cell, meaning that in some cases it is not possible to have more than one lake in a grid. Given the relative sparsity of the North American data, all of the North American grids were retained regardless of the number of lakes represented in each grid. This selective procedure left 168 grids evenly divided between Scandinavia and North America, as well as four grids in Siberia. Lake area data were also aggregated to a resolution of 0.5°, and $\log_{10}$ transformed prior to computing the mean $\log_{10}$ lake area per grid.

2.3 | Predictors of pCO_2

A set of environmental and climatic variables were selected and sourced, largely from publically available geodatabases (Table 1). The choice of retained variables was guided by two principles: firstly, we only chose data that were global in coverage with a resolution of at least 0.5°; secondly, we prioritized variables that have previously been shown to drive the variation of pCO_2 in inland water bodies. Among all variables, pCO_2 and average slope gradient of the catchment clearly showed a skewed distribution, and were thus log-transformed before regression analysis.

Firstly, the Pearson correlation coefficients between $\log_{10}$ (pCO_2) and all of the variables, as well as between variables themselves, were analyzed using the software package STATISTICA™. Secondly, we fitted multiple linear regression models with all possible combinations of three of the 17 predictors described in Table 1 using the software package R (R Core Team 2013). Initially, we placed no limit on the number of retained predictors, but found that adding a fourth predictor added little in the way of additional descriptive power. Therefore, we placed a limit of three predictors to ensure a parsimonious model.

2.4 | Upscaling of pCO_2 data

The fitted regression equation was applied in a geographical information system (GIS) to build a high-resolution map (0.5°) of pCO_2 from the drivers selected from the statistical treatment of the geodata. Terrestrial NPP was taken from the MODIS satellite-derived dataset described in Zhao, Heinsch, Nemani, and Running (2005), while precipitation was taken from Hijmans, Cameron, Parra, Jones, and Jarvis (2005), a high-resolution database based on interpolation of global weather station data. Log_{10} (lake area) (and in turn pCO_2) was adjusted to be representative for the total lake area within each grid, and hence, the total gas exchange flux through the aquatic-atmosphere interface, using the following equation:

$$\text{Adj. } A_{\text{lake}} = \frac{\sum \text{Log}_{10}(A_{\text{lake}}) * A_{\text{lake}}}{\sum A_{\text{lake}}} (\text{per grid}) \quad (1)$$

where A_{lake} denotes lake area.

Lake area data from the literature (Lapierre & del Giorgio, 2012; Shirokova et al., 2013; Sobek et al., 2005; Weyhenmeyer et al., 2012) were used for training the statistical model (i.e., lake area as a predictor of pCO_2) while lake area from the GLOWABO database (Verpoorter et al., 2014) was used for extrapolation of pCO_2 and calculation of FCO_2. This is because each aggregated pCO_2 value is the only representative of those lakes, which happen to have been sampled in that specific grid. When we extrapolate to the circumboreal scale, we use the GLOWABO database so that estimated pCO_2 is representative of the average size of all of the lakes in each grid. GLOWABO is a global inventory of lakes, which was developed by applying a lake extraction algorithm to high-resolution stationary satellite imagery (Verpoorter, Kutser, & Tranvik, 2012). As the database is stationary, it relies on an algorithm to filter the satellite images to minimize false detection due to mountain and cloud shadows.

The regression model was validated with 131 of the discarded grids from Scandinavia, selecting only those aggregated from a minimum of 10 samples per grid. All of the validation data were aggregated to the 0.5° grid scale before being compared to predicted pCO_2.

2.5 | Calculation of FCO_2

FCO_2 was calculated using the equation:

$$FCO_2 = A_{\text{lake}} * k * \Delta CO_2 \quad (2)$$

where FCO_2 is in mol/day, A_{lake} (m^2) denotes lake area, k (m/day) the gas exchange coefficient, and ΔCO_2 (mol/m^3) difference between water and air pCO_2, assuming an atmospheric pCO_2 of 390 μatm. In Equation 2, lake pCO_2 is computed according to the procedure described in the preceding sections. pCO_2 was then converted to CO_2 concentrations using Henry's constant K_H, corrected for temperature.

TABLE 1 Environmental geodata evaluated as potential predictors of $\log_{10}(pCO_2)$ (training data): Basic statistics

Variable	Unit	Mean	Median	Min	Max	Source	Resolution
Lake area (A_{lake})	km^2	10.95	0.52	0.01	82,200.00	Sobek et al. (2005), Weyhenmeyer et al. (2012), Lapierre & del Giorgio (2012)	30″
Air temperature (*T*, Apr–Nov monthly mean)	°C	9.01	9.30	1.20	12.80	Hijmans et al. (2005)	30″
Precipitation (*P*, Apr–Nov monthly mean)	mm/month	71.90	73.46	41.61	100.99	Hijmans et al. (2005)	30″
Wind speed (Apr–Nov monthly mean)	m/s	3.94	3.83	2.09	5.89	Hijmans et al. (2005)	30″
Soil pH index (top 5 cm)	–	5.14	5.12	4.68	6.18	SoilGrids (2014)	30″
Soil carbon content (top 5 cm)	g/kg	94.99	94.99	63.54	126.44	SoilGrids (2014)	30″
Terrestrial net primary productivity (NPP)	g C m^{-2} $year^{-1}$	486.73	504.15	60.20	746.89	Zhao et al., 2005;	30″
Population density	Inh./km^2	3.05	2.35	0.02	23.79	CIESIN & CIAT (2005)	2.5″
Elevation	m	240.45	226.60	18.60	973.81	GLOBE Task Team (1999)	30″
Catchment slope gradient	Degrees	0.988	0.670	0.0945	8.744	Lauerwald et al. (2015) GLOBE Task Team (1999)	30″
Runoff	mm/year	379.88	359.54	98.00	929.29	Fekete, Vörösmarty, & Grabs (2002)	30″
% cover evergreen trees	—	28	29	0	74	Global Land Cover 2000 database (2003)	2″
% cover mixed trees	—	22	18	0	83	Global Land Cover 2000 database (2003)	2″
% cover cultivated areas	—	10	02	0	88	Global Land Cover 2000 database (2003)	2″
% silt of soil	—	34	34	0	43	FAO/IIASA/ISRIC/ISSCAS/ JRC (2012)	30″
% sand of soil	—	52	53	0	62	FAO/IIASA/ISRIC/ISSCAS/ JRC (2012)	30″
% clay of soil	—	13	13	0	21	FAO/IIASA/ISRIC/ISSCAS/ JRC (2012)	30″

Lake area was calculated from the GLOWABO database (Verpoorter et al., 2014). The temperature-adjusted *k* was derived from k_{600} for each 0.5° grid using the Schmidt number, calculated from the mean water temperature over the April–November period. Water temperature was in turn calculated from mean air temperature over the same period using the equation reported by Lauerwald et al. (2015):

$$T_{water}[^\circ C] = 3.941 + 0.818\, T_{air}\,[^\circ C]\ (r^2 = .88) \quad (3)$$

There are numerous methods for estimating k_{600} (m/day) in the literature. In this study, we have compared estimated boreal lake FCO_2 using three different methodologies for calculating k_{600}. The first method is based on lake area, where four different k_{600} values (0.54, 1.16, 1.32, and 1.90 m/day) were applied for four corresponding lake area bins (<0.1, 0.1–1, 1–10, and >10 km^2). This approach is taken from Raymond et al. (2013), following the relationship between k_{600} and lake area proposed by Read et al. (2012). The second method is based on the relationship between k_{600} and wind speed given by Cole and Caraco (1998):

$$k_{600} = 2.07 + 0.251U_{10}^{1.7} \quad (4)$$

where U_{10} is the average wind speed in m/s at 10 m (Hijmans et al., 2005). This gives a k_{600} range between 0.58 and 2.04 m/day.

The third is taken from Vachon and Prairie (2013) and incorporates the effects of both wind speed and lake area:

$$k_{600} = 2.51 + 1.48 * U_{10} + 0.39 * U_{10} * \log_{10} A_{lake} \quad (5)$$

where A_{lake} is in km^2 and U_{10} in m/s.

k_{600} has also been found to depend on buoyancy flux in addition to wind speed and lake area. MacIntyre et al. (2010) found that higher k_{600} values occur during convective cooling compared to heating, and that during overnight low wind conditions, k_{600} depends on buoyancy flux rather than wind speed. We were unable to take these additional considerations into account as we lacked the necessary data. Despite this, we are confident that with the use of three methods to calculate k_{600}, and by explicitly

incorporating its variation in our Monte Carlo analysis, we adequately account for the uncertainty associated with gas exchange velocity.

FCO_2 was then converted to g C $year^{-1}$ by multiplying by the molar mass of carbon (12.01 g/mol) and the number of days per year, 365. FCO_2, was first calculated per grid and for the four lake size categories, <0.1, 0.1–1, 1–10, and >10 km^2 described previously (Read et al., 2012), before being amalgamated.

Total FCO_2 was calculated by summing FCO_2 from each size class. After pCO_2 and FCO_2 had been extrapolated, various masks were applied in ArcGIS in order to estimate values for different regions and countries. The boreal forest land cover region (BF) was taken from Potapov et al. (2008) (Fig. S13).

2.6 | Future projections of boreal pCO_2 and FCO_2

We acquired data from four Earth System Models used in the fifth Coupled Model Intercomparison Project (CMIP5) to predict pCO_2 and FCO_2 over the twenty-first century for the IPCC scenarios RCP2.6 and RCP8.5 (Intergovernmental Panel on Climate Change (IPCC), Climate Change, 2013): the Canadian Centre for Climate Modeling and Analysis (CCCma) CanESM2 model, the Met Office (UKMO) HadGEM2-ES model, the Institute Pierre-Simon Laplace (IPSL) IPSL-CM5A-LR model, and the Max Planck Institute for Meteorology (MPI) MPIESM-LR model. We selected these firstly, as they are all used in the CMIP5 project, and secondly because projected values of the environmental drivers of pCO_2 identified for the present day (terrestrial NPP and precipitation), as well as air temperature for adjusting future K_H, and k via the Schmidt number (as described in preceding section), were easily accessible online. The data were downloaded from the Earth System Grid Federation (ESGF) node hosted by the IPSL, for RCP2.6 and RCP8.5 model runs, as well as historical runs (1850–2005). All data were taken from the r1i1p1 realization of the simulations. In order to account for interannual variation, the data were aggregated in time to obtain projections for 10-year periods centered on the years 2000, 2030, 2050, and 2100. For example, total annual terrestrial NPP for the year 2030 was based on aggregating across the years 2025–2034. Some of the models only provide projections up to the year 2100 so for this specific period we aggregated the data across the preceding 10 years, 2090–2099. After aggregation, the projections of terrestrial NPP, air temperature, and precipitation were aggregated in a GIS to ensure a uniform spatial resolution of 0.5°. Finally, for each 0.5° grid, the simulated terrestrial NPP, air temperature, and precipitation was adjusted using the difference between modeled and observed data for the 2000 period. For example, the terrestrial NPP for the year 2030 was adjusted according to:

$$NPP_{2030,A} = NPP_{2000,O} + (NPP_{2030,M} - NPP_{2000,M}) \quad (6)$$

where A denotes adjusted, O denotes observed, and M denotes modeled. $NPP_{2000,O}$ was taken from Zhao et al. (2005), which was also used for training the prediction equation.

This procedure is similar to the one applied in regional downscaling of future projections, to reflect the fact the Earth System models are mostly designed to predict future changes and not to capture spatial patterns at the (sub)-regional scale.

After calculating future pCO_2, we then added the projected increase in atmospheric pCO_2 (above the present day value of 390 μatm used in this analysis) for the equivalent scenario and year. For example, under RCP8.5, atmospheric CO_2 is projected to be 936 μatm by 2100 (Meinshausen et al., 2011), and thus we added 546 (936–390) to our projected estimate of boreal lake pCO_2 for the same year and scenario. With this approach, we account for the fact that we do not use present day atmospheric pCO_2 in our prediction equation of lake pCO_2, but we assume it to be implicitly represented in the intercept of that equation, and we further assume that the increase in atmospheric pCO_2 will lead to a corresponding, total increase in lake pCO_2, preserving the total delta pCO_2. In our calculation of FCO_2, no such step is required as it is calculated from delta pCO_2, not absolute lake pCO_2.

2.7 | Uncertainty estimates based on Monte Carlo simulation

We calculated the uncertainty associated with our pCO_2 and FCO_2 estimates (5th and 95th percentiles) using a Monte Carlo simulation comprising 10,000 runs. Here, we calculated a probability density function for pCO_2 based on varying the b-estimates for each of the three predictors (terrestrial NPP, precipitation, and lake area) retained in the multiple linear regression, assuming a normal distribution constrained by the standard errors of each b-estimate (Table 4). Uncertainty associated with our future projections of pCO_2 and FCO_2 were similarly calculated using a Monte Carlo simulation forced by future projections of terrestrial NPP and precipitation. However, to account for the variation of terrestrial NPP and precipitation projections resulting from using four different Earth System models, we performed 4× 2,500 runs (2,500 for each model). Uncertainty was propagated to the calculation of FCO_2 using Equation 2, where k values were randomly selected between the minimum and maximum values calculated for each grid by the three methods outlined in the previous section. Monte Carlo runs (10,000) were again performed, assuming a uniform distribution of k. It was also assumed that pCO_2 and k vary independently of one another. The Monte Carlo analyses were performed using the statistical software R 3.2.2 (R Core Team, 2013).

3 | RESULTS

3.1 | Controls of spatial variation in pCO_2

Table 2 presents the correlations between aggregated (0.5°) $\log_{10}(pCO_2)$ and 17 variables derived from environmental geodata (Table 2).

Highly significant ($p < .001$) positive correlations were found for wind speed (Apr–Nov monthly mean) ($r = .58$), annual terrestrial net

TABLE 2 Pearson correlation (*r*) between the variables tested

	P	$\log_{10}$ (A_{lake})	Wind speed	Terrestrial NPP	Soil pH	*T*	$\log_{10}$ (slope gradient)	Elevation	Pop. density	Runoff	Soil org. C content	% cover evergreen trees	% cover mixed trees	% cover cultivated areas	% silt of soil	% sand of soil	% clay of soil
$\log_{10}$ (pCO_2)	−0.42***	−0.51***	0.58***	0.35***	−0.17*	0.34***	0.11	−0.30***	0.46***	−0.33***	0.041	0.36***	0.029	0.32***	−0.022	0.14	−0.18*
P		−0.043	−0.25***	0.28***	−0.11	0.33***	0.057	0.30***	−0.25**	0.74***	−0.22**	−0.42***	0.30***	−0.17*	0.10	−0.093	−0.0018
$\log_{10}$ (A_{lake})			−0.42***	−0.35***	0.092	−0.26***	−0.073	−0.13	−0.20**	0.10	0.17*	0.00	−0.15	−0.24**	−0.19*	−0.22**	−0.015
Wind speed				0.28***	0.08	0.29***	−0.31***	−0.20**	0.29**	−0.30***	−0.078	0.17*	−0.08	0.25***	−0.28***	0.25**	−0.057
Terrestrial NPP					−0.011	0.84***	−0.037	−0.24**	0.31***	0.018	−0.47***	−0.15	0.33***	0.22**	−0.11	0.10	−0.08
Soil pH						−0.07	−0.28***	−0.033	0.12	−0.14	−0.58***	−0.53***	−0.16*	0.26***	−0.25**	−0.45***	0.56***
T							−0.084	−0.41***	0.43***	0.10	−0.57***	−0.11	0.29***	0.38***	−0.15*	0.033	−0.10
$\log_{10}$ (slope gradient)								0.33***	0.03	0.29***	0.042	−0.080	0.17*	−0.061	0.51***	−0.24**	−0.37***
Elevation									−0.48***	0.15*	−0.043	−0.23**	0.074	−0.30***	0.15	0.026	−0.19*
Pop. density										−0.084	−0.17*	0.08	−0.04	0.43***	−0.037	0.021	0.18*
Runoff											−0.055	−0.30***	0.12	−0.21**	0.27***	−0.13	−0.056
Soil org. C content												0.64***	−0.25**	−0.26***	0.011	0.43***	−0.19*
% cover evergreen trees													−0.42***	−0.10	−0.16*	0.43***	−0.34***
% cover mixed trees														0.06	0.34***	−0.16*	0.0092
% cover cultivated areas															−0.052	0.012	0.15
% silt of soil																−0.13	0.0034
% sand of soil																	−0.18*

***$p < .001$, **$p < .01$, *$p < .05$, $p > .05$.

primary productivity (terrestrial NPP) (r = .35), air temperature (T, Apr–Nov monthly mean) (r = .34), population density (r = .46), the percentage of the grid covered by needle-leaved evergreen trees (GLC land cover class 4) (r = .36), and by cultivated or managed land (GLC 16) (r = .32), while highly significant negative correlations were found with $\log_{10}$lake area (A_{lake}) (r = −.51), precipitation (P, April to November monthly mean) (r = −.42), elevation (r = −.30), and runoff (r = −.33).

In order to ensure a parsimonious prediction model, we restricted the maximum number of drivers to three. All possible combinations of drivers were tested in the regression analysis. Based on this analysis, the model with both the highest r^2 and lowest Root-mean-square error (RMSE) included A_{lake}, P, and T, as predictors (Table 3). However, we chose to proceed with the second best performing model, composed of A_{lake}, terrestrial NPP, and P. This choice is justified by the more direct mechanistic link between lake pCO_2 and terrestrial NPP compared to T. T is a proxy for many different variables and controls a variety of different processes in multiple ways, meaning that T-related patterns at the spatial scale are not necessarily transferrable to the temporal scale (Weyhenmeyer et al., 2015). In addition, the allochthonous C inputs to lakes, as related to terrestrial NPP, have been shown to be a stronger constraint on CO_2 evasion from boreal lakes than the direct temperature effect on decomposition rates (Kortelainen et al., 2006). Accordingly, terrestrial NPP has been identified as a main control of CO_2 emission from lakes (Maberly et al., 2013), underpinning its strong mechanistic linkage to lake FCO_2. It is likely that the effects of terrestrial NPP are statistically represented by T and P in the training data, and that is one explanation for terrestrial NPP not being retained in the equation with the lowest RMSE. The relationships between P, T, and terrestrial NPP are, however, more complex than the statistics for the training data suggest, and the empirical relations between the three predictors in the training data are not necessarily representative of the extrapolation area or prediction periods. Thus, we assume the combination of terrestrial NPP, P, and A_{lake} to be more robust for extrapolations in space and time. Together, these predictors explain 56% (r^2 = .56, Figure 2) of the spatial variation in $\log_{10}(pCO_2)$. P has the strongest partial correlation coefficient (r = −.535) followed by A_{lake} (r = −.407) and terrestrial NPP (r = .360) (Table 4). The resulting prediction of boreal lake pCO_2 is given by the following equation:

$$\begin{aligned}\log_{10}(pCO_2\ [\mu atm]) =& 3.26 \pm 0.070 - 0.0844 \pm 0.0115 * \\ & \log_{10}(A_{lake}[km^2]) + 6.89 \pm 1.11 * \\ & NPP[10^4\ g\,C\,m^{-2}\,year^{-1}] - 8.30 \pm 0.841 * \\ & P\,[10^3 mm] (r^2 = .56).\end{aligned}$$

where the ranges (±) represent the standard errors.

Figure 2 shows the scatterplot of observed vs. predicted $\log_{10}$ (pCO_2), categorized by region. Siberia is represented by only a handful of grids with pCO_2 data. Despite this geographical bias, Figure 2 shows that there is no consistent pCO_2 prediction bias in the Siberian grids compared to those of Scandinavia and North America.

The regression model was validated with 131 of the discarded grids from Scandinavia. The resulting scatterplot (Figure 3) has a slope of 1.052 and an intercept of −0.148, suggesting that there is minimal prediction bias. Based on the multiple regression model described above, lake area weighted mean pCO_2 for the BF land cover region is estimated to be 966 μatm (Figure 4a) with an uncertainty range of 678–1,325 μatm corresponding to the 5th and 95th percentiles. The range of uncertainty was obtained from the Monte Carlo analysis, outlined in the methods. Mean pCO_2 varies inversely with lakes size category (Table 5). Indeed, our estimate of mean pCO_2 for the smallest lake size category (<0.1 km^2) of 1,558 μatm (1,110–2,208) is approximately twice that of our estimate for the largest category (>10 km^2) of 789 μatm (563–1,120).

3.2 | Estimates of FCO_2 for present day conditions

The map of FCO_2 (Figure 4c) shows a complex spatial pattern reflecting the high spatial variation of both pCO_2 (Figure 4a) and A_{lake} (Figure 4b). Integrated over the BF region, we estimated a total FCO_2 of 189 Tg C $year^{-1}$ (range of 74–347 Tg C $year^{-1}$) while for the entire 50°–70°N latitudinal band (Fig. S12), we estimated a total evasion of 272 (115–487) Tg C $year^{-1}$. Canada alone showed the

TABLE 3 The top 10 ranking multilinear regression equations composed of three drivers. Shown in descending order of ability (r^2) to predict the dependent variable $\log_{10}$ (pCO_2)

Predictors	r^2	Root-mean-square error (RMSE)
$\log_{10}\left(A_{lake}\ \left[km^2\right]\right)$, P (mm), T (°C)	.59	0.16
$\log_{10}\left(A_{lake}\ \left[km^2\right]\right)$, P (mm), terrestriaL NPP (g C m^{-2} $year^{-1}$)	.56	0.17
$\log_{10}\left(A_{lake}\ \left[km^2\right]\right)$, P (mm), Wind speed (m/s)	.53	0.18
$\log_{10}\left(A_{lake}\ \left[km^2\right]\right)$, Wind speed (m/s), T (°C)	.53	0.18
$\log_{10}\left(A_{lake}\ \left[km^2\right]\right)$, P (mm), Pop. density (Inh./km)	.52	0.18
P [mm], terrestrial NPP (g C m^{-2} $year^{-1}$), Wind speed (m/s)	.52	0.18
$\log_{10}\left(A_{lake}\ \left[km^2\right]\right)$, P (mm), Elevation (m)	.52	0.18
Wind speed (m/s), Elevation (m), $\log_{10}$ (slope gradient (°))	.51	0.18
Wind speed (m/s), $\log_{10}$ (slope gradient (°)), % cover evergreen trees	.51	0.18
$\log_{10}\left(A_{lake}\ \left[km^2\right]\right)$, Wind speed (m/s), % cover evergreen trees	.50	0.18

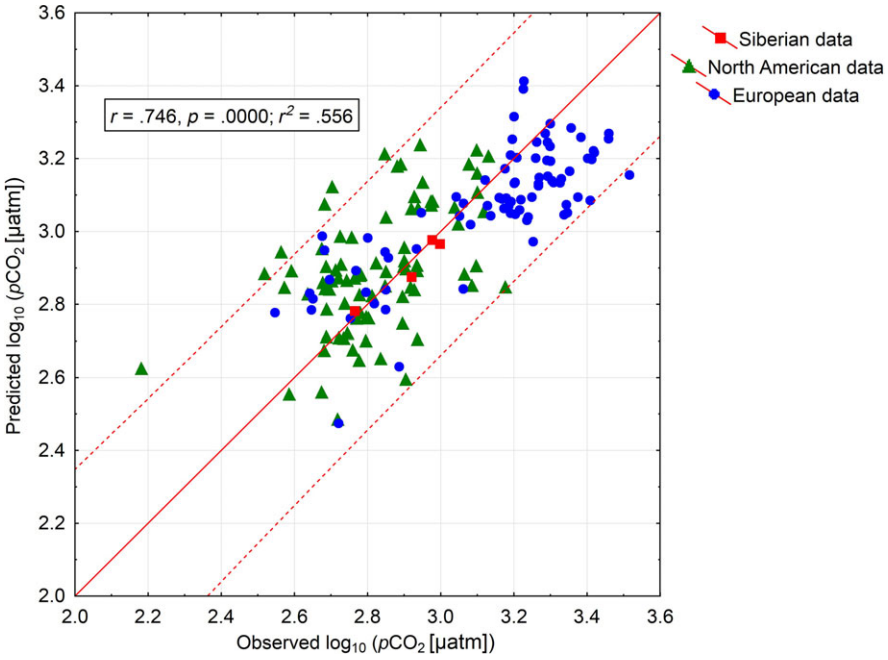

FIGURE 2 Observed vs. predicted $\log_{10}$ (pCO_2 [µatm]) for model training (168 grids) categorized by region with 95% prediction interval (dashed lines)

TABLE 4 Retained predictors with b-estimates, associated standard errors, and partial correlations to the dependent variable $\log_{10}$ (pCO_2)

Predictors	b-estimate	*SE*	*p*-value	Partial correlation
Intercept	3.26	0.0696	<.0001	
A_{lake} (km^2)	−0.0844	0.0115	<.0001	−0.407
Terrestrial NPP (10^4 g C m^{-2} $year^{-1}$)	6.89	1.11	<.0001	0.360
P (10^3 mm)	−8.30	0.841	<.0001	−0.535

highest FCO_2 with 137 (55–250) Tg C $year^{-1}$ (Table S2), mainly because Canada has the greatest total A_{lake}. We found a relatively even contribution for the four different lake size categories to total FCO_2, and this is a reflection of the contrasting relationships between lake area and pCO_2, and lake area and gas exchange velocity *k*. While the smallest lakes had the highest estimated pCO_2 values, they also had the lowest values of *k*, and this pattern is reversed in the largest lakes (Table 5).

3.3 | Projections of pCO_2 and FCO_2

We used our empirical model and projections of terrestrial NPP and *P* to estimate the change in BF lake pCO_2 and FCO_2 over the twenty-first century, under two future GHG emission scenarios, namely the lowest (RCP2.6) and highest (RCP8.5) emission scenario of the IPCC. Additionally, we used future projections of *T*, to adjust K_H and *k*, and in turn FCO_2. Based on the multimodel mean, we estimated that annual terrestrial NPP will increase by 135% from 282 to 664 g C m^{-2} $year^{-1}$ (Figure 5b) by 2100 under RCP8.5, while we estimated that *P* (April–November monthly mean) will increase by 20% from 51 to 61 mm (Figure 5d). For the years 2030 and 2050, respectively, we estimated that terrestrial NPP will increase by 49% and 67% and that *P* will increase by 5% and 10%. Under RCP2.6 (Figure 5a,c), we estimated a 45% increase in terrestrial NPP by 2050 to 408.5 g C m^{-2} $year^{-1}$ before reducing slightly to 404 g C m^{-2} $year^{-1}$ by 2100. *P* is predicted to increase by just 7.5% to 54.8 mm for both the years 2050 and 2100. In both scenarios (RCP2.6 and 8.5), and across all three time periods (2030, 2050, and 2100), all four models (CanESM2, HadGEM2-ES, IPSL-CM5A-LR, and MPIESM-LR) project an increase in both terrestrial NPP and *P* across the BF region, relative to the year 2000 (Figure 5a–d). The HadGEM2-ES model consistently projects the largest increase in terrestrial NPP while the CanESM2 model consistently projects the smallest increase. In contrast, the CanESM2 model projects the largest increase in *P* across all scenarios and time periods, with the exception of RCP2.6 for the year 2030 (for which HadGEM2-ES projects the highest *P*). In all our projections, we have assumed that A_{lake} will remain constant.

Under scenario RCP8.5, both pCO_2 (Table S3) and FCO_2 (Figure 6) are predicted to increase substantially across the BF region, peaking at 2,198 (1,303–3,761) µatm and 392 (96.5–922) Tg C $year^{-1}$, respectively, for the year 2100. This represents a 127.5% increase in pCO_2 and a 107% increase in FCO_2. For the years 2030 and 2050, respectively, we estimate that pCO_2 will increase by 31% and 48% and that FCO_2 will increase by 37% and 49%. For RCP2.6, pCO_2 and FCO_2 (Table S3; Figure 7) are estimated at 1,230 (739–1981) µatm and 260 (82.9–537) Tg C $year^{-1}$, respectively, for the year 2100. This represents a 27% increase in pCO_2 and a 38% increase in FCO_2. Note that the increases in pCO_2 and FCO_2 are not equivalent, largely due to the fact that the former is a reflection of absolute pCO_2 while the latter is calculated from delta pCO_2.

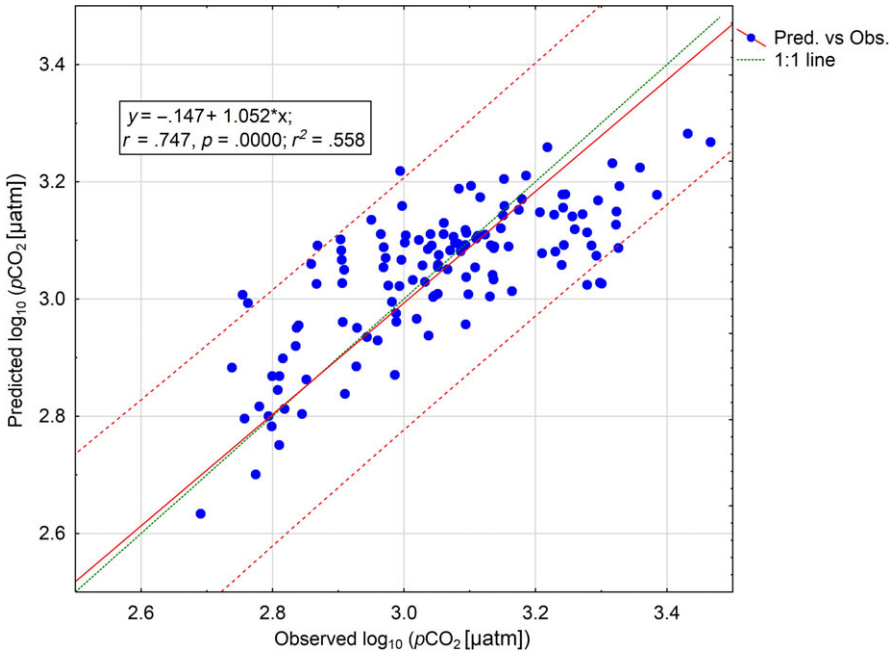

FIGURE 3 Observed vs. predicted $\log_{10}$ (pCO_2 [µatm]) for model validation (131 grids) with 95% prediction interval (dashed lines)

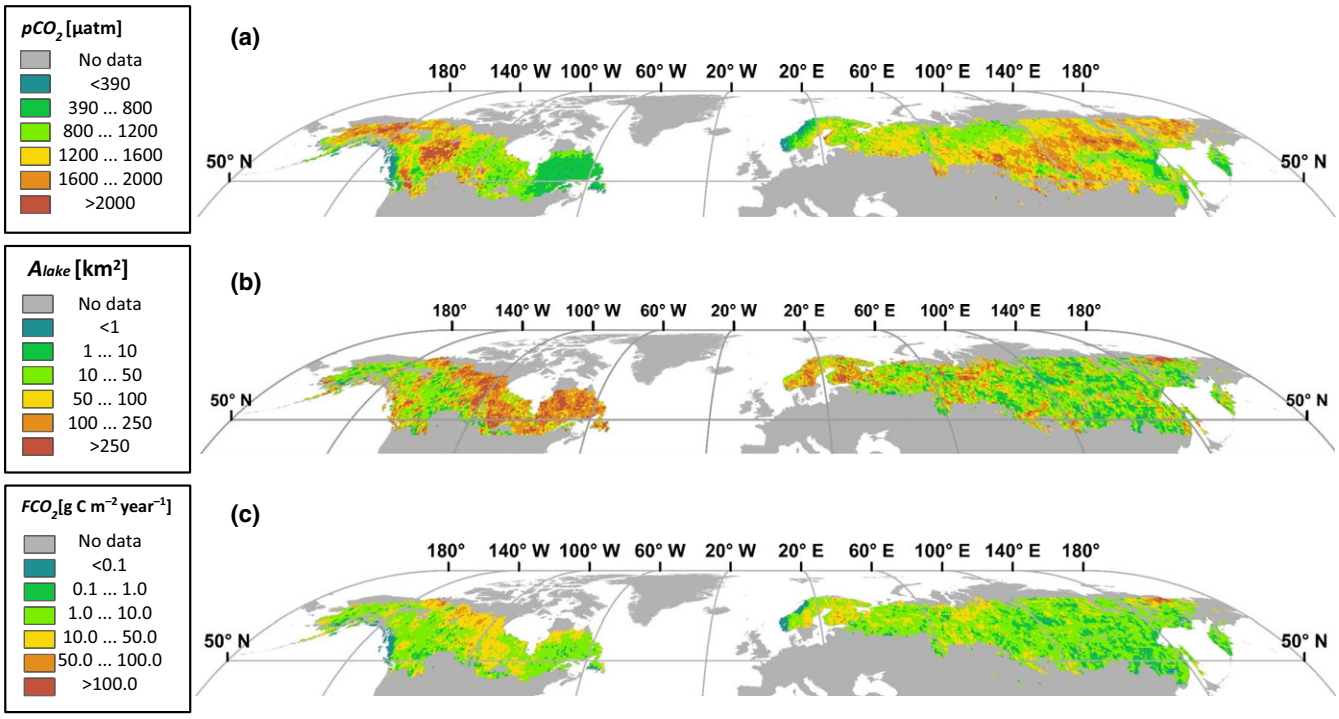

FIGURE 4 Predicted maps of (a) pCO_2, (b) A_{lake}, and (c) FCO_2 for the BF land cover region

The projected increase in lake CO_2 evasion is mainly driven by the substantial projected increase in the positive driver terrestrial NPP over the twenty-first century. In contrast, the projected increase in *P*, a negative driver of pCO_2 in our model, is very small. The smaller increase in both pCO_2 and FCO_2 predicted under the RCP2.6 scenario is reflected in the smaller increase in terrestrial NPP (Figure 5) under this scenario. By applying a linear function between our FCO_2 values for 2000, 2030, 2050, and 2100, we estimate a cumulative CO_2 evasion flux of around 29 Pg C from boreal lakes to the atmosphere over the course of the twenty-first century, under RCP8.5. Under RCP2.6, we estimate a smaller cumulative flux of 24 Pg C. See Table S1 for more detailed results of future FCO_2 projections relating to different options for the adjustment of *k* and K_H.

TABLE 5 pCO_2, FCO_2, total A_{lake}, and k values in relation to lake size classes for the BF land cover region

Lake size class (km^2)	pCO_2 (µatm)	FCO_2 (Tg C $year^{-1}$)	Total A_{lake} (km^2)	Mean k (m/day)
<0.1	1,558 (1,110–2,208)	38.17 (18.64–65.00)	208,008	0.63
0.1–1	1,237 (890–1,739)	52.36 (9.380–87.10)	282,434	0.87
1–10	1,020 (734–1,434)	44.80 (19.54–79.34)	286,624	0.99
>10	789 (563–1,120)	53.96 (9.910–116.5)	570,583	1.27

4 | DISCUSSION

4.1 | Drivers of pCO_2 and FCO_2 spatial variability

We found significant relationships between lake pCO_2 and a variety of environmental drivers (Table 2). The majority of these relationships concur with existing literature, that is, the observed positive correlation between terrestrial NPP and pCO_2 is in line with previous studies, which report on CO_2 supersaturation in boreal and temperate lakes as a result of allochthonous inputs of organic carbon (OC) and inorganic carbon (IC) from the catchment (Maberly et al., 2013; Weyhenmeyer et al., 2015; Wilkinson, Buelo, Cole, & Pace, 2016). Similarly, the presence of boreal coniferous forest has previously been shown to exert a strong positive control on pCO_2 by inducing a net heterotrophic state from elevated DOC concentrations (Chmiel et al., 2016; Hanson, Bade, Carpenter, & Kratz, 2003; Sobek, Tranvik, Prairie, Kortelainen, & Cole, 2007). There is less of a clear consensus regarding the effects of temperature and wind. Several previous studies have found pCO_2 and organic C mineralization in lakes to be strongly positively linked to temperature (Gudasz et al., 2010; Kosten et al., 2010; Marotta, Duarte, Sobek, & Enrich-Prast, 2009), while other studies have found only a weak relationship (Lapierre, Seekell, & del Giorgio, 2015; Sobek et al., 2005, 2007). In our study, temperature had a moderately positive effect on lake pCO_2, similar to that of terrestrial NPP, and indeed the two are highly intercorrelated (r = .84). At present, it is not clear in how far our observed positive relationship between T and pCO_2 is related to increased aquatic respiration or other in-lake processes or to increased terrestrial NPP. Since terrestrial NPP was almost equally powerful in explaining pCO_2 in lakes as T, we built our predictive model on terrestrial NPP to get a better mechanistic understanding.

Wind speed has previously been shown to exert a strong negative control on pCO_2 in reservoirs (Morales-Pineda, Cózar, Laiz, Úbeda, & Gálvez, 2014), the likely mechanism being higher wind speeds leading to higher FCO_2 (Cole & Caraco, 1998; Read et al., 2012; Vachon & Prairie, 2013) and in turn a decrease in pCO_2. In contrast, our study revealed a strong positive correlation between wind speed and lake pCO_2. This could be explained by higher wind speeds promoting the vertical mixing of waters, especially during summer in lakes where a thermocline has formed. Nevertheless, wind speed was not retained for the empirical model, because the

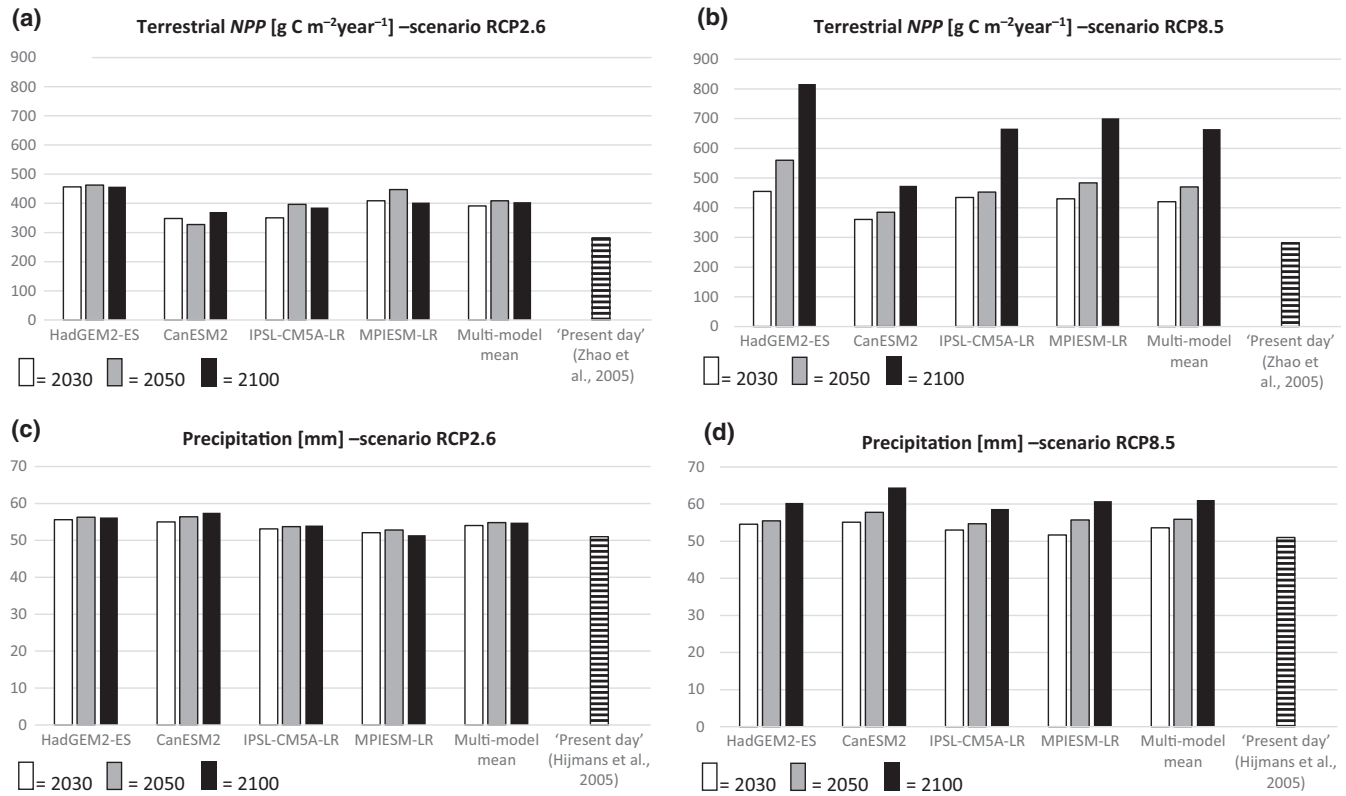

FIGURE 5 Projected BF terrestrial NPP for the years 2030, 2050 & 2100 under scenario (a) RCP2.6 (b) RCP8.5 and P under scenario (c) RCP2.6 (d) RCP8.5

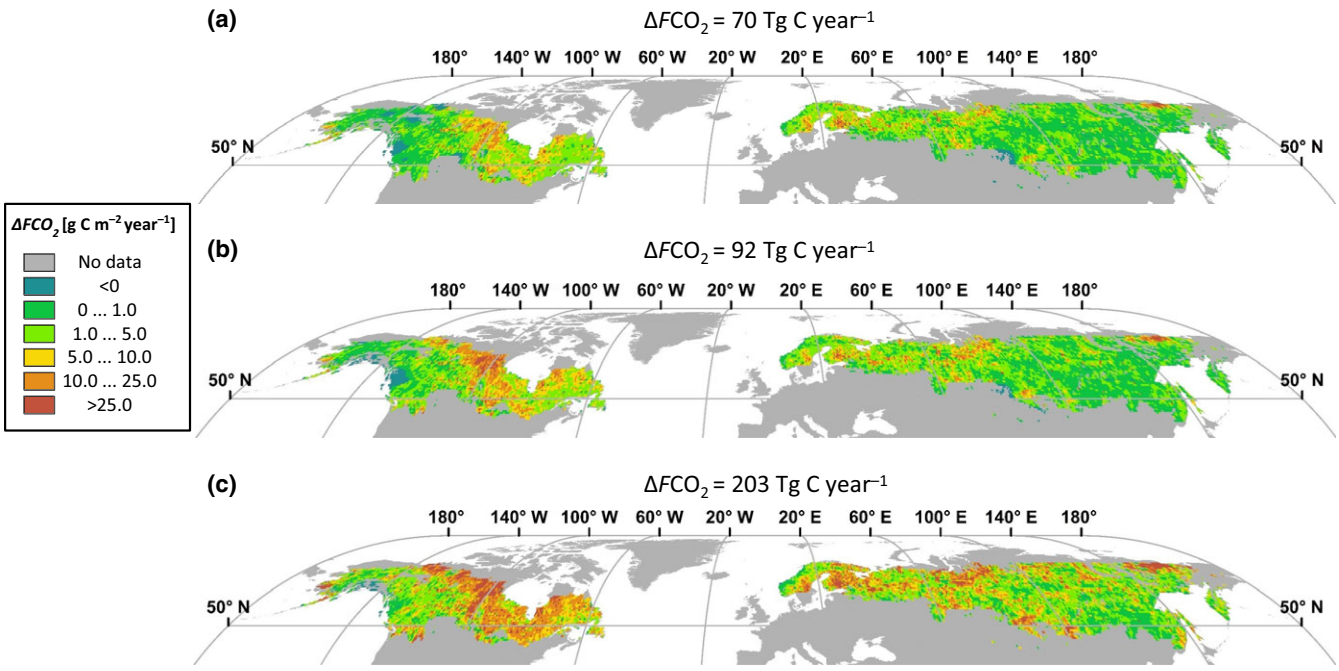

FIGURE 6 Spatially resolved (0.5°) predicted change in CO_2 evasion, ΔFCO_2 (from a year 2000 baseline) under scenario RCP8.5 for the year (a) 2030, (b) 2050, and (c) 2100 for the BF land cover region

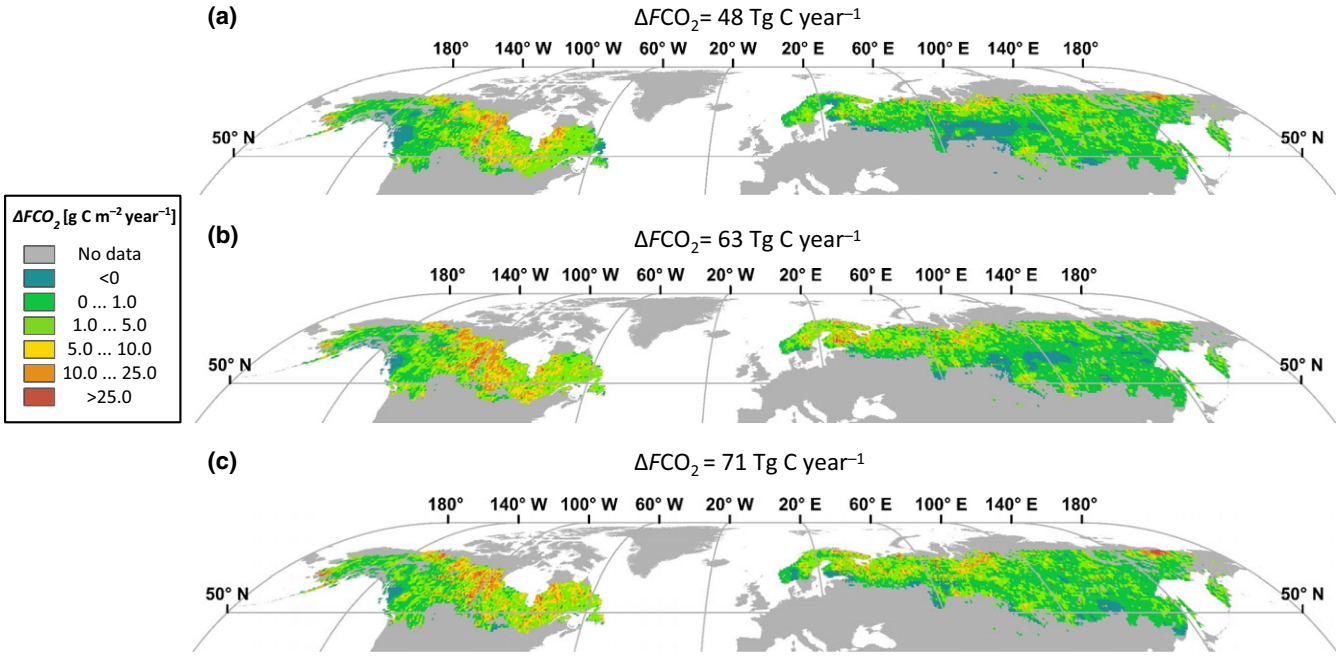

FIGURE 7 Spatially resolved (0.5°) predicted change in CO_2 evasion, ΔFCO_2 (from a year 2000 baseline) under scenario RCP2.6 for the year (a) 2030, (b) 2050, and (c) 2100 for the BF land cover region

b-estimate was associated with a very high standard error (>20% of the b-estimate).

In terms of the negative controls of pCO_2, it is well established in the literature that smaller lakes generally have higher pCO_2 values, due to their proportionately greater surface area in contact with the catchment and greater allochthonous C inputs per unit volume (Catalan, Marce, Kothawala, & Tranvik, 2016; Humborg et al., 2010; Kortelainen et al., 2006; Sobek et al., 2003). Perhaps, the most interesting result is the strong negative control of *P* on pCO_2. Although, previous large-scale studies have found a positive relationship between *P* and open water pCO_2 (Rantakari & Kortelainen, 2005; Sobek et al., 2003), a recent temporal study over a 17-year period

found a negative or no relation at all between precipitation and pCO_2 in boreal inland waters (Nydahl, Wallin, & Weyhenmeyer, 2017). Nydahl et al. (2017) suggested that increased precipitation results in a dilution of CO_2 concentrations in inland waters due to an altered balance between surface and CO_2-rich groundwater flow. In addition, *P* induced increased surface water runoff can cause a faster water flushing through the landscape giving less time for in situ CO_2 production in inland waters. It is, however, important to keep in mind that *P* is highly intercorrelated with a number of the other variables tested, most notably elevation, percentage of coniferous tree cover per grid, population density, and wind speed. As such, it may be that the relationship with *P* is also representing the effects of these environmental and physical drivers on pCO_2.

Our maps show a high degree of spatial variation and a complex pattern of pCO_2, reflecting the fact that no single driver is dominant. One region where a clear divergence in pCO_2 can be observed is Scandinavia. For instance, we estimated an average area weighted pCO_2 of 949 (637–1,345) μatm for Sweden and a substantially smaller value of 552 (372–779) μatm for Norway (Table S2), as a result of the differing topography and climate found in the two countries. Due to its close proximity to the sea and relatively high mean elevation, Norway receives a substantially greater amount of monthly precipitation (87 mm as a monthly mean over the April–November period) compared to Sweden (58 mm as a monthly mean over the April–November period), and a lower annual terrestrial NPP of 208 g C m^{-2} $year^{-1}$ compared to Sweden's total of 373 g C m^{-2} $year^{-1}$ (Fig. S14a,c). Another region where a relatively strong pattern can be seen is in Quebec, where low pCO_2 (Figure 4a) coincides with low terrestrial NPP (Fig. S14a) and relatively high precipitation (Fig. S14c). The spatial pattern in FCO_2 is even more complex, because the hotspots of pCO_2 generally do not coincide with those of A_{lake}.

4.2 | Comparison with previous global studies

In Table 6, we compare our results of FCO_2 and pCO_2 to values found in the literature, averaged across the boreal region. For an extended table with additional regional breakdowns of results, please refer to Table S2. Our estimate of total FCO_2 of 189 (74–347) Tg C $year^{-1}$ from lakes in the BF region is substantially higher (by a factor of nearly 2.5) than the estimate of 79 Tg C $year^{-1}$ proposed by Raymond et al. (2013) for the same region. Our estimate of FCO_2 over the 50°–70°N latitudinal band is also higher than the two previous estimates of Raymond et al. (103 Tg C $year^{-1}$) and Aufdenkampe et al. (110 Tg C $year^{-1}$) by a comparable factor.

There are several explanations for our relatively high estimates of FCO_2. One substantial difference in our study is the incorporation of the new GLOWABO lake database. Across the BF region, the GLOWABO database contains a total A_{lake} of 1,350,353 km^2, compared to the total of 931,619 km^2 estimated by Raymond et al. (2013). We calculated an area-specific FCO_2 of 140 g C m^{-2} $year^{-1}$, which is still 64% larger than that of Raymond et al. (2013). Indeed using Raymond's value of total A_{lake}, we would calculate a total FCO_2 of 130 Tg C $year^{-1}$. Therefore, total A_{lake} is not the only reason for the substantial difference in FCO_2 between the two studies. The greater number of the smallest, high pCO_2 lakes in GLOWABO compared to previous methods (see Verpoorter et al., 2014) is another plausible explanation for our high estimate.

In comparison with Raymond et al. (2013), we also used a substantially different methodology, as well as different data to train our model. We used additional boreal pCO_2 in the training of our model from Canada (Lapierre & del Giorgio, 2012), Sweden (Weyhenmeyer et al., 2012), and Siberia (Shirokova et al., 2013). Our methodology for estimating *k* also differed compared to previous studies. We used the same two methodologies for deriving *k* as Raymond et al. (2013) but added an additional method outlined in Vachon and Prairie (2013), which led to slightly higher *k* values.

Using only the two methods for calculating *k* (Cole & Caraco, 1998; Read et al., 2012) used in Raymond et al. (2013), we obtain a total BF evasion of 150 (67–258) Tg C $year^{-1}$, which gives an area-specific CO_2 evasion rate of 111 g C m^{-2} $year^{-1}$. If we multiply this flux density by the total A_{lake} from Raymond et al. (2013), we reach a total evasion of 104 Tg C $year^{-1}$. Thus, we conclude that the remaining discrepancy of 25 Tg C $year^{-1}$ between our results and those of Raymond et al. is due to methodological and pCO_2 data differences.

4.3 | Sources of uncertainty

4.3.1 | Upscaling

Using a statistical model to extrapolate pCO_2 in regions of minimal data coverage is suitable if the variation in the environmental parameters in the predictor equation is similar in the training areas and in

TABLE 6 pCO_2, FCO_2, total A_{lake}, and *k* values compared to previous studies

Region	pCO_2 (μatm)	FCO_2 (Tg C $year^{-1}$)	Total A_{lake} (km^2)	Mean *k* (m/day)	Source
BF	1,278 Area weighted 966 (678–1,325)	189 (74–347)	1,350,353	0.86	This study
BF	790	79	931,619	0.82	Raymond et al. (2013) (INTERPOLATED)
50°–70°N	1,305 Area weighted 1,006 (715–1,366)	272 (115–487)	1,751,985	0.88	This study
50°–70°N	812	103	1,194,701	0.84	Raymond et al. (2013) (INTERPOLATED)
50°–90°N*	1,100	110	80,000–1,650,000	0.96	Aufdenkampe et al. (2011)

the extrapolated areas. This condition is fulfilled in our study (see Figs S6–S10) where 99.6% of the variation in extrapolated terrestrial NPP and 98.8% of the variation in extrapolated *P* lies within the range recorded in the grids used for training. We are thus confident that we are not extrapolating too far beyond the statistical model boundaries. However, it is important to note that the mean values of both terrestrial NPP and *P* are substantially higher over the grids used in training the data compared to the mean values over the entire extrapolated region. For the training data, mean terrestrial NPP and *P* are 477 g C m^{-2} $year^{-1}$ and 71 mm, respectively, compared to 282 g C m^{-2} $year^{-1}$ and 51 mm across the BF land cover region.

Our mean estimated pCO_2 across the extrapolated BF region of 1,278 μatm is higher than the value of 1,133 μatm observed in our training data but more importantly, the vast majority of the variation in our extrapolated pCO_2 lies within the range of observed pCO_2. While the minimum pCO_2 of 25.6 μatm over our extrapolated grids is lower than the minimum of 152 μatm in the observed grids, the maximum value over our extrapolated grids is also lower (Figs S3 and S4), resulting in a smaller pCO_2 range over the extrapolated grids. Moreover, reducing the number of grids in our analysis from 584 to 168 could have resulted in certain geographical/climatic areas being underrepresented but grids from the vast majority of boreal latitudes remained after this edit, and therefore, we are confident that most of the variation in the original data is retained. In addition, the variation in pCO_2, as well as terrestrial NPP and *P*, is similar across the 584 grids and the 168 grids (Figs S3–S11).

In calculating annual FCO_2 across the boreal region, we multiplied our daily evasion estimates by number of days per year irrespective of location and associated ice cover duration. This choice is guided by the fact that significant CO_2 accumulation has been previously reported under ice-covered lakes and very high emissions during ice melt (Striegl et al., 2001). Such findings concur with our own preliminary analysis of the seasonality of pCO_2 at individual sampling locations, where peak pCO_2 values were often measured during spring before April and at temperatures below 4°C. These conditions were excluded from our analysis as we restricted our dataset to samples measured at a water temperature greater than 4°C and between the months of April to November. Moreover, a disproportionate percentage (45%) of our raw data was sampled during the summer (July–September) and this data had a median value of 997 μatm. We can compare this to the spring (April–June) data with a median value of 1,416 μatm, which comprised just 20% of our data, or the annual median value of 1,478 μatm (all data samples including winter and <4°C), and conclude that the data used in our final analysis likely leads to a conservative estimate of pCO_2. This choice compensates somewhat for the lack of accounting for variable ice cover duration in our estimation of FCO_2. As discussed in the methods, we only included pCO_2 data with a pH of ≤5.4 in order to filter out unreliable data. However, this could also lead to underestimation of pCO_2. Note that Raymond et al. (2013) used the same approach, meaning that the results from both studies can be compared.

4.3.2 | Lake area

There are a number of limitations associated with GLOWABO. Despite the use of a number of filters to minimize errors, some false detection of lakes due to cloud and mountain shadow is unavoidable. Other sources of errors include the elucidation of lakes from large rivers and wetlands.

While these limitations are significant, validation of GLOWABO against a high-resolution map of Sweden (Verpoorter et al., 2012), an area which encounters all of the aforementioned problems, achieved a performance index of 91% for lake area, while lake number differed by less than 3% (see Verpoorter et al., 2012 for further discussion).

4.3.3 | Gas exchange velocity k

Gas exchange velocity *k* represents one of the largest sources of uncertainty. We assessed this uncertainty by using three different methods to calculate *k* and we reported a best estimate as the average of these three *k* quantification methods. We further accounted for this uncertainty by incorporating *k* into the Monte Carlo analysis. Additionally, in order to assess the uncertainty associated with the variation in *k* alone, we undertook an extra Monte Carlo analysis in which we only accounted for variation in *k* (i.e., uncertainty associated with pCO_2 calculation was excluded). Based on this analysis, the mean FCO_2 is 185 Tg C $year^{-1}$ for the BF region, very close to the 189 Tg C $year^{-1}$ estimated in the original Monte Carlo analysis. The range of uncertainty is only moderately smaller; we estimate 5th and 95th percentile FCO_2 at 98 and 297 Tg C $year^{-1}$, respectively, compared to the original range of 74–347 Tg C $year^{-1}$. Thus, we conclude that *k* is indeed the largest source of uncertainty in our calculation of FCO_2.

4.3.4 | Future changes in lake CO_2 evasion

Our study does not account for future changes in the extent of the boreal forest, predicted as a result of increasing temperature (Gauthier et al., 2015; Koven, 2013). However, recent research estimating future changes in the boreal C stock under scenario RCP4.5, suggests that any C gained from northern expansion of the boreal forest is likely to be offset by loss from southern boreal retreat, and thus little net change is predicted (Gauthier et al., 2015). Finally, our study does not account for the future impact of permafrost thaw, which will become an increasingly important source of C. Drake, Wickland, Spencer, McKnight, and Striegl (2015) report that by 2100, between 5 and 10 Tg, C will be released annually from Yedoma soils alone.

4.4 | Present and future carbon budget for the boreal region

This study is the first to spatially resolve lake pCO_2 and FCO_2 across the boreal region, moreover, using only environmental drivers

derived from freely available geodata. The resolution of our maps is compatible with most global land surface and inversion models (Ciais et al., 2013), and thus could potentially be used for validation purposes. High-resolution estimates of C fluxes such as those reported here are crucial in deriving more reliable regional C budgets, particularly along the land-ocean aquatic continuum (LOAC) where large uncertainties remain (Regnier et al., 2013). Figure 8 integrates our results within a C budget for the boreal region using previous spatially resolved estimates of terrestrial NPP (Zhao et al., 2005), FCO_2 in rivers (Lauerwald et al., 2015), C burial in lake sediments (Heathcote et al., 2015), lateral C exports to the ocean (Mayorga et al., 2010), C accumulation in forests (Pan et al., 2011), and emissions from fires (van der Werf et al., 2017; in review). Our updated budget suggests that lakes are the most significant contributor to the LOAC budget. This is largely due to their substantially greater surface area; we estimate that it is 11 times that of rivers in the BF region. Moreover, we estimate that in the order of 3%–5% of the C fixed by terrestrial vegetation (terrestrial NPP) is leaking each year into inland water bodies. This value is comparable to the global estimate of 3.2% proposed by Regnier et al. (2013) and the 5% recently calculated for the Amazon basin by Lauerwald et al. (2017), which ignores the lateral mobilization of POC. Interestingly, the magnitude of the LOAC C flux is of the same order as the mean C storage in the boreal forest biomass and soils combined. It is also greater than the vertical flux as a result of boreal forest fires (van der Werf et al., 2017; in review) and the lateral C flux from harvested wood (Pan et al., 2011). Our findings imply that the leakage through the LOAC considerably reduces the C accumulation in boreal forests. This could particularly be true for Canada where our estimate of lake FCO_2 alone of 137 Tg C $year^{-1}$ is substantially larger than the mean (1990–2007) C storage of 20 Tg C $year^{-1}$ for the Canadian boreal forested proposed by Pan et al. (2011). Our budget is also likely to be conservative given that we do not account for methane (CH_4) fluxes. A recent study (Rasilo, Prairie, & del Giorgio, 2015) of 224 lakes in Quebec found that as much as 25% of the emissions from lakes, in terms of atmospheric warming potential, are in the form of CH_4. Moreover, there are a small number of additional C fluxes contributing to the net ecosystem exchange budget, such as emissions associated with the consumption of crop products, which we do not include but are of relatively minor importance (Ciais et al., in review).

We estimate that lake pCO_2 and FCO_2 will increase substantially over the 21st century relative to our present day estimates. Under RCP8.5, we predict a 37%, 49%, and 107% increase in boreal lake FCO_2 by 2030, 2050, and 2100, respectively, amounting to a cumulative perturbation of the lake to atmosphere CO_2 flux of about 9 Pg C over the twenty-first century. This is a significant perturbation, of a similar magnitude to predicted future changes in boreal soil organic C stocks in some land C models (Nishina et al., 2014). Our projections are largely driven by increases in terrestrial NPP of 46%, 67%, and 135% over the same period. Interestingly, even under the GHG

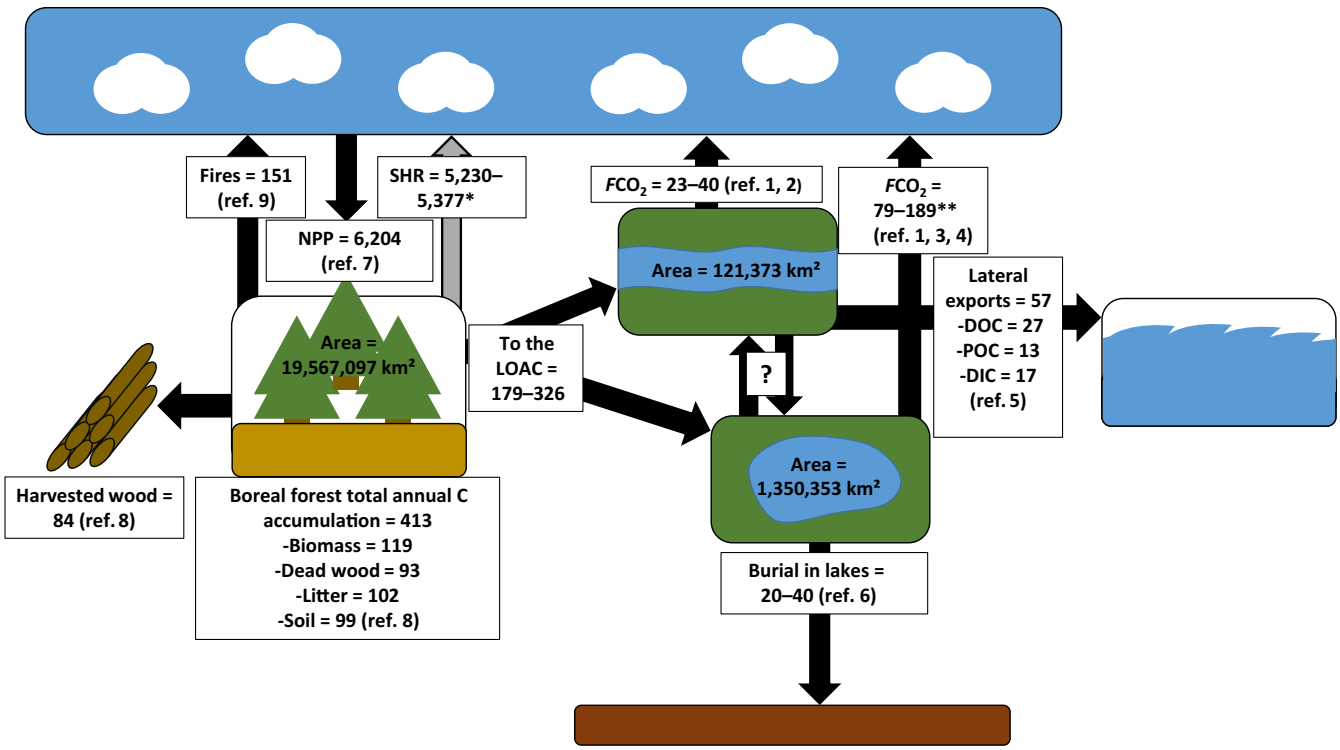

FIGURE 8 Updated carbon budget along the land-ocean aquatic continuum (LOAC) for the boreal region. Units are Tg C $year^{-1}$. Ref. 1 – Aufdenkampe et al. (2011), ref. 2 – Lauerwald et al. (2015), ref. 3 – Raymond et al. (2013), ref. 4** (this study), ref. 5 – Mayorga et al. (2010), ref. 6 – Heathcote et al. (2015), ref. 7 – Zhao et al. (2005), ref. 8 – Pan et al. (2011), and ref. 9– van der Werf et al. (2017, in review). *SHR (soil heterotrophic respiration) is derived from budget closure. This scheme does not include estuarine C fluxes which are relatively minor in this region (Laruelle et al., 2013), or the C fluxes between lakes and rivers for which no estimate could be found for the boreal region

scenario RCP2.6, we predict a 25%, 33%, and 38% increase in boreal lake FCO_2 by 2030, 2050, and 2100, respectively. This suggests that a substantial strengthening of the CO_2 evasion flux from boreal lakes is expected irrespective of the emission scenario. Our results concur with those of Larsen, Andersen, and Hessen (2011), which projected a 65% increase in TOC concentration in Norwegian lakes by 2100 under the superseded IPCC B2 scenario, an intermediate GHG emission scenario. In our study, NPP increases at an equivalent rate while the increase in precipitation is much smaller, meaning that the proportion of NPP lost from lakes to the atmosphere remains relatively constant at approximately 3% under both scenarios. Finally, our estimates of future FCO_2 are likely to be conservative, due to our lack of accounting for the impact of permafrost thaw on remobilizing, old labile C. Accounting for this substantial source of future C should be prioritized in future studies of boreal and high latitude regions, and may require the explicit representation of these processes in mechanistic Earth System models.

ACKNOWLEDGEMENTS

Financial support was received from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 643052 (C-CASCADES project). RL acknowledges funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 703813 for the Marie Sklodowska-Curie European Individual Fellowship "C-Leak." GW acknowledges funding from the Swedish Research Council (Grant No. 2016-04153) and the Knut and Alice Wallenberg Foundation (KAW project). The research leading to these results has received additional funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no 336642 to SS. Many thanks to Jean-François Lapierre and Paul del Giorgio for providing us with Canadian lake pCO_2 data. We are also grateful to the very constructive comments of two reviewers.

ORCID

Adam Hastie http://orcid.org/0000-0003-2098-3510

REFERENCES

Abril, G., Bouillon, S., Darchambeau, F., Teodoru, C. R., Marwick, T. R., Tamooh, F., … Borges, A. V. (2015). Technical note: Large overestimation of pCO_2 calculated from pH and alkalinity in acidic, organic-rich freshwaters. *Biogeosciences*, *12*(1), 67–78. https://doi.org/10.5194/bg-12-67-2015

Aufdenkampe, A. K., Mayorga, E., Raymond, P. A., Melack, J. M., Doney, S. C., Alin, S. R., … Yoo, K. (2011). Riverine coupling of biogeochemical cycles between land, oceans, and atmosphere. *Frontiers in Ecology and the Environment*, *9*(1), 53–60. https://doi.org/10.1890/100014

Butman, D., Stackpoole, S., Stets, E., McDonald, C. P., Clow, D. W., & Striegl, R. G. (2016). Aquatic carbon cycling in the conterminous United States and implications for terrestrial carbon accounting. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(1), 58–63. https://doi.org/10.1073/pnas.1512651112

Catalan, N., Marce, R., Kothawala, D. N., & Tranvik, L. J. (2016). Organic carbon decomposition rates controlled by water retention time across inland waters. *Nature Geosci*, *9*(7), 501–504. https://doi.org/10.1038/ngeo2720

Chmiel, H. E., Kokic, J., Denfeld, B. A., Einarsdóttir, K., Wallin, M. B., Koehler, B., & Sobek, S. (2016). The role of sediments in the carbon budget of a small boreal lake. *Limnology and Oceanography*, *61*(5), 1814–1825. https://doi.org/10.1002/lno.10336

Ciais, P., Gasser, T., Lauerwald, R., Peng, S., Raymond, P. A., Wang, Y., … Zhu, D. (2017). Observed regional carbon budgets imply reduced soil heterotrophic respiration. *Nature*, in review.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., … (2013). Carbon and Other Biogeochemical Cycles. In T. F.. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley (Eds.), *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change* (pp. 465–570). Cambridge University Press, Cambridge.

CIESIN, and CIAT. (2005). Gridded population of the world version 3 (GPWv3): POPULATION Grids CIESIN, Columbia University. New York, Palisades, N. Y. https://doi.org/10.7927/h4639mpp

Cole, J. J., & Caraco, N. F. (1998). Atmospheric exchange of carbon dioxide in a low-wind oligotrophic lake measured by the addition of SF6. *Limnology and Oceanography*, *43*(4), 647–656. https://doi.org/10.4319/lo.1998.43.4.0647

Cole, J. J., Caraco, N. F., Kling, G. W., & Kratz, T. K. (1994). Carbon dioxide supersaturation in the surface waters of lakes. *Science*, *265*(5178), 1568–1570. https://doi.org/10.1126/science.265.5178.1568

Cole, J. J., Prairie, Y. T., Caraco, N. F., McDowell, W. H., Tranvik, L. J., Striegl, R. G., & Melack, J. (2007). Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon budget. *Ecosystems*, *10*(1), 172–185. https://doi.org/10.1007/s10021-006-9013-8

Denfeld, B. A., Kortelainen, P., Rantakari, M., Sobek, S., & Weyhenmeyer, G. A. (2016). Regional variability and drivers of below ice CO_2 in boreal and subarctic lakes. *Ecosystems*, *19*(3), 461–476. https://doi.org/10.1007/s10021-015-9944-z

Drake, T. W., Wickland, K. P., Spencer, R. G. M., McKnight, D. M., & Striegl, R. G. (2015). Ancient low–molecular-weight organic acids in permafrost fuel rapid carbon dioxide production upon thaw. *Proceedings of the National Academy of Sciences*, *112*(45), 13946–13951. https://doi.org/10.1073/pnas.1511705112

FAO/IIASA/ISRIC/ISSCAS/JRC. (2012). *Harmonized World Soil Database* (version 1.2). Rome, Italy: FAO and Laxenburg, Austria: IIASA.

Fekete, B. M., Vörösmarty, C. J., & Grabs, W. (2002). High-resolution fields of global runoff combining observed river discharge and simulated water balances. *Global Biogeochemical Cycles*, *16*(3), 10–15. https://doi.org/10.1029/1999GB001254

Gauthier, S., Bernier, P., Kuuluvainen, T., Shvidenko, A. Z., & Schepaschenko, D. G. (2015). Boreal forest health and global change. *Science*, *349*(6250), 819–822. https://doi.org/10.1126/science.aaa9092

Global Land Cover 2000 database. (2003). European Commission, Joint Research Centre, 2003. http://bioval.jrc.ec.europa.eu/glc2000/glc2000.php.

GLOBE Task Team and others (Hastings, David A., Paula K. Dunbar, Gerald M. Elphingstone, Mark Bootz, Hiroshi Murakami, Hiroshi Maruyama, Hiroshi Masaharu, Peter Holland, John Payne, Nevin A. Bryant, Thomas L. Logan, J.-P. Muller, Gunter Schreier, and John S. MacDonald), eds. (1999). The Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, Version 1.0. National Oceanic and Atmospheric Administration, National Geophysical Data Center, 325 Broadway, Boulder, Colorado 80305-3328, U.S.A. Digital data base on the World Wide Web (http://www.ngdc.noaa.gov/mgg/topo/globe.html) and CD-ROMs.

Gudasz, C., Bastviken, D., Steger, K., Premke, K., Sobek, S., & Tranvik, L. J. (2010). Temperature-controlled organic carbon mineralization in lake sediments. *Nature*, *466*(7305), 478–481. https://doi.org/10.1038/nature09186

Hanson, P. C., Bade, D. L., Carpenter, S. R., & Kratz, T. K. (2003). Lake metabolism: Relationships with dissolved organic carbon and phosphorus. *Limnology and Oceanography*, *48*(3), 1112–1119. https://doi.org/10.4319/lo.2003.48.3.1112

Heathcote, A. J., Anderson, N. J., Prairie, Y. T., Engstrom, D. R., & del Giorgio, P. A. (2015). Large increases in carbon burial in northern lakes during the Anthropocene. *Nature Communications*, *6*, 10016. https://doi.org/10.1038/ncomms10016

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978. https://doi.org/10.1002/joc.1276

Holgerson, M. A., & Raymond, P. A. (2016). Large contribution to inland water CO_2 and CH_4 emissions from very small ponds. *Nature Geosci*, *9*(3), 222–226. https://doi.org/10.1038/ngeo2654

Humborg, C., Mörth, C.-M., Sundbom, M., Borg, H., Blenckner, T., Giesler, R., & Ittekkot, V. (2010). CO_2 supersaturation along the aquatic conduit in Swedish watersheds as constrained by terrestrial respiration, aquatic respiration and weathering. *Global Change Biology*, *16*(7), 1966–1978. https://doi.org/10.1111/j.1365-2486.2009.02092.x

Intergovernmental Panel on Climate Change (IPCC), Climate Change. (2013). The physical science basis. In T. F. Stocker, et al. (Eds.), *Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge: Cambridge University Press.

Kortelainen, P., Rantakari, M., Huttunen, J. T., Mattsson, T., Alm, J., Juutinen, S., … Martikainen, P. J. (2006). Sediment respiration and lake trophic state are important predictors of large CO_2 evasion from small boreal lakes. *Global Change Biology*, *12*(8), 1554–1567. https://doi.org/10.1111/j.1365-2486.2006.01167.x

Kortelainen, P., Rantakari, M., Pajunen, H., Huttunen, J. T., Mattsson, T., Juutinen, S., … Martikainen, P. J. (2013). Carbon evasion/accumulation ratio in boreal lakes is linked to nitrogen. *Global Biogeochemical Cycles*, *27*(2), 363–374. https://doi.org/10.1002/gbc.20036

Kosten, S., Roland, F., Da Motta Marques, D. M. L., Van Nes, E. H., Mazzeo, N., Sternberg, L. D. S., Scheffer, M., … Cole, J. J. (2010). Climate-dependent CO_2 emissions from lakes. *Global Biogeochemical Cycles*, *24*(2). https://doi.org/10.1029/2009GB003618

Koven, C. D. (2013). Boreal carbon loss due to poleward shift in low-carbon ecosystems. *Nature Geosci*, *6*(6), 452–456. https://doi.org/10.1038/ngeo1801

Lapierre, J.-F., & del Giorgio, P. A. (2012). Geographical and environmental drivers of regional differences in the lake pCO_2 versus DOC relationship across northern landscapes. *Journal of Geophysical Research: Biogeosciences*, *117*(G3). https://doi.org/10.1029/2012JG001945

Lapierre, J.-F., Seekell, D. A., & del Giorgio, P. A. (2015). Climate and landscape influence on indicators of lake carbon cycling through spatial patterns in dissolved organic carbon. *Global Change Biology*, *21*(12), 4425–4435. https://doi.org/10.1111/gcb.13031

Larsen, S., Andersen, T. O. M., & Hessen, D. A. G. O. (2011). Climate change predicted to cause severe increase of organic carbon in lakes. *Global Change Biology*, *17*(2), 1186–1192. https://doi.org/10.1111/j.1365-2486.2010.02257.x

Laruelle, G. G., Dürr, H. H., Lauerwald, R., Hartmann, J., Slomp, C. P., Goossens, N., & Regnier, P. A. G. (2013). Global multi-scale segmentation of continental and coastal waters from the watersheds to the continental margins. *Hydrology and Earth System Sciences*, *17*(5), 2029–2051. https://doi.org/10.5194/hess-17-2029-2013

Lauerwald, R., Hartmann, J., Moosdorf, N., Kempe, S., & Raymond, P. A. (2013). What controls the spatial patterns of the riverine carbonate system?—a case study for North America. *Chemical Geology*, *337–338*, 114–127. https://doi.org/10.1016/j.chemgeo.2012.11.011

Lauerwald, R., Laruelle, G. G., Hartmann, J., Ciais, P., & Regnier, P. A. G. (2015). Spatial patterns in CO_2 evasion from the global river network. *Global Biogeochemical Cycles*, *29*(5), 534–554. https://doi.org/10.1002/2014GB004941

Lauerwald, R., Regnier, P., Camino-Serrano, M., Guenet, B., Guimberteau, M., Ducharne, A., & Ciais, P. (2017). ORCHILEAK: A new model branch to simulate carbon transfers along the terrestrial-aquatic continuum of the Amazon basin. *Geoscientific Model Development Discuss.*, 1–58. https://doi.org/10.5194/gmd-2017-79

Maberly, S. C., Barker, P. A., Stott, A. W., & De Ville, M. M. (2013). Catchment productivity controls CO_2 emissions from lakes. *Nature Climate Change*, *3*(4), 391–394. http://dx.doi.org/10.1038/nclimate1748

MacIntyre, S., Jonsson, A., Jansson, M., Aberg, J., Turney, D. E., & Miller, S. D. (2010). Buoyancy flux, turbulence, and the gas transfer coefficient in a stratified lake. *Geophysical Research Letters*, *37*(24). https://doi.org/10.1029/2010GL044164

Marotta, H., Duarte, C. M., Sobek, S., & Enrich-Prast, A. (2009). Large CO_2 disequilibria in tropical lakes. *Global Biogeochemical Cycles*, *23*(4). https://doi.org/10.1029/2008GB003434

Mayorga, E., Seitzinger, S. P., Harrison, J. A., Dumont, E., Beusen, A. H. W., Bouwman, A. F., … Van Drecht, G. (2010). Global Nutrient Export from WaterSheds 2 (NEWS 2): Model development and implementation. *Environmental Modelling & Software*, *25*(7), 837–853. https://doi.org/10.1016/j.envsoft.2010.01.007

McDonald, C. P., Stets, E. G., Striegl, R. G., & Butman, D. (2013). Inorganic carbon loading as a primary driver of dissolved carbon dioxide concentrations in the lakes and reservoirs of the contiguous United States. *Global Biogeochemical Cycles*, *27*(2), 285–295. https://doi.org/10.1002/gbc.20032

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., … van Vuuren, D. P. P. (2011). The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, *109*(1), 213. https://doi.org/10.1007/s10584-011-0156-z

Meybeck, M., Dürr, H. H., & Vörösmarty, C. J. (2006). Global coastal segmentation and its river catchment contributors: A new look at land-ocean linkage. *Global Biogeochemical Cycles*, *20*(1). https://doi.org/10.1029/2005gb002540

Morales-Pineda, M., Cózar, A., Laiz, I., Úbeda, B., & Gálvez, J. Á. (2014). Daily, biweekly, and seasonal temporal scales of pCO_2 variability in two stratified Mediterranean reservoirs. *Journal of Geophysical Research: Biogeosciences*, *119*(4), 509–520. https://doi.org/10.1002/2013JG002317

Nishina, K., Ito, A., Beerling, D. J., Cadule, P., Ciais, P., Clark, D. B., … Yokohata, T. (2014). Quantifying uncertainties in soil carbon responses to changes in global mean temperature and precipitation. *Earth System Dynamics*, *5*(1), 197–209. https://doi.org/10.5194/esd-5-197-2014

Nydahl, A. C., Wallin, M. B., & Weyhenmeyer, G. A. (2017). No long-term trends in *p*CO2 despite increasing organic carbon concentrations in boreal lakes, streams, and rivers. *Global Biogeochemical Cycles*, *31*(6), 985–995. https://doi.org/10.1002/2016GB005539

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., … Hayes, D. (2011). A large and persistent carbon sink in the world's Forests. *Science*, *333*(6045), 988–993. https://doi.org/10.1126/science.1201609

Perga, M.-E., Maberly, S. C., Jenny, J.-P., Alric, B., Pignol, C., & Naffrechoux, E. (2016). A century of human-driven changes in the carbon dioxide concentration of lakes. *Global Biogeochemical Cycles*, *30*(2), 93–104. https://doi.org/10.1002/2015GB005286

Potapov, P., Hansen, M. C., Stehman, S. V., Loveland, T. R., & Pittman, K. (2008). Combining {MODIS} and Landsat imagery to estimate and

map boreal forest cover loss. *Remote Sensing of Environment*, *112*(9), 3708–3719. https://doi.org/10.1016/j.rse.2008.05.006

Price, D. T., McKenney, D. W., Joyce, L. A., Siltanen, R. M., Papadopol, P., & Lawrence, K. (2011). High-resolution interpolation of climate scenarios for canada derived from general circulation model simulations. Nat. Resour. Can., Can. For. Serv., North. For. Cent., Edmonton, AB.

R Core Team. (2013). R: A language and environment for statistical computing. [Available at http://www.r-project.org.]

Rantakari, M., & Kortelainen, P. (2005). Interannual variation and climatic regulation of the CO_2 emission from large boreal lakes. *Global Change Biology*, *11*(8), 1368–1380. https://doi.org/10.1111/j.1365-2486.2005.00982.x

Rasilo, T., Prairie, Y. T., & del Giorgio, P. A. (2015). Large-scale patterns in summer diffusive CH4 fluxes across boreal lakes, and contribution to diffusive C emissions. *Global Change Biology*, *21*(3), 1124–1139. https://doi.org/10.1111/gcb.12741

Raymond, P. A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., … Guth, P. (2013). Global carbon dioxide emissions from inland waters. *Nature*, *503*(7476), 355–359. Retrieved from https://doi.org/10.1038/nature12760

Read, J. S., Hamilton, D. P., Desai, A. R., Rose, K. C., MacIntyre, S., Lenters, J. D., … Wu, C. H. (2012). Lake-size dependency of wind shear and convection as controls on gas exchange. *Geophysical Research Letters*, *39*(9). https://doi.org/10.1029/2012GL051886

Regnier, P., Friedlingstein, P., Ciais, P., Mackenzie, F. T., Gruber, N., Janssens, I. A., … Thullner, M. (2013). Anthropogenic perturbation of the carbon fluxes from land to ocean. *Nature Geosci*, *6*(8), 597–607. https://doi.org/10.1038/ngeo1830

Roehm, C. L., Prairie, Y. T., & del Giorgio, P. A. (2009). The pCO_2 dynamics in lakes in the boreal region of northern Québec, Canada. *Global Biogeochemical Cycles*, *23*(3). https://doi.org/10.1029/2008GB003297

Shirokova, L. S., Pokrovsky, O. S., Kirpotin, S. N., Desmukh, C., Pokrovsky, B. G., Audry, S., & Viers, J. (2013). Biogeochemistry of organic carbon, CO_2, CH_4, and trace elements in thermokarst water bodies in discontinuous permafrost zones of Western Siberia. *Biogeochemistry*, *113*(1), 573–593. https://doi.org/10.1007/s10533-012-9790-4

Sobek, S., Algesten, G., Bergström, A.-K., Jansson, M., & Tranvik, L. J. (2003). The catchment and climate regulation of pCO_2 in boreal lakes. *Global Change Biology*, *9*(4), 630–641. https://doi.org/10.1046/j.1365-2486.2003.00619.x

Sobek, S., Tranvik, L. J., & Cole, J. J. (2005). Temperature independence of carbon dioxide supersaturation in global lakes. *Global Biogeochemical Cycles*, *19*(2). https://doi.org/10.1029/2004GB002264

Sobek, S., Tranvik, L. J., Prairie, Y. T., Kortelainen, P., & Cole, J. J. (2007). Patterns and regulation of dissolved organic carbon: An analysis of 7,500 widely distributed lakes. *Limnology and Oceanography*, *52*(3), 1208–1219. https://doi.org/10.4319/lo.2007.52.3.1208

SoilGrids. (2014). http://soilgrids.org/index.html

Striegl, R. G., Kortelainen, P., Chanton, J. P., Wickland, K. P., Bugna, G. C., & Rantakari, M. (2001). Carbon dioxide partial pressure and 13C content of north temperate and boreal lakes at spring ice melt. *Limnology and Oceanography*, *46*(4), 941–945. https://doi.org/10.4319/lo.2001.46.4.0941

Telmer, K., & Veizer, J. (1999). Carbon fluxes, pCO_2 and substrate weathering in a large northern river basin, Canada: Carbon isotope perspectives. *Chemical Geology*, *159*(1–4), 61–86. https://doi.org/10.1016/s0009-2541(99)00034-0

Tranvik, L. J., Downing, J. A., Cotner, J. B., Loiselle, S. A., Striegl, R. G., Ballatore, T. J., … Weyhenmeyer, G. A. (2009). Lakes and reservoirs as regulators of carbon cycling and climate. *Limnology and Oceanography*, *54*(6 part 2), 2298–2314. https://doi.org/10.4319/lo.2009.54.6_part_2.2298

Verpoorter, C., Kutser, T., Seekell, D. A., & Tranvik, L. J. (2014). A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters*, *41*(18), 6396–6402. https://doi.org/10.1002/2014GL060641

Verpoorter, C., Kutser, T., & Tranvik, L. (2012). Automated mapping of water bodies using Landsat multispectral data. *Limnology and Oceanography: Methods*, *10*(12), 1037–1050. https://doi.org/10.4319/lom.2012.10.1037

Wallin, M. B., Löfgren, S., Erlandsson, M., & Bishop, K. (2014). Representative regional sampling of carbon dioxide and methane concentrations in hemiboreal headwater streams reveal underestimates in less systematic approaches. *Global Biogeochemical Cycles*, *28*(4), 465–479. https://doi.org/10.1002/2013GB004715

Vachon, D., & Prairie, Y. T. (2013). The ecosystem size and shape dependence of gas transfer velocity versus wind speed relationships in lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, *70*(12), 1757–1764. https://doi.org/10.1139/cjfas-2013-0241

van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., … Kasibhatla, P. S. (2017). Global fire emissions estimates during 1997–2015. *Earth System Science Data Discussion*, *2017*, 1–43. https://doi.org/10.5194/essd-2016-62

Weyhenmeyer, G. A., Kortelainen, P., Sobek, S., Muller, R., & Rantakari, M. (2012). Carbon dioxide in boreal surface waters: A comparison of lakes and streams. *Ecosystems*, *15*(8), 1295–1307. http://www.jstor.org/stable/23325685

Weyhenmeyer, G. A., Kosten, S., Wallin, M. B., Tranvik, L. J., Jeppesen, E., & Roland, F. (2015). Significant fraction of CO_2 emissions from boreal lakes derived from hydrologic inorganic carbon inputs. *Nature Geoscience*, *8*(12), 933–936. https://doi.org/10.1038/ngeo2582

Wilkinson, G. M., Buelo, C. D., Cole, J. J., & Pace, M. L. (2016). Exogenously produced CO_2 doubles the CO_2 efflux from three north temperate lakes. *Geophysical Research Letters*, *43*(5), 1996–2003. https://doi.org/10.1002/2016GL067732

Zhao, M., Heinsch, F. A., Nemani, R. R., & Running, S. W. (2005). Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sensing of Environment*, *95*(2), 164–176. https://doi.org/10.1016/j.rse.2004.12.011

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Hastie A, Lauerwald R, Weyhenmeyer G, Sobek S, Verpoorter C, Regnier P. CO_2 evasion from boreal lakes: Revised estimate, drivers of spatial variability, and future projections. *Glob Change Biol*. 2017;00:1–19. https://doi.org/10.1111/gcb.13902