

Generalized additive models of climatic and metabolic controls of sub-annual variation in pCO₂ in productive hardwater lakes

E. Wiik^{1,2}, H. A. Haig¹, N. M. Hayes¹, K. Finlay¹, G. L. Simpson³, R. J. Vogt⁴, P. R. Leavitt^{1,3}

¹Limnology Laboratory, Department of Biology, University of Regina, Regina, Saskatchewan, S4S 0A2, Canada

²School of Environment, Natural Resources and Geography, Bangor University, Bangor, Gwynedd, LL57 2UW, UK

³Institute of Environmental Change and Society, University of Regina, Regina, Saskatchewan, S4S 0A2, Canada

⁴Pavillon des sciences biologiques (SB), Université du Québec à Montréal, Montréal (Québec), H2X 1Y4, Canada

Key Points:

- In Canadian hardwater prairie lakes, calculated CO₂ fluxes correlate mostly with pH, not DIC
- Intra-annual CO₂ correlates with algal abundance (-CO₂) and prolonged clearwater phases (+CO₂)
- CO₂ influx increases with drier weather conditions, and is reduced with extreme N loading

Corresponding author: Emma Wiik, e.wiik@bangor.ac.uk, e.wiik@ucl.ac.uk

Abstract

Spatio-temporal variation in climate and weather, allochthonous carbon loads, and autochthonous factors such as lake metabolism (photosynthesis and respiration) interact to regulate atmospheric CO₂ exchange of lakes. Understanding this interplay in diverse basin types at different timescales is required to adequately place lakes into the global carbon cycle, and predict CO₂ flux through space and time. We analyzed 18 years of data from seven moderately hard lakes in an agricultural prairie landscape in central Canada. We applied generalized additive models and sensitivity analyses to evaluate the roles of metabolic and climatic drivers in regulating CO₂ flux at the intra-annual scale. In all basins, at mean conditions with respect to other predictors, metabolic controls resulted in uptake of atmospheric CO₂ when surface waters exhibited moderate primary production, but released CO₂ only when primary production was very low (5 – 13 μg L⁻¹) or when dissolved nitrogen was elevated (>2000 μg L⁻¹), implying that respiratory controls offset photosynthetic CO₂ uptake under these conditions. Climatically, dry conditions increased the likelihood of ingassing, likely due to evaporative concentration of base cations and/or reduced allochthonous carbon loads. While more research is required to establish the relative importance of climate and metabolism at other time scales (diel, autumn/winter), we conclude that these hard fresh waters characteristic of continental interiors are mainly affected by metabolic drivers of pCO₂ at daily-monthly timescales, are climatically controlled at interannual intervals, and are more likely to in-gas CO₂ for a given level of algal abundance, than are softwater, boreal ecosystems.

1 Introduction

It is widely accepted that lakes are important nodes that process terrestrial carbon (C) and influence global C fluxes [Cole *et al.*, 2007; Downing *et al.*, 2008; Tranvik *et al.*, 2009]. However, improved understanding of regulatory mechanisms which underlie trends and variability among lentic systems is needed to improve predictions of how lakes will both contribute and respond to future climate change [Prairie, 2008; Tranvik *et al.*, 2009]. In particular, there remains high regional and temporal variation in the mechanisms regulating lake pCO₂, despite increasing efforts to synthesize and upscale in-lake CO₂ levels and greenhouse gas fluxes. In part, this variability reflects the wide range of analytical methods and study time frames, varying from instantaneous estimates of regional lakes [Duarte *et al.*, 2008; Lapierre and del Giorgio, 2012] to decadal analyses of individual sites [Finlay *et al.*, 2015; Perga *et al.*, 2016]. Furthermore, certain lake types (e.g., hardwater and saline) are understudied relative to softwa-

48 ter boreal systems. Variability in the importance of contrasting regulatory mechanisms (e.g.,
49 broad-scale climatic drivers vs. local metabolic factors) across temporal and spatial scales can
50 obscure the hierarchical relationships among control processes, which in turn limits insights
51 derived from upscaled, ecosystem-level comparisons and global estimates.

52 Interannual and decadal trends in lake pCO₂ are modulated by many interacting vari-
53 ables, primarily acting at the landscape scale through climatic and meteorological drivers. For
54 example, changes in precipitation affects transport of solutes such as dissolved organic (DOC)
55 and inorganic carbon (DIC), which in turn alter lake water CO₂ content [Ojala *et al.*, 2011].
56 In the case of organic forms of carbon, higher substrate supply tends to elevate microbial res-
57 piration [Maberly *et al.*, 2013; Ducharme-Riel *et al.*, 2015], whereas increased DIC can either
58 increase or reduce in situ pCO₂ in hardwater systems, depending on ambient pH and alternate
59 buffering mechanisms [e.g., Baehr and DeGrandpre, 2004; Knoll *et al.*, 2013]. Additionally,
60 landscape-scale variation in irradiance (e.g., cloud cover) or air temperature [O'Reilly *et al.*,
61 2015] can lead to evaporative concentration of lakes [Pham *et al.*, 2009] and consequent changes
62 in parameters regulating pCO₂ (DIC, DOC, nutrients, etc.). For example, in continental Cana-
63 dian hardwater lakes, interannual variability in both temperature and precipitation has affected
64 pH and CO₂ flux via effects of ice-off timing [Finlay *et al.*, 2015], DIC content [Pham *et al.*,
65 2009], and regional hydrology [Bonsal and Shabbar, 2008; van der Kamp *et al.*, 2008].

66 Metabolic processes are likely to be paramount in regulating atmospheric exchange of
67 greenhouse gases at scales of hours to days. For example, water-column pCO₂ typically in-
68 creases overnight as photosynthesis becomes light-limited and respiration continues [Raymond
69 *et al.*, 2013; Liu *et al.*, 2016]. In softwater reservoirs, these diel metabolic patterns can account
70 for ca. 30% of total variation in CO₂ flux over a summer season [Morales-Pineda *et al.*, 2014].
71 In general, larger diel amplitudes of CO₂ content are found as lake productivity increases [Han-
72 son *et al.*, 2003; Shao *et al.*, 2015; Morales-Pineda *et al.*, 2014], suggesting that multiple tem-
73 poral scales may (Replaced: ~~need to be~~ replaced with: **be needed to**) evaluate CO₂ regulation
74 in productive lakes.

← noticed this
was wrong

75 At intermediate timescales, trends in lake pCO₂ are likely to be regulated by a combi-
76 nation of metabolic and climatic mechanisms [Morales-Pineda *et al.*, 2014]. For example, metabolic
77 controls underlie seasonal trends in dimictic temperate lakes when, in winter, CO₂ accumu-
78 lates under ice [Denfeld *et al.*, 2015], causing springtime efflux of CO₂ during ice melt and
79 lake overturn. Reduced pCO₂ occurs in summer when the water column is stable and primary

80 production increases, whereas pCO₂ increases during fall as respiratory products in the hy-
81 polimnion are mixed into surface waters [Alin and Johnson, 2007; Stets et al., 2009; Ducharme-
82 Riel et al., 2015; Marcé et al., 2015]. These seasonal patterns can be disrupted by climatic or
83 meteorological events such as passing storms or heat waves [Maberly, 1996; Klug et al., 2012;
84 Audet et al., 2017], or be dampened in polymictic lakes where CO₂ exhibits more limited sea-
85 sonal variation [Jonsson et al., 2003].

86 While metabolic controls of CO₂ also operate at seasonal scales in hardwater lakes [Striegl
87 and Michmerhuizen, 1998], their influence can be overrun by landscape-level controls of so-
88 lute loading [Anderson et al., 1999; Sobek et al., 2005; Christensen et al., 2013; Knoll et al.,
89 2013; Marcé et al., 2015]. For example, lakes with strong groundwater influences can have high
90 allochthonous supplies of DIC and exhibit super-saturation of CO₂, particularly in regions close
91 to the groundwater entry points [Stets et al., 2009]. On the other hand, the high pH and alka-
92 linity of hardwater lakes also buffers against large fluctuations in pH [Duston et al., 1986; Han-
93 son et al., 2003], leading to smaller amplitudes of both pH and CO₂ than exist in softwater lakes.
94 Therefore, especially in polymictic hardwater lakes without strong stratification, hypolimnetic
95 CO₂ accumulation should be relatively low and uniform throughout the year, with the net di-
96 rection of atmospheric CO₂ exchange depending on climate effects on solute loading and metabolism.
97 Thus, seasonal patterns of CO₂ content in hard-water lakes may contrast sharply from those
98 known from dimictic boreal systems.

99 Here, we use generalized additive models (GAMs) and sensitivity analysis to quantify
100 the effects of climatic and metabolic parameters in regulating intra-annual variability in pCO₂
101 of hardwater lakes in the sub-humid Canadian interior. Using bi-weekly data for 18 years in
102 seven lakes, we sought to determine: 1) When and to what extent metabolic factors (photo-
103 synthesis and respiration) were regulating lakewater pCO₂ and CO₂ flux; 2) Whether local me-
104 teorology and global climatic factors contribute to intra-annual CO₂ flux variability, and; 3)
105 How consistent the drivers of CO₂ flux were among study lakes that varied more than 10-fold
106 in size, productivity, and catchment area. Improved understanding of the relative importance
107 of biotic and abiotic controls of CO₂ flux in hardwater lake types is critical to achieving a pre-
108 dictive understanding of the role of freshwater ecosystems in global carbon cycles.

2 Methods

2.1 Study sites

The seven study sites are situated within the Qu'Appelle River catchment (ca. 52,000 km²) in the northern Great Plains of southern Saskatchewan, Canada (Fig. 1). The region has a sub-humid continental climate and is hydrologically reliant on water originating from the Rocky Mountains as well as local snowmelt [Bonsal and Shabbar, 2008; Pham *et al.*, 2009]. The South Saskatchewan River feeds the Qu'Appelle River system via Lake Diefenbaker reservoir (D). Water flows eastward from the main reservoir through a chain of lakes including Buffalo Pound (B), Pasqua (P), Katepwa (K), and Crooked (C) Lakes. Wascana (W) and Last Mountain (L) Lakes are situated on tributaries that feed into the Qu'Appelle river system upstream of Pasqua Lake. All lakes receive diffuse nutrient sources from agriculture, with the wastewater treatment plants from the cities of Regina and Moose Jaw acting as point sources of nutrients to Pasqua and eastern basins [Hall *et al.*, 1999]. All lakes are dammed to variable extent, and Buffalo Pound and Diefenbaker are actively managed reservoirs. For simplicity, we refer to all sites as lakes.

Median nutrient concentrations are generally elevated (Table 1), including total dissolved nitrogen (TDN) (0.96 mg N L⁻¹) and total dissolved phosphorus (TDP) (106 µg L⁻¹), resulting in high algal abundance as chlorophyll *a* (Chl *a*) (median 16 µg L⁻¹) and mesotrophic to hypereutrophic conditions [Hall *et al.*, 1999; Finlay *et al.*, 2009]. Compared with saline lakes worldwide [e.g., Duarte *et al.*, 2008], Qu'Appelle lakes have moderate DIC (median = 45 mg L⁻¹) and conductivity (median = 1050 µS L⁻¹), but rather high pH (median = 8.8) (Fig. 2a). DOC concentrations are moderate (median 11.5 mg L⁻¹). Temporal variation in many major chemical variables such as pH is highly synchronous across the sites (Fig. 2b; Vogt *et al.* [2011]) (see Fig. A.1 for intra-annual variability in variables relating to nutrient status and lake metabolism).

2.2 Long-term limnological sampling

Biweekly limnological sampling of pH, temperature, dissolved oxygen, conductivity, salinity, DIC, DOC, Chl *a*, TDN and metabolic bioassay estimates (primary production, respiration) followed methods outlined in Finlay *et al.* [2009]. Briefly, pH was measured at the lake surface, while oxygen, temperature, conductivity and salinity were recorded at 1 m depth using YSI-85 multi-probe meters (YSI, Inc., Yellow Springs, OH). DIC, DOC, Chl *a*, TDN and

139 metabolic bioassay samples used depth-integrated water samples pooled from 2-L Van Dorn
140 sampler casts taken at 0.5 m intervals.

141 Filtered water (0.45 μm pore size) was used for DIC and DOC analyses using a total
142 carbon analyser (Shimadzu 500A), while TDN was measured by photocombustion, both fol-
143 lowing Environment Canada protocols [*Environment Canada*, 1979]. Chl *a* was determined
144 trichromatically from particulate organic matter (POM) collected on 1.2 μm pore Whatman
145 GF/C glass fiber filters following *Jeffrey and Humphrey* [1975] and following extraction us-
146 ing 80% acetone: 20% methanol, by volume. The wavelength-specific absorbance was quan-
147 tified using a Hewlett Packard model 8452A photodiode array spectrophotometer (1996–2004)
148 or an Agilent model 8453 UV-Visible spectrophotometer (2005–2014).

149 Metabolic estimates of gross primary production, net primary production and respira-
150 tion were based on changes in oxygen concentration following incubation of whole water sam-
151 ples in light and dark glass bottles [*Finlay et al.*, 2009]. All analyses were run in triplicate us-
152 ing screened (243 μm mesh), depth-integrated water following *Howarth and Michaels* [2000].
153 Incubations occurred for 24 h at ambient lake temperature and under a 12-hour light/dark cy-
154 cle with 450 $\mu\text{mol quanta m}^{-2} \text{ s}^{-1}$, comparable to that recorded in situ at Secchi depth us-
155 ing a profiling radiometer [*Finlay et al.*, 2009].

156 Sampling occurred primarily from May 1st to August 31st between the hours of 0900
157 and 1300, with ca. 5% of sampling dates occurring earlier in spring or later in autumn. This
158 long-term ecological research program began sampling in 1994, but for reasons related to data
159 availability, we restricted this study to data from either 1996 (most lakes) or 2004 (Pasqua)
160 to 2014, inclusive.

161 **2.3 CO₂ flux calculation**

162 In the absence of direct measurements of CO₂, we relied on calculated fluxes which ap-
163 proximate real values particularly well in high-alkalinity lakes [*Abril et al.*, 2015] (such as our
164 study sites), where there are strong chemical relationships between pH and dissolved CO₂ [*Soumis*
165 *et al.*, 2004, $R^2 = 0.81$]. (Added: Calculated values are widely applied in the absence of mea-
166 surements, particularly when long-term or broad spatial data are being examined (e.g. *Duarte*
167 *et al.* [2008]; *Seekell and Gudasz* [2016]))

168 The procedure for calculating CO₂ fluxes and pCO₂ followed *Finlay et al.* [2009]. Briefly,
 169 CO₂ concentrations ([CO₂]) were calculated based on DIC concentrations (depth-integrated
 170 samples) and pH (surface), with correction for ionic strength and water temperature measured
 171 at 1 m depth [*Stumm and Morgan*, 1996]. Partial pressure of CO₂ (Pa) was estimated using
 172 Henry's Law constant [*Kling et al.*, 1992], and chemically enhanced CO₂ flux (mmol m⁻² d⁻¹)
 173 was calculated following *Cole et al.* [1998]:

$$\text{net daily CO}_2 \text{ flux} = \alpha k ([CO_{2_{lake}}] - [CO_{2_{sat}}]) \quad (1)$$

174 where: in-lake CO₂ concentration for [CO_{2_{lake}}] refers to surface water; saturation levels [CO_{2_{sat}}]
 175 refer to equilibrium with the atmosphere; α is the chemical enhancement of CO₂ flux at high
 176 pH [*Hoover and Berkshire*, 1969], calculated following *Wanninkhof and Knox* [1996], and; k
 177 is piston velocity (cm h⁻¹) following *Cole et al.* [1998], relating k to wind speed and temper-
 178 ature [*Wanninkhof*, 1992].

179 The effect of an alternative piston velocity was evaluated by including the effect of lake
 180 surface area on piston velocity and therefore CO₂ flux in our sensitivity analysis (See Statis-
 181 tical methods) [equations for k derived from Table 2, Model B; *Vachon and Prairie*, 2013].
 182 We did not have data to account for wind direction, which would plausibly incur errors in lake
 183 area-based estimates of gas transfer for e.g. Katepwa (North-South orientation) vs Pasqua (West-
 184 East orientation). Overall, however, the influence of lake area on chemically enhanced flux was
 185 subsidiary to pH and therefore not considered further in this paper.

186 Complete data for calculating CO₂ flux were available from 1996 for all lakes except
 187 Pasqua at which sampling began in 2004. Variables included temperature, pH, conductivity,
 188 salinity, DIC, wind speed, air pressure and atmospheric pCO₂. Observations with any one miss-
 189 ing variable were omitted, leaving 991 data points for modeling. Hourly wind speed and air
 190 pressure were acquired from publicly available Environment Canada (EC) data ([http://climate.
 191 weather.gc.ca/](http://climate.weather.gc.ca/)) using Regina stations 4016560 and 4016566 (Climate IDs) which had
 192 complete records for the study period. Using one weather station location for all lakes was deemed
 193 acceptable as existing records from other weather stations were found highly correlated. Two-
 194 week average wind speed was calculated to smooth out brief effects of extreme weather events.
 195 Monthly averages of air pressure (EC), and Mauna Loa atmospheric pCO₂ (Earth System Re-
 196 search Laboratory, <http://www.esrl.noaa.gov/gmd/ccgg/trends/data.html>)
 197 were used.

198 **2.4 Statistical methods**

199 All statistical analyses were performed using R version 3.2.5 [*R Development Core Team,*
200 2016], using packages mgcv [*Wood, 2011, 2017*] and pse [*Chalom and de Prado, 2016*]. R code
201 is available at <https://github.com/simpson-lab/jgr-co2-flux>.

202 (Added: Our analytical approach follows a few key underlying considerations. Since CO₂
203 flux was estimated from water chemistry and physical variables and not measured directly, we
204 avoided any approach that would circularly include these 'calculation variables' as metabolic
205 or climatic proxy predictors of CO₂ flux. Furthermore, we were specifically interested in which
206 of these calculation variables correlate the most with CO₂ flux in our study region. In this re-
207 gard we note that, although the real, rather than estimated, relationship between these variables
208 and CO₂ flux is unknown, this step can identify which variable is key to proxy CO₂ flux in
209 our region (and conversely, which variables are not). Therefore, we first quantified the influ-
210 ence of the calculation variables on estimated CO₂ flux ('influence' here used in the regres-
211 sion sense of changes in x influencing estimates of y , rather than a directional causal sense).
212 Secondly, we regressed our metabolic and climatic variables of interest against the variable
213 that accounted for most of this variation. The second step allowed us to use a measured, rather
214 than estimated, response variable, reducing the amount of imprecision in our regression val-
215 ues. We were then able to relate these values back into CO₂ flux estimates using the results
216 from the first step, thereby avoiding presenting misleadingly precise results for CO₂ flux it-
217 self.)

218 **2.4.1 Variable selection**

219 (~~Deleted: Since pCO₂ was estimated from water chemistry parameters and not measured
220 directly, we avoided modeling CO₂ flux using any variables involved in the calculations. Instead,
221 we staggered analysis into two stages, first selecting the input variable which accounted for
222 most of the variation in CO₂ flux, and then regressing the remaining metabolic and climatic
223 variables against that selected variable.~~)

224 Metabolic variables were selected from various estimates of lake production and respi-
225 ration to achieve the greatest availability over the data period. In the case of highly correlated
226 variables, we modelled only a single variable, so in our case, respiration (R) was selected over
227 net and gross primary production, whereas TDN was retained over TDP (at most times at most
228 study sites, N limitation exceeds P limitation [*Patoine et al., 2006*]). Ultimately, five metabolic

229 variables were selected for modeling, including in situ O₂ (respiration/photosynthesis), DOC
 230 (potential effects on respiration), Chl *a* (algal biomass or production), R (respiration), and TDN
 231 (nutrient availability). Chl *a*, TDN, and DOC were log₁₀-transformed to approximate a nor-
 232 mal error distribution.

233 To capture the major climatic processes most likely to influence lake CO₂ via solute and
 234 nutrient loading (hydrological processes, evapotranspiration), we included both broad drivers
 235 of intra-annual climate and more local, instantaneous proxies for evaporation-precipitation bal-
 236 ance. Variables included the Southern Oscillation Index (SOI) and Pacific Decadal Oscillation
 237 (PDO), metrics of climate systems which strongly influence regional precipitation and tem-
 238 perature patterns, either alone or in combination [*Bonsal and Shabbar, 2008; Pham et al., 2009;*
 239 *Shabbar and Yu, 2012*]. Both indices were included as three-month averages, six months prior
 240 to sample collection, to account for the lags between the regions of observation and effect [*Pomeroy*
 241 *et al., 2007; Shabbar et al., 2011*]. Monthly values were obtained from the National Oceanic
 242 and Atmospheric Administration (NOAA) ([http://www.cpc.noaa.gov/data/indices/](http://www.cpc.noaa.gov/data/indices/soi)
 243 [soi](http://www.cpc.noaa.gov/data/indices/soi)) and the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) ([http://](http://research.jisao.washington.edu/pdo/PDO.latest)
 244 research.jisao.washington.edu/pdo/PDO.latest). Because regional precip-
 245 itation is highly localized (lake-specific; *Vogt et al. [2011]*) and weather stations were not ad-
 246 jacent to our study sites, we did not attempt to use data from weather stations to estimate rain-
 247 fall. Instead, Standardized Precipitation Evapotranspiration Index (SPEI) values for each site
 248 (0.5 degree spatial resolution) were obtained from the Consejo Superior de Investigaciones Cient-
 249 ficas (CSIC) Global SPEI database (<http://sac.csic.es/spei/database.html>)
 250 [*Vicente-Serrano et al., 2016*]. Index values were calculated using a two-month 'memory' (au-
 251 tocorrelation) to account for temporal variation in soil drying and hydration.

252 **2.4.2 Sensitivity analysis**

253 Given the absence of direct measurements, we analysed data to select the best proxy of
 254 CO₂ in our climatic-metabolic model by simulating the sensitivity of calculated CO₂ flux to
 255 changes in pH, conductivity, salinity, water temperature, DIC, wind speed, atmospheric pCO₂,
 256 and local air pressure. A sensitivity analysis was used for this purpose because it shows the
 257 magnitude of individual variable contributions to estimate CO₂ flux for multiple combinations
 258 of variables and values. Further, this method allows us to perform multi-step calculations while
 259 controlling for underlying data correlations [*Chalom and de Prado, 2015*].

260 Differences among lakes in the relative contribution of variables to calculated CO₂ flux
 261 were tested by comparing an analysis conducted for all lakes combined, with those for each
 262 lake individually. Specifically, we used a latin hypercube sampling (LHS) approach [*Chalom*
 263 *and de Prado, 2015*] and generated realistic data variations of all variables for each lake based
 264 on their observed variation over the sampling period (n = 500 per simulation). Rank corre-
 265 lations were selected, rather than a linear analysis among variables, to account for potential
 266 nonlinear relationships between predictors and responses. The output metric (partial rank cor-
 267 relation coefficient: PRCC), for any one variable, controls for the effect of all other variables
 268 by reflecting the correlation between the unexplained part of the outcome, given all other vari-
 269 ables, and the unexplained part of one variable, given all other variables (i.e., a correlation be-
 270 tween residuals).

271 **2.4.3 Generalized additive models**

272 pH was the strongest correlate with calculated CO₂ flux based on sensitivity analysis (see
 273 Results) and, therefore, was carried forward to evaluate the effects of selected metabolic and
 274 climatic variables on CO₂ flux. Here, we applied generalized additive models (GAM), which
 275 account for nonlinear relationships between predictors and responses [*Hastie and Tibshirani,*
 276 *1990; Wood, 2017*] (Section 2.4.4). GAMs also allowed us to include Year and Lake as ran-
 277 dom effects to account for between-lake and inter-annual variations known to be important [*Fin-*
 278 *lay et al., 2009, 2015*]. The resolutions of all other predictors also link with the resolution of
 279 variability they are able to explain: e.g., biweekly predictors can explain pH variation at a within-
 280 month scale, while monthly predictors can only explain pH variation occurring at a between-
 281 month scale. Temporal structure within the climatic-metabolic model was visualised by plot-
 282 ting term contributions to pH against time.

283 The first model, which evaluated the degree to which lakes differed in their relationship
 284 between CO₂ and pH was formulated as follows, for $y = \text{CO}_2$ flux,

$$y = \beta_0 + f(\text{pH}) + f_{\text{lake}}(\text{pH}) + \alpha_{\text{lake}} + \gamma_{\text{year}} + \varepsilon \quad (2)$$

285 Here, the effect of pH was modelled both globally ($f(\text{pH})$) and by lake ($f_{\text{lake}}(\text{pH})$), while terms
 286 α and γ were random effects of lake and year, respectively, and ε was the error term. The global
 287 and lake-specific effects of pH were identified via different orders of quadratic penalties on
 288 their respective basis expansions. The global function of pH ($f(\text{pH})$) was subject to the usual

289 second-order penalty whereby the wiggleness penalty was on the second derivative of a fitted
 290 spline. First-order penalties were used for the lake-specific splines so that the penalty applied
 291 to departure from a flat or zero function. This approach had the effect of making each $f_{\text{lake}}(\text{pH})$
 292 represent the departure of each lake from the global pH effect. Smoothness parameters for f
 293 and f_{lake} were chosen using restricted maximum likelihood (REML) selection [Wood, 2011].
 294 Lake-specific effects of pH on CO_2 , ($f_{\text{lake}}(\text{pH})$), were only retained when they were assessed
 295 to be significantly different from a zero (flat) function. Therefore, lake-specific splines retained
 296 reflect regional heterogeneity (objective 3) between the study sites. pH was selected for a com-
 297 bined metabolic and climatic GAM to explore sub-annual controls of CO_2 flux.

298 The second model, which quantified the influence of climatic and metabolic variables
 299 on pH, followed the principles outlined above for the first model. For $y = \text{pH}$,

$$y = \beta_0 + \sum_{j=1}^J [f(x_j) + f_{\text{lake}}(x_j)] + f(\text{PDO}, \text{SOI}) + \alpha_{\text{lake}} + \gamma_{\text{year}} + \varepsilon \quad (3)$$

300 where: x_j is the j^{th} metabolic (TDN, DOC, Chl a , O_2) or climatic (SPEI) covariate, $f(\text{PDO}, \text{SOI})$
 301 is a 2-D tensor product spline combining the main and interactive effects of PDO and SOI;
 302 α and γ are random effects of lake and year, and; ε is the error term. As above, the unique ef-
 303 fects of the x_j for each lake were incorporated through inclusion of separate difference splines
 304 for each lake ($f_{\text{lake}}(x_j)$) employing first-order wiggleness penalties. REML smoothness selec-
 305 tion was used as described above. Where model terms were marginally significant, likelihood
 306 ratio tests were used to determine whether a model including the terms was justifiable.

307 Preliminary runs suggested that co-linearity between DOC and TDN was sufficient to
 308 confound results, and argued for retaining only one predictor (DOC), based on both internal
 309 model Wald tests and Akaike and Bayesian Information Criteria (AIC; BIC). However, due
 310 to TDN being a significant correlate absent from the final model, the model replacing DOC
 311 with TDN is also used in this paper to portray the relationship between TDN and pH.

312 **3 Results**

313 **3.1 The sensitivity of CO_2 flux to variables used in its calculation**

314 Sensitivity analysis showed that pH explained the greatest amount of variation in CO_2
 315 flux (PRCC = -0.96) followed by DIC (PRCC = 0.51) for all lakes (Table 2, Fig. B.1). This
 316 sequence was also retained in the simulations for individual lakes; however, DIC was more

317 influential in some lakes (B, C, D, L) than in others (K, P, W). Overall, the importance of DIC
 318 was small (Table 2) and sensitive to which simulation data were used for analysis (not shown).

319 Generalized additive modeling echoed the results of the sensitivity analysis and showed
 320 that pH was the main correlate of CO₂ flux (Fig. 3). This model explained 97% of deviance
 321 in CO₂ flux, while the use of DIC as an additional term only explained a further 1% of vari-
 322 ation (and an equivalent model with DIC, not pH, explained only 30% of flux variation; not
 323 shown).

324 Lakes were predicted to in-gas atmospheric CO₂ above a pH of 8.8, the median pH over
 325 the whole data set, while no net atmospheric exchange occurred around pH 8.7. Generally more
 326 productive lakes (K,P,W) were significantly different from less productive sites (B, C, D, L)
 327 based on GAM analysis of the relationship between pH and CO₂, primarily at the high and
 328 low ends of pH (<10% of all observations). These groups of lake also differed in the extent
 329 to which DIC content tended to influence sensitivity analyses (Table 2).

330 **3.2 Metabolic and climatic regulation of pH**

331 GAM analysis explained 43% of historical deviance in pH, mainly due to climatic and
 332 metabolic parameters (~~Deleted: (Figs 4, 6, 3)~~ (Added: (Figs 4 – 6)). Significant predictors
 333 of pH included Chl *a* ($p < 0.001$), PDO*SOI ($p < 0.001$), Lake + Year ($p < 0.001$), oxygen
 334 ($p = 0.0108$), DOC ($p = 0.0137$) and SPEI ($p = 0.0122$). The only variable for which individ-
 335 ual lake splines were significant was Chl *a*. In all cases, R was insignificant and removed from
 336 the model. The ranges of pH over which the metabolic and climatic variables exerted control
 337 were variable, and in decreasing order included PDO*SOI (ca 8.5-10), Chl *a* (8.5-9.6), oxy-
 338 gen (8.6-8.9), DOC (8.75-8.9), and SPEI (8.9-9.1), approximately (see uncertainties at the edge
 339 of prediction: Figs 4, 6). Using all measured combinations of our predictors, i.e. the empir-
 340 ical data, our model pH predictions encompass a range from 7.8 to 10 (\pm errors), which does
 341 not capture the full range of observed pH (7 to 10.9) (Figs C.1, C.2).

342 Concentrations of Chl *a* were correlated positively with pH, with low algal abundance
 343 ($< 5 - 13 \mu\text{g L}^{-1}$) occurring when depressed pH (~~Replaced: favoured~~ replaced with: **corre-**
 344 **lates with**) out-gassing of CO₂ when all other predictors were held at their mean (Fig. 4). Re-
 345 sults from the two small, shallow lakes (W, B) were significantly different from other basins
 346 in that both increases and declines in Chl *a* had comparatively strong relationships with pH.
 347 In general, pH increased with oxygen saturation, with CO₂ in-gassing at supersaturated oxy-

348 gen concentrations $>9 - 10 \text{ mg L}^{-1}$). CO_2 efflux occurred only at low oxygen concentrations
 349 ($<5\%$ of all observations which were lower than ca 5 mg L^{-1} when all other predictors held
 350 at their mean). Finally, DOC was positively correlated with pH, particularly in the range where
 351 elevated pH favoured influx of CO_2 .

352 In the alternative model where DOC was replaced with TDN, TDN had a slight posi-
 353 tive relation with pH up to concentrations of ca $1100 \mu\text{g N L}^{-1}$ above which pH declined con-
 354 sistently (Fig. 5). Uncertainties in the effect of TDN on pH were high at both ends of the range
 355 due to low observation frequency; however, extremely high values of TDN ($> 2000 - 6500 \mu\text{g N L}^{-1}$)
 356 co-occurred with pH values that (~~Replaced: favoured~~ replaced with: **correspond with**) CO_2 ef-
 357 flux.

358 Broader-scale climate variables PDO and SOI had stronger relationships with pH than
 359 did SPEI. The highest pH values were associated with the most negative SOI and positive PDO
 360 (Figs 4 a-c), which typically indicate warm and dry conditions. In contrast variation in SPEI
 361 had a limited effect on pH (ca 0.2 units) and was associated with above-mean pH at the low
 362 and high end of its range (Fig. 6). Low pH was particularly common when PDO was low and
 363 wet conditions predominate [Bonsal and Shabbar, 2008]. PDO had a more complex multi-modal
 364 relationship with pH than did SOI, which was more linear (Figs 4 b-c). For a given PDO, in-
 365 creasing SOI shifted the position of the spline. In general, SOI had a positive relationship with
 366 pH except at high PDO when high pH occurred also at low SOI values (Figs 4 b-c). Over-
 367 all, the range in climatic index values during the observation period was similar to that recorded
 368 during the past century (PDO mostly within -2,2; SOI mostly within -2.5, 2.5, SPEI mostly
 369 within -2,2).

370 Consistent long-term intra-annual trends were apparent for the metabolic variables Chl *a*,
 371 and oxygen (Fig. 7), but not DOC or the climatic variables SPEI and PDO*SOI. Chl *a* increased
 372 in positive effect on pH over the summer in most lakes except during the clear-water phase
 373 in June. Below-average pH at low Chl *a* occurred consistently at the least productive site, Lake
 374 Diefenbaker. Oxygen effects in four lakes (C, K, B, D) were most negative towards the end
 375 of the summer.

376 **4 Discussion**

377 Given the importance of climate and ice-cover duration in determining annual mean pH
 378 and CO_2 flux in these hard-water lakes [Finlay *et al.*, 2015], we sought to determine whether

379 metabolic factors would emerge as a driving factor at an intra-annual timescale. While we found
 380 similar controls also at sub-annual timescales (high coherence within the region, pH the most
 381 significant predictor of CO₂) (Table 2, Fig. 3), metabolic controls were important in determin-
 382 ing the balance between high likelihoods of influx (pH>8.8) and efflux (pH<8.7) of CO₂ (Fig. 4).
 383 Lake metabolism, as measured using algal abundance (Chl *a*), was a key parameter control-
 384 ling whether lakes acted as C sources or sinks within any given year.

385 **4.1 The role of lake metabolism in directing pH and CO₂ flux**

386 There was strong evidence for metabolic control of pH and thereby CO₂ flux both at the
 387 high and low ends of a gradient of nutrient concentration when either primary production was
 388 insufficient to sequester CO₂ or it seemed offset by high levels of inferred respiration. Fur-
 389 ther, metabolic effects exhibited a strong intra-annual pattern, stressing the importance of short
 390 term controls of pH and thereby CO₂ flux in these lakes in calculating the annual CO₂ bud-
 391 get.

392 Elevated algal abundance increased the likelihood of net CO₂ uptake from the atmosphere.
 393 Specifically, we found that CO₂ under average conditions (all other predictors at mean) was
 394 in-gassing at moderate to high primary production (Chl *a* > 15 µg L⁻¹) while lower levels of
 395 productivity (Chl *a* 5 – 10 µg L⁻¹) could result in a net heterotrophic state and CO₂ efflux.
 396 Such low productivity values were found most frequently in the mesotrophic Lake Diefenbaker,
 397 while strongly positive relationships between pH and Chl *a* occurred often in the most shal-
 398 low lakes (Wascana, Buffalo Pound; Fig. 4). In general, the observed Chl *a* concentrations needed
 399 for net CO₂ release were low (7 – 15 µg L⁻¹) relative to those found in other eutrophic lakes
 400 where out-gassing may predominate even under the most productive conditions (Chl *a* > 40 µg L⁻¹)
 401 [Huttunen *et al.*, 2003; Reis and Barbosa, 2014], although outgassing was predicted even in
 402 our sites at similar algal production provided other predictors were set to values favouring out-
 403 gassing (e.g. low oxygen, high TDN).

404 Both Chl *a* and pH increased through the summer in most lakes suggesting a progres-
 405 sive increase in the importance of metabolic controls. However, these trends were not mono-
 406 tonic, particularly in the more productive lakes. In early summer, the more productive lakes
 407 have consistent clear-water periods [Dröscher *et al.*, 2009] caused by strong zooplankton graz-
 408 ing on phytoplankton, thus increasing pCO₂ and subsequently decreasing pH (Fig. 7). Con-
 409 versely, in late summer, the more productive lakes exhibit reduced oxygen concentrations (<5 mg L⁻¹)

410 indicative of increased respiration of organic material, which favours release of CO₂ to the at-
411 mosphere (Fig. 7). More intensive evaluation of fall metabolism is required to establish whether
412 this trend continues through to ice formation in late October or November.

413 Similar to results of annual mean data [Finlay *et al.*, 2009], rising DOC content tended
414 to co-occur with increasing pH at moderate to high DOC levels (DOC: 5 – 25 mg L⁻¹). These
415 patterns are contrary to studies from boreal lakes which tend to show that DOC mineraliza-
416 tion increases pCO₂ and reduces pH [Balmer and Downing, 2011]. Although speculative, the
417 observed positive relationship between pH and DOC may reflect recalcitrant DOC which is
418 not respired [Ostapenia *et al.*, 2009], autochthonously derived DOC during high primary pro-
419 duction [Søndergaard *et al.*, 2000], and/or a positive correlation between DOC and nutrient
420 influx [Osburn *et al.*, 2011]. The latter two are most likely given the positive correlation be-
421 tween TDN and DOC in our study lakes, however further research is required to distinguish
422 among these explanations.

423 The unimodal relationship of TDN and pH (peak ca. 1100 µg N L⁻¹) suggests that there
424 is a limit to the fertilising effect of nutrients on primary production and in turn pH. Such a
425 limit may reflect a consistent rise in bacterial decomposition of organic matter along the pro-
426 duction gradient, leading to a paramount effect of respiration under highly eutrophic condi-
427 tions [Hollander and Smith, 2001]. In our case, TDN itself may be directly utilised by heterotrophs,
428 as most (>80%) dissolved N in these lakes is in organic forms of TDN not available to au-
429 totrophs [Bogard *et al.*, 2012]. Consistent with this idea, we note that addition of organic N
430 (as urea) to mesocosm experiments in Wascana Lake increased respiration and decreased pH
431 (Replaced: favouring replaced with: corresponding with) CO₂ efflux [Bogard *et al.*, 2017]. Fi-
432 nally, we infer that the negative correlation between high TDN and pH does not reflect a change
433 in the nutrient limitation status of the lakes, as only Diefenbaker and to a lesser extent Buf-
434 falo Pound show evidence of P limitation [Vogt *et al.*, 2015; Quiñones-Rivera *et al.*, 2015] and
435 these sites generally exhibit low TDN values relative to other, more definitively N-limited sys-
436 tems [Leavitt *et al.*, 2006; Patoine *et al.*, 2006].

437 While we observed a predictable positive relationship between pH and O₂ concentration
438 when oxygen was below saturation, the relationship reversed direction when waters were super-
439 saturated with oxygen (Fig. 4b). We speculate that there are times when there may be simul-
440 taneous supersaturation of oxygen and CO₂ thereby decoupling the relationships between oxy-

441 gen and pH, as observed in other hardwater systems where excess allochthonous carbon co-
442 incides with high primary production [Stets *et al.*, 2009; McDonald *et al.*, 2013].

443 **4.2 Climatic regulation of pH**

444 The strength of the relationship between climatic variables and pH was comparable to
445 that of metabolism and pH (Figs 4, 6), a pattern which suggests that climatic mechanisms may
446 also influence intra-annual variation in regional CO₂ flux. For example, dry and warm con-
447 ditions (very high PDO and very low SOI) as well as high drought index values were asso-
448 ciated with elevated pH and increased concentrations of base cations in these and other lakes
449 lakes [Pham *et al.*, 2009; Lake, 2011]. Similarly, this pattern is consistent with findings of *Fin-*
450 *lay et al.* [2015] who demonstrated that spring and summer pH is elevated during years when
451 short duration of ice cover reduces under-ice respiration and favours increased pH in spring
452 and summer. The most likely drivers of climatic effects on pH are increased base cation con-
453 centrations due to evaporative concentration [Evans and Prepas, 1996; Pham *et al.*, 2009], el-
454 evated residence time [Knoll *et al.*, 2013], reduced allochthonous DIC loads due to longer tran-
455 sit times [Stets *et al.*, 2017], and higher reliance on groundwater contributions [Lake, 2011].
456 However, further research will be required to better refine these possibilities, including spa-
457 tial studies relating geology, landscape position, external loading and groundwater supply to
458 seasonality of lake chemistry.

459 Despite strong and significant results from our modelling exercise, our statistical approach
460 captured only ca. 43% of the deviance in pH, leaving a considerable proportion to be accounted
461 for by other factors. Because model residuals were random and normal, they provided little
462 indication of model deficiencies. In principle, model prediction might be improved through
463 distinction of DOC providence via spectrophotometric or compound-specific analyses to bet-
464 ter estimate its effect on respiration [Koehler *et al.*, 2012], while quantification of physico-chemical
465 processes such as convection and mixing may be important in identifying additional controls
466 of pH, such as seen elsewhere [Maberly, 1996; Morales-Pineda *et al.*, 2014; Liu *et al.*, 2016].
467 Thirdly, the use of more finely resolved taxonomic data (e.g., algal groups) in place of coarse
468 metrics of planktonic metabolism (Chl *a*, R) may help refine how the importance of biotic con-
469 trols varies along long limnological gradients [Felip and Catalan, 2000; George and Heaney,
470 1978; Zhang *et al.*, 2010]. Finally, we have not been able to account for alkalinity affecting
471 the buffering capacity and thus the lakes' responsiveness in pH to changes in metabolic CO₂.

472 However, the lack of overall correlation between pH and metabolic covariates suggests that
 473 alkalinity changes should be investigated for their potential contribution to pH and CO₂ flux.

474 **4.3 Regional coherence and implications for upscaling CO₂ fluxes**

475 Predicting CO₂ fluxes in these hard-water systems was simplified by the fact that DIC
 476 concentrations varied little across the lakes, and that all lakes behaved similarly with regards
 477 to metabolic and physical relationships with pH over broad spatial scales. While the lakes var-
 478 ied substantially in salinity and conductivity (Fig. 2), these parameters had relatively low im-
 479 pact on CO₂ fluxes in their respective ranges (Table 2). Conversely, while DIC concentrations
 480 are predicted to have substantial effects on atmospheric CO₂ exchange in other lake regions
 481 [Cumming *et al.*, 1995; Doctor *et al.*, 2008; Duarte *et al.*, 2008], in our study DIC levels were
 482 comparatively low, and also correlated weakly and negatively with changes in pH ($p < 0.001$,
 483 $R^2=0.014$) which implied an absence of negative effects of high DIC on CO₂ influx at high
 484 pH.

485 We found an unexpectedly strong effect of lake morphology on the role of algal abun-
 486 dance (as Chl *a*) as a determinant of pH, with the effect of Chl *a* being much greater in very
 487 shallow Buffalo Pound and Wascana lakes (<4 m mean depth) than deeper lakes, particularly
 488 at very high pH values (Fig. 4). We speculate that shallow lakes are more likely to exhibit whole-
 489 lake responses to photic-zone metabolism, and may have less vertical structure than even deep
 490 polymictic lakes (Zhang *et al.* [2010], but see George and Heaney [1978]). Fortunately, most
 491 prairie lakes are of a similar depth, many being shallow [Last, 1989], suggesting that varia-
 492 tion in morphology will not unduly affect efforts to estimate regional CO₂ fluxes [Finlay *et al.*,
 493 2015]. Overall, the high level of coherence among basins in terms of high pH and moderate
 494 DIC suggest that many lakes will act as CO₂ sinks during much of the summer, provided they
 495 are moderately to highly productive ($> 15 - 20 \mu\text{g L}^{-1}$ Chl *a*) and are not under extreme (or-
 496 ganic) TDN loads.

497 Metabolic control of CO₂ flux in these hardwater lakes does not appear to be as strong
 498 as that observed in boreal or softwater regions where microbial metabolism of DOC [Sobek
 499 *et al.*, 2005; Lapierre and del Giorgio, 2012] or photosynthesis [Maberly, 1996; Reis and Bar-
 500 bosa, 2014] regulates pCO₂, albeit with variable allochthonous contributions of respired or oth-
 501 erwise derived DIC [Weyhenmeyer *et al.*, 2015; Bogard and del Giorgio, 2016]. These results
 502 fit within the larger matrix of lake types along gradients of DIC, DOC, nutrients and alkalin-

503 ity, and suggest that moderately hardwater lakes are more likely to capture atmospheric CO₂
 504 at a given level of productivity than would dilute lakes [Reis and Barbosa, 2014], those with
 505 high DOC loads [Huttunen *et al.*, 2003], or hardwater systems with chronic oversaturation of
 506 DIC [Marcé *et al.*, 2015]. Further, because such systems often co-occur with intensively fer-
 507 tilized agricultural regions, there exists the possibility that many of these systems will fall be-
 508 low the global average estimate of lake CO₂ flux [Raymond *et al.*, 2013].

509 **5 Conclusions**

510 Based on advanced time series analysis using GAMs, we found that both metabolic and
 511 climatic factors strongly influenced factors related to pH and that variation in DIC was of only
 512 secondary importance in affecting CO₂ content. Overall, a modest degree of eutrophication
 513 was required for high rates of CO₂ uptake from the atmosphere and some less productive lakes
 514 exhibited a release of CO₂ from surface waters. These agricultural areas often exhibit high al-
 515 lochthonous loads of organic carbon and nitrogen which are likely to fertilize the lake. This
 516 increases the likelihood of CO₂ influx, but the balance may switch in favour of respiration at
 517 extreme nitrogen loads. Overall, climate appeared to have an effect on gas exchange mainly
 518 during extremes, such as regional drought, when evaporative concentration of base cations and
 519 elevated pH may favour regional influx of CO₂ into lakes. These results aid in our ability to
 520 understand and predict how future human-mediated changes to nutrient loading and climate
 521 change will impact carbon cycling in lakes.

522 **6 Tables, and figure captions**

529 Fig. 1: The seven study sites lie along the Qu'Appelle River (SK, Canada) flowing west
 530 to east, with the exception of Wascana (south tributary) and Last Mountain (north tributary).

531 Fig. 2: a: Box plots for limnological data used to calculate carbon dioxide flux in the
 532 lakes, showing medians, upper and lower quartiles, 1.5×inter-quartile ranges, and 'outliers'.
 533 b: Major patterns of annual variation in pH in all lakes, based on a generalised additive model
 534 of pH by Lake, Year, and Day of Year. Rug: annual means of pH observed over time.x

535 Fig. 3: GAM splines for pH with lake splines significantly different (see section 2.4.3)
 536 from the global spline indicated by color/hue and linetype. Dotted lines: means of y and x;
 537 Shaded area: Middle 90% of all observations. Rug: Data points

523 **Table 1.** Summary data of study lakes, showing median, minimum-to-maximum (in parentheses) values of
 524 monitoring data over the sampling period, as well as mean depth and residence time.

Lake	Residence time (yr)	Mean depth (m)	TDN ($\mu\text{g N L}^{-1}$)	Chl <i>a</i> ($\mu\text{g L}^{-1}$)	DOC (mg L^{-1})	TDP ($\mu\text{g P L}^{-1}$)
B	0.7	3	491 (218-1350)	20.1 (1.5-319)	6.1 (0.5-31)	23 (9-132)
C	0.5	8	920 (450-2090)	18 (0-237)	12 (0-41)	126 (16-650)
D	1.3	33	401 (107-1440)	4.7 (0.8-26)	4.8 (0-29)	9 (0.4-295)
K	1.34	14	1152 (418-2390)	21 (1.5-117)	12 (3.7-37)	159 (40-690)
L	12.6	8	999 (482-1510)	13 (2.3-49)	13 (0-82)	31 (14-470)
P	0.71	6	1420 (171-3100)	22 (1.2-287)	12 (0-56)	162 (5-662)
W	0.7	1.5	1309 (600-6400)	27 (2.2-309)	16 (4.8-53)	318 (33-830)

525 **Table 2.** Partial Rank Correlation Coefficients (PRCCs) following Latin Hypercube sensitivity analysis for
 526 all variables and all lakes (left panel) and the most important two variables for individual lakes (right panel).

Variable	PRCC (all lakes)	Lake	PRCC (pH)	PRCC (DIC)
pH	-0.96	Last Mountain	-0.98	0.74
DIC	0.51	Crooked	-0.99	0.69
Temperature	-0.28	Diefenbaker	-0.99	0.68
Conductivity	-0.26	Buffalo Pound	-0.99	0.65
Wind	0.20	Pasqua	-0.99	0.64
Salinity	0.10	Katepwa	-0.99	0.57
Air pressure	0.10	Wascana	-0.99	0.56
Air $p\text{CO}_2$	-0.09			

538 Fig. 4: a-c: GAM splines for significant metabolic variables. Dotted lines: means of y
 539 and x ; Shaded area: Middle 90% of all observations. Rug: Data points. a: GAM splines for
 540 chlorophyll *a*, with lakes with significantly different splines to the global spline (see section
 541 2.4.3) indicated by color/hue and linetype. b: GAM spline of oxygen, with standard errors in-
 542 dicated by shading. c: GAM spline of DOC, with standard errors indicated by shading.

543 Fig. 5: GAM spline for TDN in the alternative model without DOC. Dotted lines: means
 544 of y and x ; Shaded area: Middle 90% of all observations. Rug: Data points. Standard errors
 545 are indicated by shading.

527 **Table 3.** Summary of the climatic-metabolic model of pH, showing the estimated effects of the predictors.
 528 EDF=estimated degrees of freedom, DF=degrees of freedom. Deviance explained: 43.2%, n=991.

Predictor spline	EDF	DF	chi ²	p value
Chlorophyll <i>a</i> (global)	0.979	9	134.366	≪ 0.0001
Chlorophyll <i>a</i> (Katepwa)	0.000159	4	0	0.47556
Chlorophyll <i>a</i> (Last Mountain)	0.0000767	4	0	1
Chlorophyll <i>a</i> (Buffalo Pound)	1.80	4	11.168	0.01886
Chlorophyll <i>a</i> (Crooked)	0.277	4	0.433	0.22987
Chlorophyll <i>a</i> (Diefenbaker)	0.0380	4	0.05	0.28051
Chlorophyll <i>a</i> (Wascana)	2.65	4	66.947	≪ 0.0001
Chlorophyll <i>a</i> (Pasqua)	0.000168	4	0	0.49175
DOC	1.40	9	39.519	0.01285
Oxygen	3.07	9	28.417	0.00772
PDO*SOI	10.8	24	567	≪ 0.0001
SPEI	1.41	2	16.342	0.01158
Lake*Year	105	128	532.24	≪ 0.0001

546 Fig. 6: a-c: GAM interactions of PDO and SOI. a: Heatmap with data points. Dashed
 547 lines indicate cross sections for b-c, which show GAM splines for pH for selected combina-
 548 tions of SOI (b) and PDO (c) values. Missing line segments reflect uncertainties in prediction.
 549 d: GAM spline of SPEI, with standard errors indicated by shading. Rug: Data points.

550 Fig. 7: Contributions of each predictor to pH summarised over the months of highest
 551 data availability, averaged across lakes for weather and climate indices which were homoge-
 552 nous through the study region. Box plots show medians, upper and lower quartiles, 1.5×inter-
 553 quartile ranges, and 'outliers'. Shaded area: ± 0.05 regions to aid comparison of magnitudes
 554 across predictors.

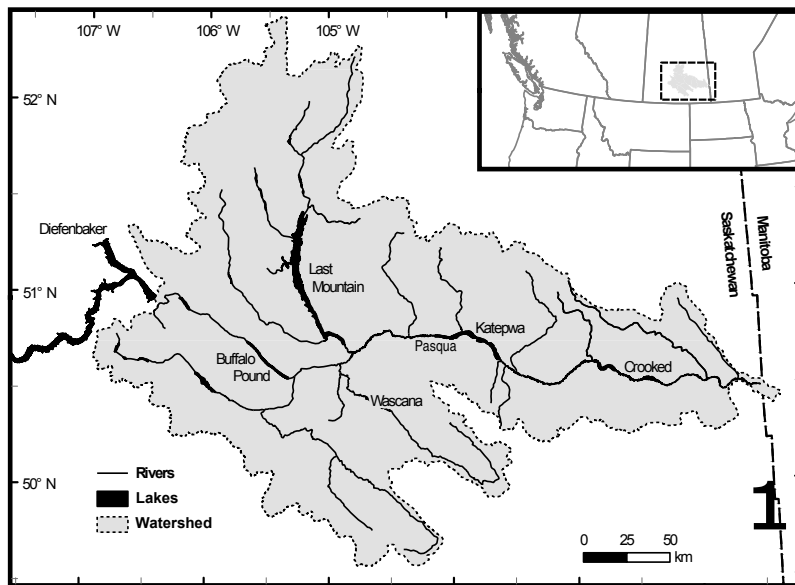
555 Fig. A1: Intra-annual variability expressed as median absolute deviation (i.e. the me-
 556 dian of the absolute deviations from the median) of key metabolic and/or nutrient status vari-
 557 ables over the LTER period over the months of most frequent observations (May-September).
 558 The data are superimposed such that the lakes with the lowest variability appear toward the
 559 centre of the figure, and lakes with higher variability contain the variability of the more cen-
 560 tral lakes plus the additional value indicated by the colouring.

561 Fig. B1: The relationship between calculated carbon dioxide flux and simulated data sets
562 (N=500) of input variables for sensitivity analysis.

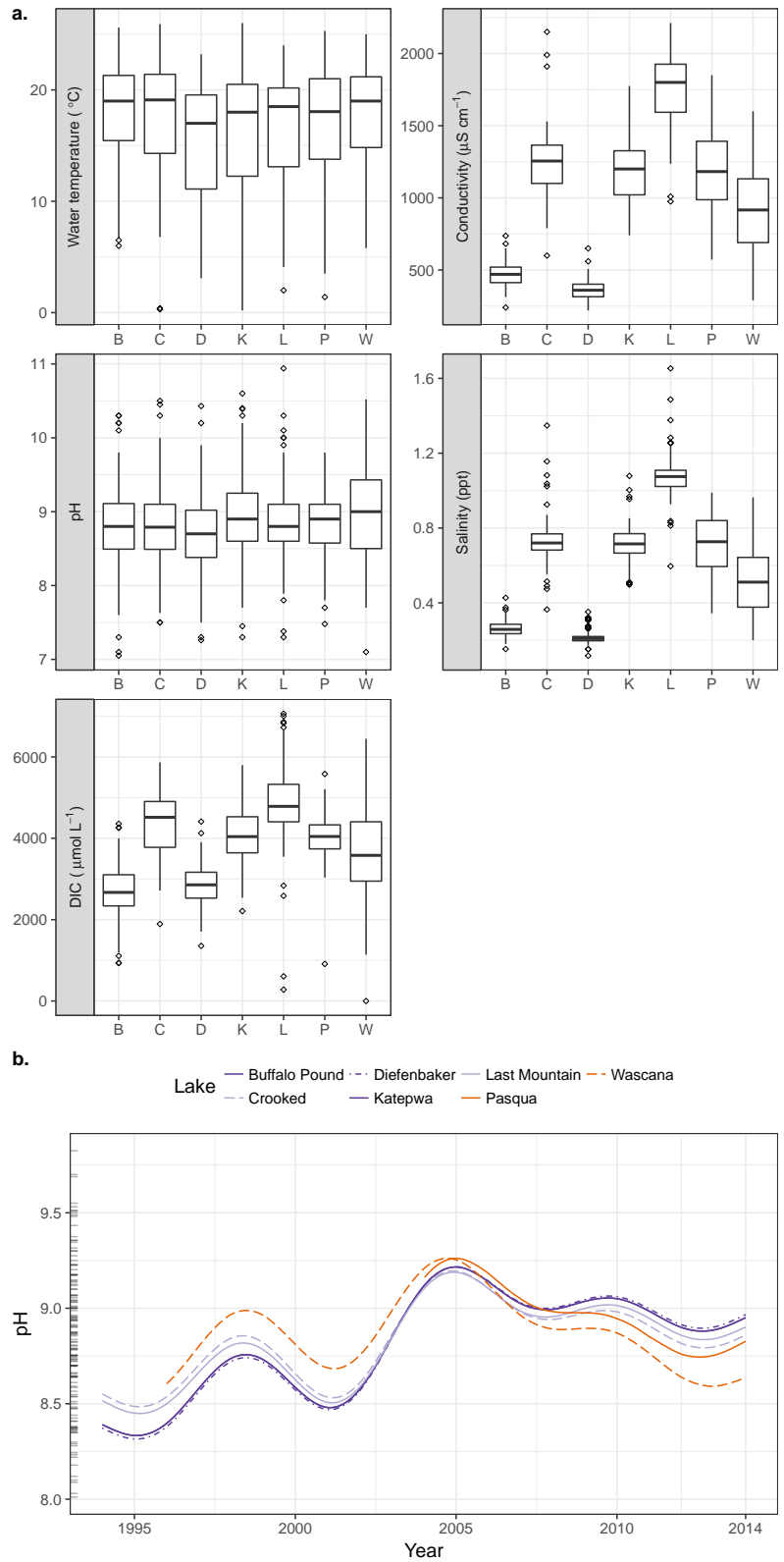
563 Fig. C1: R output for main model diagnostics.

564 Fig. C2: Measured vs predicted pH over time in the study sites, displayed as monthly
565 means over the months of the most frequent observations.

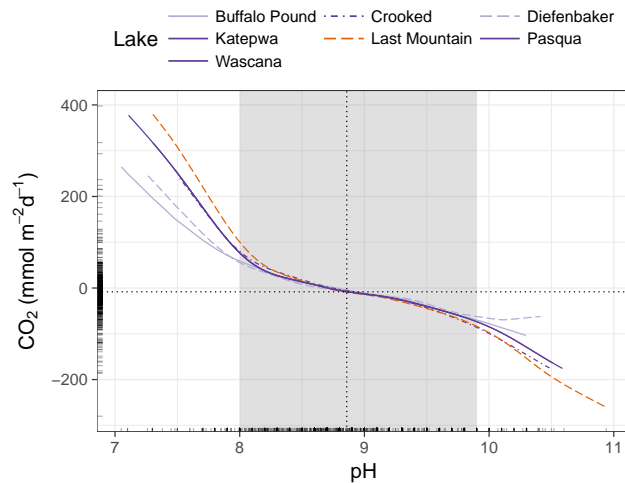
566 **7 Figures**



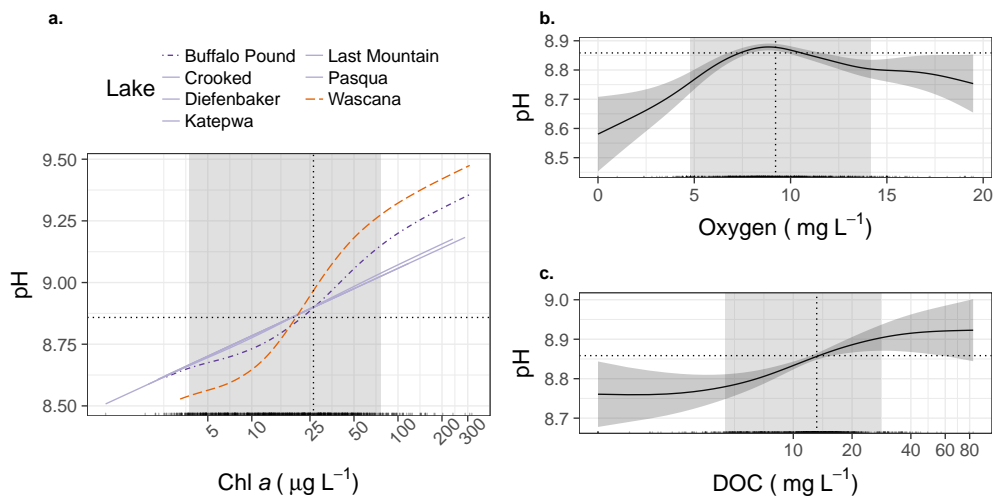
567 **Figure 1.** The seven study sites lie along the Qu'Appelle River (SK, Canada) flowing west to east, with the
568 exception of Wascana (south tributary) and Last Mountain (north tributary).



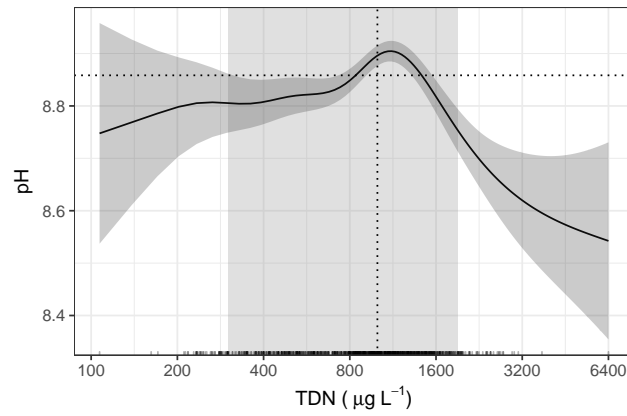
569 **Figure 2.** a: Box plots for limnological data used to calculate carbon dioxide flux in the lakes, showing
 570 medians, upper and lower quartiles, $1.5 \times$ inter-quartile ranges, and 'outliers'. b: Major patterns of annual
 571 variation in pH in all lakes, based on a generalised additive model of pH by Lake, Year, and Day of Year. Rug:
 572 annual means of pH observed over time.



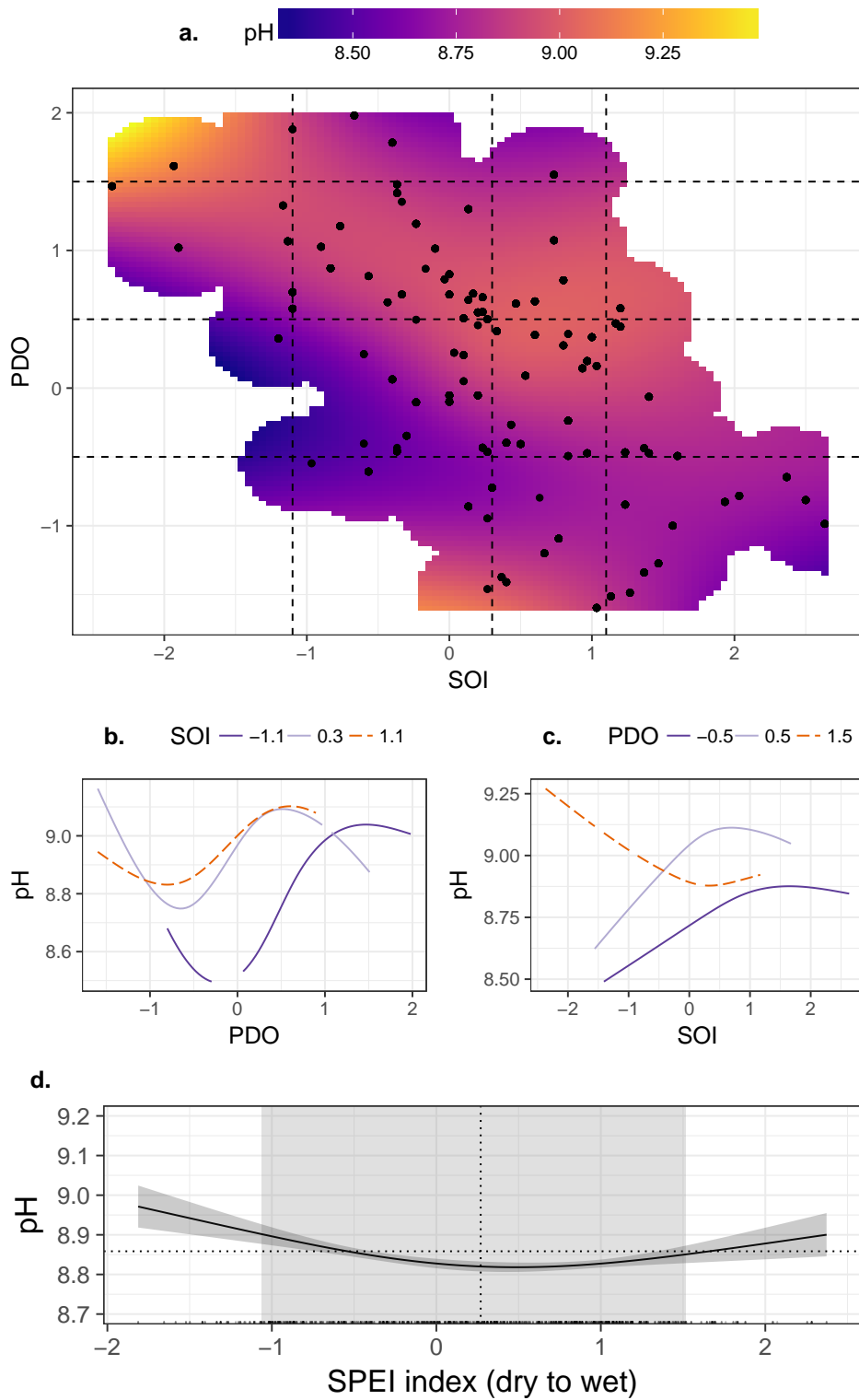
573 **Figure 3.** GAM splines for pH with lake splines significantly different (see section 2.4.3) from the global
 574 spline indicated by color/hue and linetype. Dotted lines: means of y and x; Shaded area: Middle 90% of all
 575 observations. Rug: Data points



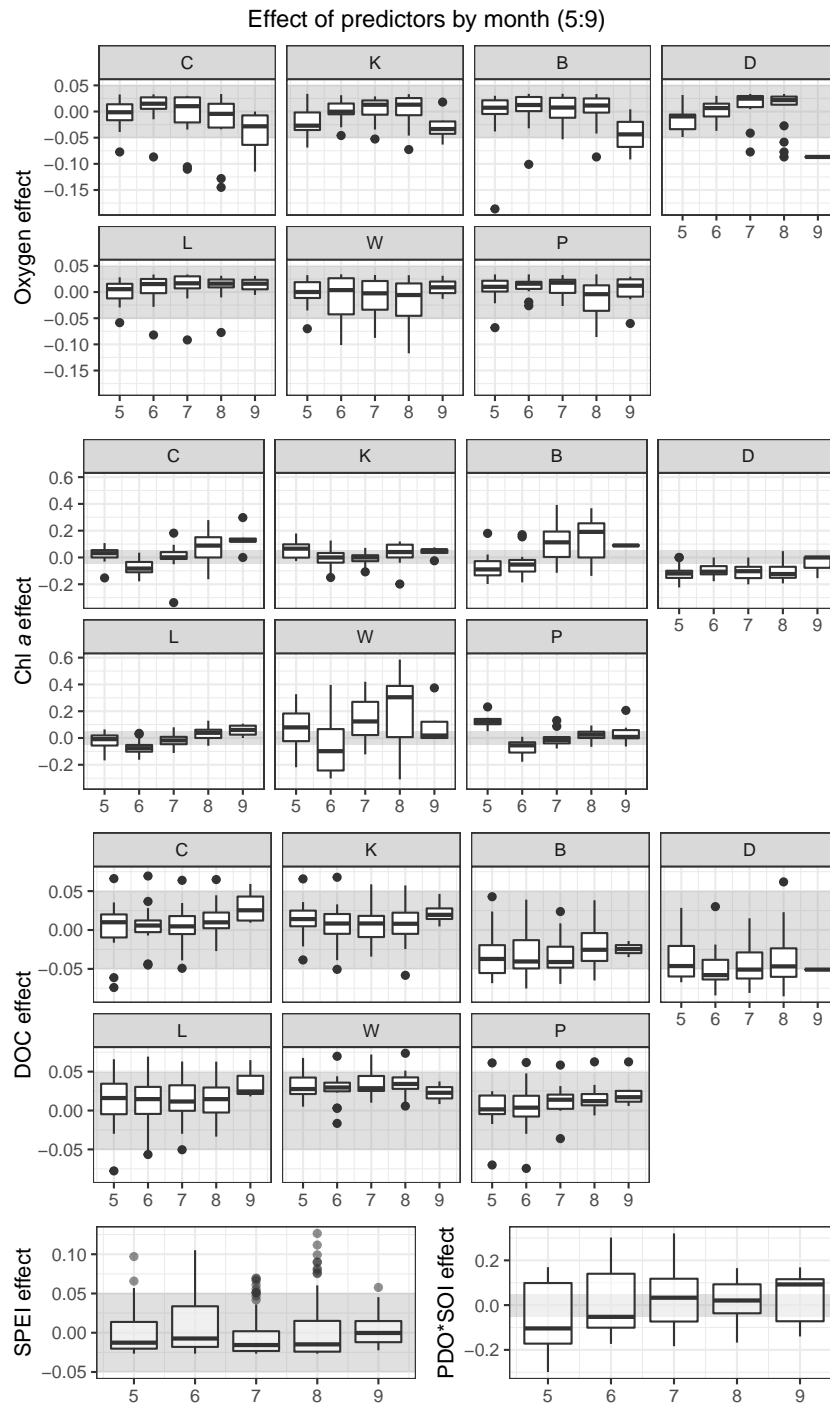
576 **Figure 4.** a-c: GAM splines for significant metabolic variables. Dotted lines: means of y and x; Shaded
 577 area: Middle 90% of all observations. Rug: Data points. a: GAM splines for chlorophyll *a*, with lakes with
 578 significantly different splines to the global spline (see section 2.4.3) indicated by color/hue and linetype. b:
 579 GAM spline of oxygen, with standard errors indicated by shading. c: GAM spline of DOC, with standard
 580 errors indicated by shading.



581 **Figure 5.** GAM spline for TDN in the alternative model without DOC. Dotted lines: means of y and x;
582 Shaded area: Middle 90% of all observations. Rug: Data points. Standard errors are indicated by shading.



583 **Figure 6.** a-c: GAM interactions of PDO and SOI. a: Heatmap with data points. Dashed lines indicate
 584 cross sections for b-c, which show GAM splines for pH for selected combinations of SOI (b) and PDO (c)
 585 values. Missing line segments reflect uncertainties in prediction. d: GAM spline of SPEI, with standard errors
 586 indicated by shading. Rug: Data points.



587 **Figure 7.** Contributions of each predictor to pH summarised over the months of highest data availability,
 588 averaged across lakes for weather and climate indices which were homogenous through the study region. Box
 589 plots show medians, upper and lower quartiles, $1.5 \times$ inter-quartile ranges, and 'outliers'. Shaded area: ± 0.05
 590 regions to aid comparison of magnitudes across predictors.

A: Summary data for all lakes

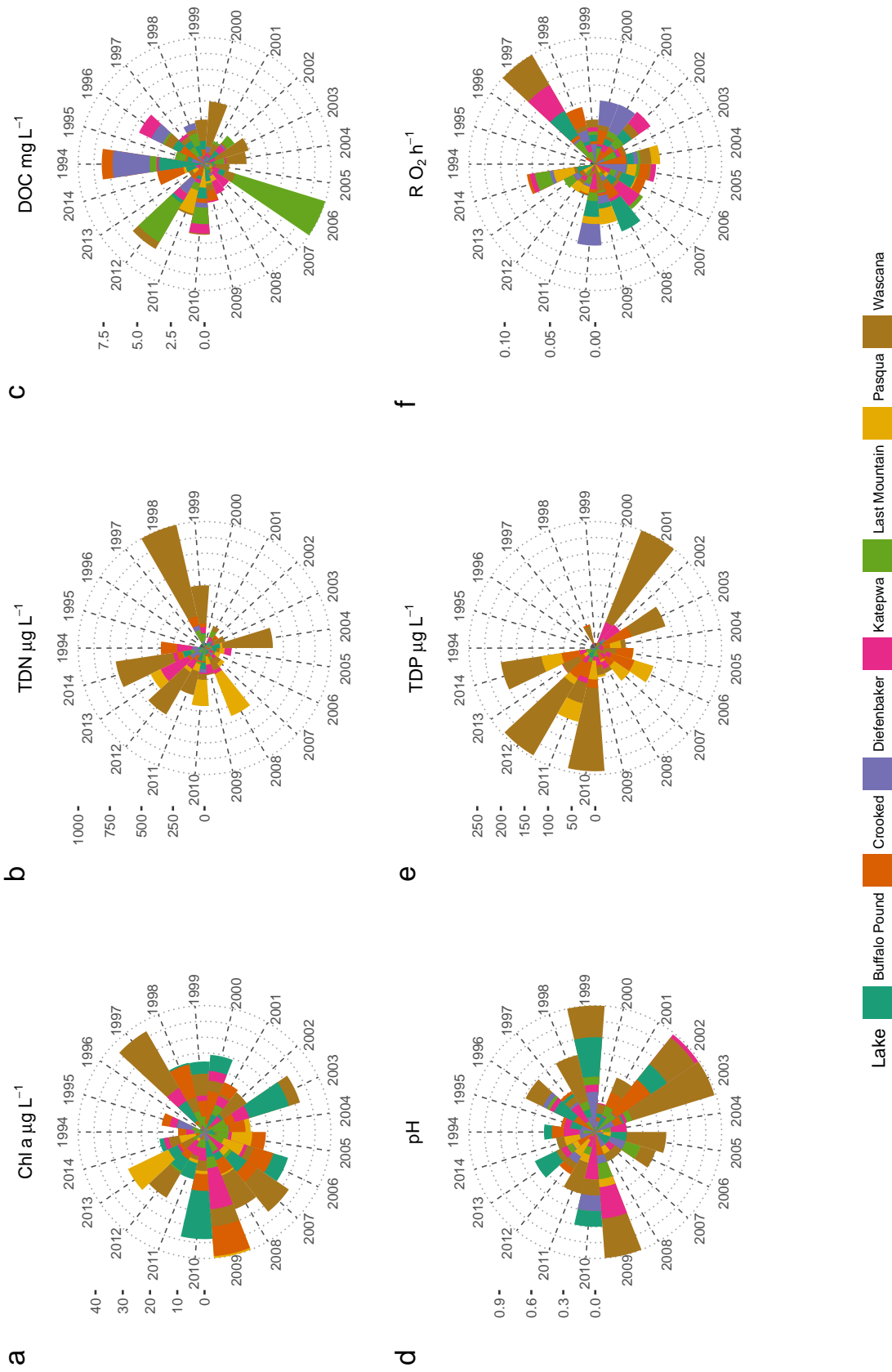
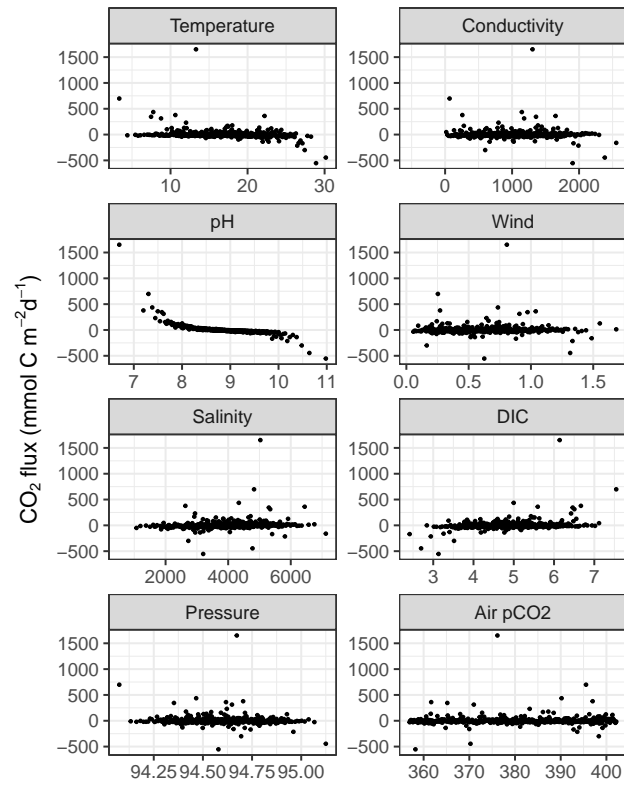


Figure A.1. Intra-annual variability expressed as median absolute deviation (i.e. the median of the absolute deviations from the median) of key metabolic and/or nutrient status variables over the LTER period over the months of most frequent observations (May-September). The data are superimposed such that the lakes with the lowest variability appear toward the centre of the figure, and lakes with higher variability contain the variability of the more central lakes plus the additional value indicated by the colouring.

592

B: Simulated relationships between predictors and CO₂ using sensitivity analysis



593

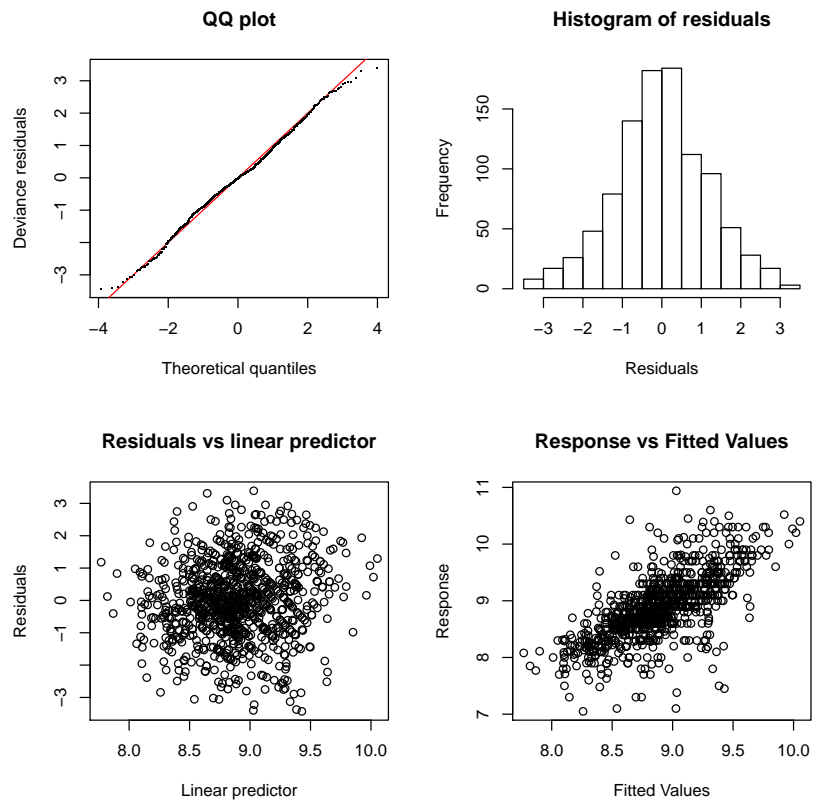
Figure B.1. The relationship between calculated carbon dioxide flux and simulated data sets (N=500) of

594

input variables for sensitivity analysis.

595

C: Model summaries and diagnostic plots



596

Figure C.1. R output for main model diagnostics.

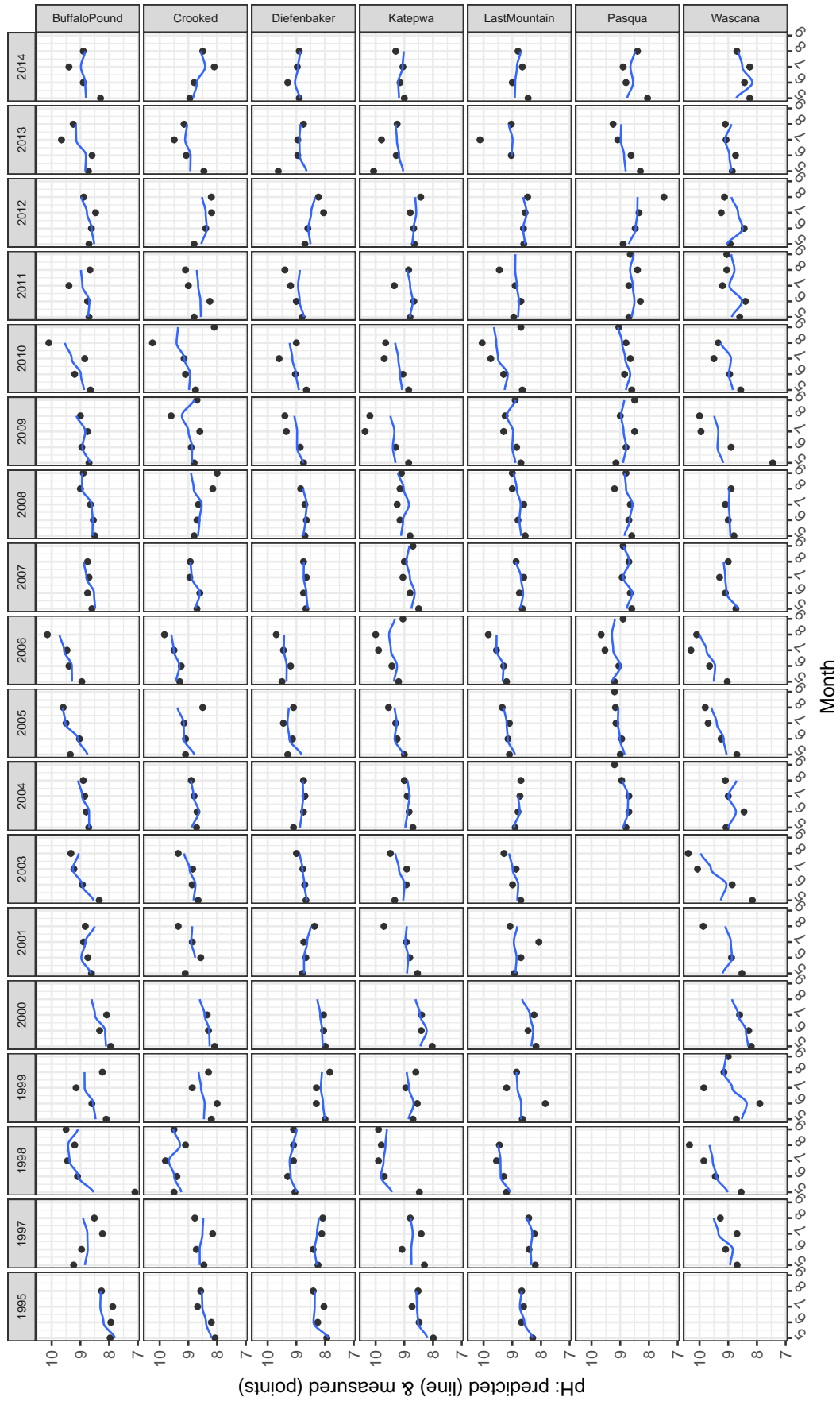


Figure C.2. Measured vs predicted pH over time in the study sites, displayed as monthly means over the months of the most frequent observations.

597 **Acknowledgments**

598 This work was funded by the NSERC Canada Discovery Grants program (PRL, GLS);
599 and Canada Research Chairs, Canada Foundation for Innovation, the Province of Saskatchewan,
600 and the University of Regina (PRL). This paper is a contribution to the Qu'Appelle Long-term
601 Ecological Research Program.

602 All R code and open data for analysis and figures has been archived at <https://github.com/simpson-lab/jgr-co2-flux>. The version containing private data has been archived
603 at <https://github.com/simpson-lab/jgr-co2-flux-private-data>. DOIs
604 for both archives will be provided upon acceptance. Queries regarding private data belonging
605 to the Qu'Appelle Long-term Ecological Research Program may be directed to Peter.Leavitt@uregina.ca.
606

607 We want to thank past and present members of the Limnology Laboratory for assistance
608 with data collection since 1994, and A Chalom for his quick response to changing code on
609 the R PSE package. B Tutolo provided very useful answers to some water chemistry questions.

610 **References**

- 611 Abril, G., S. Bouillon, F. Darchambeau, C. R. Teodoru, T. R. Marwick, F. Tamooch,
612 F. Ochieng Omengo, N. Geeraert, L. Deirmendjian, P. Polsenaere, et al. (2015), Tech-
613 nical note: Large overestimation of pCO₂ calculated from pH and alkalinity in acidic,
614 organic-rich freshwaters, *Biogeosciences*, 12(1), 67–78.
- 615 Alin, S. R., and T. C. Johnson (2007), Carbon cycling in large lakes of the world: A
616 synthesis of production, burial, and lake-atmosphere exchange estimates, *Global Biogeo-*
617 *chemical Cycles*, 21(3), doi:10.1029/2006GB002881.
- 618 Anderson, D. E., R. G. Striegl, D. I. Stannard, C. M. Michmerhuizen, T. A. Mc-
619 Connaughey, and J. W. LaBaugh (1999), Estimating lake-atmosphere CO₂ exchange,
620 *Limnology and Oceanography*, 44(4), 988–1001.
- 621 Audet, J., É. M. Neif, Y. Cao, C. C. Hoffmann, T. L. Lauridsen, S. E. Larsen,
622 M. Søndergaard, E. Jeppesen, and T. A. Davidson (2017), Heat-wave effects on green-
623 house gas emissions from shallow lake mesocosms, *Freshwater Biology*, 62(7), 1130–
624 1142.
- 625 Baehr, M. M., and M. D. DeGrandpre (2004), In situ pCO₂ and O₂ measurements in
626 a lake during turnover and stratification: observations and modeling, *Limnology and*
627 *Oceanography*, 49(2), 330–340.

- 628 Balmer, M. B., and J. A. Downing (2011), Carbon dioxide concentrations in eutrophic
629 lakes: undersaturation implies atmospheric uptake, *Inland Waters*, *1*, 125–132.
- 630 Bogard, M. J., and P. A. del Giorgio (2016), The role of metabolism in modulating
631 CO₂ fluxes in boreal lakes, *Global Biogeochemical Cycles*, *30*(10), 1509–1525, doi:
632 10.1002/2016GB005463.
- 633 Bogard, M. J., D. B. Donald, K. Finlay, and P. R. Leavitt (2012), Distribution and regula-
634 tion of urea in lakes of central north america, *Freshwater Biology*, *57*(6), 1277–1292.
- 635 Bogard, M. J., K. Finlay, M. J. Waiser, V. P. Tumber, D. B. Donald, E. Wiik, G. L. Simp-
636 son, P. A. del Giorgio, and P. R. Leavitt (2017), Effects of experimental nitrogen fer-
637 tilization on planktonic metabolism and co₂ flux in a hypereutrophic hardwater lake,
638 *PLOS ONE*, *12*(12), 1–19, doi:10.1371/journal.pone.0188652.
- 639 Bonsal, B., and A. Shabbar (2008), Impacts of Large-Scale Circulation Variability on
640 Low Streamflows over Canada: A Review, *Canadian Water Resources Journal*, *33*(2),
641 137–154.
- 642 Chalom, A., and P. I. K. L. de Prado (2015), Parameter space exploration of ecological
643 models, <https://arxiv.org/pdf/1210.6278.pdf>, pp. 1–46.
- 644 Chalom, A., and P. I. K. L. de Prado (2016), *pse: Parameter Space Exploration with Latin*
645 *Hypercubes*.
- 646 Christensen, J. P. A., K. Sand-Jensen, and P. A. Staehr (2013), Fluctuating water levels
647 control water chemistry and metabolism of a charophyte-dominated pond, *Freshwater*
648 *Biology*, *58*(7), 1353–1365.
- 649 Cole, J., J. Nina, and F. Caraco (1998), Atmospheric exchange of carbon dioxide in a low-
650 wind oligotrophic lake measured by the addition of SF₆, *Limnology and Oceanography*,
651 *43*, 647–656.
- 652 Cole, Y. T., J. J. and Prairie, N. F. Caraco, W. H. McDowell, L. J. Tranvik, R. G. Striegl,
653 C. M. Duarte, P. Kortelainen, J. A. Downing, J. J. Middelburg, and J. Melack (2007),
654 Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon
655 budget, *Ecosystems*, *10*(1), 172–185.
- 656 Cumming, B., S. Wilson, R. Hall, and J. P. Smol (1995), Diatoms from British Columbia
657 (Canada) Lakes and their Relationship to Salinity, Nutrients and Other Limnological
658 Variables, in *Bibliotheca Diatomologica*, vol. Band 31, edited by H. Lange-Bertalot, pp.
659 1–203, J Cramer, Berlin, Germany.

- 660 Denfeld, B. A., M. B. Wallin, E. Sahlée, S. Sobek, J. Kokic, H. E. Chmiel, and G. A.
661 Weyhenmeyer (2015), Temporal and spatial carbon dioxide concentration patterns in a
662 small boreal lake in relation to ice cover dynamics, *Boreal Environmental Research*, 20,
663 667–78.
- 664 Doctor, D. H., C. Kendall, S. D. Sebestyen, J. B. Shanley, N. Ohte, and E. W. Boyer
665 (2008), Carbon isotope fractionation of dissolved inorganic carbon (DIC) due to out-
666 gassing of carbon dioxide from a headwater stream, *Hydrological Processes*, 22(14),
667 2410–2423.
- 668 Downing, J. A., J. J. Cole, J. J. Middelburg, R. G. Striegl, C. M. Duarte, P. Kortelainen,
669 Y. T. Prairie, and K. A. Laube (2008), Sediment organic carbon burial in agriculturally
670 eutrophic impoundments over the last century, *Global Biogeochemical Cycles*, 22(1),
671 n/a–n/a, doi:10.1029/2006GB002854.
- 672 Dröscher, I., A. Patoine, K. Finlay, and P. R. Leavitt (2009), Climate control of the spring
673 clear-water phase through the transfer of energy and mass to lakes, *Limnology and*
674 *Oceanography*, 54(6/2), 2469–2480.
- 675 Duarte, C. M., Y. T. Prairie, C. Montes, J. J. Cole, R. Striegl, J. Melack, and J. A. Down-
676 ing (2008), CO₂ emissions from saline lakes: A global estimate of a surprisingly large
677 flux, *Journal of Geophysical Research: Biogeosciences*, 113(G4).
- 678 Ducharme-Riel, V., D. Vachon, P. A. del Giorgio, and Y. T. Prairie (2015), The relative
679 contribution of winter under-ice and summer hypolimnetic CO₂ accumulation to the
680 annual CO₂ emissions from northern lakes, *Ecosystems*, 18(4), 547–559.
- 681 Duston, N. M., R. M. Owen, and B. H. Wilkinson (1986), Water chemistry and sedimen-
682 tological observations in Littlefield Lake, Michigan: Implications for lacustrine marl
683 deposition, *Environmental Geology and Water Sciences*, 8(4), 229–236.
- 684 Environment Canada (1979), *Analytical method manual. Inland Waters Directorate*
685 *NAQUADAT 15413*, Environment Canada.
- 686 Evans, J. C., and E. E. Prepas (1996), Potential effects of climate change on ion chemistry
687 and phytoplankton communities in prairie saline lakes, *Limnology and Oceanography*,
688 41(5), 1063–1076.
- 689 Felip, M., and J. Catalan (2000), The relationship between phytoplankton biovolume
690 and chlorophyll in a deep oligotrophic lake: decoupling in their spatial and temporal
691 maxima, *Journal of Plankton Research*, 22(1), 91–106.

- 692 Finlay, K., P. Leavitt, B. Wissel, and Y. Prairie (2009), Regulation of spatial and temporal
693 variability of carbon flux in six hard-water lakes of the northern great plains, *Limnology
694 and Oceanography*, *54*(6), 2553.
- 695 Finlay, K., R. J. Vogt, M. J. Bogard, B. Wissel, B. M. Tutolo, G. L. Simpson, and P. R.
696 Leavitt (2015), Decrease in CO₂ efflux from northern hardwater lakes with increasing
697 atmospheric warming, *Nature*, *519*, 215–218.
- 698 George, D. G., and S. I. Heaney (1978), Factors Influencing the Spatial Distribution of
699 Phytoplankton in a Small Productive Lake, *Journal of Ecology*, *66*(1), 133–155.
- 700 Hall, R. I., P. R. Leavitt, R. Quinlan, A. S. Dixit, and J. P. Smol (1999), Effects of agri-
701 culture, urbanization, and climate on water quality in the northern great plains, *Limnol-
702 ogy and Oceanography*, *44*(3), 739–756.
- 703 Hanson, P. C., D. L. Bade, S. R. Carpenter, and T. K. Kratz (2003), Lake metabolism:
704 Relationships with dissolved organic carbon and phosphorus, *Limnology and Oceanog-
705 raphy*, *48*(3), 1112–1119.
- 706 Hastie, T., and R. Tibshirani (1990), *Generalized Additive Models*, Chapman and
707 Hall/CRC.
- 708 Hollander, D. J., and M. A. Smith (2001), Microbially mediated carbon cycling as a
709 control on the $\delta^{13}\text{C}$ of sedimentary carbon in eutrophic Lake Mendota (USA): New
710 models for interpreting isotopic excursions in the sedimentary record, *Geochimica et
711 Cosmochimica Acta*, *65*(23), 4321–4337.
- 712 Hoover, T. E., and D. C. Berkshire (1969), Effects of hydration on carbon dioxide ex-
713 change across an air-water interface, *Journal of Geophysical Research*, *74*(2), 456–464.
- 714 Howarth, R. W., and A. F. Michaels (2000), *The measurement of primary productivity in
715 aquatic ecosystems*, pp. 72–85, Springer.
- 716 Huttunen, J. T., J. Alm, A. Liikanen, S. Juutinen, T. Larmola, T. Hammar, J. Silvola, and
717 P. J. Martikainen (2003), Fluxes of methane, carbon dioxide and nitrous oxide in bo-
718 real lakes and potential anthropogenic effects on the aquatic greenhouse gas emissions,
719 *Chemosphere*, *52*(3), 609–621.
- 720 Jeffrey, S., and G. F. Humphrey (1975), New spectrophotometric equations for deter-
721 mining chlorophylls *a*, *b*, *c*₁ and *c*₂ in higher plants, algae and natural phytoplankton,
722 *Biochemie und Physiologie der Pflanzen*, *167*(S), 191–194.
- 723 Jonsson, A., J. Karlsson, and M. Jansson (2003), Sources of carbon dioxide supersatura-
724 tion in clearwater and humic lakes in Northern Sweden, *Ecosystems*, *6*(3), 224–235.

- 725 Kling, G., G. Kipphut, and M. Miller (1992), The flux of CO₂ and CH₄ from lakes and
726 rivers in arctic Alaska, *Hydrobiologia*, *240*(23), 23–36.
- 727 Klug, J. L., D. C. Richardson, H. A. Ewing, B. R. Hargreaves, N. R. Samal, D. Vachon,
728 D. C. Pierson, A. M. Lindsey, D. M. O’Donnell, S. W. Effler, and K. C. Weathers
729 (2012), Ecosystem Effects of a Tropical Cyclone on a Network of Lakes in Northeast-
730 ern North America, *Environmental Science & Technology*, *46*(21), 11,693–11,701.
- 731 Knoll, L. B., M. J. Vanni, W. H. Renwick, E. K. Dittman, and J. A. Gephart (2013),
732 Temperate reservoirs are large carbon sinks and small CO₂ sources: Results from high-
733 resolution carbon budgets, *Global Biogeochemical Cycles*, *27*(1), 52–64.
- 734 Koehler, B., E. von Wachenfeldt, D. Kothawala, and L. J. Tranvik (2012), Reactivity con-
735 tinuum of dissolved organic carbon decomposition in lake water, *Journal of Geophysical*
736 *Research: Biogeosciences*, *117*(G1), n/a–n/a, doi:10.1029/2011JG001793.
- 737 Lake, P. (2011), *Drought and Aquatic Ecosystems: Effects and Responses*, 1 ed., Wiley-
738 Blackwell.
- 739 Lapierre, J.-F., and P. A. del Giorgio (2012), Geographical and environmental drivers of
740 regional differences in the lake pCO₂ versus DOC relationship across northern land-
741 scapes, *Journal of Geophysical Research: Biogeosciences*, *117*(G3).
- 742 Last, W. M. (1989), Continental brines and evaporites of the northern Great Plains
743 of Canada, *Sedimentary Geology*, *64*(4), 207–221, doi:https://doi.org/10.1016/
744 0037-0738(89)90049-3.
- 745 Leavitt, P. R., C. S. Brock, C. Ebel, and A. Patoine (2006), Landscape-scale effects of
746 urban nitrogen on a chain of freshwater lakes in central North America, *Limnology and*
747 *Oceanography*, *51*(5), 2262–2277, doi:10.4319/lo.2006.51.5.2262.
- 748 Liu, H., Q. Zhang, G. G. Katul, J. J. Cole, F. S. C. III, and S. MacIntyre (2016), Large
749 CO₂ effluxes at night and during synoptic weather events significantly contribute
750 to CO₂ emissions from a reservoir, *Environmental Research Letters*, *11*(6), doi:
751 10.1088/1748-9326/11/6/064001.
- 752 Maberly, S. C. (1996), Diel, episodic and seasonal changes in pH and concentrations of
753 inorganic carbon in a productive lake, *Freshwater Biology*, *35*(3), 579–598.
- 754 Maberly, S. C., P. A. Barker, A. W. Stott, and M. M. De Ville (2013), Catchment produc-
755 tivity controls CO₂ emissions from lakes, *Nature Climate Change (Letter)*, *3*, 391–394.
- 756 Marcé, R., B. Obrador, J.-A. Morguá, J. L. Riera, P. López, and J. Armengol (2015), Car-
757 bonate weathering as a driver of CO₂ supersaturation in lakes, *Nature Geoscience*, *8*(2),

758 107–111.

759 McDonald, C. P., E. G. Stets, R. G. Striegl, and D. Butman (2013), Inorganic carbon
760 loading as a primary driver of dissolved carbon dioxide concentrations in the lakes
761 and reservoirs of the contiguous united states, *Global Biogeochemical Cycles*, 27(2),
762 285–295, doi:10.1002/gbc.20032.

763 Morales-Pineda, M., A. Cózar, I. Laiz, B. Úbeda, and J. Á. Gálvez (2014), Daily, bi-
764 weekly, and seasonal temporal scales of pCO₂ variability in two stratified mediterranean
765 reservoirs, *Journal of Geophysical Research: Biogeosciences*, 119(4), 509–520.

766 Ojala, A., J. L. Bellido, T. Tulonen, P. Kankaala, and J. Huotari (2011), Carbon gas
767 fluxes from a brown-water and a clear-water lake in the boreal zone during a sum-
768 mer with extreme rain events, *Limnology and Oceanography*, 56(1), 61–76, doi:
769 10.4319/lo.2011.56.1.0061.

770 O'Reilly, C. M., S. Sharma, D. K. Gray, S. E. Hampton, J. S. Read, R. J. Rowley,
771 P. Schneider, J. D. Lenters, P. B. McIntyre, B. M. Kraemer, G. A. Weyhenmeyer,
772 D. Straile, B. Dong, R. Adrian, M. G. Allan, O. Anneville, L. Arvola, J. Austin,
773 J. L. Bailey, J. S. Baron, J. D. Brookes, E. de Eyto, M. T. Dokulil, D. P. Hamilton,
774 K. Havens, A. L. Hetherington, S. N. Higgins, S. Hook, L. R. Izmet'seva, K. D.
775 Joehnk, K. Kangur, P. Kasprzak, M. Kumagai, E. Kuusisto, G. Leshkevich, D. M. Liv-
776 ingstone, S. MacIntyre, L. May, J. M. Melack, D. C. Mueller-Navarra, M. Naumenko,
777 P. Noges, T. Noges, R. P. North, P.-D. Plisnier, A. Rigosi, A. Rimmer, M. Rogora,
778 L. G. Rudstam, J. A. Rusak, N. Salmaso, N. R. Samal, D. E. Schindler, S. G. Schladow,
779 M. Schmid, S. R. Schmidt, E. Silow, M. E. Soylu, K. Teubner, P. Verburg, A. Vouti-
780 lainen, A. Watkinson, C. E. Williamson, and G. Zhang (2015), Rapid and highly vari-
781 able warming of lake surface waters around the globe, *Geophysical Research Letters*,
782 42(24), 10,773–10,781.

783 Osburn, C. L., C. R. Wigdahl, S. C. Fritz, and J. E. Saros (2011), Dissolved organic mat-
784 ter composition and photoreactivity in prairie lakes of the U.S. Great Plains, *Limnology
785 and Oceanography*, 56(6), 2371–2390.

786 Ostapenia, A. P., A. Parparov, and T. Berman (2009), Lability of organic carbon in lakes
787 of different trophic status, *Freshwater Biology*, 54(6), 1312–1323.

788 Patoine, A., M. D. Graham, and P. R. Leavitt (2006), Spatial variation of nitrogen fixation
789 in lakes of the northern Great Plains, *Limnology and Oceanography*, 51(4), 1665–1677,
790 doi:10.4319/lo.2006.51.4.1665.

- 791 Perga, M.-E., S. C. Maberly, J.-P. Jenny, B. Alric, C. Pignol, and E. Naffrechoux (2016),
 792 A century of human-driven changes in the carbon dioxide concentration of lakes,
 793 *Global Biogeochemical Cycles*, 30(2), 93–104.
- 794 Pham, S. V., P. R. Leavitt, S. McGowan, B. Wissel, and L. I. Wassenaar (2009), Spatial
 795 and temporal variability of prairie lake hydrology as revealed using isotopes of hydro-
 796 gen and oxygen, *Limnology and Oceanography*, 54(1), 101–118.
- 797 Pomeroy, J. W., D. de Boer, and L. Martz (2007), *Hydrology and water resources*, pp.
 798 63–80, Canadian Plains Research Centre.
- 799 Prairie, Y. T. (2008), Carbocentric limnology: looking back, looking forward, *Canadian*
 800 *Journal of Fisheries and Aquatic Sciences*, 65(3), 543–548.
- 801 Quiñones-Rivera, Z. J., K. Finlay, R. J. Vogt, P. R. Leavitt, and B. Wissel (2015), Hydro-
 802 logic, metabolic and chemical regulation of water-column metabolism and atmospheric
 803 CO₂ exchange in a large continental reservoir during spring and summer, *Journal of*
 804 *Great Lakes Research*, 41(Supp. 2), 144 – 154.
- 805 R Development Core Team (2016), *R: A Language and Environment for Statistical Comput-*
 806 *ing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.
- 807 Raymond, P. A., J. Hartmann, R. Lauerwald, S. Sobek, C. McDonald, M. Hoover, D. But-
 808 man, R. Striegl, E. Mayorga, C. Humborg, P. Kortelainen, H. Dürr, M. Meybeck,
 809 P. Ciais, and P. Guth (2013), Global carbon dioxide emissions from inland waters,
 810 *Nature*, 503, 355–359.
- 811 Reis, P., and F. Barbosa (2014), Diurnal sampling reveals significant variation in CO₂
 812 emission from a tropical productive lake, *Brazilian Journal of Biology*, 74(3), S113–
 813 S119.
- 814 Seekell, D. A., and C. Gudasz (2016), Long-term pCO₂ trends in Adirondack Lakes,
 815 *Geophysical Research Letters*, 43(10), 5109–5115.
- 816 Shabbar, A., and B. Yu (2012), Intraseasonal Canadian Winter Temperature Responses
 817 to Interannual and Interdecadal Pacific SST Modulations, *Atmosphere-Ocean*, 50(1),
 818 109–121.
- 819 Shabbar, A., B. R. Bonsal, and K. Szeto (2011), Atmospheric and Oceanic Variability
 820 Associated with Growing Season Droughts and Pluvials on the Canadian Prairies,
 821 *Atmosphere-Ocean*, 49(4), 339–355.
- 822 Shao, C., J. Chen, C. A. Stepien, H. Chu, Z. Ouyang, T. B. Bridgeman, K. P. Czajkowski,
 823 R. H. Becker, and R. John (2015), Diurnal to annual changes in latent, sensible heat,

- 824 and CO₂ fluxes over a Laurentian Great Lake: A case study in Western Lake Erie,
825 *Journal of Geophysical Research: Biogeosciences*, 120(8), 1587–1604.
- 826 Sobek, S., L. J. Tranvik, and J. J. Cole (2005), Temperature independence of carbon
827 dioxide supersaturation in global lakes, *Global Biogeochemical Cycles*, 19(2), 1–15.
- 828 Søndergaard, M., N. H. Borch, and B. Riemann (2000), Dynamics of biodegradable DOC
829 produced by freshwater plankton communities, *Aquatic Microbial Ecology*, 23(1), 73–
830 83.
- 831 Soumis, N., É. Duchemin, R. Canuel, and M. Lucotte (2004), Greenhouse gas emissions
832 from reservoirs of the western United States, *Global Biogeochemical Cycles*, 18, doi:
833 10.1029/2003GB002197.
- 834 Stets, E. G., R. G. Striegl, G. R. Aiken, D. O. Rosenberry, and T. C. Winter (2009), Hy-
835 drologic support of carbon dioxide flux revealed by whole-lake carbon budgets, *Journal*
836 *of Geophysical Research: Biogeosciences*, 114(G1), n/a–n/a, doi:10.1029/2008JG000783.
- 837 Stets, E. G., D. Butman, C. P. McDonald, S. M. Stackpoole, M. D. DeGrandpre, and
838 R. G. Striegl (2017), Carbonate buffering and metabolic controls on carbon dioxide in
839 rivers, *Global Biogeochemical Cycles*, 31(4), 663–677, doi:10.1002/2016GB005578.
- 840 Striegl, R. G., and C. M. Michmerhuizen (1998), Hydrologic influence on methane and
841 carbon dioxide dynamics at two north-central Minnesota lakes, *Limnology and Oceanog-*
842 *raphy*, 43(7), 1519–1529.
- 843 Stumm, W., and J. J. Morgan (1996), *Aquatic chemistry: Chemical equilibria and rates in*
844 *natural waters*, Wiley.
- 845 Tranvik, L. J., J. A. Downing, J. B. Cotner, S. A. Loiselle, R. G. Striegl, T. J. Bal-
846 latore, P. Dillon, K. Finlay, K. Fortino, L. B. Knoll, P. L. Kortelainen, T. Kutser,
847 S. Larsen, I. Laurion, D. M. Leech, S. L. McCallister, D. M. McKnight, J. M. Melack,
848 E. Overholt, J. A. Porter, Y. Prairie, W. H. Renwick, F. Roland, B. S. Sherman, D. W.
849 Schindler, S. Sobek, A. Tremblay, M. J. Vanni, A. M. Verschoor, E. von Wachenfeldt,
850 and G. A. Weyhenmeyer (2009), Lakes and reservoirs as regulators of carbon cycling
851 and climate, *Limnology and Oceanography*, 54(6part2), 2298–2314.
- 852 Vachon, D., and Y. T. Prairie (2013), The ecosystem size and shape dependence of gas
853 transfer velocity versus wind speed relationships in lakes, *Canadian Journal of Fisheries*
854 *and Aquatic Sciences*, 70, 1757–1764.
- 855 van der Kamp, G., D. Keir, and M. S. Evans (2008), Long-Term Water Level Changes in
856 Closed-Basin Lakes of the Canadian Prairies, *Canadian Water Resources Journal*, 33(1),

857 23–38.

858 Vicente-Serrano, S. M., S. Beguería, and J. I. López-Moreno (2016), A Multiscalar
859 Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapo-
860 transpiration Index, *Journal of Climate*, *23*, 1696–1718.

861 Vogt, R. J., J. A. Rusak, A. Patoine, and P. R. Leavitt (2011), Differential effects of en-
862 ergy and mass influx on the landscape synchrony of lake ecosystems, *Ecology*, *92*,
863 1104–1114.

864 Vogt, R. J., S. Sharma, and P. R. Leavitt (2015), Decadal regulation of phytoplankton
865 abundance and water clarity in a large continental reservoir by climatic, hydrologic and
866 trophic processes, *Journal of Great Lakes Research*, *41*, 81–90.

867 Wanninkhof, R. (1992), Relationship between wind speed and gas exchange over the
868 ocean, *Journal of Geophysical Research: Oceans*, *97*(C5), 7373–7382.

869 Wanninkhof, R., and M. Knox (1996), Chemical enhancement of CO₂ exchange in natural
870 waters, *Limnology and Oceanography*, *41*(4), 689–697.

871 Weyhenmeyer, G. A., S. Kosten, M. B. Wallin, L. J. Tranvik, E. Jeppesen, and F. Roland
872 (2015), Significant fraction of CO₂ emissions from boreal lakes derived from hydro-
873 logic inorganic carbon inputs, *Nature Geoscience*, *8*, 933–936.

874 Wood, S. (2011), Fast stable restricted maximum likelihood and marginal likelihood es-
875 timation of semiparametric generalized linear models, *Journal of the Royal Statistical*
876 *Society (B)*, *73*(1), 3–36.

877 Wood, S. (2017), *Generalized Additive Models: An Introduction with R*, 2. ed., Chapman
878 and Hall/CRC.

879 Zhang, Y., L. Feng, J. Li, L. Luo, Y. Yin, M. Liu, and Y. Li (2010), Seasonal–spatial vari-
880 ation and remote sensing of phytoplankton absorption in Lake Taihu, a large eutrophic
881 and shallow lake in China, *Journal of Plankton Research*, *32*(7), 1023–1037.

List of Changes

Replaced: ~~need-to-be~~ replaced with: **be needed to**, on page 3, line 73.

Added: ~~Calculated values are widely applied in the absence of measurements, particularly when long-term or broad spatial data are being examined (e.g. Duarte et al. [2008]; Seekell and Gudasz [2016]),~~ on page 6, line 165.

Added: ~~Our analytical approach follows a few key underlying considerations. Since CO₂ flux was estimated from water chemistry and physical variables and not measured directly, we avoided any approach that would circularly include these 'calculation variables' as metabolic or climatic proxy predictors of CO₂ flux. Furthermore, we were specifically interested in which of these calculation variables correlate the most with CO₂ flux in our study region. In this regard we note that, although the real, rather than estimated, relationship between these variables and CO₂ flux is unknown, this step can identify which variable is key to proxy CO₂ flux in our region (and conversely, which variables are not). Therefore, we first quantified the influence of the calculation variables on estimated CO₂ flux ('influence' here used in the regression sense of changes in x influencing estimates of y , rather than a directional causal sense). Secondly, we regressed our metabolic and climatic variables of interest against the variable that accounted for most of this variation. The second step allowed us to use a measured, rather than estimated, response variable, reducing the amount of imprecision in our regression values. We were then able to relate these values back into CO₂ flux estimates using the results from the first step, thereby avoiding presenting misleadingly precise results for CO₂ flux itself.,~~ on page 8, line 202.

Deleted: ~~Since pCO₂ was estimated from water chemistry parameters and not measured directly, we avoided modeling CO₂ flux using any variables involved in the calculations. Instead, we staggered analysis into two stages, first selecting the input variable which accounted for most of the variation in CO₂ flux, and then regressing the remaining metabolic and climatic variables against that selected variable.,~~ on page 8, line 219.

Deleted: ~~(Figs 4, 6, 3),~~ on page 12, line 332.

Added: ~~(Figs 4 – 6),~~ on page 12, line 332.

Replaced: ~~favoured~~ replaced with: **correlates with**, on page 12, line 343.

Replaced: ~~favoured~~ replaced with: **correspond with**, on page 13, line 356.

Replaced: ~~favouring~~ replaced with: **corresponding with**, on page 15, line 431.