



Challenges of predicting gas transfer velocity from wind measurements over global lakes

Marcus Klaus^{1,2} · Dominic Vachon¹

Received: 18 October 2019 / Accepted: 25 April 2020 / Published online: 1 May 2020
© The Author(s) 2020

Abstract

Estimating air–water gas transfer velocities (k) is integral to understand biogeochemical and ecological processes in aquatic systems. In lakes, k is commonly predicted using wind-based empirical models, however, their predictive performance under conditions that differ from their original calibration remains largely unassessed. Here, we collected 2222 published k estimates derived from various methods in 46 globally distributed lakes to (1) evaluate the predictions of a selection of six available wind-speed based k models for lakes and (2) explore and develop new empirical models to predict k over global lakes. We found that selected k models generally performed poorly in predicting k in lakes. Model predictions were more accurate than simply assuming a mean k in only 2–39% of all lakes, however, we could not identify with confidence the specific conditions in which some models outperformed others. We developed new wind-based models in which additional variables describing the spatial coverage of k estimates and the lake size and shape had a significant effect on the wind speed- k relationship. Although these new models did not fit the global dataset significantly better than previous k models, they generate overall less biased predictions for global lakes. We further provide explicit estimates of prediction errors that integrate methodological and lake-specific uncertainties. Our results highlight the potential limits when using wind-based models to predict k across lakes and urge scientists to properly account for prediction errors, or measure k directly in the field whenever possible.

Keywords Air–water gas exchange · Model assessment · Lake gas flux · Wind speed · k_{600} · Reaeration

Introduction

Estimating gas fluxes across the air–water interface in lakes is fundamental for understanding their biogeochemical, environmental and ecological functioning. Accurate estimates of lakes carbon (CO_2 and CH_4), nitrogen (N_2 and N_2O) and oxygen (O_2) fluxes with the atmosphere are key to constrain their biogeochemical cycles (Likens 2010), quantify whole lake metabolism (Dugan et al. 2016) and evaluate greenhouse gas emissions at regional and global scales (Raymond

et al. 2013; Soued et al. 2015; Hastie et al. 2017). Estimating air–water fluxes of environmental aquatic contaminants (Hornbuckle et al. 1994; Bidleman 1999; Poissant et al. 2000) and biogenic volatile organic compounds (VOC) in lakes (Fink 2007) is also critical for preserving lake ecosystem services. Gas flux across the air–water interface (F) is described by using Fick's first law of diffusion as the difference in the surface water (C_{wtr}) and air-equilibrium (C_{eq}) gas concentrations, multiplied by the air–water gas transfer velocity (k):

$$F = k \cdot (C_{wtr} - C_{eq}), \quad (1)$$

where positive F implies flux from the water to the atmosphere. There are methods to estimate gas fluxes such as the use of floating chambers (Engle and Melack 2000; Matthews et al. 2003) or by eddy covariance (Anderson et al. 1999). However, these methods can be time and/or cost consuming and may be difficult to be applied to capture potential spatial variation in multiple systems at the same time (Cole et al. 2010; Schilder et al. 2013; Erkkilä et al. 2018).

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00027-020-00729-9>) contains supplementary material, which is available to authorized users.

✉ Marcus Klaus
marcus.klaus@posteo.net

¹ Department of Ecology and Environmental Science, Umeå University, Umeå, Sweden

² Department of Forest Ecology and Management, Swedish University of Agricultural Sciences, Umeå, Sweden

Alternatively, fluxes can be modelled using Eq. 1, based on gas concentrations and k . However, while gas concentrations are usually relatively straightforward to measure, k is by far more difficult to estimate with high confidence.

The air–water gas transfer velocity of sparingly soluble gases is driven by near-surface water turbulence, which in lakes is mainly generated by wind stress over the lake surface and thermal convection when colder waters masses at the surface sink due to greater density (MacIntyre et al. 1995). Wind is the main source of near-surface turbulence in many lakes (Read et al. 2012), however, the efficiency of this wind-to-turbulence transfer may be modulated by several lake characteristics and other hydrodynamic processes. For a given wind speed, larger fetch lengths will result in larger wave heights, greater turbulence, and thus higher k values (Schilder et al. 2013; Vachon and Prairie 2013; Gålfalk et al. 2013). Surface heat flux, which determines whether a lake is warming or cooling, also affects near-surface turbulence. A warming surface water will stratify thermally and will suppress the wind-driven turbulence (negative buoyancy flux), while a cooling lake surface will generate additional turbulence from the convective movements of water masses (positive buoyancy flux) (MacIntyre et al. 2010). This effect is related to lake area and latitude (Read et al. 2012). Using wind to predict k in lakes is thus far from being direct, however, compared with turbulence measurements, wind speed is relatively easy to measure and therefore the most widely used predictor in empirical models of k for lakes.

Several empirical wind-based k models are currently available for lakes, with k often standardized to a Schmidt number of 600 (k_{600}) to characterize CO_2 transfer at 20 °C water temperature. However, each model was calibrated in a distinct system, under specific conditions, and using different methods to determine k_{600} . This results in a wide range of model parameterizations (Table 1). The most

common methods to derive k_{600} include floating chambers (e.g. Vachon and Prairie 2013), mass balance approaches (e.g. Cole and Caraco 1998) and the eddy covariance technique (e.g. MacIntyre et al. 2010). These methods are fundamentally different in their approach, in addition to having specific issues that may affect the resulting k_{600} estimates (see MacIntyre et al. (1995), Jähne and Haussecker (1998), Wanninkhof et al. (2009) and Cole et al. (2010) for more detailed discussions). A key difference between methods, however, is related to their scales of spatial and temporal integration (SIN and TIN , respectively), ranging from centimeters and minutes (floating chambers); to meters and hours (eddy covariance technique) and the whole lake and days (mass balance approach) (Fig. 1a). The different SIN and TIN will have implications for the measured k_{600} for a given wind speed. For example, when the relationship between k_{600} and wind speed has a positive curvature, the average k_{600} derived over a period of variable winds will be greater than over a period of steady winds of the same average wind speed (Wanninkhof et al. 1987; Livingstone and Imboden 1993). Longer TIN may thus increase its estimated value for a given averaged wind speed. The effect of SIN is less well understood and documented (Schilder et al. 2013; Paranaíba et al. 2018). Wind measured in the center of the lake may potentially record higher wind speeds than the average wind speed integrated over the whole lake area (Venäläinen et al. 2003; Docquier et al. 2016). This intra-lake spatial variability depends on the size and shape of the lake and whether the lake is sheltered by trees, mountains or buildings (Kwan and Taylor 1994; Markfort et al. 2010; Vachon and Prairie 2013). As a result, complex-shape lakes should, in theory, have greater within-lake spatial variability in wind-driven turbulence and thus k_{600} . The average k_{600} derived from mass balances (whole-lake, i.e. $SIN=1$) should thus be lower than the measured k_{600} at the center of the lake

Table 1 Models for predicting the air–water gas transfer velocity (k_{600} , in cm h^{-1}) based on wind speed at 10 m height (U_{10} , in m s^{-1}) alone or in combination with lake area (LA , in km^2)

Model	Equation	Number of lakes or reservoirs	Number of samples or bins	Data binned	R^2	Method used
CC98	$2.07 + 0.215 \cdot U_{10}^{1.7}$	14	10	Yes	–	Mass balance
VP13	$2.51 \pm 0.99 + 1.48 \pm 0.34 \cdot U_{10} + 0.39 \pm 0.08 \cdot U_{10} \cdot \log_{10}(LA)$	11	64	No	0.68	Floating chambers
CW03	$0.168 + 0.228 \cdot U_{10}^{2.2}$	1	17	No	0.93	Mass balance
M10	$0.16 \pm 1.65 + 2.25 \pm 0.33 \cdot U_{10}$	1	8	Yes	0.91	Eddy covariance
G07	$1.66 \pm 0.34 \cdot \exp(0.26 \pm 0.04 \cdot U_{10})$	1	7	Yes	0.92	Eddy covariance + Floating chambers
L18	$9.038 \cdot U_{10} + 3.592$	1	7	No	0.77	Floating chambers

Also shown are the extent of the training datasets, the coefficient of determination (R^2), and the method used for k_{600} estimations

CC98 Cole and Caraco (1998), VP13 Vachon and Prairie (2013), CW03 Crusius and Wanninkhof (2003), M10 MacIntyre et al. (2010), G07 Guérin et al. (2007), L18 Li (2018)

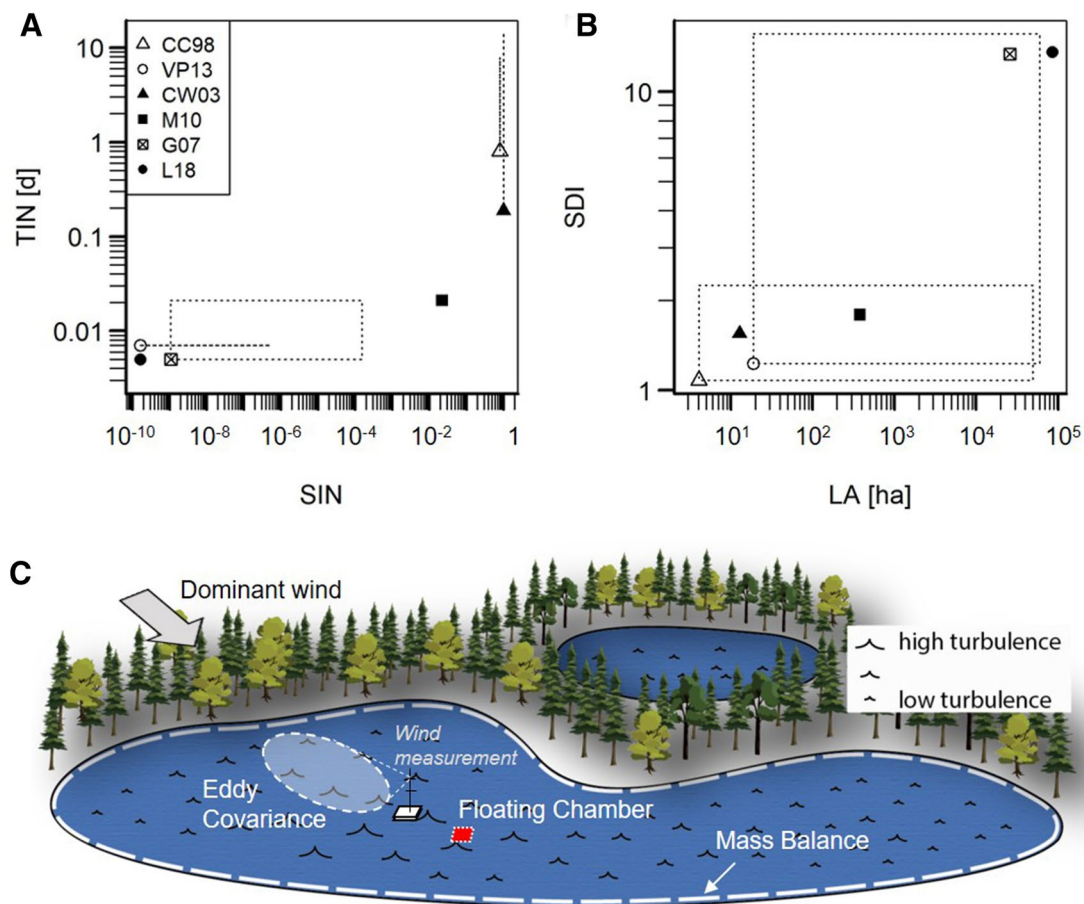


Fig. 1 Spatial and temporal scales, and the prediction space of common wind-based air–water gas transfer velocity (k) models in lakes. **a** Models vary in the space- (SIN) and time integration (TIN) of the underlying method to measure k , and **b** in the surface area (LA) and the shoreline development index (SDI) of the lakes they are calibrated for. In **a** and **b**, dashed and dotted lines cover the range of conditions in case they varied throughout the training dataset. **c** For a given dominant wind direction lake surface turbulence will be heterogene-

ous due to lake shape, size and sheltering effect. The different spatial integrations of k measurements between floating chambers (red area), eddy covariance (white shaded area), and mass balance (whole lake area) should result in different average k values. CC98 Cole and Caraco (1998), VP13 Vachon and Prairie (2013), CW03 Crusius and Wanninkhof (2003), M10 MacIntyre et al. (2010), G07 Guérin et al. (2007) and L18 Li (2018) (color figure online)

from the eddy covariance technique ($SIN < 1$) and floating chambers ($SIN \ll 1$) (Fig. 1c).

Available wind-based k_{600} models also differ in the number and geometrical dimensions of the lakes they are calibrated against, in particular the surface area (LA) and the shoreline development index (SDI), the latter describing the ratio of the shoreline length of a given lake relative to the shoreline length of a circular lake of the same area (Fig. 1b). Each model has been calibrated under specific conditions. How the models perform under different conditions still remains widely unknown. This uncertainty is typically dealt with by averaging across predictions from a number of different k_{600} models (Raymond et al. 2013; Hastie et al. 2017). However, model averaging does not necessarily reduce prediction errors nor provide more accurate estimates (Dormann et al. 2018). Existing uncertainties in

k_{600} modelling call for a systematic evaluation of the context-dependence of the bias of currently used k_{600} models and attempts to develop new models in order to test whether uncertainties can be reduced and quantified relative to environmental conditions.

Our first aim is to evaluate the predictions of a selection of currently available wind-based k_{600} models using published k_{600} measurements in global lakes. We selected six commonly used or recently published models that differ in their methods of k_{600} measurements, their SIN and TIN and the geometry of the lake under which the models were calibrated (Fig. 1 and Table 1). Specifically, we assess context-dependent model prediction performance using multivariate regression tree analyses and discuss how the environmental conditions (average wind speed), the geometry (LA and SDI) and latitude (Lat) of the lakes studied, and the

relevant scales of *SIN* and *TIN* related to the method affect the model performance. We hypothesized that none of the available models generally performs better than other models (H1), but that certain models outperform others under certain environmental conditions or system characteristics (H2). Our second aim is to explore new parametrizations using the existing lake k_{600} data that allow more flexibility by accounting for lake-specific effects and explicitly provide prediction errors. We hypothesized that the wind speed effect on k_{600} will vary among the different lakes due to their different *LA*, *SDI*, *SIN* or *Lat* (H3) and hence including these additional variables will generally improve k_{600} predictions among a wide range of lakes (H4).

Materials and methods

Data compilation and standardization

We compiled data on 2297 simultaneous estimates of k and wind speed from 79 global lakes and 10 reservoirs (Online Resources Table S1), however, only a subset was used to accommodate our analyses. The data were retrieved from all peer-reviewed scientific papers ($n = 46$) we could find via the search engine “web of science” and papers cited therein. As keywords, we used “lake”, “pond” or “reservoir” in combination with “wind”, and either “ k_{600} ”, “gas exchange”, “piston velocity”, “reaeration” or “gas transfer”. We only included k estimates that were based on sparingly soluble gases where air–water gas exchange is dominantly water-side controlled. Data were either provided by corresponding authors, extracted from tables, or digitized from figures using the web tool WebPlotDigitizer (Rohatgi 2019) as indicated in Online Resources Table S1. All k estimates were derived from the floating chamber, mass balance or eddy covariance approach and based on a variety of tracer gases (carbon dioxide, methane, oxygen, radon, helium, neon, krypton, sulfur hexafluoride, mercury and propane). The eddy covariance data were binned according to wind speed following the methodology described in the respective original papers. We documented the method-specific *SIN* and *TIN* as described in Online Resources Text S1. *SIN* was expressed relative to the total lake surface area. To make the data comparable across studies, we used k_{600} and wind speed at 10 m height above ground (U_{10}). If k values were not reported as k_{600} , we converted the estimates of k for a given Schmidt number (Sc) to a Schmidt number of 600 $k_{600} = k \left(\frac{600}{Sc} \right)^{-n}$ where n was $2/3$ for low wind speeds ($U_{10} \leq 3.7 \text{ m s}^{-1}$) and $1/2$ for high wind speeds ($U_{10} > 3.7 \text{ m s}^{-1}$) as suggested by Jähne et al. (1987). We used Schmidt number parameterizations by Wanninkhof (2014) for carbon dioxide, methane, oxygen, radon, helium (^3He), neon, krypton, and sulfur

hexafluoride, by Kuss et al. (2009) for mercury and by Witherspoon and Saraf (1965) for propane. Parameterizations were chosen for fresh- or saltwater, depending on the salinity classification in the original paper. For each lake, we also compiled *Lat*, *LA* and *SDI*, calculated following Hutchinson (1957): $SDI = \frac{p}{2\sqrt{\pi LA}}$, where p is the lake perimeter. These measures were extracted from the HydroLakes dataset (for $LA > 10$ ha, Messenger et al. (2016)), by on-screen digitization of lake surfaces in Google Maps, or, if available, as shown in maps provided in the original papers (for $LA \leq 10$ ha). For Swedish lakes, we relied on the ViVaN dataset because of its higher accuracy and resolution relative to the HydroLakes dataset (Nisell et al. 2007).

General analytical approach

We first evaluated the predictions of each of the six k_{600} models listed in Table 1 relative to observed k_{600} in our dataset. We then analyzed context-dependent model biases using multivariate regression tree (MRT) analysis to identify experimental and method-specific conditions (average U_{10} , *LA*, *SDI*, *SIN*, *TIN*, absolute value of *Lat*) under which certain k_{600} models would perform better than other models. Finally, we explored new k_{600} parametrizations using multivariate regression analysis. For these analyses, we only included those 46 lakes with at least six k_{600} observations each, and a total of 2222 observations. We chose this threshold to maximize the total number of observations and the number of lakes, but fulfill recommendations of a minimum larger than five observations per lake (Theall et al. 2011) and an average larger than 30 observations per lake (Scherbaum and Ferreter 2009, Online Resources Fig. S1).

Evaluating available wind-based k_{600} models

We evaluated models by comparing observed and predicted k_{600} based on four performance measures and summarized these in an integrated performance index. First, we calculated the root mean squared deviation (RMSD) of the predicted values relative to the 1:1 line of observed vs. predicted values (Piñeiro et al. 2008). The smaller the RMSD, the better is the model fit. Second, we calculated the coefficient of determination, $R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$, where $SS_{res} = \sum_i (y_i - f_i)^2$ is the residual sum of square and $SS_{tot} = \sum_i (y_i - \bar{y}_i)^2$ is the total sum of squares with observed values y_i , predicted values f_i and the mean of the observed data \bar{y}_i . To account for differences in model complexity, we further adjusted R^2 for the number of predictor variables v relative to the number of observations m : $R^2_{adj} = 1 - (1 - R^2)(m - 1)/(m - v - 1)$. $R^2_{adj} = 1$ implies perfect fit, $R^2_{adj} > 0$ implies the model fits the data better than

the arithmetic mean of the data, and $R^2_{adj} < 0$ implies that the arithmetic mean fits the data better than the model predictions. Third and fourth, we calculated the intercept and slope of the linear regression line of observed vs. predicted k_{600} . We tested whether the intercept was significantly different from zero based on the significance of the intercept of the linear regression of observed vs. predicted k_{600} . We also tested whether the slope was significantly different from 1 based on the significance of the slope of the linear regression of observed-minus-predicted vs. predicted k_{600} . The larger the p -value of the intercept and slope, the closer the regression line of observed vs. predicted k_{600} is to the 1:1 line (intercept=0, slope=1). Linear regressions were fit using the ‘lm’ function in R 3.6.1 (R Core Team 2019). We ranked each model, with the highest rank assigned to the lowest RMSD, highest R^2_{adj} , and highest p -values of the intercept and slope tests explained above. We used the median rank among all performance measures as an integrated performance index scaling from 1 (best performance) to 6 (worst performance).

We used MRT analysis (De’Ath 2002) to evaluate if any models perform better than others under certain conditions. MRT forms groups (leaves) of lakes by repeated splitting of the data. Each split is defined by a simple rule based on specific conditions and is selected to cluster together lakes for which the different model performance patterns are similar. For example, a cluster could be created if under specific conditions hypothetical models A and B perform better than hypothetical models C and D. In practice, MRT clusters a matrix of dependent variables under the constraints of a matrix of independent variables. As dependent variables, we chose the integrated performance index of the different k_{600} models. As independent variables, we chose average U_{10} , LA , SDI , the method-specific SIN and average TIN , and the absolute value of Lat . To account for potential variability in tree structures depending on the threshold number of lakes to be included per leaf, we fitted a series of trees with threshold numbers ranging from 1 to 23. We fitted MRTs using the ‘mvpart’ function of the R package ‘mvpart’ (De’Ath 2014), using the Euclidean distance as a dissimilarity index. We selected the best tree within one standard error of the overall best using tenfold cross-validation. We evaluated the extent to which the MRT can explain variability in model performance among lakes based on its R^2 and tenfold cross-validated relative error. As an additional evaluation, we also performed an unconstrained cluster analysis using the same number of clusters and the same metric of dissimilarity as in the MRT analysis using the ‘kmeans’ function in R (Hartigan and Wong 1979). If the unconstrained cluster analysis yields a higher R^2 than the MRT analysis, then unobserved factors account for additional variation in model performance (De’Ath 2002). We also tested if certain models perform differently within each leaf relative

to grand means using a Kruskal–Wallis one-way analysis of variance (‘kruskal.test’ function in R). In case of significant overall differences ($p < 0.05$), we tested model-specific differences using pairwise Wilcoxon rank sum tests (‘pairwise.wilcox.test’ function in R). We chose nonparametric hypothesis tests to account for the relatively small sample size.

Parametrizing new wind-based k_{600} models

To explore new k_{600} parameterizations, we fitted a series of regression models to the dataset of lakes with at least six k_{600} observations. We fitted nonlinear mixed-effects models following the multilevel approach with cross-level interactions by Bryk and Raudenbush (1992). We used a 2-level model, where U_{10} was included to explain variation in k_{600} at the (within-lake) observation level and LA , Lat , SDI and SIN to explain variation in the U_{10} - k_{600} relationships at the (among-) lake level. We tested three functional forms commonly used in the literature (Table 1): linear ($k_{600} = a \cdot U_{10} + c$), exponential ($k_{600} = a \cdot \exp(U_{10} \cdot b)$) and power ($k_{600} = a \cdot U_{10}^b + c$). For each functional form we tested to what extent the slope (a) or shape (b) of the k_{600} - U_{10} relationship is a function of either LA or SDI , based on previous evidence on the potential lake-specific transfer from wind to turbulence (Schilder et al. 2013; Vachon and Prairie 2013; Gålfalk et al. 2013). We also tested whether the offset (c) is a function of either LA or Lat , based on previous evidence of lake-specific additional wind-independent turbulence by e.g. heat fluxes (MacIntyre et al. 2010; Read et al. 2012). The effect of SIN has never been tested, therefore we allowed it to have an influence on all parameters (a , b and c). We fitted models that cover all combinations of these predictor variables, amounting to 16 (linear, exponential) or 64 (power) candidate models. We did not test for TIN effects because it would not be useful as a predictive variable and require additional but unavailable data on the frequency distribution of U_{10} for each time integration interval (Wanninkhof et al. 1987; Livingstone and Imboden 1993). We evaluated the fits of all candidate models using the Akaike Information Criterion (AIC) and selected, for each of the three model types, the model with the lowest AIC and all parameters significant as the final model. For the final models, we report the RMSD, R^2_{adj} , and slope and intercept of linear regressions of observed vs. predicted k_{600} values and present 95% confidence intervals for mean predictions, representing a typical lake, following an approach by Bolker (2008). See Online Resources Text S2 for more details on the modelling procedure.

Results

Data set

The compiled data covered k_{600} values from 0.01 to 57.62 cm h⁻¹, U_{10} from 0 to 13 m s⁻¹, LA from 181 m² to 1342 km² and SDI from 1.0 to 22.5 (Fig. 2). k_{600} increased with U_{10} and was generally < 10 cm h⁻¹ for U_{10} < 2 m s⁻¹ and > 10 cm h⁻¹ for U_{10} > 8 m s⁻¹. U_{10} increased generally with LA , but decreased for very large LA (Fig. 2b), because these had very high SDI (Online Resources Fig. S2). k_{600} was not strongly related to SDI (Fig. 2c). TIN varied from

0.0024 to 122 days and SIN varied from 5 · 10⁻¹¹ to 1 (Online Resources Table S1).

General performance of wind-based k_{600} models

Table 2 summarizes the general performance of each model in predicting k_{600} for a specific lake. Model performance was similar for five of the six models with median RMSDs of 2.2–4.3 cm h⁻¹, R^2_{adj} values of -0.9 to -0.2, intercepts of -0.9 to 1.9 cm h⁻¹ and slopes of 0.7–1.7. Negative R^2_{adj} suggest that predictions were not better than using mean k_{600} observations. Only 2–39% of all cases showed positive R^2_{adj} depending on the model. Intercepts and slopes of linear regressions of observed vs. predicted

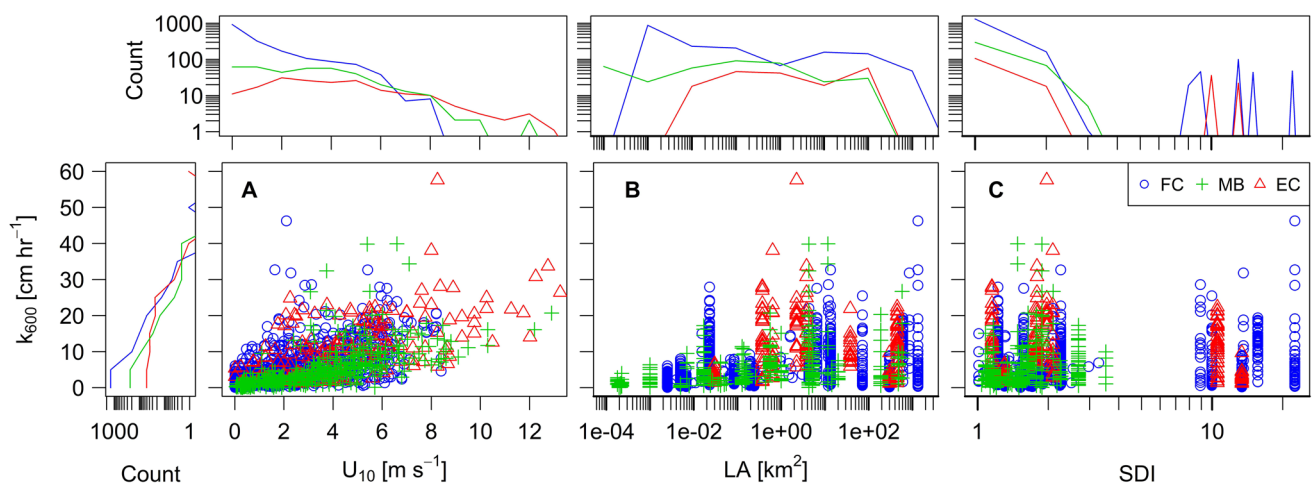


Fig. 2 Air–water gas transfer velocity (k_{600}) plotted against wind speed at 10 m height (U_{10}) (a), lake area (LA) (b), and shoreline development index (SDI) (c). Blue circles denote floating chamber measurements (FC), green crosses denote the mass balance approach

(MB), and red triangles denote the eddy covariance technique (EC). Histograms show data distributions across the plot axes (color figure online)

Table 2 Performance of empirical wind-based models for predictions of air–water gas transfer velocities (k_{600}) in 46 lakes with at least 6 observations per lake

Performance measure	CC98	VP13	CW03	M10	G07	L18
RMSD	2.2 (1.5, 7.1)	3.7 (2.1, 5.9)	3.8 (1.8, 8)	4.3 (2.2, 6.3)	3.1 (1.4, 7.5)	28.6 (12.9, 33.4)
R^2_{adj}	-0.2 (-0.9, 0.2)	-0.6 (-2.4, -0.1)	-0.9 (-2, 0.1)	-0.8 (-3.4, 0)	-0.2 (-1.3, 0.2)	-60.6 (-215.8, -29.8)
Intercept	-0.7 (-5, 2.1)	-0.9 (-7.1, 1)	1.9 (0.7, 5.5)	0.9 (-1, 2.6)	0.1 (-3.1, 2.8)	0.6 (-1.5, 1.9)
Slope	1.7 (0.9, 2.7)	1.1 (0.7, 2)	0.7 (0.3, 1.1)	0.7 (0.3, 1)	1.3 (0.5, 2.1)	0.2 (0.1, 0.2)
Prop($R^2_{adj} > 0$)	0.35	0.20	0.28	0.26	0.39	0.02
Prop(Intercept $\neq 0$)	0.41	0.30	0.61	0.48	0.46	0.33
Prop(Slope $\neq 1$)	0.46	0.41	0.52	0.48	0.50	0.96

Given are median (interquartile range) values of root mean square deviations (RMSD), coefficients of determination adjusted for the number of predictor variables (R^2_{adj}), and the intercept and slope of the linear regression of observed vs. predicted k_{600} . Given are also the proportion of positive R^2_{adj} (Prop($R^2_{adj} > 0$)), the proportion of intercepts significantly different from 0 (Prop(Intercept $\neq 0$)) and the proportion of slopes significantly different from 1 (Prop(Slope $\neq 1$))

CC98 Cole and Caraco (1998), VP13 Vachon and Prairie (2013), CW03 Crusius and Wanninkhof (2003), M10 is Macintyre et al. (2010), G07 Guerin et al. (2007), L18 Li et al. (2018)

values were significantly different from 0 and 1 in 30–61% and 41–52% of all cases, respectively. The model L18 by Li (2018) performed generally very poorly outside its training domain. Complete results for each model and lake are reported in Online Resources Table S2.

Predicting variability in model performance

The available experimental conditions (average U_{10} , LA , SDI , SIN , TIN , absolute value of Lat) were poor predictors of k_{600} model performance. This was indicated by the low explanatory power of the MRT analysis, and variable tree structures, depending on the threshold set for the minimum number of lakes included in each leaf. Depending on this threshold, two different trees were generated (Fig. 3, Online Resources Fig. S3). Cross-validated relative errors of these trees were between 1.14 and 1.17 and hence close to one, indicating poor prediction (De'Ath 2002). This was confirmed by the small proportion of variance in model performance ranks that was explained by the MRTs ($R^2 = 0.10$ – 0.11). The unconstrained cluster analyses explained a much higher proportion of variance ($R^2 = 0.32$ – 0.36), indicating that model performance ranks clustered relatively strongly and that k_{600} models performed differently under different conditions. However, the lower R^2 of MRT relative to the unconstrained cluster analysis suggests that observed conditions could

only partly account for the difference in model performance (De'Ath 2002).

Factors influencing wind-based k_{600} model performance

The two MRTs suggest that model performance was a function of TIN or LA , depending on the threshold number of lakes chosen. Apart from L18 always being the worst model, we can observe the following structures. For a threshold of 1–4 lakes per leaf, model performance was determined by LA (Online Resources Fig. S3). For very large systems ($LA \geq 375 \text{ km}^2$), the model VP13 by Vachon and Prairie (2013) performed slightly but not statistically significantly better than the other models. For systems smaller than 375 km^2 , models performed equally. For a threshold of minimum 5 lakes per leaf, model performance was primarily determined by TIN (Fig. 3). For $TIN < 105 \text{ min}$, VP13 performed slightly, but not statistically significantly better than the other models. For $TIN \geq 105 \text{ min}$, the model G07 by Guerin et al. (2007) performed significantly better than VP13.

Differences in model performance ranks were reflected by differences in the individual performance measures. For example, the higher ranks of G07 relative to VP13 for $TIN \geq 105 \text{ min}$ were due to lower RMSDs and higher R^2_{adj} (Online Resources Fig. S4). The generally poor performance

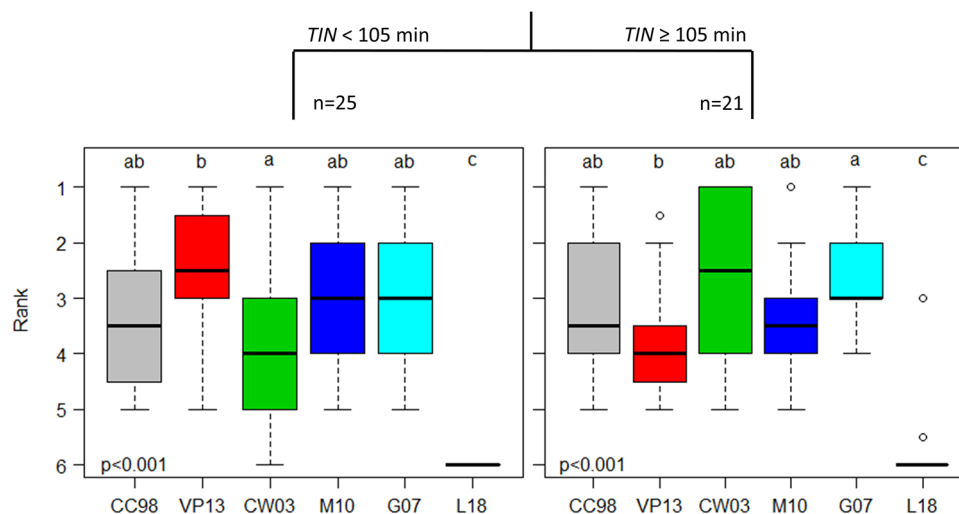


Fig. 3 Multivariate regression tree (MRT) to explain patterns in ranked model performance of air–water gas transfer velocity (k_{600}) predictions according to six published models. Ranks scale from 1 (best performance) to 6 (worst performance). The potential explanatory variables were average wind speed at 10 m height (U_{10}), lake area (LA), shoreline development index (SDI), the absolute value of latitude (Lat), space integration (SIN) and time integration (TIN). The MRT is valid for a minimum number of lakes of $n \geq 5$. The R^2 of the MRT was 0.10, and the cross-validated relative error was 1.17 ($SE = 0.12$). The boxplots show the distributions by medians (lines),

interquartile ranges (boxes) and 1.5 times the interquartile ranges (whiskers). Circles mark values outside these ranges. The p values indicate the significance of differences among models within a leaf (Kruskal–Wallis one-way analysis of variance). The letters a, b and c denote significant differences between boxplots within leaves (pairwise Wilcoxon rank sum test, $p < 0.05$). n is the number of lakes in each leaf. CC98 Cole and Caraco (1998), VP13 Vachon and Prairie (2013), CW03 Crusius and Wanninkhof (2003), M10 Macintyre et al. (2010), G07 Guerin et al. (2007), L18 Li (2018)

of L18 was reflected by extremely high RMSD, low R^2_{adj} and slopes almost always being significantly different from 1.

New k_{600} model parametrizations

Among all the tested linear, exponential and power models, we identified a suite of potential candidates for new k_{600} parameterizations in which all their parameters were statistically significant ($p < 0.05$), and which had similar statistical support (evidence ratio ≥ 0.05) as the respective models with the lowest AIC (Table 3). The candidate models all included lake- or method-specific characteristics in addition to U_{10} . Hence, including these variables significantly improved the

models relative to models with U_{10} alone. The slope of the linear model type was equivalently explained by LA , SDI or SIN , and the intercept of the power model type was equivalently explained by SIN , LA , or no predictor variable. These equivalences likely result, at least in part, from correlations between LA , SDI and SIN (Online Resources Fig. S2). The linear and power models with the lowest AIC showed a similar structure, where their slope component (a) increased with LA and their intercept component (c) decreased with SIN . We did not find any variable that significantly modulated the shape (b) component in the power model. The slope (a) and shape (b) component of the exponential model with the

Table 3 Performance characteristics of linear, exponential and power mixed-effects models to predict air–water gas transfer velocity (k_{600}) from wind speed at 10 m height (U_{10}), lake area (LA), shoreline development index (SDI) and/or space integration (SIN)

Model equation	df	logLik	AIC	Δ AIC	W (AIC)	Evidence ratio
Linear						
$(0.328 \pm 0.087 \cdot \log_{10}(LA) + 1.581 \pm 0.161) \cdot U_{10} - 0.066 \pm 0.028 \cdot \text{logit}(SIN) + 1.266 \pm 0.347$	2173	-4943.1	9906.1	0.00	0.30	0.50
$(0.338 \pm 0.087 \cdot \log_{10}(LA) + 1.504 \pm 0.159) \cdot U_{10} + 0.274 \pm 0.127 \cdot \log_{10}(LA) + 1.924 \pm 0.278$	2173	-4943.7	9907.5	1.37	0.15	0.34
$(-0.060 \pm 0.017 \cdot \text{logit}(SIN) + 1.085 \pm 0.212) \cdot U_{10} + 0.300 \pm 0.129 \cdot \log_{10}(LA) + 2.080 \pm 0.287$	2173	-4944.0	9908.1	1.97	0.11	0.27
$(1.435 \pm 0.423 \cdot \log_{10}(SDI) + 1.008 \pm 0.233) \cdot U_{10} - 0.066 \pm 0.027 \cdot \text{logit}(SIN) + 1.309 \pm 0.341$	2173	-4944.1	9908.2	2.05	0.11	0.26
$(1.586 \pm 0.425 \cdot \log_{10}(SDI) + 0.870 \pm 0.231) \cdot U_{10} + 0.315 \pm 0.125 \cdot \log_{10}(LA) + 2.007 \pm 0.273$	2173	-4944.1	9908.2	2.07	0.11	0.26
$(-0.055 \pm 0.017 \cdot \text{logit}(SIN) + 1.164 \pm 0.219) \cdot U_{10} - 0.056 \pm 0.027 \cdot \text{logit}(SIN) + 1.500 \pm 0.352$	2173	-4944.3	9908.6	2.51	0.08	0.22
$(0.339 \pm 0.088 \cdot \log_{10}(LA) + 1.509 \pm 0.160) \cdot U_{10} + 1.740 \pm 0.292$	2174	-4945.8	9909.7	3.56	0.05	0.14
$(-0.059 \pm 0.017 \cdot \text{logit}(SIN) + 1.067 \pm 0.215) \cdot U_{10} + 1.927 \pm 0.293$	2174	-4946.3	9910.7	4.56	0.03	0.09
$(1.442 \pm 0.427 \cdot \log_{10}(SDI) + 0.922 \pm 0.234) \cdot U_{10} + 1.802 \pm 0.288$	2174	-4946.9	9911.8	5.72	0.02	0.05
Exponential						
$(-0.057 \pm 0.022 \cdot \text{logit}(SIN) + 2.366 \pm 0.273) \cdot \exp(U_{10} \cdot (0.144 \pm 0.021 \cdot \log_{10}(SDI) + 0.156 \pm 0.011))$	2173	-5031.0	10,078.0	0.00	0.90	0.50
$2.667 \pm 0.258 \cdot \exp(U_{10} \cdot (0.151 \pm 0.021 \cdot \log_{10}(SDI) + 0.153 \pm 0.011))$	2174	-5034.2	10,082.4	4.41	0.10	0.10
Power						
$(0.281 \pm 0.085 \cdot \log_{10}(LA) + 1.361 \pm 0.200) \cdot U_{10}^{1.097 \pm 0.064} + (-0.072 \pm 0.029 \cdot \text{logit}(SIN) + 1.401 \pm 0.376)$	2172	-4946.6	9915.3	0.00	0.19	0.50
$(0.280 \pm 0.084 \cdot \log_{10}(LA) + 1.278 \pm 0.194) \cdot U_{10}^{1.107 \pm 0.065} + (0.326 \pm 0.137 \cdot \log_{10}(LA) + 2.114 \pm 0.316)$	2172	-4947.5	9916.9	1.69	0.08	0.30
$(0.299 \pm 0.088 \cdot \log_{10}(LA) + 1.343 \pm 0.201) \cdot U_{10}^{1.077 \pm 0.063} + 1.859 \pm 0.327$	2173	-4948.7	9917.4	2.17	0.07	0.25
$(1.226 \pm 0.374 \cdot \log_{10}(SDI) + 0.817 \pm 0.228) \cdot U_{10}^{1.127 \pm 0.063} + (-0.071 \pm 0.029 \cdot \text{logit}(SIN) + 1.492 \pm 0.372)$	2172	-4948.3	9918.6	3.38	0.04	0.16
$(1.321 \pm 0.365 \cdot \log_{10}(SDI) + 0.686 \pm 0.218) \cdot U_{10}^{1.145 \pm 0.064} + (0.380 \pm 0.134 \cdot \log_{10}(LA) + 2.252 \pm 0.308)$	2172	-4948.5	9919.0	3.79	0.03	0.13
$(-0.047 \pm 0.014 \cdot \text{logit}(SIN) + 0.917 \pm 0.215) \cdot U_{10}^{1.149 \pm 0.064} + (-0.062 \pm 0.029 \cdot \text{logit}(SIN) + 1.695 \pm 0.379)$	2172	-4949.3	9920.5	5.27	0.01	0.07

Given are the model equation (means \pm SD), estimated using the maximum likelihood algorithm, the degrees of freedom (df), log-likelihood values (logLik), Akaike Information Criterion (AIC), the difference in AIC (Δ AIC) between the current model and the model with the lowest AIC, Akaike weights (w(AIC)) and the evidence ratio, i.e. the probability that the current model is to be preferred over the model with the lowest AIC. Only models with evidence ratios ≥ 0.05 and all parameters significant ($p < 0.05$) are shown. Models are ranked from lowest to highest AIC. Note that the models including SDI are mechanistically not meaningful and should not be used for predictive purposes (for more details, see discussion section)

lowest AIC decreased with *SIN* and increased with *SDI*, respectively (Table 3).

For each functional shape, we exemplarily chose the model with the lowest AIC for further evaluation. Accordingly, the linear and power model predictions explained similar variability in k_{600} (65%) and had similar RMSD (3.35 and 3.34 cm h^{-1} , Online Resources Table S3). The linear regressions of observed vs. predicted k_{600} followed closely the 1:1 line, with intercepts and slopes not significantly different from 0 and 1, respectively (Fig. 4a,c). Accounting for lake-specific intercepts and slopes would not improve the fit of observed and predicted k_{600} (Likelihood ratio test, Online Resources Table S6), suggesting a good fit among all lakes. Model residuals were homogeneous with a mean near zero across the whole range of predicted k_{600} , U_{10} , LA , *SDI*, *SIN*, and *Lat* (Online Resources Fig. S5–S6). Finally, the linear and power model parameters were robust against the threshold minimum number of k_{600} observations per lake (Online Resources Fig. S8). In contrast, the exponential model with the lowest AIC explained only 35% of the variability in k_{600} with a RMSD of 4.89 cm h^{-1} (Online Resources Table S3). The linear regressions of observed and predicted k_{600} diverged from the 1:1 line, with intercept and slope significantly different from 0 and 1, respectively (Fig. 4b). Accounting for lake-specific intercepts and slopes would improve the fit of observed and predicted k_{600} (Likelihood ratio test, Online Resources Table S6), suggesting that the exponential model was significantly biased for some lakes. Model residuals were heterogeneous across the whole range of predicted k_{600} , U_{10} , LA , *SDI*, *SIN*, and *Lat* (Online Resources Fig. S7). The model parameter coefficients varied with the threshold minimum number of k_{600} observations per lake (Online Resources Fig. S8). Overall, these evaluations suggest that the linear and power models fitted the data

substantially better and with less bias than the exponential model.

The k_{600} predictions based on our new parameterizations were surprisingly similar for the linear and power models with the lowest AIC values, because of their similar model structure and the power exponent is close to 1 (Fig. 5). Predictions by the exponential model tended to be higher for relatively low ($< 1 \text{ m s}^{-1}$) and high ($> 7 \text{ m s}^{-1}$) U_{10} . These mixed-effect models integrate a large variability in lake-specific model parameterizations (see grey lines in Fig. 5). For example, U_{10} - k_{600} slopes (a) varied roughly between 0 and 5, and power exponents (b) varied from near 0 to up to 10 (Online Resources Fig. S10). Our mixed-effects model predictions fell largely within the range of predictions by published models (Fig. 5). For example, the intercept of our linear model was similar to M10 and CW03 for whole-lake space integrations ($SIN=1$) and similar to G07, CC98, and VP13 for the minimum space integration found in our dataset ($SIN=5 \cdot 10^{-11}$). The slopes were within the range of slopes in VP13.

The k_{600} predictions showed large uncertainties, as exemplarily shown in Fig. 6 for the linear model with the lowest AIC value. The lower and upper bounds of 95% confidence intervals were 10–250% smaller or larger than the mean predictions, respectively. This ratio was relatively small ($< 25\%$) for $U_{10} > 2 \text{ m s}^{-1}$ and $LA > 1 \text{ km}^2$ but increased drastically towards smaller LA where the density of available data was relatively scarce, and towards lower U_{10} . The increase in prediction uncertainties towards lower U_{10} was more pronounced for whole-lake integrations ($SIN=1$, Fig. 6a–c) relative to smaller space integrations of, for example, 1 m^2 ($SIN=1 \text{ m}^2/LA$, Fig. 6d–f).

Compared with previous k_{600} models, none of our new models performed better. Our linear, exponential and power

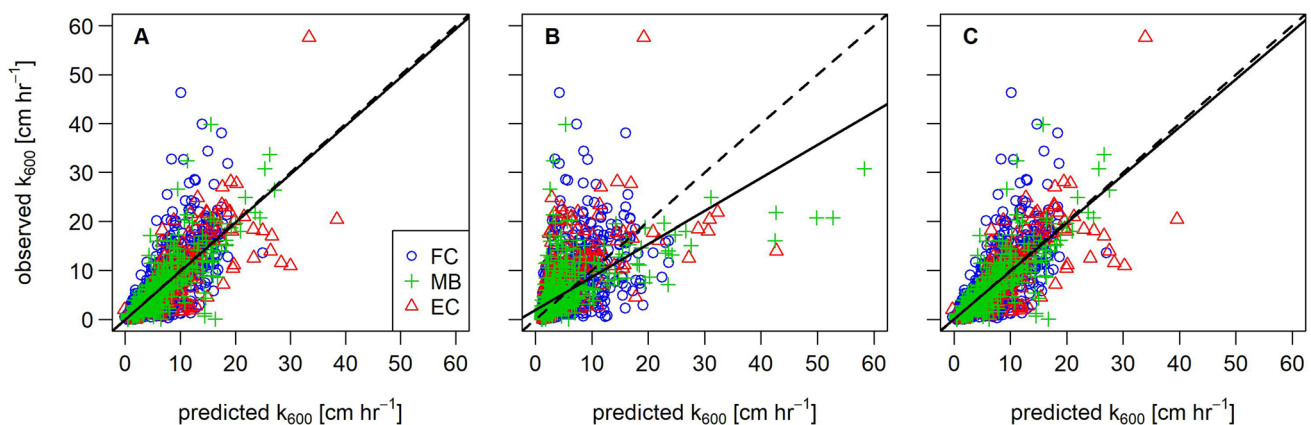


Fig. 4 Observed vs. predicted air–water gas transfer velocities (k_{600}) according to the linear (a), exponential (b) and power (c) model with the lowest Akaike Information Criterion values among all candidate models of the respective functional shape. Blue circles denote float-

ing chamber measurements (FC), green crosses denote the mass balance approach (MB), and red triangles denote the eddy covariance technique (EC). The solid line shows the linear regression line, the dashed line is the 1:1 line (color figure online)

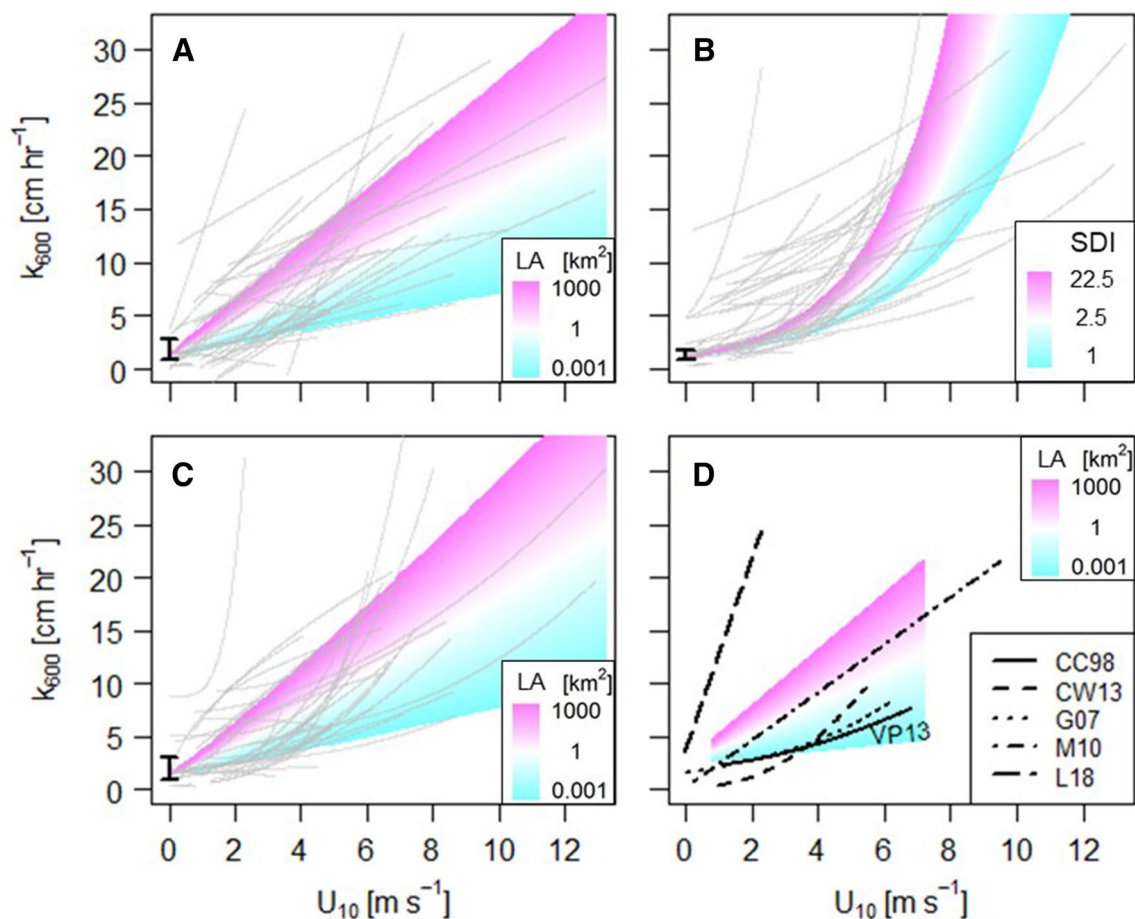


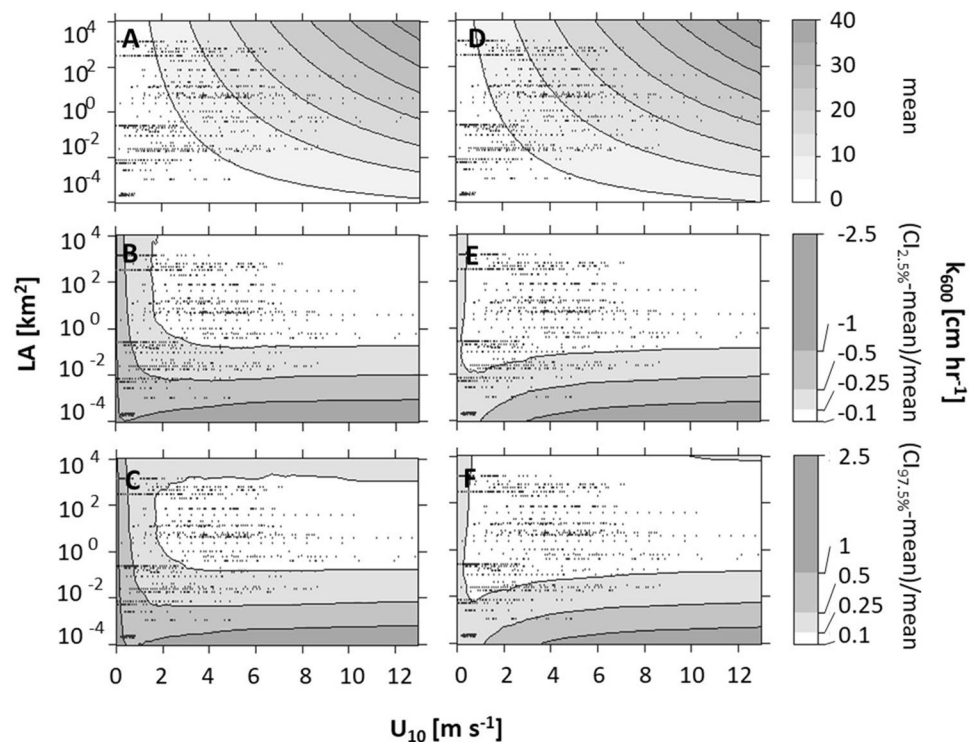
Fig. 5 Relationship between air–water gas transfer velocities (k_{600}) and wind speed at 10 m height (U_{10}) predicted by the linear (a), exponential (b) and power (c) model with the lowest Akaike Information Criterion values among all candidate models of the respective functional shape, and six published models (d). Colored shadings show predictions for different lake areas (LA) or shoreline development indices (SDI). Predictions in a–c are exemplarily shown for whole-lake space integrations ($SIN=1$). The error bar shows the variation in intercepts of predictions in this study, for a whole lake space integration ($SIN=1$, lower boundary), and the minimum space integration found in our dataset ($SIN=5 \cdot 10^{-11}$, upper boundary). Grey lines

show least-square predictions for each individual lake (in c three lakes are missing due to model fitting failure). The model equations are $k_{600} = (0.328 \cdot \log_{10}(LA) + 1.581) \cdot U_{10} - 0.066 \cdot \text{logit}(SIN) + 1.266$ (a), $k_{600} = (-0.057 \cdot \text{logit}(SIN) + 2.366) \cdot \exp(U_{10} \cdot (0.144 \cdot \log_{10}(SDI) + 0.156))$ (b), and $k_{600} = (0.281 \cdot \log_{10}(LA) + 1.361) \cdot U_{10}^{1.097} - 0.072 \cdot \text{logit}(SIN) + 1.401$ (c). Note that the model in b is not meaningful (for more details, see Discussion section). CC98 Cole and Caraco (1998), CW03 Crusius and Wanninkhof (2003), VP13 Vachon and Prairie (2013), M10 MacIntyre et al. (2010), G07 Guérin et al. (2007), L18 Li (2018) (color figure online)

models with the lowest AIC had median RMSDs of 3.3, 3.6 and 3.4, R^2_{adj} of -0.1 , -0.2 and -0.1 and intercepts and slopes of observed vs. predicted k_{600} of -0.3 , -0.7 and -0.3 , and 1.0 , 1.2 and 0.9 , respectively (c.f. Table 2). The proportion of positive R^2_{adj} was 0.41 , 0.39 and 0.37 , the proportion of intercepts significantly different from 0 was 0.28 , 0.46 and 0.28 , and the proportion of slopes significantly different from

1 was 0.41 , 0.46 and 0.41 , respectively. Overall, the integrated performance indices of our new models were not significantly different from most previous models, except for L18 (pairwise Wilcoxon rank sum test, Online Resources Fig. S9).

Fig. 6 Predicted air–water gas transfer velocities (k_{600}) as a function of wind speed at 10 m height (U_{10}), and lake area (LA) according to the linear model with the lowest Akaiques Information Criterion value. Shown are mean predictions (a, d) and the 95% confidence interval (CI), i.e. the 2.5% bound (b, e) and 97.5% bound (c, f) expressed relative to mean predictions for whole-lake space integration ($SIN=1$) (a–c) and $SIN=10^{-6}$ km²/LA (d–f). Grey dots show observations. The CIs take uncertainties from fixed effects into account



Discussion

How well can wind speed predict k_{600} over global lakes?

Previously published models showed that U_{10} alone can explain a high share of variance in measured k_{600} when applied within their calibration domain ($R^2=0.68$ – 0.93 ; Table 1), suggesting that U_{10} can be a robust predictor under specific conditions. However, several studies have also failed to identify U_{10} as a predictor of k_{600} in specific lakes (Cole and Caraco 1998; Matthews et al. 2003; Xiao et al. 2014; Podgrajsek et al. 2015; Holgerson et al. 2017), and the number of unpublished unsuccessful attempts is unknown. Our global data synthesis allowed us to evaluate U_{10} as a predictor of k_{600} over a wide range of lakes. This dataset, i.e. 2222 simultaneous k_{600} and U_{10} measurements from 46 lakes and reservoirs, is by far more extensive than the database of previously published wind-based models (Table 1). We show that applying these wind-based models on new lakes and under new conditions results in poor and arbitrary k_{600} predictability, suggesting that U_{10} parameterizations derived from one or few systems cannot always be used to predict k_{600} beyond the training data sets. The R^2_{adj} of observed vs. predicted k_{600} was on average negative, and only in a minority of cases, U_{10} -based predictions were more accurate than simply assuming a mean k_{600} independent of U_{10} . (Table 2). These results signify

that there is no wind-based model that predicts k_{600} well in all types of lakes.

Among all lakes, none of the tested published models clearly performed better than the others, in line with our first hypothesis (H1). However, one model (L18) performed worse than all other models likely because it was developed in a reservoir with significant lateral water flow as an additional source of turbulence (Li 2018). Some models seemed to perform slightly better than others under specific conditions such as very large LA or short TIN (Fig. 3 and Online Resources Fig. S3). However, these differences were small (typically <2 performance index ranks) and not statistically significant. Overall, we did not clearly identify conditions under which certain models performed better than others, and therefore reject our second hypothesis (H2). This result is further supported by the very low explanatory power of the MRTs ($R^2=0.10$ – 0.11) relative to the unconstrained cluster analysis ($R^2=0.32$ – 0.36), implying that there was little structure in model performance and that this structure was mainly determined by unobserved factors, or by secondary factors that were observed but remained unrevealed due to a lack of statistical power.

The generally poor and nearly random performance of published k_{600} models suggests that k_{600} predictions are associated with large errors, especially when models are used to extrapolate to new conditions. This finding emphasizes our second research question whether a more general model fitted to a wider range of training data, and including additional easy to obtain predictor variables could improve

k_{600} predictions or at least better account for prediction uncertainties.

Functional shape of wind-based k_{600} parametrizations

To properly assess the new k_{600} parameterizations and the effects of additional predictor variables, we first have to evaluate the shape of the U_{10} - k_{600} relationship. The U_{10} - k_{600} relationship can have many different shapes among lakes (Fig. 5), and this can be due to several factors. For example, non-linear U_{10} - k_{600} relationships can be a result of k_{600} being enhanced at low U_{10} due to convection (MacIntyre et al. 2010; Polsenaere et al. 2013) or chemical enhancement of reactive gases (Wanninkhof et al. 1987). Our models were not flexible enough to accommodate such variability, which is reflected in high uncertainties at low U_{10} (Fig. 6). k_{600} can also be enhanced by bubbles formed by breaking waves, leading to an accelerating increase at $U_{10} > 12 \text{ m s}^{-1}$ (Broecker and Siems 1984). This process was negligible in our dataset because most observations were done below the U_{10} threshold for breaking waves. Some studies hypothesized the presence of microbubbles affecting k_{600} measurements from less soluble gases (e.g. CH_4 ; Prairie and Giorgio 2013; McGinnis et al. 2015). While the exact drivers of this phenomenon are still unknown, it could indeed affect the lake-specific U_{10} - k_{600} shape. We accounted for potential variability among lakes in the shape of the U_{10} - k_{600} relationships by the exponent b in our power model. Interestingly, b varied widely among lakes but none of the tested lake- or method-related characteristics could significantly explain this high variability. This highlights the need for future studies to identify other factors that could explain high between-lake variability in U_{10} - k_{600} shapes. Our best estimate for the exponent b in the power model was near 1 (Table 3). Alternative parameterizations that assume an exponential shape resulted in poor model fits with strong biases. Therefore, we conclude that the U_{10} - k_{600} relationship of a typical lake in our database is linear and recommend the use of our linear model parameterizations for applications across a wide range of lakes. We provide code implemented in the program R to estimate k_{600} and associated uncertainties as a function of U_{10} , LA , and SIN based on our linear model with the lowest AIC value (Online Resources Code S1, Data S1, and Data S2). We emphasize that this model is one of several other linear models that fitted the data equally well.

Modulators of wind-speed effect on k_{600}

Our new multilevel modelling procedure showed that LA , SDI and/or SIN can significantly explain between-lake differences in the U_{10} - k_{600} —relationships, which supports our third hypothesis (H3). The U_{10} effect on k_{600} increased with

LA likely as a result of progressively larger fetch length, wave build-up and hence turbulence (Schilder et al. 2013; Vachon and Prairie 2013). The modulating effect of LA in our dataset was less strong than the effect found by Vachon and Prairie (2013) and Guérin et al. (2007) (Online Resources Fig. S11, but note absence of effects in Wanninkhof et al. (1987)). This variability in the effect of LA on U_{10} can arise from differences in shoreline sheltering of the lakes considered. Relatively many small lakes included in our dataset (Sebacher et al. 1983; Leuning et al. 1984; Boyd and Teichert-Coddington 1992; Denmead and Freney 1992) had only little shoreline sheltering which allowed k_{600} to respond relatively strongly to U_{10} . As a result, the LA effect on U_{10} - k_{600} relationships was rather weak (Online Resources Fig. S11). In contrast, the very large LA effects reported by Guérin et al. (2007) may be due to the long effective fetch length in the elongated estuaries included in this study, allowing larger wave heights for a given U_{10} . The larger LA effects reported by Vachon and Prairie (2013) and Guérin et al. (2007) could also be a result of estimating k_{600} locally in the lake center, for which fetch length may matter more than for whole-lake integrated data. Hence, the effect of LA on U_{10} - k_{600} relationships is far from universal and may need further investigations. Our dataset covered both sheltered and unsheltered systems, and also a wide range of SDI . We argue that all these effects were partly captured by the addition of LA as a modulator of the U_{10} effect, making our model less specific and more generic than previous models.

Our modelling procedure quantified, for the first time, the effect of SIN on k_{600} estimates. This scale is strongly related to the method of estimating k_{600} . Predicted k_{600} was around 2.4 cm h^{-1} higher for small-scale integrations ($10^{-6} \text{ km}^2/LA$) in the lake center (typical for floating chambers) relative to integrations over the whole lake (typical for mass balances). With the mass balance approach, k_{600} are integrated over the whole lake surface while floating chambers and the eddy covariance technique usually integrate smaller areas near the lake center (Fig. 1c). Here, U_{10} and hence k_{600} is usually higher than near-shore (Venäläinen et al. 2003; Schilder et al. 2013; Vachon and Prairie 2013; Docquier et al. 2016). Therefore smaller SIN resulted in higher k_{600} in our model. It is important to note, however, that our models predict local k_{600} only for areas around the lake center and that higher or lower k_{600} should be expected in near-shore areas, depending on the wind direction and lake shape (Vachon and Prairie 2013). The SIN effect should raise the awareness among k_{600} model users about which proportion of the lake the anticipated k_{600} values should integrate (Fig. 1c). This finding has implications for calculations of local vs. whole-lake gas fluxes where the scale of concentration and k_{600} estimates should match. For example, flux calculations should be based on $SIN \ll 1$ if gas concentrations are measured in the center of the lake, and on $SIN = 1$, if gas concentrations are

measured at (multiple) points that are representative for the whole lake surface.

Some of the selected best candidate models included *SDI*, suggesting the U_{10} effect on k_{600} to increase in lakes with more complex shoreline geometry. This finding is rather counter-intuitive as one would think that k_{600} should decrease with *SDI*, given the increased shoreline sheltering simulating the effect of a small lake (Schilder et al. 2013; Vachon and Prairie 2013; Gålfalk et al. 2013). The positive effect of *SDI* likely resulted from *SDI* being highly correlated with *LA* (Online Resources Fig. S2A) and *LA* having a positive effect on k_{600} . We, therefore, conclude that sheltering due to complex shorelines is not a dominant modulator of the U_{10} - k_{600} relationship and do not recommend using our models that include *SDI*.

Performance and applicability of new parametrization on global lakes

Accounting for the effects of *LA* or *SIN* on lake-level U_{10} - k_{600} relationships improved model fits relative to single-level models based on U_{10} alone. However, despite this improvement, even our best k_{600} models did not clearly outperform the previously published k_{600} models most of which only included U_{10} (Online Resources Fig. S9), which provides support to falsify our fourth hypothesis (H4). However, even if including additional variables in addition to U_{10} did not improve the statistical model fit, it may contribute to a more accurate geographic explanation of variations in k_{600} among a wide range of lakes.

Our best linear model is designed to fit average conditions across a wide range of global lakes and to account for their variability by estimating the error. We regard our model to be the globally least biased k_{600} model available, as it averages out method and system-specific errors across a wide range of conditions and explicitly predicts these errors. This approach fills an important gap, relative to previous k_{600} models which were developed under limited conditions (Table 1). Our new model also greatly expands previous limits of the calibration data sets in terms of U_{10} and *LA* and should include most conditions encountered in the world ($U_{10}=0\text{--}13\text{ m s}^{-1}$, $LA=183\text{ m}^2$ to 1342 km^2 , $SDI=1\text{--}22.5$; (c.f. Verpoorter et al. 2014)). Hence, our model yields predictions to represent, to the best possible, an average global lake, and error predictions to account for potential spatiotemporal variability. With these characteristics, our new model is suitable for large-scale applications such as upscaling gas fluxes to regional and global scales.

Within our prediction domain, prediction errors were large (up to 10 cm h^{-1} or $>200\%$ of mean predictions) and must be accounted for in large-scale or global gas flux estimates. Interestingly, prediction errors varied non-linearly with U_{10} and *LA*. While k_{600} was rather well constrained

under conditions that have previously been relatively commonly sampled (intermediate U_{10} and *LA*), large uncertainties still exist in relative terms in small lakes ($LA < 1\text{ km}^2$) and for low wind speeds ($U_{10} < 2\text{ m s}^{-1}$), conditions that are common globally (Verpoorter et al. 2014, <https://globalwindatlas.info/>). Under these conditions, k_{600} can be dominantly driven by many other factors in addition to U_{10} (Crusius and Wanninkhof 2003; Read et al. 2012; Holgerson et al. 2017). Future data collections should focus on low wind speeds and small lakes, to identify underlying drivers of high variability in k_{600} and by that reduce prediction uncertainties.

Limitations and way forward for better wind-based k_{600} models

Several factors that are known to potentially influence k_{600} were not accounted for in our new k_{600} models. Such factors include surface films, turbulence due to convective cooling, bubble-mediated gas transfer, gas-specific behavior (e.g. chemical enhancement), boundary layer stability, stratification, variability in wind stress and the wave field for given U_{10} levels due to shoreline sheltering, or method-specific issues beyond spatial and temporal integrations (Wanninkhof et al. 1987; MacIntyre et al. 1995; Jähne and Haussecker 1998). To capture these drivers is often difficult and labor-intensive, hence, relevant data are only available for a limited number of systems. Accounting for environmental factors beyond wind may improve k_{600} predictions in specific lakes (MacIntyre et al. 2010; Polsenaere et al. 2013; Heiskanen et al. 2014). Here, a way forward is to develop new models based on mechanistic first principles, relating environmental factors to turbulent kinetic energy dissipation as the primary driver of k_{600} (Zappa et al. 2007; Vachon et al. 2010). However, it remains to be assessed whether these factors would matter and to what extent mechanistic models that account for these factors could improve k_{600} predictions at larger spatial scales. If turning out important, simple proxies of otherwise difficult to measure processes would need to be found to develop k_{600} models that are applicable over many lakes.

Our analysis indicates that our ability to predict k_{600} based on empirical wind-based models could approach an upper limit that is not necessarily determined by a lack of understanding of the controls of k_{600} , but by a lack of methodological consistency. First, models with additional predictor variables or relatively many lakes included (e.g. CC98, VP13), did not perform significantly better than single lake/single variable models (e.g. G07, M10). Second, there was also only little structure in the model's performance relative to the experimental or environmental conditions under which the data were collected (Fig. 3 and Online Resources Fig. S3). These observations could be either explained by a true lack of structure or that this structure is masked by high

noise in the collected k_{600} data due to measurement errors or inconsistencies in methodologies.

One important methodological inconsistency with consequences for the predictability of k_{600} is the way U_{10} is measured. In our data set, U_{10} was mainly measured on or within 3 km of the lake surface (76% or 89% of observations from all lakes and 42% or 85% of observations from lakes with at least 6 observations, Online Resources Table S1), but some were measured even farther inland. Measurements were also scaled to 10 m height from different measurement heights. Wind speed point measurements are not always representative of the whole lake mean U_{10} (Fig. 1c; Venäläinen et al. 2003; Docquier et al. 2016). Wind height scaling is also associated with uncertainties (Large and Pond 1981). To reduce these uncertainties, the spatial resolution of wind speed measurements should be increased to match the scale and extent of k_{600} estimates [e.g. several anemometers over the lake, (Wanninkhof et al. 1987)].

Many other methodological and environmental issues that could lead to noise in U_{10} - k_{600} relationships have been discussed in the literature. In essence, every approach addresses air–water gas transfer from a different angle, with characteristic scales and with method-specific advantages and disadvantages (MacIntyre et al. 1995; Jähne and Haussecker 1998). Our new k_{600} models account for most of these uncertainties as they integrate the variability in previous studies carried out under widely different conditions. However, despite the wide variety of studies included, more efforts are needed to measure k_{600} in lakes types that go beyond our dataset and to assess how representative our data collection is for global conditions of lake-atmosphere gas transfer.

Implications and conclusions

With the currently available set of k_{600} models with their limited calibration domain, researchers have been in need to extrapolate k_{600} to their system of interest without being able to properly quantify and account for resulting uncertainties. This practice would strongly limit their ability to gain insights into ecological and biogeochemical processes in specific systems (Dugan et al. 2016; Kiuru et al. 2019) and to upscale the lakes' air–water gas fluxes to the globe (Raymond et al. 2013). Based on an evaluation of existing wind-based k_{600} models against an extensive set of published U_{10} and k_{600} estimates, we conclude that extrapolation can lead to significant biases in k_{600} predictions.

Building on the growing awareness that “the gas transfer velocity is not simply a function of the wind speed” (Jähne and Haußecker 1998), we found here that U_{10} is generally a poor universal predictor of k_{600} over lakes or reservoirs, no matter which of the existing model parameterizations are applied. Prediction errors remain high even in new

parameterizations calibrated against the global data set. Therefore, we conclude (in agreement with Cole et al. 2010) that wind-based models are currently very limited in their use to scale k_{600} across lakes and advise better measurements rather than models of k_{600} when accurate estimations for specific lakes are needed. Similar challenges of scaling k_{600} across systems remain in streams and rivers (Hall and Ulseth 2019).

For the future development of lake models (c.f. Tan and Zhuang 2015; Stepanenko et al. 2016) or larger scale applications (c.f. Zwart et al. 2018), we emphasize accounting for large uncertainties when modelling k_{600} . To do so, we propose a new k_{600} model that provides estimates of means and uncertainties in k_{600} as a function of U_{10} , LA , and SIN . Large uncertainties in k_{600} models may be overcome by developing mechanistic models from first principles. Until this is achieved, the best approach remains to calibrate empirical constants against extensive data sets. Progress on these lines will improve the development of coupled atmosphere–land surface–lake models and incorporation of lakes in earth system models (MacKay et al. 2009).

The results from this study emphasize careful thought about the strength and limitations, in particular the calibration domain, and spatial scale of integration of the anticipated means to estimate k_{600} , no matter if k_{600} is estimated empirically or modelled. By an informed choice of the most suitable k_{600} model, researchers should be able to limit uncertainties in k_{600} predictions within acceptable boundaries, or as George Box phrased it: “Essentially, all models are wrong, but some are useful”.

Acknowledgements Open access funding provided by Umea University. This paper is based on research supported by the Knut and Alice Wallenberg Foundation (dnr: 2016.0083). Dominic Vachon was also supported by postdoctoral fellowships from the Natural Sciences and Engineering Research Council of Canada. We thank Anders Jonsson, Werner Eugster, Erik Sahlee, Eva Podgrajsek, and Kenneth Thoro-Martinsen for providing original published data. We thank Rik Wanninkhof, Peter Liss and one anonymous reviewer for constructive comments on a previous version of this manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anderson DE, Striegl RG, Stannard DI et al (1999) Estimating lake-atmosphere CO₂ exchange. *Limnol Oceanogr* 44:988–1001. <https://doi.org/10.4319/lo.1999.44.4.0988>
- Bidleman TF (1999) Atmospheric transport and air-surface exchange of pesticides. In: Van Dijk HFG, Van Pul WAJ, De Voogt P (eds) Fate of pesticides in the atmosphere: implications for environmental risk assessment. Springer, Dordrecht
- Bolker B (2008) Ecological models and data in R. Princeton University Press, Princeton
- Boyd CE, Teichert-Coddington D (1992) Relationship between wind speed and reaeration in small aquaculture ponds. *Aquac Eng* 11:121–131. [https://doi.org/10.1016/0144-8609\(92\)90014-O](https://doi.org/10.1016/0144-8609(92)90014-O)
- Broecker HC, Siems W (1984) The role of bubbles for gas transfer from water to air at higher windspeeds. Experiments in the wind-wave facility in Hamburg. In: Brutsaert W, Jirka GH (eds) Gas transfer at water surfaces. Springer, New York, pp 229–236
- Bryk AS, Raudenbush SW (1992) Hierarchical linear models: applications and data analysis methods. Sage Publications, New York
- Clark JF, Schlosser P, Wanninkhof R, Simpson HJ, Schuster WSF, Ho DT (1995) Gas transfer velocities for SF₆ and ³He in a small pond at low wind speeds. *Geophys Res Lett* 22(2):93–96. <https://doi.org/10.1029/94GL02410>
- Cole J, Bade D, Bastviken D (2010) Multiple approaches to estimating air-water gas exchange in small lakes. *Limnol Oceanogr Methods*. <https://doi.org/10.4319/lom.2010.8.285>
- Cole JJ, Caraco NF (1998) Atmospheric exchange of carbon dioxide in a low-wind oligotrophic lake measured by the addition of SF₆. *Limnol Oceanogr* 43:647–656. <https://doi.org/10.4319/lo.1998.43.4.0647>
- Crusius J, Wanninkhof R (2003) Gas transfer velocities measured at low wind speed over a lake. *Limnol Oceanogr* 48:1010–1017. <https://doi.org/10.4319/lo.2003.48.3.1010>
- De'Ath G (2002) Multivariate regression trees : a new technique for modelling species-environment relationships. *Ecology* 83:1105–1117. [https://doi.org/10.1890/0012-9658\(2002\)083\[1105:MRTANT\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[1105:MRTANT]2.0.CO;2)
- De'Ath G (2014) mvpart 1.6–2 Multivariate regression trees. <https://mran.microsoft.com/snapshot/2014-12-11/web/packages/mvpart/index.html>. Accessed 06 Apr 2020).
- Denmead OT, Freney JR (1992) Transfer coefficients for water-air exchange of ammonia, carbon dioxide and methane. *Ecol Bull* 42:31–41
- Docquier D, Thiery W, Lhermitte S, van Lipzig N (2016) Multi-year wind dynamics around Lake Tanganyika. *Clim Dyn* 47:3191–3202. <https://doi.org/10.1007/s00382-016-3020-z>
- Dormann CF, Calabrese JM, Guillera-Arroita G et al (2018) Model averaging in ecology: a review of Bayesian, information-theoretic, and tactical approaches for predictive inference. *Ecol Monogr* 88:485–504. <https://doi.org/10.1002/ecm.1309>
- Dugan HA, Iestyn Woolway R, Santoso AB et al (2016) Consequences of gas flux model choice on the interpretation of metabolic balance across 15 lakes. *Inl Waters* 6:581–592. <https://doi.org/10.5268/IW-6.4.836>
- Emerson S (1975) Gas exchange rates in small Canadian Shield lakes. *Limnol Oceanogr* 20:754–761. <https://doi.org/10.4319/lo.1975.20.5.0754>
- Emerson S, Broecker W, Schindler DW (1973) Gas-exchange rates in a small lake as determined by radon method. *J Fish Res Board Canada* 30:1475–1484. <https://doi.org/10.1139/f73-237>
- Engle D, Melack JM (2000) Methane emissions from an Amazon floodplain lake: enhanced release during episodic mixing and during falling water. *Biogeochemistry* 51:71–90. <https://doi.org/10.1023/A:1006389124823>
- Erkkilä K-M, Ojala A, Bastviken D et al (2018) Methane and carbon dioxide fluxes over a lake: comparison between eddy covariance, floating chambers and boundary layer method. *Biogeosciences* 15:429–445. <https://doi.org/10.5194/bg-15-429-2018>
- Fink P (2007) Ecological functions of volatile organic compounds in aquatic systems. *Mar Freshw Behav Physiol* 40:155–168. <https://doi.org/10.1080/10236240701602218>
- Frost T, Upstill-Goddard RC (2002) Meteorological controls of gas exchange at a small English lake. *Limnol Oceanogr* 47:1165–1174. <https://doi.org/10.4319/lo.2002.47.4.1165>
- Gålfalk M, Bastviken D, Fredriksson S, Arneborg L (2013) Determination of the piston velocity for water-air interfaces using flux chambers, acoustic Doppler velocimetry, and IR imaging of the water surface. *J Geophys Res Biogeosci* 118:770–782. <https://doi.org/10.1002/jgrg.20064>
- Gelda RK, Auer MT, Effler SW et al (1996) Determination of reaeration coefficients: whole-lake approach. *J Environ Eng*. [https://doi.org/10.1061/\(ASCE\)0733-9372\(1996\)122](https://doi.org/10.1061/(ASCE)0733-9372(1996)122)
- Guérin F, Abril G, Serça D et al (2007) Gas transfer velocities of CO₂ and CH₄ in a tropical reservoir and its river downstream. *J Mar Syst* 66:161–172. <https://doi.org/10.1016/j.jmarsys.2006.03.019>
- Hall RO, Ulseth AJ (2019) Gas exchange in streams and Rivers. *WIREs Water* 1391:1–18. <https://doi.org/10.1002/wat2.1391>
- Hartigan JA, Wong MA (1979) Algorithm AS 136: a K-means clustering algorithm. *J R Stat Soc Ser C* 28:100–108. [https://doi.org/10.2307/2346830\(Applied Stat\)](https://doi.org/10.2307/2346830(Applied Stat))
- Hastie A, Lauerwald R, Weyhenmeyer G et al (2017) CO₂ evasion from boreal lakes: revised estimate, drivers of spatial variability, and future projections. *Glob Chang Biol* 2:711–728. <https://doi.org/10.1111/gcb.13902>
- Heiskanen JJ, Mammarella I, Haapanala S et al (2014) Effects of cooling and internal wave motions on gas transfer coefficients in a boreal lake. *Tellus, Ser B Chem Phys Meteorol* 66:1–16. <https://doi.org/10.3402/tellusb.v66.22827>
- Hesslein RH, Broecker WS, Quay PD, Schindler DW (1980) Whole-Lake radiocarbon experiment in an oligotrophic lake at the experimental Lakes Area, Northwestern Ontario. *Can J Fish Aquat Sci* 37:454–463. <https://doi.org/10.1139/f80-059>
- Ho DT, Engel VC, Ferrón S et al (2018) On factors influencing air-water gas exchange in emergent Wetlands. *J Geophys Res Biogeosciences* 123:178–192. <https://doi.org/10.1002/2017JG004299>
- Holgerson MA, Farr ER, Raymond PA (2017) Gas transfer velocities in small forested ponds. *J Geophys Res Biogeosci* 122:1011–1021. <https://doi.org/10.1002/2016JG003734>
- Hornbuckle KC, Jeremiason JD, Sweet CW, Elsenreich SJ (1994) Seasonal variations in air-water exchange of polychlorinated biphenyls in Lake Superior. *Environ Sci Technol* 28:1491–1501. <https://doi.org/10.1021/es00057a018>
- Howard EM, Forbrich I, Giblin AE et al (2018) Using noble gases to compare parameterizations of air–water gas exchange and to constrain oxygen losses by ebullition in a shallow aquatic environment. *J Geophys Res Biogeosci* 123:2711–2726. <https://doi.org/10.1029/2018jg004441>
- Hutchinson GE (1957) A treatise on limnology geography, physics, and chemistry. Part. 1. Geography and physics of lakes, vol 1. Wiley, New York
- Jähne B, Fischer KH, Imberger J, et al (1984) Parametrization of air/lake gas exchange. In: Brutsaert W, Jirka GH (eds) Gas Transfer at Water Surfaces. D. Reidel Publishing Company, pp 459–466
- Jähne B, Haussecker H (1998) Air-water gas exchange. *Annu Rev Fluid Mech* 30:443–468. <https://doi.org/10.1146/annurev.fluid.30.1.443>
- Jähne BJ, Münnich KOM, Börsinger R et al (1987) On the parameters influencing air-water gas exchange. *J Geophys Res* 92:1937–1949. <https://doi.org/10.1029/JC092iC02p01937>

- Jean-Baptiste P, Poisson A (2000) Gas transfer experiment on a lake (Kerguelen Islands) using ^3He and SF_6 . *J Geophys Res* 105:1177–1186. <https://doi.org/10.1029/1999JC900088>
- Jonsson A, Åberg J, Lindroth A, Jansson M (2008) Gas transfer rate and CO_2 flux between an unproductive lake and the atmosphere in northern Sweden. *J Geophys Res* 113:G04006. <https://doi.org/10.1029/2008JG000688>
- Kiuru P, Ojala A, Mammarella I et al (2019) Applicability and consequences of the integration of alternative models for CO_2 transfer velocity into a process-based lake model. *Biogeosci Discuss* 16:3297–3317. <https://doi.org/10.5194/bg-2019-95>
- Kuss J, Holzmann J, Ludwig R (2009) An elemental mercury diffusion coefficient for natural waters determined by molecular dynamics simulation. *Environ Sci Technol* 43:3183–3186. <https://doi.org/10.1021/es8034889>
- Kwan J, Taylor PA (1994) On gas fluxes from small lakes and ponds. *Boundary-Layer Meteorol* 68:339–356
- Large WG, Pond S (1981) Open ocean momentum flux measurements in moderate to strong winds. *J Phys Oceanogr* 11:324–336. [https://doi.org/10.1175/1520-0485\(1981\)011%3c0324:OOMFM1%3e2.0.CO;2](https://doi.org/10.1175/1520-0485(1981)011%3c0324:OOMFM1%3e2.0.CO;2)
- Leuning R, Denmead O, Simpson J, Freney J (1984) Processes of ammonia loss from shallow floodwater. *Atmos Environ* 18:1583–1592. [https://doi.org/10.1016/0004-6981\(84\)90380-9](https://doi.org/10.1016/0004-6981(84)90380-9)
- Li S (2018) CO_2 oversaturation and degassing using chambers and a new gas transfer velocity model from the Three Gorges Reservoir surface. *Sci Total Environ* 640–641:908–920. <https://doi.org/10.1016/j.scitotenv.2018.05.345>
- Likens GE (ed) (2010) *Biogeochemistry of Inland Waters*, 1st Edition. Academic Press, Cambridge
- Livingstone DM, Imboden DM (1993) The non-linear influence of wind-speed variability on gas transfer in lakes. *Tellus B* 45:275–295. <https://doi.org/10.1034/j.1600-0889.1993.t01-2-00005.x>
- López Bellido J, Tulonen T, Kankaala P, Ojala A (2009) CO_2 and CH_4 fluxes during spring and autumn mixing periods in a boreal lake (Pääjärvi, southern Finland). *J Geophys Res Biogeosci* 114:1–12. <https://doi.org/10.1029/2009JG000923>
- MacIntyre S, Jonsson A, Jansson M et al (2010) Buoyancy flux, turbulence, and the gas transfer coefficient in a stratified lake. *Geophys Res Lett* 37:L24604. <https://doi.org/10.1029/2010GL014164>
- MacIntyre S, Wanninkhof R, Chanton JP (1995) Trace gas exchange across the air-water interface in freshwater and coastal marine environments. In: Matson PA, Harriss RC (eds) *Biogenic trace gases: measuring emissions from soil and water*. Wiley, New York, pp 52–97
- MacKay MD, Neale PJ, Arp CD et al (2009) Modeling lakes and reservoirs in the climate system. *Limnol Oceanogr* 54:2315–2329. https://doi.org/10.4319/lo.2009.54.6_part_2.2315
- Markfort CD, Perez ALS, Thill JW et al (2010) Wind sheltering of a lake by a tree canopy or bluff topography. *Water Resour Res*. <https://doi.org/10.1029/2009WR007759>
- Martinsen KT, Kragh T, Sand-Jensen K (2020) Carbon dioxide efflux and ecosystem metabolism of small forest lakes. *Aquat Sci* 82:1–17. <https://doi.org/10.1007/s00027-019-0682-8>
- Matthews CJD, St Louis VL, Hesslein RH (2003) Comparison of three techniques used to measure diffusive gas exchange from sheltered aquatic surfaces. *Environ Sci Technol* 37:772–780. <https://doi.org/10.1021/es0205838>
- McGinnis DF, Kirillin G, Tang KW et al (2015) Enhancing surface methane fluxes from an oligotrophic lake: exploring the microbubble hypothesis. *Environ Sci Technol* 49:873–880. <https://doi.org/10.1021/es503385d>
- Messenger ML, Lehner B, Grill G et al (2016) Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nat Commun* 7:13603. <https://doi.org/10.1038/ncomms13603>
- Nisell J, Lindsjö A, Temmerud J (2007) Rikstäckande virtuellt vattendrags nätverk för flödesbaserad modellering ViVaN (in Swedish with an English summary). Department of Aquatic Science and Assessment. Swedish University of Agricultural Sciences. Uppsala. Sveriges Lantbruksuniversitet Rep 17
- Paranaíba JR, Barros N, Mendonça R et al (2018) Spatially resolved measurements of CO_2 and CH_4 concentration and gas-exchange velocity highly influence carbon-emission estimates of reservoirs. *Environ Sci Technol* 52:607–615. <https://doi.org/10.1021/acs.est.7b05138>
- Piñeiro G, Perelman S, Guerschman JP, Paruelo JM (2008) How to evaluate models: observed vs. predicted or predicted vs. observed? *Ecol Modell* 216:316–322. <https://doi.org/10.1016/j.ecolmodel.2008.05.006>
- Podgrajsek E, Sahlée E, Rutgersson A (2015) Diel cycle of lake-air CO_2 flux from a shallow lake and the impact of waterside convection on the transfer velocity. *J Geophys Res Biogeosci*. <https://doi.org/10.1002/2014JG002781>. Received
- Poissant L, Amyot M, Pilote M, Lean D (2000) Mercury water–air exchange over the upper St. Lawrence River and Lake Ontario. *Environ Sci Technol* 34:3069–3078. <https://doi.org/10.1021/es990719a>
- Polsenaere P, Deborde J, Detandt G et al (2013) Thermal enhancement of gas transfer velocity of CO_2 in an Amazon floodplain lake revealed by eddy covariance measurements. *Geophys Res Lett* 40:1734–1740. <https://doi.org/10.1002/grl.50291>
- Poulain AJ, Orihel DM, Amyot M et al (2006) Relationship between the loading rate of inorganic mercury to aquatic ecosystems and dissolved gaseous mercury production and evasion. *Chemosphere* 65:2199–2207. <https://doi.org/10.1016/j.chemosphere.2006.05.066>
- Prairie YT, Giorgio PA (2013) A new pathway of freshwater methane emissions and the putative importance of microbubbles. *Inl waters* 3:311–320. <https://doi.org/10.5268/IW-3.3.542>
- R Core Team (2019) R: a language and environment for statistical computing. <https://www.r-project.org>. Accessed 03 Apr 2020
- Rantakari M, Heiskanen J, Mammarella I et al (2015) Different apparent gas exchange coefficients for CO_2 and CH_4 : comparing a brown-water and a clear-water lake in the boreal zone during the whole growing season. *Environ Sci Technol* 49:11388–11394. <https://doi.org/10.1021/acs.est.5b01261>
- Raymond PAP, Hartmann J, Lauerwald R et al (2013) Global carbon dioxide emissions from inland waters. *Nature* 503:355–359. <https://doi.org/10.1038/nature12760>
- Read JS, Hamilton DP, Desai AR et al (2012) Lake-size dependency of wind shear and convection as controls on gas exchange. *Geophys Res Lett* 39:L09405. <https://doi.org/10.1029/2012GL051886>
- Rohatgi A (2019) WebPlotDigitizer 4.2. <https://automeris.io/WebPlotDigitizer> Accessed 16 Apr 2019
- Rosentreter JA, Maher DT, Ho DT et al (2017) Spatial and temporal variability of CO_2 and CH_4 gas transfer velocities and quantification of the CH_4 microbubble flux in mangrove dominated estuaries. *Limnol Oceanogr* 62:561–578. <https://doi.org/10.1002/lno.10444>
- Scherbaum CA, Ferreter JM (2009) Estimating statistical power and required sample sizes for organizational research using multi-level modeling. *Organ Res Methods* 12:347–367. <https://doi.org/10.4135/9780857028228>
- Schilder J, Bastviken D, Van Hardenbroek M et al (2013) Spatial heterogeneity and lake morphology affect diffusive greenhouse gas emission estimates of lakes. *Geophys Res Lett* 40:5752–5756. <https://doi.org/10.1002/2013GL057669>
- Schilder J, van Hardenbroek M, Bastviken D, Heiri O (2016) Spatio-temporal patterns in methane flux and piston velocity at low wind

- speed: implications for upscaling studies on small lakes. *J Geophys Res Biogeosci*. <https://doi.org/10.1002/2016JG003346>
- Sebacher DI, Harriss RC, Bartlett KB (1983) Methane flux across the air-water interface: air velocity effects. *Tellus B* 35B:103–109. <https://doi.org/10.1111/j.1600-0889.1983.tb00014.x>
- Sollberger S, Wehrli B, Schubert CJ et al (2017) Minor methane emissions from an Alpine hydropower reservoir based on monitoring of diel and seasonal variability. *Environ Sci Process Impacts* 19:1278–1291. <https://doi.org/10.1039/c7em00232g>
- Soued C, del Giorgio PA, Maranger R (2015) Nitrous oxide sinks and emissions in boreal aquatic networks in Québec. *Nat Geosci*. <https://doi.org/10.1038/ngeo2611>
- Soumis N, Canuel R, Lucotte M (2008) Evaluation of two current approaches for the measurement of carbon dioxide diffusive fluxes from lentic ecosystems. *Environ Sci Technol* 42:2964–2969. <https://doi.org/10.1021/es702361s>
- Southworth G, Lindberg S, Hintelmann H et al (2007) Evasion of added isotopic mercury from a northern temperate lake. *Environ Toxicol Chem* 26:53–60. <https://doi.org/10.1897/06-148r.1>
- Stepanenko V, Mammarella I, Ojala A et al (2016) LAKE 2.0: a model for temperature, methane, carbon dioxide and oxygen dynamics in lakes. *Geosci Model Dev* 9:1977–2006. <https://doi.org/10.5194/gmd-9-1977-2016>
- Tan Z, Zhuang Q (2015) Methane emissions from pan-Arctic lakes during the 21st century: an analysis with process-based models of lake evolution and biogeochemistry. *J Geophys Res Biogeosci* 120:1–13. <https://doi.org/10.1002/2015JG003184>
- Theall KP, Scribner R, Broyles S et al (2011) Impact of small group size on neighbourhood influences in multilevel models. *J Epidemiol Community Health* 65:688–695. <https://doi.org/10.1136/jech.2009.097956>
- Torgersen T, Mathieu G, Hesslein RH, Broecker WS (1982) Gas exchange dependency on diffusion coefficient: direct ^{222}Rn and ^3He comparisons in a small lake. *J Geophys Res* 87:546–556. <https://doi.org/10.1029/JC087iC01p00546>
- Upstill-Goddard RC, Watson AJ, Liss PS, Liddicoat MI (1990) Gas transfer velocities in lakes measured with SF_6 . *Tellus B* 42:364–377. <https://doi.org/10.1034/j.1600-0889.1990.t01-3-00006.x>
- Vachon D, Langenegger T, Donis D et al (accepted) Methane emission offsets carbon dioxide uptake in a small productive lake. *Limnol Oceanogr Lett*
- Vachon D, Prairie YT (2013) The ecosystem size and shape dependence of gas transfer velocity versus wind speed relationships in lakes. *Can J Fish Aquat Sci* 70:1757–1764. <https://doi.org/10.1139/cjfas-2013-0241>
- Vachon D, Prairie YT, Cole JJ (2010) The relationship between near-surface turbulence and gas transfer velocity in freshwater systems and its implications for floating chamber measurements of gas exchange. *Limnol Oceanogr* 55:1723–1732. <https://doi.org/10.4319/lo.2010.55.4.1723>
- Venäläinen A, Sahlgrén V, Podsechin V, Huttula T (2003) Small-scale variability of the wind field over a typical Scandinavian lake. *Boreal Environ Res* 8:71–81
- Verpoorter C, Kutser T, Seekell DA, Tranvik LJ (2014) A global inventory of lakes based on high-resolution satellite imagery. *Geophys Res Lett*. <https://doi.org/10.1002/2014GL060641>
- Wanninkhof R (2014) Relationship between wind speed and gas exchange over the ocean revisited. *Limnol Oceanogr Methods* 12:351–362. <https://doi.org/10.4319/lom.2014.12.351>
- Wanninkhof R, Asher WE, Ho DT et al (2009) Advances in quantifying air-sea gas exchange and environmental forcing. *Ann Rev Mar Sci* 1:213–244. <https://doi.org/10.1146/annurev.marine.010908.163742>
- Wanninkhof R, Ledwell J, Crusius J (1991) Gas transfer velocities on lakes measured with sulfur hexafluoride. In: Wilhelms SC, Gulliver JS (eds) *Proceedings of the Second International Symposium on Gas Transfer at Water Surfaces*. American Society of Civil Engineers, New York, pp 441–458
- Wanninkhof R, Ledwell JR, Broecker WS (1985) Gas exchange-wind speed relation measured with sulfur hexafluoride on a lake. *Science* (80-) 227:1224–1226. <https://doi.org/10.1126/science.227.4691.1224>
- Wanninkhof R, Ledwell JR, Broecker WS, Hamilton M (1987) Gas exchange on Mono Lake and Crowley Lake, California. *J Geophys Res Ocean* 92:14567–14580. <https://doi.org/10.1029/JC092iC13p14567>
- Witherspoon PA, Saraf DN (1965) Diffusion of methane, ethane, propane, and n-butane in water from 25 to 43°. *J Phys Chem* 69:3752–3755. <https://doi.org/10.1021/j100895a017>
- Xiao S, Yang H, Liu D et al (2014) Gas transfer velocities of methane and carbon dioxide in a subtropical shallow pond. *Tellus, Ser B Chem Phys Meteorol*. <https://doi.org/10.3402/tellusb.v66.23795>
- Yu SL, Tuffey TJ, Lee D-S (1977) Atmospheric reaeration in a lake. Water problems in an urbanizing state. Rutgers. The State University of New Jersey. Water Resources Research Institute. 1–50
- Zappa CJ, McGillis WR, Raymond PA et al (2007) Environmental turbulent mixing controls on air-water gas exchange in marine and aquatic systems. *Geophys Res Lett* 34:1–6. <https://doi.org/10.1029/2006GL028790>
- Zwart JA, Hanson ZJ, Vanderwall J et al (2018) Spatially explicit, regional-scale simulation of lake carbon fluxes. *Global Biogeochem Cycles* 32:1276–1293. <https://doi.org/10.1002/2017GB005843>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.