

JGR Atmospheres

RESEARCH ARTICLE

10.1029/2020JD033417

Key Points:

- By testing against a highly diverse lake observation data set, Arctic Lake Biogeochemistry Model is shown to be capable of simulating global lake thermal dynamics
- Turbulent heat fluxes, wind-driven mixing, water clarity, and snow density are the main uncertainty sources for modeled lake thermal regimes
- The relative importance of the key processes depends on lake latitude and depth

Supporting Information:

- Supporting Information S1

Correspondence to:

Q. Zhuang and Z. Tan,
qzhuang@purdue.edu;
zeli.tan@pnnl.gov

Citation:

Guo, M., Zhuang, Q., Yao, H., Golub, M., Leung, L. R., Pierson, D., & Tan, Z. (2021). Validation and sensitivity analysis of a 1-D lake model across global lakes. *Journal of Geophysical Research: Atmospheres*, 126, e2020JD033417. <https://doi.org/10.1029/2020JD033417>

Received 30 JUN 2020

Accepted 21 DEC 2020

Author Contributions:

Conceptualization: Mingyang Guo, Zeli Tan

Funding acquisition: Qianlai Zhuang

Methodology: Mingyang Guo, Huaxia Yao, Zeli Tan

Project Administration: Malgorzata Golub








Supervision: Qianlai Zhuang

Visualization: Mingyang Guo

Writing – original draft: Mingyang Guo

Writing – review & editing: Qianlai Zhuang, Huaxia Yao, Malgorzata Golub, L. Ruby Leung, Don Pierson, Zeli Tan

Validation and Sensitivity Analysis of a 1-D Lake Model Across Global Lakes

Mingyang Guo¹ , Qianlai Zhuang^{1,2} , Huaxia Yao³ , Malgorzata Golub⁴ , L. Ruby Leung⁵ , Don Pierson⁴ , and Zeli Tan⁵ 

¹Department of Earth, Atmospheric and Planetary Sciences, Purdue University, West Lafayette, IN, USA, ²Purdue Climate Change Research Center, West Lafayette, IN, USA, ³Dorset Environmental Science Centre, Ontario Ministry of Environment, Conservation and Parks, Dorset, ON, Canada, ⁴Department of Ecology and Genetics/Limnology, Uppsala University, Uppsala, Sweden, ⁵Pacific Northwest National Laboratory, Richland, WA, USA

Abstract Lakes have important influence on weather and climate from local to global scales. However, their prediction using numerical models is notoriously difficult because lakes are highly heterogeneous across the globe, but observations are sparse. Here, we assessed the performance of a 1-D lake model in simulating the thermal structures of 58 lakes with diverse morphometric and geographic characteristics by following the phase 2a local lake protocol of the Intersectoral Impact Model Intercomparison Project (ISIMIP2a). After calibration, the root-mean-square errors (RMSE) were below 2°C for 70% and 75% of the lakes for epilimnion and full-profile temperature simulations, with an average of 1.71°C and 1.43°C, respectively. The model performance mainly depended on lake shape rather than location, supporting the possibility of grouping model parameters by lake shape for global applications. Furthermore, through machine-learning based parameter sensitivity tests, we identified turbulent heat fluxes, wind-driven mixing, and water transparency as the major processes controlling lake thermal and mixing regimes. Snow density was also important for modeling the ice phenology of high-latitude lakes. The relative influence of the key processes and the corresponding parameters mainly depended on lake latitude and depth. Turbulent heat fluxes showed a decreasing importance in affecting epilimnion temperature with increasing latitude. Wind-driven mixing was less influential to lake stratification for deeper lakes while the impact of light extinction, on the contrary, showed a positive correlation with lake depth. Our findings may guide improvements in 1-D lake model parameterizations to achieve higher fidelity in simulating global lake thermal dynamics.

1. Introduction

Lakes are a critical component of the earth system that strongly influence the global water cycle and energy balance. Lakes modulate local atmospheric boundary layer conditions by affecting fluxes of heat, moisture, and momentum. For example, the presence of lakes was found to produce a warming effect on regional climate (Samuelsson et al., 2010), enhance water cycling between adjacent land grids, and alter the behavior of severe weather events (King et al., 2003; Thiery et al., 2016). It is thus important to represent lakes in global climate models to improve the accuracy of air temperature and precipitation prediction, as shown in previous modeling studies (Balsamo et al., 2012; Bates et al., 1993; Dutra et al., 2010; Hostetler et al., 1993; Long et al., 2007; Mironov et al., 2008). Also, studies have shown that lakes are sensitive to environmental changes such as warmer temperature and thus are likely important sentinel of climate change (Adrian et al., 2009; Carpenter et al., 2011; O'Reilly et al., 2015; Schmid et al., 2014; Woolway & Merchant, 2019). Therefore, it is of great scientific and societal importance to understand the historical and future climate change impacts on lakes and modeling is a powerful tool. Moreover, accurate modeling of lake thermodynamic properties and processes is essential for simulating lake carbon dynamics, such as CO₂ and CH₄ emissions, of which lakes have been identified as an important source (Bastviken et al., 2011; Guo, Zhuang, Tan, et al., 2020; Prairie et al., 2013; Saunio et al., 2016; Tan & Zhuang, 2015a, 2015b).

Based on the research purpose, lake models varying in complexity and formulations have been developed and selected for modeling lakes of different sizes, regions, and mixing regimes (e.g., General Lake Model-GLM, Hipsey et al., 2019; FLake, Mironov et al., 2008; LAKE, Stepanenko et al., 2010, 2016). Although more advanced three-dimensional (3-D) lake models are now available, 1-D models are and will still be

widely used because of their simplicity in parameterization and computational efficiency in applications that require long-term simulations. Also, small lakes (<0.1 km²), which can be assumed to be horizontally homogeneous and can thus be reasonably represented by a 1-D model, dominate the lake systems on earth (Downing et al., 2006; McDonald et al., 2012; Paltan et al., 2015) and are hotspots of carbon emissions (Guo, Zhuang, Tan, et al., 2020; Holgerson & Raymond, 2016; Tan & Zhuang, 2015a, 2015b; Wik et al., 2016). As such, improving the capability of 1-D lake models in representing the thermal regimes of diverse lake systems is of great value. Previous lake model validations mostly focused on individual lakes because data for calibration and verification of a large variety of lakes are not readily available (Guseva et al., 2020; Stepanenko et al., 2010). Also, previous studies did not identify which parameters or model processes should be the focus for improving model performance at the global scale, making it difficult to determine whether the developed numerical treatments are appropriate for regional or global applications is unclear. It is thus important to test model across diverse lakes to help develop robust parameterization for global-scale simulations.

The Arctic Lake Biogeochemistry Model (ALBM) is a 1-D process-based lake biogeochemical model originally developed to simulate the thermal and carbon dynamics of Arctic lakes (Tan et al., 2015, 2017) and later successfully applied to boreal lakes (Guo, Zhuang, Tan, et al., 2020) and temperate lakes (Guseva et al., 2020; Tan et al., 2018). The Intersectoral Impact Model Intercomparison Project (ISIMIP) lake sector provides a framework for multimodel ensemble and model intercomparison, allowing for improved predictions of changes in the thermal and biological properties of global inland water bodies under climate warming. The eight-model ensemble (Table S1) shows high robustness in the reported climate impacts on lakes. As ALBM is used in the ISIMIP lake sector (Vanderkelen et al., 2020), continued testing and improvement would be of great values to both the ISIMIP project and the scientific community. In this study, we evaluated the performance of ALBM over 58 lakes across the globe following the ISIMIP2a local lake protocol. By taking advantage of the highly diverse data set provided by the ISIMIP project, we aim to (1) assess the ALBM accuracy in simulating lake thermal profiles and investigate the dependence of model performance on lake characteristics, (2) analyze parameter sensitivities and identify the key processes as well as the parameters that affect the model simulations, and (3) discuss a possible strategy to improve the accuracy of ALBM at the global scale. To provide a comprehensive evaluation of the modeling practice, we analyzed the model outputs from several aspects including the epilimnion and full-profile temperature, lake stratification, and ice phenology.

2. Data and Method

2.1. Study Sites

Our study examined 58 freshwater lakes varying in shape, landscape and climate (Figure 1 and Table S2). Among the 58 lakes, 17 are located north to 50°N, 35 are within latitudes 25–50°N, two are in the tropics and the remaining four are in the Southern Hemisphere within 35–40°S. Lake surface areas range from 0.14 to 2,700 km², maximum depths from 6.4 to 501 m and elevations from 210 m below to 4,300 m above sea level. For presentation convenience, lakes are referred to by abbreviations (as in Table S2) in the figures. Sufficient water temperature at multilayers with high frequency or at least 5 years of monthly observations were available for each lake and were used for model calibration (Text S1). All lakes have observational data during the most recent decade while the earliest can be traced back to the 1902. Furthermore, bathymetry measurements are available for all lakes.

2.2. ALBM Model Configuration

In ALBM, a water body is divided into vertical layers and the thermal dynamics within each layer consists of three processes: heat exchange with the air at the water surface, heat transfer and mixing in the water column driven by wind-induced eddies and molecular diffusion and heat exchange with the sediment at the bottom. The governing equation of thermal diffusion is

$$\frac{\partial T_w}{\partial t} = \frac{1}{A} \frac{\partial}{\partial z} \left(A \left(D_m + ktscale D_e \right) \frac{\partial T_w}{\partial z} \right) + \frac{1}{A} \frac{1}{\rho_w c_{pw}} \frac{\partial [\Phi A]}{\partial z}, \quad (1)$$

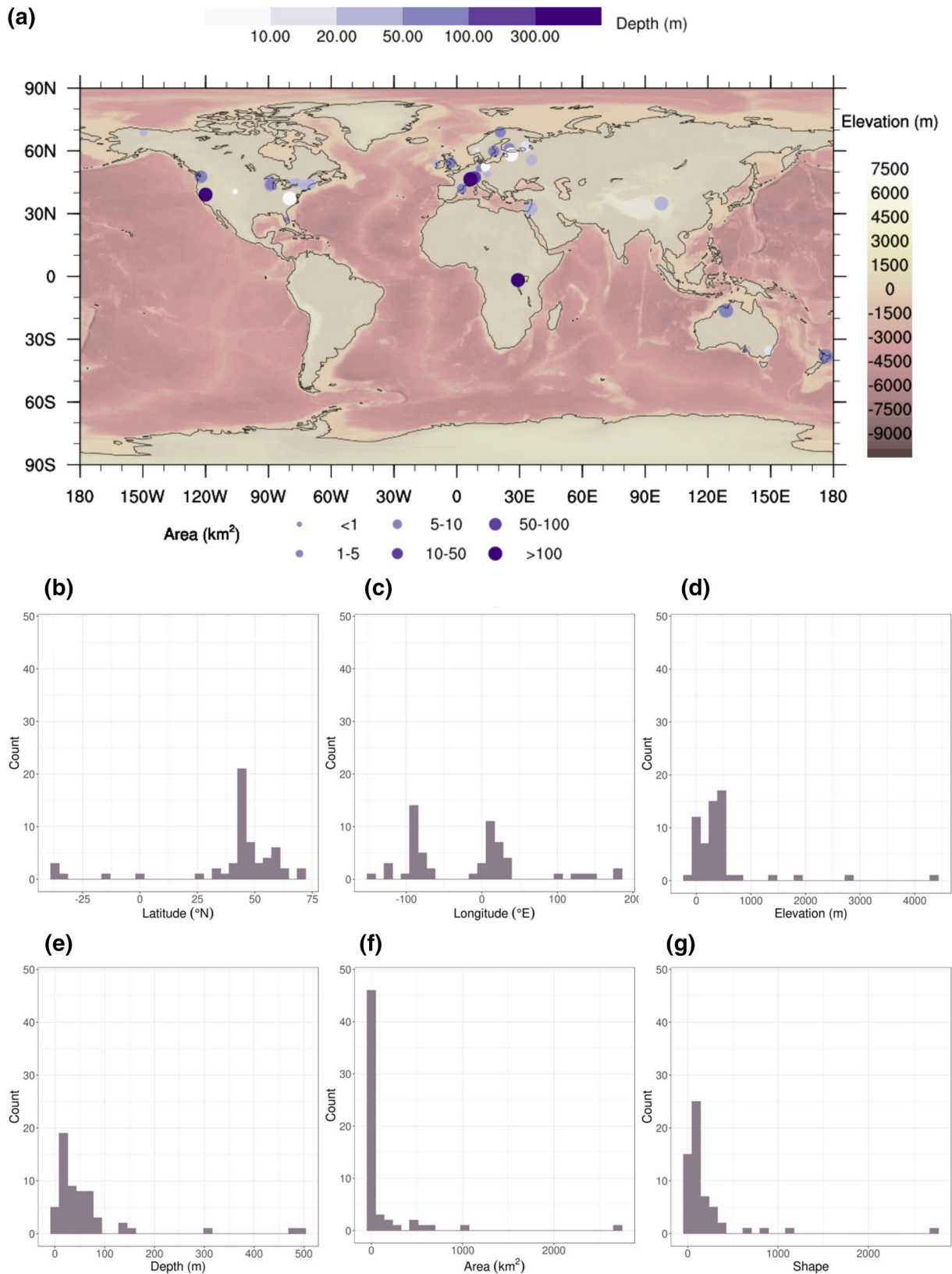


Figure 1. A map of all lakes in the dataset with background elevation contours, lake area denoted by the size of the points and depth by the colors (a) and histograms of lake latitude (b), longitude (c), elevation (d), depth (e), area (f), and shape factor (g).

where T_w is the water temperature (K) at depth z (m), t is time (s), A is the lake cross-section area at depth z (m^2), D_m is the molecular diffusivity (m s^{-1}), D_e is the wind-driven eddy diffusivity (m s^{-1}) (Hostetler & Bartlein, 1990), $ktscale$ is a scaling factor of D_e that compensates the underestimation of D_e and the unresolved hydrodynamical processes by Hostetler and Bartlein (1990) for large lakes, ρ_w is the water density (kg m^{-3}), c_{pw} is the specific heat of water ($\text{J kg}^{-1} \text{K}^{-1}$) and $\Phi = (1 - \beta)L_0e^{-\eta z}$ is the solar radiation (W m^{-2}) penetrating to depth z . β is the fraction reflected by water surface, L_0 is the incident solar radiation (W m^{-2}) at lake surface and η is the light extinction coefficient (m^{-1}) calculated by a modified equation from Subin et al. (2012):

$$\eta = feta \cdot 1.1925d^{-0.424}, \quad (2)$$

where $feta$ is a correction factor of η for shallow lakes and d is the lake depth (m). The bulk expressions are used to calculate the sensible heat (Q_H) and latent heat (Q_E) by

$$Q_H = \rho_a c_{pa} hwt C_H u (T_w - T_a), \quad (3)$$

$$Q_E = \rho_a L hwt C_E u (q_w - q_a), \quad (4)$$

where ρ_a is the air density (kg m^{-3}), c_{pa} is the specific heat of air ($\text{J kg}^{-1} \text{K}^{-1}$), $C_H = C_E = 1.75 \times 10^{-3}$ are the bulk transfer coefficients for the sensible and latent heat, respectively, hwt is a correction factor of C_H and C_E , u is the 2-m wind speed (m s^{-1}), and T_w and T_a are temperatures (K) and q_w and q_a are specific humidity of the surface water and of air at 2 m, respectively. The model represents the convective mixing as the balance of wind-induced kinetic energy and thermal and chemical stratification induced potential energy (Saloranta & Andersen, 2007) with kinetic energy adjusted by the wind shielding coefficient $F_{wstr} = wstr(1 - e^{-0.3A_s})$, where $wstr$ is a correction factor for irregular lakes and A_s is the lake surface area (km^2). Notably, convective mixing can be triggered in the model due to the occurrence of negative water density profiles even when wind-induced turbulence is small. The snow and ice dynamics of lakes are explicitly represented in ALBM by one snow layer, one gray ice layer formed when the weight of snow depress the ice and leads to the seepage of lake water into the snowpack, and multiple ice layers (Tan et al., 2018). The ice cover grows from top to bottom and melts from bottom to top. When the ice thickness of a lake layer equals to the layer's thickness, this layer is then counted as an ice layer. The thickness of ice cover equals to the sum of thickness of all ice layers plus the ice thickness of the partially frozen water layer. ALBM treats the reflection of solar radiation for five types of lake surface (i.e., water, ice, gray ice, fresh snow, and melting snow) differently. Furthermore, the albedo for each type of lake surface is calculated separately for direct and diffusive radiation following the method of Koehler et al. (2014). For fresh snow and melting snow, we use prescribed albedo profiles covering the wavelength band from the Simple Model of the Atmospheric Radiative Transfer of Sunshine (Gueymard, 1995). For gray ice, the albedo is defined as 0.4 and 1.0 for direct and diffusive radiation, respectively. For water and ice, the albedo is fixed at 0.066 and 0.55, respectively, for diffusive radiation but calculated using Fresnel's equation for direct radiation (Koehler et al., 2014). The governing diffusion equation for sediment heat transfer is:

$$c_{vs} \frac{\partial T_s}{\partial t} = \frac{\partial}{\partial z} \left(k_s \frac{\partial T_s}{\partial z} \right), \quad (5)$$

where $c_{vs} = \rho_w c_{pw} \theta + \rho_i c_{pi} \theta_i + \rho_s c_{ps} (1 - pi)$ is the volumetric heat capacity of the sediment ($\text{J m}^{-3} \text{K}^{-1}$), ρ_i and ρ_s are the density of sediment ice and soil, respectively, c_{pi} and c_{ps} are the specific heat ($\text{J kg}^{-1} \text{K}^{-1}$), θ and θ_i are the water and ice content, respectively, and pi is the sediment porosity. k_s is the sediment heat conductivity ($\text{W m}^{-1} \text{K}^{-1}$) and T_s is the sediment temperature (K).

Daily climate forcing data were retrieved from the global gridded data set of historical bias-corrected climate input data provided by ISIMIP2a (Frieler et al., 2017; Lange et al., 2019) for each lake from 1979 to 2016 at a daily time step (Text S1). All simulations were run for the same period with spin-up periods ranging from 2 to 50 years depending on the mixing regime of the lake. The initial temperature of the lake water column was set to be 4°C uniformly. We used 50 vertical grid levels for all the lakes and applied an

Table 1
Abbreviations, Descriptions, and Value Ranges of the Parameters in Model Calibration

Parameter	Description	Value range
k_s	Thermal conductivity of sediment ($\text{W m}^{-1} \text{K}^{-1}$)	[0.25,0.29]
c_{ps}	Heat capacity of sediment soil ($\text{J kg}^{-1} \text{K}^{-1}$)	[750,1930]
pi	Porosity of sediment	[0.3,0.6]
ρ_s	Density of sediment solid particle (kg m^{-3})	[1500,2700]
ρ_n	Density of snow (kg m^{-3})	[100,400]
$feta$	Light attenuation correction factor	[0.1,10]
$wstr$	Wind shielding factor of mixing	[0.1,10]
$kt scale$	Scaling factor of turbulence diffusivity	[0.5,2]
hwt	Scaling factor of bulk coefficients for turbulent latent/sensible heat transfer	[0.5,5]

exponential-stretched spacing with finer grid resolutions in the upper layers to accurately resolve the complex radiation, heat transfer and mixing, and ice processes. Inflow or outflow data are not available for most of the lakes, and consequently, we decided to keep constant water levels for all lakes in this study.

2.3. Model Performance Evaluation

Based on previous studies using ALBM (Guo, Zhuang, Tan, et al., 2020; Tan et al., 2015, 2017), we calibrated nine parameters related to lake sediment properties (solid particle thermal conductivity, k_s ; solid particle heat capacity, c_{ps} ; sediment porosity, pi ; solid particle density, ρ_s), radiative transfer (light extinction coefficient, $feta$), turbulent heat and momentum fluxes (wind shielding factor, $wstr$; turbulence diffusivity scaling factor, $kt scale$; bulk heat transfer factor, hwt), and snow density (ρ_n), respectively. Detailed explanation and value ranges for each parameter can be found in Table 1. For each lake, 6 years of the observation data were used for calibration and the rest were used for model performance evaluation. Some lakes have observation shorter than 6 years, so we only show the calibration results for these lakes to represent the model performance. Model calibration was conducted in the following two steps: (1) we first ran Monte Carlo simulations for each lake using a perturbed parameter ensemble (PPE) of 10,000 samples with nine parameters generated by Sobol sequence sampling method (Sobol, 1967) varying all together; (2) the root-mean-square errors (RMSEs) of epilimnion and full-profile temperature simulations were then calculated separately and the sample generating the minimum mean of the two RMSE values was selected for each lake as the optimum parameter set. Full-profile temperature is defined as the volumetrically weighted average whole lake temperature. Two other metrics, correlation coefficient (r) and percentage relative error (PRE) were also calculated to assess the model performance. We further analyzed the correlation between the simulation evaluation metrics and the lake characteristics including lake latitude, depth, surface area, shape factor ($=\sqrt{\text{surface area}/\text{depth}}$), elevation, and stability. The overall parameter uncertainties were calculated for each lake as the standard deviations of the model output using the PPE.

2.4. Model Sensitivity Analysis

Model sensitivity was tested for the nine parameters mentioned above. In addition, we evaluated the sensitivities of two categorical parameters: water-sediment heat flux on/off ($sedh$) and lake bathymetry construction scheme ($bthmtry$). By default, the water-sediment heat flux was enabled but in the $sedh$ sensitivity test this heat flux was turned off to investigate how this flux term would affect the thermal dynamics of the study lakes. In the $bthmtry$ sensitivity test, three bathymetry schemes were considered: the rectangular-shape profile, the conic-shape profile, and the observed profile. Although the observed bathymetry profiles are available for all the lakes in this study, they are usually hard to obtain, especially for small lakes. Therefore, we would like to investigate whether using a prescribed shape would compromise the simulation accuracy.

To analyze model sensitivity for a mixture of numerical and categorical parameters, we adopted a machine-learning algorithm, classification and regression trees (CART) (Breiman et al., 1984) method which

has been found adequate (Choubin et al., 2018; Rodrigues & de la Riva, 2014) and widely used to detect the relationships between variables in large complex data and to make predictions (Gou, Zhuang, Tan, et al., 2020; W. Zhang et al., 2016). The CART method recursively splits the target data set, in our study, the model outputs, by selecting the explanatory variable and value that best distinguish the two subgroups at each node and resulting in a decision tree. The pruning process is to build a tree with the best size and the lowest misclassification rate. The CART method is able to examine the complicated interaction among parameters in the final tree and to quantify the relative influence (RI) of each parameter. In our study, the model output with 10,000 sample sets run for 1 year was used to train the regression trees for each lake. To justify using of only 1 year of simulation, we trained the regression trees with 5-year and 20-year simulations, respectively, for several lakes and found that the resulting parameter RI values were highly consistent (Figure S1). The optimal number of iterations was estimated using the cross-validation method. A tree size of 500 was found to be the most appropriate. Data training and sensitivity analysis were done using the R package “gbm” (Greenwell et al., 2019).

To gauge the parameter sensitivity comprehensively, the CART models were trained to evaluate the sensitivities of several lake features from three aspects: water temperature (lake epilimnion and full-profile temperature), lake stability (center of buoyancy, N , and Schmidt stability, SSI), and ice phenology (ice-on and ice-off day, ice-cover duration, ice thickness, and snow thickness). Water temperature is calculated for open water periods only. The ice-on day is simply defined as the first day that ice appears at surface, the ice-off day is the first day that the ice cover is gone, and the ice duration is the period between these two days. The calculation of lake stratification and stability were all conducted using the R package “rLakeAnalyzer” (Winslow et al., 2019).

3. Results

3.1. Model Performance

In the model calibration, for the epilimnion and full-profile temperatures of the 58 lakes, the average RMSEs were 1.61°C and 1.24°C, the average r values were 0.95 and 0.94, the average PREs were 0.08% and -0.13%, and the average absolute values of PREs were 0.14% and 0.20%, respectively (Figure 2). The highest RMSEs (>2.5°C) of epilimnion temperature simulations occurred at Eagle, Erken, Monona, and Toolik. Lakes with the highest RMSEs (>2°C) in full-profile temperatures are Allequash, BigMuskellunge, Eagle, Erken, Monona, Rotorua, and Trout.

In the model performance evaluation, for the epilimnion and full-profile temperatures of the 58 lakes, the average RMSEs were 1.71°C and 1.43°C, the average r values were 0.95 and 0.94, the average PREs were 0.05% and -0.19%, and the average absolute values of PREs were 0.17% and 0.23%, respectively (Figure S2). The average parameter uncertainty was 2.10°C and 1.86°C for epilimnion and full-profile temperature simulations, respectively (Figure S3). The model generally performed better in predicting the full-profile temperature than the epilimnion temperature, consistent with our previous assessment on a temperate lake (Guseva et al., 2020).

In both calibration and performance evaluation, all lakes had r values over 0.8 except for Lake Kivu with exceptionally low r values of 0.38 and 0.31 for the epilimnion and the full-profile temperature, respectively. Lake Kivu is located near the equator, which results in stable water temperatures all year round with small epilimnion and full-profile temperature deviations (only 0.93°C and 0.22°C, respectively) (Figure S4). As such, slight discrepancies in the model simulation would compromise the r values. For this reason, we neglected this outlier in our analysis of the correlation between lake characteristics and model performance. Argyle, Laramie, Rotorua, and Toolik also have relatively low r values (<0.9) for both epilimnion and full-profile temperature simulations. Argyle and Rotorua are similar in the thermal regime to Kivu with stable warm temperatures all year round (not shown). Laramie only had observations for 6 days and the calculation of r was thus not as informative. For Toolik, the model overestimated the temperature fluctuation. All PREs were within $\pm 1\%$ and the values indicated that the model tended to slightly overestimate the epilimnion temperature but underestimate the full-profile temperature.

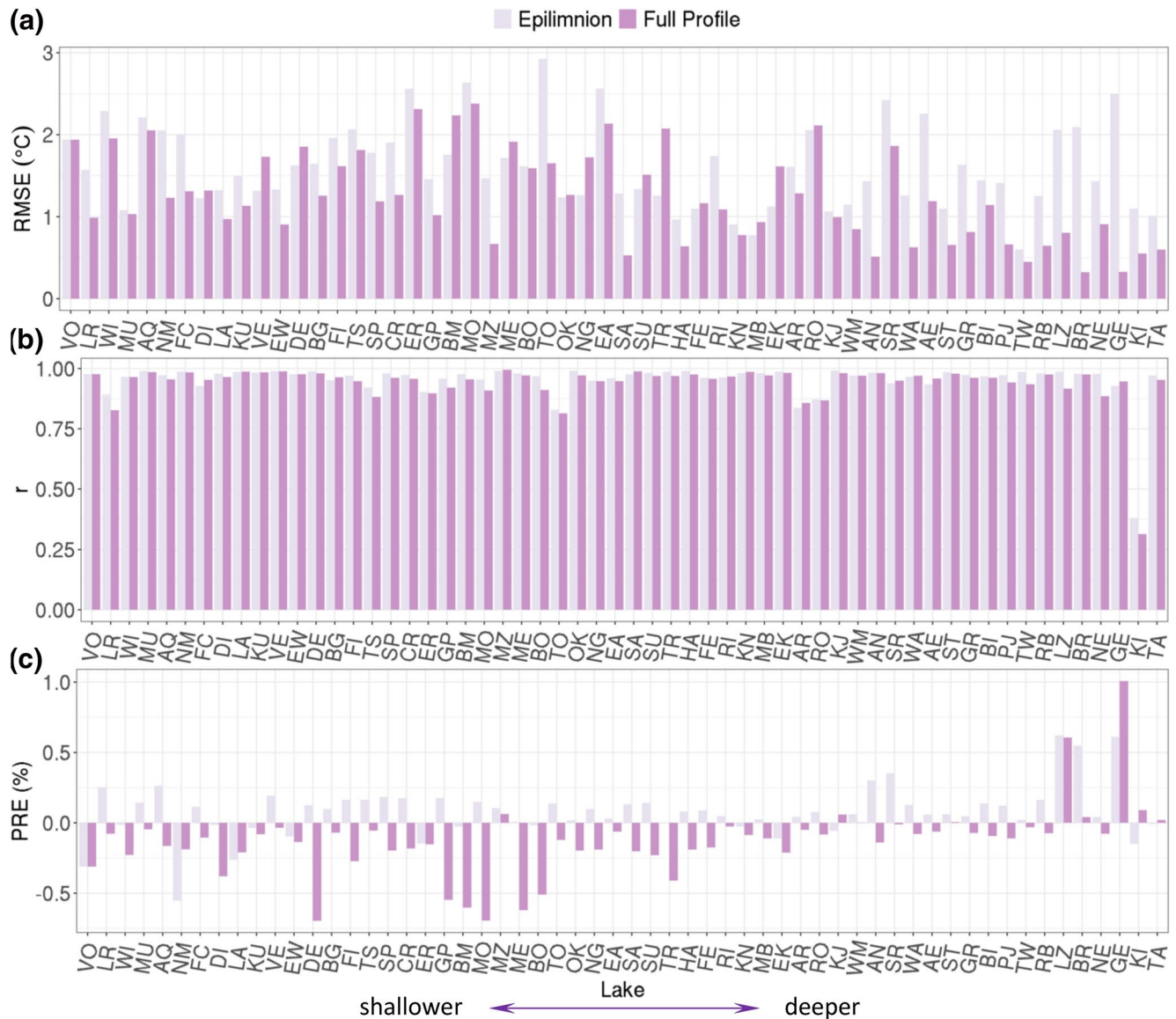


Figure 2. Root-mean-square error (RMSE) (a), correlation coefficient (r) (b), and percentage relative error (PRE) (c) of lake epilimnion and full-profile temperature simulations in model calibration. Lakes are arranged in the order of the maximum depth.

The correlation coefficients (r) between model performance and lake properties for epilimnion temperature simulations decreased for larger lakes (Figure 3). The prediction of full-profile temperatures was found to increase in accuracy for deeper lakes. The lake stability was also found to slightly impact the simulation accuracy such that with higher stability, the RMSEs of full-profile temperature simulations decreased.

We also compared the simulated turbulent heat fluxes at six lakes, including Annie, Esthwaitewater, Feeagh, Rotorua, Tahoe, and Windermere, with the calculated values in Woolway et al. (2018). Overall, the model performs well in simulating turbulent energy exchange while for lakes Feeagh, Rororua, and Windermere reproduced lower values with weaker seasonal cycles relative to Woolway et al. (2018) (Figure S5). This could be mainly because Woolway et al. (2018) calculated heat fluxes from the observed water temperature and meteorological variables using a Monin-Obukhov similarity-based method with empirical parameters, which has been found to overestimate turbulent heat fluxes (Charusombat et al., 2018). Also, they used field observed wind speeds which likely have higher accuracy and more fluctuations, leading to stronger seasonal cycles than our results.

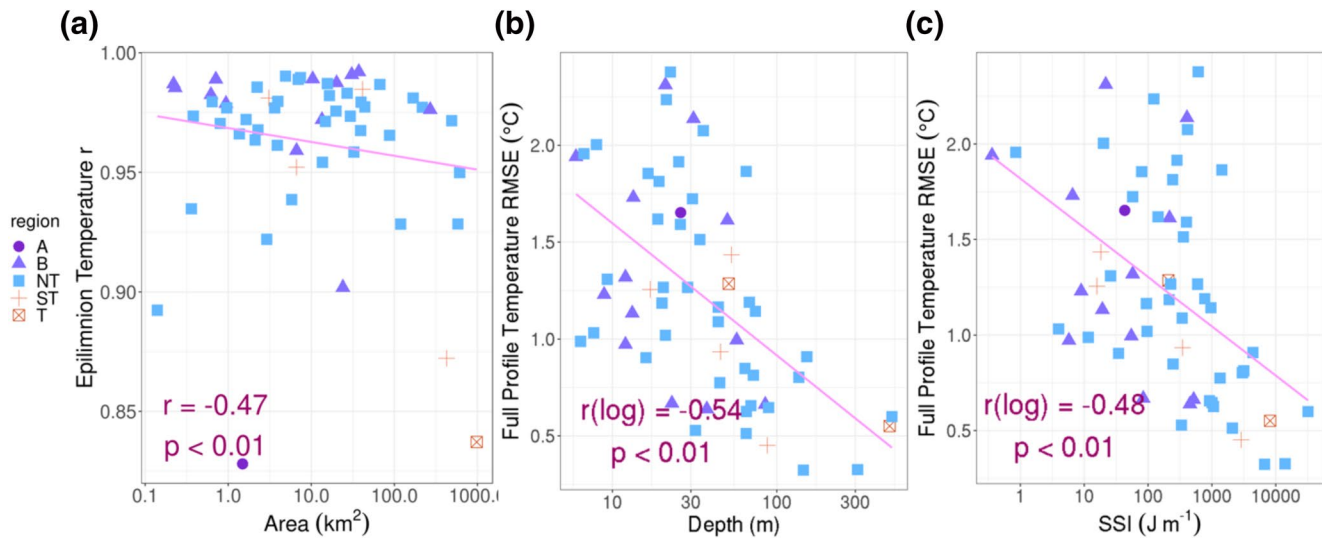


Figure 3. Correlation between epilimnion temperature correlation coefficients and lake surface areas (a), full-profile temperature RMSEs and logarithmic lake depths (b), and full-profile temperature RMSEs and logarithmic Schmidt stabilities (c). The lines are regression fits. The colors denote the Arctic (A), boreal (B), northern temperate (NT), southern temperate (ST) and tropic (T) regions, respectively. Lake Kivu is not included in (a). RMSE, Root-mean-square error.

3.2. Sensitivity Analysis

3.2.1. Lake Temperature Profile

The parameter sensitivities are highly consistent among all the lakes for epilimnion temperature simulations (Figure 4) where hwt is the dominant parameter with an average RI value of 76.81, followed by $wstr$ with RI of 19.29 and $ktscale$ of 2.1. These three parameters together explained over 97.5% of the change in modeled epilimnion temperature for all lakes except two Arctic lakes. For Kilpisjarvi and Toolik in high latitudes, ρ_n was also important with RI of 6.66 and 11.37, respectively. For the full-profile temperature, the pattern became diverse with hwt , $wstr$, and $feta$ as the three most important parameters with mean RI values of 45.39, 20.45, and 31.15, respectively, but the relative influence varied significantly among the lakes. The importance of hwt showed a slightly decreasing trend from tropics to northern high latitudes in the prediction of epilimnion temperature (Figure 5) indicating that the turbulent heat exchange process has a larger impact on low-latitude lakes. This can be explained by the decreasing energy inputs to lakes toward the poles, so turbulent heat fluxes have less impacts on the lake thermal properties. For some lakes, such as Kilpisjarvi and Toolik that were most obviously less affected by hwt , the stronger surface winds at the high latitudes would amplify the impact of wind-induced heat mixing and the corresponding parameters, in our model, $wstr$ and $ktscale$. The RI of $wstr$ that controls wind-induced turbulent mixing at lake surface was affected by lake depth in the full-profile temperature prediction due to the limited impact of surface energy transfer on the water column at deeper layers and thus the thermal regime of the whole lake.

3.2.2. Lake Stability

The modeled center of buoyancy and Schmidt stability were most sensitive to hwt , $wstr$, and $feta$ for most of the lakes, which is consistent with the modeled full-profile water temperature (Figure 6). The average RI values of these three parameters in center of buoyancy (N) simulations were 11.18, 15.75, and 68.65, respectively, and the values were 41.35, 30.69, and 24.03, respectively, in SSI simulations. The influence of $feta$ was slightly less for lakes with higher shape factors, that is, flatter shapes, but robustly stronger for deeper lakes while it is the opposite for $wstr$ (Figure 7). The parameter $feta$ controls the solar radiation penetration in the water column and the vertical temperature gradient that defines the thermal stratification of the lake. Shallower and larger lakes often experience more frequent mixing events that prevent stratification (Gorham & Boyce, 1989), so they usually have lower Schmidt stability values and are more strongly influenced by wind-induced turbulence. In contrast, the influence of wind-induced turbulence is smaller while the influence of thermal stratification is larger for the mixing of deeper lakes. These explain the correlation of $feta$ and $wstr$ to lake bathymetry. Parameters related to lake sediment properties, including c_{ps} and ρ_s , played a

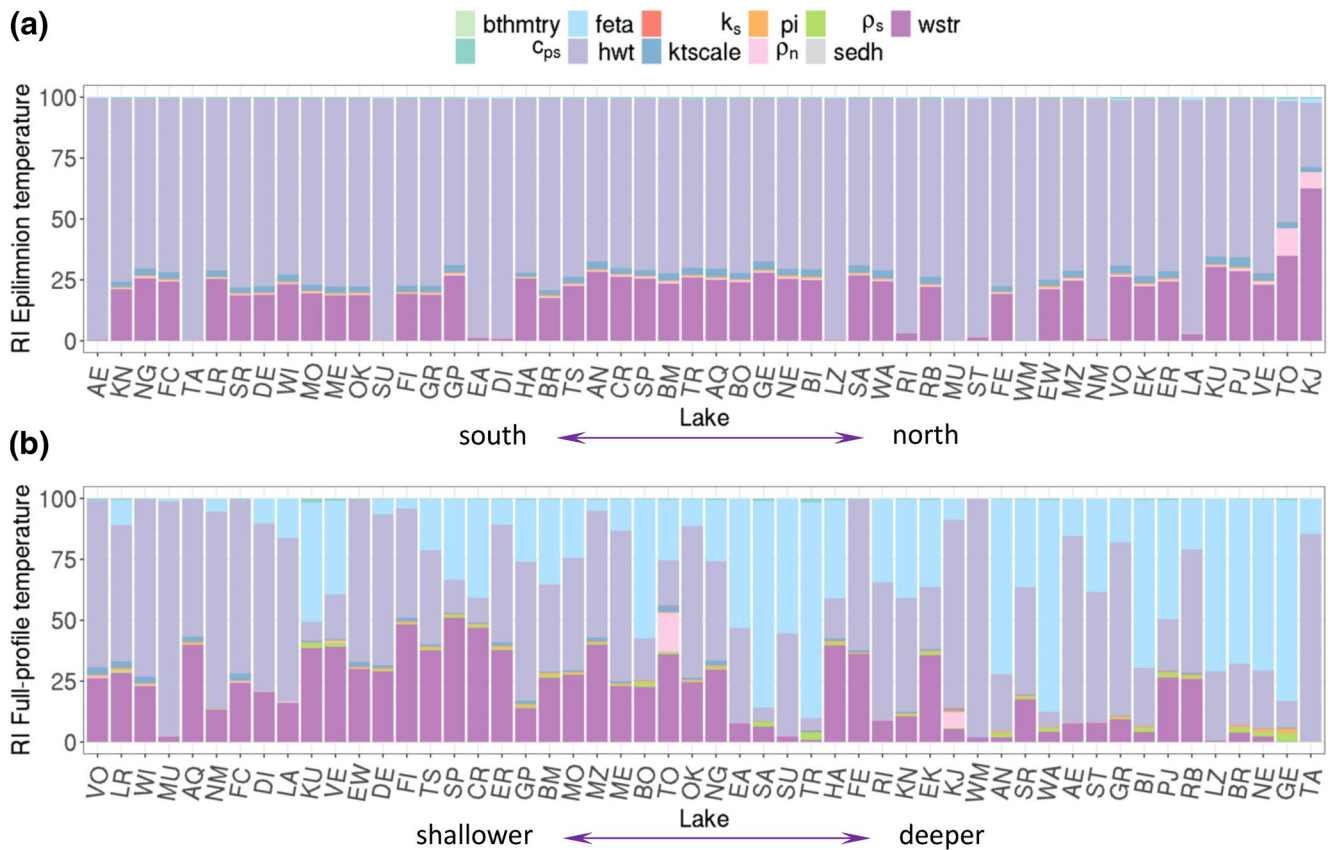


Figure 4. Parameter relative importance (RI) values for the modeled lake epilimnion temperature (a) and full-profile temperature (b). Lakes are arranged in the order of the latitude in (a) and maximum depth in (b).

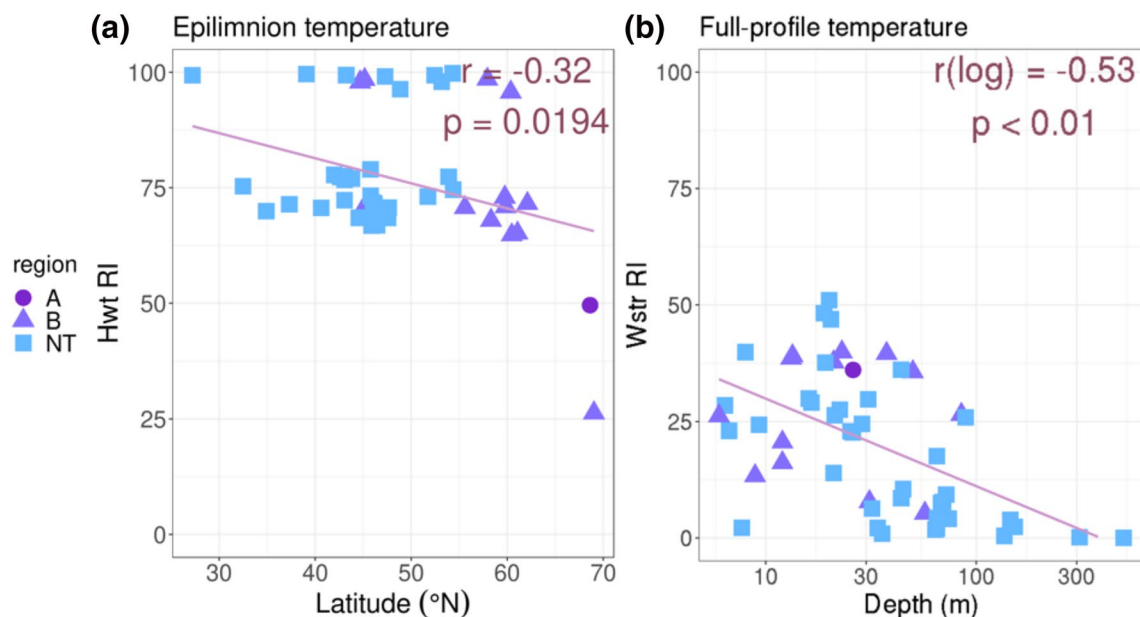


Figure 5. Correlation between RI values of hwt and lake latitudes in lake epilimnion temperature simulations (a) and correlation between RI values of $wstr$ and logarithmic lake depths in full-profile temperature simulations (b). The lines are regression fits. RI, relative importance.

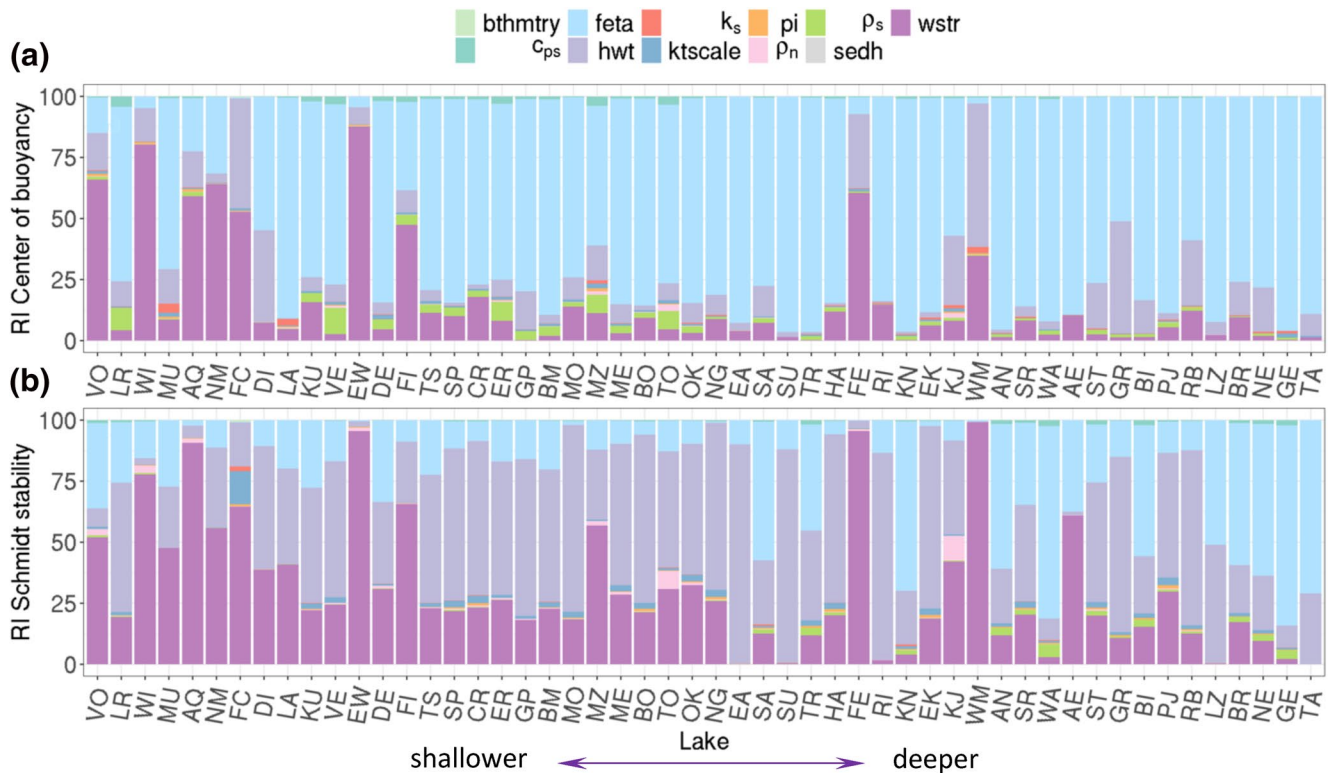


Figure 6. Same as Figure 4 but for center of buoyancy (a) and Schmidt stability (b). Lakes are arranged in the order of maximum depth.

slightly more important role in the prediction of lake stratification state than the temperature profile because of their modulating effects on the lake vertical temperature gradient.

3.2.3. Lake Ice Phenology

Similar to the epilimnion temperature, *hwt* and *wstr* were the most important parameters in the prediction of the ice-on date with RI values of 61.59 and 32.3, respectively (Figure 8), because the water temperature at lake surface directly determines ice-cover initiation. For the ice-off day and ice duration, apart from *hwt*

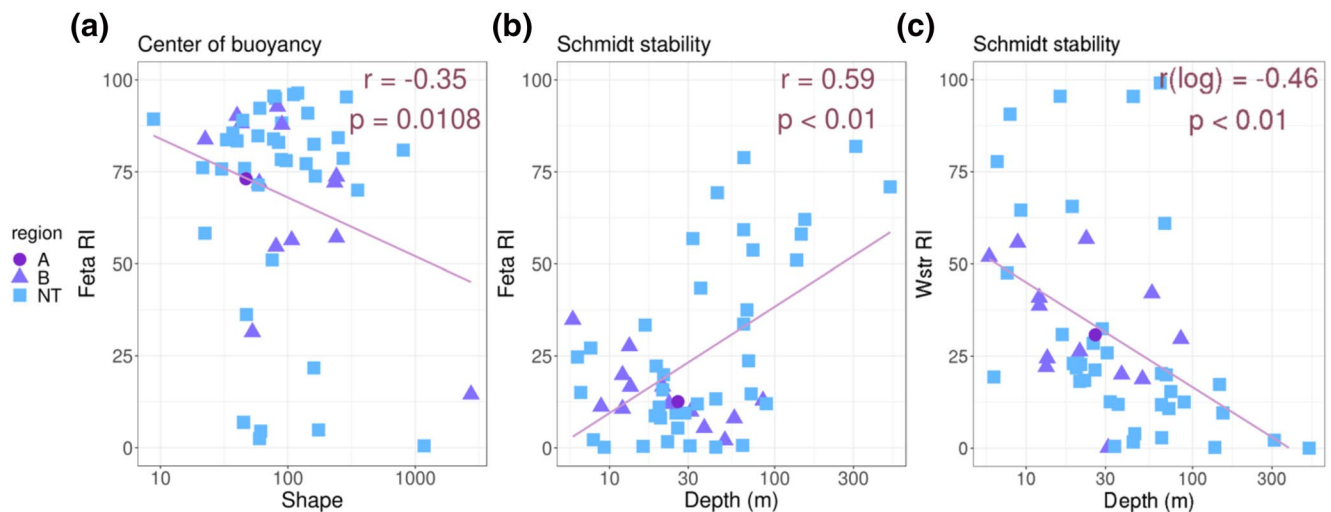


Figure 7. Same as Figure 5 but for correlation between *feta* and lake shape factors (a) in center of buoyancy (N) simulations, *feta* and depths (b) and *wstr* and depths (c) in Schmidt stability (SSI) simulations.

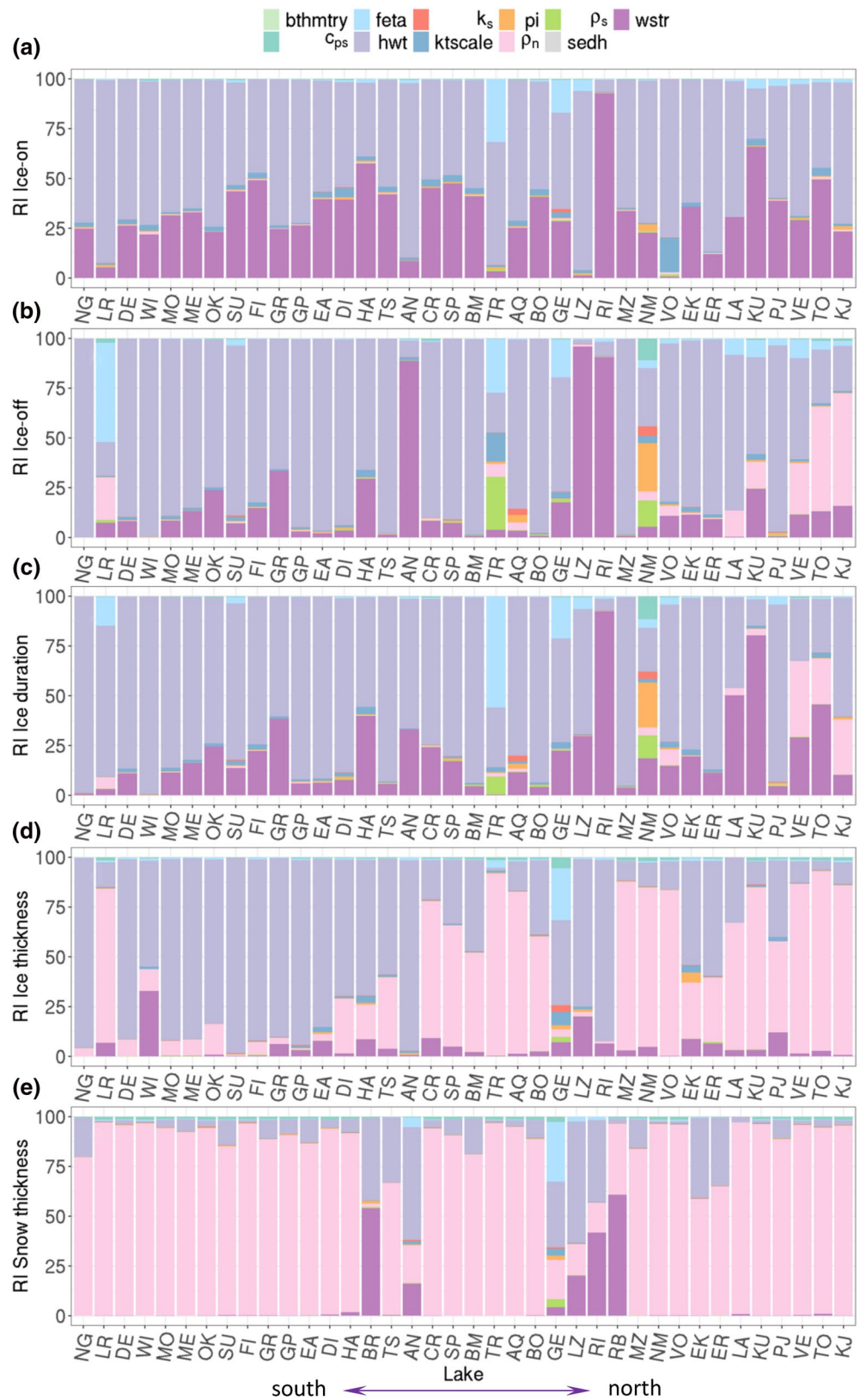


Figure 8. Same as Figure 4 but for ice-on day (a), ice-off day (b), ice duration (c), ice thickness (d), and snow thickness (e). Lakes are arranged in the order of latitude.

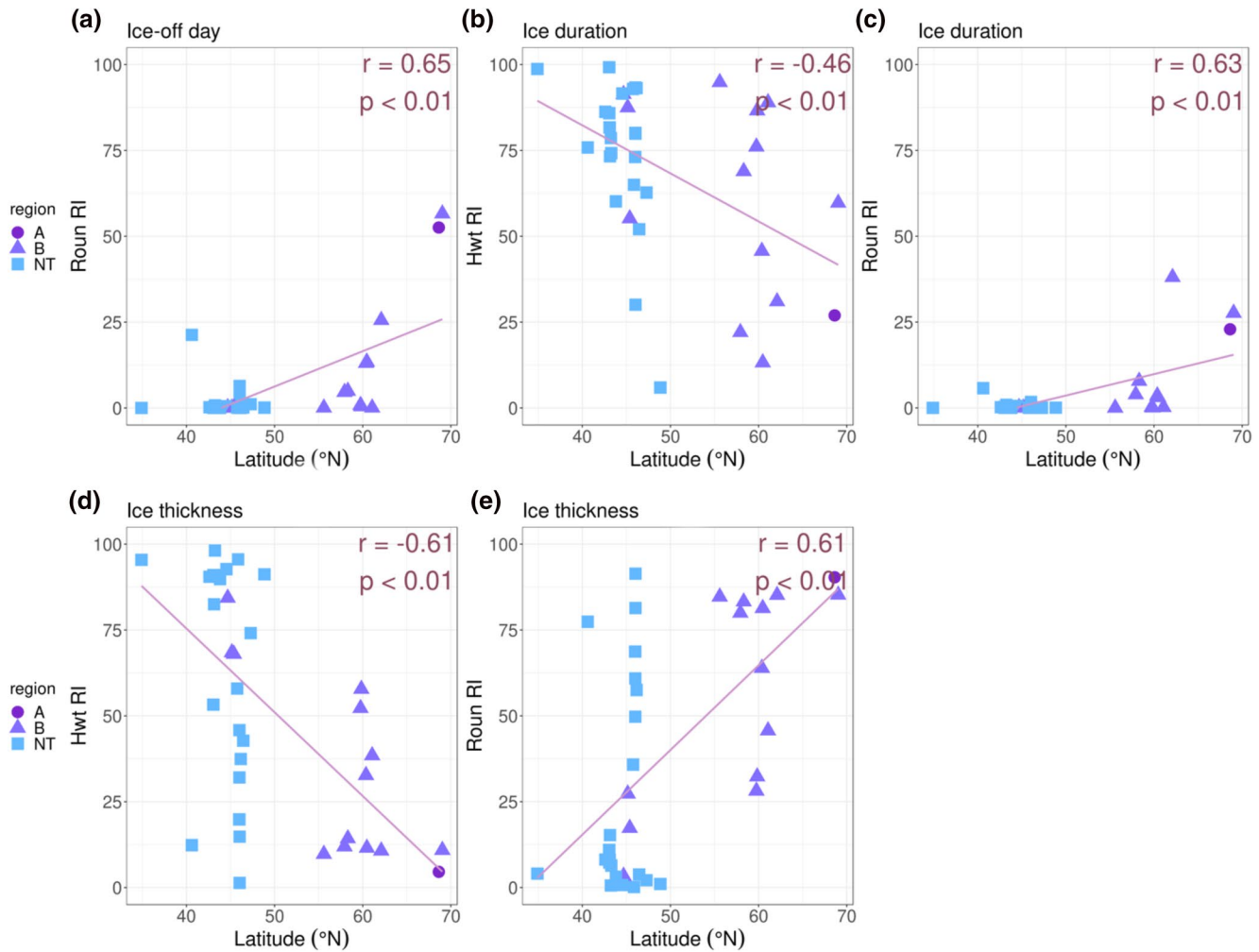


Figure 9. Same as Figure 5 but for ρ_n and lake latitudes in ice-off day simulations (a), hwt and lake latitudes (b) and ρ_n and lake latitudes (c) in ice duration simulations, hwt and the latitude (d), and ρ_n and the latitude (e) in ice-cover thickness simulations.

(mean RI = 69.07 and 68.93) and $wstr$ (mean RI = 16.06 and 20.39), the importance of ρ_n (mean RI = 5.79 and 3.31) became more evident for some northern lakes with RI as high as 56.57. The decreasing RI of hwt with latitude for ice duration and ice-cover thickness (Figure 9) was primarily because surface turbulent heat fluxes are small in ice-over periods. Lakes at high latitudes are usually frozen in winter and affected by snow, which easily explains the increase of the RI values of ρ_n with latitude in the ice-off date, ice-cover duration, and ice-over thickness simulations.

4. Discussion

4.1. Lake Characteristic Impacts on Model Accuracy

Overall, the model accuracy shows a dependence on the size of the lake rather than the location or elevation. This is why even though the ALBM was developed for Arctic lakes, it can be confidently applied to lakes in different climate zones. The investigation of three different error assessment metrics indicates that the accuracy of full-profile temperature simulations is enhanced with increasing lake depth, consistent with the result using the GLM model (Bruce et al., 2018). Here, we present the by-layer temperature simulation results of eight lakes of diverse lake properties (Figure S6) as examples of the model performance in simulating vertical temperature profiles and annual and interannual thermal evolution. Except for the tropical and shallow (<10 m) lakes, the water temperature below the thermocline hardly changes over

the time, which reduces the complexity in simulating full-profile temperature. The model underestimated deeper-layer temperature in Lake FallingCreek which is the largest in the shape factor, consistent with the Wedderburn number criteria. The Wedderburn number (Patterson et al., 1984) is the ratio of the wind friction to the pressure gradient established by the lake stratification and can be used as the criteria considering the lake as 1-D or not for modeling. According to the definition, for lakes of the same area, those with larger thermocline depths would have larger Wedderburn number values and thus can be more reasonably represented by 1-D models. The fact that deeper lakes usually have deeper thermocline depths also explains the better model performance in full-profile temperature simulations. It is also found that the model performance may deteriorate for lakes with larger surface areas in simulating epilimnion temperatures because the 1-D model assumes a horizontally homogeneous condition whereas, for larger lakes, horizontal and 3-D mixing become more important in heat circulation (particularly for the middle and bottom layers), resulting in simulation biases. The decrease of lake stability also has a slightly negative influence on the simulation accuracy of full-profile temperature, which is consistent with the negative relationship between the Wedderburn number and the lake turbulence.

For the lakes with relatively poor performance, all except Lake Rotorua were discontinuous cold polymictic or dimictic lakes with moderate depths ranging from 8 to 35.7 m. Rotorua is a discontinuous warm polymictic lake with maximum depth of 52.9 m. Some of these lakes (i.e., BigMuskellunge, Eagle, Rotorua, and Trout) have very irregular shapes with relatively large islands and peninsulas dividing the whole water bodies. Rotorua has several connecting rivers and streams with large volume of water flow, which was not represented in the simulation due to lack of measurements. Also, depending on the dominant wind direction, there can be seiches in some lakes that cannot be resolved in the 1-D model. There was underestimation of water temperature at mid-to-bottom layers for most of these lakes (Figures S6c and S7), reflecting the possibly unresolved seiches and the inefficiency of heat transfer from the water surface downward. Also, using the daily average wind speed as forcing can lead to underestimation of wind-driven mixing due to the smoothed wind speed. Despite the varying model performances, the RMSEs of all lakes are within an acceptable range of [0.6, 2.92] and [0.32, 2.38]°C for epilimnion and full-profile temperatures, respectively. Therefore, we conclude that the ALBM is also appropriate for modeling lakes of different sizes and mixing regimes. Possible model development strategies for the lakes with relatively poor model performances are discussed later.

4.2. Key Processes and Parameters, and Calibration Strategy for Global Applications

The lake thermal structure is dominantly determined by the parameters related to surface processes, including surface heat and momentum exchange, mixing efficiency, light extinction, and snow density, while the impact of the parameters related to lake sediment processes is minor. In ALBM, the corresponding control parameters are scaling factor for turbulent heat transfer hwt , wind shielding factor of mixing $wstr$, light attenuation correction factor $feta$, and snow density ρ_n . The sum of these parameters is responsible for over 95% of the variations in the modeled lake temperature structure, stability and ice phenology for lakes ranging broadly in size, climate zone, and mixing regime. The relative importance of each process and parameter varies with different regions and lake depths. The four key parameters in the ALBM can be divided into two types based on the RI values: the latitude-dependent parameters hwt and ρ_n , and the lake depth-dependent ones $wstr$ and $feta$.

The parameterization of turbulent heat fluxes has already been identified and discussed by local studies as the main source of uncertainties in the modeling of lake temperature (Charusombat et al., 2018). Here our study reemphasizes that this parameterization has huge impacts on the evolution of lake thermal structures globally and thus should be the focus of future model improvement. We further demonstrate that for both lake temperature and ice phenology modeling, hwt has a decreasing impact with increasing latitudes northward.

In the simulation of water temperature and thermal stratification, $wstr$ is found to have less impacts on deeper lakes. However, Bruce et al. (2018) investigated 32 lakes globally using GLM and concluded that the sensitivity of modeled thermocline depth to parameters controlling wind-driven mixing increases for deeper lakes, which is opposite to our result. This is probably because in Bruce et al. (2018) the sensitivity indices

of these parameters were generally around zero for most of the lakes with only five to six lakes showing values obviously larger than zero. The largely uneven distribution of sensitivity index values would bias the statistical calculation and thus the generalization may not be representative.

The light extinction coefficient parameterizes water transparency and directly determines the depth to which solar radiation penetrates in the water column. By applying two 1-D models, LAKE and FLake, to a boreal lake, Heiskanen et al. (2015) found that the simulated thermal stratification is highly sensitive to the light extinction coefficient, consistent with our result. They further concluded that the parameter is more important and thus requires lake-specific calibration for clear lakes compared to relatively humic lakes. However, the trophic state and lake color information were not available for some of the lakes in our study and therefore, we did not investigate the correlation of lake chemical and optical characteristics with *feta*. Instead, we use more accessible lake physical characteristics and demonstrate that the importance of *feta* decreases with lake shape factor and increases with depth.

It is logical that ρ_n plays an important role in the prediction of ice-related variables for high-latitude lakes. Snow density modulates the lake ice-cover thickness and duration, which strongly affects local hydrology and climate in particular for the climate-sensitive northern regions. Snow cover impacts the underlying ice in several ways (Duguay et al., 2003; Vavrus et al., 1996; Yao & Fu, 2019). First, the thermal insulation effect of snow cover delays lake ice growth and melting in winter and spring, respectively. Lower snow density would lead to a more porous snow cover structure and enhance the insulation effects. Also, snow cover increases surface albedo and the light extinction coefficient, both contributing to more persistent ice cover in spring. In addition, in converting snowfall provided by climate models to snow thickness, lower snow density also means larger snow thickness. Therefore, given the close coupling of snow and ice dynamics, the parameter ρ_n may need a more explicit parameterization (e.g., accounting for snow aging) instead of being treated as a constant to better model lake ice phenology.

The calibrated parameter values were quite variable between lakes and diversely distributed within the sampling ranges (Table 1, Figure S8). There is no significant correlation between parameter values and accessible lake characteristics, so the optimal parameter values are highly lake-dependent and model calibration is still necessary to obtain reliable model results. To minimize the recalibration effort in global lake simulations, one strategy is to conduct calibration by dividing lakes into groups according to the controlling parameters. Within each subdivision, we can then aim at achieving convergence of the major parameters, lower the requirement on constraining the value ranges of the secondary ones, and safely set the other parameters to default values. Based on our findings, generally for *hwt*, the calibrated value should be more constrained for temperate lakes compared to Arctic lakes. The parameter ρ_n , on the contrary, requires relatively high accuracy for the Arctic and sub-Arctic lakes and can be set to default for lakes that only freeze for a short period annually or rarely freeze. The opposite correlations of *wstr* and *feta* to lake depth indicate that there should be different calibration emphases within different depth subdivisions. Under this framework, the detailed grouping criteria could be adjusted according to the number of the lakes and the value distribution of the two lake properties.

4.3. Data Availability Impacts

Our data set covers lakes over broad ranges in latitude and climate zones, but the observation sites tend to be denser at midlatitudes and temperate zones than the tropical and Arctic regions, resulting in possible biases in the analysis. An important thing to note here is that although the conclusion on the key parameters was valid for all the lakes, we noticed that lake location may affect the correlation between parameter sensitivities and lake characteristics. Including lakes in the Southern Hemisphere and near the equator (hereafter, southern lakes) seemed to reduce the correlation and statistical significance. To compare the correlation coefficients, we applied the Fisher's z transformation and confirmed that the differences of the r values were statistically insignificant between groups with and without the southern lakes for all the correlations shown in Figures 5 and 7 (Table S3). Due to the underrepresentation of southern lakes in the data set, the differences in the parameter sensitivity patterns from temperate and Arctic lakes may be overwhelmed in such bulk analysis. As a result, we excluded southern lakes in the correlation analysis and the calibration strategy discussion. The nonuniform distribution of lake latitude can also obscure the correlation between lake

characteristics and parameter relative importance. For example, it has been found that lake geographical features (latitude, longitude, and altitude) may have significant impacts on the lake trophic level (Faithful et al., 2011; Wood et al., 2017). However, the relationship does not appear in our analysis. It indicates that a more complete analysis would require further collection of lake data in the underrepresented regions.

4.4. Model Improvement

The relatively large biases in the simulation of some lakes require future model improvement. The discrepancies for intermediate water layers and the increasing RMSEs for lakes with weaker stratifications all indicated that the unresolved horizontal and 3-D mixing processes can lead to inefficiency in simulating heat transfer and mixing and thus biased vertical thermal structures (Fang & Stefan, 1996). As a solution, adding an enhanced diffusion term has been proposed and supported (Huang et al., 2019; Subin et al., 2012). However, according to our result, $wstr$ is a much more important parameter than $ktscale$, meaning that adjusting the wind-driven mixing process has a dominant effect on compensating for the underestimated mixing. Therefore, for models with similar configurations to ALBM, scaling eddy diffusion will likely lead to marginal improvement in simulating lake turbulent mixing. Moreover, this approach enhances mixing across all water columns equally, which may result in undesired overestimation of mixing below the lake boundary layer. At present, there is no consensus on the best lake turbulent mixing scheme, but our results indicate its importance in correctly simulating lake vertical temperature profiles and therefore, is a process that could be targeted for future improvement. The ISIMIP2a protocol provides a unique opportunity for a consistent evaluation of an ensemble of models. The intercomparison with different parameterizations will help selecting a better solution to this demand.

Furthermore, the light extinction coefficient has also been found to have distinct effects on the thermocline depth and thus the vertical thermal structure (Heiskanen et al., 2015; Zolfaghari et al., 2017). Shatwell et al. (2016) found that lakes with moderate depths and relatively high light extinction can switch between polymictic and dimictic states and thus experience seasonal shifts in water transparency and mixing regime due to fluctuations of phytoplankton biomass. As a result, giving greater weight to $feta$ in spring and summer for lakes deeper than 5 m would be expected to lead to improved calibration and model performance.

5. Conclusions

We conducted a study on the validation and sensitivity analysis of a 1-D lake model ALBM using an unprecedentedly diverse data set. In the assessment of RMSE, r , and PRE values, the model showed an overall good performance at all 58 lakes. The simulation biases were mainly related to lake size but not the location. Therefore, we concluded that the ALBM is robust for applications on a global scale. We further investigated the key processes and the associated parameters that determine model performance. For the prediction of lake water temperature, stratification, and ice phenology, surface thermodynamic processes have the most dominant contribution to the simulation results for most of the lakes while lakes sediment-related processes only played a negligible role. The turbulent heat flux process was found to be the most crucial, followed by wind-driven convective mixing, light extinction, and for northern lakes, snow density. These key processes regulate how lakes respond to climate change and thus should be of primary concern in the modeling practice. Parameters related to these processes were together responsible for over 95% of the variations in the predicted variables mentioned above. The relative influence of each parameter varied mainly with the lake depth and latitude. The scaling factors of bulk coefficients for turbulent latent/sensible heat transfer hwt and snow density ρ_n were found to be mainly latitude-sensitive while the wind shielding factor of mixing $wstr$ and light extinction attenuation factor $feta$ are lake depth-sensitive. Taking the patterns into consideration, modelers can largely reduce the calibration complexity in global lake simulations by grouping lakes based on the two properties and implement different calibration procedures accordingly. Due to the underrepresentation of southern lakes in our data set, whether the patterns are appropriate for lakes in the tropical and the Southern Hemisphere remains to be verified, highlighting the need for more long-term observation data for these regions. Further improving the turbulence mixing scheme and incorporating the seasonality of the light extinction coefficient in the calculation can also benefit the simulation accuracy and

adapt the model to lakes with even more diverse characteristics. Our study provides a direction to the improvement of the ALBM and other 1-D process-based lake models across diverse lake ecosystems, allowing for improved predictions of climate impacts on the physical, chemical, and biological processes in global lakes. Furthermore, the more accurate representation of lake thermal regimes will in return enhance the performance of weather and climate forecast systems from regional to global scales.

Data Availability Statement

The model output and codes for data analysis in this study are available at <https://purr.purdue.edu/publications/3529/1>.

Acknowledgments

This study is supported through a projected funded to Qianlai Zhuang by NASA (NNX17AK20G) and a project from the United States Geological Survey (G17AC00276). Zeli Tan and L. Ruby Leung are supported by the US DOE's Earth System Modeling program through the Energy Exascale Earth System Model (E3SM) project. The Pacific Northwest National Laboratory is operated by Battelle for the US Department of Energy under Contract DE-AC05-76RLO1830. The authors thank Wim Thiery and Rafael Marce for coordinating the ISIMIP lake sector. The supercomputing is provided by the Rosen Center for Advanced Computing at Purdue University.

References

Adrian, R., O'Reilly, C. M., Zagarese, H., Baines, S. B., Hessen, D. O., Keller, W., et al. (2009). Lakes as sentinels of climate change. *Limnology & Oceanography*, 54(6), 2283–2297. https://doi.org/10.4319/lo.2009.54.6_part_2.2283

Balsamo, G., Salgado, R., Dutra, E., Boussetta, S., Stockdale, T., & Potes, M. (2012). On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model. *Tellus A: Dynamic Meteorology and Oceanography*, 64(1), 15829. <https://doi.org/10.3402/tellusa.v64i0.15829>

Bastviken, D., Tranvik, L. J., Downing, J. A., Crill, P. M., & Enrich-Prast, A. (2011). Freshwater methane emissions offset the continental carbon sink. *Science*, 331(6013), 50. <https://doi.org/10.1126/science.1196808>

Bates, G. T., Giorgi, F., & Hostetler, S. W. (1993). Toward the simulation of the effects of the Great Lakes on regional climate. *Monthly Weather Review*, 121(5), 1373–1387. [https://doi.org/10.1175/1520-0493\(1993\)121<1373:TTSOTE>2.0.CO;2](https://doi.org/10.1175/1520-0493(1993)121<1373:TTSOTE>2.0.CO;2)

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. Boca Raton, FL: Chapman & Hall/CRC press.

Bruce, L. C., Frassl, M. A., Arhonditsis, G. B., Gal, G., Hamilton, D. P., Hanson, P. C., et al. (2018). A multi-lake comparative analysis of the General Lake Model (GLM): Stress-testing across a global observatory network. *Environmental Modelling & Software*, 102, 274–291. <https://doi.org/10.1016/j.envsoft.2017.11.016>

Carpenter, S. R., Stanley, E. H., & Vander Zanden, M. J. (2011). State of the world's freshwater ecosystems: Physical, chemical, and biological changes. *Annual Review of Environment and Resources*, 36, 75–99. <https://doi.org/10.1146/annurev-environ-021810-094524>

Charusombat, U., Fujisaki-Manome, A., Gronewold, A. D., Lofgren, K., B. M., Anderson, E. J., Blanken, P., et al. (2018). Evaluating and improving modeled turbulent heat fluxes across the North American Great Lakes. *Hydrology and Earth System Sciences*, 22(10), 5559–5578. <https://doi.org/10.5194/hess-22-5559-2018>

Choubin, B., Zehtabian, G., Azareh, A., Rafiei-Sardooi, E., Sajedi-Hosseini, F., & Kişi, Ö. (2018). Precipitation forecasting using classification and regression trees (CART) model: a comparative study of different approaches. *Environmental Earth Sciences*, 77(8), 314. <https://doi.org/10.1007/s12665-018-7498-z>

Downing, J. A., Prairie, Y. T., Cole, J. J., Duarte, C. M., Tranvik, L. J., Striegl, R. G., et al. (2006). The global abundance and size distribution of lakes, ponds, and impoundments. *Limnology & Oceanography*, 51, 2388–2397. <https://doi.org/10.4319/lo.2006.51.5.2388>

Duguay, C. R., Flato, G. M., Jeffries, M. O., Ménard, P., Morris, K., & Rouse, W. R. (2003). Ice-cover variability on shallow lakes at high latitudes: model simulations and observations. *Hydrological Processes*, 17(17), 3465–3483. <https://doi.org/10.1002/hyp.1394>

Dutra, E., Stepanenko, V. M., Balsamo, G., Viterbo, P., Miranda, P. M., Mironov, D., & Schär, C. (2010). An offline study of the impact of lakes on the performance of the ECMWF surface scheme. *Boreal Environment Research*, 15(2), 100–112.

Faithfull, C. L., Bergström, A. K., & Vrede, T. (2011). Effects of nutrients and physical lake characteristics on bacterial and phytoplankton production: A meta-analysis. *Limnology & Oceanography*, 56(5), 1703–1713. <https://doi.org/10.4319/lo.2011.56.5.1703>

Fang, X., & Stefan, H. G. (1996). Long-term lake water temperature and ice cover simulations/measurements. *Cold Regions Science and Technology*, 24(3), 289–304. [https://doi.org/10.1016/0165-232X\(95\)00019-8](https://doi.org/10.1016/0165-232X(95)00019-8)

Frieler, K., Lange, S., Piontek, F., Reyer, C. P. O., Schewe, J., Warszawski, L., et al. (2017). Assessing the impacts of 1.5°C global warming—Simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). *Geoscientific Model Development*, 10, 4321–4345. <https://doi.org/10.5194/gmd-10-4321-2017>

Gorham, E., & Boyce, F. M. (1989). Influence of lake surface area and depth upon thermal stratification and the depth of the summer thermocline. *Journal of Great Lakes Research*, 15(2), 233–245. [https://doi.org/10.1016/S0380-1330\(89\)71479-9](https://doi.org/10.1016/S0380-1330(89)71479-9)

Greenwell, B., Boehmke, B., Cunningham, J., Developers, G. B. M., & Greenwell, M. B. (2019). *Package 'gbm'*. R Package Version, 2(5).

Gueymard, C. (1995). *SMARTS2, Simple model of the atmospheric radiative transfer of sunshine: Algorithms and performance assessment* (Rep. FSEC-PF-270-95). Cocoa, FL: Florida Solar Energy Center.

Guo, M., Zhuang, Q., Tan, Z., Shurpali, N., Juutinen, S., Kortelainen, P., & Martikainen, P. (2020). Rising methane emissions from boreal lakes due to increasing ice-free days. *Environmental Research Letters*, 15(6), 064008. <https://doi.org/10.1088/1748-9326/ab8254>

Guseva, S., Bleninger, T., Jöhnk, K., Polli, B. A., Tan, Z., Thiery, W., et al. (2020). Multimodel simulation of vertical gas transfer in a temperate lake. *Hydrology and Earth System Sciences*, 24(2), 697–715. <https://doi.org/10.5194/hess-24-697-2020>

Heiskanen, J. J., Mammarella, I., Ojala, A., Stepanenko, V., Erkkilä, K.-M., Miettinen, H., et al. (2015). Effects of water clarity on lake stratification and lake-atmosphere heat exchange. *Journal of Geophysical Research: Atmospheres*, 120, 7412–7428. <https://doi.org/10.1002/2014JD022938>

Hipsey, M. R., Bruce, L. C., Boon, C., Busch, B., Carey, C. C., Hamilton, D. P., et al. (2019). A General Lake Model (GLM 3.0) for linking with high-frequency sensor data from the Global Lake Ecological Observatory Dataset (GLEON). *Geoscientific Model Development*, 12, 473–523. <https://doi.org/10.5194/gmd-12-473-2019>

Holgerson, M., & Raymond, P. (2016). Large contribution to inland water CO₂ and CH₄ emissions from very small ponds. *Nature Geoscience*, 9, 222–226. <https://doi.org/10.1038/ngeo2654>

- Hostetler, S. W., & Bartlein, P. J. (1990). Simulation of lake evaporation with application to modeling lake level variations of Harney-Malheur Lake, Oregon. *Water Resources Research*, 26(10), 2603–2612.
- Hostetler, S. W., Bates, G. T., & Giorgi, F. (1993). Interactive coupling of a lake thermal model with a regional climate model. *Journal of Geophysical Research*, 98(D3), 5045–5057. <https://doi.org/10.1029/92JD02843>
- Huang, A., Lazhu, Wang, J., Dai, Y., Yang, K., Wei, N., et al. (2019). Evaluating and improving the performance of three 1-D Lake models in a large deep Lake of the central Tibetan Plateau. *Journal of Geophysical Research: Atmospheres*, 124, 3143–3167. <https://doi.org/10.1029/2018JD029610>
- King, P. W., Leduc, M. J., Sills, D. M., Donaldson, N. R., Hudak, D. R., Joe, P., & Murphy, B. P. (2003). Lake breezes in southern Ontario and their relation to tornado climatology. *Weather and Forecasting*, 18(5), 795–807. [https://doi.org/10.1175/1520-0434\(2003\)018<0795:LBISOA>2.0.CO;2](https://doi.org/10.1175/1520-0434(2003)018<0795:LBISOA>2.0.CO;2)
- Koehler, B., Landelius, T., Weyhenmeyer, G. A., Machida, N., & Tranvik, L. J. (2014). Sunlight-induced carbon dioxide emissions from inland waters. *Global Biogeochemical Cycles*, 28, 696–711. <https://doi.org/10.1002/2014GB004850>
- Lange, S. (2019). *EartH2Observe, WFDEI and ERA-interim data merged and bias-corrected for ISIMIP (EWEMBI)* (V. 1.1): GFZ Data Services. <https://doi.org/10.5880/pik.2019.004>
- Long, Z., Perrie, W., Gyakum, J., Caya, D., & Laprise, R. (2007). Northern lake impacts on local seasonal climate. *Journal of Hydrometeorology*, 8(4), 881–896. <https://doi.org/10.1175/JHM591.1>
- McDonald, C. P., Rover, J. A., Stets, E. G., & Striegl, R. G. (2012). The regional abundance and size distribution of lakes and reservoirs in the United States and implications for estimates of global lake extent. *Limnology & Oceanography*, 57, 597–606. <https://doi.org/10.4319/lo.2012.57.2.0597>
- Mironov, D. V. (2008). *Parameterization of lakes in numerical weather prediction: Description of a lake model* (p. 41). Offenbach am Main, Germany: DWD.
- O'Reilly, C. M., Sharma, S., Gray, D. K., Hampton, S. E., Read, J. S., Rowley, R. J., et al. (2015). Rapid and highly variable warming of lake surface waters around the globe. *Geophysical Research Letters*, 42(24), 10–773. <https://doi.org/10.1002/2015GL066235>
- Paltan, H., Dash, J., & Edwards, M. (2015). A refined mapping of Arctic lakes using Landsat imagery. *International Journal of Remote Sensing*, 36(23), 5970–5982. <https://doi.org/10.1080/01431161.2015.1110263>
- Patterson, J. C., Hamblin, P. F., & Imberger, J. (1984). Classification and dynamic simulation of the vertical density structure of lakes. *Limnology & Oceanography*, 4, 845–861. <https://doi.org/10.4319/lo.1984.29.4.0845>
- Prairie, Y. T., & del Giorgio, P. A. (2013). A new pathway of freshwater methane emissions and the putative importance of microbubbles. *Inland Waters*, 3(3), 311–320. <https://doi.org/10.5268/IW-3.3.542>
- Rodrigues, M., & de la Riva, J. (2014). An insight into machine-learning algorithms to model human-caused wildfire occurrence. *Environmental Modelling & Software*, 57, 192–201. <https://doi.org/10.1016/j.envsoft.2014.03.003>
- Saloranta, T. M., & Andersen, T. (2007). MyLake—A multi-year lake simulation model code suitable for uncertainty and sensitivity analysis simulations. *Ecological Modelling*, 207(1), 45–60. <https://doi.org/10.1016/j.ecolmodel.2007.03.018>
- Samuelsson, P., Kourzeneva, E., & Mironov, D. (2010). The impact of lakes on the European climate as stimulated by a regional climate model. *Boreal Environment Research*, 15(2), 113–129.
- Saunio, M., Bousquet, P., Poulter, B., Peregon, A., Ciais, P., Canadell, J. G., et al. (2016). The global methane budget: 2000–2012. *Earth System Science Data*, 8(2), 697–751. <https://doi.org/10.5194/essd-2016-25>
- Schmid, M., Hunziker, S., & Wüest, A. (2014). Lake surface temperatures in a changing climate: A global sensitivity analysis. *Climatic Change*, 124, 301–315. <https://doi.org/10.1007/s10584-014-1087-2>
- Shatwell, T., Adrian, R., & Kirillin, G. (2016). Planktonic events may cause polymictic-dimictic regime shifts in temperate lakes. *Scientific Reports*, 6, 24361. <https://doi.org/10.1038/srep24361>
- Sobol', I. M. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, 7(4), 86–112. [https://doi.org/10.1016/0041-5553\(67\)90144-9](https://doi.org/10.1016/0041-5553(67)90144-9)
- Stepanenko, V. M., Goyette, S., Martynov, A., Perroud, M., Fang, X., & Mironov, D. (2010). First steps of a lake model intercomparison Project: LakeMIP. *Boreal Environment Research*, 15, 191–202.
- Stepanenko, V. M., Mammarella, I., Ojala, A., Miettinen, H., Lykosov, V., & Vesala, T. (2016). LAKE 2.0: A model for temperature, methane, carbon dioxide and oxygen dynamics in lakes. *Geoscientific Model Development*, 9(5), 1977–2006. <https://doi.org/10.5194/gmd-9-1977-2016>
- Subin, Z. M., Riley, W. J., & Mironov, D. (2012). An improved lake model for climate simulations: Model structure, evaluation, and sensitivity analyses in CESM1. *Journal of Advances in Modeling Earth Systems*, 4(M02001). <https://doi.org/10.1029/2011ms000072>
- Tan, Z., Yao, H., & Zhuang, Q. (2018). A small temperate lake in the 21st century: Dynamics of water temperature, ice phenology, dissolved oxygen, and chlorophyll a. *Water Resources Research*, 54(7), 4681–4699. <https://doi.org/10.1029/2017WR022334>
- Tan, Z., & Zhuang, Q. (2015a). Arctic lakes are continuous methane sources to the atmosphere under warming conditions. *Environmental Research Letters*, 10(5), 054016. <https://doi.org/10.1088/1748-9326/10/5/054016>
- Tan, Z., & Zhuang, Q. (2015b). Methane emissions from pan-Arctic lakes during the 21st century: An analysis with process-based models of lake evolution and biogeochemistry. *Journal of Geophysical Research: Biogeosciences*, 120, 2641–2653. <https://doi.org/10.1002/2015JG003184>
- Tan, Z., Zhuang, Q., Shurpali, N. J., Marushchak, M. E., Biasi, C., Eugster, W., & Anthony, K. W. (2017). Modeling CO₂ emissions from Arctic lakes: Model development and site-level study. *Journal of Advances in Modeling Earth Systems*, 9, 2190–2213. <https://doi.org/10.1002/2017MS001028>
- Tan, Z., Zhuang, Q., & Walter Anthony, K. (2015). Modeling methane emissions from arctic lakes: Model development and site-level study. *Journal of Advances in Modeling Earth Systems*, 7(2), 459–483. <https://doi.org/10.1002/2014MS000344>
- Thiery, W., Davin, E. L., Seneviratne, S. I., Bedka, K., Lhermitte, S., & Van Lipzig, N. P. (2016). Hazardous thunderstorm intensification over Lake Victoria. *Nature Communications*, 7(1), 1–7. <https://doi.org/10.1038/ncomms12786>
- Vavrus, S. J., Wynne, R. H., & Foley, J. A. (1996). Measuring the sensitivity of southern Wisconsin lake ice to climate variations and lake depth using a numerical model. *Limnology & Oceanography*, 41, 822–831. <https://doi.org/10.4319/lo.1996.41.5.0822>
- Vanderkelen, I., van Lipzig, N. P. M., Lawrence, D. M., Droppers, B., Golub, M., Gosling, S. N., et al. (2020). Global heat uptake by inland waters. *Geophysical Research Letters*, 47, e2020GL087867. <https://doi.org/10.1029/2020GL087867>
- Wik, M., Varner, R. K., Anthony, K. W., MacIntyre, S., & Bastviken, D. (2016). Climate-sensitive northern lakes and ponds are critical components of methane release. *Nature Geoscience*, 9(2), 99–105. <https://doi.org/10.1038/ngeo2578>
- Winslow, L., Read, J., Woolway, R., Brenttrup, J., Leach, T., Zwart, J., & Collinge, D. (2019). *Package 'rLakeAnalyzer' R package version: 1.11.4.1*. Retrieved from <https://cran.r-project.org/web/packages/rLakeAnalyzer/>

- Wood, S. A., Maier, M. Y., Puddick, J., Pochon, X., Zaiko, A., Dietrich, D. R., & Hamilton, D. P. (2017). Trophic state and geographic gradients influence planktonic cyanobacterial diversity and distribution in New Zealand lakes. *FEMS Microbiology Ecology*, *93*(2), fiw234. <https://doi.org/10.1093/femsec/fiw234>
- Woolway, R. I., & Merchant, C. J. (2019). Worldwide alteration of lake mixing regimes in response to climate change. *Nature Geoscience*, *12*(4), 271–276. <https://doi.org/10.1038/s41561-019-0322-x>
- Woolway, R. I., Verburg, P., Lenters, J. D., Merchant, C. J., Hamilton, D. P., Brookes, J., et al. (2018). Geographic and temporal variations in turbulent heat loss from lakes: A global analysis across 45 lakes. *Limnology & Oceanography*, *63*, 2436–2449. <https://doi.org/10.1002/lno.10950>
- Yao, H., & Fu, C. (2019). Explaining the non-significant changes in ice-off date over six decades at Lake of Bays and Lake Nipissing, south-central Ontario. *International Journal of Environment and Climate Change*, *9*(1), 29–43. <https://doi.org/10.9734/ijecc/2019/v9i130095>
- Zhang, W., Du, Z., Zhang, D., Yu, S., & Hao, Y. (2016). Boosted regression tree model-based assessment of the impacts of meteorological drivers of hand, foot and mouth disease in Guangdong, China. *Science of the Total Environment*, *553*, 366–371. <https://doi.org/10.1016/j.scitotenv.2016.02.023>
- Zolfaghari, K., Duguay, C. R., & Kheyrollah Pour, H. (2017). Satellite-derived light extinction coefficient and its impact on thermal structure simulations in a 1-D lake model. *Hydrology and Earth System Sciences*, *21*, 377–391. <https://doi.org/10.5194/hess-21-377-2017>