Water Resources Research

Key Points:

- A set of optimal algorithms and parameterization strategies is provided to better simulate global lake thermal regimes with 1-D lake models
- More complicated algorithms can lead to larger biases and require more recalibration efforts for global application
- Combination of algorithms for lakes of different characteristics can optimize the model performance

Correspondence to:

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

Intercomparison of Thermal Regime Algorithms in 1-D Lake Models

Mingyang Guo[1] [ID]**, Qianlai Zhuang[1,2]** [ID]**, Huaxia Yao[3]** [ID]**, Malgorzata Golub[4]** [ID]**, L. Ruby Leung[5]** [ID]**, and Zeli Tan[5]** [ID]

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

Abstract Lakes are an important component of the global weather and climate system, but the modeling of their thermal regimes has shown large uncertainties due to the highly diverse lake properties and model configurations. Here, we evaluate the algorithms of four key lake thermal processes including turbulent heat fluxes, wind-driven mixing, light extinction, and snow density, using a highly diverse lake data set provided by the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) 2a lake sector. Algorithm codes are configured and run separately within the same parent model to rule out any interference from factors apart from the algorithms examined. Evaluations are based on both simulation accuracy and recalibration complexity for application to global lakes. For turbulent heat fluxes, the non-Monin–Obukhov similarity (MOS) based, more simplified algorithms perform better in predicting lake epilimnion temperatures and achieve high convergence in the values of the calibrated parameters. For wind-driven mixing, a two-algorithm strategy considering lake shape and season is suggested with the regular mixing algorithm used for spring and earlier summer and the mixing-enhanced algorithm for summer steady stratification and fall overturn periods. There are no evident differences in the simulated thermocline depths using different light extinction algorithms or the observation. Finally, for lake ice phenology, an optimal algorithm is decided for most northern lakes while the Arctic lakes require separate consideration. Our study provides highly practical guides for improving 1-D lake models and feasible parameterization strategies to better simulate global lake thermal regimes.

1. Introduction

Lakes modulate weather and climate at different scales, from regional to global (Bates et al., 1993; Brown & Duguay, 2010; Dutra et al., 2010; Rouse et al., 2005). Lake thermal regime is also a crucial control for lake biogeochemistry including greenhouse gas emissions (DelSontro et al., 2016; Matthews et al., 2020; Wik et al., 2016). With the increasing recognition of the importance of lakes in weather and climate prediction, the development of lake models has been advancing over the past two decades. Modelers have acknowledged the unprecedented diversity in model configurations and applications and emphasized the importance to exploit this diversity by model intercomparison and coupling to improve lake models (Janssen et al., 2015; Mooij et al., 2010; Trolle et al., 2012). One-dimensional lake models, with relatively low complexity and high flexibility, are commonly incorporated into numerical weather and climate prediction systems (MacKay et al., 2009; Mironov et al., 2010; Samuelsson et al., 2010; Subin et al., 2012). However, a comprehensive evaluation and comparison of 1-D lake models has not been done yet. Previous lake model intercomparison studies were often conducted at the model level (Guseva et al., 2020; Huang et al., 2019; Stepanenko et al., 2013; Yao et al., 2014), so how different parameterizations of individual processes affect the modeling results has not been investigated. The results were also inevitably affected by nonphysical factors such as differences in numerical solvers. Therefore, limited perspectives are available to guide the improvement of individual models. Also, these comparison studies were made on only one lake. Charusombat et al. (2018) examined five turbulent heat flux algorithms by isolating the related codes in each model but only focused on the Laurentian Great Lakes. Thus, whether the conclusions from these studies could be generalized to lakes in other regions or of different sizes remain to be verified.

To address these gaps in 1-D lake model evaluation and improvement, we investigate the lake thermal regime from an algorithm perspective using a highly diverse lake data set including 50 lakes varying in shape, landscape setting, and climate provided by the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) 2a lake sector. Testing on a diverse set of lakes is the only way to develop a robust parameterization of global lakes, for which there is no data product available for model calibration. Results from previous modeling studies (Bruce et al., 2018; Charusombat et al., 2018; Guo, Zhuang, Golub, Yao, Leung, Pierson, et al., 2020; Heiskanen et al., 2015) have shown that simulated lake temperature is highly sensitive to turbulent heat fluxes and wind-driven mixing while light extinction plays an important role in lake thermal stratification. Also, snow density was found to have strong impacts on the ice phenology of high-latitude lakes (Duguay et al., 2003; Guo, Zhuang, Golub, Yao, Leung, Pierson, et al., 2020; Vavrus et al., 1996). Therefore, we focus on the parameterization schemes currently used in different lake models for these four key processes. For a variable-controlling comparison, we extract the codes of various algorithms and run them individually within a parent model. The turbulent heat flux, wind-driven mixing, and light extinction algorithms are tested against the measured water temperatures while the snow density algorithms are tested against ice-off day observations. Algorithms are evaluated for both the simulation accuracy and the applicability to lakes of various characteristics.

We use the Arctic Lake Biogeochemical Model (ALBM) (Tan et al., 2015, 2017) as the framework to test the algorithms. The model has shown good performances in simulating lake thermal dynamics in previous studies and ISIMIP global simulations (Guo, Zhuang, Tan, et al., 2020; Guseva et al., 2020; Tan et al., 2018). We expect the conclusions drawn from this study are of universal reference value. For the simulation of eddy diffusion, all 1-D lake models mainly fall into two categories: the Hostetler-based models, such as ALBM, CLM4-LISS (Subin et al., 2012), MyLake (Saloranta & Andersen, 2007), Minlake (Fang & Stefan, 1996), and WRF-Lake (Gu et al., 2015), and k–ε models, including GLM (Hipsey et al., 2019), Simstrat (Goudsmit et al., 2002), and LAKE (Stepanenko et al., 2016). The S12 and K-profile parameterization (KPP) algorithms are only able to be incorporated into the Hostetler-based models and likewise, the k–ε related wind-driven mixing algorithms are only feasible for the k–ε models. Therefore, ALBM is representative for one type of models and the conclusion drawn from the comparison between S12 and KPP is transferable to all models within the same category. The parameterization of other major processes is independent of the eddy diffusion categories and thus the comparison of the algorithms has referential value for all 1-D lake models. In particular, the knowledge gained here is likely transferable to current and future ISIMIP lake models (e.g., CLM4_LISS, MyLake, GLM, Simstrat, and LAKE) because they are under similar configurations of atmospheric forcing, resolution, and lake bathymetry in ISIMIP simulations.

Through the experiments, we aim to show (1) capability differences of algorithms in modeling lake thermal regime, (2) whether an optimal algorithm for each key thermal process can be determined, and (3) the parameterization strategy in the global application of 1-D lake models.

2. Algorithms

In this section, we introduce the algorithms briefly and focus on their differences. The algorithms are denoted by the authors and the publication year if not otherwise named. The parameter descriptions, value ranges, and references for all the algorithms evaluated in this study are listed in Table 1. We selected algorithms that are currently used in different lake models as well as algorithms that have been proposed for performance enhancement (Text S1). Although some algorithms are not typically calibrated in applications, we still calibrated the empirical parameters for a fair comparison.

2.1. Turbulent Heat Flux

For turbulent heat flux, we tested eight algorithms that are currently used in different 1-D lake models. CW08 (Cole & Wells, 2008) is a highly parameterized algorithm, with the sensible heat flux (Q_H, W m^{-2}) and latent heat flux (Q_E, W m^{-2}) calculated by

$$Q_H = a_{CW}\beta U\left(\theta_w - \theta_a\right), \tag{1}$$

Table 1
Parameters and Value Ranges for Calibration in Algorithms for All Processes

Algorithm	Parameter	Unit	Description	Range	Reference
Turbulent heat flux					
PW79	hwt	–	Scaling factor of bulk transfer coefficient for turbulent heat fluxes	[0.5, 5]	Parkinson and Washington (1979)
LS87	α	–	Charnock number	[0.001, 0.028]	Liu and Schwab (1987)
C89	α	–	Charnock number	[0.001, 0.028]	Croley (1989)
HB90	hwt_a	–	Scaling factor of bulk transfer coefficient for turbulent heat fluxes	[0.5, 5]	Hostetler and Bartlein (1990)
	hwt_b	–		[−0.1, −0.01]	
Z98	z_L	–	Scaling factor of the roughness length scale for momentum	[0.2, 5]	Zeng et al. (1998)
J99	a_J	–	Empirical parameters for bulk transfer coefficient for wind stress	$[5.4 \times 10^{-4}, 1.25 \times 10^{-2}]$	Jordan et al. (1999), Hunke et al. (2015)
	b_J	–		$[2.84 \times 10^{-5}, 7.1 \times 10^{-4}]$	
	c_J	–		$[1.53 \times 10^{-5}, 3.82 \times 10^{-4}]$	
COARE	a_{CO}	–	Empirical parameters for Charnock number	$[3.4 \times 10^{-4}, 5.8 \times 10^{-3}]$	Fairall et al. (1996), Edson et al. (2013)
	b_{CO}	–		[−0.025, −0.001]	
CW08	a_{CW}	W m^{-2} mb^{-1} m^{-1} s	Empirical parameter for the Lake Hefner formula	[1.37, 34.27]	Cole and Wells (2008)
Wind-driven mixing					
S12	$wstr$	–	Wind shielding factor of mixing	[0.1,10]	Henderson-Sellers (1985), Subin et al. (2012)
KPP	a_J	–	Empirical parameters for bulk transfer coefficient for wind stress	$[5.4 \times 10^{-4}, 1.25 \times 10^{-2}]$	Large et al. (1994), Zhang et al. (2019)
	b_J	–		$[2.84 \times 10^{-5}, 7.1 \times 10^{-4}]$	
	c_J	–		$[1.53 \times 10^{-5}, 3.82 \times 10^{-4}]$	
Light extinction coefficient					
Obs	–	–	–	–	–
H95	–	–	–	–	Håkanson (1995)
S19	–	–	–	–	Shatwell et al. (2019)
H95c	$feta$	–	Scaling factor of light extinction coefficient	[0.1, 10]	–
Snow density					
Calib	$roun$	kg m^{-3}	Snow density	[100, 400]	–
Y81	$C1$	m s^{-1}	Empirical parameters for snow compression	$[7.22 \times 10^{-4}, 2.5 \times 10^{-3}]$	Yen (1981)
	$C2$	m^3 kg^{-1}		$[4.2 \times 10^{-3}, 1.05 \times 10^{-1}]$	
H03	$\Gamma_{\rho s}$	kg m^{-4}	Empirical parameter for snow density	[100, 400]	Heise et al. (2003)
	$\Gamma_{\kappa s}$	W m^{-2} K^{-1}	Empirical parameter for snow heat conductivity	[0.2, 1.5]	
H19	f_R	–	Scaling factor of rainfall	[0.2, 5]	Hipsey et al. (2019)
	f_S	–	Scaling factor of snowfall	[0.2, 5]	

Note. Except for parameters used in ALBM and γ, other parameter values are perturbed between 0.2 and 5 times of the default values in the original algorithm.

$$Q_E = a_{CW}U\left(e_w - e_a\right), \tag{2}$$

where a_{CW} (W m^{-3}mb^{-1} s) is an empirical parameter used in the Lake Hefner formula, $\beta = 62.66$ mb K^{-1} is Bowen's coefficient, U (m s^{-1}) is the surface wind speed, θ_w and θ_a are the potential temperatures (K) of surface water and air, respectively, and e_w and e_a are the saturation vapor pressure (mb).

Turbulent heat flux in all other algorithms are determined by

$$Q_H = \rho_a c_p C_H S(\theta_w - \theta_a),$$
(3)

$$Q_E = \rho_a L_E C_E S(q_w - q_a),$$
(4)

where ρ_a is the air density (kg m^{-3}), c_p is the specific heat of air (J kg^{-1} K^{-1}), L_E is the latent heat of vaporization (J kg^{-1}), C_H and C_E are the bulk transfer coefficients of sensible and latent heat, respectively, $S = \sqrt{U^2 + G^2}$ is the total wind speed (m s^{-1}), G is the gustiness (m s^{-1}), and q_w and q_a are the specific humidity. The seven algorithms differ in the calculation of C_H and C_E, and S.

PW79 (Parkinson & Washington, 1979) is the algorithm currently used in ALBM which adopts a simple empirical approach:

$$C_H = hwt\, c_H,$$
(5)

$$C_E = hwt\, c_E,$$
(6)

where hwt is a scaling factor, $c_H = c_E = 1.75 \times 10^{-3}$.

HB90 (Hostetler & Bartlein, 1990) further relates the bulk transfer coefficients to lake surface area A_s, leading to

$$C_H = hwt_a\, c_H A_s^{hwt_b},$$
(7)

$$C_E = hwt_a\, c_E A_s^{hwt_b}.$$
(8)

Other algorithms adopt the Monin–Obukhov similarity (MOS) (Monin & Obukhov, 1954) method in which

$$C_H = \frac{\kappa}{Pr} \frac{1}{\ln\frac{z}{z_0} - \Psi_M(\xi)} \frac{1}{\ln\frac{z}{z_{0\theta}} - \Psi_\theta(\xi)},$$
(9)

$$C_E = \frac{\kappa}{Pr} \frac{1}{\ln\frac{z}{z_0} - \Psi_M(\xi)} \frac{1}{\ln\frac{z}{z_{0q}} - \Psi_q(\xi)},$$
(10)

where $\kappa = 0.41$ is the von Karman constant; $Pr = 1.0$ is the turbulent Prandtl number; z is the observation height (m); z_0, $z_{0\theta}$, and z_{0q} are the roughness lengths for momentum, temperature, and humidity, respectively; and Ψ_M, Ψ_θ, and Ψ_q are MOS profile functions assumed to be the same in all the algorithms. The MOS-based algorithms differ in the calculation of the roughness lengths. Except for the Coupled Ocean-Atmosphere Response Experiment (COARE) algorithm, $z_0 = z_{0\theta} = z_{0q}$ is assumed.

For LS87 (Liu & Schwab, 1987),

$$z_0 = \alpha \frac{u*^2}{g} + 0.11 \frac{\upsilon}{u^*},$$
(11)

where α is the Charnock number, $u^* = S C_D^{0.5}$ is the friction velocity (m s^{-1}), g is gravity (m s^{-2}), and υ is the kinematic viscosity (m^2 s^{-1}).

For C89 (Croley, 1989),

$$z_0 = \alpha \frac{u*^2}{g},$$
(12)

where u^* is the same as in LS87.

For Z98 (Zeng et al., 1998),

$$z_0 = 0.001 \cdot z_L, \tag{13}$$

where z_L is a scaling factor of the roughness length for momentum.

For J99 (Hunke et al., 2015; Jordan et al., 1999),

$$z_0 = z \exp\left(-\frac{\kappa}{\dfrac{a_J}{U} + b_J + c_J U}\right), \tag{14}$$

where a_J, b_J, and c_J are empirical parameters.

For COARE (Edson et al., 2013; Fairall et al., 1996),

$$z_0 = \alpha \frac{u^{*2}}{g} + 0.11 \frac{\upsilon}{u^*}, \tag{15}$$

$$z_{0\theta} = z_{0q} = \min\left(1.6 \times 10^{-4}, 5.8 \times 10^{-5} Rr^{-0.72}\right), \tag{16}$$

where $\alpha = a_{CO} \max(U, 19) - b_{CO}$ is a function of wind speed, υ is kinetic viscosity (m^2 s^{-1}), and $Rr = \frac{u^* z_0}{\upsilon}$ is the roughness Reynolds number.

Only Z98 and COARE consider nonzero gustiness velocity such that

$$G = \beta \frac{g}{\rho_a}\left(\frac{Q_H}{c_p T} + 0.61 \frac{Q_E}{L_E}\right), \tag{17}$$

where $\beta = 1.0$ in Z98 and 1.2 in COARE, and T is air temperature (K).

2.2. Wind-Driven Mixing

Two wind-driven mixing algorithms compatible to 1-D lake models were evaluated: S12 (Subin et al., 2012) which is highly parameterized and KPP (Large et al., 1994) which has been primarily used for ocean mixing simulation and was proposed by Zhang et al. (2019) for lake applications to improve the model efficiency in lake mixing. ALBM adopts the first algorithm. Both algorithms are based on the governing equation of water temperature (T_w) defined by

$$\frac{\partial T_w}{\partial t} = \frac{1}{A}\frac{\partial}{\partial z}\left(AK\frac{\partial T_w}{\partial z}\right) + R, \tag{18}$$

where t is the length of time step (s), z is the depth (m), A is lake cross-section area (m^2), $K = \kappa_m + \kappa_e$ is the total thermal diffusivity that consists of molecular diffusivity κ_m (m^2 s^{-1}) and wind-driven eddy diffusivity κ_e (m^2 s^{-1}), and R is a solar radiation term.

The algorithms differ in the calculation of the eddy diffusivity. In S12, $\kappa_m = 1.4 \times 10^{-7}$ m^2 s^{-1} is assumed to be a constant, and κ_e is defined by

$$\kappa_e = \frac{\kappa w^* z}{P_r \left(1 + 37 R_i^2\right)} e^{-k^* z}, \tag{19}$$

where w^* is the surface friction velocity (m s^{-1}), R_i is the Richardson number, and k^* is a latitude-dependent parameter.

In the KPP scheme, κ_e is considered separately for the lake boundary layer and the lake interior. In the lake boundary layer, κ_e follows the MOS and is calculated as

$$\kappa_e(\sigma) = h w(\sigma) G(\sigma), \tag{20}$$

where h is the boundary layer depth (m), $\sigma = d\,/\,h$ is a dimensionless vertical coordinate that varies from 0 to 1 within the boundary layer, $w(\sigma)$ is the velocity scale, and $G(\sigma)$ is a shape function. Variable $w(\sigma)$ is a function of the friction velocity (u^*) which is calculated using the same method as in J99 and therefore, the same parameters are included in the calibration. In the lake interior, κ_e is defined as

$$\kappa_e = k_s + k_w, \tag{21}$$

where k_s is the diffusivity due to shear instability (m^2 s^{-1}), and $k_w = 10^{-7}$ m^2 s^{-1} is the internal wave diffusivity. The calculation of k_s is related to the local gradient Richardson number. The equations for each term on the RHS of Equations 18 and 19 can be found in Large et al. (1994) and Zhang et al. (2019).

The convective mixing in ALBM is represented as the balance of wind-induced kinetic energy and stratification-induced potential energy (Saloranta & Andersen, 2007). The kinetic energy (F_{wstr}) is modified by the wind shielding coefficient: $F_{wstr} = wstr\left(1 - e^{-0.3 A_s}\right)$, where $wstr$ is a correction factor for irregular lakes and A_s is the lake surface area (km^2).

2.3. Light Extinction

We compared calculated light extinction coefficients (η, m^{-1}) generated by three algorithms with observed values and evaluated the corresponding model performance. The first algorithm was proposed by Håkanson (1995) and was derived from 88 Swedish glacial lakes:

$$\eta = 1.1925 \, d^{-0.424}, \tag{22}$$

where d is the maximum lake depth (m). The second algorithm proposed by Shatwell et al. (2019) was derived from a collection of 1,258 lakes:

$$\eta = 5.681 \, d^{-0.795}. \tag{23}$$

The third, which is used in ALBM, is a modified method of H95 by further applying a scaling factor *feta* to the equation (hereafter denoted as H95c):

$$\eta = feta \cdot 1.1925 \, d^{-0.424}. \tag{24}$$

2.4. Snow Density

Snow and ice processes can be complicated, involving formation of snow, gray ice, and white ice and the transformation between water, snow, and ice species. Here, we compared four algorithms of snow density (ρ_s, kg m^{-3}), snow heat conductivity (κ_s, W m^{-1} K^{-1}), and snow cover compaction (if considered) while kept the modeling of snow-ice transformation and snow melting the same as in ALBM.

The first algorithm, which is used in ALBM, directly calculates snow density in a simple way:

$$\rho_s = \text{roun}, \tag{25}$$

where roun is calibrated and does not change over time. Snow heat conductivity $\kappa_s = 0.27 \, \text{W m}^{-1} \, \text{K}^{-1}$ is also considered to be constant.

Y81 (Yen, 1981) simulates the densification of snow cover ($\Delta\rho_s$) by $\Delta\rho_s = \rho_s \, C_1 \, W_s e^{-C_2\rho_s} \, \Delta t$ where $Ws = H_{s,\text{new}} \dfrac{\rho_s}{\rho_{w,\text{max}}}$ is the water equivalent of new snow (m), $H_{s,\text{new}} = \text{snowfall} \cdot \Delta t$ is the height of new snow (m) accumulated during time Δt, $\rho_{w,\text{max}} = 1{,}000 \, \text{kg m}^{-3}$ is the maximum density of water, and $C1$ and $C2$ are empirical parameters. Then, the old snow cover thickness is updated to $H_s = H_s \dfrac{\rho_s}{\rho_s + \Delta\rho_s}$, and old snow density to $\rho_s = \rho_s + \Delta\rho_s$. Finally, the total snow density is achieved as the weighted average of new and old snow:

$$\rho_s = \frac{\rho_s \, H_s + \rho_{s,\text{new}} H_{s,\text{new}}}{H_s + H_{s,\text{new}}}, \tag{26}$$

$$\rho_s = \min\left(\rho_{s,\text{max}}, \rho_s\right), \tag{27}$$

where $\rho_{s,\text{new}} = 250 \, \text{kg m}^{-3}$ is the density of new snow, and $\rho_{s,\text{max}} = 400 \, \text{kg m}^{-3}$ is the maximum density of snow. Snow heat conductivity is treated as a constant that equals to $0.2 \, \text{W m}^{-1} \, \text{K}^{-1}$.

H03 (Heise et al., 2003) uses empirical approximations relating ρ_s and κ_s to the snow cover thickness:

$$\rho_s = \min\left(\rho_{s,\text{max}}, \frac{\rho_{s,\text{min}}}{\left| 1 - \dfrac{H_s \Gamma_{\rho_s}}{\rho_{w,\text{max}}} \right|} \right), \tag{28}$$

$$\kappa_s = \min\left(\kappa_{s,\text{max}}, \kappa_{s,\text{min}} + H_s \Gamma_{\kappa_s} \frac{\rho_s}{\rho_{w,\text{max}}} \right), \tag{29}$$

where Γ_{ρ_s} (kg m^{-4}) and Γ_{κ_s} ($\text{W m}^{-2} \, \text{K}^{-1}$) are empirical parameters, $\rho_{s,\text{max}} = 400 \, \text{kg m}^{-3}$ and $\rho_{s,\text{min}} = 100 \, \text{kg m}^{-3}$ are the maximum and minimum snow densities, respectively, and $\kappa_{s,\text{max}} = 1.5 \, \text{W m}^{-1} \, \text{K}^{-1}$ and $\kappa_{s,\text{min}} = 0.2 \, \text{W m}^{-1} \, \text{K}^{-1}$ are the maximum and minimum snow heat conductivities, respectively.

H19 (Hipsey et al., 2019) is the most complex algorithm that considers the impact of rainfall and snowfall on the compaction of snow cover under different air temperatures. The compaction rate (compact) is simulated as

$$\text{compact} = \begin{cases} 0.166 + 0.834 \cdot \left(1 - e^{-f_R \, \text{rainfall} \, \Delta t} \right), & T > 0°C \\[3mm] 0.088 + 0.912 \cdot \left(1 - e^{-\frac{f_R \, \text{rainfall}}{f_S \, \text{snowfall}}} \right), & \text{snowfall} > 0, \quad T \leq 0°C \\[3mm] 0, \quad \text{rainfall} > 0, \text{snowfall} = 0, & T \leq 0°C \end{cases} \tag{30}$$

where f_R and f_S are scaling factors of rainfall and snowfall rate, respectively. Based on the meteorological conditions, the impact of compaction on old snow density ($\rho_{s,\text{old}}$) is simulated as

$$\rho_{s,\text{old}} = \begin{cases} \rho_{s,\max}, \text{rainfall} > 0, \text{snowfall} = 0, & T \leq 0°C \\ \rho_s + (\rho_{s,\max} - \rho_s)\text{compact}, & \text{otherwise} \end{cases}, \tag{31}$$

where $\rho_{s,\max} = 300 \text{ kg m}^{-3}$, different from Y81. The snow cover height is then updated to $H_s = \frac{H_s \rho_s}{\rho_{s,\text{old}}}$. The final snow density at the current time step is calculated as

$$\rho_s = \begin{cases} \dfrac{\rho_{s,\text{old}} H_s + 1{,}000 \cdot f_R \text{ rainfall } \Delta t}{f_S \text{ snowfall } \Delta t + H_s}, & \text{rainfall} > 0, \text{snowfall} > 0 \\ \rho_{s,\max}, \quad \text{rainfall} > 0, \text{snowfall} = 0, & T \leq 0°C \\ \rho_{s,\text{old}}, & \text{otherwise} \end{cases}, \tag{32}$$

The snow heat conductivity is defined as an empirical function of snow density (Ashton, 1986):

$$\kappa_s = 0.021 + 0.0042\rho_s + 2.2 \times 10^{-9}\rho_s^3. \tag{33}$$

3. Data and Method

3.1. Data

We examined 50 lakes from the ISIMIP2a data set, including 1 Arctic, 13 boreal, and 36 northern temperate lakes (Table S1). The surface areas range from 0.14 to 611 km^2, maximum depths from 6.4 to 501 m, and elevations from 210 m below to 4,300 m above sea level. Bathymetry profiles and light extinction measurements are available for all lakes. Water temperature measurements were taken at multilayers at least monthly for over 5 years for most of the lakes. For snow density parameterization, we only investigated five lakes that are evidently affected by snow phenology (Guo, Zhuang, Golub, Yao, Leung, Pierson, et al., 2020) and have long-term ice season observations. The observations are from the Global Lake and River Ice Phenology Dataset (Benson et al., 2000, updated 2020) for lakes Kilpisjarvi, Kuivajarvi, and Trout, from the Toolik Field Station (Environmental Data Center Team, 2020) for Toolik Lake and Zdorovennov et al. (2013) for Lake Vendyurskoe. The ice period observations range from 11 to 32 years with ice-on day, ice-off day, and ice duration recorded annually.

Meteorological forcing data supplied by the ISIMIP2a bias-corrected climate input data (Frieler et al., 2017; Lange, 2019) were organized at daily time step from 1979 to 2016. The required variables include surface air temperature, surface minimum and maximum air temperature, surface air pressure, relative humidity, surface wind speed, precipitation, snowfall, surface downward shortwave radiation, and surface downward longwave radiation. Inflow and outflow observations were not available and therefore, we assumed static water levels for all lakes.

3.2. ALBM

The ALBM is a 1-D process-based lake biogeochemistry model that has been successfully applied to lake thermal dynamics simulations around the world (Guo, Zhuang, Golub, Yao, Leung, Pierson, et al., 2020; Guo, Zhuang, Tan, et al., 2020; Guseva et al., 2020; Tan et al., 2018). The model uses the fourth-order adaptive Runge-Kutta scheme and an adjustable time step that varies with the numerical solver convergence rate to avoid instability. For spatial representation, the water column is divided into 50 vertical layers following an exponential distribution with the layer thickness increasing from the top to the bottom to resolve complicated processes at the water surface, and the sediment is divided into 10 layers equally. The mixing and convection of momentum, energy, and mass between adjacent layers are explicitly simulated.

Table 2
Experiment Setup

Test group	Turbulent heat flux	Wind-driven mixing	Light extinction	Snow density (kg m^{-3})
Turbulent heat flux	–	S12	Observation	200
Wind-driven mixing	Calibrated	–	Observation	200
Light extinction	Calibrated	S12	–	200
Snow density	Calibrated	S12	Observation	–

Note. The best-performing turbulent heat flux algorithm is selected and calibrated first and was used for the simulations of other groups.

3.3. Experiment Setup

One main objective of this study is to isolate the impact of each algorithm and avoid irrelevant influence factors. Therefore, algorithm codes were incorporated into the parent model and run individually while the algorithm groups not being tested were set to a default setup (Table 2). The turbulent heat flux algorithm group was run first to determine the best-performing algorithm and then this calibrated algorithm was used as the default algorithm when evaluating other groups. For the other algorithms, S12 *wstr* = 1, the observed light extinction coefficients, and ρ_s = 200 kg m^{-3} were used as the default setup. For the wind-driven mixing group, we first calibrated *wstr* for each lake under the S12 algorithm and then ran KPP simulations with the calibrated *wstr* values. For calibration, Monte Carlo simulations were run with perturbed parameter ensembles (PPEs). The dimension of the PPE equaled 1,000 times the number of parameters in each algorithm.

The algorithms were assessed by two criteria: the simulation accuracy and the applicability to lakes of diverse characteristics. The simulation accuracy was evaluated by the error metrics of modeled variables compared to measured values. For different algorithm groups, different reference variables along with different error metrics were considered. In total, six error metrics were calculated for algorithm evaluation, including the root mean square errors (RMSEs), the correlation coefficient (*r*), the mean bias error $\left(\text{MBE} = \frac{1}{n}\sum_{i=1}^{n}\left(P_i - O_i\right) \right)$, the mean absolute bias error $\left(\text{MABE} = \frac{1}{n}\sum_{i=1}^{n}\left|P_i - O_i\right| \right)$, the reliability index $\left(\text{RI} = \exp\left(\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{O_i}{P_i}\right)^2 \right) \right)$, and the Nash-Sutcliffe efficiency $\left(\text{NSE} = 1 - \frac{\sum_{i=1}^{n}\left(P_i - O_i\right)^2}{\sum_{i=1}^{n}\left(O_i - \bar{O}\right)^2} \right)$, where *n* is the number of observations, P_i and O_i are the *i*th prediction and observation, respectively, and $\bar{O}$ is the observation mean.

The algorithm applicability to global lakes was assessed by the distribution of parameter values in the individual and the ensemble calibrations of the lakes. For the individual calibration, we simply selected the optimum parameter value with the smallest simulation RMSE for each lake. For the ensemble calibration, we considered all the lakes together by calculating the mean RMSE for all lakes using each parameter set and selecting 20 parameter sets with the lowest values. The algorithm with more converged parameter values requires less recalibration complexity when applied to global lakes.

4. Results

4.1. Turbulent Heat Flux

For turbulent heat flux algorithms, the error metrics for the simulated epilimnion temperature were calculated (Table 3) and the full results are provided in Tables S4 and S5 for the individual and the ensemble calibration, respectively. For the individual calibration, the non-MOS based algorithms, PW79, HB90, and CW08, showed relatively better performances with overall higher NSE values compared to others (Figure 1a). Especially for the Arctic lake, Toolik, other algorithms all led to large biases with NSE below 0 and RMSE over 5°C while the NSEs are around 0.5 and RMSEs were around 3°C for the non-MOS based algorithms. The non-MOS based algorithms also performed better for large lakes such as Lake Tahoe with NSEs around 0.84 while the values using other algorithms were near or below 0. All algorithms led to a negative NSE for Lake Laramie because only six observations were available, biasing the calculation. All algorithms tended to underestimate the epilimnion temperature. Algorithm COARE showed the worst performance with an MBE of −1.7°C. There was an outlier of a boreal lake, Lake Eagle (Figure 1b), with highly underestimated epilimnion temperatures using all algorithms due to the very irregular shape of the lake with peninsulas dividing the water body. The optimal parameter values all varied largely within the sampling ranges (Figure 1c).

Table 3

All-Lake Mean Error Metrics: Epilimnion Temperature (°C) for Turbulent Heat Flux, Epilimnion Temperature (°C)/ Thermocline Depth (m) for Wind-Driven Mixing, Summer Thermocline Depth (m) for Light Extinction, and Ice-Off Day (Day) for Snow Density Algorithms

	Turbulent hear flux							
Algorithm	PW97	LS87	C89	HB90	Z98	J99	COARE	CW08
NSE	0.76	0.62	0.58	0.76	0.61	0.69	0.63	0.76
MBE	−0.76	−0.69	−0.13	−0.77	−0.13	−0.49	−1.7	−0.73
RMSE	2.91	3.53	3.61	2.91	3.38	3.17	3.62	2.92
	Wind-driven mixing			Light extinction			Snow density		
Algorithm	S12	KPP	Obs	H95	S19	H95c	Calib	Y81	H03	H19
NSE	0.77/−0.11	0.74/−0.23	−0.13	−0.11	−0.07	0	0.2	0.09	0.2	−0.52
MBE	−0.46/−11.2	−0.97/2.6	−1.16	−0.96	−1.13	−0.95	2.39	4.46	2.99	12.87
RMSE	2.68/13.1	2.87/12.5	3.92	3.97	3.95	3.9	8.38	8.35	7.96	14.59

For the ensemble calibration, the mean NSE for each algorithm decreased by about 0.1 and the RMSE increased by about 0.4°C (Tables S4 and S5). Overall, PW79, HB90, J99, and CW08 showed similar performances with mean NSEs above 0.6 and mean RMSEs below 3.5°C, better than other algorithms, and the accuracy in the Arctic lake was still the highest using PW79, HB90, and CW08 (Figure 2a). Interestingly, the *hwt* values in PW79 and the a_{CW} values in CW08 both achieved convergence at around 1.36 and 6.21, respectively (Figure 2c), indicating high applicability to global lakes with minimum recalibration effort. C89 also resulted in a highly converged Charnock number with the values concentrated at the upper boundary of the sampling range which is around 5 times the default values in the original model. Heiskanen et al. (2015) also found the calculated Charnock number from the measurements of a small boreal lake to be larger than that in the original parameterization. These findings suggest that the default Charnock number may not be appropriate for lakes with strong mixing events and should be calibrated for models adopting the MOS method in future applications. Except for J99, the parameters of all other MOS algorithms converged within 20% of the sample ranges but the relatively larger sacrifice of simulation accuracy indicates that they may not be applicable for modeling global lakes.

4.2. Wind-Driven Mixing

In the investigation of wind-driven mixing algorithms, the performances in both epilimnion temperature and thermocline depth simulations were evaluated and the full sets of calculated error metrics are provided in Tables S6–S9. In general, S12 outperformed KPP in epilimnion temperature simulations in the individual calibration. The mean NSE, MBE, and RMSE for individual epilimnion temperature calibrations were 0.74, −0.46°C, and 2.68°C using S12, respectively, and 0.77, −0.97°C, and 2.87°C using KPP, respectively (Table 3 and Figure 3a). Algorithm KPP resulted in larger underestimation than S12 (Figure 3b). However, S12 was less accurate in predicting the thermocline depth with mean RMSE of 13.1 m and MBE of −11.2 m, while the values were 12.5 and 2.6 m using KPP (Table 3; Figures 3c and 3d). Lake Geneva showed to be an outlier with RMSEs over 60 m using both algorithms because it is a very deep lake of 310 m. Both algorithms showed relatively poor performances on reservoirs Mozaisk, Rappbode, Rimov, and SauReservoir (RMSE > 10 m, NSE < 0) because river discharge to deep layers of reservoirs, such as metalimnetic or hypolimnetic discharge, can largely affect the water thermal profile and even outrun wind-induced mixing. Algorithm S12 generally underestimated the thermocline depth while KPP on the contrary overestimated the depth. NSEs were negative for both algorithms, showing their deficiency in predicting the seasonal movements of thermocline for 46% and 54% of the lakes using S12 and KPP, respectively (Table S7). This is further explained below and discussed in Section 5.2 about simulation strategies. The parameters for both algorithms achieved relatively high convergence in both individual and ensemble calibration (Figures 3e and 3f) and therefore, the model accuracy only slightly decreased in the ensemble calibration (Figure S1). This indicates that both algorithms could be applied to larger-scale simulations without extensive recalibration efforts.
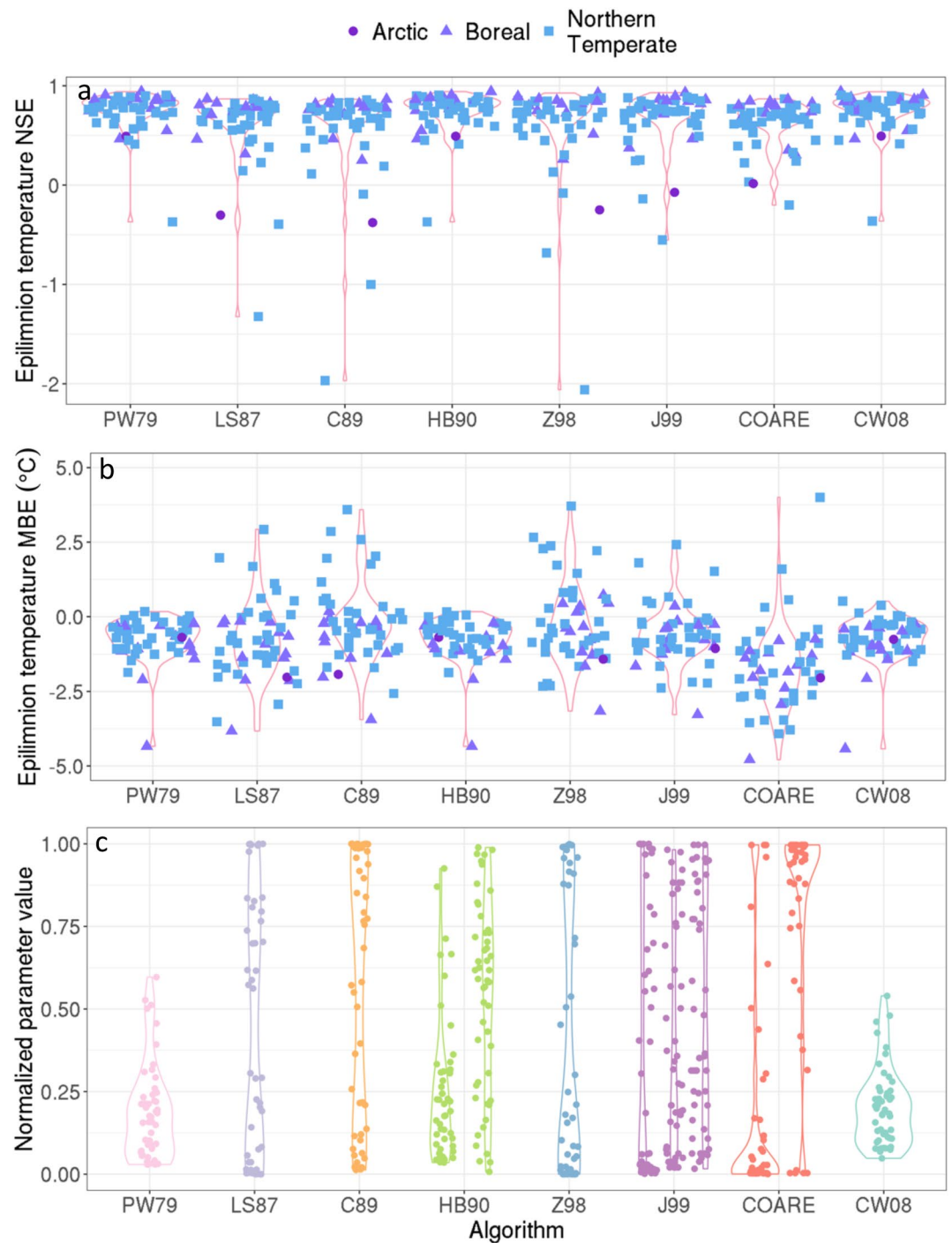
Figure 1. NSEs (a) and MBEs (b) of simulated lake epilimnion temperatures, and the normalized values of each parameter (c) in the individual model calibration of turbulent heat fluxes. The parameter values are normalized with respect to the total range of parameter values in Table 1. NSE, Nash-Sutcliffe efficiency; MBE, mean bias error.

The performances of wind-driven mixing algorithms were related to lake shape and depth (Figure S2). KPP showed increasing accuracy as the shape factor ($= \sqrt{area} / depth$) increases while it was the opposite for S12. The mean MBEs for lakes with shape factors larger than 100 were -6.3 and 0.6 m, using Algorithm S12 and KPP, respectively. Large shallow lakes are typically affected by horizontal and 3-D mixing which is not well resolved in 1-D lake models because of the assumption of horizontal homogeneity. The better predicted
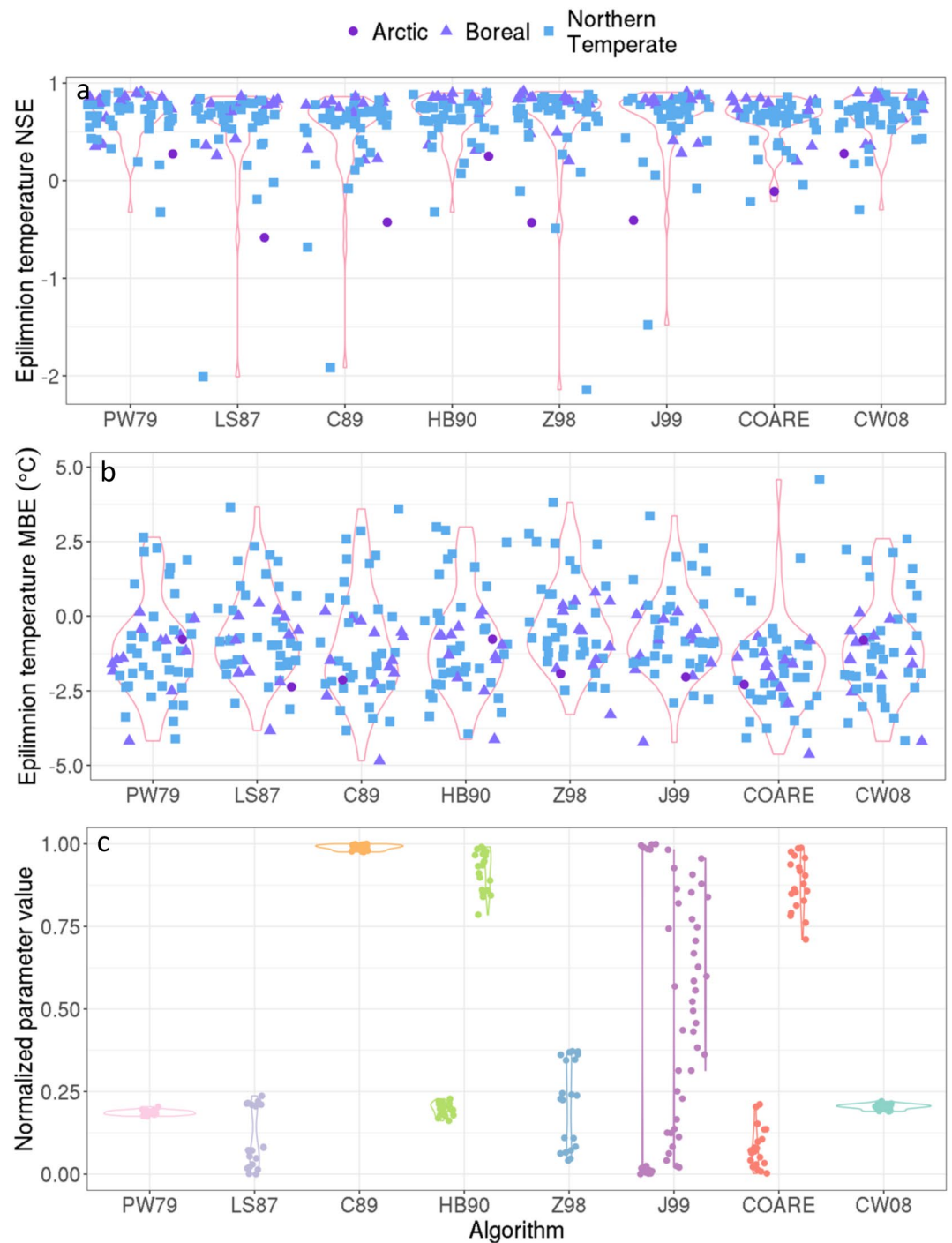
Figure 2. Same as Figure 1 but for the ensemble model calibration in which we calculate the mean RMSEs of the simulated epilimnion temperatures for all lakes. (c) The 20 sets of parameter values with the smallest mean RMSEs for each algorithm. The range which the parameter values vary within defines the degree of convergence. RMSE, root mean square error.

stratification and mixing of these lakes with KPP compared to the widely used S12 proves the practicability and reliability of 1-D lake models in simulating lakes of various types with improvement in the parameterization schemes.
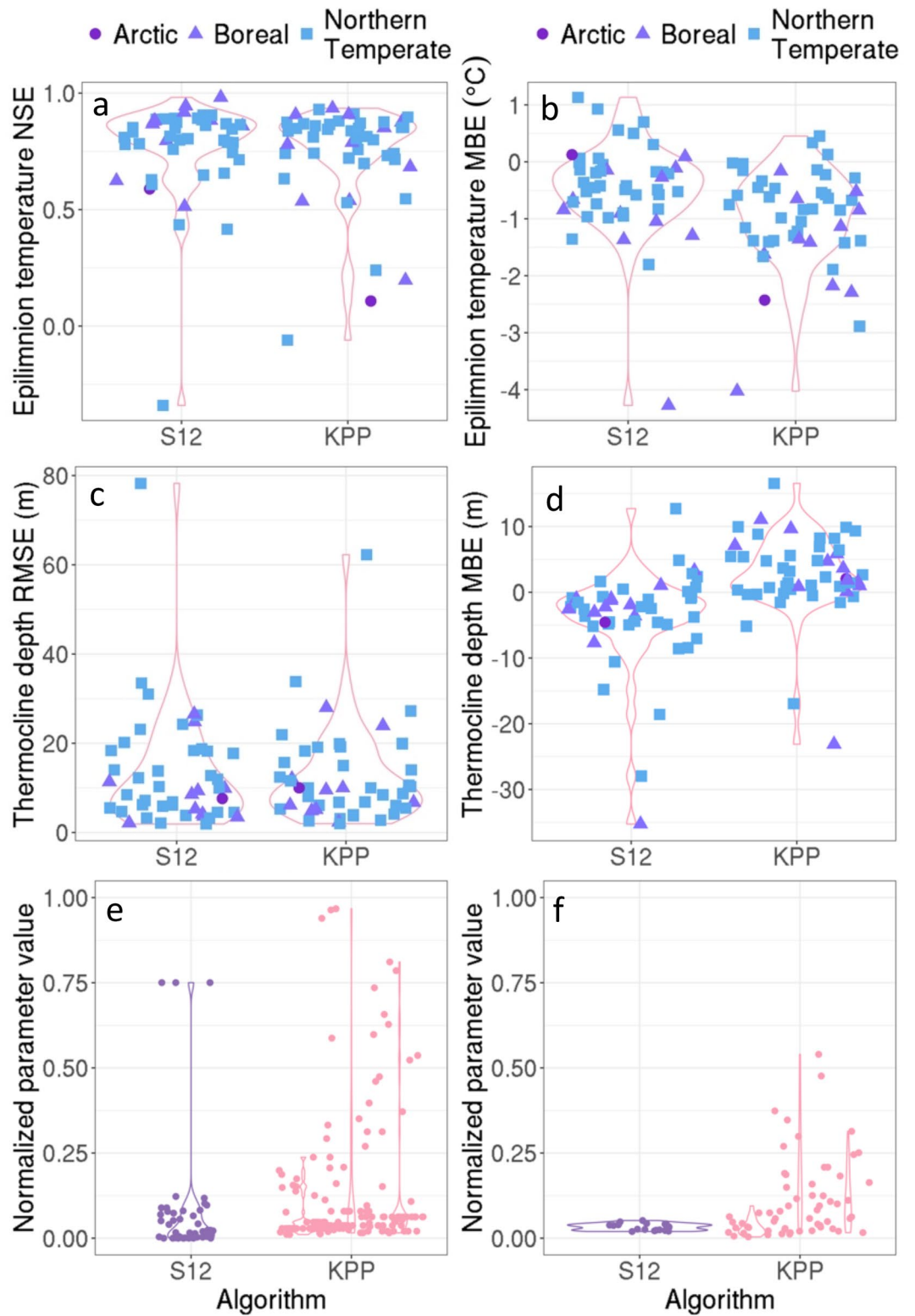
Figure 3. NSEs and MBEs for modeled epilimnion temperatures (a and b) and lake thermocline depths (c and d) in the individual calibration, and the normalized parameter values in the individual calibration (e) and the ensemble calibration (f). NSE, Nash-Sutcliffe efficiency; MBE, mean bias error.

The comparison of modeled epilimnion temperatures and thermocline depths by month also revealed a seasonal pattern (Figure 4). Generally, the largest simulation biases of epilimnion temperatures occurred in May and declined toward October for both algorithms. KPP was less accurate than S12 in May and June as it largely underestimated the epilimnion temperatures while the performance improved from July to October and became better than S12 (Figures 4a and 4b). In the prediction of thermocline depths, the RMSEs decreased from May to summer and increased into fall for both algorithms (Figures 4c and 4d). KPP only showed larger biases than S12 in May and June because of overestimation of the thermocline depth. The advantage of KPP was especially evident in fall when S12 largely underestimated lake mixing.

4.3. Light Extinction

Because wind-driven mixing is likely the main controlling factor of lake thermocline depths during spring and fall overturn, we compared modeled summer (JJA) thermocline depths only in the evaluation of light extinction algorithms. The summer thermocline is overall stable with little variation and thus, r and NSE were not calculated. Other error metrics can be found in Table S10. The simulation results showed no dramatic differences among different algorithms (Figure 5). All algorithms generally underestimated the lake thermocline depth with H95c being the worst, indicating possible inefficiency in the modeling of water mixing. Compared to the observed light extinction, the calculated values using H95 had the lowest RMSE of 0.55 (Figure 6). The uncalibrated algorithms (H95 and S19) tended to overestimate η at lower values while underestimating it at higher values. The calibrated algorithm (H95c) showed low coherence to the observed light extinction values.

4.4. Snow Density

The ice-off day alone was used as the reference variable for the snow density group. Ice-on day was not considered here because it was found to be mainly affected by turbulent heat transfer and wind-driven mixing (Guo, Zhuang, Golub, Yao, Leung, Pierson, et al., 2020; Tan et al., 2018). Since the annual ice-off day is not a continuous event, only RMSE, MBE, and MABE were calculated as shown in Tables S11 and S12. For the individual calibration, the mean RMSE and MBE of modeled ice-off day were 8.38 and 2.4 days, respectively, using algorithm Calib, 8.36 and 4.46 days using Y81, 7.97 and 3.0 days using H03, and 14.59 and 12.88 using H19, respectively (Figure 7). All algorithms predicted a later ice-off day for Lake Kuivajarvi by at least 13 days. H19 produced the largest discrepancies and overestimated ice-off days for all lakes. For Calib, the calibrated snow densities were around 105 kg m^{-3} except for Toolik Lake for which the optimal density was 268 kg m^{-3} (Figure S5). For Y81, snow densities for all lakes showed a similar pattern with a value that slightly increased above 250 kg m^{-3} as snow accumulated and dropped back at the end. H03 varied in the same pattern as Y81 during the ice-on season but started from 100 kg m^{-3}. The simulated density fluctuated sharply between the minimum and the maximum values during the same season using H19. For snow heat conductivity, all algorithms except for H19 resulted in values around 0.25 W m^{-1} K^{-1} with the value ~0.04 W m^{-1} K^{-1} lower using H03, and not much different among lakes (Figure S6). Following the snow density pattern, H19 simulated values varying frequently between 1.5 and 0.25 W m^{-1} K^{-1}. In the ensemble calibration, only Calib showed evident differences in the error metrics from the individual calibration. Both Calib and H03 achieved convergence in parameter values. Overall, H03 showed the best performance except for underestimating the ice-off day of Toolik Lake.

5. Discussion

5.1. Turbulent Heat Flux

We examined the impact of the algorithm choice on the simulation performance for lakes of different characteristics including climate (annual mean air temperature), latitude, surface area, maximum depth, and shape factor but found no significant correlations, indicating that no algorithm is specifically the best or the worst for a certain type of lakes and thus, one optimal algorithm can be used for global lakes. Overall, the single-parameter non-MOS algorithms, PW79 and CW08, had similar performances and although they
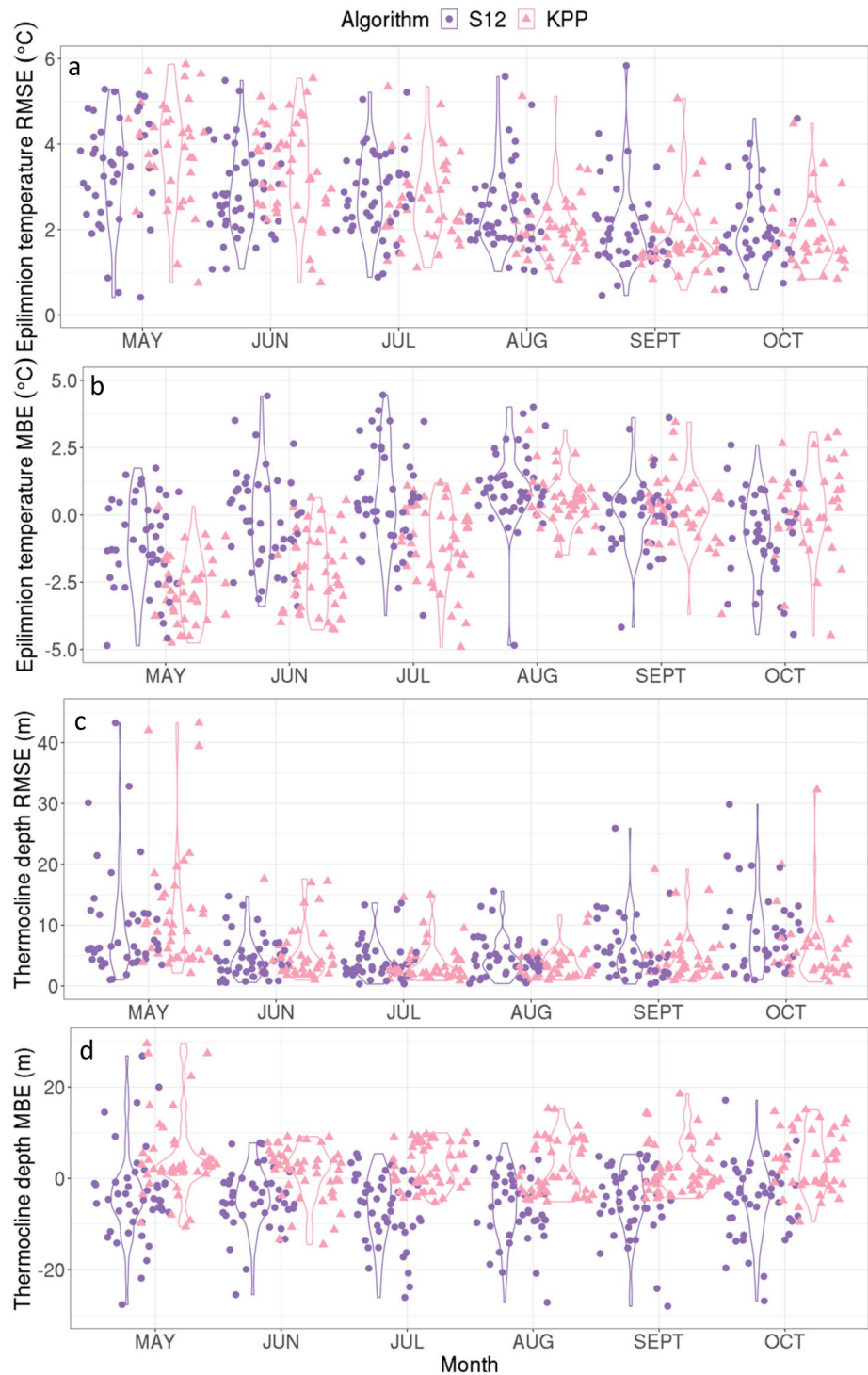
Figure 4. RMSEs and MBEs of modeled epilimnion temperatures (a and b) and thermocline depths (c and d) by month. RMSE, root mean square error; MBE, mean bias error.

are the most simplified, they had higher NSEs and lower RMSEs than other algorithms and were also the most applicable for larger-scale simulations. One explanation is that more parameters would introduce more uncertainties and thus, HB90 showed no advantage over PW79 even though in HB90 the bulk transfer coefficients are formulated as functions of lake surface area for a more realistic representation. Also, the MOS methods would result in more computation errors by calculating the logarithmic functions and the
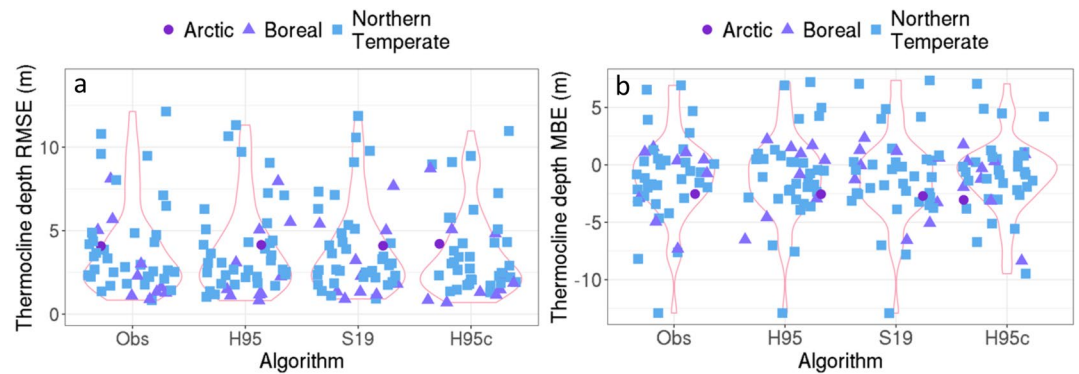
Figure 5. RMSEs (a) and MBEs (b) for lake thermocline depths from June to August in model calibration of light extinction coefficient algorithms. RMSE, root mean square error; MBE, mean bias error.

fractional numbers with small values (e.g., roughness length scale) in the denominator. Both PW79 and CW08 algorithms tended to underestimate the epilimnion temperature by about 1.1°C, slightly larger than other algorithms except for COARE. This can be compensated by using a calibrated wind-driven mixing algorithm (Table S6).

Our results differ from Charusombat et al. (2018) that tested the five MOS algorithms on four Great Lakes and found the COARE to be the most accurate and others to overestimate turbulent heat fluxes in fall and winter. This is probably attributed to the lack of calibration of the algorithms and thus simulation discrepancies resulting from the inappropriate default parameter values but rather the algorithm configurations. Also, the study focused on the simulated heat fluxes but epilimnion temperatures, which is not fully comparable with our findings. Further testing using calibrated parameters is needed to verify the cause of different results.

For all algorithms, we found no dependence of parameter values on lake characteristics and therefore, the same parameter values should be assigned to all lakes. The optimal value of parameter *hwt* in algorithm PW79 was 1.33 and the value of a_{CW} in CW08 was 6.08, both higher than the default values. Combining
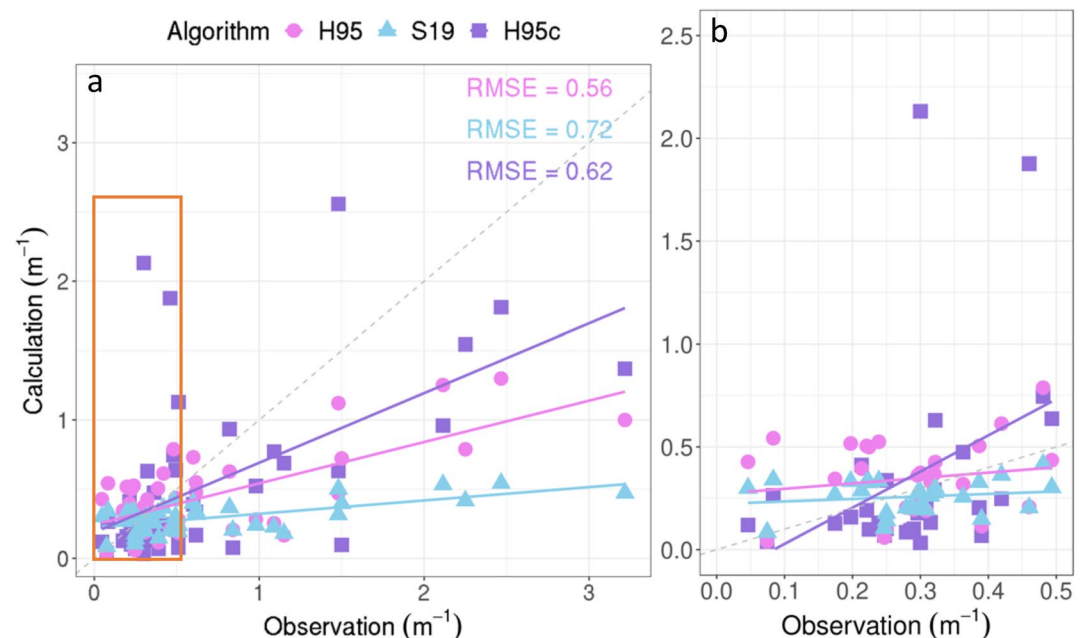
Figure 6. Calculated light extinction coefficients versus the observed value (a) and the zoomed-in plot of observed values lower than 0.5 m^{-1}, that is, the points within the orange box (b).
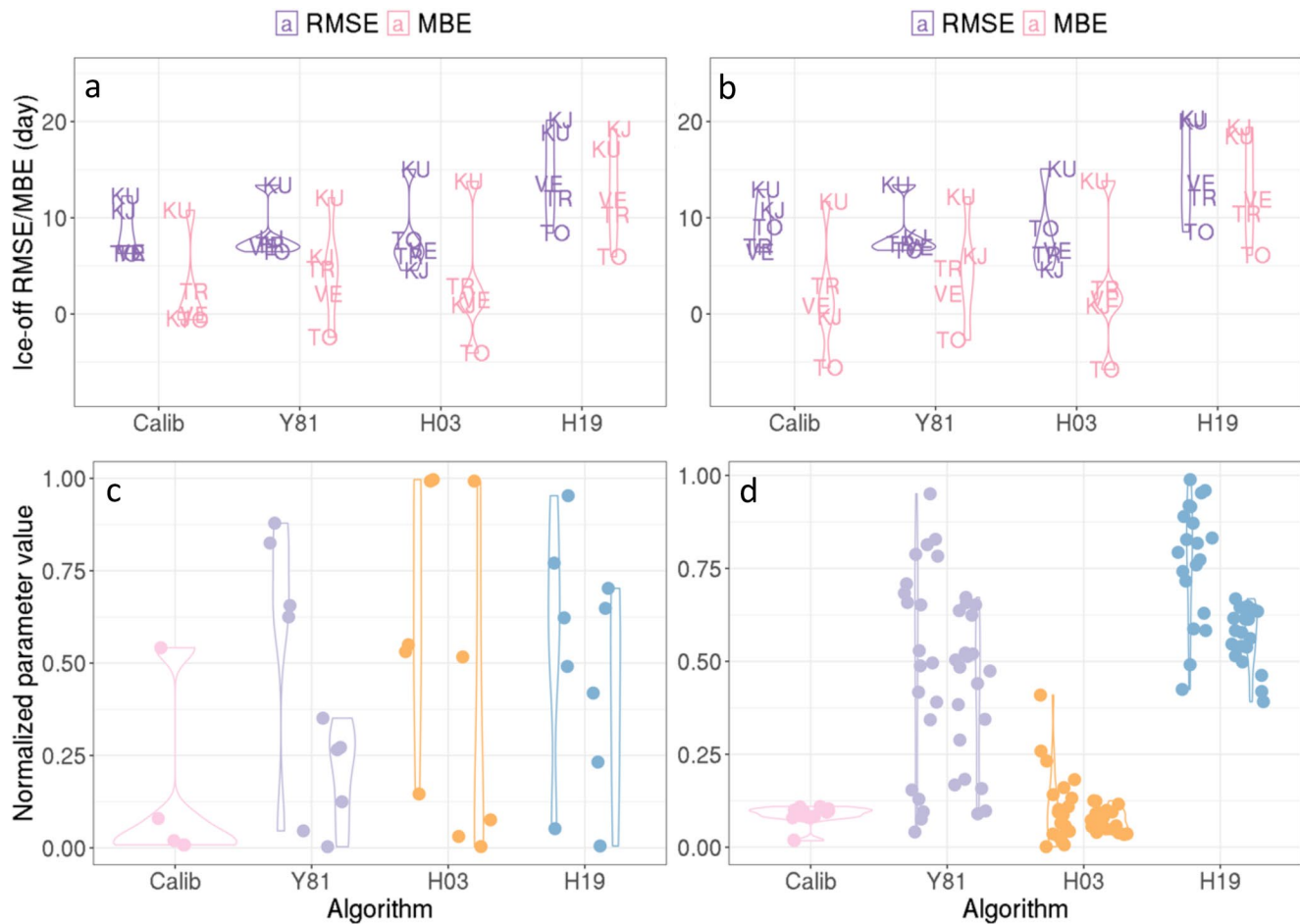
Figure 7. RMSEs and biases for ice-off days (a), and the distributions of normalized parameters (c) in the individual calibration, and in the ensemble calibration (b and d) of snow density. The lakes are denoted with abbreviations as in Table S1. RMSE, root mean square error.

the calibrated PW79 and the calibrated S12 algorithm alone will result in a mean epilimnion temperature NSE of 0.77, RMSE of 2.68°C, and MBE of −0.46°C, which is adequate for global simulations. The model performance can be further improved by implementing a two-algorithm strategy for wind-driven mixing as discussed below.

5.2. Wind-Driven Mixing

The KPP algorithms inherently enhances mixing, which leads to lower epilimnion temperatures and deeper thermocline depths compared to S12. For high latitude lakes, in spring and early summer after lake ice is just thawed, the water surface usually warms up quickly by solar radiation. However, by simulating unrealistically strong lake mixing that brings heat downward quickly and prevents heat accumulation, KPP substantially underestimates the lake epilimnion temperature and overestimates the thermocline depth. Examples include lakes Allequash, Harp, Langtjern, and Toolik (Figure S3). When the lake becomes stably stratified in summer, S12 is inefficient in modeling lake mixing and thus predicts slightly higher epilimnion temperature and shallower thermocline depth. This deficiency has larger impacts on simulation accuracy in the fall when cooling at the water surface is faster than the hypolimnetic layers, leading to overturning. S12 fails to reproduce the strong mixing during this period and thus largely overestimates epilimnion temperature and underestimates thermocline depth. This phenomenon can be found in several lakes such as BigMuskellunge, Erken, Paarjarvi, and Vendyurskoe (Figure S4). Our finding agrees with Stepanenko et al. (2013) and Huang et al. (2019) that models using S12 failed to predict summer stratification correctly due to insufficient wind-induced mixing. Huang et al. (2019) proposed to multiply the total thermal diffu-

sivity K with an enhancing factor of 40 to improve the model performance, which is too arbitrary to apply to other lakes and can also cause unrealistically strong mixing below the lake boundary layer.

Based on the performances in modeling lake temperature and stratification, we suggest a two-algorithm strategy for wind-driven mixing parameterization to improve simulation accuracy. Depending on lake shape, KPP can largely improve the simulation of lakes with shape factors larger than 100 while S12 can be used for small deep lakes that are well represented under the 1-D assumption. S12 is more accurate for spring and early summer while KPP can reproduce the lake thermal profile better during the steady stratification and fall overturn periods. The actual months of each period may vary by the lake with later spring thawing and earlier overturning for higher latitude lakes.

5.3. Light Extinction

Among the four algorithms, the modeled summer thermocline depth had slightly lower RMSE and lower PAE using the observed values. Heiskanen et al. (2015) found that the simulated lake thermal stratification is especially sensitive to the light extinction coefficient when it is lower than 0.5 m^{-1} and mapping of the coefficient values would result in better model accuracy while for lakes over the threshold, a default value could be safely assumed. At this value range, S19 provided the closest approximation of the observation. Therefore, when observation is not available, S19 can better represent oligotrophic and mesotrophic lakes that are sensitive to the calculated light extinction coefficient.

In the model, shortwave radiation penetrating to water depth z (m) is modeled using Beer's law, $\Phi = \left(1 - \beta\right)L_0 e^{-\eta z}$, where 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. We acknowledge that a more process-based lake model should include the effect of solids and aquatic organisms on light penetration and ALBM actually has such a capability as described in Tan et al. (2017). However, this effect cannot be evaluated in the current study because the ISIMIP data set does not provide the data of either solids and aquatic organisms or nutrient levels and hydrological fluxes required to simulate solids and organisms.

5.4. Snow Density

Through comparing the simulated snow density using Calib and H03 (Figure S6), we noticed that although H03 models similar snow densities to Calib but more realistic snow processes for most lakes, it failed to model the high snow density for Toolik Lake and thus underestimated the ice-off day. This indicates that H03 can be applied to most northern temperate and boreal lakes while for Arctic lakes, a higher value may be necessary. Ice phenology records for more Arctic lakes will help to verify this inference.

The difference in modeled snow density is the main source of the different simulation biases and snow heat conductivity explains the rest (Figures S6 and S7). Generally, higher snow density leads to later ice-off day while higher snow heat conductivity leads to earlier ice-off day. All algorithms showed a relatively high discrepancy for Lake Kuivajarvi, overestimating the ice-off day by 10–20 days, which was due to the biases in surface heat flux simulations. For Y81 and H19, the high snow density leads to more gray ice formulation. Calib and H03 simulate similar snow density at around 100 kg m^{-3} while the snow heat conductivity is lower using H03, resulting in more heat loss at the snow surface instead of transfer downward and thus slower melting. The algorithm behaviors for other lakes can be explained following the same logic. Considering the simulation performances, the 0.27 W m^{-1} K^{-1} used in Calib may be appropriate for all lakes.

6. Conclusions

We conduct a comprehensive evaluation of algorithms used for modeling four key thermal processes in lakes based on the criteria of both simulation accuracy and global applicability. By isolating the algorithms within the same parent model, we can focus on the merits of the algorithms individually and rule out all other influential factors in the comparison within each group. Moreover, the 50 lakes used to test the algorithms cover a wide range of size, shape, latitude, and climate zone, and therefore, our findings are informative for global lakes. For turbulent heat flux, although the MOS-based algorithms consider more details of physical processes, the discrepancy in simulated epilimnion temperature becomes larger due to

complex parameterization and potential computing errors. The non-MOS-based algorithms achieve highly convergent parameters in the ensemble calibration without sacrificing the overall accuracy. The two wind-driven mixing algorithms evaluated can both be applied to global lakes without complicated recalibration. We find that using the two algorithms in combination based on lake shape and season leads to optimum results in modeled lake temperature and stratification. The model performances are similar using different light extinction algorithms while in the absence of observation, S19 results in the closest values and thus is the best to represent the oligotrophic and mesotrophic lakes that are especially affected by calculated light extinction coefficient. Finally, H03 is found to behave well for simulating boreal and northern temperature lake snow densities while a higher value of $\sim$300 kg m^{-3} is likely more appropriate for Arctic lakes. Meanwhile, the snow heat conductivity at 0.27 W m^{-1} K^{-1} is accurate enough for all lakes. Our findings provide informative guides to 1-D lake model improvement for global applications, which enhances the practicability of coupling these models with weather and climate prediction systems to improve prediction accuracy at a global scale.

Data Availability Statement

The model output and codes for data analysis in this study are available at https://purr.purdue.edu/publications/3603/1 (Guo, Zhuang, Golub, Yao, Leung, & Tan, 2020).

References

Ashton, G. D. (Ed.). (1986). *River and lake ice engineering*. Littleton, CO: Water Resources Publications.

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

Benson, B., Magnuson, J., & Sharma, S. (2000). *Global lake and river ice phenology database, Version 1 (indicate subset used) updated 2020*. Boulder, CO: NSIDC National Snow and Ice Data Center. https://doi.org/10.7265/N5W66HP8

Brown, L. C., & Duguay, C. R. (2010). The response and role of ice cover in lake–climate interactions. *Progress in Physical Geography*, *34*(5), 671–704. https://doi.org/10.1177/0309133310375653

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

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

Cole, T. M., & Wells, S. A. (2008). *CE-QUAL-W2: A two-dimensional, laterally averaged, hydrodynamic and water quality model, version 3.6 user manual*. Vicksburg, MS: Environmental Laboratory U.S. Army Corps of Engineers Waterways Experiment Station.

Croley, T. E. (1989). Verifiable evaporation modeling on the Laurentian Great Lakes. *Water Resources Research*, *25*(5), 781–792. https://doi.org/10.1029/WR025i005p00781

DelSontro, T., Boutet, L., St-Pierre, A., del Giorgio, P. A., & Prairie, Y. T. (2016). Methane ebullition and diffusion from northern ponds and lakes regulated by the interaction between temperature and system productivity. *Limnology & Oceanography*, *61*(S1), S62–S77. https://doi.org/10.1002/lno.10335

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., 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*, 100–112.

Edson, J. B., Jampana, V., Weller, R. A., Bigorre, S. P., Plueddemann, A. J., Fairall, C. W., et al. (2013). On the exchange of momentum over the open ocean. *Journal of Physical Oceanography*, *43*(8), 1589–1610. https://doi.org/10.1175/JPO-D-12-0173.1

Environmental Data Center Team. (2020). *Toolik naturalist's journal*. Fairbanks, AK: Toolik Field Station, Institute of Arctic Biology, University of Alaska Fairbanks. Retrieved from http://toolik.alaska.edu/edc/journal/index.php

Fairall, C. W., Bradley, E. F., Rogers, D. P., Edson, J. B., & Young, G. S. (1996). Bulk parameterization of air-sea fluxes for tropical ocean-global atmosphere coupled-ocean atmosphere response experiment. *Journal of Geophysical Research*, *101*(C2), 3747–3764. https://doi.org/10.1029/95JC03205

Fang, X., & Stefan, H. G. (1996). Dynamics of heat exchange between sediment and water in a lake. *Water Resource Research*, *32*(6), 1719–1727. https://doi.org/10.1029/96WR00274

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

Goudsmit, G. H., Burchard, H., Peeters, F., & Wüest, A. (2002). Application of k–ε turbulence models to enclosed basins: The role of internal seiches. *Journal of Geophysical Research*, *107*(C12), 3230. https://doi.org/10.1029/2001JC000954

Gu, H., Jin, J., Wu, Y., Ek, M. B., & Subin, Z. M. (2015). Calibration and validation of lake surface temperature simulations with the coupled WRF-lake model. *Climatic Change*, *129*, 471–483. https://doi.org/10.1007/s10584-013-0978-y

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

Guo, M., Zhuang, Q., Yao, H., Golub, M., Leung, L. R., Pierson, D., & Tan, Z. (2020). Validation and sensitivity analysis of a 1-D lake model across global Lakes. *Journal of Geophysical Research: Atmospheres*, *125*, e2020JD033417. https://doi.org/10.1029/2020JD033417

Guo, M., Zhuang, Q., Yao, H., Golub, M., Leung, L. R., & Tan, Z. (2020). *Data for intercomparison of thermal regime algorithms in 1-D lake models*. Purdue University Research Repository. https://doi.org/10.4231/8246-C724

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

Håkanson, L. (1995). Models to predict Secchi depth in small glacial lakes. *Aquatic Sciences*, *57*(1), 31–53. https://doi.org/10.1007/BF00878025

Heise, E., Lange, M., Ritter, B., & Schrodin, R. (2003). *Improvement and validation of the multi- layer soil model. COSMO Newsletter* (No. 3, pp. 198–203). Offenbach am Main, Germany: Consortium for small-scale Modelling, Deutscher Wetterdienst. Retrieved from https://www.cosmo-model.org

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

Henderson-Sellers, B. (1985). New formulation of Eddy diffusion thermocline models. *Applied Mathematical Modelling*, *9*(6), 441–446. https://doi.org/10.1016/0307-904X(85)90110-6

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 Network (GLEON). *Geoscientific Model Development*, *12*, 473–523. https://doi.org/10.5194/gmd-12-473-2019

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. https://doi.org/10.1029/WR026i010p02603

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

Hunke, E. C., Lipscomb, W. H., Turner, A. K., Jeffery, N., & Elliott, S. (2015). *CICE: The Los Alamos Sea Ice Model documentation and software user's manual (LA-CC-06-012, pp. 1–116)*. Los Alamos, NM: Los Alamos National Laboratory.

Janssen, A. B. G., Arhonditsis, G. B., Beusen, A., Bolding, K., Bruce, L., Bruggeman, J., et al. (2015). Exploring, exploiting and evolving diversity of aquatic ecosystem models: A community perspective. *Aquatic Ecology*, *49*, 513–548. https://doi.org/10.1007/s10452-015-9544-1

Jordan, R. E., Andreas, E. L., & Makshtas, A. P. (1999). Heat budget of snow-covered sea ice at North Pole 4. *Journal of Geophysical Research*, *104*(C4), 7785–7806. https://doi.org/10.1029/1999JC900011

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

Large, W. G., McWilliams, J. C., & Doney, S. C. (1994). Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Reviews of Geophysics*, *32*(4), 363–403. https://doi.org/10.1029/94RG01872

Liu, P. C., & Schwab, D. J. (1987). A comparison of methods for estimating u* from given uz and air–sea temperature differences. *Journal of Geophysical Research*, *92*(C6), 6488–6494. https://doi.org/10.1029/JC092iC06p06488

MacKay, M. D., Patrick, J., De Senerpont, D., Fang, X., Gal, G., Jöhnk, K. D., et al. (2009). Modeling lakes and reservoirs in the climate system. *Limnology & Oceanography*, *54*(6–2), 2315–2329.

Matthews, E., Johnson, M. S., Genovese, V., Du, J., & Bastviken, D. (2020). Methane emission from high latitude lakes: Methane-centric lake classification and satellite-driven annual cycle of emissions. *Scientific Reports*, *10*(1), 12465. https://doi.org/10.1038/s41598-020-68246-1

Mironov, D., Heise, E., Kourzeneva, E., Ritter, B., Schneider, N., & Terzhevik, A. (2010). Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO. *Boreal Environmental Research*, *15*, 218–230.

Monin, A. S., & Obukhov, A. M. (1954). Basic laws of turbulent mixing in the surface layer of the atmosphere. *Geophysical Institute of the Academy of Sciences of the USSR*, *151*(163), e187.

Mooij, W. M., Trolle, D., Jeppesen, E., Arhonditsis, G., Belolipetsky, P. V., Chitamwebwa, D. B. R., et al. (2010). Challenges and opportunities for integrating lake ecosystem modelling approaches. *Aquatic Ecology*, *44*, 633–667. https://doi.org/10.1007/s10452-010-9339-3

Parkinson, C. L., & Washington, W. M. (1979). A large-scale numerical model of sea ice. *Journal of Geophysical Research*, *84*(C1), 311–337. https://doi.org/10.1029/JC084iC01p00311

Rouse, W. R., Oswald, C. J., Binyamin, J., Spence, C., Schertzer, W. M., Blanken, P. D., et al. (2005). The role of Northern lakes in a regional energy balance. *Journal of Hydrometeorology*, *6*, 291–305. https://doi.org/10.1175/JHM421.1

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 simulated by a regional climate model. *Boreal Environmental Research*, *15*, 113–129.

Shatwell, T., Thiery, W., & Kirillin, G. (2019). Future projections of temperature and mixing regime of European temperate lakes. *Hydrology and Earth System Sciences*, *23*(3), 1533–1551. https://doi.org/10.5194/hess-23-1533-2019

Stepanenko, V., 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*, 1977–2006. https://doi.org/10.5194/gmd-9-1977-2016

Stepanenko, V., Martynov, A., Jöhnk, K. D., Subin, Z. M., Perroud, M., Fang, X., et al. (2013). A one-dimensional model intercomparison study of thermal regime of a shallow, turbid midlatitude lake. *Geoscientific Model Development*, *6*(4), 1337–1352. https://doi.org/10.5194/gmd-6-1337-2013

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*, 4681–4699. https://doi.org/10.1029/2017WR022334

Tan, Z., Zhuang, Q., Shurpali, N. J., Marushchak, M. E., Biasi, C., Eugster, W., & Anthony, K. W. (2017). Modeling CO_2 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*, 459–483. https://doi.org/10.1002/2014MS000344

Trolle, D., Hamilton, D. P., Hipsey, M. R., Bolding, K., Bruggeman, J., Mooij, W. M., et al. (2012). A community-based framework for aquatic ecosystem models. *Hydrobiologia*, *683*, 25–34. https://doi.org/10.1007/s10750-011-0957-0

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*(5), 822–831. https://doi.org/10.4319/lo.1996.41.5.0822

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

Yao, H., Samal, N. R., Joehnk, K. D., Fang, X., Bruce, L. C., Pierson, D. C., et al. (2014). Comparing ice and temperature simulations by four dynamic lake models in Harp Lake: Past performance and future predictions. *Hydrological Processes*, *28*(16), 4587–4601. https://doi.org/10.1002/hyp.10180

Yen, Y. C. (1981). *Review of thermal properties of snow, ice, and sea ice* (Vol. 81, No. 10). Hanover, NH: Cold Regions Research and Engineering Laboratory, US Army, Corps of Engineers.

Zdorovennov, R., Palshin, N., Zdorovennova, G., Efremova, T., & Terzhevik, A. (2013). Interannual variability of ice and snow cover of a small shallow lake. *Estonian Journal of Earth Sciences*, *62*(1). https://doi.org/10.3176/earth.2013.03

Zeng, X., Zhao, M., & Dickinson, R. E. (1998). Intercomparison of bulk aerodynamic algorithms for the computation of sea surface fluxes using TOGA COARE and TAO data. *Journal of Climate*, *11*, 2628–2644. https://doi.org/10.1175/1520-0442(1998)011<2628:IOBAAF>2.0.CO;2

Zhang, Q., Jin, J., Wang, X., Budy, P., Barrett, N., & Null, S. E. (2019). Improving lake mixing process simulations in the Community Land Model by using K profile parameterization. *Hydrology and Earth System Sciences*, *23*(12), 4969. https://doi.org/10.5194/hess-23-4969-2019