RESEARCH ARTICLE

10.1029/2022JG007137

Key Points:
- Current global lake methane release estimated to account for 11% of the global total natural source emissions
- Future global lake emissions are estimated to increase 58%–86% under the severe climate warming scenario RCP8.5
- Under warming conditions, methanogenesis will increase, but enhanced methane oxidation in water column can dampen the net methane emission

Abstract

Freshwater ecosystem contributions to the global methane budget remains the most uncertain among natural sources. With warming and accompanying carbon release from thawed permafrost and thermokarst lake expansion, the increase of methane emissions could be large. However, the impact and relative importance of various factors related to warming remain uncertain. Based on diverse lake characteristics incorporated in modeling and observational data, we calibrate and verify a lake biogeochemistry model. The model is then applied to estimate global lake methane emissions and examine the impacts of temperature increase for the first and the last decades of the 21st century under different climate scenarios. We find that current emissions are 24.0 ± 8.4 Tg CH₄ yr⁻¹ from lakes larger than 0.1 km², accounting for 11% of the global total natural source as estimated based on atmospheric inversion. Future projections under the RCP8.5 scenario suggest a 58%–86% growth in emissions from lakes. Our model sensitivity analysis indicates that additional carbon substrates from thawing permafrost may enhance methane production under warming in the Arctic. Warming enhanced methane oxidation in lake water can be an effective sink to reduce the net release from global lakes.

Plain Language Summary

Methane is the second most important greenhouse gas after carbon dioxide but with much stronger warming potential. Providing accurate estimates and future projections of methane emissions are challenging but essential for supporting climate mitigation strategies. Freshwater ecosystems including lakes and reservoirs contribute large uncertainty in the quantification of global methane sources. In this study, with an improved lake modeling approach, we estimate global lake methane emissions and examine the impacts of temperature increase, permafrost thaw, and thermokarst lake area changes in the future. Our modeling suggests that the current magnitude and the future increase of global lake methane emissions are lower than previously suggested. Under warming conditions, the increase of methane oxidation reduces the net lake methane emissions.

1. Introduction

Over the past two decades, freshwater ecosystems including lakes, ponds, and reservoirs have received increasing attention as observations indicate that they are a disproportionately large methane source compared to their surface areas (Saunois et al., 2016, 2020). Recent bottom-up estimates of global freshwater methane emissions (Johnson et al., 2022; Saunois et al., 2020) are 2–4 times larger than top-down estimates of all natural sources excluding wetlands (Saunois et al., 2016, 2020). For the region north of 50°N with the largest lake coverage and climate sensitivity, statistical upscaling (Walter et al., 2007; Wik et al., 2016) and process-based modeling (Tan & Zhuang, 2015) suggests that over 60% of the total natural methane emission is from lakes, which disagrees with some previous estimates suggesting a dominant contribution from wetlands (Fung et al., 1991; Melton et al., 2013). These disparities might be central to explaining the differences between top-down emission estimates based on atmospheric measurements and bottom-up estimated net emissions (Ronsentretre et al., 2021; Saunois et al., 2016, 2020). Both top-down and bottom-up results can be biased due to (a) low temporal resolution of most observations of methane fluxes; (b) over-representation or under-representation of lakes from certain regions and of certain properties, such as the bias toward more frequent sampling of small northern lakes.
Climate warming is likely to increase methane production from lakes and reservoirs, especially in the circum-polar permafrost region where gradual thaw can expose previously frozen carbon to microbial decomposition and abrupt thaw can accelerate the formation and expansion of thermokarst lakes (Schuur et al., 2009, 2015; Walter Anthony et al., 2016, 2018). Measurements of 37 permafrost-affected lakes over the past 60 years have shown 0%-4,000% higher methane ebullitive emissions in the expansion zones than in open water (Walter Anthony et al., 2016). Lakes formed through glacial and periglacial processes during the last deglaciation are estimated to represent around 78% of northern lake surface area (Wik et al., 2016). These lakes were projected to increase their emissions by 5 and 26 Tg CH4 yr⁻¹ due to newly thawed carbon within this century under the representative concentration pathway (RCP) 2.6 and RCP8.5 scenarios, respectively (Von Deimling et al., 2015). However, the contributions of arctic lakes to the global methane budget in a warmer future are still uncertain.

Here, we present a process-based modeling analysis of global lake methane emissions for the first and the last decades of the 21st century under different climate scenarios. We adapt a one-dimensional (1-D) process-based lake biogeochemical model (Tan et al., 2015, 2017) by modifying the methane module for various climate zones and lake and sediment properties. The model calibration and validation data sets are organized based on the temporal and spatial coverages, frequency of observations, and measurement techniques. We use the Hydro-LAKES database (Messager et al., 2016) which includes all lakes and reservoirs larger than 0.1 km² as the basis for our simulations. We discuss reasons behind the different estimates by statistical upscaling and process-based modeling. We then investigate the significance of different climate-induced factors, including direct warming, permafrost sediment organic carbon (SOC) input, and area changes in thermokarst lakes. We aim to reduce the uncertainties in freshwater methane emissions and improve our understanding of how lake ecosystems respond to climate change.

2. Methods

2.1. Model Configuration

The Arctic Lake Biogeochemical Model (ALBM) is a 1-D process-based model that simulates water and sediment thermal conditions and carbon and nutrient dynamics (Guo, Zhuang, Yao, Golub, Leung, & Tan, 2021; Tan et al., 2015, 2017; Figure S1 in Supporting Information S1). The model has been applied to lakes in various regions apart from the arctic (Guo et al., 2020; Guseva et al., 2020; Tan et al., 2018). The thermal module simulates radiative heating, turbulent heat transfer, ice phenology, heat diffusion and convection within water columns, sediment-water interface heat exchange, and sediment heat diffusion. More details can be found in Guo, Zhuang, Yao, Golub, Leung, Pierson, et al. (2021) and Guo, Zhuang, Yao, Golub, Leung, and Tan (2021) who evaluated the module against a variety of lakes, improved the algorithms through model intercomparisons, and calibrated the parameters to suit global-scale simulations. Here for simplicity, we briefly describe the algorithms for methane related processes which are the foci of this study. Detailed biogeochemical processes, including phytoplankton and other microbial activities and oxygen dynamics can be found in Tan et al. (2015, 2017). Full parameter definitions of the methane module are listed in Table 1. Some of the parameter values vary by lake type as explained in the next section.

Methane production (P) in the lake sediment is calculated as

$$P = R_c \cdot C_{labile} \cdot PQ_{10}^{\frac{T_e - T_{ref}}{10}}$$

(1)

where \(R_c\) is the carbon decomposition rate (s⁻¹), \(C_{labile}\) is the labile carbon in upper 0.3 m of sediment (kg C m⁻²), \(PQ_{10}\) is the Q10 factor representing the sensitivity to temperature, \(T_e\) is the sediment temperature (°C), and \(T_{ref}\) is the reference temperature (°C). \(C_{labile} = C_{pool} \cdot a_{labile}\), where \(C_{pool}\) is the total organic carbon and \(a_{labile}\) is the fraction of labile carbon. The carbon pool is considered constant over the simulation period. Methane is transported through the water column by diffusion and ebullition. The concentration in the water column (\(C_{CH4}\)) is calculated as

$$\frac{\partial C_{CH4}}{\partial t} = \frac{\partial}{\partial z} \left( D_{CH4} \frac{\partial C_{CH4}}{\partial z} \right) - O + E_{CH4}$$

(2)
where $D_{CH_4}$ is eddy diffusivity or entrainment, $O$ is aerobic methane oxidation, and $E_{CH_4}$ is the gas exchange with bubbles. Methane exchanges at the sediment-water and water-air interfaces are added to the bottom and top boundaries, respectively. Methane oxidation ($O$) in the water column is a function of water temperature and methane concentrations

$$O = O_{\text{max}} \frac{T_w - T_{\text{ref}}}{k_{O_2} + C_{O_2} + k_{CH_4} + C_{CH_4}}$$

where $O_{\text{max}}$ is the oxidation potential ($\mu$mol m$^{-3}$ s$^{-1}$), $OQ_{10}$ is the Q10 factor, $T_w$ is the water temperature (C), $T_{\text{ref}}$ is the reference temperature (C), $C_{O_2}$ and $C_{CH_4}$ are gas concentrations ($\mu$mol m$^{-3}$), and $k_{O_2}$ and $k_{CH_4}$ are Michaelis-Menten constants ($\mu$mol m$^{-3}$). The gas diffusive flux ($F$) at the water-air interface is calculated as

$$F = k_{600}(C_w - C_{\text{ref}})$$

where $k_{600}$ is gas transfer velocity normalized to $CO_2$ at 20°C, $C_w$ is near-surface concentration of dissolved $CH_4$, and $C_{\text{ref}}$ is near-surface concentration in equilibrium with the atmosphere. The calculation of gas transfer velocity follows Heiskanen et al. (2014): $k_{600} = \sqrt{(c_1 U)^2 + (c_2 w_s)^2 Sc^{-0.5}}$, where $c_1$ and $c_2$ are empirical constants, $U$ is the near-surface wind speed (m s$^{-1}$), $w_s = (\beta z_{AML})^{1/3}$, $\beta$ is buoyancy flux, $z_{AML}$ is the mixing layer depth, and $Sc$ is the Schmidt number. The governing equations of bubble transfer in the water column are adapted from ocean models (Liang et al., 2011; Woolf & Thorpe, 1991). The gas concentration in bubbles of certain size at certain depth and time step is calculated by a continuity equation in which the impacts of buoyant rising, gas exchange with ambient water, and bubble expansion are included (Liang et al., 2011). Gases in bubbles are assumed to release instantly to the atmosphere upon reaching the lake surface. Detailed formulation can be found in Tan et al. (2015).

### 2.2. Observational Data on Methane Emissions

We filtered flux measurement data to avoid short observation periods, inadequate spatial coverage, and under-represented seasonal variability. For the tropical regions with shallow lakes that experience large variations in water level and methane fluxes throughout the year, observations that cover at least 1 year and a range of water levels were selected (Amaral et al., 2020; Attermeyer et al., 2016; Marani & Alvalá, 2007). Therefore, observations that cover the hydrological cycle for at least 1 year were selected. For all other regions, a minimum of

---

**Table 1**

*Model Parameters in the Methane Module*

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Lake type</th>
<th>Value</th>
<th>Units</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_{\text{H}2}$</td>
<td>Pan-Arctic permafrost-affected</td>
<td>20</td>
<td>%</td>
<td>Schädal et al. (2014)</td>
</tr>
<tr>
<td></td>
<td>Oligo/mesotrophic, shallow warm water</td>
<td>15</td>
<td></td>
<td>Ostapenia et al. (2009) and Klump et al. (1989)</td>
</tr>
<tr>
<td></td>
<td>Eutrophic</td>
<td>22</td>
<td></td>
<td>Ostapenia et al. (2009)</td>
</tr>
<tr>
<td></td>
<td>Hypereutrophic</td>
<td>45</td>
<td></td>
<td>Keaveney et al. (2020)</td>
</tr>
<tr>
<td>$T_{\text{ref}}$</td>
<td>–</td>
<td>–</td>
<td></td>
<td>Tan et al. (2015)</td>
</tr>
<tr>
<td>$T_o$</td>
<td>–</td>
<td>–</td>
<td>–5.5</td>
<td>Zhuang et al. (2004)</td>
</tr>
<tr>
<td>$k_{O_2}$</td>
<td>–</td>
<td>1.4 $\times$ 10$^4$</td>
<td>$\mu$mol m$^{-3}$</td>
<td>Segers (1998) and van Bodegom et al. (2001)</td>
</tr>
</tbody>
</table>

*Note.* Calibrated parameters are listed with ranges.
3 months, at least one ice-free season and a sampling frequency that exceeded 1 day per month were required for lake observations. For lakes larger than 10 km², data from at least two methods, such as floating chambers and bubble traps, were required. If the boundary layer method was used to calculate diffusive fluxes, studies using wind speed measured simultaneously onsite instead of taken from more distant meteorological stations were selected. As a result, out of the 838 lakes and reservoirs with published data, 60, varying in size and climate zone (Figure S1 and File S1 in Supporting Information S1), were used for model calibration and validation. Fifty-five have observed diffusive emissions and 45 include ebullition observations. The physical properties (coordinates, surface area, and depth) and water and sediment properties (trophic status, thermokarst or not, yedoma or not) of the lakes, information of the measurements (time period, number of measurements, and technique), and the mean daily/annual fluxes are included in (File S1 in Supporting Information S1). Here, lakes and reservoirs were not treated as separate categories due to the lack of sufficient reservoir observations. The selected sites were first classified by climate zone and then by lake water nutrient level and sediment type. For tropical and subtropical regions, due to the low number of observation sites, they were divided into shallow warm water (30°N–30°S, mean depth <10 m) and other lakes (Table S1 in Supporting Information S1).

2.3. Model Calibration and Validation

The calibration of parameters in the thermal module was done in a previous study (Guo, Zhuang, Yao, Golub, Leung, & Tan, 2021) and thus, we calibrated the model for methane fluxes here. Observations with both diffusive and ebullitive emissions were used for parameter calibration and the remainder of observations for validation. For each site, the model was run over the observation period and parameters were calibrated against the mean daily fluxes during that period. For the lakes with only mean annual fluxes provided, we divided the values by the annual mean modeled length of ice-free days to obtain the mean daily fluxes.

The calibration was carried out using a Bayesian approach in which a uniform distribution was used for all parameters, and Monte-Carlo-based calibrations were conducted more than once by sampling from the posterior distribution obtained from the previous round (Guo, Zhuang, Yao, Golub, Leung, & Tan, 2021; Tang & Zhuang, 2009). The number of the sampling rounds was case-sensitive and decided by the model performance. The size of the perturbed parameter ensemble (PPE) equals the number of parameters multiplied by 1,000 in each round. The root-mean-square errors (RMSEs) of the simulated emissions during the same periods as the observations were calculated.

Lakes in the same category were calibrated together, meaning that the optimal parameters with the minimum mean RMSEs of all lakes in that category were selected. The calibration stops when RMSE is <10%. We did not calibrate the model against the time series of measurements because: (a) some observational data were only daily or annual means; (b) the temporal resolutions vary among sites, and therefore, using the time series will bias the model toward the observation-dense sites; and (c) the focus of this study is the annual mean emissions because long-term continuous measurements are very rare.

2.4. Global Lake Classification and Simulation

We used the HydroLAKES database (Messager et al., 2016) for global lake and reservoir mapping with information on coordinates, areas, and depths, which remained constant through all the simulation periods. Due to lack of bathymetric data, the horizontal cross-sectional area is assumed to decrease linearly for all lakes. Lakes were classified into various categories for model calibration (Figure S2 in Supporting Information S1), and the calibrated parameter values were applied to all lakes for the same category. The global SOC content is approximated from the global soil organic carbon map (GSOCmap, http://54.229.242.119/GSOCmap/). For the pan-Arctic region, the yedoma map was based on the Ice-Rich Yedoma Permafrost (IRYP, Strauss et al., 2016) data set, the permafrost map from Brown et al. (1997), and the thermokarst lake distribution from Olefeldt et al. (2016). For northern temperate lakes, the trophic status was determined by a combination of the Lake Water Quality (LWQ) data set that includes a trophic state index, chlorophyll concentrations (Filazzola et al., 2020), and the Global River Chemistry Database (GLORICH; Hartmann et al., 2014) that includes total nitrogen and phosphorus concentrations. Chlorophyll and solute concentrations were converted into trophic state indices based on the OECD (OECD, 1992) classification scheme. Lakes not included in these data sets were classified as oligotrophic.

Simulations were run from 2004 to 2006 (hereafter refer to as the 2000s), and 2094 to 2096 (hereafter the 2090s) under RCP8.5 scenarios with a spinup period of 10 years using climate forcing data from 1994 to 2003 and...
2084 to 2093, respectively. The required meteorological forcing data include surface air temperature, minimum and maximum surface air temperature, surface air pressure, relative humidity, 10-m wind speed, precipitation, snowfall, surface downward shortwave radiation, and surface downward longwave radiation. Inflow and outflow observations were not available at the global scale; therefore, we assumed that water levels are static. The forcing data for present and future scenarios were supplied by the ISIMIP2a bias-corrected climate input data by HadGEM2-ES and GFDL-ESM2M climate models (Frieler et al., 2017; Lange, 2019), which represent one of the severest and the mildest warming scenarios under RCP8.5, respectively, among climate models. The data are at 0.5° × 0.5° grids and daily time step from 1979 to 2016 and 2006 to 2099, respectively.

The ALBM is a site-level model, meaning that lakes are simulated individually. Given the large number of lakes globally (≈1.4 million), we reduced the total computational time by avoiding repetitive simulations of similar lakes. Specifically, lakes of the same type and similar depths and shapes (=√area/depth) within each 0.5° × 0.5°, the same resolution as the climate forcing data, were considered as a cohort and simulated as such. The mean coordinates, depth, surface area, and elevation were used as the simulation input for each cohort, and modeled fluxes were considered uniform for all lakes in the same cohort. Depths were binned with an interval of 1 m within (0,10] m and 10 m within (10,100] m. Shape factors are binned by (0, 50, 100, 200, 300, 400, 500, 600, 800, 1,000]. This manipulation helped to reduce about 27% of the total simulated lakes, especially in regions dense with small lakes (Figure S3 in Supporting Information S1).

The site uncertainty is quantified as the normalized root-mean-square-error NRMSE = RMSE/\(\bar{P}\). \(\bar{P}\) is model prediction. We assume the NRMSE values for calibrated lakes are constants for each lake type, which are then used to generate regional emission uncertainties based on lake types and their simulated emissions.

### 2.5. Experiments on Permafrost Dynamics Impacts

The large stocks of frozen organic carbon in permafrost soils might be exposed to microbial decomposition under warming-induced thaw. To estimate this impact on methane emissions from arctic lakes, we estimate lake area changes in yedoma and nonyedoma continuous and discontinuous permafrost regions based on the predicted permafrost maps under the RCP4.5 and RCP8.5. The permafrost maps were estimated based on the simulated soil temperature by MRI-CGCM3 (Yukimoto et al., 2012). We followed the method in Koven et al. (2013) to estimate both the historical and future extent of lake area. With the lake area having permafrost reduced by about 96% under the RCP8.5 (Table S4 in Supporting Information S1), lake methanogenesis might be significantly fueled by old organic carbon. Without regional-scale measurements or modeling results of the amount of permafrost organic carbon adding to lakes, we cannot explicitly simulate the impacts on lake methane emissions. Instead, we conducted a sensitivity test by doubling the lake SOC concentration under the RCP8.5 using the HadGEM2-ES climate model output. We acknowledge this is a hypothetical modeling experiment, although previous modeling and observational-based inventory studies all suggest increasing carbon availability due to thawing permafrost under future warming scenarios (Hugelius et al., 2020; Koven et al., 2015).

Thermokarst lakes in the carbon-rich yedoma and western Siberia regions have high methane emissions (Serikova et al., 2019; Walter Anthony et al., 2012). These lakes often persist for short periods compared to other lakes, and therefore, are more variable under climate change (van Huisstden et al., 2011). Abrupt permafrost thaw can cause rapid lake formation and large-scale gradual thaw will facilitate expansion and deepening of the lake (Turetsky et al., 2020; Walter Anthony et al., 2016). Alternatively, the changing hydrology may promote lateral breaching and thus, drainage of the lakes (Jones et al., 2011; Nitze et al., 2020). Both expansion and drainage have been observed across the pan-Arctic permafrost region whereas the dominant processes, magnitudes, and time scales are heterogeneous (Table S5 in Supporting Information S1). To evaluate the impacts on methane emissions, we assumed a change in the total thermokarst lake area by ±30% by the end of the century, which covers most of the observations over 30–60 years and is also consistent with Wik et al. (2016) for comparison purpose.

### 3. Results and Discussion

#### 3.1. Model Calibration and Validation

Overall, the model showed good performance with NRMSEs of 0.5 (0.66), 1.07 (0.27), and 1.28 (0.27) mmol m⁻² day⁻¹ in calibration (validation) and MBEs (mean bias error) of 0.26 (0.11), −0.15 (−0.14),
and 0.11 (−0.06) mmol m⁻² day⁻¹ for simulated diffusive, ebullitive, and total emissions, respectively (Figure 1). The metrics for each lake are listed in Tables S2 and S3 in Supporting Information S1. The model tends to produce lower diffusive fluxes than the observations. The calculation of vertical mixing and gas transfer velocity at water-atmosphere interface may contribute to the underestimation. Furthermore, floating chambers may occasionally capture bubbles, leading to overestimated diffusive fluxes.

### 3.2. Current Emissions and Their Spatial Distribution

The simulated mean annual methane emissions from global lakes and reservoirs larger than 0.1 km² are 24.0 ± 8.4 Tg CH₄ yr⁻¹, 2.6 ± 1.2 Tg CH₄ yr⁻¹ is from diffusion and 21.4 ± 7.2 Tg CH₄ yr⁻¹ is from ebullition. The temperate zone (23.5°–50°N, 23.5°–60°S) is the largest source contributing 9.7 ± 3.4 Tg CH₄ yr⁻¹, followed by 7.4 ± 2.6 Tg CH₄ yr⁻¹ and 6.9 ± 2.4 Tg CH₄ yr⁻¹ for the pan-Arctic (>50°N) and the tropics (23.5°N–23.5°S), respectively (Table 2 and Figure 2). The emission per unit lake area has different latitudinal distributions with 16.9 g CH₄ m⁻² yr⁻¹ for the tropics, 8.0 g CH₄ m⁻² yr⁻¹ for the temperate region, and 5.6 g CH₄ m⁻² yr⁻¹ for the pan-Arctic. The areal flux peaks between 0°S and 30°S and decreases toward the poles. The distribution patterns in the total emission and areal fluxes are related to the right-skewed distribution of lake surface areas across the globe (Figure 2b). The pan-Arctic total lake surface area (1.3 × 10⁶ km²)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Value</th>
<th>Pan-arctic</th>
<th>Temperate</th>
<th>Tropics</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>Eₗ₄₄</td>
<td>7.4</td>
<td>9.7</td>
<td>6.9</td>
<td>24.0</td>
</tr>
<tr>
<td>2090sHadGEM2-ES</td>
<td>Eₗ₄₄</td>
<td>23.5</td>
<td>12.8</td>
<td>8.3</td>
<td>44.6</td>
</tr>
<tr>
<td></td>
<td>ΔEₗ₄₄</td>
<td>217.6</td>
<td>31.3</td>
<td>21.0</td>
<td>85.8</td>
</tr>
<tr>
<td>2090sGFDL-ESM2M</td>
<td>Eₗ₄₄</td>
<td>19.1</td>
<td>11.6</td>
<td>7.1</td>
<td>37.8</td>
</tr>
<tr>
<td></td>
<td>ΔEₗ₄₄</td>
<td>166.1</td>
<td>19.4</td>
<td>2.6</td>
<td>57.5</td>
</tr>
</tbody>
</table>

Figure 1. Biases for lakes in calibration and validation for total methane fluxes (a, b), diffusive fluxes (c, d), and ebullitive fluxes (e, f). Lakes with both observed diffusive and ebullitive fluxes were used for calibration and those with only one observed pathway or only total fluxes were used for validation.
is larger than the areas in temperate (1.2 × 10^6 km^2) and tropical regions (0.4 × 10^6 km^2), and thus the total emissions from the pan-Arctic lakes are high even with the lower areal fluxes.

Our estimate is on the lower end of the 16.9–67.4 Tg yr⁻¹ (first and third quartiles, median = 32.1 Tg yr⁻¹) of the estimate by Rosentreter et al. (2021) for lakes larger than 0.1 km^2 and all reservoirs; the estimated emission from reservoirs smaller than 0.1 km^2 was only 0.4 Tg yr⁻¹. The estimated total for the pan-Arctic is about 55% lower than 16.5 ± 9.2 (mean ± 0.5 interquartile) Tg yr⁻¹ estimated by Wik et al. (2016) which includes all lakes larger than 0.002 km^2. Lakes smaller than 0.1 km^2, although minor in the total surface area, were identified as a disproportionately major, but especially uncertain, source in Rosentreter et al. (2021).

Both Wik et al. (2016) and Rosentreter et al. (2021) used the global water bodies (GLOWABO) data set (Verpoorter et al., 2014) for lakes larger than 0.002 km^2, and Rosentreter et al. (2021) included smaller lakes by statistically extrapolating the abundance and area. Messager et al. (2016) developed the HydroLAKES data set by manually correcting for fluvial features, and reducing the uncertainties caused by double counting small lakes and wetlands. For lakes larger than 0.1 km², GLOWABO reports higher lake abundance than HydroLAKES especially in the 0.1–1 km² size classes with 204% more lakes, and the discrepancy decreases with increasing size class. If we

Figure 2. Mean annual lake methane fluxes weighted by lake area (a). Latitudinal distributions of total number and surface area of lakes in HydroLAKES with 1° intervals (b), annual total methane emissions (c), and annual methane fluxes averaged by lake area (d). The smoothed curves are fitted by polynomial regression and shades are the uncertainty ranges.
assume the same discrepancy, these additional lakes would add 12.7 Tg yr\(^{-1}\) (median, 5.3–34.4 Tg yr\(^{-1}\), first and third quartiles) more emissions, leading to total 36.7 Tg yr\(^{-1}\) from all lakes.

Reservoirs are significant methane sources due, in part, to flooding of terrestrial organic carbon, enhanced ebullition during drawdown, and nutrients from anthropogenic activities (Deemer et al., 2016). In our study, although reservoirs and natural lakes were not differentiated, six reservoirs were included in the calibration and validation. The model overestimated the total fluxes by 0.1–0.5 mmol m\(^{-2}\) day\(^{-1}\) for four of the sites and underestimated the fluxes by 0–0.2 mmol m\(^{-2}\) day\(^{-1}\) for the other two. HydroLAKES includes 6,687 reservoirs covering 2.49 \times 10^5 \text{km}^2, about 9\% of the total surface area of lakes and reservoirs. The modeled reservoir emissions are 2.1 Tg CH\(_4\) yr\(^{-1}\), 0.4 Tg CH\(_4\) yr\(^{-1}\) by diffusion and 1.7 Tg CH\(_4\) yr\(^{-1}\) by ebullition, contributing ~9\% of the total emission. Johnson et al. (2021) estimated total reservoir emissions of 10.1 Tg CH\(_4\) yr\(^{-1}\) (7.2–12.9 Tg CH\(_4\) yr\(^{-1}\)) using a statistical upscaling approach, almost five times higher than our estimates, based on a data set of 42,000 reservoirs summing up to 2.97 \times 10^5 \text{km}^2. Johnson et al. (2021) also used HydroLAKES for areas and abundances but included more small reservoirs with a total area about 20\% greater than that in our modeling. Furthermore, to provide estimates of annual fluxes when the observational data did not include all seasons, Johnson et al. (2021) used two statistically fitted equations from previous studies to relate diffusive and ebullitive emissions to surface air temperature, respectively. Applying the equations developed for a few specific sites to the global lakes can lead to biases. Also, the relationship between air and water temperature is not well represented.

Our estimates of global lake and reservoir emissions are consistent with current global top-down estimates of natural emissions ranging from 21 to 47 Tg CH\(_4\) yr\(^{-1}\) excluding wetlands (Saunois et al., 2020). The top-down emission estimates optimize a variety of bottom-up estimates to be consistent with atmospheric CH\(_4\) observations. Global total microbial CH\(_4\) emissions, including emissions from natural wetlands, lakes, reservoirs and wild animals, and anthropogenic sources of livestock, waste/landfills, and rice cultivation, are further constrained by the stable carbon isotope of CH\(_4\) (delta-13CH\(_4\)), suggesting a total emission ranging from 360 to 388 Tg CH\(_4\) yr\(^{-1}\) for 2000–2009 (Lan et al., 2021). Wetland emissions from the top-down CH\(_4\) inversions are estimated to be in the range of 153–196 CH\(_4\) yr\(^{-1}\), while anthropogenic microbial emissions are in the range of 198–219 CH\(_4\) yr\(^{-1}\) for 2000–2009. If we assume other sink and source estimates are accurate, the lake and reservoirs emissions need to be small to match the atmospheric burden constrained by CH\(_4\) and delta-13CH\(_4\) observations. The large estimates of 70.9 Tg CH\(_4\) yr\(^{-1}\) by Rosentreter et al. (2021) and 122–159 Tg CH\(_4\) yr\(^{-1}\) by Saunois et al. (2020) for global freshwater methane emissions will need to have larger sink terms or smaller source terms in the global methane cycling to reconcile the atmospheric observations and modeled concentrations.

### 3.3. Warming Effects Under Future Scenarios

Future warming enhances microbial activities, contributing to increased methane emissions (Guo et al., 2020; Yvon-Durocher et al., 2014; Zhu et al., 2020). The estimated total emissions by the end of the 21st century are 44.6 ± 15.1 Tg CH\(_4\) yr\(^{-1}\) and 37.8 ± 11.2 Tg CH\(_4\) yr\(^{-1}\), increases by 86\% and 58\% compared to the 2000s, under the RCP8.5 using climate outputs from HadGEM2-ES and GFDL-ESM2M, respectively (Table 2 and Figure 3). Under both warming conditions, the pan-Arctic has the fastest increase, over 160\% within the century, followed by temperate (20\%–30\%) and tropical regions (3\%–20\%). The ratio of diffusive to ebullitive emissions decreases by 10\% and 5\% using HadGEM2-ES and GFDL-ESM2M, respectively. Ebullition is more sensitive to warming, as suggested in Aben et al. (2017), but underlying causes are uncertain.

Our estimated increase in global lake emissions tends to be high. A recent projection suggested that future warming of bottom waters will likely increase methane production by 13\%–40\% by the end of the century (Janssen et al., 2022), but they also indicated that many low-latitude lakes experience an increase of up to 17 times the historical (1970–1999) global average under RCP8.5. For the pan-Arctic, our estimate of total emissions is lower than 28.1 Tg CH\(_4\) yr\(^{-1}\) for the region north of 60°N by Tan et al. (2015). The higher estimate by Tan et al. (2015) is probably because they calibrated the model against five lakes, four of which are small thermokarst lakes whereas the less productive large glacial lakes cover most area in this region (Matveev et al., 2016; Walter Anthony et al., 2016; Wik et al., 2016). We predict a higher emission increase than the increase of 29\% estimated by Wik et al. (2016). This difference might be due to the differences in ice-free season between our study and Wik et al. (2016). Our previous study suggested that ice-free days play a role in rising methane emissions from boreal lakes (Guo et al., 2020). Furthermore, an increase from enhanced microbial production is included in our estimate. This reflects the importance of explicit representation of lake thermal and carbon dynamics for future projections.
In the model, the average summer (July, August, and September—JAS) methane production rate increased by 19% and 36% for the pan-Arctic, 25% and 40% for the temperate, and 8% and 22% for the tropics using GFDL-ESM2M and HadGEM2-ES, respectively. It has been found that methanogens are more climate sensitive than methanotrophs (Zhu et al., 2020), which is also shown in the calibrated $A_{\text{meth}}$ values which range within $1.7-5.5$ and $1.0-1.2$, respectively. However, the simulations show the opposite for the temperate and the tropics (Table 2). Based on modeled process rates, methane oxidation increased at a comparable rate to methane production for most of the lakes in these two regions (Figure 4). Based on Equation 3, this may be caused by higher dissolved oxygen concentrations, higher methane concentrations, or both. The oxygen concentration is likely to decrease because solubility decreases with increasing water temperature. Also, according to Woolway and Merchant (2019), global lake mixing regimes are shifting toward stronger and longer stratification under warming, leading to weaker mixing of oxygen to lower depths. Thus, increase in water column oxygen is unlikely the cause of higher methane oxidation rates while increasing methane concentration resulting from enhanced
production is more likely the major cause. Our modeling result is consistent with several experimental studies that measured process rates and feedbacks under controlled warming and found methanotrophy is enhanced similarly to increased methanogenesis (Lofton et al., 2014; Shelley et al., 2015).

### 3.4. Uncertainties in Permafrost Dynamics Impacts

The doubling in SOC results in a 56% increase of 5.5 Tg CH$_4$ yr$^{-1}$ in the total methane emissions. Diffusive and ebullitive fluxes account for 47% and 57% of the increase, respectively. By analyzing the increases of methane oxidation versus production rates with doubled SOC (same as Figure 4, not shown), we find a consistent pattern with the warming-induced methanogenesis increase. The extra SOC input by thawing permafrost is not fully translated to methane emission to the atmosphere, which again demonstrates that despite the growing methane source due to warming and old carbon input, methane oxidation in lake water can serve as an important sink that counteracts part of the climate impacts. In this scenario, the lake classification remained the same as the present because landscape evolution is not within the scope of this study.

Our modeled current total methane emission from thermokarst lakes larger than 0.1 km$^2$ is $2.3 \pm 0.7$ Tg CH$_4$ yr$^{-1}$. Wik et al. (2016) gave an estimate of $4.1 \pm 2.2$ Tg CH$_4$ yr$^{-1}$ from thermokarst lakes larger than 0.002 km$^2$ using statistical upscaling. If the difference between these results is due to lakes smaller than 0.1 km$^2$ with the same number of small lakes from these studies, the total emissions from thermokarst lakes becomes $2.9$ Tg CH$_4$ yr$^{-1}$. With the thermokarst lake area changing by $\pm 30\%$, a change of $\pm 1.65$ Tg CH$_4$ yr$^{-1}$ is calculated from the modeling. If lakes smaller than 0.1 km$^2$ are included, the total variation would be $\pm 2.1$ Tg CH$_4$ yr$^{-1}$, about 28% and 9% of the current pan-Arctic and global total, respectively. Therefore, the thermokarst lake area change can have an impact at the regional scale but has a small role in global lake methane emissions.

### 3.5. Research Limitations

There are several limitations in the study. First, difficulties in characterizing lake properties introduce uncertainties to our simulations. Classifying all lakes without water quality information as oligotrophic is likely to understate emissions. Sediment carbon stocks are heterogeneous, and a full representation would involve local carbon...
inputs and outputs, and simulation of the full carbon cycle that includes algae, plants and microbial activities, and carbon sedimentation and burial. These processes are seldom fully quantified, and parameter values are poorly constrained due to lack of measurements. Although the model performed well in both calibration and validation (Figure S2 in Supporting Information S1), using parameters derived from current states to make future projections is based on the assumption of parameter stationarity. Riley et al. (2011) noted this as a universal issue with prediction by process-based modeling. Furthermore, several algorithms, such as those used to model ebullition, gas exchange at the air-water interface and vertical mixing, can be improved.

Second, our predictions are significantly limited by the quality and quantity of observational data for diverse lakes. This uncertainty could be from our calibration for various processes of lake methane cycling including methane production, oxidation, and transport pathways. In addition, our calibration and verification are conducted primarily for individual lakes smaller than 20 km² with a few exceptions, while the global lake area includes large lakes. This will likely overestimate our regional and global simulations when applying parameters obtained from relatively small lakes to large lakes. Future analysis shall be based on more emission data for large lakes when they are available.

Third, lacking regional-scale measurements of permafrost thaw impacts on pan-Arctic lakes is a major uncertainty in predicting lake methane emissions by the end of the century, especially when over 90% of the permafrost is estimated to disappear under RCP8.5 by multiple climate models (Koven et al., 2013). Previous studies estimating the amount of thawed old carbon can shed lights on the impacts on lakes (Hugelius et al., 2020; Shurr et al., 2015), but to incorporate the process into models, lake-specific projections are needed. Site-level measurements have shown diverse changes in lake SOC input (Walter Anthony et al., 2016) and thus, it requires further understanding from the mechanistic aspect to gain more insights into the issue. Our sensitivity analysis with the doubled SOC input to lakes is a relatively simple treatment. More detailed data on lateral flow of SOC from land to aquatic ecosystems considering permafrost thaw and thermokarst will improve our future lake methane emission analysis.

Another important uncertainty is the global lake distribution data, especially for small lakes (<0.1 km²). These small lakes could be regenerated, expanded, or disappeared due to thawing permafrost and thermokarst in the Arctic and their emission rates are large. For instance, Johnson et al. (2021) using a new spatially explicit data set of lakes >50°N, included lakes with areas smaller than 0.1 km². Matthews et al. (2020) estimated that the area of lakes north of 50°N is 10.95 × 10⁵ km². While our estimates are based on HydroLAKES data with lake and water reservoir area of 2.49 × 10⁵ km² only. Thus, including small lakes' areal dynamics should improve our future regional estimates.

Finally, the water column methanotrophs are simulated for three water layers and their significant role in affecting the lake methane emissions are modeled. However, the methanotroph parameterization is compromised because it is not based on observed methane oxidation rate, rather based on the net lake emission data. Methane oxidation rates could vary significantly from 0.25 to 81 mg of C m⁻² day⁻¹ for different lakes (Bastviken et al., 2002). Additionally, oxidation rates at different lake depths within the same lake could also vary (Langenegger et al., 2022). Thus, parameterizing the model with depth oxidation profiles for various lakes will constrain our model uncertainty.

4. Conclusions

We estimate that global methane emissions from lakes and reservoirs larger than 0.1 km² using an improved modeling approach are 24.0 ± 8.4 Tg CH₄ yr⁻¹ for the present, which is lower than previous statistical upscaling results. Under the RCP8.5 warming scenario, our sensitivity analysis suggests the global total emissions will increase by 58%–86% with an assumption that the carbon substrate for lake methanogenesis is doubled. We further investigate the emission sensitivity to different factors, including warming, carbon input from thawed permafrost, and changes in thermokarst lake areas, on lake methane emissions under changing climate conditions. We find that, although carbon from thawed permafrost can increase methane production, methane oxidation would also be enhanced due to increasing methane concentration. Our hypothetical analysis on thermokarst areal changes, either shrinking or expanding, suggests they exert a minor impact on global total methane emissions.

Our study suggests that explicitly incorporating processes related to lake thermal and carbon dynamics (e.g., lateral flow of carbon from land to aquatic ecosystems) is important to future projections while developing and modeling lake distribution dynamics is a research priority.
Data Availability Statement

Research data generated in this study are available via doi: https://doi.org/10.4231/JZJ10-FH54 and the model code is available via doi: https://doi.org/10.4231/67YG-V518 in Purdue University Research Repository.

References


