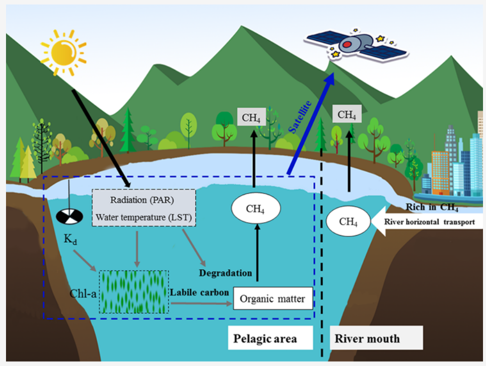
Environmental Science & Technology
pubs.acs.org/est
Article

Quantification of Diffusive Methane Emissions from a Large Eutrophic Lake with Satellite Imagery

Hongtao Duan,* Qitao Xiao,* Tianci Qi, Cheng Hu, Mi Zhang, Ming Shen, Zhenghua Hu, Wei Wang, Wei Xiao, Yinguo Qiu, Juhua Luo, and Xuhui Lee

Cite This: *Environ. Sci. Technol.* 2023, 57, 13520−13529
Read Online
ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Lakes are major emitters of methane (CH_4); however, a longstanding challenge with quantifying the magnitude of emissions remains as a result of large spatial and temporal variability. This study was designed to address the issue using satellite remote sensing with the advantages of spatial coverage and temporal resolution. Using Aqua/MODIS imagery (2003−2020) and *in situ* measured data (2011−2017) in eutrophic Lake Taihu, we compared the performance of eight machine learning models to predict diffusive CH_4 emissions and found that the random forest (RF) model achieved the best fitting accuracy (R^2 = 0.65 and mean relative error = 21%). On the basis of input satellite variables (chlorophyll *a*, water surface temperature, diffuse attenuation coefficient, and photosynthetically active radiation), we assessed how and why they help predict the CH_4 emissions with the RF model. Overall, these variables mechanistically controlled the emissions, leading to the model capturing well the variability of diffusive CH_4 emissions from the lake. Additionally, we found climate warming and associated algal blooms boosted the long-term increase in the emissions via reconstructing historical (2003−2020) daily time series of CH_4 emissions. This study demonstrates the great potential of satellites to map lake CH_4 emissions by providing spatiotemporal continuous data, with new and timely insights into accurately understanding the magnitude of aquatic greenhouse gas emissions.



KEYWORDS: *inland lakes, algal blooms, CH_4 emissions, spatial−temporal variability, satellite estimation*

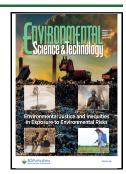
1. INTRODUCTION

The current increase of the atmospheric methane (CH_4) concentration is motivating research to investigate the magnitude of CH_4 emissions from natural sources.[1−3] Lakes produce a large amount of CH_4 via anaerobic degradation of organic carbon and, thus, are key natural emitters of CH_4.[1,4,5] Global lakes are suffering eutrophication with algal blooms as a result of nutrient enrichment and climate warming,[6−8] which can enhance CH_4 production and emissions as a result of large algae-derived organic carbon.[5,9,10] Thus, quantifying the magnitude of CH_4 emissions from lakes has become a top priority in the natural CH_4 budget estimate.[11−13]

While several studies were dedicated to quantifying global lake CH_4 emissions, the current estimates are still poorly constrained.[4,14] It is estimated that global lakes emit 8−200 Tg of CH_4 year^{-1} (diffusion plus ebullition) to the atmosphere.[1,3,15] The diffusive CH_4 emissions alone spans more than 1 order of magnitude.[13,15] These estimates are mostly based on traditional analysis via scaling limited local/regional measurements across the sample population to the globe without being spatially and temporally explicit. However, evidence showed that CH_4 emissions from lakes varied considerably across time and regions as a result of large differences in lake properties.[4,16,17] Even CH_4 emissions can vary greatly within a single lake as a result of highly heterogeneous environmental variables.[10,18,19] Unfortunately, the majority of CH_4 emission measurements were conducted with limited time coverage or low time resolution, likely leading to large uncertainty in lake CH_4 emission estimation.

Lake environmental variables have a significant impact on CH_4 emissions mechanistically, therefore, which are often estimated by various models via relating emissions to lake variables.[11,13,15] For example, lake CH_4 emissions at regional to global scales can be estimated by the water temperature,[20,21] primary production or eutrophic status,[11,22] and surface area.[13,16] However, a longstanding problem remains with the proposal as a result of the current limited availability of spatially and temporally explicit data sets on lake properties.[4,15] Satellite remote sensing can map dynamic lake properties with acceptable accuracy at large spatial coverage and high temporal resolution,[23−25] providing a powerful tool in obtaining

Received: July 16, 2023
Revised: August 20, 2023
Accepted: August 21, 2023
Published: August 31, 2023



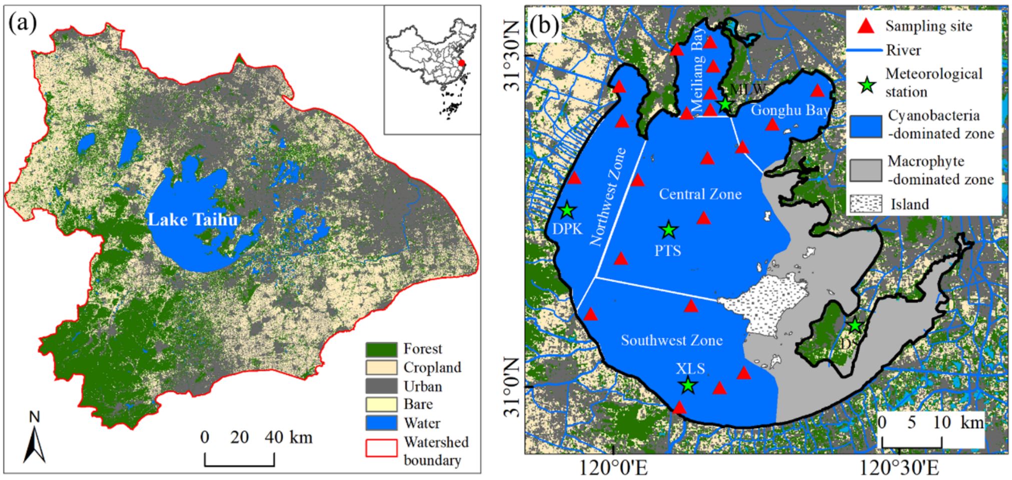


Figure 1. Map showing the geographic location of study sites in the cyanobacteria-dominated region of Lake Taihu. Note that the study region was divided into five subzones (Northwest Zone, Meiliang Bay, Gonghu Bay, Central Zone, and Southwest Zone). These abbreviations of MLW, DPK, PTS, XLS, and DS, which disturbed in the subzones of the lake, indicated the meteorological station.

unbiased CH_4 emission estimation.[11,26] Meantime, machine learning algorithms have been used successfully in carbon gas emission prediction of inland waters;[27−29] specifically, better performance and adaptation in estimating inland water CH_4 emissions with a machine learning algorithm have been exhibited in comparison to traditional statistical methods (e.g., linear regression).[30] This provides a more powerful tool and greater potential for lake CH_4 emission mapping using satellite imagery.

Lake Taihu, the third largest freshwater lake in China, is a typical eutrophic lake with algal blooms. Frequent algal blooms and associated eco-environmental issues have made Lake Taihu a popular research site for remote sensing[31,32] and CH_4 emission studies[10,33] at a long time. In the attempt to better understand diffusive CH_4 emissions from eutrophic lakes, we first developed a machine learning model to obtain more temporal and spatial coverage of diffusive CH_4 emissions from satellite data in the lake. There are numerous satellite systems providing key eutrophic lake properties (e.g., algal blooms, chlorophyll *a*, water clarity, and dissolved organic matter);[23,25,34] moderate-resolution imaging spectroradiometer (MODIS) data have the advantage with a long data set and daily revisit periods.[35]

The main objectives of our study were (1) to develop machine-learning-approach-based models of diffusive CH_4 emissions in Lake Taihu from available MODIS/Aqua images and ground data, (2) to establish a long-term time series of CH_4 emissions from satellite data (2003−2020) with high spatial and temporal coverage, and (3) to evaluate the response of CH_4 emissions to lacustrine environmental changes. This study, to the best of our knowledge, is the first to reconstruct a long-term record of diffusive CH_4 emissions from a eutrophic lake via satellite remote sensing, providing new insight into obtaining large-scale lake diffusive CH_4 emissions with daily time scales.

2. MATERIALS AND METHODS

2.1. Study Site and *In Situ* Data Set. Lake Taihu, with a mean depth of 1.9 m and surface area of 2338 km^2, is located in the developed Yangtze River Delta of China. There are 172 rivers entering into the lake,[36] and the local climate is a subtropical climate with high water temperatures in the summer and low temperatures in the winter.[33] Large external nutrient loading stimulates frequent algal blooms in Lake Taihu.[37] The lake is divided into Meiliang Bay, Northwest Zone, Gonghu Bay, Central Zone, Southwest Zone, and vegetation-dominated zone (Figure 1) according to nutrient concentrations and aquatic vegetation distribution.[10,33,38] The spatial differences in the chlorophyll *a* and nutrient concentrations also support the division,[37,39] and more information is shown in Text S1 of the Supporting Information.

From September 2011 to November 2017, field measurements were carried out to determine diffusive CH_4 emissions of eutrophic Lake Taihu. Water samples in each site of Meiliang Bay, Northwest Zone, Gonghu Bay, and Central Zone were collected monthly to obtain the CH_4 emissions before August 2013, and then, water samples were collected seasonally after August 2013. Field surveys in the Southwest Zone were conducted seasonally from 2011 to 2017. Sampling and calculation for the CH_4 emissions used in this study were detailed in the study of Xiao et al.,[10] which are also presented in Text S2 of the Supporting Information.

2.2. MODIS-Derived Data Acquisition and Processing. Satellite data collected by MODIS/Aqua with a maximum spatial resolution of 250 m (bands 1 and 2) and a very short revisit interval (1 image/day) were used to obtain sufficient spatial−temporal coverage of Lake Taihu (Text S3 of the Supporting Information). We used 1257 high-quality cloud-free MODIS/Aqua images in this study (Table S1 of the Supporting Information). MODIS-derived environmental variables, including the chlorophyll *a* concentration (Chl *a*), lake surface temperature (LST), diffuse attenuation coefficient (K_d), and photosynthetically active radiation (PAR), were used as inputs to estimate the diffusive CH_4 emissions. These variables controlled the CH_4 emissions from lakes, and they are often used as either a proxy for CH_4 emission estimation[11,15] or as solution to biased field measurements of emissions.[20,40]

Chl *a* and LST are important predictors for CH_4 emission in lakes, especially for eutrophic lakes, mostly as a result of their effects on organic matter decomposition.[4,10,11] K_d and PAR characterize the process of light penetration and heat transfer in a lake,[41] which was associated with CH_4 variability. Wind speed influences the CH_4 exchange velocity across the lake−air

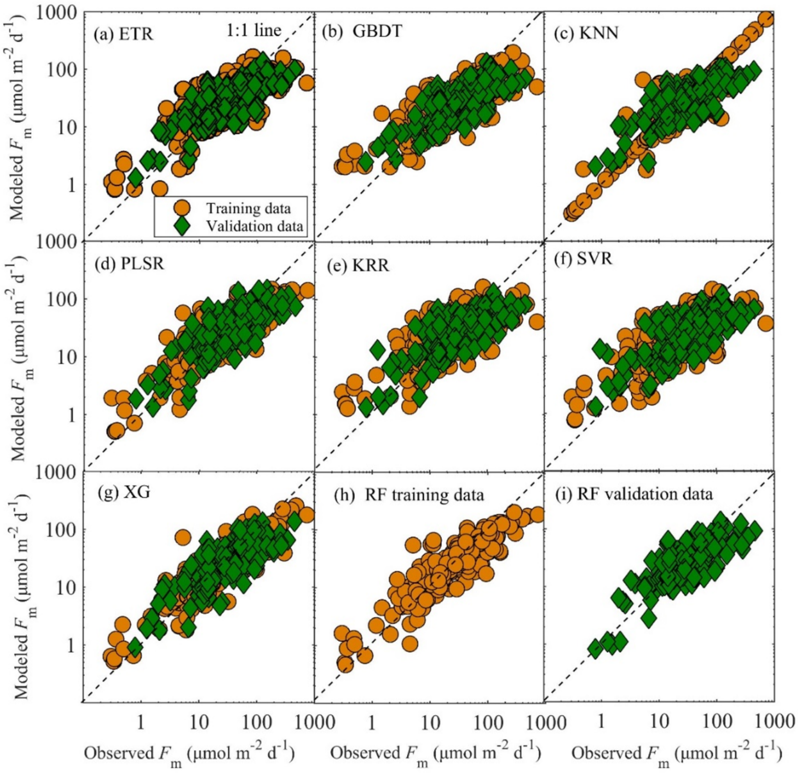


Figure 2. Performance evaluation of the diffusive CH_4 flux (F_m) estimation using the machine learning algorithms (e.g., ETR, GBDT, KNN, PLSR, KRR, SVR, XG, and RF) and performance of the final RF algorithm development and validation are shown separately. Training and validation results (e.g., R^2, MRE, $RMSE_{log}$, and RRMSE) obtained for the CH_4 flux using different machine learning algorithms are shown in Table S3 of the Supporting Information.

interface and could also act as a proxy of CH_4 exchange velocity, which had been demonstrated in our previous study.[10] Thus, wind speed was also used as an input variable. It was worth noting that the diffusive CH_4 emissions was governed by the dissolved CH_4 concentration of the lake (Figure S1 of the Supporting Information). The details of these input variable calculations/acquisitions are described by Qi et al.[35] Briefly, Shi et al.[32] and Huang et al.[41] have developed the algorithm for Chl *a* and K_d at 490 nm calculations in Lake Taihu, respectively, which were used in this study.

2.3. Model Development. Machine learning algorithms are used in this study as a result of their prominent advantage in data mining. A flowchart was presented to show the satellite data process and associated CH_4 emission estimation model development (Figure S2 of the Supporting Information). To obtain the best performing model to predict the diffusive CH_4 emissions, we compared eight popular machine learning algorithms with different complexity. Extra trees regression (ETR), gradient boosting decision tree (GBDT), *K*-nearest neighbor (KNN), partial least squares regression (PLSR), kernel ridge regression (KRR), support vector regression (SVR), XGBoost (XG), and random forest (RF) were explored to construct the predicting model. Before model training, the response variable CH_4 emission was $\log_{10}$-transformed to ensure the normality of the distribution. Each input variable was standardized by subtracting the mean and dividing the standard deviation. The eight machine learning models were tested using the same input variables. The important parameters to define the structure and tuning of these models were shown in Table S2 of the Supporting Information. All of the scripts for developing these models were written in Python and are available at https://github.com/qitaolake/Diffusive-CH4.

For model training and validation, field-measured CH_4 emissions and satellite-derived variables were paired first. In total, there are 662 satellite-to-ground synchronization matchup data, which were separated into two groups (the training data sets and independent validation data sets); more details are shown in Text S4 of the Supporting Information. Standard statistical metrics, including the coefficient of determination (R^2), mean relative error (MRE), root-mean-square error in log form ($RMSE_{log}$), and relative root-mean-square error (RRMSE) were used to quantify the accuracy of the estimated diffusive CH_4 emissions (Text S4 of the Supporting Information). To better understand how the modeled CH_4 emissions respond to the uncertainty of satellite-derived input independent variables, a model sensitivity analysis was carried out. The analysis incorporated the uncertainty of estimated diffusive CH_4 emissions with the satellite date by selecting from a normal distribution with the same mean absolute relative error as the corresponding input variables. The uncertainty for each input variable was obtained from previous studies.[32,35,41] Individual partial dependence plots are useful to determine whether the model is appropriate; each partial dependence plot is carried out by varying the selected input variables across the range and keeping others predictors constant.

3. RESULTS

3.1. Model Comparison and Selection. To predict diffusive CH_4 emissions from Lake Taihu, the performance of eight machine learning models was compared (Figure 2 and

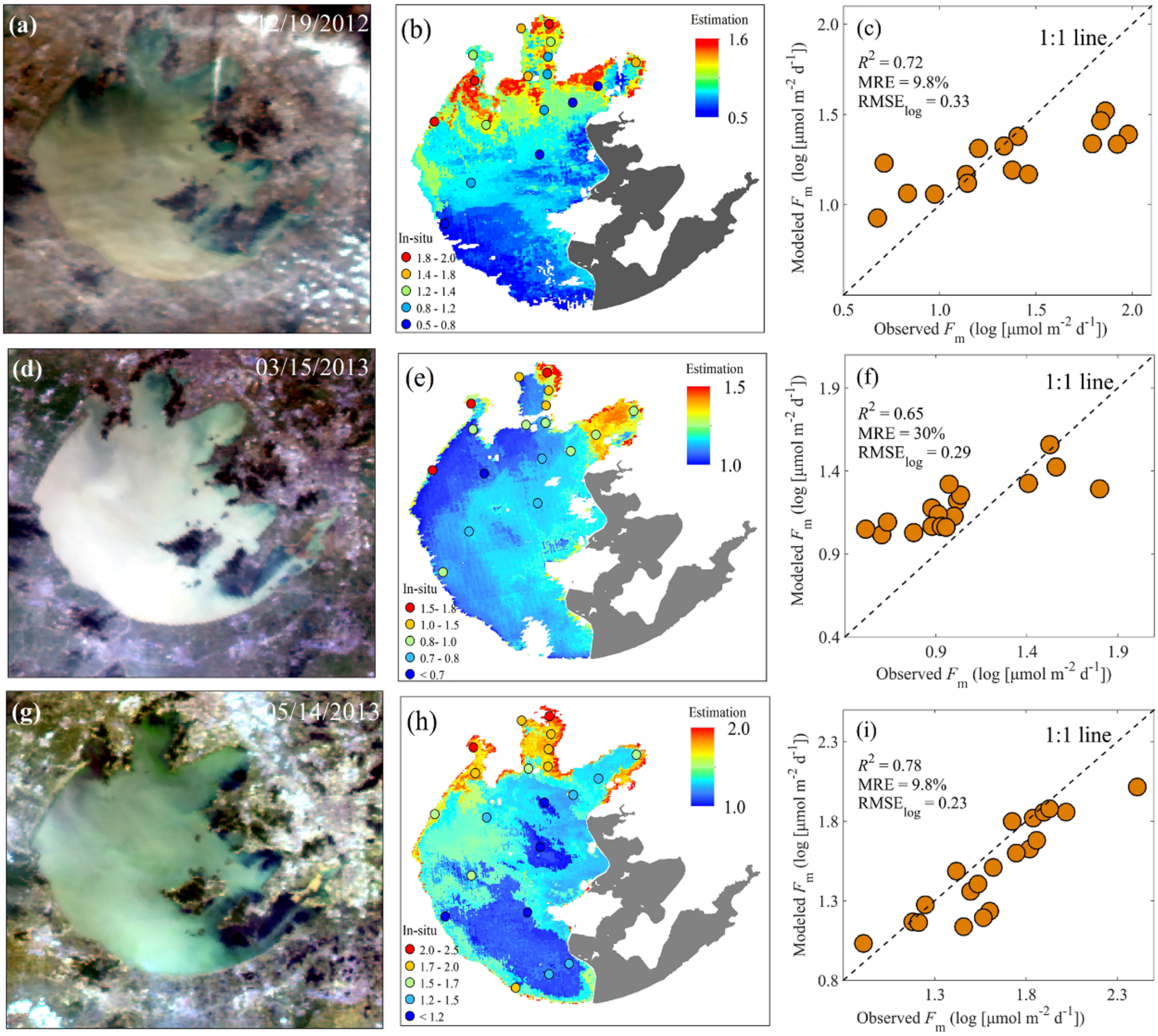


Figure 3. MODIS-derived diffusive CH_4 emissions (F_m) compared to *in situ* measured emissions on (a−c) December 19, 2012, (d−f) March 15, 2013, and (g−i) May 14, 2013. Note that sometimes field measurements were completed in 2 consecutive day; the spatial distribution of modeled F_m was obtained on the basis of pixel-by-pixel estimations [log (μmol m^{-2} day^{-1})] with the same spatial resolution of MODIS. The first column was the true color composite image; the second column was satellite-estimated F_m with *in situ* measured F_m overlaid as circle points; and the third column was the comparison between *in situ* measured and satellite-estimated F_m.

Table S3 of the Supporting Information). Overall, these models showed good fitting accuracy, especially in the pelagic area. It should be noted that most of the input variables were correlated with each other (Table S4 of the Supporting Information), but the model (e.g., RF) performance significantly decreases (Table S5 of the Supporting Information) if the water temperature and chlorophyll *a* were removed.

Using the same training and validation data sets from the pelagic area, the eight machine learning models for diffusive CH_4 emission were tested. Results showed that KNN, PLSR, RF, and XG had better performance in the training results (R^2 = 0.75−0.87, MRE = 7−21%, $RMSE_{log}$ = 0.20−0.27, and RRMSE = 21−26%). Meanwhile, the RF model had the best fitting accuracy (R^2 = 0.65, MRE = 21%, $RMSE_{log}$ = 0.33, and RRMSE = 28%; panels h and i of Figure 2) in the validation data, suggesting that the RF model had the highest stability among them. Most of the data pairs of *in situ* measured CH_4 emissions and modeled CH_4 emissions clustered along the 1:1 line, with few outliers in the RF model trained and tuned on the training data set. Thus, RF led to better predictions.

The trained RF model can well predict the variability of the lake diffusive CH_4 emissions under different seasons, which was evidenced from some examples with available satellite images (Figure 3). Comparisons of the model to *in situ* field measurements at the specific locations showed agreement with the CH_4 emission variability. The MODIS-derived CH_4 emissions are consistent with the *in situ* field measurements across time and space,[10] showing that the approach captured well the spatial and temporal variability of CH_4 emissions.

3.2. Temporal and Interannual Variations of Satellite-Estimated CH_4 Emissions. The long-term (2003−2020) time series of daily diffusive CH_4 emissions was modeled on the basis of available satellite images. The modeled CH_4 emissions are highly consistent with *in situ* measured CH_4 emission flux, which was seasonally variable, with peaks occurring in the summer and minima in the winter (Figure

https://doi.org/10.1021/acs.est.3c05631
Environ. Sci. Technol. 2023, 57, 13520−13529

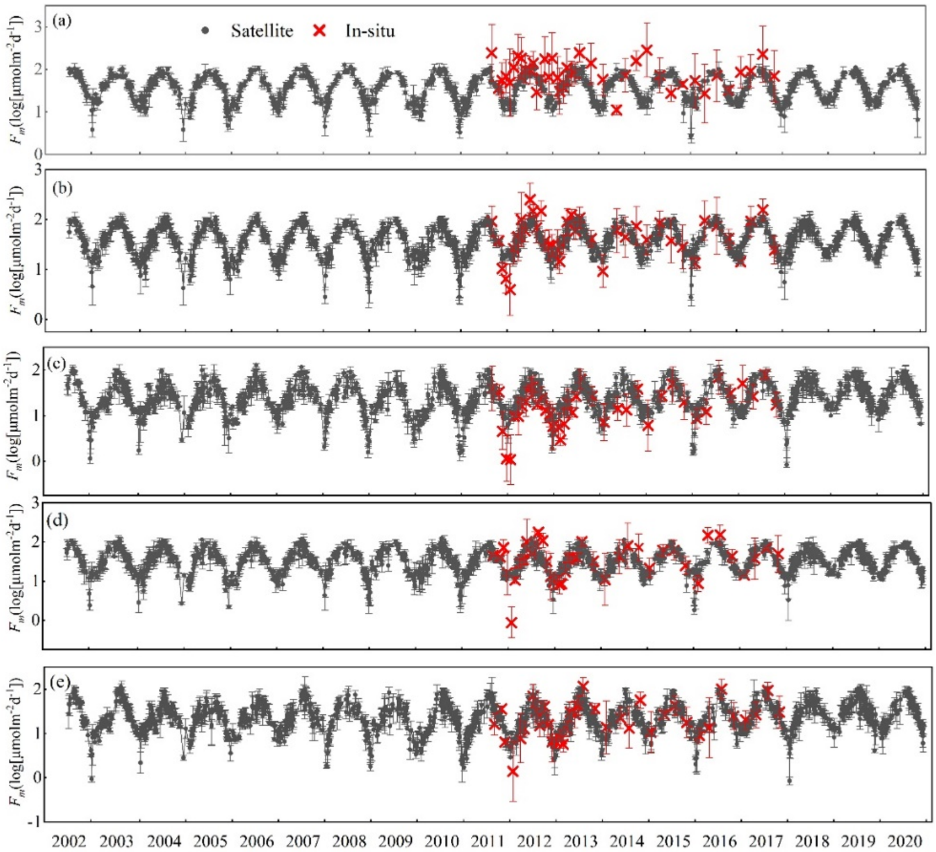


Figure 4. Time series of the MODIS-estimated (black circles, 2002−2020) and *in situ* measured (red crosses, 2011−2017) diffusive CH_4 emission flux (F_m) at (a) Northwest Zone, (b) Meiliang Bay, (c) Gonghu Bay, (d) Central Zone, and (e) Southwest Zone of Lake Taihu. The error bars indicated standard deviations. The zonal mean value was calculated using measurements made at the sites within that subzone.

4). The satellite-to-ground synchronization matchup data (*in situ* measured versus satellite estimated) were also highly correlated in Meiliang Bay (R^2 = 0.86 and p < 0.01), Gonghu Bay (R^2 = 0.60 and p < 0.01), Central Zone (R^2 = 0.71 and p < 0.01), and Southwest Zone (R^2 = 0.66 and p < 0.01). However, poor correlation was found in the Northwest Zone (R^2 = 0.06 and p > 0.01) with large external loadings input via river discharge. This finding indicates that the model can well capture the temporal variability in diffusive CH_4 emissions, except that in the subregion (Northwest Zone) with large external loadings. However, it should be noted the hot spots of the subregion (Northwest Zone) in the emissions were successfully predicted by the model (Figures 3 and 4).

Monthly and annual mean diffusive CH_4 emissions in the lake are also presented (Figure 5). Overall, Meiliang Bay and Northwest Zone were "hot spots" for CH_4 emission, and CH_4 emissions were generally higher than that in the Central Zone. The satellite-derived diffusive CH_4 emissions also showed interannual variations from 2003 to 2020 (Figure 5). The annual CH_4 emissions flux in Meiliang Bay and Central Zone both showed a significantly increasing trend (Meiliang Bay, R^2 = 0.47 and p < 0.01; Central Zone, R^2 = 0.42 and p < 0.01). The lake flux was calculated as the area-weighted mean of the five subzonal fluxes, which also exhibited an obvious increasing trend between 2003 and 2020 (R^2 = 0.33 and p < 0.05). The annual CH_4 emission during the algal bloom period (May−July) exhibited a stronger increasing trend (R^2 = 0.56 and p < 0.05; Figure 5).

4. DISCUSSION

4.1. CH_4 Emission Estimation and Environmental Controls.

This study aimed to predict the CH_4 emissions by combining satellite data and a machine-learning-based model. Satellite remote sensing can retrieve lake variables at high spatial−temporal resolutions.[23,24,32] Machine learning is well-known for its greater advantage in coping with complex regression, which might represent optimal options for CH_4 emission prediction. Sure enough, the RF model achieved acceptable accuracy ($RMSE_{log}$ = 0.33 and R^2 = 0.65; Table S3 of the Supporting Information). Results showed that the model captured well the spatial−temporal variability of CH_4 emissions (Figures 4 and 5), because the variability of modeled emissions was consistent with the *in situ* field measurements, as shown in a previous study.[10] As shown in other studies,[27,42] the RF model was a suitable model for carbon gas emission prediction in an aquatic ecosystem.

A sensitivity analysis showed that the model was sensitive to Chl *a* and water temperature estimations (Table S6 of the Supporting Information). The conservative larger uncertainties of 30% for Chl *a* and 15% for water temperature led to the MRE of CH_4 emission estimation being 8.19 and 10.12%, respectively. Sensitivity analysis also showed that Chl *a* and water temperature containing variable combinations always had a large MRE. The results of sensitivity analysis are expected, because Chl *a* and water temperature are two important variables in determining the CH_4 emissions from lakes, as shown in previous studies.[11,20,43]

https://doi.org/10.1021/acs.est.3c05631
Environ. Sci. Technol. 2023, 57, 13520−13529

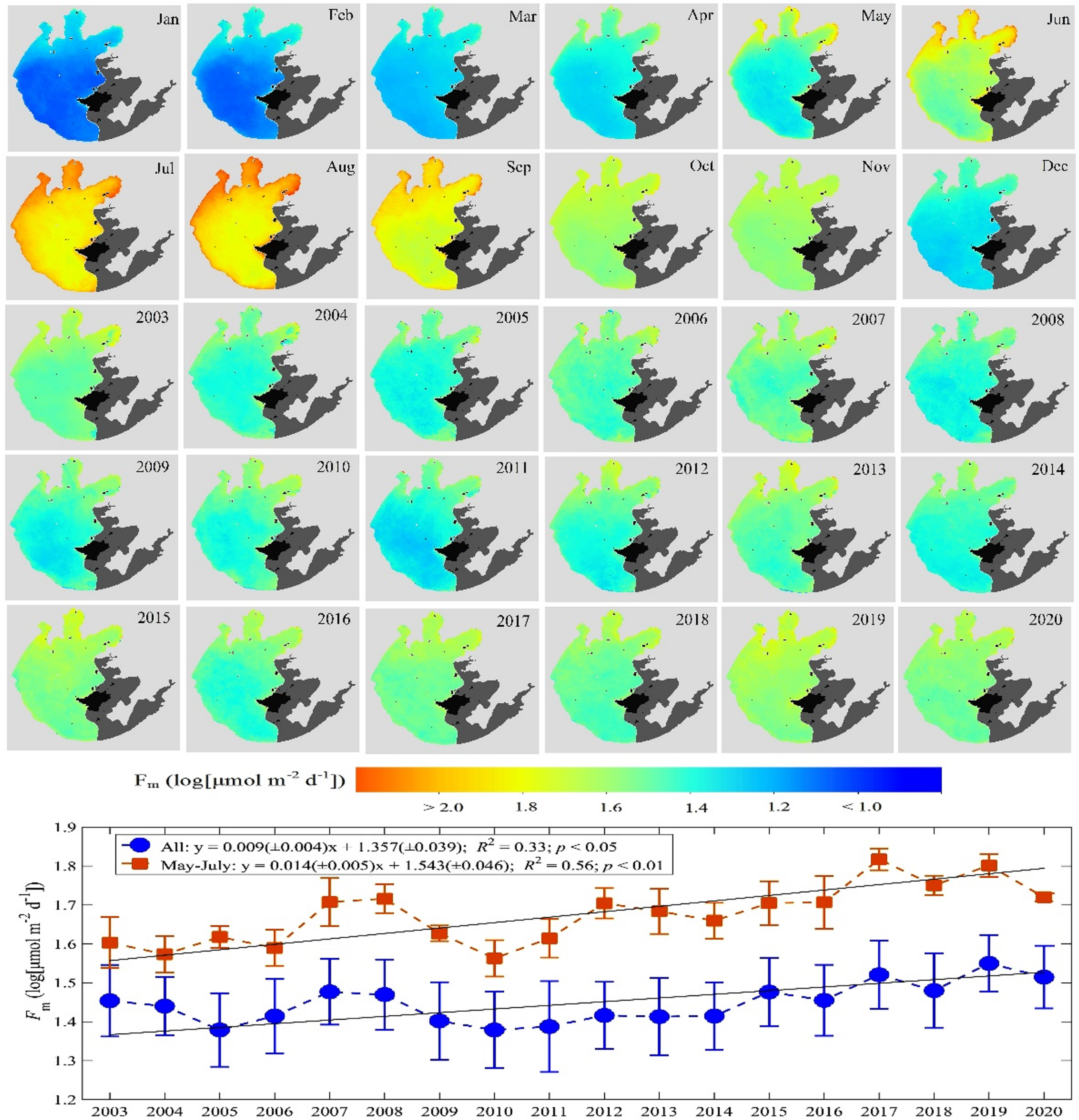


Figure 5. MODIS-estimated monthly mean diffusive CH_4 flux (F_m) from January to December and annual mean diffusive F_m from 2003 to 2020. The dark gray zone (island) and light gray zone (macrophyte-dominated zone) were excluded. Pixel-by-pixel CH_4 emission predictions were conducted, and the long-term trends of whole-lake F_m from 2003 to 2020 during all seasons and the algal bloom period (May–July) were shown in the bottom. The error bars indicated standard deviations.

Partial dependence plots showed that the CH_4 emissions varied as expected (Figure S4 of the Supporting Information). Specifically, the CH_4 emissions increase as Chl *a* and water temperature increase, as expected for increased CH_4 production rates. The majority of CH_4 in lakes are produced by organic carbon decomposition.[44] High Chl *a* indicated algal blooms with abundant phytoplankton biomass, which would fuel labile carbon for CH_4 production.[9,45,46] The organic carbon decomposition and associated CH_4 production are temperature-dependent,[21,47] which led to the CH_4 emissions to increase considerably with evaluating the temperature. Importantly, the input variables (Chl *a* and water temperature) have generally been used as a proxy for lake CH_4 emission estimations in previous studies.[11,20] PAR and K_d can influence the emissions via indirect ways of controlling the water temperature and algal blooms.[41] For example, the water temperature increases with high PAR and leads to the expected increase in emissions, and high K_d can hinder light and heat into the lake water and then suppress CH_4 production and emissions.

https://doi.org/10.1021/acs.est.3c05631
Environ. Sci. Technol. 2023, 57, 13520−13529

The CH_4 emissions can be estimated by a satellite, and estimated emissions varied spatially with hot spots in the Northwest Zone (Figure 5). However, the temporal variation of emissions was not well-captured in the Northwest Zone (Figures 3 and 4). The Northwest zone received large external loading input via river discharge (Figure 1).[37] Previous studies have found that CH_4 emissions, especially in the river inflowing zone of a lake, were driven by external loading input,[18,48,49] showing less dependency upon the input variables of our estimation model.[10,22] Indeed, the seasonal CH_4 emissions in the river inflowing zone of the Northwest Zone were unrelated to Chl *a* but were controlled by dissolved CH_4 concentrations of inflowing rivers (Figure S3 of the Supporting Information). Meanwhile, a satellite could not predict the CH_4 emission variability at the river mouth with a small area as a result of the coarse resolution of MODIS (~250 m). Although hot spots of the subregion (Northwest Zone) in the emissions were successfully predicted by the model, further studies should consider the external loading input from the catchment (e.g., river CH_4 input) to better predict the temporal variability of lake CH_4 emissions.[46,48,49]

Caution should be taken when considering the model section. First, all of these machine learning approaches are commonly characterized to be able to address complex questions, and the data processing reported here was not necessary for each machine learning tool.[50] Here, the data processing aimed at reducing the effects of outliers, which may be more suitable to some machine learning, showing sensitivity to data that span multiple orders of magnitude,[51] such as our case of CH_4 emission prediction. Second, each machine learning approach has its advantages and disadvantages, and there are different kernel functions for parametrization of the models. One obvious advantage of the RF-based machine learning is that it can reflect the nonlinearity between the responsive variable and predictive variables without clearly knowing their functional dependence.[51] However, the RF-based machine learning can be replaced with other algorithms, such as SVR, which is sensitive to variables that span multiple orders of magnitude.[51]

4.2. Long-Term Trend of Diffusive CH_4 Flux and Its Potential Drivers. Quantifying the long-term variations in diffusive CH_4 emissions is essential for predicting the response of emissions to future environmental changes. Many studies have proposed that aquatic CH_4 emissions can vary considerably at interannual time scales as a result of the changes in lacustrine environmental variables,[48,52,53] but only a few long-term data series (more than 10 years) have been reported to date,[22] posing a challenge to better understate the contribution of lakes to the global CH_4 budget. To address this challenge, we reconstructed the long-term (2003−2020) diffusive CH_4 emissions of Lake Taihu with fine spatial and temporal resolutions, using satellite data for the first time. Overall, our results revealed that the annual CH_4 flux increased, which was more significant during the algal bloom period in the eutrophic lake (Figure 5).

Previous studies implied that warming would stimulate CH_4 emissions from lakes.[17,47,53] Significant warming with 0.35 °C per decade was recorded over the lake,[45] which likely accounts for the increase in annual CH_4 emissions on longer time scales. Notably, measured Chl *a* was significantly correlated with emissions (Figure S5 of the Supporting Information), which may explain the increasing trend why it was more significant during the algal bloom period. It should be noted that an increasing temperature contributed to the algal blooms of the lake based on the long-term measurements,[31,45] which, in turn, increased the CH_4 emissions.[11] A long-term warming experiment in mesocosms showed that a rising temperature can significantly increase aquatic CH_4 emissions.[53] Thus, warming together with increasing eutrophication probably contributes to the increase in the diffusive CH_4 emissions during the past 2 decades.

4.3. Implication of the Lake CH_4 Budget. Field measurements suffer from large uncertainties associated with spatial and temporal variation in CH_4 emissions, which is highly relevant for estimating the CH_4 budget of lakes. Although lake CH_4 emission varied markedly on a seasonal basis,[21] a review showed that 80% of the study in northern lakes measures diffusive CH_4 emission with a short time and mostly in the summer, leading to large uncertainties.[17,40] To reduce sampling bias and resolve seasonal variability, previous studies suggest that diffusive CH_4 emission should be measured over 11−22 days scattered throughout the ice-free period.[20,40] Our approach obtained emissions on at least 51 days scattered in a year cycle based on the distribution of the available satellite images (Table S1 of the Supporting Information). Thus, this approach has the potential to fill gaps in sampling bias associated with temporal variation in CH_4 emissions.

Lake CH_4 emission estimates generally rely on extrapolations from a limited number of sites.[16,18,19] The distribution of CH_4 and environmental constituents is unevenly distributed,[19,54,55] likely leading to the whole lake estimation with large uncertainty and/or inaccurate effluxes. Our approach with satellite images may act as a solution to biased sampling associated with spatial resolution, because previous studies suggest sampling with high spatial resolution is needed to obtain reliable lake CH_4 emissions. For example, a study in small lakes recommended measurements should be made at at least three locations,[40] and increasing sampling sites can reduce the uncertainty of diffusive emission estimation.[18] Fortunately, each pixel of the MODIS image with a spatial resolution of 250 × 250 m^2 is equivalent to setting a sample site in our study, which has far exceeded the requirements for the number of sampling sites.

On the basis of the satellite-estimated daily CH_4 emission flux, the effects of sampling scenarios on CH_4 emission estimations were evaluated via bootstrap analysis (Text S5 of the Supporting Information). The standard deviation of estimated diffusive CH_4 emissions decreased significantly with increasing numbers of sampling days (Figure S6 of the Supporting Information). On the basis of the study by Wik et al.,[40] we simulated hypothetical sampling scenarios (Text S6 of the Supporting Information) and found that measurements should be made at least 30 days scattered throughout the year to achieve unbiased CH_4 emission estimation. Seasonal or monthly measurements were carried out generally to determine the annual mean diffusive CH_4 emissions;[10,18,19] our hypothetical sampling scenarios showed that the measurements yielded high risks (up to 78%) of over- and underestimating of the flux potential of the lake (Figure S7 of the Supporting Information). Cautions should be taken because the over- or underestimation may depend upon the lake size and heterogeneity.

https://doi.org/10.1021/acs.est.3c05631
Environ. Sci. Technol. 2023, 57, 13520−13529

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.3c05631.

Detailed information on methods, figures showing the flow char, relationship between CH_4 and Chl *a*, partial dependence plots, and effects of sampling days on CH_4 estimation, and tables showing the temporal distribution of available satellite images, candidate model structure and performances, correlation between the input variables, and model sensitivity analysis (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Hongtao Duan − *Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, Jiangsu 210008, People's Republic of China; University of Chinese Academy of Sciences, Nanjing, Jiangsu 211135, People's Republic of China*; orcid.org/0000-0002-1985-2292; Email: htduan@niglas.ac.cn

Qitao Xiao − *Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, Jiangsu 210008, People's Republic of China; Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science & Technology, Nanjing, Jiangsu 210044, People's Republic of China*; orcid.org/0000-0003-4255-3183; Phone: +86-25-86882161; Email: qtxiao@niglas.ac.cn

Authors

Tianci Qi − *Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, Jiangsu 210008, People's Republic of China*

Cheng Hu − *College of Biology and the Environment, Joint Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing, Jiangsu 210037, People's Republic of China*

Mi Zhang − *Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science & Technology, Nanjing, Jiangsu 210044, People's Republic of China*

Ming Shen − *Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, Jiangsu 210008, People's Republic of China*

Zhenghua Hu − *Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science & Technology, Nanjing, Jiangsu 210044, People's Republic of China*

Wei Wang − *Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science & Technology, Nanjing, Jiangsu 210044, People's Republic of China*

Wei Xiao − *Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science & Technology, Nanjing, Jiangsu 210044, People's Republic of China*

Yinguo Qiu − *Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, Jiangsu 210008, People's Republic of China*

Juhua Luo − *Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, Jiangsu 210008, People's Republic of China*

Xuhui Lee − *School of the Environment, Yale University, New Haven, Connecticut 06511, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.3c05631

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The study was supported jointly by the National Natural Science Foundation of China (42271114, 42271377, 41971309, and 42101433), the Provincial Science and Technology Innovative Program for Carbon Peak and Carbon Neutrality of Jiangsu of China (BK20220018 and BE2022422), the National Key R&D Program of China (2020YFA0607501), the Provincial Natural Science Foundation of Jiangsu of China (BK20220055 and BK20201100), and the Youth Innovation Promotion Association of Chinese Academy of Sciences (CAS) (2023329).

■ REFERENCES

(1) Rosentreter, J. A.; Borges, A. V.; Deemer, B. R.; Holgerson, M. A.; Liu, S.; Song, C.; Melack, J.; Raymond, P. A.; Duarte, C. M.; Allen, G. H.; Olefeldt, D.; Poulter, B.; Battin, T. I.; Eyre, B. D. Half of global methane emissions come from highly variable aquatic ecosystem sources. *Nat. Geosci.* **2021**, *14* (4), 225−230.

(2) Borges, A. V.; Deirmendjian, L.; Bouillon, S.; Okello, W.; Lambert, T.; Roland, F. A. E.; Razanamahandry, V. F.; Voarintsoa, N. R. G.; Darchambeau, F.; Kimirei, I. A.; Descy, J.-P.; Allen, G. H.; Morana, C. Greenhouse gas emissions from African lakes are no longer a blind spot. *Sci. Adv.* **2022**, *8* (25), No. eabi8716.

(3) Bastviken, D.; Tranvik, L. J.; Downing, J. A.; Crill, P. M.; Enrich-Prast, A. Freshwater methane emissions offset the continental carbon sink. *Science* **2011**, *331* (6013), 50−50.

(4) Deemer, B. R.; Holgerson, M. A. Drivers of methane flux differ between lakes and reservoirs, complicating global upscaling efforts. *J. Geophys. Res.: Biogeosci.* **2021**, *126* (4), No. e2019JG005600.

(5) Townsend-Small, A.; Disbennett, D.; Fernandez, J. M.; Ransohoff, R. W.; Mackay, R.; Bourbonniere, R. A. Quantifying emissions of methane derived from anaerobic organic matter respiration and natural gas extraction in Lake Erie. *Limnol. Oceanogr.* **2016**, *61* (S1), S356−S366.

(6) Ho, J. C.; Michalak, A. M.; Pahlevan, N. Widespread global increase in intense lake phytoplankton blooms since the 1980s. *Nature* **2019**, *574* (7780), 667−670.

(7) Paerl, H. W.; Paul, V. J. Climate change: Links to global expansion of harmful cyanobacteria. *Water. Res.* **2012**, *46* (5), 1349−1363.

(8) Ma, J.; Loiselle, S.; Cao, Z.; Qi, T.; Shen, M.; Luo, J.; Song, K.; Duan, H. Unbalanced impacts of nature and nurture factors on the phenology, area and intensity of algal blooms in global large lakes: MODIS observations. *Sci. Total Environ.* **2023**, *880*, 163376.

(9) Bartosiewicz, M.; Maranger, R.; Przytulska, A.; Laurion, I. Effects of phytoplankton blooms on fluxes and emissions of greenhouse gases in a eutrophic lake. *Water. Res.* **2021**, *196*, 116985.

(10) Xiao, Q.; Zhang, M.; Hu, Z.; Gao, Y.; Hu, C.; Liu, C.; Liu, S.; Zhang, Z.; Zhao, J.; Xiao, W.; Lee, X. Spatial variations of methane emission in a large shallow eutrophic lake in subtropical climate. *J. Geophys. Res.: Biogeosci.* **2017**, *122* (7), 1597−1614.

(11) Beaulieu, J. J.; DelSontro, T.; Downing, J. A. Eutrophication will increase methane emissions from lakes and impoundments during the 21st century. *Nat. Commun.* **2019**, *10* (1), 1375.
(12) Li, M.; Peng, C.; Zhu, Q.; Zhou, X.; Yang, G.; Song, X.; Zhang, K. The significant contribution of lake depth in regulating global lake diffusive methane emissions. *Water. Res.* **2020**, *172*, 115465.
(13) Holgerson, M. A.; Raymond, P. A. Large contribution to inland water CO_2 and CH_4 emissions from very small ponds. *Nat. Geosci.* **2016**, *9* (3), 222−226.
(14) Johnson, M. S.; Matthews, E.; Du, J.; Genovese, V.; Bastviken, D. Methane emission from global lakes: New spatiotemporal data and observation-driven modeling of methane dynamics indicates lower emissions. *J. Geophys. Res.: Biogeosci.* **2022**, *127* (7), No. e2022JG006793.
(15) DelSontro, T.; Beaulieu, J. J.; Downing, J. A. Greenhouse gas emissions from lakes and impoundments: Upscaling in the face of global change. *Limnol. Oceanogr. Lett.* **2018**, *3* (3), 64−75.
(16) Rasilo, T.; Prairie, Y. T.; Del Giorgio, P. A. Large-scale patterns in summer diffusive CH_4 fluxes across boreal lakes, and contribution to diffusive C emissions. *Global Change Biol.* **2015**, *21* (3), 1124−1139.
(17) Wik, M.; Varner, R. K.; Anthony, K. W.; MacIntyre, S.; Bastviken, D. Climate-sensitive northern lakes and ponds are critical components of methane release. *Nat. Geosci.* **2016**, *9* (2), 99−105.
(18) Yang, P.; Yang, H.; Sardans, J.; Tong, C.; Zhao, G.; Penuelas, J.; Li, L.; Zhang, Y.; Tan, L.; Chun, K. P.; Lai, D. Y. F. Large spatial variations in diffusive CH_4 fluxes from a subtropical coastal reservoir affected by sewage discharge in southeast China. *Environ. Sci. Technol.* **2020**, *54* (22), 14192−14203.
(19) Loken, L. C.; Crawford, J. T.; Schramm, P. J.; Stadler, P.; Desai, A. R.; Stanley, E. H. Large spatial and temporal variability of carbon dioxide and methane in a eutrophic lake. *J. Geophys. Res.: Biogeosci.* **2019**, *124* (7), 2248−2266.
(20) Jansen, J.; Thornton, B. F.; Wik, M.; MacIntyre, S.; Crill, P. M. Temperature proxies as a solution to biased sampling of lake methane emissions. *Geophys. Res. Lett.* **2020**, *47* (14), No. e2020GL088647.
(21) Yvon-Durocher, G.; Allen, A. P.; Bastviken, D.; Conrad, R.; Gudasz, C.; St-Pierre, A.; Thanh-Duc, N.; del Giorgio, P. A. Methane fluxes show consistent temperature dependence across microbial to ecosystem scales. *Nature* **2014**, *507* (7493), 488−491.
(22) Davidson, T. A.; Audet, J.; Svenning, J. C.; Lauridsen, T. L.; Søndergaard, M.; Landkildehus, F.; Larsen, S. E.; Jeppesen, E. Eutrophication effects on greenhouse gas fluxes from shallow lake mesocosms override those of climate warming. *Global Change Biol.* **2015**, *21* (12), 4449−4463.
(23) Kutser, T. The possibility of using the Landsat image archive for monitoring long time trends in coloured dissolved organic matter concentration in lake waters. *Remote Sens. Environ.* **2012**, *123*, 334−338.
(24) Duan, H.; Tao, M.; Loiselle, S. A.; Zhao, W.; Cao, Z.; Ma, R.; Tang, X. MODIS observations of cyanobacterial risks in a eutrophic lake: Implications for long-term safety evaluation in drinking-water source. *Water. Res.* **2017**, *122*, 455−470.
(25) Seegers, B. N.; Werdell, P. J.; Vandermeulen, R. A.; Salls, W.; Stumpf, R. P.; Schaeffer, B. A.; Owens, T. J.; Bailey, S. W.; Scott, J. P.; Loftin, K. A. Satellites for long-term monitoring of inland U.S. lakes: The MERIS time series and application for chlorophyll-a. *Remote Sens. Environ.* **2021**, *266*, 112685.
(26) Hondula, K. L.; DeVries, B.; Jones, C. N.; Palmer, M. A. Effects of using high resolution satellite-based inundation time series to estimate methane fluxes from forested wetlands. *Geophys. Res. Lett.* **2021**, *48* (6), No. e2021GL092556.
(27) Martinsen, K. T.; Kragh, T.; Sand-Jensen, K. Carbon dioxide partial pressure and emission throughout the scandinavian stream network. *Global Biogeochem. Cycles* **2020**, *34* (12), No. e2020GL088647.
(28) Ludwig, S. M.; Natali, S. M.; Mann, P. J.; Schade, J. D.; Holmes, R. M.; Powell, M.; Fiske, G.; Commane, R. Using machine learning to predict inland aquatic CO_2 and CH_4 concentrations and the effects of wildfires in the Yukon-Kuskokwim Delta, Alaska. *Global Biogeochem. Cycles* **2022**, *36* (4), No. e2021GB007146.
(29) Beaulieu, J. J.; Waldo, S.; Balz, D. A.; Barnett, W.; Hall, A.; Platz, M. C.; White, K. M. Methane and carbon dioxide emissions from reservoirs: Controls and Upscaling. *J. Geophys. Res.: Biogeosci.* **2020**, *125* (12), No. e2019JG005474.
(30) Li, G.; Yang, M.; Zhang, Y.; Grace, J.; Lu, C.; Zeng, Q.; Jia, Y.; Liu, Y.; Lei, J.; Geng, X.; Wu, C.; Lei, G.; Chen, Y. Comparison model learning methods for methane emission prediction of reservoirs on a regional field scale: Performance and adaptation of methods with different experimental datasets. *Ecol. Eng.* **2020**, *157*, 105990.
(31) Duan, H. T.; Ma, R. H.; Xu, X. F.; Kong, F. X.; Zhang, S. X.; Kong, W. J.; Hao, J. Y.; Shang, L. L. Two-decade reconstruction of algal blooms in China's Lake Taihu. *Environ. Sci. Technol.* **2009**, *43* (10), 3522−3528.
(32) Shi, K.; Zhang, Y.; Xu, H.; Zhu, G.; Qin, B.; Huang, C.; Liu, X.; Zhou, Y.; Lv, H. Long-term satellite observations of microcystin concentrations in Lake Taihu during cyanobacterial bloom periods. *Environ. Sci. Technol.* **2015**, *49* (11), 6448−56.
(33) Lee, X.; Liu, S.; Xiao, W.; Wang, W.; Gao, Z.; Cao, C.; Hu, C.; Hu, Z.; Shen, S.; Wang, Y.; Wen, X.; Xiao, Q.; Xu, J.; Yang, J.; Zhang, M. The Taihu Eddy Flux Network: An observational program on energy, water, and greenhouse gas fluxes of a large freshwater lake. *Bull. Am. Meteorol. Soc.* **2014**, *95* (10), 1583−1594.
(34) Shen, M.; Duan, H.; Cao, Z.; Xue, K.; Qi, T.; Ma, J.; Liu, D.; Song, K.; Huang, C.; Song, X. Sentinel-3 OLCI observations of water clarity in large lakes in eastern China: Implications for SDG 6.3.2 evaluation. *Remote Sens. Environ.* **2020**, *247*, 111950.
(35) Qi, T.; Xiao, Q.; Cao, Z.; Shen, M.; Ma, J.; Liu, D.; Duan, H. Satellite Eestimation of dissolved carbon dioxide concentrations in China's Lake Taihu. *Environ. Sci. Technol.* **2020**, *54* (21), 13709−13718.
(36) Qin, B.; Xu, P.; Wu, Q.; Luo, L.; Zhang, Y. Environmental issues of Lake Taihu, China. *Hydrobiologia* **2007**, *581* (1), 3−14.
(37) Xiao, Q.; Xu, X.; Duan, H.; Qi, T.; Qin, B.; Lee, X.; Hu, Z.; Wang, W.; Xiao, W.; Zhang, M. Eutrophic Lake Taihu as a significant CO_2 source during 2000−2015. *Water. Res.* **2020**, *170*, 115331.
(38) Luo, J.; Pu, R.; Duan, H.; Ma, R.; Mao, Z.; Zeng, Y.; Huang, L.; Xiao, Q. Evaluating the influences of harvesting activity and eutrophication on loss of aquatic vegetations in Taihu Lake, China. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *87*, 102038.
(39) Xiao, Q.; Xu, X.; Zhang, M.; Duan, H.; Hu, Z.; Wang, W.; Xiao, W.; Lee, X. Coregulation of nitrous oxide emissions by nitrogen and temperature in China's third largest freshwater lake (Lake Taihu). *Limnol. Oceanogr.* **2019**, *64*, 1070−1086.
(40) Wik, M.; Thornton, B. F.; Bastviken, D.; Uhlbaeck, J.; Crill, P. M. Biased sampling of methane release from northern lakes: A problem for extrapolation. *Geophys. Res. Lett.* **2016**, *43* (3), 1256−1262.
(41) Huang, C. C.; Yao, L.; Huang, T.; Zhang, M. L.; Zhu, A. X.; Yang, H. Wind and rainfall regulation of the diffuse attenuation coefficient in large, shallow lakes from long-term MODIS observations using a semianalytical model. *J. Geophys. Res.: Atmos.* **2017**, *122* (13), 6748−6763.
(42) Tu, Z.; Le, C.; Bai, Y.; Jiang, Z.; Wu, Y.; Ouyang, Z.; Cai, W. J.; Qi, D. Increase in CO_2 uptake capacity in the Arctic Chukchi Sea during summer revealed by satellite-based estimation. *Geophys. Res. Lett.* **2021**, *48* (15), No. e2021GL093844.
(43) Natchimuthu, S.; Sundgren, I.; Gålfalk, M.; Klemedtsson, L.; Crill, P.; Danielsson, Å.; Bastviken, D. Spatio-temporal variability of lake CH_4 fluxes and its influence on annual whole lake emission estimates. *Limnol. Oceanogr.* **2016**, *61* (S1), S13−S26.
(44) Schrier-Uijl, A. P.; Veraart, A. J.; Leffelaar, P. A.; Berendse, F.; Veenendaal, E. M. Release of CO_2 and CH_4 from lakes and drainage ditches in temperate wetlands. *Biogeochemistry* **2011**, *102* (1−3), 265−279.
(45) Xiao, Q.; Duan, H.; Qin, B.; Hu, Z.; Zhang, M.; Qi, T.; Lee, X. Eutrophication and temperature drive large variability in carbon

dioxide from China's Lake Taihu. *Limnol. Oceanogr.* **2022**, *67*, 379–391.
(46) Zhang, L.; Xia, X.; Liu, S.; Zhang, S.; Li, S.; Wang, J.; Wang, G.; Gao, H.; Zhang, Z.; Wang, Q.; Wen, W.; Liu, R.; Yang, Z.; Stanley, E. H.; Raymond, P. A. Significant methane ebullition from alpine permafrost rivers on the East Qinghai-Tibet Plateau. *Nat. Geosci.* **2020**, *13* (5), 349–354.
(47) Zhu, Y.; Purdy, K. J.; Eyice, Ö.; Shen, L.; Harpenslager, S. F.; Yvon-Durocher, G.; Dumbrell, A. J.; Trimmer, M. Disproportionate increase in freshwater methane emissions induced by experimental warming. *Nat. Clim. Change* **2020**, *10* (7), 685–690.
(48) Colas, F.; Chanudet, V.; Daufresne, M.; Buchet, L.; Vigouroux, R.; Bonnet, A.; Jacob, F.; Baudoin, J. M. Spatial and temporal variability of diffusive CO_2 and CH_4 fluxes from the Amazonian Reservoir Petit-Saut (French Guiana) reveals the importance of allochthonous inputs for long-term C emissions. *Global Biogeochem. Cycles* **2020**, *34* (12), No. e2020GB006602.
(49) Paranaiba, J. R.; Barros, N.; Mendonca, R.; Linkhorst, A.; Isidorova, A.; Roland, F.; Almeida, R. M.; Sobek, S. Spatially resolved measurements of CO_2 and CH_4 concentration and gas-exchange velocity highly influence carbon-emission estimates of reservoirs. *Environ. Sci. Technol.* **2018**, *52* (2), 607–615.
(50) Chen, S.; Hu, C.; Barnes, B. B.; Wanninkhof, R.; Cai, W.-J.; Barbero, L.; Pierrot, D. A machine learning approach to estimate surface ocean pCO_2 from satellite measurements. *Remote Sens. Environ.* **2019**, *228*, 203–226.
(51) Dávila-Santiago, E.; Shi, C.; Mahadwar, G.; Medeghini, B.; Insinga, L.; Hutchinson, R.; Good, S.; Jones, G. D. Machine learning applications for chemical fingerprinting and environmental source tracking using non-target chemical data. *Environ. Sci. Technol.* **2022**, *56* (7), 4080–4090.
(52) Jansen, J.; Thornton, B. F.; Cortés, A.; Snöälv, J.; Wik, M.; MacIntyre, S.; Crill, P. M. Drivers of diffusive CH_4 emissions from shallow subarctic lakes on daily to multi-year timescales. *Biogeosciences* **2020**, *17* (7), 1911–1932.
(53) Yvon-Durocher, G.; Hulatt, C. J.; Woodward, G.; Trimmer, M. Long-term warming amplifies shifts in the carbon cycle of experimental ponds. *Nat. Clim. Change* **2017**, *7* (3), 209–213.
(54) Crawford, J. T.; Loken, L. C.; Casson, N. J.; Smith, C.; Stone, A. G.; Winslow, L. A. High-speed limnology: Using advanced sensors to investigate spatial variability in biogeochemistry and hydrology. *Environ. Sci. Technol.* **2015**, *49* (1), 442–50.
(55) Schilder, J.; Bastviken, D.; van Hardenbroek, M.; Kankaala, P.; Rinta, P.; Stötter, T.; Heiri, O. Spatial heterogeneity and lake morphology affect diffusive greenhouse gas emission estimates of lakes. *Geophys. Res. Lett.* **2013**, *40* (21), 5752–5756.

https://doi.org/10.1021/acs.est.3c05631
Environ. Sci. Technol. 2023, 57, 13520–13529