
Article:
McNicol, G., Fluet-Chouinard, E., Ouyang, Z. et al. (2023) Upscaling wetland methane emissions from the FLUXNET-CH4 eddy covariance network (UpCH4 v1.0): model development, network assessment, and budget comparison. AGU Advances, 4 (5). ISSN: 2576-604X

https://doi.org/10.1029/2023av000956

AGU ADVANCING EARTH AND SPACE SCIENCES

AGU Advances

RESEARCH ARTICLE
10.1029/2023AV000956

Upscaling Wetland Methane Emissions From the FLUXNET-CH4 Eddy Covariance Network (UpCH4 v1.0): Model Development, Network Assessment, and Budget Comparison

Gavin McNicol[1,2], Etienne Fluet-Chouinard[1,3], Zutao Ouyang[1], Sara Knox[4], Zhen Zhang[5], Tuula Aalto[6], Sheel Bansal[7], Kuang-Yu Chang[8], Min Chen[9], Kyle Delwiche[10], Sarah Feron[11], Mathias Goeckede[12], Jinxun Liu[13], Avni Malhotra[14], Joe R. Melton[15], William Riley[8], Rodrigo Vargas[16], Kunxiaojia Yuan[8], Qing Ying[17], Qing Zhu[8], Pavel Alekseychik[18], Mika Aurela[6], David P. Billesbach[19], David I. Campbell[20], Jiquan Chen[21], Housen Chu[22], Ankur R. Desai[23], Eugenie Euskirchen[24], Jordan Goodrich[20], Timothy Griffis[25], Manuel Helbig[26], Takashi Hirano[27], Hiroki Iwata[28], Gerald Jurasinski[29,30], John King[31], Franziska Koebsch[32], Randall Kolka[33], Ken Krauss[34], Annalea Lohila[6,35], Ivan Mammarella[35], Mats Nilson[36], Asko Noormets[37], Walter Oechel[38], Matthias Peichl[36], Torsten Sachs[39], Ayaka Sakabe[40], Christopher Schulze[41], Oliver Sonnentag[42], Ryan C. Sullivan[43], Eeva-Stiina Tuittila[44], Masahito Ueyama[45], Timo Vesala[35,46], Eric Ward[33], Christian Wille[39], Guan Xhuan Wong[47], Donatella Zona[38,48], Lisamarie Windham-Myers[49], Benjamin Poulter[50], and Robert B. Jackson[1,51,52]

[1]Department of Earth System Science, Stanford University, Stanford, CA, USA, [2]Now at Department of Earth and Environmental Sciences, University of Illinois Chicago, Chicago, IL, USA, [3]Department of Environmental Systems Science, Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland, [4]Department of Geography, The University of British Columbia, Vancouver, BC, Canada, [5]Department of Geographical Sciences, University of Maryland, College Park, MD, USA, [6]Finnish Meteorological Institute, Climate Change, Helsinki, Finland, [7]U.S. Geological Survey, Northern Prairie Wildlife Research Center, Jamestown, ND, USA, [8]Earth and Environmental Sciences Area, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, [9]Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, Madison, WI, USA, [10]Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA, USA, [11]Knowledge Infrastructure, University of Groningen, Groningen, The Netherlands, [12]Department for Biogeochemical Signals, Max Planck Institute for Biogeochemistry, Jena, Germany, [13]U.S. Geological Survey, Western Geographic Science Center, Moffett Field, CA, USA, [14]Department of Geography, University of Zurich, Zurich, Switzerland, [15]Climate Research Division, Environment and Climate Change Canada, Victoria, BC, Canada, [16]Department of Plant and Soil Sciences, University of Delaware, Newark, DE, USA, [17]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA, [18]Natural Resources Institute Finland (LUKE), Helsinki, Finland, [19]Department of Biological Systems Engineering, University of Nebraska-Lincoln, Lincoln, NE, USA, [20]School of Science, University of Waikato, Hamilton, New Zealand, [21]Center for Global Change and Earth Observations, Michigan State University, East Lansing, MI, USA, [22]Climate and Ecosystem Sciences Division, Lawrence Berkeley National Lab, Berkeley, CA, USA, [23]Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, Madison, WI, USA, [24]University of Alaska Fairbanks, Institute of Arctic Biology, Fairbanks, AK 99775, USA, [25]Department Soil, Water, and Climate, University of Minnesota Twin Cities, St. Paul, MN, USA, [26]Department of Physics and Atmospheric Science, Dalhousie University, Halifax, NS, Canada, [27]Research Faculty of Agriculture, Hokkaido University, Sapporo, Japan, [28]Department of Environmental Science, Faculty of Science, Shinshu University, Matsumoto, Japan, [29]Landcape Ecology, University of Rostock, Rostock, Germany, [30]Peatland Science, University of Greifswald, Greifswald, Germany, [31]Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, USA, [32]Bioclimatology, University of Göttingen, Göttingen, Germany, [33]USDA Forest Service Northern Research Station, Grand Rapids, MN, USA, [34]USGS Wetland and Aquatic Research Center, Lafayette, LA, USA, [35]Institute for Atmospheric and Earth System Research/Physics, Faculty of Science, University of Helsinki, Helsinki, Finland, [36]Department of Forest Ecology and Management, Swedish University of Agricultural Sciences, Umeå, Sweden, [37]Department of Ecology and Conservation Biology, Texas A&M University, College Station, TX, USA, [38]Department Biology, San Diego State University, San Diego, CA, USA, [39]GFZ German Research Centre for Geosciences, Potsdam, Germany, [40]Hakubi center, Kyoto University, Kyoto, Japan, [41]University of Alberta, Renewable Resources; Université de Montréal, Département de géographie, Université de Montréal, Montréal, QC, Canada, [42]Département de géographie, Université de Montréal, Montréal, QC, Canada, [43]Environmental Science Division, Argonne National Laboratory, Lemont, IL, USA, [44]School of Forest Sciences, University of Eastern Finland, Joesnuu, Finland, [45]Graduate School of Agriculture, Osaka Metropolitan University, Sakai, Japan, [46]Faculty of Agriculture and Forestry, Institute for Atmospheric and Earth System Research/Forest Sciences, University of Helsinki, Helsinki, Finland, [47]Sarawak Tropical Peat Research Institute, Kota Samarahan, Malaysia, [48]Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK, [49]U.S. Geological Survey, Water

Peer Review The peer review history for this article is available as a PDF in the Supporting Information.

Key Points:

- Random forest models trained on FLUXNET-CH4 methane fluxes reproduced spatiotemporal patterns in extra-tropical wetlands (R^2: 0.59–0.64)
- Globally upscaled annual wetland methane emissions (146 $TgCH_4$ y^{-1}) overlapped with land surface and inversion model ensemble estimates
- Humid/monsoon tropics dominate upscaled wetland methane emissions (~68%) and uncertainties (~78%) due to limited FLUXNET-CH4 site coverage

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

G. McNicol and E. Fluet-Chouinard,
gmcnicol@uic.edu;
efluet@stanford.edu

Citation:

McNicol, G., Fluet-Chouinard, E., Ouyang, Z., Knox, S., Zhang, Z., Aalto, T., et al. (2023). Upscaling wetland methane emissions from the FLUXNET-CH4 eddy covariance network (UpCH4 v1.0): Model development, network assessment, and budget comparison. *AGU Advances*, *4*, e2023AV000956. https://doi.org/10.1029/2023AV000956

Received 12 JUL 2022
Accepted 15 MAY 2023

Author Contributions:

Conceptualization: Gavin McNicol, Etienne Fluet-Chouinard, Zutao Ouyang, Sara Knox, Benjamin Poulter, Robert B. Jackson

Data curation: Gavin McNicol, Etienne Fluet-Chouinard, Zutao Ouyang, Sara Knox
Formal analysis: Gavin McNicol
Funding acquisition: Sara Knox, Lisamarie Windham-Myers, Benjamin Poulter, Robert B. Jackson
Methodology: Zhen Zhang
Project Administration: Sara Knox, Robert B. Jackson
Supervision: Benjamin Poulter, Robert B. Jackson
Validation: Zhen Zhang
Visualization: Etienne Fluet-Chouinard
Writing – original draft: Gavin McNicol, Etienne Fluet-Chouinard
Writing – review & editing: Gavin McNicol, Etienne Fluet-Chouinard, Zutao Ouyang, Sara Knox, Zhen Zhang, Lisamarie Windham-Myers, Benjamin Poulter, Robert B. Jackson

Mission Area, Menlo Park, CA, USA, [50]Biospheric Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA, [51]Woods Institute for the Environment, Stanford University, Stanford, CA, USA, [52]Precourt Institute for Energy, Stanford University, Stanford, CA, USA

Abstract Wetlands are responsible for 20%–31% of global methane (CH_4) emissions and account for a large source of uncertainty in the global CH_4 budget. Data-driven upscaling of CH_4 fluxes from eddy covariance measurements can provide new and independent bottom-up estimates of wetland CH_4 emissions. Here, we develop a six-predictor random forest upscaling model (UpCH4), trained on 119 site-years of eddy covariance CH_4 flux data from 43 freshwater wetland sites in the FLUXNET-CH4 Community Product. Network patterns in site-level annual means and mean seasonal cycles of CH_4 fluxes were reproduced accurately in tundra, boreal, and temperate regions (Nash-Sutcliffe Efficiency ~0.52–0.63 and 0.53). UpCH4 estimated annual global wetland CH_4 emissions of 146 ± 43 $TgCH_4\ y^{-1}$ for 2001–2018 which agrees closely with current bottom-up land surface models (102–181 $TgCH_4\ y^{-1}$) and overlaps with top-down atmospheric inversion models (155–200 $TgCH_4\ y^{-1}$). However, UpCH4 diverged from both types of models in the spatial pattern and seasonal dynamics of tropical wetland emissions. We conclude that upscaling of eddy covariance CH_4 fluxes has the potential to produce realistic extra-tropical wetland CH_4 emissions estimates which will improve with more flux data. To reduce uncertainty in upscaled estimates, researchers could prioritize new wetland flux sites along humid-to-arid tropical climate gradients, from major rainforest basins (Congo, Amazon, and SE Asia), into monsoon (Bangladesh and India) and savannah regions (African Sahel) and be paired with improved knowledge of wetland extent seasonal dynamics in these regions. The monthly wetland methane products gridded at 0.25° from UpCH4 are available via ORNL DAAC (https://doi.org/10.3334/ORNLDAAC/2253).

Plain Language Summary Wetlands account for a large share of global methane emissions to the atmosphere, but current estimates vary widely in magnitude (~30% uncertainty on annual global emissions) and spatial distribution, with diverging predictions for tropical rice growing (e.g., Bengal basin), rainforest (e.g., Amazon basin), and floodplain savannah (e.g., Sudd) regions. Wetland methane model estimates could be improved by increased use of land surface methane flux data. Upscaling approaches use flux data collected across globally distributed measurement networks in a machine learning framework to extrapolate fluxes in space and time. Here, we train and evaluate a methane upscaling model (UpCH4) and use it to generate monthly, globally gridded wetland methane emissions estimates for 2001–2018. The UpCH4 model uses only six predictor variables among which temperature is dominant. Global annual methane emissions estimates and associated uncertainty ranges from upscaling fall within state-of-the-art model ensemble estimates from the Global Carbon Project (GCP) methane budget. In some tropical regions, the spatial pattern of UpCH4 emissions diverged from GCP predictions, however, inclusion of flux measurements from additional ground-based sites, together with refined maps of tropical wetlands extent, could reduce these prediction uncertainties.

1. Introduction

The post-industrial rise in atmospheric methane (CH_4) concentrations has had a large climate warming effect, 60% the size of that for carbon dioxide (CO_2) (IPCC, 2021). The short atmospheric lifetime of CH_4 also promises relatively fast climate change mitigation effects following CH_4 emissions reductions, rather than century-or-more timescales for CO_2 reductions (Abernethy et al., 2021; Turner et al., 2019). However, current emissions trajectories more closely track high emissions scenarios (Zhang et al., 2023). Since 2014, there has been an accelerating increase in the CH_4 growth rate that reached a record level in 2022, at 18.2 $ppb\ y^{-1}$ (Lan et al., 2023), and these increases could continue as global temperatures rise (Bansal et al., 2023; Zhang et al., 2017).

Large uncertainties around total (8%–39%) and individual CH_4 sources (Table 1) prevent CH_4 budget closure at regional-to-global scales. Better constrained CH_4 budgets are needed to more accurately attribute and mitigate the sources causing the accelerating rise in CH_4 emissions (Nisbet et al., 2022). Improving wetland CH_4 emissions estimates will help constrain the global CH_4 budget as wetlands comprise both the largest natural CH_4 emissions source (20%–30% of total emissions) and the second largest uncertainty in the CH_4 budget (Saunois et al., 2020).

Table 1
Natural and Anthropogenic Methane Emissions by Source Sector for the Decade 2008–2017 From Global Carbon Project Top-Down (TD) and Bottom-Up (BU) Model Ensembles (Saunois et al., 2020)

Source sector	Ensemble	Average emissions [ensemble range] ($TgCH_4\ y^{-1}$)	Absolute range ($TgCH_4\ y^{-1}$)	Normalized range
Total	BU	737 [594–881]	287	39%
	TD	576 [550–594]	44	8%
	BU—TD	161		
Natural wetlands	BU	149 [102–182]	80	54%
	TD	181 [159–200]	41	23%
	BU - TD	−32		
Other natural	BU	222 [143–306]	163	73%
	TD	37 [21–50]	29	78%
	BU—TD	185		
Agriculture and waste	BU	206 [191–223]	32	16%
	TD	217 [207–240]	33	15%
	BU—TD	−11		
Fossil fuels	BU	128 [113–154]	41	32%
	TD	111 [81–131]	50	45%
	BU—TD	17		
Biomass and biofuel burning	BU	30 [26–40]	14	47%
	TD	30 [22–36]	14	47%
	BU—TD	0		

Note. "Other natural" combines open (non-vegetated) freshwaters, geological (on and offshore), wild animal, termite, wildfire, permafrost, and biological oceanic sources. Normalized ranges are reported as absolute ranges divided by ensemble average emissions.

Wetlands, as defined here, include both seasonally and permanently inundated soils that are vegetated including riparian and floodplain forests, inland marsh systems, and peat-forming wetlands (peatlands) but exclude rice agriculture, tidal and non-vegetated waterbodies such as ponds, lakes, streams, rivers, and estuaries (Zhang et al., 2021).

Current methods to estimate global wetland CH_4 emissions generally fall into one of two approaches: top-down (TD) atmospheric observation-based inversions and bottom-up (BU) land surface models. Although both approaches involve state-of-the-art methods, emissions estimates vary significantly within and between the two approaches (Saunois et al., 2020). For the decade 2008–2017, a comprehensive TD model ensemble, compiled as part of the Global Carbon Project (GCP) methane budget activity, estimated wetland emissions of 159–200 (mean 181) $TgCH_4\ yr^{-1}$. TD inversions use approaches such as Bayesian inference or ensemble Kalman filters to estimate the (posterior) wetland emissions required to reproduce in-situ and satellite atmospheric CH_4 concentration retrievals. Since observations are limited, wetland source attribution strongly depends on prior assumptions as additional sources of information, constraining sectoral contributions (i.e., natural, industrial, agricultural and waste emissions) in space and time (Jacob et al., 2022). Considerable variability exists between TD models, primarily due to differences in assumptions and whether satellite retrievals are included as constraints (Kirschke et al., 2013; Saunois et al., 2020).

In contrast to TD models, BU model estimates are not constrained by atmospheric CH_4 concentration data and attempt to directly represent wetland CH_4 fluxes and underlying flux processes with varying complexity (Riley et al., 2011; Ueyama et al., 2023). For the same decade (2008–2017), the GCP's 13-member BU model ensemble estimated emissions of 102–182 (mean 149) $TgCH_4\ yr^{-1}$, ~20% lower than the TD ensemble mean (Saunois et al., 2020). The widespread observed among BU models arises from differences in model parameterization, which is informed by process knowledge and literature estimates of parameter values and sometimes by calibration to observed wetland CH_4 fluxes at a limited number of sites (e.g., 3 northern wetlands for

Wetland-DNDC in Zhang et al., 2002). However, Chang et al. (2021) showed a large variability in CH_4 flux temperature dependency across a network of eddy covariance tower sites (FLUXNET-CH4; Delwiche et al., 2021; Knox et al., 2019), indicating that upscaling from few sites is likely to introduce errors during global extrapolation. To date, process-based models have yet to use networked data, for instance FLUXNET-CH4, for multi-site calibration, in part because this approach is more technically challenging than for machine learning models. Equally large uncertainties (~50% of total uncertainty) are introduced when BU models simulate independent wetland extents in prognostic runs versus using prescribed global wetland extents in diagnostic runs (Melton et al., 2013). Notably, the substantial GCP BU spread (±30%-50% of ensemble mean emissions) was observed even in diagnostic model runs where all models were prescribed a common Wetland Area and Dynamics for Methane Modeling (WAD2M) wetland extent (Zhang et al., 2021), underscoring the need to reduce wetland CH_4 flux uncertainties as well as wetland extent uncertainties.

No global benchmark data set exists to favor or falsify, with a strong degree of confidence, any BU or TD model (Saunois et al., 2020). Insights about model parameterization and sources of uncertainty can only be gained at present via model intercomparisons, such as the BU Wetland CH_4 Inter-Comparison of Models Project (WETCHIMP) activity (Melton et al., 2013) and the TD Wetland CH_4 emission and uncertainty ensemble data set for Atmospheric Chemistry and Transport modeling (WetCHARTS) activity (Bloom et al., 2017), the GCP wetland CH_4 synthesis (Poulter et al., 2017), by regional scale evaluation of converging or diverging TD and BU estimates (Stavert et al., 2021; example in Figure 1), or by comparison to satellite retrievals (Parker et al., 2018). Independent estimates of global wetland CH_4 emissions, incorporating new data for calibration and model constraints, and implementing new modeling approaches, such as machine learning algorithms, are emerging alternatives for refining models and reducing uncertainties around wetland CH_4 sources (Saunois et al., 2020).

One data stream that can improve wetland CH_4 models and global emission estimates is the growing availability of CH_4 fluxes measured near the land surface. In situ eddy covariance flux towers provide long-term, semi-automated, and quasi-continuous fluxes at ecosystem scales (<1 km^2) with minimal disturbance to soils or canopy structure/function (Baldocchi, 2014; Chu et al., 2021). Although BU models have been parameterized using CH_4 flux data at individual sites, network CH_4 data has not been fully utilized. Since the late-1990s, FLUXNET has provided standardized CO_2 flux data measured using eddy covariance across hundreds of locations around the world, enabling independent benchmarking of satellite measurements and Earth system models (Jung et al., 2020; Pastorello et al., 2020). Upscaling is a workflow combining statistical models and data to transfer information across scales, often using machine learning, and has been used by projects such as FLUXCOM to extrapolate FLUXNET data from 224 sites (~850 site years) and predict global terrestrial ecosystem carbon and energy fluxes (Bodesheim et al., 2018; Jung et al., 2020; Tramontana et al., 2016). The FLUXNET-CH4 data set now provides similar opportunities to refine model parameterization and generate independent, data-driven estimates of regional-to-global CH_4 emissions (Chang et al., 2021; Delwiche et al., 2021; Knox et al., 2019). Peltola et al. (2019) independently acquired eddy covariance data from 25 high-latitude sites and developed a wetland CH_4 flux upscaling workflow to predict monthly, regional (>45°N) wetland CH_4 emissions for 2012–2013. Annual emissions of 31–38 $TgCH_4\ y^{-1}$ agreed well with previous bottom-up estimates (e.g., Chen et al., 2015; Treat et al., 2018; Zhang et al., 2017) but were higher than those of top-down estimates (23–28 $TgCH_4\ y^{-1}$) for the region (e.g., Bruhwiler et al., 2014; Spahni et al., 2011). Peltola et al. (2019) thus demonstrated that upscaling estimates from eddy covariance data could produce plausible CH_4 fluxes at regional scales. To date, however, no upscaling project has taken advantage of the full FLUXNET-CH4 site network to make and evaluate global wetland CH_4 emissions predictions.

Here, we develop a wetland CH_4 upscaling workflow (UpCH4) that combines FLUXNET-CH4 and globally gridded predictor data to train random forest model ensembles, including validation and test routines optimized for spatial prediction applications. Given that surface flux measurement may networks take decades to grow, our first goal is to robustly evaluate the ability of machine learning models to extrapolate beyond training conditions in space and time. We then use this CH_4 flux upscaling workflow (UpCH4) to predict wetland CH_4 emissions globally and compare them to current top-down and bottom-up model estimates. Given knowledge of model structure and network coverage, our second goal is to identify regions of convergence and diagnose regions of divergence between UpCH4 and existing estimates. Finally, we use knowledge of the model structure to conduct a FLUXNET-CH4 network dissimilarity analysis with the goal of informing strategic improvements in eddy covariance site coverage.

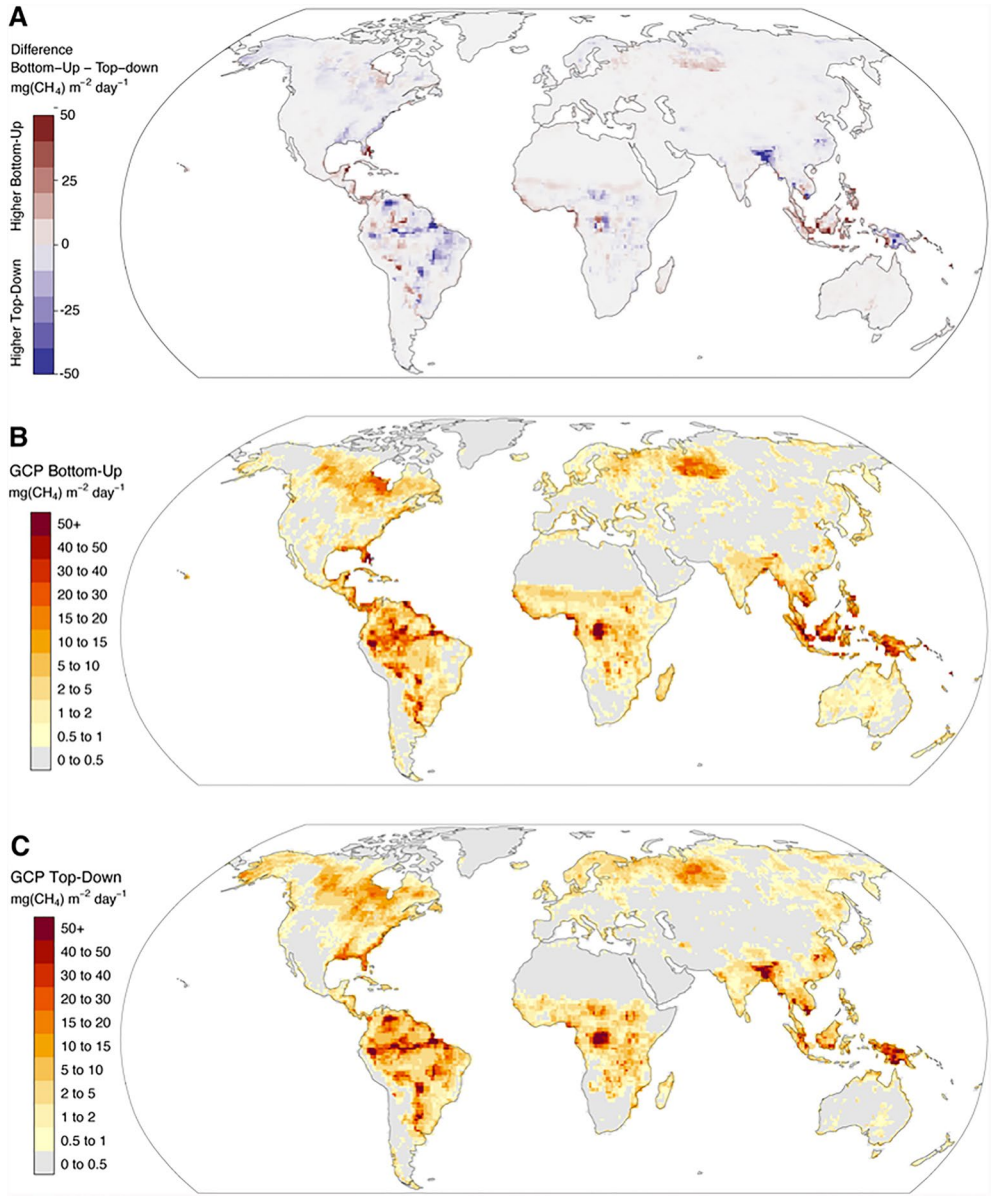

Figure 1. (a) Large regional discrepancies exist between (b) bottom-up (BU) and (c) top-down (TD) model estimates of wetland CH_4 emissions, in addition to different global totals. Mean daily natural wetland CH_4 emissions for 2010–2017 estimated from a 17-member TD ensemble are subtracted from the mean from a 13-member BU process model ensemble (Saunois et al., 2020). High northern latitude bounding boxes correspond to locations of Hudson Bay Lowland (left) and West Siberian Lowland (right) wetland complexes.

2. Methods

2.1. Predictive Modeling

2.1.1. Eddy Covariance CH_4 Flux Data

UpCH4 was designed to predict CH_4 fluxes globally at inundated and non-inundated (i.e., shallow water table) freshwater vegetated wetlands. Forty-five natural and restored freshwater wetland sites from the FLUXNET-CH4 database qualified for model training (Delwiche et al., 2021). Two sites (RU-VrK and SE-St1) were excluded after quality control filtering and 1 year of data was excluded from a restored wetland site (US-Sne; Table 2) that had not yet developed vegetation cover. One site (DE-Hte) is coastal but was freshwater dominated during the observation period. The final eddy covariance tower data set consisted of 43 freshwater wetland sites covering

AGU ADVANCING EARTH AND SPACE SCIENCES
AGU Advances
10.1029/2023AV000956

Table 2
Cluster Assignments, Number of Weeks and Percent of Total Data Set, and Wetland Site ID, Location, Class, Climate, and Digital Object Identifiers (DOI) for the 43 Wetlands Included in Upscaling

Cluster	Weeks	Percent	ID	Site name	Country	Latitude	Longitude	Class	MAT (C)	MAP (mm)	DOI references
1	140	2.2	BR-Npw	Northern Pantanal Wetland	Brazil	−16.498	−56.412	Swamp	25.2	1,318	Vourlitis et al. (2020)
2	34	0.5	BW-Gum	Guma	Botswana	−18.964722	22.3711111	Swamp	23.1	459	Helfter (2020a)
2	33	0.5	BW-Nxr	Nxaraga	Botswana	−19.548056	23.1791667	Swamp	23.5	433	Helfter (2020b)
3	114	1.8	CA-SCB	Scotty Creek Bog	Canada	61.3089	−121.2984	Bog	−2.8	414	Sonnentag and Helbig (2020a)
3	116	1.8	CA-SCC	Scotty Creek Landscape	Canada	61.3079	−121.2992	Bog	−2.9	414	Sonnentag and Helbig (2020b)
4	379	6.0	DE-Hte	Huetelmoor	Germany	54.210278	12.176111	Fen	8.5	584	Koebsch and Jurasinski (2020)
4	248	4.0	DE-Zrk	Zarnekow	Germany	53.87594	12.88901	Fen	8.3	580	Sachs and Wille (2020)
5	115	1.8	DE-SfN	Schechenfilz Nord	Germany	47.80639	11.3275	Bog	8.3	1,123	Schmid and Klatt (2020)
6	257	4.1	FI-Lom	Lompolojankka	Finland	67.99724	24.20918	Fen	−1.0	512	Lohila et al. (2020)
7	125	2.0	FI-Si2	Siikaneva-2 Bog	Finland	61.8372	24.1967	Bog	3.2	664	Vesala, Tuittila, Mammarella, and Alekseychik (2020)
7	257	4.1	FI-Sii	Siikaneva	Finland	61.83265	24.19285	Fen	3.2	666	Vesala, Tuittila, Mammarella, and Rinne (2020)
8	37	0.6	FR-LGt	La Guette	France	47.322917	2.284102	Fen	11.0	707	Jacotot et al. (2020)
9	62	1.0	ID-Pag	Palangkaraya undrained forest	Indonesia	−2.32	113.9	Swamp	27.4	2386	Sakabe et al. (2020)
10	114	1.8	JP-BBY	Bibai bog	Japan	43.3230056	141.810697	Bog	6.7	1,153	Ueyama et al. (2020)
11	81	1.3	MY-MLM	Maludam National Park	Malaysia	1.453575	111.149492	Swamp	26.9	3,401	Wong et al. (2020)
12	206	3.3	NZ-Kop	Kopuatai	New Zealand	−37.3879	175.5539	Bog	13.9	1,343	Campbell and Goodrich (2020)
13	117	1.9	RU-Ch2	Chersky reference	Russia	68.61689	161.35089	Wet tundra	−12.3	172	Goeckede (2020)
14	78	1.2	RU-Cok	Chokurdakh	Russia	70.82914	147.49428	Wet tundra	−14.1	210	Dolman et al. (2020)
15	246	3.9	SE-Deg	Degero	Sweden	64.182029	19.556539	Fen	1.7	620	Nilsson and Peichl (2020)
16	58	0.9	US-A03	ARM-AMF3-Oliktok	United States	70.495328	−149.8823	Wet tundra	−11.9	144	Billesbach and Sullivan (2020a)
16	72	1.1	US-ICs	Imnavait Creek Watershed Wet Sedge Tundra	United States	68.6058	−149.311	Wet tundra	−8.9	242	Euskirchen et al. (2020c)
17	36	0.6	US-A10	ARM-NSA-Barrow	United States	71.3242	−156.6149	Wet tundra	−12.0	107	Billesbach and Sullivan (2020b)
17	84	1.3	US-Atq	Atqasuk	United States	70.4696	−157.4089	Wet tundra	−10.3	133	Zona and Oechel (2020a)
17	74	1.2	US-Beo	Barrow Environmental Observatory (BEO) tower	United States	71.281	−156.6123	Wet tundra	−11.9	109	Zona and Oechel (2020b)

Table 2
Continued

Cluster	Weeks	Percent	ID	Site name	Country	Latitude	Longitude	Class	MAT (C)	MAP (mm)	DOI references
17	128	2.0	US-Bes	Barrow-Bes (Biocomplexity Experiment South tower)	United States	71.2809	−156.5965	Wet tundra	−12.0	109	Zona and Oechel (2020c)
17	131	2.1	US-NGB	NGEE Arctic Barrow	United States	71.280044	−156.60918	Wet tundra	−11.9	109	Torn and Dengel (2020a)
18	97	1.5	US-BZB	Bonanza Creek Thermokarst Bog	United States	64.695547	−148.32084	Bog	−2.4	292	Euskirchen and Edgar (2020a)
18	91	1.5	US-BZF	Bonanza Creek Rich Fen	United States	64.703733	−148.31333	Fen	−2.5	294	Euskirchen and Edgar (2020b)
18	206	3.3	US-Uaf	University of Alaska, Fairbanks	United States	64.86627	−147.85553	Bog	−2.8	298	Iwata et al. (2020)
19	177	2.8	US-DPW	Disney Wilderness Preserve Wetland	United States	28.05206	−81.43611	Marsh	22.1	1,223	Hinkle and Bracho (2020)
20	165	2.6	US-Ivo	Ivotuk	United States	68.4865	−155.7503	Wet tundra	−8.5	247	Zona and Oechel (2020d)
21	92	1.5	US-LA2	Salvador WMA Freshwater Marsh	United States	29.8587	−90.2869	Marsh	20.0	1,616	Holm et al. (2020)
22	270	4.3	US-Los	Lost Creek	United States	46.0827	−89.9792	Fen	4.1	833	Desai (2020)
23	478	7.6	US-Myb	Mayberry Wetland	United States	38.049861	−121.76498	Marsh	15.4	346	Matthes et al. (2020)
23	114	1.8	US-Sne	Sherman Island Restored Wetland	United States	38.0369	−121.7547	Marsh	15.5	340	Short et al. (2020)
23	285	4.5	US-Tw1	Twitchell Wetland West Pond	United States	38.1074	−121.6469	Marsh	15.4	371	Valach, Szutu, et al. (2020)
23	304	4.8	US-Tw4	Twitchell East End Wetland	United States	38.1027436	−121.64133	Marsh	15.4	370	Eichelmann et al. (2020)
23	34	0.5	US-Tw5	East Pond Wetland	United States	38.107155	−121.64257	Marsh	15.4	371	Valach, Kasak, et al. (2020)
24	224	3.6	US-NC4	NC_AlligatorRiver	United States	35.7879	−75.9038	Swamp	16.5	1,322	Noormets et al. (2020)
25	33	0.5	US-NGC	NGEE Arctic Council	United States	64.8614	−163.7008	Wet tundra	−3.1	413	Torn and Dengel (2020b)
26	191	3.0	US-ORv	Olentangy River Wetland Research Park	United States	40.0201	−83.0183	Marsh	11.0	954	Bohrer and Morin (2020)
26	31	0.5	US-OWC	Old Woman Creek	United States	41.3795167	−82.512467	Marsh	9.9	898	Bohrer et al. (2020)
26	138	2.2	US-WPT	Winous Point North Marsh	United States	41.464639	−82.996157	Marsh	9.9	881	Chen and Chu (2020)

Note. Mean annual temperature (MAT; °C) and mean annual precipitation (MAP; mm) were extracted from WorldClim 2.0 (Fick & Hijmans, 2017).

AGU Advances
10.1029/2023AV000956

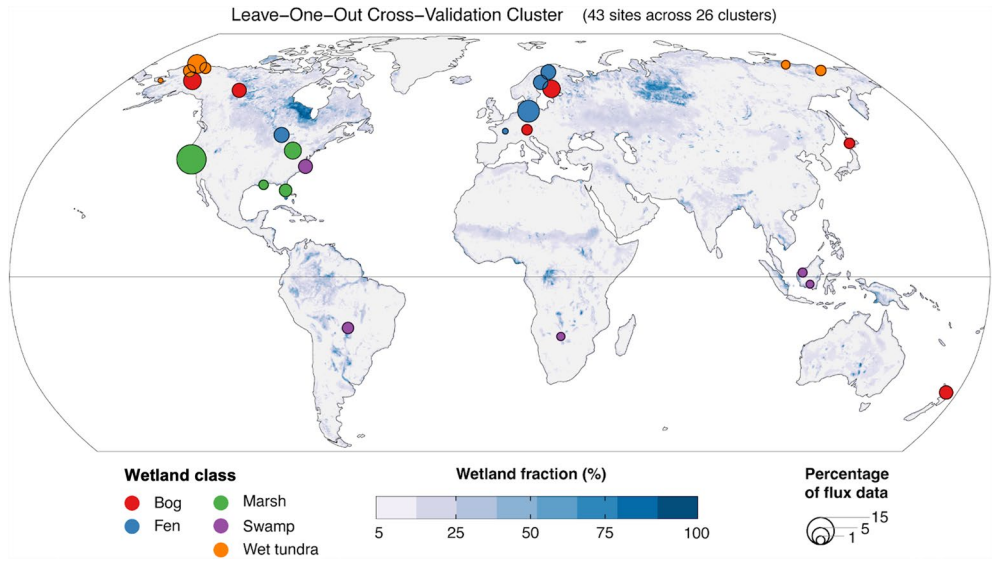

Figure 2. Location, class, and size of 26 globally distributed freshwater wetland clusters. Symbol sizes reflect the number of weekly CH_4 fluxes in each cluster expressed as a percent of the total data set considered in this study. Clusters combine data from sites that occur within 300 km of each other. At least one site was available from each major climate zone (Arctic-boreal, temperate, and tropical) and all major wetland classes were represented. Mean annual maximum wetland area fraction over 2000–2017 is shown from the Wetland Area Dynamics for Methane Modeling (WAD2M) product (Zhang et al., 2021).

bog (8), fen (8), marsh (10), swamp (6), and wet tundra (11) wetland classes and distributed across Arctic-boreal (20), temperate (16), and (sub)tropical (7) climate zones.

Weekly mean CH_4 fluxes were computed from half-hourly FLUXNET-CH4 Version 1.0 fluxes, which are available as a standardized gap-filled product (Delwiche et al., 2021). Although current wetland CH_4 emission products are resolved at monthly timesteps, a finer-resolution (here, weekly) increases training data size and helps capture sub-monthly functional dependencies between predictors and flux (Jung et al., 2020; Tramontana et al., 2016). Weekly fluxes were only retained when there was a minimum of 1 day (48 half-hours or ∼14%) of CH_4 observations. This gap-filling threshold was chosen to retain as much training data as possible while minimizing the errors introduced by filling long gaps (Dengel et al., 2013; Peltola et al., 2019). Most gaps were less than 5 hr in length, and the maximum possible gap-length was 12 days (Figure S1 in Supporting Information S1). A detailed study of machine learning-based gap-filling of eddy covariance CH_4 fluxes found that bias introduced by gap-filling remains small and consistent across gap-lengths of 12 days or less (Irvin et al., 2021). After applying the gap-filling threshold, the final flux data set consisted of 6,210 weekly observations spanning 2006–2018, with 96% of the data recorded after 2010, and 38% recorded after 2015.

Sites within 300 km of each other were grouped together, resulting in 26 clusters that were used for spatial leave-one-out cross-validation (LOOCV) of the machine learning model, where each training/validation fold consisted of all data except one hold-out cluster (Meyer et al., 2019) (Figure 2). Spatial LOOCV has previously been applied to evaluate models used in global upscaling of CO_2 and energy fluxes (Tramontana et al., 2016) and is suitable for making spatio-temporal predictions from spatially sparse time series data (Roberts et al., 2017). Further information on the wetland site class, geolocation, climate, site investigators, and data source is provided in Table 2. Additional information about the FLUXNET-CH4 sites considered for this study, including data digital object identifiers, site references, and source locations, are detailed in (Table S1), on the FLUXNET website and in Delwiche et al. (2021).

2.1.2. Predictor Data

A total of 140 candidate predictors were considered for data-driven upscaling. These candidate predictors were organized into five broad classes: climatic (e.g., temperature, precipitation), biometeorological (i.e., flux tower-measured air temperature, and ecosystem carbon and energy fluxes), land cover class and properties (e.g., vegetation class) (Tuanmu & Jetz, 2014, 2015), soil physical and chemical properties (e.g., clay content) (Hengl

et al., 2017; Lamarque et al., 2013), topography (e.g., slope), and vegetation indices (e.g., vegetation greenness) (Gao, 1996; Hall & Riggs, 2016; Huete et al., 2002; Jensen & Mcdonald, 2019; Myneni et al., 2015; Vermote, 2015; Wan et al., 2015). Each predictor class contains data from different sources (e.g., models, tower-observations, and/or remote sensing) and information content (e.g., spatial only, temporal only, or spatio-temporal). Gridded data products were extracted at pixels corresponding to tower locations and used directly in model development whereas tower-measured data, when available, were used preferentially in model training and then substituted by a globally gridded product to force the model during upscaling runs (e.g., tower-measured air temperature was mapped to the MERRA-2 reanalysis air temperature gridded product). Predictor data sources for each class are described in detail in (Text S1 and Table S2 in Supporting Information S1; Table S3).

Lead and lag times of one, two, or 3 weeks or months (for weekly and monthly data, respectively) were imposed on all temporally-resolved predictors, corresponding to multi-day and seasonal lead and lag timescales identified between wetland CH_4 fluxes, and temperature, eddy covariance-derived gross primary production (GPP), and soil moisture-related drivers (Delwiche et al., 2021; Knox et al., 2021). For MODIS data, monthly mean seasonal cycles (MSC) and other annual metrics (e.g., site-year mean, minimum, maximum, amplitude) were also derived. Quality control figures (example shown in Figure S2 in Supporting Information S1) were generated for all predictors and used primarily to identify and replace outliers with values from a proximate site within a cluster, an adjacent pixel (when sites were isolated within a cluster), or the site-year median (when varying in time). After deriving predictors, a total of 273 predictors were available for model training.

2.1.3. Predictor Selection

We used forward feature selection (FFS) to identify the optimal predictor subset from across all possible predictors to use in the final model (Gregorutti et al., 2017; Meyer et al., 2018). For each FFS step, we trained a random forest model algorithm (Breiman, 2001) with all possible predictors and computed the cost function (here, mean absolute error (MAE)) on validation data to identify the predictor(s) associated with the smallest MAE. In the first FFS step, we evaluated all possible predictor *pairs* (33,670 possible combinations) and selected the pair that resulted in the smallest MAE. In the second FFS step, we evaluated all possible *single* predictors (of 273), and selected the predictor that, when combined with the first pair, resulted in the smallest MAE. The second step was repeated 10 times to ensure the identification of a global MAE minima. More details are provided in Text S2 in Supporting Information S1 and predictors identified via FFS are visualized in Figure S4 in Supporting Information S1. FFS is suitable for spatial prediction tasks because it adds predictors when they reduce the cost function computed on held-out data. In contrast, recursive selection relies on importance rankings generated from the training data itself, which increases chances of overfitting (Meyer et al., 2018).

2.1.4. Cross Validation

After FFS, random forest models were re-trained using the final predictor set and their predictive performance was evaluated using leave one (cluster) out cross validation (cross validation, hereafter) (Meyer et al., 2018). During model training, a full hyperparameter grid-search was performed, which allowed for deeper, more complex trees (with a small minimum leaf node size). Model performance was evaluated by comparing predicted and observed CH_4 fluxes using the coefficient of determination (R^2), MAE normalized by flux standard deviation (nMAE), and bias, computed as the mean of residuals. Nash-Sutcliffe Efficiency (NSE) was also computed as an integrative measure of model performance. NSE is equal to R^2 when model bias is zero, and a NSE > 0 corresponds to a model performance better than simply taking the average of the data. Performance was evaluated with respect to three data set components: site mean flux, mean seasonal cycle (MSC) calculated as the average monthly anomaly from site mean, and interannual monthly anomalies from the MSC (Jung et al., 2020; Peltola et al., 2019; Tramontana et al., 2016). These components distinguish spatial prediction performance (annual site means) from monthly mean seasonal cycles (MSC), and interannual variability (weekly or monthly anomalies).

2.2. Upscaling Global CH_4 Flux

2.2.1. Final Model Ensemble

A final random forest model ensemble was trained to propagate uncertainties in training data to a gridded product. First, Monte Carlo simulations were used to create 1,000 simulated training datasets (each composed of CH_4 flux plus the final predictor variables) where each weekly observation was drawn from a normal (Gaussian) distribution,

except CH_4 flux for which the measurement uncertainty is defined as the variance of a double-exponential (Laplace) distribution (Irvin et al., 2021; Knox et al., 2019). For gridded products (e.g., static WorldClim data), dispersion around the true observations (distribution mean) was parameterized as a standard deviation of a 0.25° bounding box around each extracted pixel, for unitless MODIS enhanced vegetation index (EVI) products with the overall measurement uncertainty (0.015), for CH_4 flux as the weekly mean uncertainty (incorporating both random and gap-filling uncertainties (Irvin et al., 2021; Knox et al., 2019), and for tower-measured air temperature as a conservative estimate of standard temperature sensor precision (0.5°C) (Campbell Scientific, Utah). Second, 500 datasets were bootstrap sampled with replacement from the 1,000 simulations, and in each sample one site cluster was dropped at random. The site US-OWC was also always excluded due to its exceptionally high fluxes which could not be reproduced accurately by the best model (Figure 3a) and could introduce unnecessary bias error at other sites. Training on each bootstrap data set resulted in a final 500-member random forest model ensemble. Full details on Monte Carlo simulations are provided in Table S4 in Supporting Information S1.

2.2.2. Model Forcing

We applied the 500-member final model ensemble to an 18-year (216-month) time series of global grids of final predictor variables covering 2001–2018 from which the mean and standard deviation of predictions at each pixel globally was used as mean CH_4 flux and data-driven uncertainty. The reconstruction period (2001–2018) aligns with the current Global Carbon Project CH_4 modeling protocol and the monthly timestep of the WAD2M wetland extent product (Saunois et al., 2020; Zhang et al., 2021), though weekly data were used to train the machine learning model. For MODIS data, Google Earth Engine (Gorelick et al., 2017) was first used to prepare global monthly grids at 10-km resolution, excluding low quality observations, aggregating from 8-day values to monthly averages using the average of all good 8-day observations within a month, and aggregating from 500-m resolution to 10-km resolution using the average of all good quality 500-m or 1000-m observations within the 10-km pixel. Gaps of time-series of MODIS images were filled using the same methods as for site-based MODIS time series. All global grids were then resampled to a common 0.25° resolution and cropped to exclude Antarctica. Data were reprojected to WGS-84 geographic coordinates. Monthly positive and negative lags were imposed on grid stacks by shifting the stack by whole-month time steps, and linear interpolation between these shifts was used to create weekly time shifts. For each temporal predictor, a mean seasonal cycle stack was created by averaging rasters for each month across all available years for use in the global dissimilarity and tower constituency analyses (see Sections 2.2.3 and 2.2.4).

2.2.3. Wetland Area Products

We weighted each grid cell CH_4 flux prediction by fractional grid cell wetland extent to estimate CH_4 emissions using the primary WAD2M (Zhang et al., 2021) product and an alternate Global Inundation Estimate from Multiple Satellites GIEMS version 2; Prigent et al., 2020) global wetland map. We used WAD2M as the primary map because it was also used in the GCP bottom-up model ensemble allowing for direct flux prediction comparisons (Saunois et al., 2020). However, global wetland area is the largest source of uncertainty in wetland CH_4 emissions along with flux rates (Bloom et al., 2017; Melton et al., 2013), and the two maps enable a preliminary illustration of the sensitivity of our predicted emissions to different wetland area products. A robust evaluation of wetland extent uncertainties on upscaled emission estimates would require a detailed intercomparison such as that of Melton et al. (2013). Both WAD2M and GIEMS-2 maps were modified with several correction data layers to represent the monthly area covered by vegetated wetlands, excluding open water and coastal wetlands (Text S3 in Supporting Information S1) (Pekel et al., 2016). The maps were generated based on distinct multi-sensor methodologies estimating monthly inundated wetland area, to which a set of tailored correction layers and steps were applied to isolate only vegetated wetland area, following the methodology of Zhang et al. (2021).

2.2.4. Global Applicability and Tower Constituency

Similar to the challenges faced in global wetland CH_4 prediction using BU process models upscaled model predictions were extrapolated across a much larger spatial (Stell et al., 2021) and temporal (Chu et al., 2017) domain than that captured in the training data. Model extrapolation is likely to reduce the accuracy of flux predictions and can distort uncertainty estimates (Stell et al., 2021). To measure extrapolation during CH_4 upscaling, we first computed point-based dissimilarity (Hoffman et al., 2013) globally, defined as the minimum Euclidean distance between each grid cell-to-flux tower pair in predictor space, normalized by the mean distance among flux towers (Meyer & Buchta, 2020; Meyer et al., 2018, 2019). We then used the dissimilarity map to define a

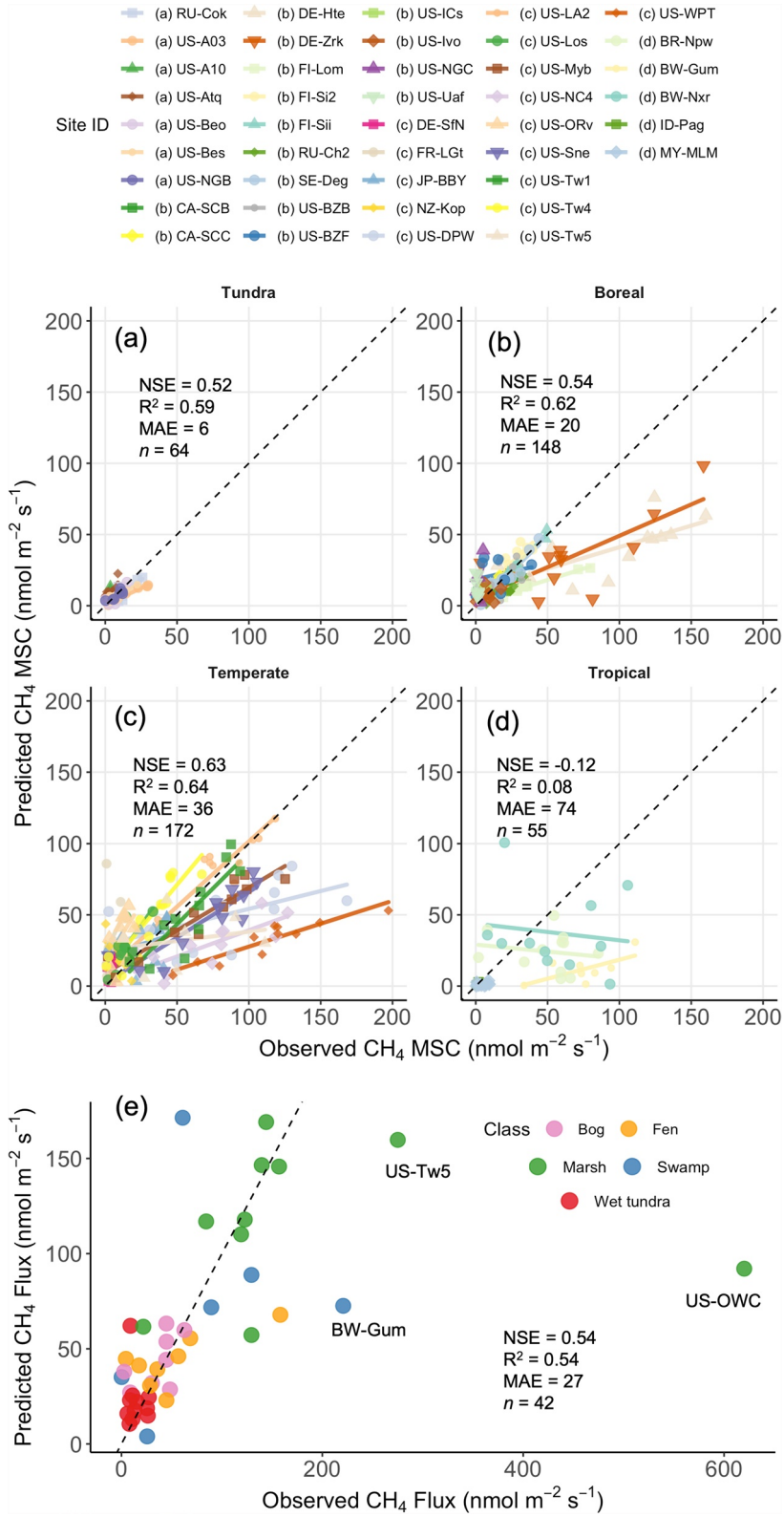

Figure 3. Random forest model predicted versus observed values for(a–d) the mean seasonal cycle (MSC) of methane (CH_4) flux for sites in (a) tundra, (b) boreal, (c) temperate, and (d) tropical climate regions and (e) the overall sites mean CH_4 flux, during cross validation. Although the US-OWC site is plotted in (a), it is excluded from calculation of Nash-Sutcliffe Efficiency (NSE), Coefficient of Determination (R^2), and mean absolute error (MAE; nmol m^{-2} s^{-1}) performance metrics, and sample count (n). The 1:1 fit is shown as a dashed black line.

monthly, global area of model applicability (AOA) and corresponding area of extrapolation, using a dissimilarity threshold (Meyer & Pebesma, 2022). Finally, we defined the global constituency of each site cluster to identify which training conditions dominate global predictions and further evaluate the plausibility of global extrapolations. Each pixel was assigned as a constituent of the site cluster that was closest in predictor space (Hargrove & Hoffman, 2004). We propose that combining constituencies with AOA provides a semi-quantitative and informative approach to evaluating global representativeness and extrapolation confidence in upscaling models. We demonstrate this approach by identifying the regions that are most like training conditions and *which* training conditions each region is most similar to. This approach allows us to diagnose CH_4 flux prediction patterns in extrapolations.

3. Results

3.1. Predictive Modeling

3.1.1. Model Predictors

Only the first six predictors from the FFS were used in the final upscaling model, as they accounted for ~85% of the MAE reduction. These predictors were tower-measured air temperature, with and without a 2-week lag, MODIS EVI with a 3-week lag, mean temperature of the driest quarter, precipitation of the wettest month, and vegetation canopy height (Figure S4 in Supporting Information S1). It is notable that three of the final six predictors were temperature related. An additional five predictors not included in the final model extended the FFS to the MAE minimum (Table S5 in Supporting Information S1) and included MODIS snow cover, tower-measured GPP with a 2-week lead, and the annual minimum in MODIS EVI. The random forest variable importance rankings deviated slightly in order from FFS, with air temperature and the two static climatological predictors ranked as the most important final predictors, while canopy height and MODIS EVI were less important (Figure S5 in Supporting Information S1). A strong exponential dependency was observed between CH_4 flux and air temperature, likely explaining its dominance in the variable importance, while more complex and/or less tightly correlated dependencies were observed between CH_4 flux and the other predictors (Figure S6 in Supporting Information S1). Temperature hysteresis was not reproduced in the model in many sites where it has been observed in the CH_4 flux observations (Chang et al., 2021) (Figure S7 in Supporting Information S1).

3.1.2. Cross Validation Performance

Model residuals (errors) were normally distributed around zero at bogs, fens, swamps, and wet tundra sites (Figure S8 in Supporting Information S1). At freshwater marshes, residuals displayed more negative outliers due to one site (US-OWC) that displayed exceptionally high CH_4 fluxes (>10× higher than overall median) that the model did not reproduce. The fluxes at US-OWC are plausible because the site is situated in a eutrophic estuarine marsh on the southern shore of Lake Erie, USA, (Rey-Sanchez et al., 2018), which displays very high rates of sediment methanogenesis (Angle et al., 2017). However, evaluating the global scale emissions from eutrophic wetland is beyond the scope of this first wetland upscaling effort and therefore, hereafter, cross validation metrics are reported with the exclusion of the exceptionally high fluxes at site US-OWC.

When predicting site mean CH_4 fluxes during cross validation (Figure 3a), the model achieved an NSE of 0.54 and nMAE of 0.42, and low model bias (2.6 nmol m^{-2} s^{-1}) relative to the overall site mean CH_4 flux (61.5 nmol m^{-2} s^{-1}). Site mean CH_4 flux errors were not spread evenly among wetland classes. Mean absolute errors (MAE) increased from wet tundra (11.7 nmol m^{-2} s^{-1}) to bogs (13.2 nmol m^{-2} s^{-1}), fens (25.7 nmol m^{-2} s^{-1}), marshes (35.2 nmol m^{-2} s^{-1}), and swamps (62.2 nmol m^{-2} s^{-1}). However, after normalizing by flux standard deviation, nMAE increased from marshes (0.53) to fens (0.54), bogs (0.64), swamps (0.78), and wet tundra (1.48), reflecting low CH_4 flux variability at wet tundra and high flux variability at marshes (Table S6 in Supporting Information S1). Mean seasonal cycles were, overall, predicted by the model with comparable accuracy to the site means (NSE = 0.53 and nMAE = 0.41). However, model prediction performance on MSC differed greatly by climate region (Figures 3b–3e), and decreased from higher at temperate (R^2 = 0.64; NSE = 0.63; nMAE = 0.37), boreal (R^2 = 0.62; NSE = 0.54; nMAE = 0.37), and tundra sites (R^2 = 0.59; NSE = 0.52; nMAE = 0.53), to lower at tropical sites (R^2 = 0.08; NSE = −0.12; nMAE = 0.81). Although all tropical sites were swamps, the pattern was less clear when sites were grouped by wetland class rather than by climate region, because the model achieved high MSC performance at a temperate swamp, US-NC4 (R^2 = 0.78; NSE = 0.50; nMAE = 0.58) (Table S6 in Supporting Information S1). Finally, the model was unable to predict interannual monthly anomalies from the MSC (NSE = −5.11; $R^2 = 2e^{-3}$; nMAE = 37.2).

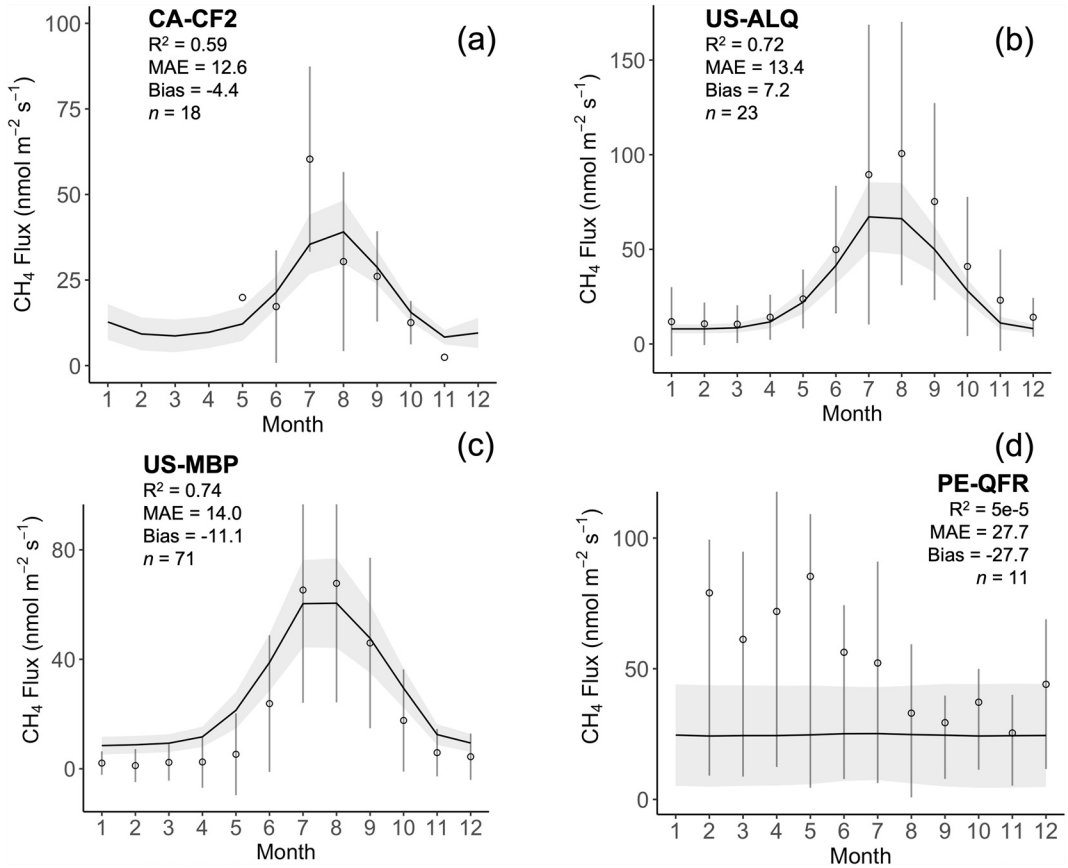

Figure 4. The mean seasonal cycle of model-predicted CH_4 flux (solid black line) and uncertainty range (gray ribbon) compared against monthly observed mean fluxes (open circles) and (a–c) standard deviation or (d) 25th–75th percentiles (vertical bars) for four AmeriFlux test sites. The model reproduced mean fluxes and the seasonal cycle best at (a) a boreal fen (CA-CF2; Tenuta, 2020), (b) a temperate fen (US-ALQ; Olson, 2018), and (c) a boreal bog site (US-MBP; Roman et al., 2021), whereas seasonal cycle performance was not reproduced at (d) one humid tropical forest site (PE-QFR; Roman et al., 2020), which is an upland site that experiences very wet soil conditions and supports substantial but highly variable CH_4 fluxes.

3.2. Global Upscaling

3.2.1. Unweighted Global Methane Flux Predictions

Gridded freshwater wetland CH_4 flux predictions, before being weighted by wetland extent (Figure S9 in Supporting Information S1), were compared pixel-wise to the original 43 training and four additional test flux tower sites (Figure 4). Globally, the model achieved an R^2 of 0.53, NSE of 0.51, nMAE of 0.35, and Bias of −6.0 n mol m^{-2} s^{-1}. The slight improvement compared to cross-validation performance is expected as all the training data appears many times across the 500 bootstrap-sampled data sets models. Examples of gridded product predictions at training sites and their uncertainties are shown in Figure S10 in Supporting Information S1. Notably, the model also performed well at three of four additional test sites (Figure 4), reproducing the site mean with nMAE ranging from 0.39 to 0.49 at the boreal and temperate sites. However, as was observed in training evaluations, predictions at the tropical forest test site (PE-QFR) did not reproduce the seasonal signal and exhibited the largest nMAE (1.35).

3.2.2. Global Model Applicability and Tower Constituency

Confidence in gridded model predictions was also evaluated semi-quantitatively using global dissimilarity and tower constituency analyses. Global dissimilarity was low, even in areas geographically distant from existing towers (Figure S11 in Supporting Information S1), suggesting that the model was not forced to extrapolate far from training conditions, even when making global predictions. The most dissimilar regions were Eastern

Siberia, South Asia (e.g., India), the Sahel and Congo regions of sub-Saharan Africa, and the Amazon basin. The most dissimilar sites were in tropical and temperate regions, with lower dissimilarity in boreal and tundra regions (Figure S11 in Supporting Information S1). The tower constituency map (Figure S12 in Supporting Information S1) visualized *how* the model is likely to extrapolate from training data at sites to geographically distant regions, given their similar predictor conditions. The Amazon and Congo basins, which lack flux towers, fall within the constituencies of two SE Asian towers (ID-Pag and MY-MLM). The constituency of the Brazilian Pantanal tower, BR-Npw, encompasses other wet savanna and monsoon regions of South America, Africa, and Asia. Much of central and western Australia and semi-arid tropics of South America and the Sahel are assigned to the subtropical Botswana site constituencies (BW-Nxr and BW-Gum) or Mediterranean California (USA) site constituencies (US-Myb, US-Tw1). These semi-arid or arid tropical constituency assignments are reasonable on the basis of climate; however, it is important to note that the wetland sites in these environments where the aforementioned towers are located are associated with large inland deltas (i.e., the Okavango Delta, Botswana, or the Sacramento Delta, California), which provide water to support marsh and swamp wetlands. The final model did not include surface hydrological variables and therefore extrapolation to these regions may not be easily evaluated by the dissimilarity and constituency analysis employed here. No relationship was observed between model error or variance and site-month dissimilarity (Figure S13 in Supporting Information S1) that could be used to scale errors in regions of extrapolation (Jung et al., 2020).

3.2.3. Wetland Area-Weighted CH_4 Fluxes

Time series of mean freshwater wetland CH_4 fluxes from UpCH4 (2001–2018), weighted by WAD2M pixel wetland area, displayed regional patterns that reflected the interaction of wetland area and the model's flux predictions (Figure 5a). The highest wetland area-weighted fluxes (>30 mg CH_4 m^{-2} d^{-1}) were predicted in both high- and low-latitude regions with extensive wetland area (e.g., Hudson Bay Lowlands (HBL), Congo Basin) and in semi-arid regions where wetland cover is low but model flux predictions were very high (Figure S9 in Supporting Information S1). Relative uncertainties (Figure 5b) were the smallest for high-latitude high emission (>20 mgCH_4 m^{-2} d^{-1}) hotspots associated with the HBL and West Siberian Lowlands (WSL) wetland complexes and were largest in monsoon regions of relatively low flux (<5 mgCH_4 m^{-2} d^{-1}) sandwiched between the semi-arid and humid tropics.

Upscaled (UpCH4) fluxes were compared to three alternative global wetland CH_4 emission datasets using either the same WAD2M wetland area (i.e., GCP BU ensemble; Figure 5c), or variable wetland products (i.e., GCP TD ensemble; Figure 5e). At high-latitude wetland complexes (HBL and WSL), UpCH4 predicted slightly lower emissions than the mean of the GCP BU ensemble, similar to the GCP TD ensemble, but higher than WetCHARTS. Given the same WAD2M wetland area was applied in upscaling as in the GCP BU ensemble, this difference can be attributed to lower predicted CH_4 fluxes by UpCH4. At mid-to-low latitudes, UpCH4 predicted 10–30 mgCH_4 m^{-2} d^{-1} higher fluxes than the other products from the semi-arid tropics, including the Sahel and Horn of Africa, central and western Australia, and western Asia, while also predicting 20–30 mgCH_4 m^{-2} d^{-1} lower fluxes from the humid tropics, including the large wetland complexes of the Amazon and Congo Basins, and SE Asia, especially in Indonesia and Malaysia. Again, as the GCP BU ensemble used the same wetland area, these product differences can be attributed to wetland CH_4 flux rates rather than extent. The regional pattern of tropical emissions in WetCHARTS was more similar to the GCP TD ensemble pattern than the GCP BU ensemble pattern, which differ as described in Figure 1.

3.2.4. Temporal Trends and Spatial Patterns in UpCH4 Emissions

Upscaling model (UpCH4) emissions using both WAD2M and GIEMS-2 wetland area products are seasonal, with a JJA peak that corresponds with the expansion of wetland area, the warmest soil temperatures, and peak productivity at northern high-latitudes, (Figure 6). As indicated by globally integrated fluxes, UpCH4-WAD2M emissions correspond well with the average of the GCP BU ensemble range, whereas UpCH4-GIEMS-2 emissions are significantly lower, although the amplitude of the seasonal cycle is larger. No clear long term annual trend is predicted by UpCH4 (or other models), though interannual variability is apparent, driven by wetland extent changes.

The latitudinal pattern of UpCH4 emissions using either wetland extent lacks the year-round elevated equatorial band found in the GCP products (Figure 7). The temporal pattern of this band varies slightly between the GCP BU and TD ensemble, which show a clear seasonal peak in AMJ. The UpCH4-WAD2M upscaling still

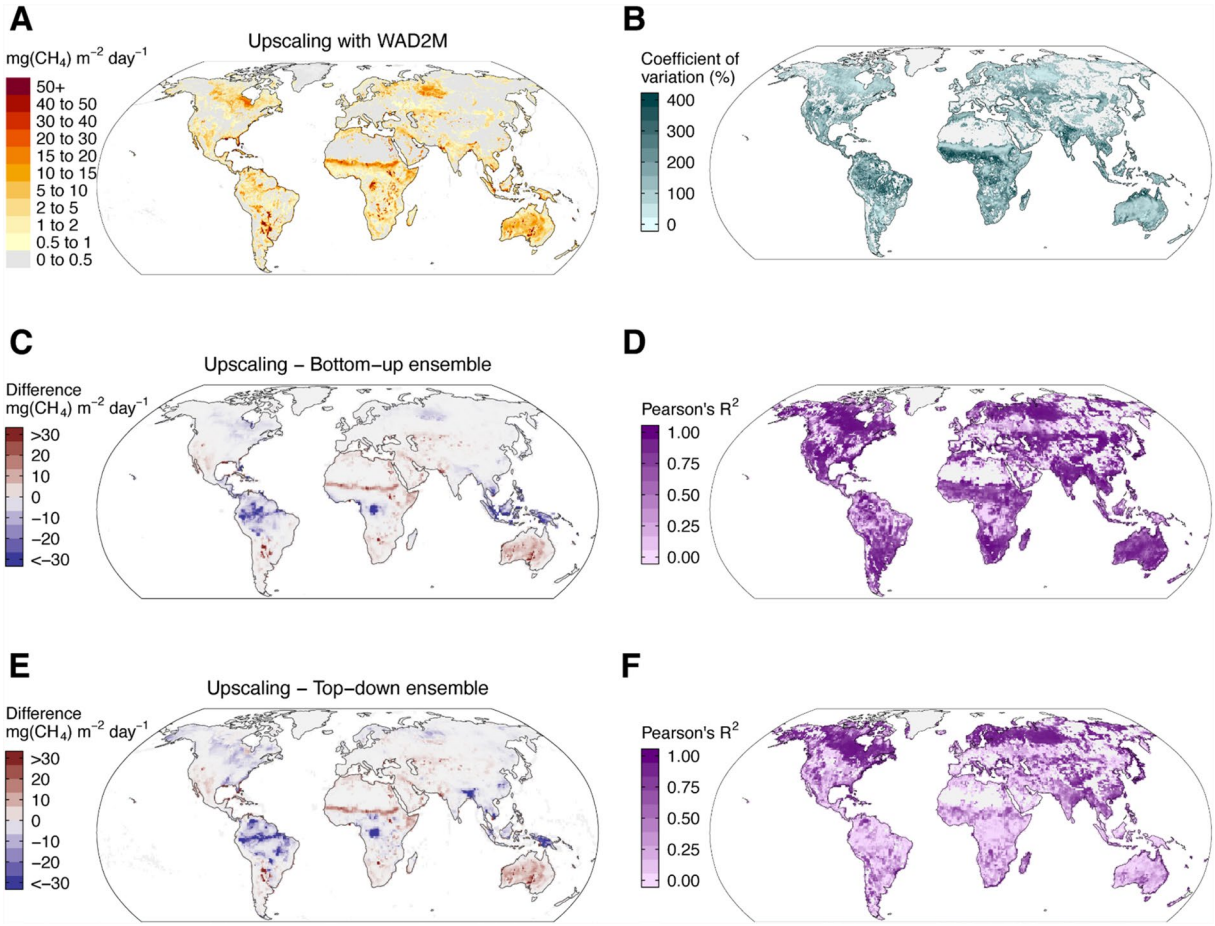

Figure 5. Global maps of: (a) Upscaled (UpCH4) mean 2001–2018 CH_4 flux using WAD2M wetland area (b) CH_4 flux uncertainty computed as 1 standard deviation of the random forest ensemble expressed as a percent of the mean (i.e., coefficient of variation) 2001–2018 flux; (c, e) the mean 2001–2018 CH_4 flux from the Global Carbon Project (GCP) bottom-up/top-down process model ensemble (also using WAD2M) subtracted from the UpCH4 mean (a); (d, f) the correlation, expressed as the correlation coefficient of determination (R^2), between the mean seasonal cycle (MSC) of UpCH4 and the GCP bottom-up/top-down ensemble.

produces a similar global total to the GCP BU ensemble from a much larger latitudinal range for moderate fluxes. Increased UpCH4 emissions are observed for narrow tropical N hemisphere (10–20°N) and subtropical S hemisphere (20–30°S) bands, likely related to high flux predictions in the semi-arid climate whereas GCP BU and TD ensembles show enhanced fluxes around 30°N and 10°N.

4. Discussion

4.1. Model Performance

UpCH4 achieved comparable metrics based on our spatial cross-validation (global $R^2 = 0.50$; Northern site $R^2 = 0.48$) to a recent northern high-latitude upscaling of CH_4 fluxes from 25 eddy covariance towers (Peltola et al., 2019), despite a larger and more variable global flux data set. UpCH4 also achieved better performance than global upscaling models for net ecosystem exchange of CO_2 ($R^2 < 0.5$) (FLUXCOM; Tramontana et al., 2016). As noted by Peltola et al. (2019), net CH_4 fluxes measured with eddy covariance may be difficult for machine learning models to reproduce with limited datasets as they display complex and non-linear behavior, and are subject to storage effects and lags (Knox et al., 2021; Sturtevant et al., 2016) due to underlying CH_4 production, oxidation, and transport processes (Bridgham et al., 2013; Chang et al., 2021) identified substantial hysteresis in the seasonal temperature dependency of wetland CH_4 flux and Delwiche et al. (2021) identified lags and leads of various lengths between peak growing season air temperature and peak CH_4 flux. Temperature lags may be due to the substrate control of CH_4 production and would agree with coherence with ecosystem production (Knox

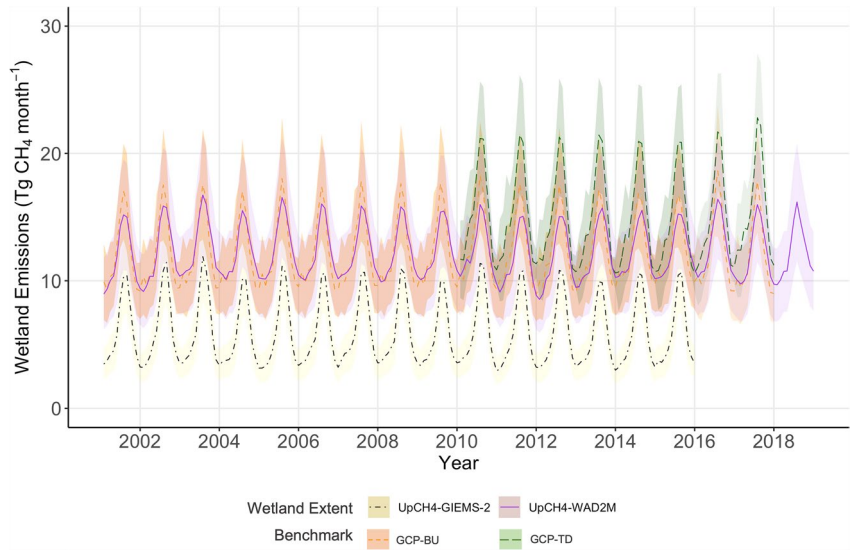

Figure 6. Monthly time series of upscaled wetland CH_4 emissions (UpCH4) using WAD2M (2001–2018) and GIEMS-2 (2001–2015) compared with the 13-member GCP BU ensemble (2001–2017) and 17-member subset of GCP TD ensemble (2010–2017). The UpCH4 mean and spread are difficult to distinguish from those of the GCP BU ensemble due to small differences between the two estimates.

et al., 2021; Mitra et al., 2020). In UpCH4, leads and lags of various lengths were imposed on all temporal predictors to address part of this complexity and two lagged predictors were selected in the final model (i.e., lagged temperature and EVI), indicating their utility when using non-temporal machine learning algorithms such as random forests. However, observed hysteresis between temperature and CH_4 flux was not reproduced (Figure S7 in Supporting Information S1), suggesting that simply imposing lags on predictors cannot capture the complex biogeochemical processes that drive intra-seasonal variability (Chang et al., 2021). Deep learning models able to learn temporal dependencies, such as Long Short Term Memory (LSTM) neural networks, and/or able to incorporate process knowledge as constraints, could be considered as the core algorithms for CH_4 upscaling; however, this would require adapting model architecture for spatio-temporal predictions (Reichstein et al., 2019).

Performance improvements may also be expected if wetland CH_4 fluxes measured using eddy covariance can be partitioned into ebullition, diffusion, and plant-mediated transport pathways as has been done to partition diffusion and ebullition in lakes and bogs (Iwata et al., 2018; Ueyama et al., 2022). Each of these transport

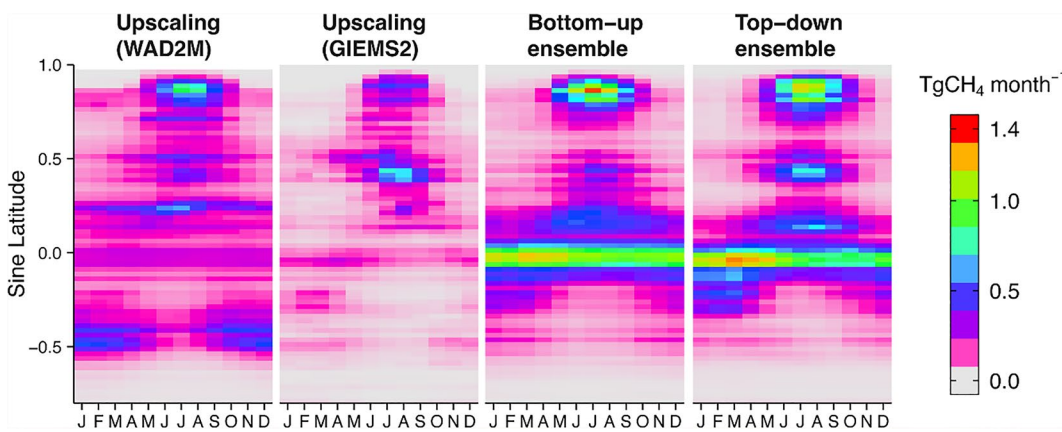

Figure 7. Average monthly freshwater wetland CH_4 emissions ($TgCH_4$ $month^{-1}$) for ~1°latitude bands from UpCH4 for the two wetland maps (WAD2M and GIEMS-2) over 2001–2017; three alternative global datasets (Bottom-up GCP ensemble over 2001–2017, Top-down ensemble over 2010–2017, and WetCHARTS v1.0 over 2000–2010) are also shown. Average CH_4 flux estimates from five data sources showing differences in total emissions.

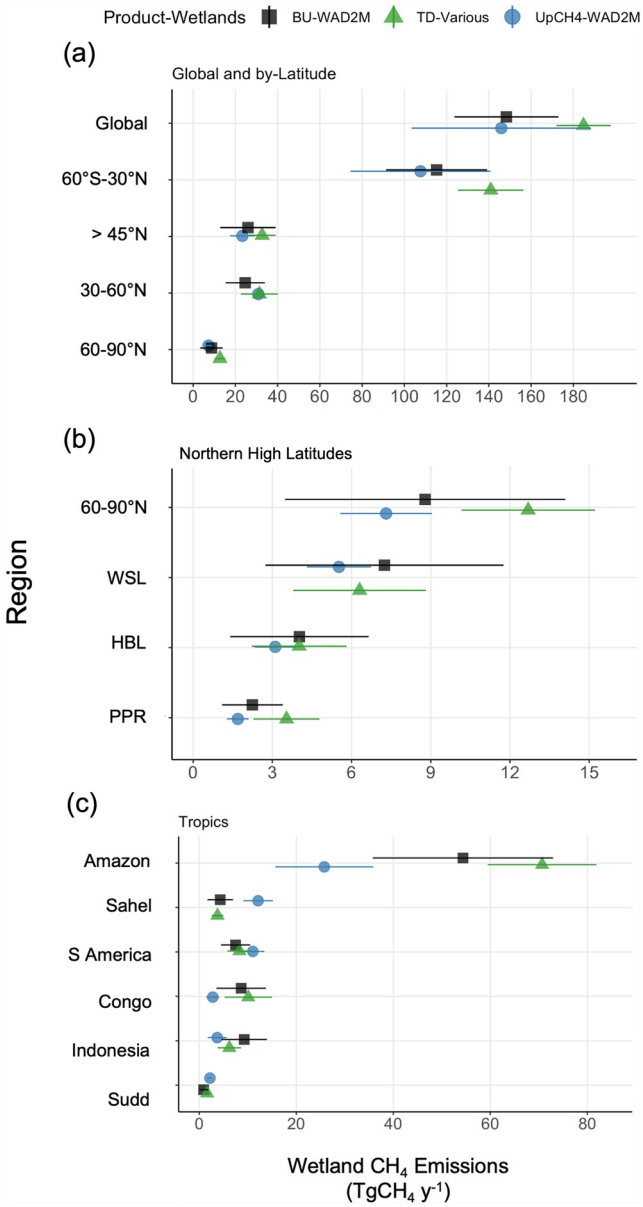

Figure 8. (a) Global and (b, c) regional comparisons of annual wetland CH_4 emissions from upscaling (UpCH4; blue circles), GCP BU ensemble (bottom-up; black square), and TD inversion ensemble (top-down; green triangle). Regional wetland complex initialisms for West Siberian Lowlands (WSL; 52–74°N, 60–95°W), Hudson Bay Lowlands (HBL; 50–60°N, 75–96°W); Prairie Pothole Region (PPR; 42–55°N, 92–115°W).

processes is regulated by distinct drivers and produces flux signals that may be more directly attributed to these processes when separated from each other. Improved models may also be possible if we can reconcile current wetland classifications with real wetland differences in the mean and/or variance of methane fluxes.

4.2. Budget Comparisons

Global freshwater wetland CH_4 emissions and overall uncertainties during the period 2001–2018 from UpCH4 were 146 ± 42.7 $TgCH_4\ y^{-1}$ using WAD2M wetland area (Figure 8a), which closely matches emissions from the GCP bottom-up model ensemble for 2007–2018 (mean 149 $TgCH_4\ y^{-1}$; range 102–182 $TgCH_4\ y^{-1}$) but is substantially lower (31–44 $TgCH_4\ y^{-1}$) than GCP top-down emissions (mean 181 $TgCH_4\ y^{-1}$; range 159–200 $TgCH_4\ y^{-1}$) (Saunois et al., 2020) and WETCHIMP emissions (190 ± 39 $TgCH_4\ y^{-1}$) (Melton et al., 2013). Notably, UpCH4-WAD2M global emissions were also similar to that of Nzotungicimpaye et al. (2021) (158.6 $TgCH_4\ y^{-1}$) who implemented WETMETH—a CH_4 process model, within the UVic Earth System Climate Model, and Ma et al. (2021) (148 $TgCH_4\ y^{-1}$) who used satellite-based observations to refine estimates from 42 BU process models. Using GIEMS-2 corrected for only vegetated wetland area, rather than WAD2M, cut UpCH4 global emissions by more than half (71.8 ± 22.6 $TgCH_4\ y^{-1}$), highlighting the sensitivity of total emissions to the wetland map used (Figures S14–S16 in Supporting Information S1).

At the global scale, UpCH4 mean CH_4 emissions agreed more closely with GCP BU emissions for tropical (60S–30N), temperate (30–60 N), northern (>45 N), and Arctic (60–90 N) latitudinal ranges compared to GCP TD estimates, which were consistently higher, most notably in the tropics. However, UpCH4 uncertainty ranges overlapped with both TD and BU estimates and supported the previously reported observation that ~68% of wetland CH_4 emissions originate from tropical wetlands (Saunois et al., 2020). Within major wetland complexes at northern latitudes (Figure 8b), UpCH4 agreed more closely with lower TD estimates for the WSL, and UpCH4 emissions for the HBL and Prairie Pothole Region were also lower than both BU and TD estimates. For tropical wetland regions, estimates diverged more significantly between UpCH4 and the GCP ensembles (Figure 8c) with larger uncertainties likely due to lack of data in tropics. Although UpCH4 agreed closely with BU estimates for the tropical latitude total, UpCH4 emissions were much (~3x) higher for the semi-arid monsoon Sahel compared to either GCP ensemble product, and much lower emissions estimated for the humid tropical forested wetlands of the Amazon, Congo, and the Indonesian archipelago.

4.3. Interpreting Product Differences

Interpreting similarities and differences in spatial patterns between global estimates of wetland CH_4 emissions is challenging because they arise from differences in both wetland fluxes and wetland area (akin to Melton et al., 2013). However, some broad conclusions can be drawn. GCP BU ensemble comparisons are informative at subtropical and tropical latitudes because in this study the shared use of the WAD2M wetland area enables model differences to be fully attributed to flux processes. As UpCH4 global upscaling emissions agreed most closely with GCP BU ensemble, it is notable that tropical and subtropical regional patterns diverged substantially (Figure 5c). The semi-arid to humid tropics gradient is inverted in UpCH4 when compared with GCP and thus these models produce similar global totals but different regional distributions. Given that inversion and observational studies in the Amazon Basin indicate a very large CH_4 source (Devol et al., 1988; Gauci et al., 2022; Pangala et al., 2017), it is plausible that

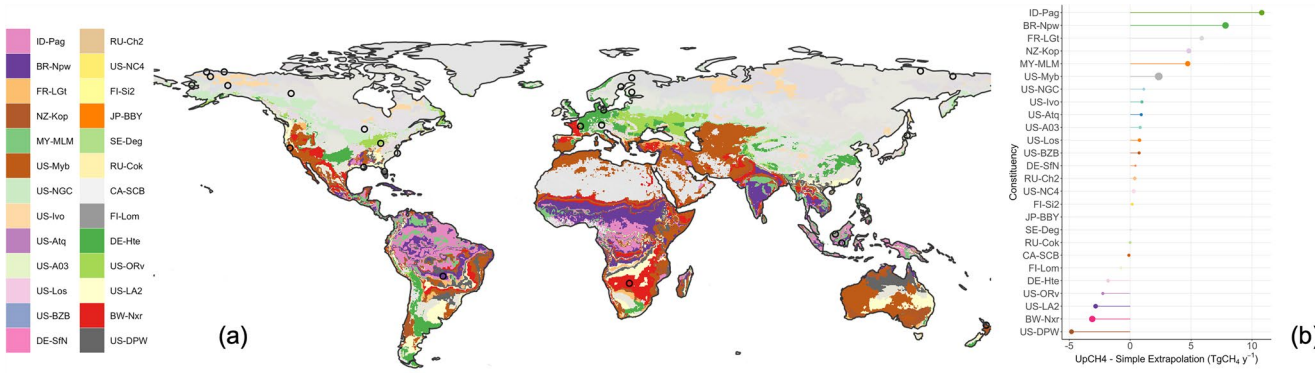

Figure 9. (a) Global wetland CH_4 flux site-cluster constituencies for which pixels are assigned to one of 26 site-clusters (colors) based on similarity to UpCH4 predictor conditions at the sites within that cluster. (b) The ranked difference in predicted CH_4 emissions ($TgCH_4\ y^{-1}$) between UpCH4 predictions for a given constituency (colors) and a simple extrapolation of the monthly mean flux for the site cluster to the entire constituency (ignoring flux data from other sites). Constituencies and their upscaling-based emission estimates are likely to be insensitive to adding additional sites (more transparent colors in map (a)) when emissions differences are close to zero (short segments in (b)). In contrast, constituencies and emission estimates are likely to be more sensitive to additional sites (opaque colors in map (a)) where absolute emissions differences are large (long segments in (b)). The size of the segment end points is proportional to the annual constituency CH_4 flux. A full constituency map without variable color transparency is provided in Figure S12 in Supporting Information S1.

the upscaling pattern is biased, and that semi-arid and humid tropical wetland emissions are over- and under-estimated, respectively. However, observations of exceptional CH_4 emissions from seasonal wetlands in the hot semi-arid tropics (e.g., BR-Nxr and BR-Gum, the Okavango Delta wetlands in this study) underscore that more measurements could support more robust validation of these patterns. If semi-arid/monsoon wetland CH_4 sources are currently being underestimated in BU models, this missing source would significantly close the gap between current BU and TD global wetland CH_4 emissions estimates.

4.4. Data Limitations

The training data contained a geographic bias with many more temperate and boreal northern hemisphere sites than tropical and southern hemisphere sites. Only 9% of the training data (5 sites) were acquired from south of 23°N, despite the tropics accounting for the vast majority of the non-frozen global wetland area and an estimated two-thirds of global CH_4 emissions (Melton et al., 2013; Saunois et al., 2020). As a result, tropical CH_4 flux predictions were based on flux data from a few towers that are unlikely to be representative of the entire tropical region and, correspondingly, tropical prediction uncertainties from UpCH4 were high (Figure 5b). For instance, training data for semi-arid subtropical regions were dominated by site clusters in the Sacramento Delta, United States (US-Myb, US-Sne, US-Tw1, US-Tw4, US-Tw5), and the Okavango Delta, Botswana (BW-Gum, BW-Nxr), which together accounted for ~20% of all training data. In both regions, minerotrophic deltaic wetlands are dependent on large seasonal or permanent allochthonous riverine inputs from regional-scale drainage basins, and thus sustain very productive marsh or swamp conditions conducive to high CH_4 fluxes that contrast sharply from the surrounding dryland environments (Hemes et al., 2019; Knox et al., 2015). The lack of wetland data from more varied hydrologic classes of wetlands (e.g., riverine, isolated or rainfed, i.e., ombrotrophic) under similar climate conditions, combined with the biases described above, may have led to the high predicted CH_4 fluxes for the semi-arid subtropics.

The CH_4 flux predictions and uncertainties (Figure 5), combined with the tower constituency and model applicability maps (Figures S10 and S11 in Supporting Information S1), provide the first global survey for situating new eddy covariance measurement sites based on both CH_4 flux and environmental information. Moreover, an expanded site constituency representativeness analysis confirmed that global wetland CH_4 upscaling will benefit most from additional tropical wetland flux data (Figure 9). The large humid tropical site constituencies (e.g., ID-Pag—a SE Asian peat swamp forest) imply very wide model extrapolation across the Amazon and Congo Basins, where there is currently no available CH_4 flux data and where fluxes could be quite different. Similarly, the largest CH_4 flux uncertainties were observed for transitional climate regions between the semi-arid and humid tropics (Figure 5) and establishing towers to bracket or traverse these regions could help capture important gradients in tropical wetland conditions relevant to CH_4 flux variability. Overall, the lack of observations for

large regions of the tropics combined with the distinctive hydrological regimes characterizing different tropical wetland regions (Apers et al., 2022; Dalmagro et al., 2018), likely account for the upscaling discrepancy with past and typically higher tropical emission estimates, such as for the Amazon Basin (Figure 6). Refined methods to evaluate CH_4 flux tower network representativeness along different dimensions of variability could result in improved estimates, as has been undertaken at regional scales (Malone et al., 2022; Villarreal & Vargas, 2021). Similarly, use of more finely resolved spatial forcing data can more accurately represent wetland conditions and may improve model functional responses (e.g., 30-m resolution in Bansal et al., 2023).

In addition to geographic bias at global scales, wetland flux towers are also likely biased with respect to average grid cell conditions. Wetlands are often a minority cover type at landscape scales meaning that training data based on grid cell averages alone, for example, MODIS vegetation products may not be representative of tower conditions (Chu et al., 2021). Applied to the case described above of minerotrophic swamp or marsh surrounded by drylands, scale-mismatch will likely result in erroneous wetland vegetation indices, such as greenness and phenology metrics, as well as derived products such as GPP, which many studies, including this study, indicate as important for predicting CH_4 fluxes(e.g., Bridgham et al., 2013; Chang et al., 2021; Delwiche et al., 2021; Knox et al., 2019; Whiting & Chanton, 1993).

Spatial biases at global and landscape scales could be addressed in future work by improving the geographic and grid cell representativeness of CH_4 flux and predictor data, respectively. Developing methods to reconcile and integrate chamber flux data with tower flux data could be prioritized to gain information from the large amount of existing chamber data, such as in Bansal et al. (2023), Kuhn et al. (2021), and Turetsky et al. (2014). Chamber methods offer a relatively inexpensive and accessible means to gather data in underrepresented regions (Harriss & Matson, 2009), may provide insights into patch-scale effects when paired with tower data via environmental response functions methods (Xu et al., 2017), and can extend the temporal representativeness of flux data (Chu et al., 2017).

5. Conclusions

We develop a wetland CH_4 flux upscaling workflow (UpCH4) for eddy covariance flux data and evaluate global CH_4 emission predictions. Extratropical estimates from UpCH4 can provide insights from comparisons to existing CH_4 model predictions. The use of UpCH4 tropical wetland emissions estimates should include consideration of uncertainties and joint use of additional regional data constraints is strongly encouraged. UpCH4 estimates average annual freshwater wetland CH_4 emissions of 146 $\pm$ 42.7 $TgCH_4\ y^{-1}$ for 2001–2018 which aligns closely with the most recent GCP BU estimates (149 $TgCH_4\ y^{-1}$) (Saunois et al., 2020) and a hybrid study that constrained a large ensemble of process models with satellite data (148 $TgCH_4\ y^{-1}$) (Ma et al., 2021). UpCH4 emission uncertainties were larger than, and overlapped with, both GCP BU and TD estimates, and the sensitivity to wetland extent products is illustrated by the halving of the global emissions total (71.8 $\pm$ 22.6 $TgCH_4\ y^{-1}$) when using GIEMS-2 corrected for only vegetated wetland area. All gridded emissions products are available via ORNL DAAC (https://doi.org/10.3334/ORNLDAAC/2253). UpCH4 is most suitable for comparison to other bottom-up and top-down models within temperate, boreal, and arctic climate zones from 2010 onwards, and will improve in tropical regions as EC data coverage is expanded over time.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

Different gridded CH_4 flux outputs were generated with and without applying the WAD2M or GIEMS-2 wetland extent masks (Text S1 and Figure S3 in Supporting Information S1). Thus, final gridded products are: (a) gridded unweighted wetland CH_4 fluxes in nmol $CH_4\ m^{-2}\ s^{-1}$ and g $C\text{-}CH_4\ m^{-2}\ d^{-1}$; (b) wetland area-weighted fluxes in mg $CH_4\ m^{-2}\ d^{-1}$; and (c) wetland area-weighted fluxes in $TgCH_4$ grid $cell^{-1}\ month^{-1}$. Main data products are available via DOE ORNL DAAC and Zenodo (UpCH4: https://doi.org/10.3334/ORNLDAAC/2253; WAD2M: https://doi.org/10.5281/zenodo.5553187). The GIEMS-2 wetland extent product is available at request from Catherine Prigent (catherine.prigent@obspm.fr). Code notebooks for random forest model development

and validation are available via Zenodo (https://doi.org/10.5281/zenodo.7978099). The Global Carbon Project methane budget is available at https://doi.org/10.18160/GCP-CH4-2019.

Acknowledgments
Primary support for the project came from the Gordon and Betty Moore Foundation (Grant GBMF5439, "Advancing Understanding of the Global Methane Cycle") to Stanford University, with additional support from the John Wesley Powell Center for Analysis and Synthesis of the U.S. Geological Survey (USGS "Wetland FLUXNET Synthesis for Methane" working group). Further support came from the National Aeronautics and Space Administration Carbon Monitoring System award to Lawrence Berkeley National Laboratory and the University of Illinois Chicago (Zhu Prototyping a monitoring system of global wetland CH4 emissions with machine learning and satellite remote sensing) (20-CMS20-0039; NNH20ZDA001N). Observations from the Atmospheric Radiation Measurement (ARM) user facility are supported by the U.S. Department of Energy (DOE) Office of Science user facility managed by the Biological and Environmental Research Program. Work at Argonne National Laboratory was supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under contract DEAC0206CH11357. Sarah Feron acknowledges the support of ANID (ANILLO ACT210046) and CORFO (19BP-117358). Work at USGS was additionally supported by the Biological Carbon Sequestration program. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. ICOS Finland was funded by the Academy of Finland, the Ministry of Transport and Communications, and University of Helsinki. We thank Dario Papale, Markus Reichstein, Martin Jung, Martijn Pallandt, Hanna Meyer, and Olli Peltola for consultations on upscaling methodology. We thank Catherine Prigent for the use of the GIEMS-2 wetland data set (Prigent et al., 2020). Selected data of this study are available through the USGS ScienceBase repository (Krauss et al., 2018). We thank Tamara Wilson and Geneva Chong (USGS) for signed reviews and two further anonymous reviewers, and the research teams contributing to the FLUXNET-CH4 Synthesis Activity.

References

Abernethy, S., O'Connor, F. M., Jones, C. D., & Jackson, R. B. (2021). Methane removal and the proportional reductions in surface temperature and ozone. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *379*(2210), 20210104. https://doi.org/10.1098/rsta.2021.0104

Angle, J. C., Morin, T. H., Solden, L. M., Narrowe, A. B., Smith, G. J., Borton, M. A., et al. (2017). Methanogenesis in oxygenated soils is a substantial fraction of wetland methane emissions. *Nature Communications*, *8*(1), 1567. https://doi.org/10.1038/s41467-017-01753-4

Apers, S., De Lannoy, G. J. M., Baird, A. J., Cobb, A. R., Dargie, G. C., Pasquel, J., et al. (2022). Tropical peatland hydrology simulated with a global land surface model. *Journal of Advances in Modeling Earth Systems*, *14*(3), e2021MS002784. https://doi.org/10.1029/2021ms002784

Baldocchi, D. (2014). Measuring fluxes of trace gases and energy between ecosystems and the atmosphere—The state and future of the eddy covariance method. *Global Change Biology*, *20*(12), 3600–3609. https://doi.org/10.1111/gcb.12649

Bansal, S., Post van der Burg, M., Fern, R. R., Jones, J. W., Lo, R., McKenna, O. P., et al. (2023). Large increases in methane emissions expected from North America's largest wetland complex. *Science Advances*, *9*(9), eade1112. https://doi.org/10.1126/sciadv.ade1112

Billesbach, D., & Sullivan, R. (2020a). FLUXNET-CH4 US-A03 ARM-AMF3-Oliktok [Dataset]. FluxNet; Argonne National Laboratory. https://doi.org/10.18140/FLX/1669661

Billesbach, D., & Sullivan, R. (2020b). FLUXNET-CH4 US-A10 ARM-NSA-Barrow [Dataset]. FluxNet; Argonne National Laboratory. https://doi.org/10.18140/FLX/1669662

Bloom, A., Bowman, W. K., Lee, M., Turner, J. A., Schroeder, R., Worden, R. J., et al. (2017). A global wetland methane emissions and uncertainty dataset for atmospheric chemical transport models (WetCHARTs version 1.0). *Geoscientific Model Development*, *10*(6), 2141–2156. https://doi.org/10.5194/gmd-10-2141-2017

Bodesheim, P., Jung, M., Gans, F., Mahecha, M. D., & Reichstein, M. (2018). Upscaled diurnal cycles of land--atmosphere fluxes: A new global half-hourly data product. *Earth System Science Data*, *10*(3), 1327–1365. https://doi.org/10.5194/essd-10-1327-2018

Bohrer, G., Kerns, J., Morin, T., Rey-Sanchez, A., Villa, J., & Ju, Y. (2020). FLUXNET-CH4 US-OWC old woman creek [Dataset]. FluxNet; Old Woman Creek National Estuarine Research Reserve; The Ohio State University. https://doi.org/10.18140/FLX/1669690

Bohrer, G., & Morin, T. (2020). FLUXNET-CH4 US-ORv Olentangy River wetland research park [Dataset]. FluxNet; The Ohio State University. https://doi.org/10.18140/FLX/1669689

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Bridgham, S. D., Cadillo-Quiroz, H., Keller, J. K., & Zhuang, Q. (2013). Methane emissions from wetlands: Biogeochemical, microbial, and modeling perspectives from local to global scales. *Global Change Biology*, *19*(5), 1325–1346. https://doi.org/10.1111/gcb.12131

Bruhwiler, L., Dlugokencky, E., Masarie, K., Ishizawa, M., Andrews, A., Miller, J., et al. (2014). CarbonTracker-CH_4: An assimilation system for estimating emissions of atmospheric methane. *Atmospheric Chemistry and Physics*, *14*(16), 8269–8293. https://doi.org/10.5194/acp-14-8269-2014

Campbell, D., & Goodrich, J. (2020). FLUXNET-CH4 NZ-Kop Kopuatai [Dataset]. FluxNet; University of Waikato. https://doi.org/10.18140/FLX/1669652

Chang, K.-Y., Riley, W. J., Knox, S. H., Jackson, R. B., McNicol, G., Poulter, B., et al. (2021). Substantial hysteresis in emergent temperature sensitivity of global wetland CH4 emissions. *Nature Communications*, *12*(1), 2266. https://doi.org/10.1038/s41467-021-22452-1

Chen, J., & Chu, H. (2020). FLUXNET-CH4 US-WPT winous point North Marsh [Dataset]. FluxNet; University of Toledo / Michigan State University. https://doi.org/10.18140/FLX/1669702

Chen, X., Bohn, T. J., & Lettenmaier, D. P. (2015). Model estimates of climate controls on pan-Arctic wetland methane emissions. *Biogeosciences*, *12*(21), 6259–6277. https://doi.org/10.5194/bg-12-6259-2015

Chu, H., Baldocchi, D. D., John, R., Wolf, S., & Reichstein, M. (2017). Fluxes all of the time? A primer on the temporal representativeness of FLUXNET. *Journal of Geophysical Research: Biogeosciences*, *122*(2), 289–307. https://doi.org/10.1002/2016JG003576

Chu, H., Luo, X., Ouyang, Z., Chan, W. S., Dengel, S., Biraud, S. C., et al. (2021). Representativeness of Eddy-Covariance flux footprints for areas surrounding AmeriFlux sites. *Agricultural and Forest Meteorology*, *301–302*, 108350. https://doi.org/10.1016/j.agrformet.2021.108350

Dalmagro, H. J., Lathuillière, M. J., Hawthorne, I., Morais, D. D., Pinto, O. B., Jr., Couto, E. G., & Johnson, M. S. (2018). Carbon biogeochemistry of a flooded Pantanal forest over three annual flood cycles. *Biogeochemistry*, *139*(1), 1–18. https://doi.org/10.1007/s10533-018-0450-1

Delwiche, K. B., Knox, S. H., Malhotra, A., Fluet-Chouinard, E., McNicol, G., Feron, S., et al. (2021). FLUXNET-CH4: A global, multi-ecosystem dataset and analysis of methane seasonality from freshwater wetlands. *Earth System Science Data*, *13*(7), 3607–3689. https://doi.org/10.5194/essd-13-3607-2021

Dengel, S., Zona, D., Sachs, T., Aurela, M., Jammet, M., Parmentier, F. J. W., et al. (2013). Testing the applicability of neural networks as a gap-filling method using CH_4 flux data from high latitude wetlands. *Biogeosciences Discussions*, *10*(12), 8185–8200. https://doi.org/10.5194/bg-10-8185-2013

Desai, A. (2020). FLUXNET-CH4 US-Los lost creek [Dataset]. FluxNet; University of Wisconsin. https://doi.org/10.18140/FLX/1669682

Devol, A. H., Richey, J. E., Clark, W. A., King, S. L., & Martinelli, L. A. (1988). Methane emissions to the troposphere from the Amazon floodplain. *Journal of Geophysical Research, D: Atmospheres*, *93*(D2), 1583–1592. https://doi.org/10.1029/JD093iD02p01583

Dolman, H., Maximov, T., Parmentier, F., & Budishev, A. (2020). FLUXNET-CH4 RU-Cok Chokurdakh [Dataset]. FluxNet; Vrije Universiteit Amsterdam. https://doi.org/10.18140/FLX/1669656

Eichelmann, E., Knox, S., Sanchez, C., Valach, A., Sturtevant, C., Szutu, D., et al. (2020). FLUXNET-CH4 US-Tw4 Twitchell East end wetland [Dataset]. FluxNet; University of California. https://doi.org/10.18140/FLX/1669698

Euskirchen, E., Bret-Harte, M., & Edgar, C. (2020). FLUXNET-CH4 US-ICs imnavait creek watershed wet sedge tundra [Dataset]. FluxNet; Marine Biological Laboratory; University of Alaska Fairbanks. https://doi.org/10.18140/FLX/1669678

Euskirchen, E., & Edgar, C. (2020a). FLUXNET-CH4 US-BZB bonanza creek thermokarst bog [Dataset]. FluxNet. https://doi.org/10.18140/FLX/1669668

Euskirchen, E., & Edgar, C. (2020b). FLUXNET-CH4 US-BZF bonanza creek rich fen [Dataset]. FluxNet; University of Alaska Fairbanks, Institute of Arctic Biology. https://doi.org/10.18140/FLX/1669669

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Gao, B.-C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, *58*(3), 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3

Gauci, V., Figueiredo, V., Gedney, N., Pangala, S. R., Stauffer, T., Weedon, G. P., & Enrich-Prast, A. (2022). Non-flooded riparian Amazon trees are a regionally significant methane source. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *380*(2215), 20200446. https://doi.org/10.1098/rsta.2020.0446

Goeckede, M. (2020). FLUXNET-CH4 RU-Ch2 Chersky reference [Dataset]. FluxNet; Max Planck Institute for Biogeochemistry. https://doi.org/10.18140/FLX/1669654

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth engine: Planetary-scale geospatial analysis for everyone.

Gregorutti, B., Michel, B., & Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Statistics and Computing*, *27*(3), 659–678. https://doi.org/10.1007/s11222-016-9646-1

Hall, D. K., & Riggs, G. A. (2016). MODIS/Terra snow cover daily L3 global 500m SIN grid, version 6 [Dataset]. National Snow and Ice Data Center. https://doi.org/10.5067/MODIS/MOD10A1.006

Hargrove, W. W., & Hoffman, F. M. (2004). Potential of multivariate quantitative methods for delineation and visualization of ecoregions. *Environmental Management*, *34*(1), S39–S60. https://doi.org/10.1007/s00267-003-1084-0

Harriss, R. C., & Matson, P. A. (2009). *Biogenic trace gases: Measuring emissions from soil and water*. John Wiley & Sons.

Helfter, C. (2020a). FLUXNET-CH4 BW-Gum Guma [Dataset]. FluxNet; UK Centre for Ecology and Hydrology. https://doi.org/10.18140/FLX/1669370

Helfter, C. (2020b). FLUXNET-CH4 BW-Nxr Nxaraga [Dataset]. FluxNet; UK Centre for Ecology and Hydrology. https://doi.org/10.18140/FLX/1669518

Hemes, K. S., Chamberlain, S. D., Eichelmann, E., Anthony, T., Valach, A., Kasak, K., et al. (2019). Assessing the carbon and climate benefit of restoring degraded agricultural peat soils to managed wetlands. *Agricultural and Forest Meteorology*, *268*(January), 202–214. https://doi.org/10.1016/j.agrformet.2019.01.017

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., et al. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PloS One*, *12*(2), e0169748. https://doi.org/10.1371/journal.pone.0169748

Hinkle, C., & Bracho, R. (2020). FLUXNET-CH4 US-DPW Disney wilderness preserve wetland [Dataset]. FluxNet; University of Central Florida; University of Central Florida (UCF). https://doi.org/10.18140/FLX/1669672

Hoffman, F. M., Kumar, J., Mills, R. T., & Hargrove, W. W. (2013). Representativeness-based sampling network design for the State of Alaska. *Landscape Ecology*, *28*(8), 1567–1586. https://doi.org/10.1007/s10980-013-9902-0

Holm, G., Perez, B., McWhorter, D., Krauss, K., Raynie, R., & Killebrew, C. (2020). FLUXNET-CH4 US-LA2 Salvador WMA freshwater marsh [Dataset]. FluxNet; US Geological Survey; USGS-Wetland and Aquatic Research Center. https://doi.org/10.18140/FLX/1669681

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*(1), 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2

IPCC. (2021). Summary for policymakers. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, et al. (Eds.), *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change* (pp. 3–32). Cambridge University Press. https://doi.org/10.1017/9781009157896.001

Irvin, J., Zhou, S., McNicol, G., Lu, F., Liu, V., Fluet-Chouinard, E., et al. (2021). Gap-filling eddy covariance methane fluxes: Comparison of machine learning model predictions and uncertainties at FLUXNET-CH4 wetlands. *Agricultural and Forest Meteorology*, *308–309*, 108528. https://doi.org/10.1016/j.agrformet.2021.108528

Iwata, H., Hirata, R., Takahashi, Y., Miyabara, Y., Itoh, M., & Iizuka, K. (2018). Partitioning eddy-covariance methane fluxes from a shallow lake into diffusive and ebullitive fluxes. *Bound. -Layer Meteorol.*, *169*(3), 413–428. https://doi.org/10.1007/s10546-018-0383-1

Iwata, H., Ueyama, M., & Harazono, Y. (2020). FLUXNET-CH4 US-Uaf University of Alaska, Fairbanks [Dataset]. FluxNet; Osaka Prefecture University; Shinshu University. https://doi.org/10.18140/FLX/1669701

Jacob, D. J., Varon, D. J., Cusworth, D. H., Dennison, P. E., Frankenberg, C., Gautam, R., et al. (2022). Quantifying methane emissions from the global scale down to point sources using satellite observations of atmospheric methane. *Atmospheric Chemistry and Physics*, *22*(14), 9617–9646. https://doi.org/10.5194/acp-22-9617-2022

Jacotot, A., Gogo, S., & Laggoun-Défarge, F. (2020). FLUXNET-CH4 FR-LGt La Guette [Dataset]. FluxNet; Observatoire des Sciences de l'Univers en région Centre. https://doi.org/10.18140/FLX/1669641

Jensen, K., & Mcdonald, K. (2019). Surface water microwave product series version 3: A near-real time and 25-year historical global inundated area fraction time series from active and passive microwave remote sensing. *IEEE Geoscience and Remote Sensing Letters*, *16*(9), 1402–1406. https://doi.org/10.1109/LGRS.2019.2898779

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., et al. (2020). Scaling carbon fluxes from eddy covariance sites to globe: Synthesis and evaluation of the FLUXCOM approach. *Biogeosciences*, *17*(5), 1343–1365. https://doi.org/10.5194/bg-17-1343-2020

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J. G., Dlugokencky, E. J., et al. (2013). Three decades of global methane sources and sinks. *Nature Geoscience*, *6*(10), 813–823. https://doi.org/10.1038/ngeo1955

Knox, S. H., Bansal, S., McNicol, G., Schafer, K., Sturtevant, C., Ueyama, M., et al. (2021). Identifying dominant environmental predictors of freshwater wetland methane fluxes across diurnal to seasonal time scales. *Global Change Biology*, *27*(15), 3582–3604. https://doi.org/10.1111/gcb.15661

Knox, S. H., Jackson, R. B., Poulter, B., McNicol, G., Fluet-Chouinard, E., Zhang, Z., et al. (2019). FLUXNET-CH4 synthesis activity: Objectives, observations, and future directions. *Bulletin of the American Meteorological Society*, *100*(12), 2607–2632. https://doi.org/10.1175/BAMS-D-18-0268.1

Knox, S. H., Sturtevant, C., Matthes, J. H., Koteen, L., Verfaillie, J., & Baldocchi, D. (2015). Agricultural peatland restoration: Effects of land-use change on greenhouse gas (CO_2 and CH_4) fluxes in the Sacramento-San Joaquin delta. *Global Change Biology*, *21*(2), 750–765. https://doi.org/10.1111/gcb.12745

Koebsch, F., & Jurasinski, G. (2020). FLUXNET-CH4 DE-Hte Huetelmoor [Dataset]. FluxNet; Landscape Ecology, University of Rostock. https://doi.org/10.18140/FLX/1669634

Krauss, K. W., Raynie, R., Killebrew, C., McWhorter, D. E., Holm, G. O., Jr., & Perez, B. C. (2018). Net ecosystem exchange of CO_2 and CH_4 from two Louisiana coastal marshes [Dataset]. U.S. Geological Survey. https://doi.org/10.5066/F7MG7NSV

Kuhn, M. A., Varner, R. K., Bastviken, D., Crill, P., MacIntyre, S., Turetsky, M., et al. (2021). BAWLD-CH 4: A comprehensive dataset of methane fluxes from boreal and arctic ecosystems. *Earth System Science Data*, *13*(11), 5151–5189. https://doi.org/10.5194/essd-13-5151-2021

Lamarque, J.-F., Dentener, F., McConnell, J., Ro, C.-U., Shaw, M., Vet, R., et al. (2013). Multi-model mean nitrogen and sulfur deposition from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): Evaluation of historical and projected future changes [Dataset]. EGU, *13*, 7997–8018. https://doi.org/10.5194/acp-13-7997-2013

Lan, X., Thoning, K. W., & Dlugokencky, E. J. (2023). Trends in globally-averaged CH4, N_2O, and SF6 determined from NOAA global monitoring laboratory measurements. Version 2023-08. https://doi.org/10.15138/P8XG-AA10

Lohila, A., Aurela, M., Tuovinen, J.-P., Laurila, T., Hatakka, J., Rainne, J., & Mäkelä, T. (2020). FLUXNET-CH4 FI-Lom Lompolojankka [Dataset]. FluxNet; Finnish Meteorological Institute. https://doi.org/10.18140/FLX/1669638

Ma, S., Worden, J. R., Bloom, A. A., Zhang, Y., Poulter, B., Cusworth, D. H., et al. (2021). Satellite constraints on the latitudinal distribution and temperature sensitivity of wetland methane emissions. *AGU Advances*, *2*(3). https://doi.org/10.1029/2021av000408

Malone, S. L., Oh, Y., Arndt, K. A., Burba, G., Commane, R., Contosta, A. R., et al. (2022). Gaps in network infrastructure limit our understanding of biogenic methane emissions for the United States. *Biogeosciences*, *19*(9), 2507–2522. https://doi.org/10.5194/bg-19-2507-2022

Matthes, J., Sturtevant, C., Oikawa, P., Chamberlain, S., Szutu, D., Ortiz, A., et al. (2020). FLUXNET-CH4 US-Myb Mayberry wetland [Dataset]. FluxNet; University of California. https://doi.org/10.18140/FLX/1669685

Melton, J. R., Wania, R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., et al. (2013). Present state of global wetland extent and wetland methane modelling: Conclusions from a model inter-comparison project (WETCHIMP). *Biogeosciences*, *10*(2), 753–788. https://doi.org/10.5194/bg-10-753-2013

Meyer, D., & Buchta, C. (2020). Distance and similarity measures [R package proxy version 0.4-24]. Retrieved from https://CRAN.R-project.org/package=proxy

Meyer, H., & Pebesma, E. (2022). Machine learning-based global maps of ecological variables and the challenge of assessing them. *Nature Communications*, *13*(1), 2208. https://doi.org/10.1038/s41467-022-29838-9

Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., & Nauss, T. (2018). Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environmental Modelling & Software*, *101*, 1–9. https://doi.org/10.1016/j.envsoft.2017.12.001

Meyer, H., Reudenbach, C., Wöllauer, S., & Nauss, T. (2019). Importance of spatial predictor variable selection in machine learning applications -- Moving from data reproduction to spatial prediction. *Ecological Modelling*, *411*, 108815. https://doi.org/10.1016/j.ecolmodel.2019.108815

Mitra, B., Minick, K., Miao, G., Domec, J.-C., Prajapati, P., McNulty, S. G., et al. (2020). Spectral evidence for substrate availability rather than environmental control of methane emissions from a coastal forested wetland. *Agricultural and Forest Meteorology*, *291*, 108062. https://doi.org/10.1016/j.agrformet.2020.108062

Myneni, R., Knyazikhin, Y., & Park, T. (2015). MCD15A2H MODIS/Terra+Aqua leaf area index/FPAR 8-day L4 global 500m SIN grid V006 [Dataset]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MCD15A2H.006

Nilsson, M., & Peichl, M. (2020). FLUXNET-CH4 SE-Deg Degero [Dataset]. FluxNet; Department of Forest Ecology and Management; Swedish University of Agricultural Sciences. https://doi.org/10.18140/FLX/1669659

Nisbet, E. G., Jones, A. E., Pyle, J. A., & Skiba, U. (2022). Rising methane: Is there a methane emergency? *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *380*(2215), 20210334. https://doi.org/10.1098/rsta.2021.0334

Noormets, A., King, J., Mitra, B., Miao, G., Aguilos, M., Minick, K., et al. (2020). FLUXNET-CH4 US-NC4 NC_AlligatorRiver [Dataset]. FluxNet; Texas A&M University. https://doi.org/10.18140/FLX/1669686

Nzotungicimpaye, C.-M., Zickfeld, K., MacDougall, A. H., Melton, J. R., Treat, C. C., Eby, M., & Lesack, L. F. W. (2021). WETMETH 1.0: A new wetland methane model for implementation in Earth system models. *Geoscientific Model Development*, *14*(10), 6215–6240. https://doi.org/10.5194/gmd-14-6215-2021

Olson, B. (2018). AmeriFlux AmeriFlux US-ALQ Allequash creek site [Dataset]. AmeriFlux; USGS. https://doi.org/10.17190/AMF/1480323

Pangala, S. R., Enrich-Prast, A., Basso, L. S., Peixoto, R. B., Bastviken, D., Hornibrook, E. R. C., et al. (2017). Large emissions from floodplain trees close the Amazon methane budget. *Nature*, *552*(7684), 230–234. https://doi.org/10.1038/nature24639

Parker, R. J., Boesch, H., McNorton, J., Comyn-Platt, E., Gloor, M., Wilson, C., et al. (2018). Evaluating year-to-year anomalies in tropical wetland methane emissions using satellite CH4 observations. *Remote Sensing of Environment*, *211*, 261–275. https://doi.org/10.1016/j.rse.2018.02.011

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., et al. (2020). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific Data*, *7*(1), 225. https://doi.org/10.1038/s41597-020-0534-3

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422. https://doi.org/10.1038/nature20584

Peltola, O., Vesala, T., Gao, Y., Räty, O., Alekseychik, P., Aurela, M., et al. (2019). Monthly gridded data product of northern wetland methane emissions based on upscaling eddy covariance observations. *Earth System Science Data*, *11*(3), 1263–1289. https://doi.org/10.5194/essd-11-1263-2019

Poulter, B., Bousquet, P., Canadell, J. G., Ciais, P., Peregon, A., Saunois, M., et al. (2017). Global wetland contribution to 2000–2012 atmospheric Methane growth rate dynamics. *Environmental Research Letters: ERL [Web Site]*, *12*(9), 094013. https://doi.org/10.1088/1748-9326/aa8391

Prigent, C., Jimenez, C., & Bousquet, P. (2020). Satellite-derived global surface water extent and dynamics over the last 25 years (GIEMS-2). *Journal of Geophysical Research*, *125*(3). https://doi.org/10.1029/2019jd030711

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

Rey-Sanchez, A. C., Morin, T. H., Stefanik, K. C., Wrighton, K., & Bohrer, G. (2018). Determining total emissions and environmental drivers of methane flux in a Lake Erie estuarine marsh. *Ecological Engineering*, *114*, 7–15. https://doi.org/10.1016/j.ecoleng.2017.06.042

Riley, W. J., Subin, Z. M., Lawrence, D. M., Swenson, S. C., Torn, M. S., Meng, L., et al. (2011). Barriers to predicting changes in global terrestrial methane fluxes: Analyses using CLM4Me, a methane biogeochemistry model integrated in CESM. *Biogeosciences*, *8*(7), 1925–1953. https://doi.org/10.5194/bg-8-1925-2011

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913–929. https://doi.org/10.1111/ecog.02881

Roman, T., Griffis, T., Kolka, R., Wayson, C., Lilleskov, E., Torres, D., et al. (2020). AmeriFlux AmeriFlux PE-QFR Quistococha forest reserve [Dataset]. AmeriFlux; University of Minnesota; USDA-Forest Service-International Programs. https://doi.org/10.17190/AMF/1671889

Roman, T., Kolka, R., Griffis, T., & Deventer, J. (2021). AmeriFlux AmeriFlux US-MBP Marcell bog lake peatland [Dataset]. AmeriFlux; University of Minnesota; USDA- Forest Service. https://doi.org/10.17190/AMF/1767835

Sachs, T., & Wille, C. (2020). FLUXNET-CH4 DE-Zrk Zarnekow [Dataset]. FluxNet; GFZ German Research Centre for Geosciences. https://doi.org/10.18140/FLX/1669636

Sakabe, A., Itoh, M., Hirano, T., & Kusin, K. (2020). FLUXNET-CH4 ID-Pag Palangkaraya undrained forest [Dataset]. FluxNet; Hokkaido University; Kyoto University; University of Hyogo; University of Palangkaraya. https://doi.org/10.18140/FLX/1669643
Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., et al. (2020). The global methane budget 2000--2017. *Earth Syst. Sci. Data*, *12*(3), 1561–1623. https://doi.org/10.5194/essd-12-1561-2020
Schmid, H., & Klatt, J. (2020). FLUXNET-CH4 DE-SfN Schechenfilz Nord [Dataset]. FluxNet; Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research (IMK-IFU). https://doi.org/10.18140/FLX/1669635
Shortt, R., Hemes, K., Szutu, D., Verfaillie, J., & Baldocchi, D. (2020). FLUXNET-CH4 US-Sne Sherman Island restored wetland [Dataset]. FluxNet; University of California. https://doi.org/10.18140/FLX/1669693
Sonnentag, O., & Helbig, M. (2020a). FLUXNET-CH4 CA-SCB Scotty creek bog [Dataset]. FluxNet; Université de Montréal; Wilfrid Laurier University. https://doi.org/10.18140/FLX/1669613
Sonnentag, O., & Helbig, M. (2020b). FLUXNET-CH4 CA-SCC Scotty creek landscape [Dataset]. FluxNet; Université de Montréal; Wilfrid Laurier University. https://doi.org/10.18140/FLX/1669628
Spahni, R., Wania, R., Neef, L., van Weele, M., Pison, I., Bousquet, P., et al. (2011). Constraining global methane emissions and uptake by ecosystems. *Biogeosciences*, *8*(6), 1643–1665. https://doi.org/10.5194/bg-8-1643-2011
Stavert, A. R., Saunois, M., Canadell, J. G., Poulter, B., Jackson, R. B., Regnier, P., et al. (2021). Regional trends and drivers of the global methane budget. *Global Change Biology*, *28*(1), 182–200. https://doi.org/10.1111/gcb.15901
Stell, E., Warner, D., Jian, J., Bond-Lamberty, B., & Vargas, R. (2021). Spatial biases of information influence global estimates of soil respiration: How can we improve global predictions? *Global Change Biology*, *27*(16), 3923–3938. https://doi.org/10.1111/gcb.15666
Sturtevant, C., Ruddell, B. L., Knox, S. H., Verfaillie, J., Matthes, J. H., Oikawa, P. Y., & Baldocchi, D. (2016). Identifying scale-emergent, nonlinear, asynchronous processes of wetland methane exchange. *Journal of Geophysical Research: Biogeosciences*, *121*(1), 188–204. https://doi.org/10.1002/2015jg003054
Tenuta, M. (2020). AmeriFlux AmeriFlux CA-CF2 Churchill fen site 2 [Dataset]. AmeriFlux; University of Manitoba. https://doi.org/10.17190/AMF/1634879
Torn, M., & Dengel, S. (2020a). FLUXNET-CH4 US-NGB NGEE Arctic Barrow [Dataset]. FluxNet; Lawrence Berkeley National Laboratory. https://doi.org/10.18140/FLX/1669687
Torn, M., & Dengel, S. (2020b). FLUXNET-CH4 US-NGC NGEE Arctic Council [Dataset]. FluxNet; Berkeley Lab; Lawrence Berkeley National Laboratory. https://doi.org/10.18140/FLX/1669688
Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., et al. (2016). Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. *Biogeosciences*, *13*(14), 4291–4313. https://doi.org/10.5194/bg-13-4291-2016
Treat, C. C., Bloom, A. A., & Marushchak, M. E. (2018). Nongrowing season methane emissions-a significant component of annual emissions across northern ecosystems. *Global Change Biology*, *24*(8), 3331–3343. https://doi.org/10.1111/gcb.14137
Tuanmu, M.-N., & Jetz, W. (2014). A global 1-km consensus land-cover product for biodiversity and ecosystem modelling: Consensus land cover. *Global Ecology and Biogeography: A Journal of Macroecology*, *23*(9), 1031–1045. https://doi.org/10.1111/geb.12182
Tuanmu, M.-N., & Jetz, W. (2015). A global, remote sensing-based characterization of terrestrial habitat heterogeneity for biodiversity and ecosystem modelling: Global habitat heterogeneity. *Global Ecology and Biogeography: A Journal of Macroecology*, *24*(11), 1329–1339. https://doi.org/10.1111/geb.12365
Turetsky, M. R., Kotowska, A., Bubier, J., Dise, N. B., Crill, P., Hornibrook, E. R. C., et al. (2014). A synthesis of methane emissions from 71 northern, temperate, and subtropical wetlands. *Global Change Biology*, *20*(7), 2183–2197. https://doi.org/10.1111/gcb.12580
Turner, A. J., Frankenberg, C., & Kort, E. A. (2019). Interpreting contemporary trends in atmospheric methane. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(8), 2805–2813. https://doi.org/10.1073/pnas.1814297116
Ueyama, M., Hirano, T., & Kominami, Y. (2020). FLUXNET-CH4 JP-BBY Bibai bog [Dataset]. https://doi.org/10.18140/FLX/1669646. FluxNet; Osaka Prefecture Univeristy
Ueyama, M., Knox, S. H., Delwiche, K. B., Bansal, S., Riley, W. J., Baldocchi, D., et al. (2023). Modeled production, oxidation, and transport processes of wetland methane emissions in temperate, boreal, and Arctic regions. *Global Change Biology*, *29*(8), 2313–2334. https://doi.org/10.1111/gcb.16594
Ueyama, M., Yazaki, T., Hirano, T., & Endo, R. (2022). Partitioning methane flux by the eddy covariance method in a cool temperate bog based on a Bayesian framework. *Agricultural and Forest Meteorology*, *316*, 108852. https://doi.org/10.1016/j.agrformet.2022.108852
Valach, A., Kasak, K., Szutu, D., Verfaillie, J., & Baldocchi, D. (2020). FLUXNET-CH4 US-Tw5 East Pond wetland [Dataset]. FluxNet; University of California. https://doi.org/10.18140/FLX/1669699
Valach, A., Szutu, D., Eichelmann, E., Knox, S., Verfaillie, J., & Baldocchi, D. (2020). FLUXNET-CH4 US-Tw1 Twitchell wetland West Pond [Dataset]. FluxNet; University of California. https://doi.org/10.18140/FLX/1669696
Vermote, E. (2015). MOD09A1 MODIS/Terra surface reflectance 8-day L3 global 500m SIN grid V006 [Dataset]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MOD09A1.006
Vesala, T., Tuittila, E.-S., Mammarella, I., & Alekseychik, P. (2020). FLUXNET-CH4 FI-Si2 Siikaneva-2 bog [Dataset]. FluxNet; University of Eastern Finland; University of Helsinki. https://doi.org/10.18140/FLX/1669639
Vesala, T., Tuittila, E.-S., Mammarella, I., & Rinne, J. (2020). FLUXNET-CH4 FI-Sii Siikaneva [Dataset]. FluxNet; University of Eastern Finland; University of Helsinki. https://doi.org/10.18140/FLX/1669640
Villarreal, S., & Vargas, R. (2021). Representativeness of FLUXNET sites across Latin America. *Journal of Geophysical Research: Biogeosciences*, *126*(3), e2020JG006090. https://doi.org/10.1029/2020jg006090
Vourlitis, G., Dalmagro, H., de Nogueira, J., Johnson, M., & Arruda, P. (2020). FLUXNET-CH4 BR-Npw northern Pantanal wetland [Dataset]. FluxNet; California State University, San Marcos; Universidade de Cuiabá; Universidade Federal de Mato Grosso; University of British Columbia. https://doi.org/10.18140/FLX/1669368
Wan, Z., Hook, S., & Hulley, G. (2015). MOD11A2 MODIS/Terra land surface temperature/Emissivity 8-day L3 global 1km SIN grid V006 [Dataset]. USGS. https://doi.org/10.5067/MODIS/MOD11A2.006
Whiting, G. J., & Chanton, J. P. (1993). Primary production control of methane emission from wetlands. *Nature*, *364*(6440), 794–795. https://doi.org/10.1038/364794a0
Wong, G., Melling, L., Tang, A., Aeries, E., Waili, J., Musin, K., et al. (2020). FLUXNET-CH4 MY-MLM Maludam national park [Dataset]. FluxNet; Sarawak Tropical Peat Research Institute. https://doi.org/10.18140/FLX/1669650
Xu, K., Metzger, S., & Desai, A. R. (2017). Upscaling tower-observed turbulent exchange at fine spatio-temporal resolution using environmental response functions. *Agricultural and Forest Meteorology*, *232*, 10–22. https://doi.org/10.1016/j.agrformet.2016.07.019
Zhang, Y., Li, C., Trettin, C. C., Li, H., & Sun, G. (2002). An integrated model of soil, hydrology, and vegetation for carbon dynamics in wetland ecosystems. *Global Biogeochemical Cycles*, *16*(4), 9-1–9-17. https://doi.org/10.1029/2001gb001838

Zhang, Z., Fluet-Chouinard, E., Jensen, K., McDonald, K., Hugelius, G., Gumbricht, T., et al. (2021). Development of the global dataset of wetland area and dynamics for methane modeling (WAD2M). *Earth System Science Data*, *13*(5), 2001–2023. https://doi.org/10.5194/essd-13-2001-2021

Zhang, Z., Poulter, B., Feldman, A. F., Ying, Q., Ciais, P., Peng, S., & Li, X. (2023). Recent intensification of wetland methane feedback. *Nature Climate Change*, *13*(5), 430–433. https://doi.org/10.1038/s41558-023-01629-0

Zhang, Z., Zimmermann, N. E., Stenke, A., Li, X., Hodson, E. L., Zhu, G., et al. (2017). Emerging role of wetland methane emissions in driving 21st century climate change. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(36), 9647–9652. https://doi.org/10.1073/pnas.1618765114

Zona, D., & Oechel, W. (2020a). FLUXNET-CH4 US-Atq Atqasuk [Dataset]. FluxNet; San Diego State University. https://doi.org/10.18140/FLX/1669663

Zona, D., & Oechel, W. (2020b). FLUXNET-CH4 US-Beo Barrow environmental observatory (BEO) tower [Dataset]. FluxNet; San Diego State University. https://doi.org/10.18140/FLX/1669664

Zona, D., & Oechel, W. (2020c). FLUXNET-CH4 US-Bes Barrow-Bes (Biocomplexity Experiment South tower) [Dataset]. FluxNet; San Diego State University. https://doi.org/10.18140/FLX/1669665

Zona, D., & Oechel, W. (2020d). FLUXNET-CH4 US-Ivo Ivotuk [Dataset]. FluxNet; San Diego State University. https://doi.org/10.18140/FLX/1669679

References From the Supporting Information

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015. *Scientific Data*, *5*(1), 170191. https://doi.org/10.1038/sdata.2017.191

Amatulli, G., Domisch, S., Tuanmu, M.-N., Parmentier, B., Ranipeta, A., Malczyk, J., & Jetz, W. (2018). A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. *Scientific Data*, *5*(1), 180040. https://doi.org/10.1038/sdata.2018.40

Amatulli, G., McInerney, D., Sethi, T., Strobl, P., & Domisch, S. (2019). Geomorpho90m - global high-resolution geomorphometry layers: Empirical evaluation and accuracy assessment. *PeerJ Preprints*. https://doi.org/10.7287/peerj.preprints.27595

Behrens, T., Schmidt, K., Viscarra Rossel, R. A., Gries, P., Scholten, T., & MacMillan, R. A. (2018). Spatial modelling with Euclidean distance fields and machine learning. *European Journal of Soil Science*, *69*(5), 757–770. https://doi.org/10.1111/ejss.12687

Kuhn, M. (2020). caret: Classification and regression training. Retrieved from https://CRAN.R-project.org/package=caret

Morin, T. H. (2019). Advances in the eddy covariance approach to CH_4 monitoring over two and a half decades. *Journal of Geophysical Research: Biogeosciences*, *124*(3), 453–460. https://doi.org/10.1029/2018JG004796

R Core Team. (2022). R: A language and environment for statistical computing.

Schroeder, R., McDonald, K. C., Chapman, B. D., Jensen, K., Podest, E., Tessler, Z. D., et al. (2015). Development and evaluation of a multi-year fractional surface water data set derived from active/passive microwave remote sensing data. *Remote Sensing*, *7*(12), 16688–16732. https://doi.org/10.3390/rs71215843

Sexton, J. O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., et al. (2013). Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. *International Journal of Digital Earth*, *6*(5), 427–448. https://doi.org/10.1080/17538947.2013.786146

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research*, *116*(4), 1–12. https://doi.org/10.1029/2011JG001708

Wright, M. N., & Ziegler, A. (2017). ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, *77*(1). https://doi.org/10.18637/jss.v077.i01

Xiao, X., Boles, S., Frolking, S., Salas, W., Moore, B., Li, C., et al. (2002). Observation of flooding and rice transplanting of paddy rice fields at the site to landscape scales in China using VEGETATION sensor data. *International Journal of Remote Sensing*, *23*(15), 3009–3022. https://doi.org/10.1080/01431160110107734

Zarco-Tejada, P. J., & Ustin, S. L. (2001). Modeling canopy water content for carbon estimates from MODIS data at land EOS validation sites. In *IGARSS 2001. Scanning the present and resolving the future. Proceedings. IEEE 2001 international geoscience and remote sensing symposium (Cat. No.01CH37217)* (Vol. 1, pp. 342–344). https://doi.org/10.1109/IGARSS.2001.976152