https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

Gaps in Network Infrastructure limit our understanding of biogenic methane emissions for the United States

Sparkle L. Malone[1], Youmi Oh[2], Kyle A. Arndt[3], George Burba[4,5], Roisin Commane[6], Alexandra R. Contosta[7], Jordan P. Goodrich[8], Henry W. Loescher[9,10], Gregory Starr[11], Ruth K. Varner[3,12]

[1]Institute of the Environment & Department of Biological Sciences, Florida International University, 11200 S.W. 8th Street, Miami, FL 33199, USA.
[2]NOAA Global Monitoring Laboratory, Boulder, CO, 80305, USA
[3]Earth Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, 8 College Rd, Durham, NH, 03824, USA
[4]LI-COR Biosciences, 4421 Superior St., Lincoln, NE, 68504, USA
[5]The Robert B. Daugherty Water for Food Global Institute and School of Natural Resources, University of Nebraska, Lincoln, NE, 68583
[6]Department of Earth & Environmental Sciences, Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964, USA
[7]Earth Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, 8 College Rd, Durham, NH, 03824, USA
[8]School of Science, University of Waikato, Gate 1 Knighton Rd, Hillcrest 3240, Hamilton, New Zealand
[9] Battelle, National Ecological Observatory Network (NEON), Boulder, CO USA 80301
[10] Institute of Alpine and Arctic Research, University of Colorado, Boulder, CO USA 80301
[11]Department of Biological Sciences, University of Alabama, Tuscaloosa, AL. 35487, USA
[12]Department of Earth Sciences, University of New Hampshire, 56 College Rd, Durham, NH, 03824, USA

Correspondence to: Sparkle L. Malone (smalone@fiu.edu)

Abstract. Understanding the sources and sinks of CH_4 is critical to both predicting and mitigating future climate change. There are large uncertainties in the global budget of atmospheric CH_4, but natural emissions are estimated to be of a similar magnitude to total anthropogenic emissions. The largest sources of uncertainty in scaling bottom-up CH_4 estimates stem from limited ground-based measurements and the misalignment between drivers of CH_4 fluxes and current land use classifications. To understand the CH_4 flux potential of natural ecosystems and agricultural lands in the United States (US) of America, a multi-scale CH_4 observation network focused on CH_4 flux rates, processes, and scaling methods is required. This can be achieved with a network of ground-based observations that are distributed based on climatic regions and landcover. To determine the gaps in physical infrastructure for developing this network, we need to understand the representativeness of current measurements. We focus here on eddy covariance (EC) flux towers because they are essential for a bottom-up framework that bridges the gap between point-based chamber measurements and airborne or satellite platforms, informing the remote sensing and modelling communities and policy decisions, all the way to IPCC reports. Using multidimensional scaling and a cluster analysis, the US was divided into 10 clusters that were distributed across temperature and wetness gradients. We evaluated the distance to the medoid condition within each cluster for research sites with EC tower infrastructure to identify the gaps in existing infrastructure that limit our ability to constrain the contribution of US biogenic CH_4 emissions to the global budget. These gaps occurred across all EC flux tower networks and independently managed sites as well as in some environmental clusters. Through our analysis using climate, land

cover, and location variables, we have identified priority areas to target for research infrastructure to provide a more complete understanding of the CH_4 flux potential of ecosystem types across the US.

1 Introduction

The 21st century is characterized by ongoing changes in Earth's climate system that result from increasing concentrations of radiatively important trace gases in the atmosphere. Unlike the relatively steady increases of atmospheric CO_2 and N_2O, atmospheric CH_4 concentrations show dynamic trends with a rapid increase of ~10 ppb yr^{-1} since 2014 (Nisbet et al., 2019a). The annual increase of atmospheric CH_4 in 2020 was the largest on record at ~15 ppb yr^{-1} (Dlugokencky, 2021), despite the global pandemic reducing energy demand (Le Quéré et al., 2021). Increasing atmospheric CH_4 concentrations (Nisbet et al., 2019a) is of concern because CH_4 is 34 times more effective at trapping heat in the atmosphere compared to an equivalent mass of CO_2 over a 100-year timeframe, and accounts for ~42% of warming since the pre-industrial period (IPCC, 2021). These rapid increases in atmospheric CH_4 challenge us to reach the goals of the Paris Agreement (Nisbet et al., 2019a) but also provide an opportunity given the relatively short atmospheric residence time (~9 years) of CH_4. Understanding the sources and sinks of CH_4 is therefore critical to both predicting and mitigating future climate change.

There are large uncertainties in the sources and sinks of the global budget of atmospheric CH_4 (Saunois et al., 2020; Bruhwiler et al., 2021). Methane is emitted from a variety of often co-located biogenic, thermogenic, and pyrogenic sources (IPCC, 2013; Nisbet et al., 2019b). Biogenic emissions are thought to be of a similar magnitude to total anthropogenic emissions, yet biogenic CH_4 emissions remain the most uncertain source of the global CH_4 budget (Saunois et al., 2020). Surface-atmosphere exchange from biogenic sources and sinks, the biological and environmental processes driving these fluxes (e.g., ebullition, aerenchyma pumping), and how CH_4 sources and sinks change over space and time, including interannual variability (Michalak et al., 2009; Kirschke et al., 2013; Knox et al., 2019; Nisbet et al., 2019b), are not well constrained. Finally, the vast areas with relatively very small uptakes and emissions (e.g., deserts, grasslands, forests, water bodies) and well transport of CH_4 by the lake-ocean water continuum (e.g., fens, stems, and rivers) have been largely understudied could contribute significantly to regional and global budgets (Hutchins et al, 2019; Rosentreter et al, 2021; Zhou et al, 2021). These uncertainties hinder our ability to predict future climate change due to the complex feedbacks between biological processes (e.g., microbial production and consumption) (Sherwood et al., 2017; Zhang et al., 2017; Oh et al., 2020), climate change (Zhang et al., 2017), and landcover change (Kirschke et al., 2013; Knox et al., 2019; Saunois et al., 2020).

By far the largest source of uncertainty in scaling bottom-up CH_4 estimates are in the current land use classification (LUC) products (Kirschke et al., 2013; Knox et al., 2019; Saunois et al., 2020), which are not designed to estimate the potential CH_4 source/sink status, particularly from aquatic, wetland, and agricultural land cover. Aquatic ecosystems contribute significantly to global CH_4 emissions, with emissions increasing from natural to impacted aquatic ecosystems and from coastal to freshwater ecosystems (Rosentreter et al., 2021). Aquatic emissions are likely to change in the future due to

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

urbanization, eutrophication and positive climate feedbacks, yet current wetland classifications for land use data products are not suitable to capture these effects. Wetland classifications are often generalized too broadly in current LUC schemas to accurately scale and predict CH_4 flux rates and processes. Small changes in the delineation or characterization of LUC can result in changing the source/sink status of whole regions (Kirschke et al., 2013; Barkley et al., 2017; Knox et al., 2019). For wetlands these include (i) delineation of wetland area, the largest natural CH_4 source, especially in regions like Alaska and Florida, (ii) conflation of fluxes from wetlands and fresh waters leading to double counting (Thornton et al., 2016), (iii) classification of saturated soils as non-wetland, possibly missing strong CH_4 emission potential. For agricultural lands, we must also consider (iv) deforestation for agricultural use, which reduces the soil CH_4 sink potential (Robertson et al., 2000), or (v) accurate representation of agricultural land CH_4 potential when land use includes a complex mixture of ruminants feedlots, manure and pastures (Lassey, 2008).These large uncertainties in biogenic CH_4 fluxes cannot be addressed with the land cover maps currently used to scale CH_4 fluxes and the existing distribution of CH_4 observation sites and types (Rosentreter et al., 2021).

Large uncertainties in the global CH_4 budget also result from limited ground-based measurements. When scaling bottom-up measurements to landscape and regional scales, measurements tend not to be sufficiently geographically distributed to capture the true spatial variation that is innate to the production and consumption of CH_4, and are compounded by large source/sink strengths in small areas (e.g., periodic wetting/drying of seasonal wetlands, saturated soils) (IPCC, 2013; Knox et al., 2019; Thornton et al., 2016) and by very small source/sink strengths in very large areas. In addition, bottom-up CH_4 process-level estimates have historically been limited to short periods (<1-2 years), are discontinuous (grab sampling), and/or occur only during the growing season at middle and high latitudes (though see Groffman et al., 2006; Arndt et al., 2019 for notable exceptions).

In addition to the current uncertainty in basic ecosystem-level CH_4 processes and the way they spatially scale, the backdrop of climate change is also changing the rates of CH_4 production and consumption, as well as CH_4 transport pathways. For example, arctic regions are warming faster than most other regions of the world (Serreze and Barry, 2011), turning permafrost into wetlands and changing traditional CH_4 sinks to sources on short time scales (Chadburn et al., 2017; Schaefer, 2019; Yumashev et al., 2019). In temperate areas, higher climate-change-induced variability in precipitation (e.g., higher moisture of upland forested soils, prolonged droughts, etc.) results in a reduction of soil CH_4 uptake and a reduced global CH_4 sink (Ni and Groffman, 2018) Sea-level rise, which leads to the inundation of coastal regions turning previously dry upland environments into saturated, anoxic areas, which can in some cases increase CH_4 production and emission rates (Lu et al., 2018).

To understand the CH_4 flux potential of natural ecosystems and agricultural lands in the United States of America (US), a multi-scale CH_4 observation network focused on CH_4 flux rates, processes, and scaling methods is required. This can be achieved with a network of ground-based observations whose distribution is based on climatic region and landcover. Eddy covariance (EC) tower observations of surface-atmosphere fluxes, which provide direct measurements for specific ecosystems year-round, can be strategically placed to reduce uncertainties in the current US CH_4 budget. Automated static chamber

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

measurements within flux tower footprints would allow for measurements of specific sources or sinks within these ecosystems, such as soils or first- and second-order streams. In addition, concurrent measurements of CH_4 concentrations will allow the scientific community to determine fluxes from local (chamber; <1 m^2), ecosystem (EC flux tower; ~1 km^2) and landscape scales (tower concentrations; ~100s km^2). At a larger scale, airborne observations of atmospheric concentrations can be used to calculate surface-atmosphere fluxes and provide greater spatial coverage than towers. However, they cannot resolve fine spatial details (e.g., sources) and are limited to daytime snapshots during fair weather, and thus cannot measure concentrations during high-latitude winter and during other times of limited visibility, such as at night-time (Chang et al., 2014; Zona et al., 2016). Likewise, satellite observations of total column CH_4 are limited to sunlight time periods (missing night and polar winters) without cloud cover, require a complex modelling framework to calculate fluxes from the total column observations, and have large uncertainties from the tropospheric transport of CH_4 (Lu et al., 2021).

The primary goal of this work is to evaluate the representativeness of currently available ground-based research infrastructure to understand where gaps in CH_4 data collection exist and to provide guidance on how the research community could direct their resources to best reduce uncertainties in the US biogenic CH_4 budget. In addressing this goal, we will use a combination of land cover and climate data along with a multidimensional cluster analysis to guide the scientific community on how we can develop a distributed CH_4 observing network for the US and provide a template for the development of similar networks in other regions. Developing a distributed network of CH_4 observations will provide an opportunity for the scientific community to reduce uncertainties of the biogenic CH_4 budget of the US now and into the future as anthropogenic pressures continue to alter the carbon cycle. This network will also provide long-term measurements to potentially track the impact of policy interventions to mitigate CH_4 emissions.

2 Methods

2.1 Overview

To determine the gaps in physical infrastructure for ecosystem-scale CH_4 fluxes, we need to understand what the current tower infrastructure is capturing. We focus here on EC flux towers because they are essential for a bottom-up framework that bridges the gap between point-based chamber measurements and airborne platforms and are therefore a useful basis for identifying gaps in the current network of CH_4 observations. The US AmeriFlux network of EC towers was launched in 1996 and grew from about 15 sites in 1997 to more than 110 active sites registered today. It was originally a network of PI-managed sites measuring ecosystem CO_2, H_2O, and energy fluxes. The network was established to connect research on field sites representing major climatic and ecological biomes, including tundra, grasslands, savanna, crops, and coniferous, deciduous, and tropical forests. The AmeriFlux community tailored instrumentation to suit each unique ecosystem but now also includes towers that are a part of the standardized network, the National Ecological Observatory Network (NEON). AmeriFlux also includes tower sites from the Long-Term Ecological Research Network (LTER). In 2012, the US Department of Energy established the AmeriFlux Management Project (AMP) at Lawrence Berkeley

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

National Laboratory (LBNL) to support the broad AmeriFlux community and the AmeriFlux sites. The AMP standardizes, post-processes, and makes flux data available to the research community. More recently, flux towers began measuring CH_4 (81 towers) in freshwater, coastal, upland, natural, and managed ecosystems. In this evaluation we included EC towers that are a part of AmeriFlux (200), NEON (47), LTER (23), and known, independent PI-managed sites (141). In addition to our analysis of EC sites, we also evaluated the distribution of 161 network aquatic sites across the US. The networks included aquatic sites in AmeriFlux, NEON, LTER and the Global Lake Ecological Observatory Network (GLEON). While all the towers included in the analysis measure CO_2 fluxes, only a handful currently quantify CH_4. To understand the representativeness of the current tower infrastructure, we identified variation in climate and ecosystem type, as these two factors together are characteristic of regional resource availability and disturbance regimes. It is important to note that typically a tower is representative of just the ecosystem type and the region where it is stationed (Desai, 2010; Jung et al., 2011; Xiao et al., 2012; Chu et al., 2021); however, this representativeness analysis uses a coarser classification of ecosystem types that is more emblematic of regional disturbance regimes and resource availability, not the specific ecosystem type where the tower is situated.

2.2 Climate and dominant land cover types

We used the National Land Cover Database (NLCD; www.mrlc.gov) to create a land cover layer for the contiguous US (Jin et al., 2019). The NLCD has a 30-m resolution with a 16-class legend based on a modified Anderson Level II classification system. We reclassified the NLCD into 8 major ecotypes (water, developed, barren, forest, scrub, herbaceous, crop, and wetland). Where the NLCD was not available (Alaska, Hawaii, and Puerto Rico), we used the Moderate Resolution Imaging Spectroradiometer (MODIS; 1 km) Land Cover (type 5 - vegetation functional types) for vegetation functional type (MCD12Q1.006) (Sulla-Menashe and Friedl, 2018), which was also reclassified to the 8 major ecotypes (Table 1). The crop ecotype was expanded to non-irrigated and irrigated classes using agricultural information from the US Department of Agriculture's CropScape and Cropland Data layer (Boryan et al., 2011), and the wetland class was expanded using information from the US Fish and Wildlife Service's National Wetland Inventory. Expanded wetland classes were emergent coastal, emergent freshwater and forest freshwater) (Wilen and Bates, 1995). Climate data were obtained from DAYMET (Thornton et al., 2017). We used five climate variables to characterize the climatic conditions across the US: annual mean daily minimum, daily average, and daily maximum temperature, annual total precipitation, and mean annual daily vapor pressure deficit from 2010-2020. Understanding that these patterns are changing with climate change, we chose a shorter time period than the commonly used 30-year climate normal to better represent current conditions (Bessembinder et al., 2021). All spatial layers (ecotype and climate) were resampled to match the DAYMET climate data (1-km), and all pre-processing was done in R version 4.0.4 (R Core Team, 2021) with the *raster* package (Hijmans, 2021). This approach allowed us to create a land cover layer of the dominant ecotypes at 1 km resolution that was expanded in categories of interest for CH_4. The land cover and climate layers were chosen to represent the primary environmental conditions that are often indicative of a combination of resource availability and disturbance regimes.

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

Table 1: Ecotype and data source used in analysis. The blended land cover product comprises the National Land Cover Database (NLCD) and Moderate Resolution Imaging Spectroradiometer (MODIS). The crop category is enhanced with CropScape and the wetland category with the National Wetland Inventory (NWI) to identify areas dominated by ecotypes with expected CH_4 source potential.

Ecotype	Expanded Ecotypes	Data Source
Water	NA	NLCD, MODIS
Developed		
Barren		
Forest		
Scrub		
Herbaceous		
Crop	Crops-non irrigated	NLCD, Crop Layer
	Crops-irrigated	NLCD, Crop Layer
Wetlands	Emergent Coastal	NLCD, MODIS, NWI
	Emergent Freshwater	
	Forested Freshwater	

2.3 Defining the state space of the US.

Climate, ecotype, and location (latitude/longitude) were used in a multivariate distance analysis (Venables and Ripley, 2002; Ripley, 2007; Cox and Cox, 2008) to define the state space of the US (all 50 states & Puerto Rico) and divide it into ecological clusters using information that is important for capturing continental patterns in CH_4. The purpose of this analysis is to identify the interrelatedness of all ecological components—biotic, abiotic, terrestrial, and aquatic within a dynamic landscape (Ippoliti et al., 2019). We included location (latitude/longitude) to incorporate the interaction between climate, ecotypes, and most importantly, seasonality. We used multidimensional scaling (MDS) to condense the ecotype, climate, and location information (ecotype, five climate variables, and location) for the US down to two dimensions using the *MASS* package in R (Venables and Ripley, 2002). We subsampled the US (n = 20,000 1-km pixels) randomly, maintaining the distribution of ecotypes and climate for the MDS analysis. We measured the correlation between the ecotypes, climate layers, and locations (latitude/longitude) using the *envfit* function in the library *vegan* in R (Oksanen, 2016). This was followed by a cluster analysis to determine the optimal number of clusters using the library *cluster* in R, which partitions data around medoids (PAM algorithm), using the Gower dissimilarity matrix (Gower, 1971; Huang, 1997; Podani, 1999; Ahmad and Dey, 2007; Harikumar and Pv, 2015). We fit an increasing number of clusters from 2 to 20 to construct a silhouette plot and choose the number of clusters that maximizes the average silhouette width to determine an optimal number of clusters. This approach finds the optimal number of clusters and places a representative cell at the center of each cluster, called the medoid. Once we determined the medoid of the cluster, we measured the difference between every location within the cluster to the medoid to understand the distance, or how representative each location was from the medoid condition of each cluster. This approach provides a

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

unit-less relative measure of representativeness between a location defined as the medoid of each cluster and every other location within that cluster. Although studies of this nature often define representativeness relative to a research site (Pallandt et al., 2021), this study identified positions within the landscape that are more representative of the medoid condition of the cluster and measured the distance from this medoid to understand how representative any location within a cluster is to the medoid condition.

To extrapolate the cluster and distance layers across the entire US beyond the 20,000-pixel subsample, we fit a Random Forest model with the package *randomForest* (Liaw and Wiener, 2002) to model the first and second MDS dimension using the ecotype and climate layers as predictors. We then created a Random Forest model of the cluster layer using the first and second dimension as the explanatory variables. All models were then projected spatially to produce a spatially explicit cluster layer and a distance to medoid layer beyond the 20,000 sample points that we used in the MDS analysis. Regions more similar to the medoid were considered more representative within their given cluster. A major assumption of this approach is that the medoid is more representative of the cluster than locations that are less similar. At the same time, to understand the variation within a cluster, it is essential to capture the distribution of distances from the medoid. This approach of defining the medoid of a cluster and measuring the distance to the medoid for each location in space allows us to understand representativeness of the location to the rest of the cluster.

We extracted the cluster and distance to the medoid for 411 reported towers to evaluate the representativeness of tower infrastructure across the US. This was done first for all EC tower infrastructure to capture the representativeness of all EC infrastructure and then just for tower infrastructure measuring CH_4 (n = 94). We then repeated this analysis for the 161 aquatic sites included in our study. We show the type of sites measured within clusters and we extract the cluster for each site and the distance to the medoid to evaluate the representativeness of aquatic network infrastructure. We recognize that different measurements are made in these networks, and few of these sites are making continuous measurements of CH_4 and other greenhouse gases.

3 Results and Discussion

3.1 Defining the state space of the US

The US was divided into 10 clusters (Figure 1) that were distributed across temperature and wetness gradients (Table 2). Latitude (R^2 = 0.95; p <0.001), mean annual temperature (R^2 = 0.84; p <0.001), maximum temperature (R^2 = 0.83; p <0.001), vapor pressure deficit (R^2 = 0.83; p <0.001), minimum temperature (R^2 = 0.82; p <0.001), longitude (R^2 = 0.63; p <0.001) had strong effects on clustering, whereas precipitation (R^2 = 0.10; p <0.001), and ecotype (R^2 = 0.03; p <0.001) showed low correlations. The coldest zones were in Alaska and included clusters Na and Nb. Cool to temperate clusters in the midwestern and western US include clusters NW, W, and NEa. Temperate clusters extend from the midwestern to the eastern US and include clusters NEb and Ea. Warm regions were distributed across clusters Eb, SW and SE. Dry clusters (Na, SW, W, & Nb) were distributed across the western US and Alaska, and wet clusters (Ea, Eb, and SE) are in the south-eastern US and Hawaii. Individual clusters

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

represent 7-16% of the US each by area (Table 2) with cluster NW as the largest cluster in the pacific northwest, and the smallest cluster being cluster Nb in the northern half of Alaska. We found the size of the cluster is not correlated to the number of towers when all towers are included in the analysis but was slightly negatively correlated with the number of EC towers that include CH_4 measurements (Figure 2).

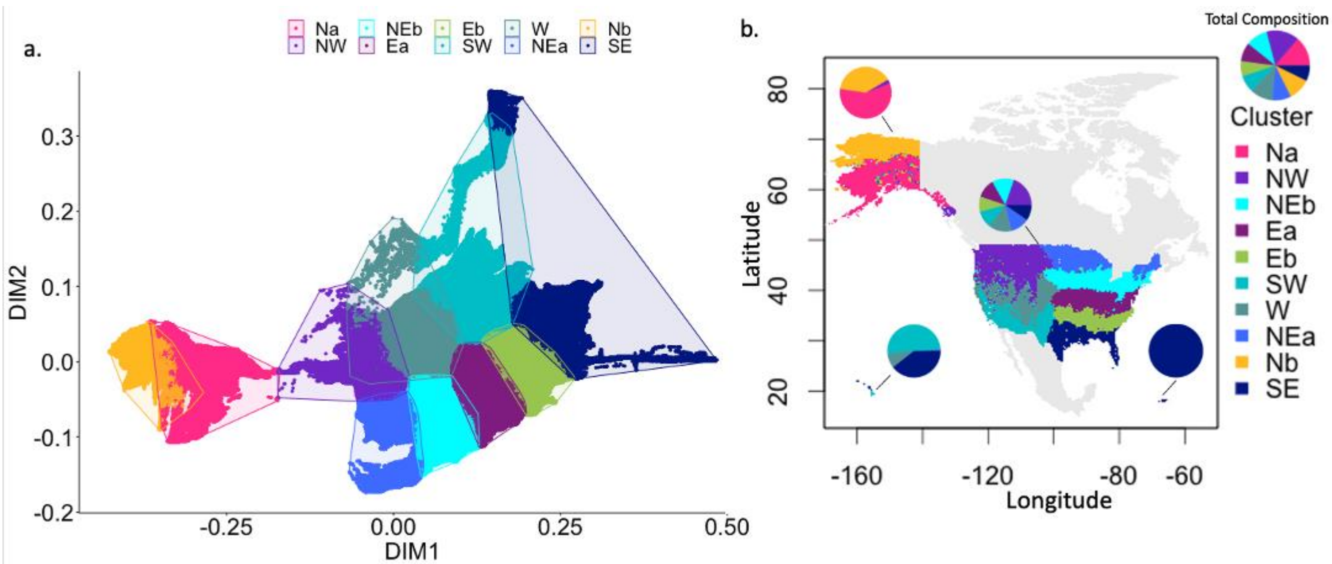

Figure 1: (a) Multidimensional scaling across the United States (US) into ten clusters using ecotype (Table1), climate, and location (lat/long). (b) Spatial distribution of the identified clusters.

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

Table 2: Ecotype, climate, and number of towers of 10 clusters in the US. Percent coverage (% Cov) is the area occupied by dominant ecotypes. We also show the total number of eddy covariance (EC) towers, the number of EC towers measuring CH_4, and the number of aquatic network (AQ) research sites in each cluster.

Cluster	Dominant Ecotypes (% Cov)					Climate	% Cov	EC	EC CH_4	AQ sites
	Forest	Scrub	Herb	Crop	Wet					
Na	27	39	4	0	2	Cold-cool (Dry)	11	14	6	8
NW	28	33	23	7	0.3	Cool-Temperate (Mild)	16	42	10	10
NEb	24	0.4	9	34	2	Temperate (Mild-Wet)	10	62	9	25
Ea	39	1	8	24	2	Temperate (Wet)	9	52	5	20
Eb	38	4	10	15	7	Warm (Wet)	8	30	9	9
SW	3	59	17	8	0.2	Warm (dry)	9	59	20	10
W	19	43	21	7	0.1	Cool-Temperate (Dry)	12	35	3	7
NEa	27	2	6	24	8	Cool-Temperate (Mild-Wet)	9	50	17	27
Nb	9	52	16	0	0.4	Cold (Dry)	7	24	9	18
SE	19	19	8	8	10	Hot (Wet)	9	44	6	27

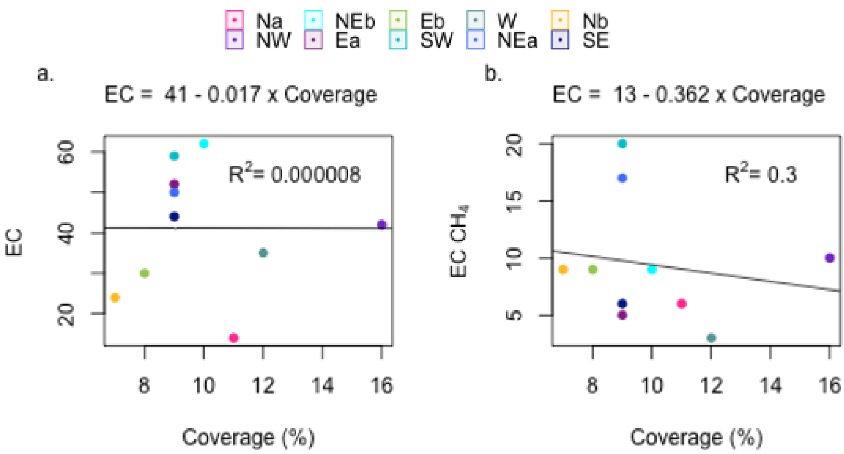

Figure 2: Percent coverage for each cluster versus the number of EC towers in a cluster for (a) all towers and (b) towers with CH_4. Lines denote linear trends.

The distance from the cluster medoid ranged from 0.01 to 0.33 (Figure 3). The mean distance was 0.04, and most areas within a cluster were less than or equal to the mean. Southern Alaska (cluster Na), Hawaii (clusters SE and Eb), Florida (cluster SE), Puerto Rico (cluster SE), and the northeast (cluster NEa) had greater distances to the medoids.

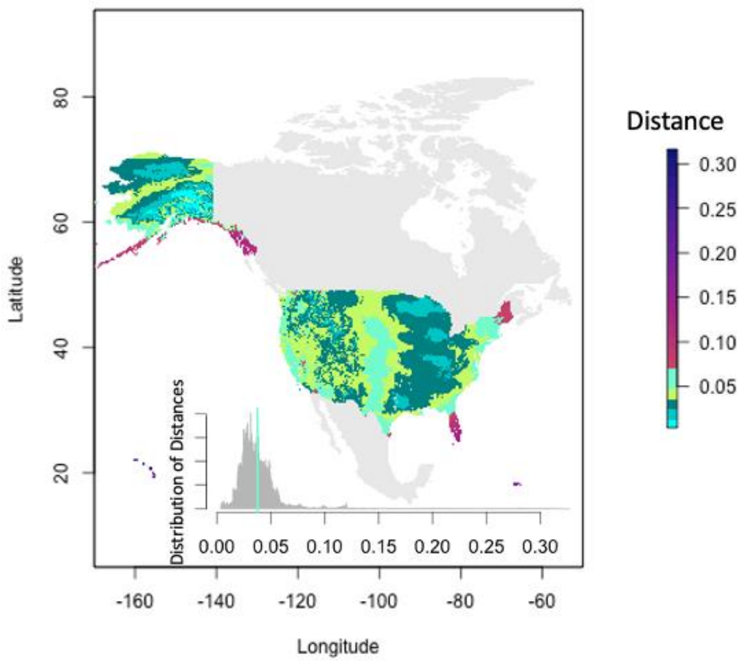

Figure 3. Distance to the cluster medoid represented spatially. Inset: the distributions of distances across all clusters shown in a histogram, in which the line denotes the mean distance across all clusters.

3.2 Representativeness of Existing Tower Infrastructure and CH_4 Tower Infrastructure

There are currently 411 reported EC towers distributed across forest (136 towers), agriculture (105 towers), wetland (69 towers), scrub (93 towers), and herbaceous (92 towers) ecotypes. Although most EC tower infrastructure was not a part of a single organized network designed to be representative of the climate and dominant land cover classes that exist within the US, EC tower infrastructure was distributed across all 10 clusters and across observed distances to the medoid for clusters NEb, Ea, SW, and SE (Table 1, Figure 4). Distance to the medoid coverage was lacking for clusters Na, W, and Nb, and currently clusters Na, W, Eb, and Nb are substantially under sampled (Figure 4c). The number of aquatic and cropland EC towers across clusters is also lacking. There is an insufficient number of towers measuring CH_4, and the distribution of these sites across distance to the medoid is poor for all clusters.

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

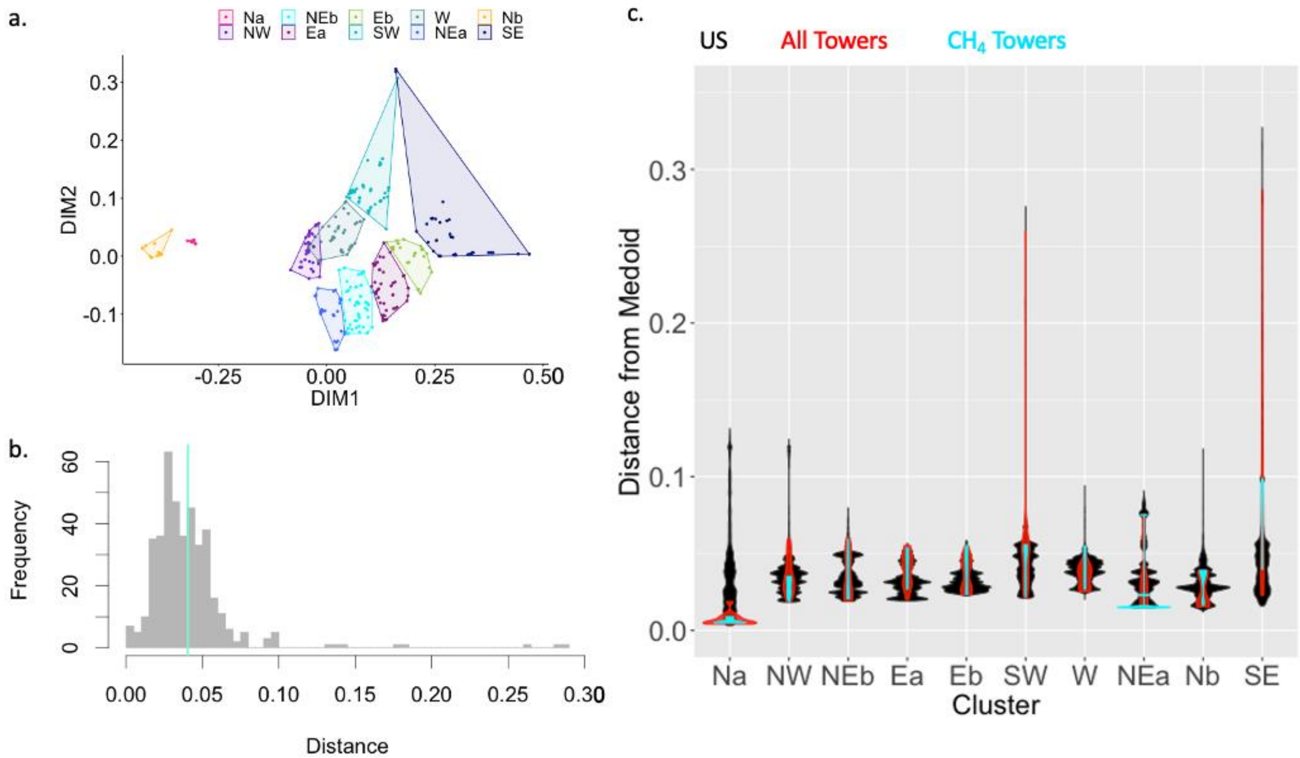

Figure 4: (a) Multidimensional scaling dimensions of the CH_4 tower infrastructure of the United States. (b) The distribution of distance from the medoid for CH_4 tower infrastructure. (c) The distribution of distance to the medoid for the US (black), for tower infrastructure (red), and EC tower infrastructure with CH_4 (cyan) by cluster.

Gaps in existing EC tower infrastructure limit our ability to constrain the contribution of US CH_4 emissions or uptakes to the global budget. These gaps occurred across all networks and independently managed sites as well as in some environmental clusters. The largest gaps in representation occurred for sites with CH_4 measurements and in aquatic sites. These gaps in representation have been pointed out in numerous investigations of CH_4 flux dynamics and budgets as a part of global CH_4 studies (Saunois et al., 2020) and FLUXNET CH_4 flux syntheses (Knox et al., 2019; Delwiche et al., 2021). In fact, the call for more measurements of CH_4 from natural sites is not new (Matthews and Fung, 1987; Bartlett and Harriss, 1993; Dlugokencky et al., 2011; Nisbet et al., 2014) and has been touted as necessary to lower the uncertainty of the CH_4 budget from natural ecosystems (Peltola et al., 2019), which is among the largest uncertainty in the global CH_4 budget (Saunois et al., 2020). There is a strong need for a continental CH_4 observatory to aid in reducing uncertainties in the natural sources and sinks of this potent greenhouse gas.

One reason for gaps in CH_4 flux tower infrastructure may be the lag in technological capability behind that of CO_2 fluxes. CH_4 gas analyzers with sufficient measurement frequency for EC were not common before the late 1990s and early 2000s (Shurpali et al., 1993; Billesbach et al., 1998; Rinne et al., 2007), and the number of commercial options has expanded only more recently (Peltola et al., 2013; Niemitz et

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

al, 2018; Burba et al, 2019, 2021). Therefore, as the flux tower infrastructure for measuring CH_4 has expanded, decisions on the locations of measurement sites have largely been tied to CO_2 and water vapor exchange research (Baldocchi, 2014) and to the availability of the grid power and suitable infrastructure (McDermitt et al, 2011). In addition to technological limitations, the environments where we expect CH_4 fluxes to be highest complicates considerations for where best to place instrumentation. Large sources of natural biogenic CH_4 can sometimes originate from small, heterogeneous components within a landscape, such as patchy wetlands within an otherwise upland forested region, causing the area to be a net source of CH_4 (Desai et al., 2015). In contrast, some systems covering large areas that are known to be important CH_4 sources, such as Arctic tundra ecosystems and shallow lakes (Wik et al., 2016; Elder et al., 2020), are simply remote and difficult to instrument. Finally, vast areas of land have been traditionally thought to have a negligible CH_4 emission or consumption rates, although these can get significant when multiplied by the area. Evidence of this can be seen in our analysis where clusters Na, W, and Nb are some of the most poorly represented clusters, corresponding to Alaska (Na & Nb) and the Rocky Mountains (W), where more rugged and remote areas exist. One limitation of this study is that a non-negligible portion of the existing CH_4 measurements, including both towers and chambers, is positioned not where CH_4 sources or sinks are but where the grid power is available to run such measurements. Positioning on the margins of the ecosystems may results in partial captures of uptakes and emissions. Our analysis does not account for this factor due to lack of coherent information available at individual site level.

3.3 Representativeness of Existing Aquatic Network Research Sites

There were 161 aquatic network sites that were distributed across the US (Table 3 and Figure 5). These sites occur in all 10 clusters, with lakes and wetlands most frequently represented. Compared to EC tower sites, there were fewer aquatic sites, and these sites were representative of the cluster medoids and not the variation within a cluster, for all clusters except cluster SE The mean distance to the medoid for aquatic network sites was 0.04, matching the mean distance for the US. In cluster SE, aquatic network sites were further from the cluster medoid, and there were no sites similar to the medoid within this cluster. Although all clusters have a coastline, only clusters Neb, Nb, and SE have coastal aquatic network sites, making this class the least sampled of aquatic network sites.

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

Table 3: The number of aquatic network sites distributed across aquatic ecosystems in the US by cluster.

Cluster	AQ sites	Ponds/ Lakes	Streams/ Rivers	Wetlands	Coastal
Na	8	1	2	5	0
NW	10	2	6	2	0
Neb	25	14	2	8	1
Ea	20	3	7	10	0
Eb	9	2	3	4	0
SW	10	0	2	8	0
W	7	1	5	1	0
NEa	27	19	1	7	0
Nb	18	2	2	13	1
SE	27	5	5	15	2

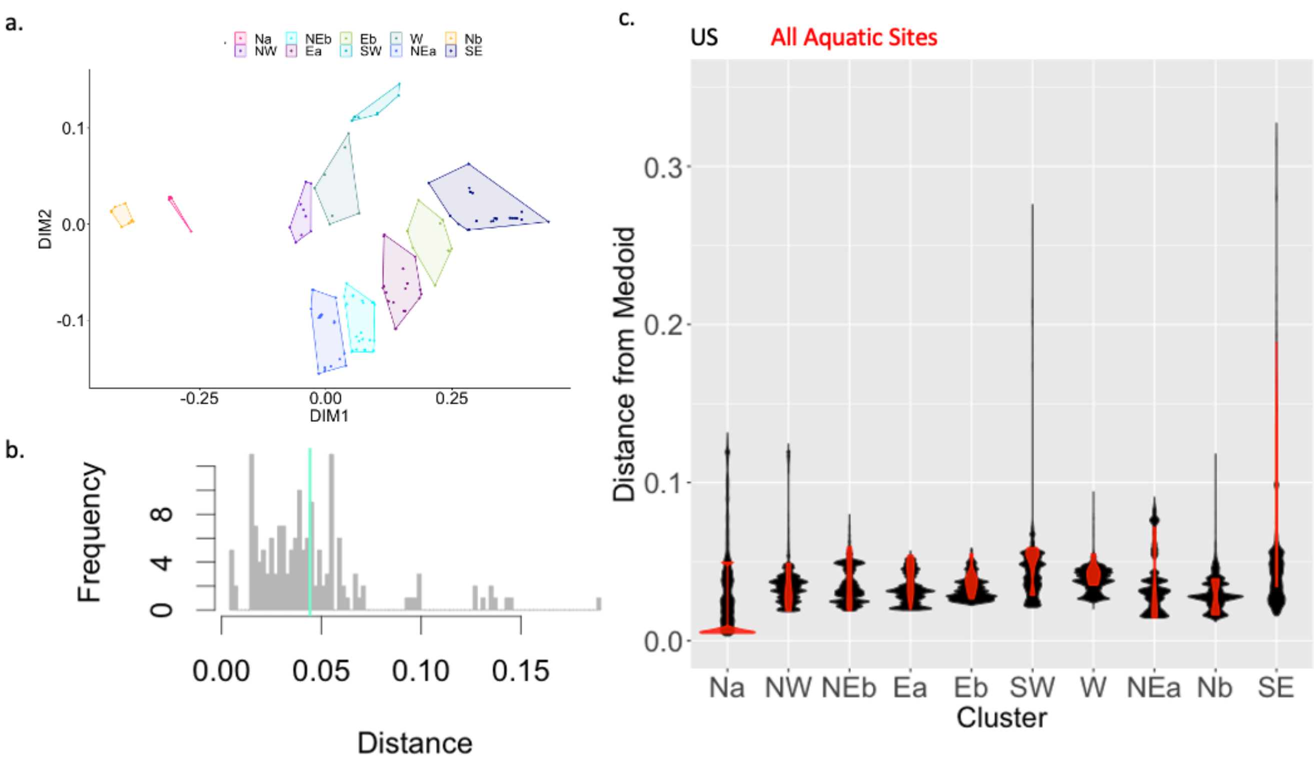

Figure 5: (a) Multidimensional scaling dimensions of the aquatic network sites in the United States (US). (b)The distribution of distance from the medoid for aquatic network sites. (c) The distribution of distance to the medoid for the US (black) and aquatic networks sites (red) by cluster.

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

Biogeosciences Discussions
Open Access

Of particular concern is the lack of representation we found in aquatic network habitats (including wetlands) given that they are responsible for roughly half of total global CH_4 emissions from anthropogenic and natural sources (Rosentreter et al., 2021). In fact, our analysis found only a few of each type of aquatic site within each cluster, and many clusters with no coastal sites. A continental methane observation network would prioritize the construction of sites in aquatic ecosystems in order to fill these measurement gaps, particularly in coastal areas given the many proposed remediation strategies for reducing CH_4 emissions from coastal wetlands (Kroeger et al., 2017). As with terrestrial ecosystems, technological and logistical challenges may help explain the gaps in monitoring CH_4 emissions from aquatic areas. Some aquatic ecosystems call for specialized approaches such as floating chambers, given that they are not often conducive to EC tower measurements, particularly smaller bodies of water or aquatic areas situated within forests that may not meet proper boundary layer and turbulence conditions. Further, aquatic ecosystems also have highly variable CH_4 flux rates due to different transport pathways including diffusion, production and transport in vegetation (Lai, 2009; Maier et al., 2018), and ebullition (Joyce and Jewell, 2003) that efflux CH_4 at different frequencies and rates, and partitioning between these processes can be difficult (Wik et al., 2016; Iwata et al., 2018).

Uncertainties in large-scale CH_4 budgets is also related to poor data coverage in measurements focused on the soil sink of CH_4. This uncertainty is due to the lack of distributed infrastructure resulting in small uptake rates assumed to occur over extensive areas of upland ecosystems (Smith et al., 2000). Increasing understanding of upland ecosystem CH_4 biogeochemistry suggests that crude scaling methods for soil CH_4 uptake may be insufficient (Covey and Megonigal, 2019). It is therefore essential that research sites are distributed both within and across clusters to capture the high variability of CH_4 sources and sinks to understand and spatially scale fluxes, and utilize the methods appropriate for measurements of such small quantities (e.g., chambers) with sufficient resolution. To do this, infrastructure must be deployed along the variation of distances from the medoid in order to capture the variation within a cluster.

The current LUC contributes additional uncertainty to the CH_4 budget, particularly the classification of inland waters and wetlands. Both these ecosystem types are heterogeneous across the landscape and can be ephemeral, challenging our ability to scale CH_4 fluxes across space and over time. Increased attention has been given to creating more accurate and finer scale maps of aquatic ecosystems, which should help improve estimates of CH_4 fluxes (Hondula et al., 2021). However, these classifications may still result in double-counting or under-counting given the apparent ephemeral conditions like variable inundation or tide levels (Parker et al., 2018). Inundation and moisture status are often used as a proxy for the redox state in soils to decide the sink or source state of ecosystems, which is also difficult to map over a regional scale, especially at scales fine enough to be meaningful in upscaling (Chu et al., 2021). The creation of a CH_4 potential map with a spatial resolution detailed enough to account for smaller lakes and ponds (<1 km) would help to better constrain and model CH_4 emissions (Zhang et al., 2021).

Besides the lack of spatial representation, CH_4 flux research often only lasts a few years at a time due to most sites being primarily operated by individual investigators with time-limited research grants. Our study includes towers that were identified based on records of sites from various networks and investigators, but that does not mean that all sites listed and used in the analysis are currently operational or have up to date and continuous data. The absence of long-term data streams can create issues in constraining CH_4 budgets in that interannual variability across ecosystems and space cannot be accounted for, which is particularly important given ongoing changes in the climate system. In order to enable more accurate and constrained estimates of CH_4 from natural ecosystems, a long-term and targeted observatory network is necessary to capture spatial and temporal variation.

4 Conclusions

Evaluating the strengths and limitations of existing measurement infrastructure is critical for strategic augmentation to provide the most valuable information toward reducing uncertainties in future large-scale budget estimations. To accurately understand the impact of climate and land cover change on biogenic CH_4 emissions, we need a long-term, calibrated, and strategic continental-scale CH_4 observatory network. Current gaps in existing measurement infrastructure limit our ability to capture the spatial and temporal variations of biogenic CH_4 fluxes and therefore limit our ability to predict future CH_4 emissions. Maps of potential CH_4 emissions require land cover classification targeted at land cover types like wetlands that are important sources of CH_4 to the atmosphere. Aquatic ecosystems like streams and lakes as well as coastal ecosystems are significant and variable sources of CH_4 not well studied on a long-term basis. Through our analysis using climate, land cover, and location variables, we have identified priority areas to enhance research infrastructure to provide a more complete understanding of the CH_4 flux potential of ecosystem types in the US. For EC tower locations, distance to the medoid coverage was lacking for clusters Na, W, and Nb, and currently clusters Na, W, Eb, and Nb are substantially under sampled. All aquatic sites are under sampled within each cluster. An enhanced network would allow for us to monitor both the response of CH_4 fluxes to climate and land use change as well as the impact of future policy interventions and mitigation strategies.

Code/Data availability: The data products produced are available on the Knowledge Network for Biocomplexity (https://knb.ecoinformatics.org; Malone 2021)

Author contribution: All authors contributed to conceptualization; SLM, KAA, RC, ARC, JPG, and RKV designed the manuscript; SLM, KAA, JPG, GS, and RKV wrote parts of the manuscript; RKV, SLM and JPG supervised the project/manuscript; SLM, KAA, and YO contributed to data curation and formal analysis, and all authors performed critical reviews.

Competing interests: The authors report no competing interests.

Acknowledgements

The authors would like to acknowledged the following researchers and support staff for their contributions to discussions that led to the idea for this manuscript: Melissa Genazzio, Amy Lafreniere, Kim Nitschke, Lori Bruhwiler, Julia Bryce, Patrick Crill, Amarnath Gupta, Ilya Zaslavsky, Ilkay Altintas, Stephen Hale, Mike Stewart, Michael Thomson and Mark Milutinovich. HWL acknowledges

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

the National Science Foundation (NSF) for ongoing support. NEON is a project sponsored by the NSF and managed under cooperative support agreement (EF-1029808) to Battelle. RKV and AC acknowledge UNH's Collaborative Research Excellence (CoRE) grant. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of our sponsoring agencies.

References

Ahmad, A. and Dey, L.: A k-mean clustering algorithm for mixed numeric and categorical data, Data Knowl. Eng., 63, 503–527, https://doi.org/10.1016/j.datak.2007.03.016, 2007.

Baldocchi, D.: Measuring fluxes of trace gases and energy between ecosystems and the atmosphere - the state and future of the eddy covariance method, Glob. Chang. Biol., 20, 3600–3609, https://doi.org/10.1111/gcb.12649, 2014.

Barkley, Z. R., Lauvaux, T., Davis, K. J., Deng, A., Miles, N. L., Richardson, S. J., Cao, Y., Sweeney, C., Karion, A., Smith, M., and Others: Quantifying methane emissions from natural gas production in north-eastern Pennsylvania, Atmos. Chem. Phys., 17, 13941–13966, 2017.

Bartlett, K. B. and Harriss, R. C.: Review and assessment of methane emissions from wetlands, Chemosphere, 26, 261–320, https://doi.org/10.1016/0045-6535(93)90427-7, 1993.
Bessembinder, J., Overbeek, B., and Siegmund, P.: Climate normals and climate change: how to communicate these together?, EGU21–4032, 2021.

Billesbach, D. P., Kim, J., Clement, R. J., Verma, S. B., and Ullman, F. G.: An Intercomparison of Two Tunable Diode Laser Spectrometers Used for Eddy Correlation Measurements of Methane Flux in a Prairie Wetland, https://doi.org/10.1175/1520-0426(1998)015<0197:aiottd>2.0.co;2, 1998.

Boryan, C., Yang, Z., Mueller, R., and Craig, M.: Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program, Geocarto Int., 26, 341–358, https://doi.org/10.1080/10106049.2011.562309, 2011.

Bruhwiler, L., Parmentier, F.-J. W., Crill, P., Leonard, M., and Palmer, P. I.: The Arctic Carbon Cycle and Its Response to Changing Climate, 7, 14–34, https://doi.org/10.1007/s40641-020-00169-5, 2021.

Burba, G., Atmospheric Flux Measurements. In W. Chen, D. Venables, and M. Sigris (Eds.). Advances in Spectroscopic Monitoring of the Atmosphere. Elsevier Science, 618 pp. ISBN: 978-0-12815014-6, 2021

Burba G., Anderson T., Komissarov A.: Accounting for spectroscopic effects in laser-based open-path eddy covariance flux measurements. Global Change Biology. 25(6):2189-202, 2019

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

Chadburn, S. E., Burke, E. J., Cox, P. M., Friedlingstein, P., Hugelius, G., and Westermann, S.: An observation-based constraint on permafrost loss as a function of global warming, Nat. Clim. Chang., 7, 340–344, https://doi.org/10.1038/nclimate3262, 2017.

Chang, R. Y.-W., Miller, C. E., Dinardo, S. J., Karion, A., Sweeney, C., Daube, B. C., Henderson, J. M., Mountain, M. E., Eluszkiewicz, J., Miller, J. B., Bruhwiler, L. M. P., and Wofsy, S. C.: Methane emissions from Alaska in 2012 from CARVE airborne observations, Proc. Natl. Acad. Sci. U. S. A., 111, 16694–16699, https://doi.org/10.1073/pnas.1412953111, 2014.

Chu, H., Luo, X., Ouyang, Z., Chan, W. S., Dengel, S., Biraud, S. C., Torn, M. S., Metzger, S., Kumar, J., Arain, M. A., Arkebauer, T. J., Baldocchi, D., Bernacchi, C., Billesbach, D., Black, T. A., Blanken, P. D., Bohrer, G., Bracho, R., Brown, S., Brunsell, N. A., Chen, J., Chen, X., Clark, K., Desai, A. R., Duman, T., Durden, D., Fares, S., Forbrich, I., Gamon, J. A., Gough, C. M., Griffis, T., Helbig, M., Hollinger, D., Humphreys, E., Ikawa, H., Iwata, H., Ju, Y., Knowles, J. F., Knox, S. H., Kobayashi, H., Kolb, T., Law, B., Lee, X., Litvak, M., Liu, H., Munger, J. W., Noormets, A., Novick, K., Oberbauer, S. F., Oechel, W., Oikawa, P., Papuga, S. A., Pendall, E., Prajapati, P., Prueger, J., Quinton, W. L., Richardson, A. D., Russell, E. S., Scott, R. L., Starr, G., Staebler, R., Stoy, P. C., Stuart-Haëntjens, E., Sonnentag, O., Sullivan, R. C., Suyker, A., Ueyama, M., Vargas, R., Wood, J. D., and Zona, D.: Representativeness of Eddy-Covariance flux footprints for areas surrounding AmeriFlux sites, Agric. For. Meteorol., 301-302, 108350, https://doi.org/10.1016/j.agrformet.2021.108350, 2021.

Covey, K. R. and Megonigal, J. P.: Methane production and emissions in trees and forests, New Phytol., 222, 35–51, https://doi.org/10.1111/nph.15624, 2019.

Cox, M. A. A. and Cox, T. F.: Multidimensional Scaling, in: Handbook of Data Visualization, edited by: Chen, C.-H., Härdle, W., and Unwin, A., Springer Berlin Heidelberg, Berlin, Heidelberg, 315–347, https://doi.org/10.1007/978-3-540-33037-0_14, 2008.

Delwiche, K. B., Knox, S. H., Malhotra, A., Fluet-Chouinard, E., McNicol, G., Feron, S., Ouyang, Z., Papale, D., Trotta, C., Canfora, E., Cheah, Y.-W., Christianson, D., Alberto, M. C. R., Alekseychik, P., Aurela, M., Baldocchi, D., Bansal, S., Billesbach, D. P., Bohrer, G., Bracho, R., Buchmann, N., Campbell, D. I., Celis, G., Chen, J., Chen, W., Chu, H., Dalmagro, H. J., Dengel, S., Desai, A. R., Detto, M., Dolman, H., Eichelmann, E., Euskirchen, E., Famulari, D., Fuchs, K., Goeckede, M., Gogo, S., Gondwe, M. J., Goodrich, J. P., Gottschalk, P., Graham, S. L., Heimann, M., Helbig, M., Helfter, C., Hemes, K. S., Hirano, T., Hollinger, D., Hörtnagl, L., Iwata, H., Jacotot, A., Jurasinski, G., Kang, M., Kasak, K., King, J., Klatt, J., Koebsch, F., Krauss, K. W., Lai, D. Y. F., Lohila, A., Mammarella, I., Belelli Marchesini, L., Manca, G., Matthes, J. H., Maximov, T., Merbold, L., Mitra, B., Morin, T. H., Nemitz, E., Nilsson, M. B., Niu, S., Oechel, W. C., Oikawa, P. Y., Ono, K., Peichl, M., Peltola, O., Reba, M. L., Richardson, A. D., Riley, W., Runkle, B. R. K., Ryu, Y., Sachs, T., Sakabe, A., Sanchez, C. R., Schuur, E. A., Schäfer, K. V. R., Sonnentag, O., Sparks, J. P., Stuart-Haëntjens, E., Sturtevant, C., Sullivan, R. C., Szutu, D. J., Thom, J. E., Torn, M. S., Tuittila, E.-S., Turner, J., Ueyama, M., Valach, A. C., Vargas, R., Varlagin, A., et al.: FLUXNET-CH4: a global, multi-ecosystem dataset and analysis of methane seasonality from freshwater wetlands, Earth Syst. Sci. Data, 13, 3607–3689, https://doi.org/10.5194/essd-13-3607-2021, 2021.

Desai, A. R.: Climatic and phenological controls on coherent regional interannual variability of carbon dioxide flux in a heterogeneous landscape, J. Geophys. Res., 115, https://doi.org/10.1029/2010jg001423, 2010.

Desai, A. R., Xu, K., Tian, H., Weishampel, P., Thom, J., Baumann, D., Andrews, A. E., Cook, B. D., King, J. Y., and Kolka, R.: Landscape-level terrestrial methane flux observed from a very tall tower, Agric. For. Meteorol., 201, 61–75, https://doi.org/10.1016/j.agrformet.2014.10.017, 2015.

Dlugokencky, E.: Trends in Atmospheric Methane Global CH4 Monthly Means, NOAA, 2021.
Dlugokencky, E. J., Nisbet, E. G., Fisher, R., and Lowry, D.: Global atmospheric methane: budget, changes and dangers, Philos. Trans. A Math. Phys. Eng. Sci., 369, 2058–2072, https://doi.org/10.1098/rsta.2010.0341, 2011.

Elder, C. D., Thompson, D. R., Thorpe, A. K., Hanke, P., Walter Anthony, K. M., and Miller, C. E.: Airborne mapping reveals emergent power law of arctic methane emissions, Geophys. Res. Lett., 47, https://doi.org/10.1029/2019gl085707, 2020.

Gower, J. C.: A General Coefficient of Similarity and Some of Its Properties, Biometrics, 27, 857–871, https://doi.org/10.2307/2528823, 1971.

Harikumar, S. and Pv, S.: K-Medoid Clustering for Heterogeneous DataSets, Procedia Comput. Sci., 70, 226–237, https://doi.org/10.1016/j.procs.2015.10.077, 2015.

Hijmans, R. J.: Geographic Data Analysis and Modeling [R package raster version 3.4-13], 2021.

Hondula, K. L., DeVries, B., Jones, C. N., and Palmer, M. A.: Effects of using high resolution satellite-based inundation time series to estimate methane fluxes from forested wetlands, Geophys. Res. Lett., 48, https://doi.org/10.1029/2021gl092556, 2021.

Huang, Z.: Clustering large data sets with mixed numeric and categorical values, in: Proceedings of the 1st pacific-asia conference on knowledge discovery and data mining,(PAKDD), 21–34, 1997.

Hutchins D.A., Jansson J.K., Remais J.V., Rich V.I., Singh B.K., Trivedi P.: Climate change microbiology—problems and perspectives. Nature Reviews Microbiology. 17(6): 391-6, 2019**.**

IPCC: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change, USA: Cambridge University Press, 2013.

IPCC: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2021.

Ippoliti, C., Candeloro, L., Gilbert, M., Goffredo, M., Mancini, G., Curci, G., Falasca, S., Tora, S., Di Lorenzo, A., Quaglia, M., and Conte, A.: Defining ecological regions in Italy based on a multivariate clustering approach: A first step towards a targeted vector borne disease surveillance, PLoS One, 14, e0219072, https://doi.org/10.1371/journal.pone.0219072, 2019.

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

Jin, S., Homer, C., Yang, L., Danielson, P., Dewitz, J., Li, C., Zhu, Z., Xian, G., and Howard, D.: Overall Methodology Design for the United States National Land Cover Database 2016 Products, Remote Sensing, 11, 2971, https://doi.org/10.3390/rs11242971, 2019.

Joyce, J. and Jewell, P. W.: Physical Controls on Methane Ebullition from Reservoirs and Lakes, Environ. Eng. Geosci., 9, 167–178, https://doi.org/10.2113/9.2.167, 2003.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Altaf Arain, M., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, https://doi.org/10.1029/2010jg001566, 2011.

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J. G., Dlugokencky, E. J., Bergamaschi, P., Bergmann, D., Blake, D. R., Bruhwiler, L., Cameron-Smith, P., Castaldi, S., Chevallier, F., Feng, L., Fraser, A., Heimann, M., Hodson, E. L., Houweling, S., Josse, B., Fraser, P. J., Krummel, P. B., Lamarque, J.-F., Langenfelds, R. L., Le Quéré, C., Naik, V., O'Doherty, S., Palmer, P. I., Pison, I., Plummer, D., Poulter, B., Prinn, R. G., Rigby, M., Ringeval, B., Santini, M., Schmidt, M., Shindell, D. T., Simpson, I. J., Spahni, R., Steele, L. P., Strode, S. A., Sudo, K., Szopa, S., van der Werf, G. R., Voulgarakis, A., van Weele, M., Weiss, R. F., Williams, J. E., and Zeng, G.: Three decades of global methane sources and sinks, Nat. Geosci., 6, 813, https://doi.org/10.1038/ngeo1955, 2013.

Knox, S. H., Jackson, R. B., Poulter, B., McNicol, G., Fluet-Chouinard, E., Zhang, Z., Hugelius, G., Bousquet, P., Canadell, J. G., Saunois, M., and Others: FLUXNET-CH 4 Synthesis Activity: Objectives, Observations, and Future Directions, Bull. Am. Meteorol. Soc., 100, 2607–2632, 2019.

Kroeger, K. D., Crooks, S., Moseman-Valtierra, S., and Tang, J.: Restoring tides to reduce methane emissions in impounded wetlands: A new and potent Blue Carbon climate change intervention, Sci. Rep., 7, 11914, https://doi.org/10.1038/s41598-017-12138-4, 2017.

Lai, D. Y. F.: Methane Dynamics in Northern Peatlands: A Review, Pedosphere, 19, 409–421, https://doi.org/10.1016/S1002-0160(09)00003-4, 2009.

Lassey, K. R.: Livestock methane emission and its perspective in the global methane cycle, Aust. J. Exp. Agric., 48, 114–118, https://doi.org/10.1071/EA07220, 2008.

Liaw, A. and Wiener, M.: Classification and regression by randomForest, 2, 18–22, 2002.

Lu, X., Zhou, Y., Zhuang, Q., Prigent, C., Liu, Y., and Teuling, A.: Increasing methane emissions from natural land ecosystems due to sea-level rise, J. Geophys. Res. Biogeosci., 123, 1756–1768, https://doi.org/10.1029/2017jg004273, 2018.

Lu, X., Jacob, D. J., Zhang, Y., Maasakkers, J. D., Sulprizio, M. P., Shen, L., Qu, Z., Scarpelli, T. R., Nesser, H., Yantosca, R. M., Sheng, J., Andrews, A., Parker, R. J., Boesch, H., Anthony Bloom, A., and Ma, S.: Global methane budget and trend, 2010–2017: complementarity of inverse analyses using in situ (GLOBALVIEWplus CH4 ObsPack) and satellite (GOSAT) observations, https://doi.org/10.5194/acp-21-4637-2021, 2021.

Maier, M., Machacova, K., Lang, F., Svobodova, K., and Urban, O.: Combining soil and tree-stem flux measurements and soil gas profiles to understand CH4 pathways in Fagus sylvatica forests, J. Plant Nutr. Soil Sci., 181, 31–35, https://doi.org/10.1002/jpln.201600405, 2018.

Malone, S. L. (2021). Gaps in Network Infrastructure limit our understanding of biogenic methane emissions in the United States [Data set]. In *Gaps in Network Infrastructure limit our understanding of biogenic methane emissions in the United States.* https://doi.org/urn:uuid:cb1e9014-92f3-49e4-a853-180e215fdebe.

Matthews, E. and Fung, I.: Methane emission from natural wetlands: Global distribution, area, and environmental characteristics of sources, Global Biogeochem. Cycles, 1, 61–86, https://doi.org/10.1029/GB001i001p00061, 1987.
Michalak, A. M., Jackson, R., Marland, G., and Sabine, C.: A U.S. Carbon Cycle Science Plan: First Meeting of the Carbon Cycle Science Working Group; Washington, D. C., 17-18 November 2008, https://doi.org/10.1029/2009eo120003, 2009.

McDermitt D., Burba G., Xu L., Anderson T., Komissarov A., Riensche B., Schedlbauer J., Starr G., Zona D., Oechel W., Oberbauer S.: A new low-power, open-path instrument for measuring methane flux by eddy covariance. Applied Physics B. 102(2):391-405, 2011

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

Nemitz E., Mammarella I., Ibrom A., Aurela M., Burba G., Dengel S., Gielen B., Grelle A., Heinesch B., Herbst M., Hörtnagl L. Standardisation of eddy-covariance flux measurements of methane and nitrous oxide. International Agrophysics. 2018.

Nisbet, E. G., Dlugokencky, E. J., and Bousquet, P.: Methane on the Rise—Again, Science, 343, 493–495, https://doi.org/10.1126/science.1247828, 2014.

Nisbet, E. G., Manning, M. R., Dlugokencky, E. J., Fisher, R. E., Lowry, D., Michel, S. E., Myhre, C. L., Platt, S. M., Allen, G., Bousquet, P., Brownlow, R., Cain, M., France, J. L., Hermansen, O., Hossaini, R., Jones, A. E., Levin, I., Manning, A. C., Myhre, G., Pyle, J. A., Vaughn, B. H., Warwick, N. J., and White, J. W. C.: Very strong atmospheric methane growth in the 4 years 2014–2017: Implications for the Paris agreement, Global Biogeochem. Cycles, 33, 318–342, https://doi.org/10.1029/2018gb006009, 2019a.

Nisbet, E. G., Manning, M. R., Dlugokencky, E. J., Fisher, R. E., Lowry, D., Michel, S. E., Myhre, C. L., Platt, S. M., Allen, G., Bousquet, P., Brownlow, R., Cain, M., France, J. L., Hermansen, O., Hossaini, R., Jones, A. E., Levin, I., Manning, A. C., Myhre, G., Pyle, J. A., Vaughn, B. H., Warwick, N. J., and White, J. W. C.: Very strong atmospheric methane growth in the 4 years 2014–2017: Implications for the Paris agreement, Global Biogeochem. Cycles, 33, 318–342, https://doi.org/10.1029/2018gb006009, 2019b.

Ni, X. and Groffman, P. M.: Declines in methane uptake in forest soils, Proc. Natl. Acad. Sci. U. S. A., 115, 8587–8590, https://doi.org/10.1073/pnas.1807377115, 2018.

Oh, Y., Zhuang, Q., Liu, L., Welp, L. R., Lau, M. C. Y., Onstott, T. C., Medvigy, D., Bruhwiler, L., Dlugokencky, E. J., Hugelius, G., D'Imperio, L., and Elberling, B.: Reduced net methane emissions due to microbial methane oxidation in a warmer Arctic, Nat. Clim. Chang., 10, 317–321, https://doi.org/10.1038/s41558-020-0734-z, 2020.

Oksanen, J.: Vegan: an introduction to ordination, ftp://ftp.rediris.net/mirror/R/web/packages/vegan/vignettes/intro-vegan.pdf, 2016.

Pallandt, M., Kumar, J., Mauritz, M., Schuur, E., Virkkala, A.-M., Celis, G., Hoffman, F., and Göckede, M.: Representativeness assessment of the pan-Arctic eddy-covariance site network, and optimized future enhancements, Biogeosciences, 1–42, https://doi.org/10.5194/bg-2021-133, 2021.

Parker, R. J., Boesch, H., McNorton, J., Comyn-Platt, E., Gloor, M., Wilson, C., Chipperfield, M. P., Hayman, G. D., and Bloom, A. A.: Evaluating year-to-year anomalies in tropical wetland methane emissions using satellite CH4 observations, Remote Sens. Environ., 211, 261–275, https://doi.org/10.1016/j.rse.2018.02.011, 2018.

Peltola, O., Mammarella, I., Haapanala, S., Burba, G., and Vesala, T.: Field intercomparison of four methane gas analyzers suitable for eddy covariance flux measurements, https://doi.org/10.5194/bg-10-3749-2013, 2013.

Peltola, O., Vesala, T., Gao, Y., Räty, O., Alekseychik, P., Aurela, M., Chojnicki, B., Desai, A. R., Dolman, A. J., Euskirchen, E. S., Friborg, T., Göckede, M., Helbig, M., Humphreys, E., Jackson, R. B., Jocher, G., Joos, F., Klatt, J., Knox, S. H., Kowalska, N., Kutzbach, L., Lienert, S., Lohila, A., Mammarella, I., Nadeau, D. F., Nilsson, M. B., Oechel, W. C., Peichl, M., Pypker, T., Quinton, W., Rinne, J., Sachs, T., Samson, M., Schmid, H. P., Sonnentag, O., Wille, C., Zona, D., and Aalto, T.: Monthly gridded data product of northern wetland methane emissions based on upscaling eddy covariance observations, https://doi.org/10.5194/essd-11-1263-2019, 2019.

Podani, J.: Extending Gower's general coefficient of similarity to ordinal characters, Taxon, 48, 331–340, https://doi.org/10.2307/1224438, 1999.

R Core Team: R: A language and environment for statistical computing, 2021.

Rinne, J., Riutta, T., Pihlatie, M., Aurela, M., Haapanala, S., Tuovinen, J.-P., Tuittila, E.-S., and Vesala, T.: Annual cycle of methane emission from a boreal fen measured by the eddy covariance technique, Tellus B Chem. Phys. Meteorol., 59, 449–457, https://doi.org/10.1111/j.1600-0889.2007.00261.x, 2007.

Ripley, B. D.: Pattern Recognition and Neural Networks, Cambridge University Press, 403 pp., 2007.
Robertson, G. P., Paul, E. A., and Harwood, R. R.: Greenhouse gases in intensive agriculture: contributions of individual gases to the radiative forcing of the atmosphere, Science, 289, 1922–1925, https://doi.org/10.1126/science.289.5486.1922, 2000.

Rosentreter, J. A., Borges, A. V., Deemer, B. R., Holgerson, M. A., Liu, S., Song, C., Melack, J., Raymond, P. A., Duarte, C. M., Allen, G. H., Olefeldt, D., Poulter, B., Battin, T. I., and Eyre, B. D.: Half of global methane emissions come from highly variable aquatic ecosystem sources, Nat. Geosci., 14, 225–230, https://doi.org/10.1038/s41561-021-00715-2, 2021.

Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond, P. A., Dlugokencky, E. J., Houweling, S., Patra, P. K., Ciais, P., Arora, V. K., Bastviken, D., Bergamaschi, P., Blake, D. R., Brailsford, G., Bruhwiler, L., Carlson, K. M., Carrol, M., Castaldi, S., Chandra, N., Crevoisier, C., Crill, P. M., Covey, K., Curry, C. L., Etiope, G., Frankenberg, C., Gedney, N., Hegglin, M. I., Höglund-Isaksson, L., Hugelius, G., Ishizawa, M., Ito, A., Janssens-Maenhout, G., Jensen, K. M., Joos, F., Kleinen, T., Krummel, P. B., Langenfelds, R. L., Laruelle, G. G., Liu, L., Machida, T., Maksyutov, S., McDonald, K. C., McNorton, J., Miller, P. A., Melton, J. R., Morino, I., Müller, J., Murguia-Flores, F., Naik, V., Niwa, Y., Noce, S., O'Doherty, S., Parker, R. J., Peng, C., Peng, S., Peters, G. P., Prigent, C., Prinn, R., Ramonet, M., Regnier, P., Riley, W. J., Rosentreter, J. A., Segers, A., Simpson, I. J., Shi, H., Smith, S. J., Steele, L. P., Thornton, B. F., Tian, H., Tohjima, Y., Tubiello, F. N., Tsuruta, A., Viovy, N., Voulgarakis, A., Weber, T. S., van Weele, M., van der Werf, G. R., Weiss,

https://doi.org/10.5194/bg-2021-256
Preprint. Discussion started: 5 October 2021

R. F., Worthy, D., Wunch, D., Yin, Y., Yoshida, Y., Zhang, W., Zhang, Z., Zhao, Y., Zheng, B., Zhu, Q., Zhu, Q., and Zhuang, Q.: The global methane budget 2000–2017, Earth Syst. Sci. Data, 12, 1561–1623, https://doi.org/10.5194/essd-12-1561-2020, 2020.

Schaefer, H.: On the Causes and Consequences of Recent Trends in Atmospheric Methane, 5, 259–274, https://doi.org/10.1007/s40641-019-00140-z, 2019.

Serreze, M. C. and Barry, R. G.: Processes and impacts of Arctic amplification: A research synthesis, Glob. Planet. Change, 77, 85–96, https://doi.org/10.1016/j.gloplacha.2011.03.004, 2011.

Sherwood, O. A., Schwietzke, S., Arling, V. A., and Etiope, G.: Global inventory of gas geochemistry data from fossil fuel, microbial and burning sources, version 2017, Earth Syst. Sci. Data, 9, 639–656, https://doi.org/10.5194/essd-9-639-2017, 2017.

Shurpali, N. J., Verma, S. B., Clement, R. J., and Billesbach, D. P.: Seasonal distribution of methane flux in a Minnesota peatland measured by eddy correlation, J. Geophys. Res., 98, 20649, https://doi.org/10.1029/93jd02181, 1993.

Smith, K. A., Dobbie, K. E., Ball, B. C., Bakken, L. R., Sitaula, B. K., Hansen, S., Brumme, R., Borken, W., Christensen, S., Priemé, A., Fowler, D., Macdonald, J. A., Skiba, U., Klemedtsson, L., Kasimir-Klemedtsson, A., Degórska, A., and Orlanski, P.: Oxidation of atmospheric methane in Northern European soils, comparison with other ecosystems, and uncertainties in the global terrestrial sink, Glob. Chang. Biol., 6, 791–803, https://doi.org/10.1046/j.1365-2486.2000.00356.x, 2000.

Sulla-Menashe, D. and Friedl, M. A.: User guide to collection 6 MODIS land cover (MCD12Q1 and MCD12C1) product, 1–18, 2018.

Thornton, B. F., Wik, M., and Crill, P. M.: Double-counting challenges the accuracy of high-latitude methane inventories, Geophys. Res. Lett., 43, 12,569–12,577, https://doi.org/10.1002/2016gl071772, 2016.

Thornton, M. M., Thornton, P. E., Wei, Y., Vose, R. S., and Boyer, A. G.: Daymet: Station-level inputs and model predicted values for North America, Version 3, 2017.

Venables WNRipley, B. D.: Modern applied statistics with S, 2002.
Wik, M., Varner, R. K., Anthony, K. W., MacIntyre, S., and Bastviken, D.: Climate-sensitive northern lakes and ponds are critical components of methane release, https://doi.org/10.1038/ngeo2578, 2016.

Wilen, B. O. and Bates, M. K.: The US fish and wildlife service's national wetlands inventory project, in: Classification and Inventory of the World's Wetlands, Springer Netherlands, Dordrecht, 153–169, https://doi.org/10.1007/978-94-011-0427-2_13, 1995.

Xiao, J., Chen, J., Davis, K. J., and Reichstein, M.: Advances in upscaling of eddy covariance measurements of carbon and water fluxes, J. Geophys. Res., 117, https://doi.org/10.1029/2011jg001889, 2012.

Yumashev, D., Hope, C., Schaefer, K., Riemann-Campe, K., Iglesias-Suarez, F., Jafarov, E., Burke, E. J., Young, P. J., Elshorbany, Y., and Whiteman, G.: Climate policy implications of nonlinear decline of Arctic land permafrost and other cryosphere elements, Nat. Commun., 10, 1900, https://doi.org/10.1038/s41467-019-09863-x, 2019.

Zhang, Z., Zimmermann, N. E., Stenke, A., Li, X., Hodson, E. L., Zhu, G., Huang, C., and Poulter, B.: Emerging role of wetland methane emissions in driving 21st century climate change, Proc. Natl. Acad. Sci. U. S. A., 114, 9647–9652, https://doi.org/10.1073/pnas.1618765114, 2017.

Zhang, Z., Fluet-Chouinard, E., Jensen, K., McDonald, K., Hugelius, G., Gumbricht, T., Carroll, M., Prigent, C., Bartsch, A., and Poulter, B.: Development of the global dataset of Wetland Area and Dynamics for Methane Modeling (WAD2M), Earth Syst. Sci. Data, 13, 2001–2023, https://doi.org/10.5194/essd-13-2001-2021, 2021.

Zhou X., Zhang M., Krause S.M., Bu X., Gu X., Guo Z., Jia Z., Zhou X., Wang X., Chen X., Wang Y.: Soil aeration rather than methanotrophic community drives methane uptake under drought in a subtropical forest. Science of The Total Environment. Jun 5:148292, 2021.

Zona, D., Gioli, B., Commane, R., Lindaas, J., Wofsy, S. C., Miller, C. E., Dinardo, S. J., Dengel, S., Sweeney, C., Karion, A., Chang, R. Y.-W., Henderson, J. M., Murphy, P. C., Goodrich, J. P., Moreaux, V., Liljedahl, A., Watts, J. D., Kimball, J. S., Lipson, D. A., and Oechel, W. C.: Cold season emissions dominate the Arctic tundra methane budget, Proc. Natl. Acad. Sci. U. S. A., 113, 40–45, https://doi.org/10.1073/pnas.1516017113, 2016.