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[3], Jordan P. Goodrich[7], Henry W. Loescher[8,9], Gregory Starr[10], Ruth K. Varner[3,11]

[1] Institute of the Environment & Department of Biological Sciences, Florida International University, 11200 S.W. 8th Street, Miami, FL 33199, USA
[2] Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309, United States
[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] School of Science, University of Waikato, Gate 1 Knighton Rd, Hillcrest 3240, Hamilton, New Zealand
[8] Battelle, National Ecological Observatory Network (NEON), Boulder, CO USA 80301
[9] Institute of Alpine and Arctic Research, University of Colorado, Boulder, CO USA 80301
[10] Department of Biological Sciences, University of Alabama, Tuscaloosa, AL. 35487, USA
[11] 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 methane (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 anthropogenic emissions. To understand CH_4 flux from biogenic sources 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 landscape representativeness of the current infrastructure. 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 that inform policy decisions and global climate agreements. Using dissimilarity, multidimensional scaling, and cluster analysis, the US was divided into 10 clusters distributed across temperature and precipitation gradients. We evaluated dissimilarity within each cluster for research sites with active CH_4 EC towers to identify gaps in existing infrastructure that limit our ability to constrain the contribution of US biogenic CH_4 emissions to the global budget. Through our analysis using climate, land cover, and location variables, we identified priority areas for research infrastructure to provide a more complete understanding of the CH_4 flux potential of ecosystem types across the US. Clusters corresponding to Alaska and the Rocky Mountains, that are inherently difficult to capture, are the most poorly represented and all clusters require a greater representation of vegetation types.

Style Definition ... [23]
Style Definition ... [22]
Style Definition ... [21]
Style Definition ... [20]
Style Definition ... [19]
Style Definition ... [18]
Style Definition ... [17]
Style Definition ... [16]
Style Definition ... [15]
Style Definition ... [14]
Style Definition ... [13]
Style Definition ... [12]
Style Definition ... [11]
Style Definition ... [10]
Style Definition ... [9]
Style Definition ... [8]
Style Definition ... [7]
Style Definition ... [6]
Style Definition ... [5]
Style Definition ... [4]
Style Definition ... [3]
Style Definition ... [2]
Style Definition ... [1]
Formatted ... [24]
Deleted: Contosa[7]…ontosta[3], Jordan P. Goodrich[8]…oodr ... [25]
Deleted: [1]Institute··· Institute of the Environment & Depar ... [26]
Deleted: [2]NOAA Global Monitoring Laboratory, Boulder ... [27]
Moved down [1]: -COR Biosciences, 4421 Superior St.,
Deleted: [5]The
Moved down [2]: Robert B. Daugherty Water for Food Global
Deleted: [6]Department of Earth &…esearch in Environme ... [28]
Deleted: [7]Earth
Moved (insertion) [1]
Moved (insertion) [2]
Deleted: [8]School
Deleted: [9]
Deleted: [10]
Deleted: [11]Department
Deleted: [12]Department
Deleted: total …nthropogenic emissions. The largest sour ... [29]
Formatted ... [30]

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 carbon dioxide (CO_2) and nitrous oxide (N_2O), atmospheric methane (CH_4) concentrations show dynamic trends with a rapid increase of ~10 ppb yr^{-1} since 2014 (Nisbet et al., 2019). 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., 2019) 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., 2019) 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 in predicting and mitigating future climate change.

Quantifying the national CH_4 budget is important for assessing realistic pathways to mitigate climate change, yet uncertainties in the magnitude, size, and location of sources and sinks limit budget development (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., 2019). 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., 2019), are not well constrained. Finally, both the vast areas with relatively small uptake and emission rates (e.g., deserts, grasslands, forests) and the lake-ocean water continuum that transports CH_4 (e.g., fens, streams, and rivers) have been largely understudied but could contribute significantly to regional and global budgets (Hutchins et al., 2019; Rosentreter et al., 2021; Zhou et al., 2021). These unknowns 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 land cover change (Kirschke et al., 2013; Knox et al., 2019; Saunois et al., 2020).

To understand the biogenic CH_4 flux potential of 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. When scaling bottom-up measurements to the landscape and regional scale, measurements of CH_4 from existing infrastructure 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 is 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 biogenic

Deleted: ,

Deleted: (Nisbet et al., 2019a).

Deleted: (Dlugokencky, 2021),

Deleted: (Le Quéré et al., 2021).

Deleted: (Nisbet et al., 2019a)

Deleted: (IPCC, 2021).

Deleted: (Nisbet et al., 2019a)

Deleted: to both

Deleted: 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).

Deleted: (IPCC, 2013; Nisbet et al., 2019b).

Deleted: (Saunois et al., 2020).

Deleted: (Michalak et al., 2009; Kirschke et al., 2013; Knox et al., 2019; Nisbet et al., 2019b),

Deleted: very

Deleted: uptakes

Deleted: emissions

Deleted: , water bodies

Deleted: well transport of CH_4 by

Deleted: stems

Deleted: (Hutchins et al, 2019; Rosentreter et al, 2021; Zhou et al, 2021).

Deleted: uncertainties

Deleted: (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).

Deleted: 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 urbanization, eutrophication and positive climate feedbacks, yet current wetland classifications for land use data products are not suitable to capture these effects.

Moved down [3]: 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

Deleted: (Kirschke et al., 2013; Barkley et al., 2017; Knox et al., 2019). For wetlands these include (i) delineation of wetland ... [31]

Deleted: are

Deleted: (IPCC, 2013; Knox et al., 2019; Thornton et al., 2016)

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) and Arndt et al., (2019) for notable exceptions).

There is a pressing need to assess the capacity of existing infrastructure for current and future applications (Lovett et al., 2007; Kumar et al., 2016; Jongman et al., 2017; Novick et al., 2018; Villarreal et al., 2018; Chu et al., 2021). The representativeness of research infrastructure is often described in terms of the extent to which the measurements collected at any given location and time represent the conditions at any other location and time, and this is often driven by ecological and climatic conditions (Sulkava et al., 2011; Chu et al., 2021). Representativeness is also measured across landscapes, and studies have evaluated how well tower infrastructure captures the variability observed at specific sites (Chu et al., 2021). These approaches seek to understand the representativeness of the measurements for a broader landscape, which is critical for upscaling point measurements to regional and global scales. These types of assessments inform the scientific community on how to increase their utility and are often designed to support network design, upscaling, and bias estimation (Chen et al., 2011; Ciais et al., 2014; Jongman et al., 2017; Schimel and Keller, 2015; Villarreal et al., 2018; Kumar et al., 2016). There have been many attempts to assess the representativeness of existing EC tower networks for various purposes. To date, no study has focused on CH_4 infrastructure across the US, though many studies have used clustering and ecoregions (Sulkava et al., 2011; Hargrove and Hoffman, 2003), dissimilarity (Yang et al., 2008), and distance measures (Hargrove and Hoffman, 2003; Yang et al., 2008; He et al., 2015; Hoffman et al., 2013) on climatic (Novick et al., 2018) and vegetation type structure and function (Chu et al., 2021) to measure the representativeness of existing research infrastructure. The primary goal of this work is to fill this key knowledge gap by determining the regions where biogenic CH_4 infrastructure is needed within the US in order to constrain both the national and global CH_4 budget.

To determine key regions where biogenic CH_4 infrastructure is needed within the US we statistically identify gaps in active research infrastructure and evaluate areas where infrastructure can be augmented to include new CH_4 measurements. We use a combination of climate data and dominant land cover to guide the scientific community on how we can develop a distributed observational network for the US by leveraging existing infrastructure. While this analysis does not capture the heterogeneity of the conditions that drive CH_4 fluxes at the ecosystem scale, it is designed to evaluate the sampling intensity of research sites at the landscape scale. This coarse resolution influences the capacity to scale ecosystem-level results to the landscape, regional, and national level, which is required for the development of CH_4 budgets and emission reduction strategies.

Deleted: Groffman et al., 2006; Arndt et al., 2019

Deleted: 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).

Moved down [4]: In temperate areas, higher climate-change-induced variability in precipitation (e.g., higher moisture of upland forested soils, prolonged droughts, etc.)

Deleted: 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 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. ... [32]

2 Methods

2.1 Overview

To determine the gaps in physical research infrastructure for ecosystem-scale CH_4 fluxes, we need to understand how the current infrastructure is distributed across the US. We focus here on EC flux towers given their capabilities for continuous measurements and use in upscaling flux estimates and are therefore a useful basis for identifying gaps in the current network of CH_4 observations. The 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). In 2012, the US Department of Energy established the AmeriFlux Management Project (AMP) at Lawrence Berkeley 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 in freshwater, coastal, upland, natural, and managed ecosystems. Although we have information on the location of existing EC tower infrastructure that is a part of AmeriFlux (n=223), NEON (n=47), and known, independent PI-managed sites (n=141). We focus this analysis on the towers measuring CH_4 (n=100) and we distinguish between towers providing data to AmeriFlux (yes = 49, no = 51) and tower activity (active = 70; inactive = 30). We understand that additional towers exist within the US, but because these towers are not reporting or providing data to the flux community, we cannot include them in this analysis.

To understand the landscape representativeness across geographic clusters, we measured dissimilarity based on climate and land cover type, as these two factors together are characteristic of regional resource availability and disturbance regimes. First, we developed a dissimilarity matrix that was condensed down to a two-dimensional ordination to determine regional clusters and calculate cluster dissimilarity for each location within a cluster (Figure 1). It is important to note that a tower should be representative of 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, the landscape representativeness analysis done here uses a coarser classification of land cover types that are more emblematic of resource availability, and factors that influence how ecosystems function, not the specific ecosystem type where the tower is situated. Chu et al. (2021) examined the land-cover composition and vegetation characteristics of 214 AmeriFlux tower site footprints. They found that most sites do not represent the dominant land-cover type of the ecosystems they exist within, and when paired with common model-data integration approaches this mismatch introduces biases on the order of 4%–20% for EVI and 6%–20% for the dominant land cover percentage (Chu et al., 2021), making it essential to consider landscape characteristics in the design and evaluation of network infrastructure. Infrastructure

Formatted: Heading 2
Deleted: what
Deleted: tower
Deleted: capturing.
Deleted: because they are essential
Deleted: a bottom-up framework that bridges the gap between point-based chamber
Deleted: airborne platforms
Deleted: US
Deleted: AmeriFlux also includes tower sites from the Long-Term Ecological Research Network (LTER).
Deleted: (81 towers)
Deleted: In this evaluation
Deleted: included
Deleted: towers
Deleted: are
Deleted: 200
Deleted: LTER (23),
Deleted: 141). In addition
Deleted: our analysis of EC sites, we also evaluated the distribution of 161 network aquatic sites across the US. The networks included aquatic sites in
Deleted: , NEON, LTER and the Global Lake Ecological Observatory Network (GLEON). While all
Deleted: towers included in the
Deleted: measure CO_2 fluxes, only a handful currently quantify CH_4.
Deleted: of the current tower infrastructure, we identified variation in
Deleted: ecosystem
Deleted: typically
Deleted: is
Deleted: just
Deleted: (Desai, 2010; Jung et al., 2011; Xiao et al., 2012; Chu et al., 2021);
Deleted: this
Deleted: ecosystem
Deleted: is
Deleted: regional disturbance regimes and

representativeness at the landscape scale is indicative of the capacity to upscale information by climate and the dominant ecosystems of locations within a landscape.

525

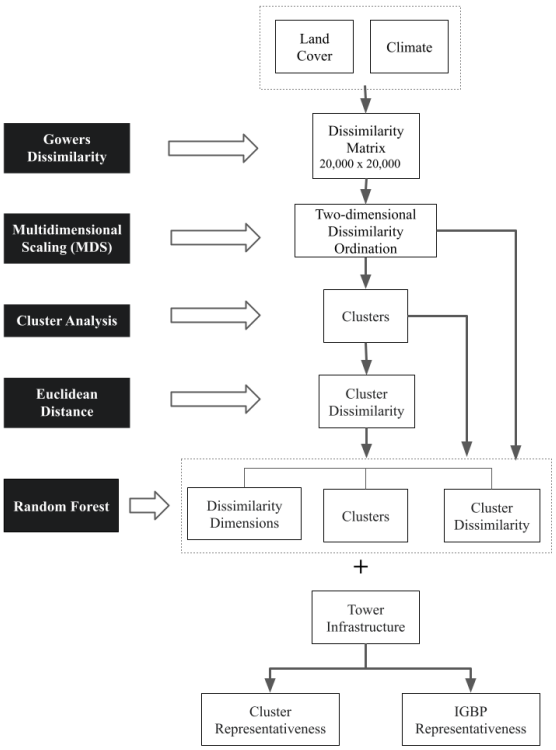


Figure 1: To determine the gaps in physical research infrastructure for CH_4 fluxes we measured landscape cover and climate dissimilarity across the US and evaluated the current distribution of CH_4 tower infrastructure.

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 land cover types (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 land cover types (Table 1). The crop land cover type 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 temperatures, 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). Land cover was 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 land cover types 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. These coarse layers are essential for considering the landscape and large-scale climate effects that can influence how ecosystems within landscapes function. While the available land cover information is appropriate for the coarse, landscape-scale analysis done here, it is important to note that the products used here are not designed to estimate the potential CH_4 source/sink status, particularly from the aquatic, wetland, and agricultural land cover types.

Formatted: Heading 2

Deleted: (Jin et al., 2019).

Deleted: ecotypes

Deleted: Land Cover

Deleted: (Sulla-Menashe and Friedl, 2018),

Deleted: ecotypes

Deleted: ecotype

Deleted: (Boryan et al., 2011),

Deleted:) (Wilen and Bates, 1995).

Deleted: (Thornton et al., 2017).

Deleted: temperature

Deleted: (Bessembinder et al., 2021). All spatial layers (ecotype and climate) were

Deleted: (R Core Team, 2021)

Deleted: (Hijmans, 2021).

Deleted: ecotypes

Formatted: Font: 9 pt, Bold

Table 1: Land cover and data sources. 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 land cover types with additional classes added for types with expected CH_4 source potential.

Land Cover	Expanded Land Cover	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 Measuring landscape dissimilarity across clusters within the US.

Climate, land cover, and location (latitude/longitude) were used in a multivariate distance analysis (Venables and Ripley, 2002; Ripley, 2007; Cox and Cox, 2008) to measure the dissimilarity across the US (all 50 states & Puerto Rico) at the landscape scale and divide it into ecological clusters. The purpose of this is to identify the interrelatedness of ecological components within a landscape (Ippoliti et al., 2019). We included location (latitude/longitude) to incorporate the interaction between climate, land cover, and most importantly, seasonality. The US was subsampled because of limitations in the maximum number of points that can be evaluated in the cluster analysis. To measure dissimilarity, we first randomly sampled (n = 20,000 1-km pixels) the US, maintaining the distribution of land cover and climate to define dissimilarity between observations. Although there were more than 8 million 1-km pixels available for the US, there are limits to the number of samples that can be analyzed by the functions used for the multidimensional scaling (MDS) analysis. We first developed a dissimilarity matrix by calculating Gower dissimilarity (Gower, 1971; Huang, 1997; Podani, 1999; Ahmad and Dey, 2007; Harikumar and Pv, 2015) using the function distmix from the package *kmed* in R. We used Gower dissimilarity because it can handle mixed data types. For each variable type in the data set, the dissimilarity metric that works well for that type is used and scaled to fall between 0 and 1. Then, a linear combination featuring user-specified weights (most simply an average) is calculated to create the final dissimilarity matrix. This approach measures the dissimilarity for each location within the US using land cover, climate, and

Deleted: Ecotype
Deleted: source used in analysis
Deleted: ecotypes
Deleted: Ecotype
Formatted: Font: Times New Roman
Formatted Table
Formatted: Font: Times New Roman
Deleted: Ecotypes
Formatted: Font: Times New Roman, Font color: Auto
Formatted: Font: Times New Roman
Formatted: Font: Times New Roman
Formatted: Font: Times New Roman
Formatted: Font: Times New Roman
Formatted: Font: Times New Roman
Formatted: Font: Times New Roman
Formatted: Font: Times New Roman
Formatted: Font: Times New Roman
Formatted: Caption
Deleted: Defining the state space of
Deleted: ecotype
Deleted: (Venables and Ripley, 2002; Ripley, 2007; Cox and Cox, 2008) to define the state space of the US
Deleted: using information that is important for capturing continental patterns in CH_4.
Deleted: analysis
Deleted: all
Deleted: —biotic, abiotic, terrestrial, and aquatic
Deleted: dynamic landscape (Ippoliti et al., 2019).
Deleted: ecotypes
Deleted: We
Deleted: to condense the ecotype, climate, and location information (ecotype…
Formatted: Highlight
Formatted: Font color: Custom Color(RGB(28,29,30)), Highlight
Formatted: Highlight

location information (land cover, five climate variables, and location) and creates a dissimilarity matrix (20,000 x 20,000) that indicates dissimilarity for a location to every other location in the US.

Once we created the dissimilarity matrix, we used MDS to generate a two-dimensional ordination showing landscape dissimilarity with the *MASS* package in R (Venables and Ripley, 2002). The MDS makes it possible to evaluate dissimilarity in two dimensions, which is essential to our goal to evaluate representativeness. We used the Kruskal method of non-metric scaling with the IsoMDS function in the *MASS* package (Venables and Ripley, 2002). IsoMDS works best when applied to metric variables (Torgerson, 1958). Torgerson (1958) initially developed this method, which assumes that the data obey distance axioms. It uses eigendecomposition of the dissimilarity to identify major components and axes, and represents any point as a linear combination of dimensions. This is very similar to principal component analysis (PCA) or factor analysis, but it uses the dissimilarity matrix rather than a correlation matrix as input. Furthermore, the included dimensions are the most important dimensions produced, like PCA which is able to identify all of the dimensions that exist in the original data up to N−1, but will retain only the most important ones.

Knowing that regional patterns in climate and land cover will be important for scaling CH_4 to the regional and national scale, we divided the US into clusters to evaluate representativeness using the first and second dimension from the MDS. This cluster analysis also allows us to summarize our results within a geographical context, an approach that has been used to delineate spatial sampling domains, to assess the spatial representativeness of networks, and to suggest arrangements of study sites (Sulkava et al., 2011; Kumar et al., 2016). We used cluster analysis to determine the optimal number of clusters using the library *cluster* and the function pam in R (Reynolds et al., 2006; Schubert and Rousseeuw, 2019, 2021). This approach uses the k-medoids algorithm, which partitions data set into k groups or clusters and is a robust alternative to k-means clustering (Kaufman and Rousseeuw, 2009). Each cluster is represented by one of the data points in the cluster named the cluster medoid. The medioid has the lowest average dissimilarity between it and all other objects in the cluster. The medoid can be considered a representative example of the members of that cluster. The k-medoid algorithm is less sensitive to noise and outliers, compared to k-means, because it uses medoids as cluster centers. The k-medoids algorithm requires the user to specify k, the number of clusters to be generated. A useful approach to determine the optimal number of clusters is the silhouette method. We fit an increasing number of clusters from 2 to 50 to construct a silhouette plot and choose the number of clusters that maximized the average silhouette width. Once we determined the number of clusters and the medoid of the cluster, we calculated the dissimilarity between every location within the cluster to the medoid to create a measure of how different each location was from the medoid condition of each cluster. We utilized the pointDistance function in the *raster* package, which provided a unit-less relative measure of dissimilarity that was determined by measuring the difference between the first and second dimensions produced by the isoMDS of each point in a cluster to the dimensions of the medoid. This analysis was repeated 10 times to ensure that the 20,000-pixel subsample would produce similar results in the dimensions and clustering. For simplicity, we show the results of the first analysis.

Deleted: down to two dimensions using

Formatted: Font color: Gray-75%, Pattern: Clear (Custom Color(RGB(252,252,252)))

Deleted: (Venables and Ripley, 2002).

Deleted: subsampled the US (n = 20,000 1-km pixels) randomly, maintaining

Deleted: distribution

Formatted: Font color: Custom Color(RGB(35,38,41)), Highlight

Deleted: ecotypes and climate for

Formatted: Font color: Custom Color(RGB(35,38,41)), Highlight

Deleted: MDS

Formatted: Pattern: Clear (Custom Color(RGB(252,252,252)))

Deleted: . We measured the

Deleted: between the ecotypes,

Deleted: layers, and locations (latitude/longitude) using the *envfit* function in the library *vegan* in R (Oksanen, 2016). This was followed by a

Deleted: in R

Deleted: around medoids (PAM algorithm), using the Gower dissimilarity matrix (Gower, 1971; Huang, 1997; Podani, 1999; Ahmad and Dey, 2007; Harikumar and Pv, 2015).

Deleted: 20

Deleted: maximizes

Deleted: 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.

Deleted: measured

Deleted: difference

Deleted: understand the distance, or

Deleted: representative

Deleted: This approach provides

Deleted: 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

Deleted: to understand how representative any location within a cluster is to the medoid condition

Formatted: Font color: Custom Color(RGB(32,33,36))

To extrapolate the cluster and dissimilarity layers across the entire US beyond the 20,000-pixel subsample, we employed the machine learning algorithm Random Forest (RF) with the package *randomForest* (Liaw and Wiener, 2002) to model the first and second dimensions using the land cover 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 dissimilarity layer beyond the 20,000 sample points that were used in the MDS analysis. The RF algorithm was first introduced by (Breiman, 2001), and uses an ensemble of regression trees to predict target values. In RF, a series of bootstrapped datasets are used to generate independent regression trees; at each node, a random sample of predictor variables is selected for use. The RF prediction is the ensemble of multiple individual trees. We created 500 trees for each year and site, using 80% of the data for model fitting and 20% for model validation. The fit of each RF was evaluated with the out-of-bag mean square error (OOB MSE), and variable importance was computed as the amount of the prediction error increased when a particular predictor was permuted. Initially, 500 RF trees were generated. Overall model fit was evaluated with the average of the 500 OOB MSEs from the final model for each year and site, and variable importance was calculated as the average rank of each predictor variable for the 500 models. This approach allowed us to measure the importance of the original data on the first and second dimensions defined by the MDS and how the MDS leads to cluster and dissimilarity patterns. This step was essential to producing a spatially explicit cluster and dissimilarity layers for the entire US, since the MDS analysis limits the number of observations that can be analysed.

2.4. Measuring the landscape representativeness of research infrastructure.

Representativeness studies discern when, where, and at what frequency networks are measuring ecological processes (Baldocchi et al., 2012; Jongman et al., 2017; Vaughan et al., 2001; Villarreal et al., 2018). To understand the representativeness of current CH_4 infrastructure, we defined clusters (Sulkava et al., 2011) and measured the dissimilarity between each location in a cluster to the medoid. We extracted the cluster and dissimilarity for all active tower sites measuring CH_4 that were distributed across the US and measured the tower cluster representativeness ($TR_{cluster}$) as the percent overlap between the range of dissimilarity sampled by the infrastructure ($r_{cluster}$) divided by the range of dissimilarity observed in the entire cluster (r; Eq. 1).

$$TR_{cluster} = (r_{cluster} / r)*100 \quad \text{Eq. 1}$$

We recognize that it is essential to capture the distribution of dissimilarity across an entire cluster to upscale ecosystem measurements. We also report the sampling intensity of the major ecosystem types within the cluster and report the ecosystem representativeness (TR_{IGBP}) by the IGBP vegetation types of the towers (Eq. 2).

$$TR_{IGBP} = (r_{IGBP} / r)*100 \quad \text{Eq. 2}$$

Deleted: distance
Deleted: fit a
Formatted: Font: 14 pt
Deleted: model
Deleted: (Liaw and Wiener, 2002)
Deleted: MDS dimension
Deleted: ecotype
Deleted: distance to medoid
Deleted: 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,
Deleted: distances from
Deleted: medoid. This approach
Deleted: defining
Deleted: medoid of a
Deleted: measuring the distance to
Deleted: medoid for each location in space allows us to understand
Deleted: of
Deleted: location to
Deleted: rest of the cluster.
Formatted: Font: Italic
Deleted: ¶
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. ¶

This approach allows the evaluation of representativeness that is not based on a specific research site, but on the dissimilarity of a location to other locations in the landscape and we use the range, to indicate a capacity to scale within a cluster which is based on both the effects of landscape dominant land cover, climate, and the specific ecosystems measured (IGBP).

3 Results

3.1 Measuring landscape dissimilarity across clusters within the US

Land cover, climate, and location were condensed down to two dissimilarity dimensions (Figure 2a). Both climatic factors and location were the most important variables for determining dimensions and explained 99% of the variance in dimensions (Figure S1). Using the first and second dimensions, the US was divided into 10 clusters (Figure S2) that were distributed across temperature and wetness gradients (Figure 2; Table 2). The coldest zones were in Alaska and included clusters Na and Nb. Cool to temperate clusters in the midwestern and western US include 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) were in the south-eastern US and Hawaii. Individual clusters represented 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.

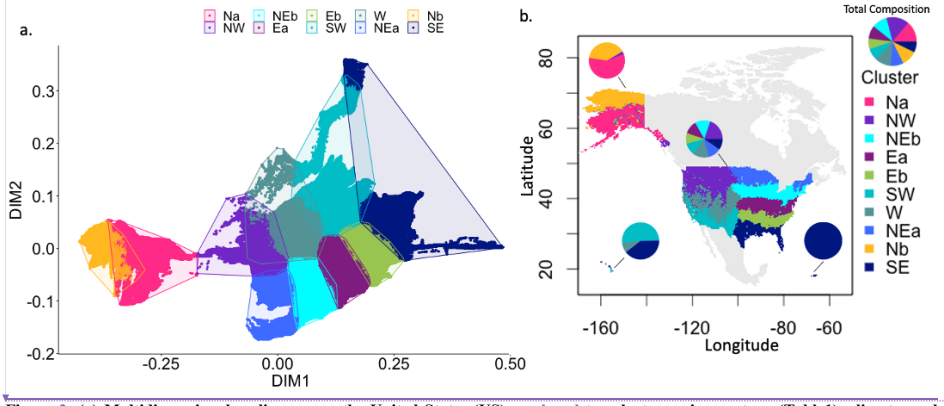


Figure 2: (a) Multidimensional scaling across the United States (US) produced ten clusters using ecotype (Table1), climate, and location (lat/long). (b) Spatial distribution of the identified clusters.

Across all clusters

Moved up [5]: 3 Results

Moved (insertion) [5]

Deleted: and Discussion¶

Deleted: Defining the state space of the US¶
The US was divided into 10

Deleted: 1

Deleted: 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.

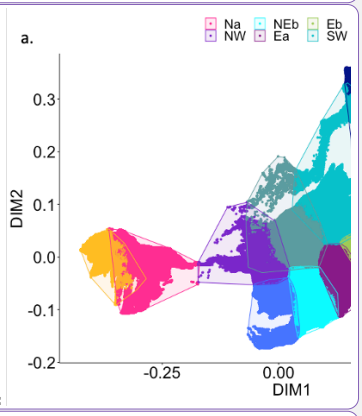
Deleted: clusters

Deleted: are

Deleted: represent

Deleted: 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).

Formatted: Normal



Deleted:

Deleted: 1

Deleted: into

Deleted: ¶

Page Break

¶

¶

Table 2: Ecotype, climate, and number of towers of 10

, dissimilarity ranged from 0.01 to 0.33 (Figure 3). The mean dissimilarity 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 than average dissimilarity in their respective clusters.

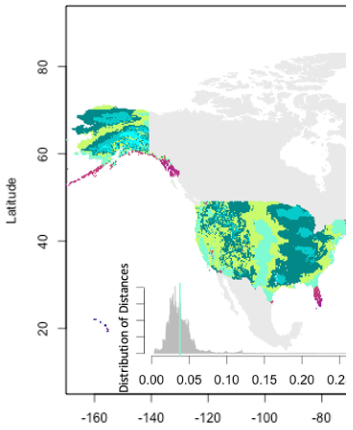
Deleted: distance

Deleted: distances to the medoids.

Dominant landscape land cover types also varied across clusters, with forests, scrub, and herbaceous ecosystems dominating clusters (> 20% coverage; Table 2). Although irrigated croplands did not have high coverage rates across any cluster, non-irrigated croplands had high coverage rates in NEb, Ea, and NEb. Wetlands did not have high coverage rates in any cluster.

Table 2: The land cover and climate of the 10 clusters in the US. Crops were divided into irrigated ($Crop_I$) and non-irrigated ($Crop_{NI}$) and wetlands into emergent coastal (Wet_{EC}), emergent freshwater (Wet_{EF}), and freshwater forest (Wet_{FF}). Percent coverage (% Cov) is the percent area occupied by a cluster and R is the range in dissimilarity for each cluster.

Cluster	Dominant Landscape Land cover (% Cov)								Climate	% Cov	R
	Forest	Scrub	Herb	$Crop_I$	$Crop_{NI}$	Wet_{EC}	We_{EF}	Wet_{FF}			
Na	27.4	39.0	3.8	0	0	0.1	1.6	1.9	Cold-cool (Dry)	11	0.13
NW	28.4	33.2	23.0	0	6.5	0.0	0.6	0.2	Cool-Temperate (Mild)	16	0.11
NEb	24.4	0.4	9.3	0	34.2	0.0	0.7	1.9	Temperate (Mild-Wet)	10	0.06
Ea	39.0	0.9	7.7	0.0	24.3	0.3	0.2	1.4	Temperate (Wet)	9	0.04
Eb	37.8	3.8	9.9	0.9	14.0	0.2	0.3	6.7	Warm (Wet)	8	0.04
SW	2.6	58.8	17.3	0.2	7.4	0.0	0.2	0.2	Warm (dry)	9	0.25
W	19.4	42.6	20.8	0	6.9	0.0	0.4	0.1	Cool-Temperate (Dry)	12	0.07
NEa	27.2	1.5	6.1	0	23.5	0.0	2.1	7.6	Cool-Temperate (Mild-Wet)	9	0.08
Nb	9.0	51.7	16.0	0	0	0.1	2.0	0.3	Cold (Dry)	7	0.11
SE	19.4	19.3	8.2	0.6	7.8	1.6	2.2	7.9	Hot (Wet)	9	0.31



Deleted:

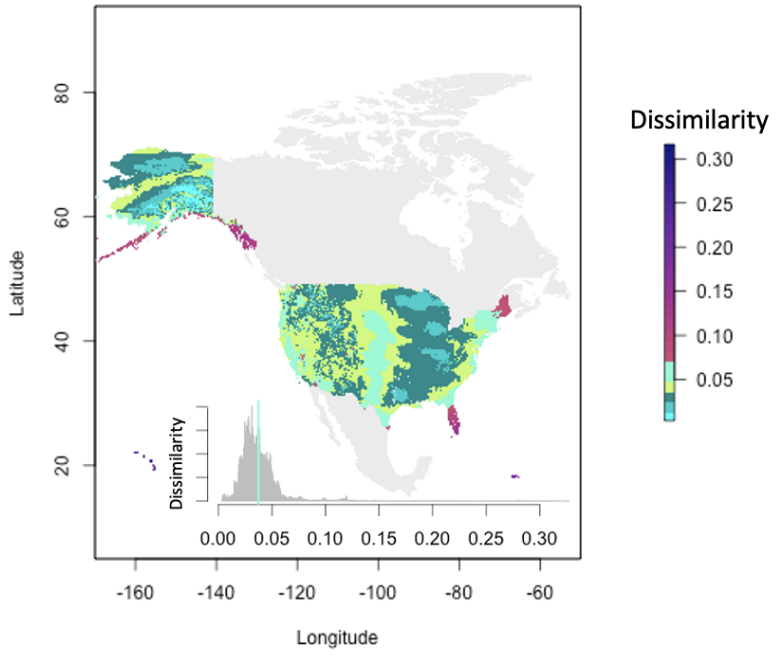


Figure 3: Cluster dissimilarity for the US. Inset: the distributions of dissimilarity across all clusters shown in a histogram, in which the line denotes the mean dissimilarity across all clusters.

3.2 Landscape Representativeness of Existing CH_4 Tower Infrastructure

There were 70 active EC towers measuring CH_4 distributed across forest (3 towers), grasslands (4 towers), shrublands (1 tower), agriculture (19 towers), wetlands (37 towers), barren (2 towers), and aquatic (4 towers) IGBP vegetation classes. Less than half of the active towers (43%) were providing data to the community through Ameriflux, limiting the development of CH_4-derived products. For this reason, we will first focus this analysis on the active towers providing data to AmeriFlux. Although CH_4 EC tower infrastructure was not a part of a single organized network designed to be representative of the climate, landscape, and dominant IGBP vegetation classes that exist within the US, EC tower infrastructure that was providing data to AmeriFlux was distributed across 8 of the 10 clusters (Table 3), with clusters NW and SE without any active towers providing data to the community. Tower representativeness ($TR_{cluster}$) of clusters ranged from 0 to 88%. The greatest $TR_{cluster}$ was for Eb and NEa and the lowest $TR_{cluster}$ was for NW and SE which had no towers. $TR_{cluster}$ was low (<50%) for most clusters and high coverage was not associated with a higher frequency of towers. A high $TR_{cluster}$ was found in clusters where towers were

Deleted: . Distance to the cluster medoid represented spatially.
Deleted: distances
Deleted: distance
Deleted: ¶
Deleted: Tower Infrastructure and
Deleted: are currently 411 reported
Deleted: 136
Deleted: 105 towers), wetland (69
Deleted: scrub (93
Deleted: and herbaceous (92
Deleted:) ecotypes. Although most
Deleted: land cover
Deleted: 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
Deleted: Na, W, Eb, and Nb are

dispersed across IGBP vegetation classes and where towers in wetlands, forests, or the arctic tundra (barren) were distributed across the observed range in the dissimilarity of clusters. Most clusters were substantially under-sampled (Table 3, Figure 4) due to an insufficient number of towers measuring CH_4 and poor distribution across the cluster.

The representativeness of IGBP vegetation types within clusters was poor for all vegetation types, excluding forests in the NEa. TR_{IGBP} ranged from 0 to 79% and wetlands were the only IGBP class to be sampled across 8 clusters. Ideally, IGBP classes should be distributed both within and across clusters where the classes exist. There was not a single cluster with towers in all of the IGBP classes (forest, scrub, aquatic ecosystems, crops, wetlands, barren tundra, and grasslands) that are found within that cluster.

Table 3: The total number of eddy covariance (EC) towers measuring CH_4 and providing data to Ameriflux. The tower frequency by dominant landscape type, the total cluster representativeness ($TR_{cluster}$), and cluster representativeness by major ecosystem types are shown (TR_{IGBP}). For $TR_{cluster}$ and TR_{IGBP} values of 0.01 were assigned where a single tower was present.

Cluster	EC CH_4 towers	Tower Frequency by Dominant Landscape Land Cover								$TR_{cluster}$ (%)	TR_{IGBP} (%)						
		Forest	Scrub	Herb	Crop	Wet	Urban	Barren	Water		Forest	Scrub	AQ	Crop	Wet	Barren	Grass
Na	4	2	1	-	-	1	-	-	-	3.0	0.01	0.01	-	-	0.02	-	-
NW	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
NEb	2	1	-	-	1	-	-	-	-	19.8	-	-	-	0.01	0.01	-	-
Ea	1	-	-	-	-	1	-	-	-	0.01	-	-	-	-	0.01	-	-
Eb	3	-	-	-	1	2	-	-	-	88	-	-	-	0.01	42.1	-	-
SW	7	-	-	1	3	2	-	1	-	2.0	-	-	-	0.14	2.0	-	0.01
W	1	-	-	-	-	-	-	-	1	0.01	-	-	-	-	0.01	-	-
NEa	7	4	-	-	-	3	-	-	-	79.3	79.3	-	0.02	-	13.4	-	-
Nb	8	-	2	4	-	2	-	-	-	21.3	-	-	0.01	-	21.3	6.3	0.01
SE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

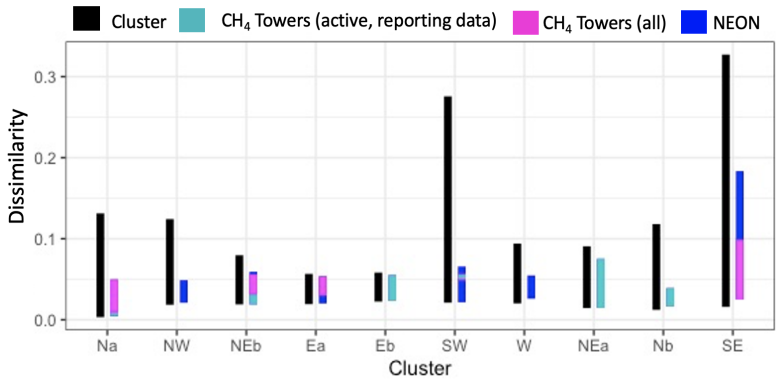


Figure 4: The range in dissimilarity for clusters (black bar), active CH_4 towers providing CH_4 data to Ameriflux (cyan), all active CH_4 towers (magenta), and for NEON towers (blue). The black lines show the range in dissimilarity observed for a cluster and greater overlap between the cluster range and the tower range is important for landscape representativeness.

Table 4: The $TR_{cluster}$ for CH_4 towers that are active and providing data to Ameriflux, the $TR_{cluster}$ for all active CH_4 towers, and the $TR_{cluster}$ for all active towers in addition to NEON towers.

Cluster	CH_4 Towers (Data Providing)	CH_4 Towers (All)	NEON Towers
Na	3.0	34.9	35.5
NW	-	0.1	26.3
NEb	19.8	60.6	65.9
Ea	0.01	63.1	89.4
Eb	88.1	88.1	88.1
SW	2.0	3.3	17.3
W	0.01	0.01	38.8
NEa	79.3	79.3	79.3
Nb	21.3	21.3	21.3
SE	-	23.6	50.8

There were important gains in the $TR_{cluster}$ when considering all CH_4 towers regardless of if they were providing data to Ameriflux (Table 4 and Figure 4). The clusters with substantial gains in representativeness (> 10%) include Na, NEb, Ea, and

Moved (insertion) [6]

Formatted: Normal

Deleted: 4c). The number of aquatic and cropland EC towers across clusters is also lacking. There

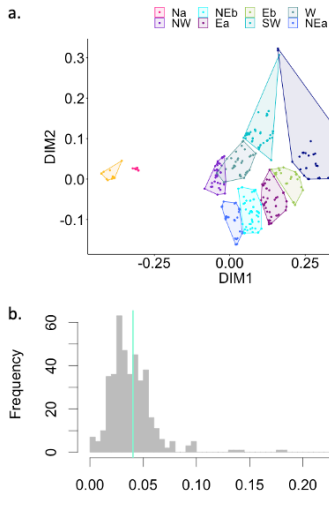
the SE. The TR $_{cluster}$ of the NW, Ea, SW, W, and the SE would be further enhanced by more than 10% with the addition of CH_4 instrumentation at NEON tower sites.

4 Discussion

To determine key regions where biogenic CH_4 infrastructure is needed within the US we identified gaps in active research infrastructure. We found that there is an insufficient number of towers measuring CH_4, and the distribution of these sites across the range in dissimilarity observed is poor for all clusters. Current EC towers measuring CH_4 are in ecosystems known to be sources of CH_4. This is extremely limiting when trying to upscale CH_4 fluxes because it leads to a serious bias towards CH_4 emissions in model results, and constrains our capacity to appropriately model ecosystems that are CH_4 sinks. In this analysis, we include NEON towers because they are purposefully distributed across climate zones and ecosystem types, they provide consistent and standardized measurements, and existing infrastructure at these sites could be quickly adapted to measure CH_4 and all data is made publicly available. We understand that for PI managed infrastructure the placement of towers is driven by the scientific question being asked and research funding priorities (Papale et al., 2015; Mahecha et al., 2017; Villarreal et al., 2018; Knox et al., 2019), but as the number of towers measuring CH_4 fluxes continues to grow, consideration for key underrepresented regions where towers are needed or where more efforts are needed for existing but nonreporting towers to contribute to AmeriFlux is of utmost importance. Making all data available must become the standard of the trace gas flux and biogeochemistry communities. Notable infrastructure gaps were in clusters Na, NW, SW, W, Nb, and the SE, and all clusters require a greater representation of IGBP vegetation types. Our analysis shows that the Na, W, and Nb clusters are the most poorly represented regions, corresponding to Alaska (Na & Nb) and the Rocky Mountains (W), where large elevational changes in the landscape are inherently difficult to capture.

One reason for gaps in CH_4 flux tower infrastructure may be the lag in technological capability behind that of CO_2 flux measurements. Methane 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; Nemitz et al., 2018; Burba et al., 2019; Burba, 2021). Therefore, as the flux tower infrastructure has expanded to measure CH_4, 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 suitable infrastructure (McDermitt et al., 2011), and not necessarily to address CH_4 hypotheses. 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 too remote and difficult to instrument. When they are instrumented, towers

Deleted: distance to the medoid is poor for all clusters.



Deleted:
Figure 4: (a) Multidimensional scaling dimensions of the ... [33]

Moved down [7]: ¶
Gaps in

Deleted: 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 ... [34]

Deleted: fluxes. CH_4

Deleted: (Shurpali et al., 1993; Billesbach et al., 1998; Rinne et al., 2007),

Deleted: (Peltola et al., 2013; Niemitz et al, 2018; Burba et al, 2019, 2021).

Deleted: for measuring CH_4

Deleted: (Baldocchi, 2014)

Deleted: the grid power and

Deleted: (McDermitt et al, 2011).

Deleted: (Desai et al., 2015).

Deleted: Arctic

Deleted: (Wik et al., 2016; Elder et al., 2020),

Deleted: Finally, vast areas of land have been traditionally thought to have a negligible CH_4 emission or consumption rates, alth ... [35]

Deleted: some of the most poorly represented clusters, corresponding to Alaska (Na & Nb) and the Rocky Mountai ... [36]

are often clustered together regionally, resulting in incremental changes in landscape representativeness. A non-negligible portion of the existing CH_4 measurements, including both towers and chambers, are not placed where CH_4 sources or sinks are but where the grid power is available to run such measurements. The likely incomplete quantification of CH_4 fluxes within heterogeneous sites and the measurement of CH_4 fluxes at sites that were established to measure CO_2 and energy fluxes, together introduce an inherent source of site level bias in existing data and our analyses. Hence, we interpret our results as a best-case scenario, as this bias likely would reduce even further our reported degree of representativeness.

Gaps in our US infrastructure and current capability to measure CH_4 were most noted when considering only the AmeriFlux sites that provide CH_4 data. When evaluating all sites with CH_4 infrastructure with the addition of the measurement capability from NEON sites, there were great improvements in landscape representativeness. Still, the largest gaps in infrastructure capability to measure expected CH_4 sources were from aquatic sites. These gaps in representation have been noted in other investigations of CH_4 flux and budget studies, as a part of larger global CH_4 analyses (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 advocated as necessary to reduce the uncertainty in CH_4 budget estimates from natural ecosystems (Peltola et al., 2019), which is among the largest uncertainty in the global CH_4 budget (Saunois et al., 2020). Even areas that have been traditionally thought to have negligible CH_4 emission or consumption rates should be monitored because their contribution to CH_4 budgets may be significant when considering their large spatial extent. There is also a strong need for a continental CH_4 observatory to aid in reducing these uncertainties in the natural CH_4 sources and sinks.

A large 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). Specific ecosystems within the landscape can contribute significantly to landscape-level and regional CH_4 source/sink estimates. Aquatic emissions are likely to change in the future due to an increase in urbanization, eutrophication, and positive climate feedbacks (IPCC, 2021). Yet current wetland classifications from land use data products are not suitable to capture these potential changes, or the potential feedbacks they may have on CH_4 processes. 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,

Deleted: is positioned

Deleted: 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

Deleted: individual

Deleted:

Deleted: .

Moved (insertion) [7]

Moved (insertion) [3]

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 potential large sources of uncertainties in biogenic CH_4 flux estimates cannot be addressed with the land cover maps currently used to scale CH_4 fluxes and the existing distribution of CH_4 observation sites (Rosentreter et al., 2021). Hence, if we are to build a US CH_4 budget using a scaled-up land use classification scheme (as is done for CO_2), we need both better representation of CH_4 measurement sites and better identification and quantification of the CH_4 source/sink potential of the land use classes themselves, i.e., specific development of land use classes based on CH_4 potential.

Ideally, CH_4 measurement infrastructure should have representation of all IGBP vegetation classes within and across clusters, where appropriate, and address the scale of spatial heterogeneity that reduces uncertainty in a national CH_4 budget with confidence limits that can inform both research objectives and mitigation policy. Thus, the incorporation of representative CH_4 sources and sink strength is essential to develop national CH_4 budgets. Neglecting sinks would further bias models that suggest sources occur where we are confident they do not. Advancing research and our process-level understanding of biogenic CH_4, we need to determine the measurement scales to assess the degree of spatial heterogeneity required to reduce uncertainty within and among sites. One means to address the within-site scale of spatial uncertainty is from automated chamber measurements within flux tower footprints, such as that found in soils, or first- and second-order streams. This would also allow the scientific community to determine the within site CH_4 source/sink strength from local (chamber; <1 m^2), ecosystem (EC flux tower; ~1 km^2) and landscape scales (tower concentrations; ~100s km^2). At even a larger scale, airborne observations of atmospheric CH_4 concentrations can be used to estimate boundary-level surface-atmosphere CH_4 fluxes and potentially provide greater spatial coverage than towers (Chang et al., 2014; Zona et al., 2016), and provide a mechanistic link between tower-based and satellite-derived CH_4 estimates.

The rate of global climate change re-enforces the urgency to establish a continental-scale CH_4 observatory network that can enable the first national CH_4 budget. As it stands, we currently do not know the scale, location, or the magnitude of site-based biogenic CH_4 source/sinks to estimate a national budget. For example, we lack the quantitative information about specific processes (particularly those that are stochastic, *e.g.*, temperature sensitivity, susceptibility to drought and flooding, tipping points, etc.) from representative ecosystems that would scale and inform a national CH_4 budget. 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 the 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,

Moved (insertion) [4]

Deleted: 3.3 Representativeness of Existing Aquatic Network Research Sites

anoxic areas, can in some cases increase CH_4 production and emission rates (Lu et al., 2018). Hence, we do not have a baseline US CH_4 budget to establish a starting point now and to compare to in the future, and as a baseline to estimate the efficacy of any mitigation decision (policy) made today. As scientists, we are often asked what is the most likely future state of an ecological system, and what is the most likely state of a system given if a decision or action is made today? The current state of CH_4 research and its ability to inform these questions are still nascent.

5

Conclusions

We used landscape dissimilarity to assess gaps in current CH_4 infrastructure at the landscape scale in the US. 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. This analysis complements previous studies based on climatic or vegetation characteristics (Hargrove and Hoffman, 2003; Yang et al., 2008; Villarreal et al., 2018), and identifies regions within the US where gaps are limiting the development of upscaling techniques. 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, dissimilarity coverage was lacking for clusters Na, W, and Nb, and currently clusters Na, W, Eb, and Nb are substantially undersampled. All aquatic sites were undersampled 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 to assess the impact of future policy 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

Moved up [6]: ¶
There were

Formatted: Heading 1

Deleted: distance to the medoid

Deleted: under sampled.

Deleted: are under sampled

Deleted: interventions

Deleted: https://knb.ecoinformatics.org

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 acknowledge 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 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. SLM acknowledges support provided by NSF grant #2047687. 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.

Arndt, K. A., Oechel, W. C., Goodrich, J. P., Bailey, B. A., Kalhori, A., Hashemi, J., Sweeney, C., and Zona, D.: Sensitivity of methane emissions to later soil freezing in arctic tundra ecosystems, J. Geophys. Res. Biogeosci., 124, 2595–2609, https://doi.org/10.1029/2019jg005242, 2019.

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.

Baldocchi, D., Reichstein, M., and Papale, D.: The role of trace gas flux networks in the biogeosciences, Eos Trans. Amer. Geophys. Union, 2012.

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

Deleted: ¶

Formatted: English (UK)

Deleted: acknowledged

Formatted: Font color: Black

Formatted: Heading 1

Deleted: 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¶

¶

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.: Meth ... [37]

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.

Breiman, L.: Random Forests, Mach. Learn., 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

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.: 9 - Atmospheric flux measurements, in: Advances in Spectroscopic Monitoring of the Atmosphere, edited by: Chen, W., Venables, D. S., and Sigrist, M. W., Elsevier, 443–520, https://doi.org/10.1016/B978-0-12-815014-6.00004-X, 2021.

Burba, G., Anderson, T., and Komissarov, A.: Accounting for spectroscopic effects in laser-based open-path eddy covariance flux measurements, Glob. Chang. Biol., 25, 2189–2202, https://doi.org/10.1111/gcb.14614, 2019.

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.

Chen, B., Coops, N. C., Fu, D., Margolis, H. A., Amiro, B. D., Barr, A. G., Black, T. A., Arain, M. A., Bourque, C. P.-A., Flanagan, L. B., Lafleur, P. M., McCaughey, J. H., and Wofsy, S. C.: Assessing eddy-covariance flux tower location bias across the Fluxnet-Canada Research Network based on remote sensing and footprint modelling, Agric. For. Meteorol., 151, 87–100, https://doi.org/10.1016/j.agrformet.2010.09.005, 2011.

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.

Ciais, P., Dolman, A. J., Bombelli, A., Duren, R., Peregon, A., Rayner, P. J., Miller, C., Gobron, N., Kinderman, G., Marland, G., Gruber, N., Chevallier, F., Andres, R. J., Balsamo, G., Bopp, L., Bréon, F.-M., Broquet, G., Dargaville, R., Battin, T. J., Borges, A., Bovensmann, H., Buchwitz, M., Butler, J., Canadell, J. G., Cook, R. B., DeFries, R., Engelen, R., Gurney, K. R., Heinze, C., Heimann, M., Held, A., Henry, M., Law, B., Luyssaert, S., Miller, J., Moriyama, T., Moulin, C., Myneni, R. B., Nussli, C., Obersteiner, M., Ojima, D., Pan, Y., Paris, J.-D., Piao, S. L., Poulter, B., Plummer, S., Quegan, S., Raymond, P., Reichstein, M., Rivier, L., Sabine, C., Schimel, D., Tarasova, O., Valentini, R., Wang, R., van der Werf, G., Wickland, D., Williams, M., and Zehner, C.: Current systematic carbon-cycle observations and the need for implementing a policy-relevant carbon observing system, Biogeosciences, 11, 3547–3602, https://doi.org/10.5194/bg-11-3547-2014, 2014.

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.

Groffman, P. M., Hardy, J. P., Driscoll, C. T., and Fahey, T. J.: Snow depth, soil freezing, and fluxes of carbon

dioxide, nitrous oxide and methane in a northern hardwood forest, Glob. Chang. Biol., 12, 1748–1760, https://doi.org/10.1111/j.1365-2486.2006.01194.x, 2006.

Hargrove, W. W. and Hoffman, F. M.: New analysis reveals representativeness of the AmeriFlux network, Eos Trans. Amer. Geophys. Union, 2003.

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.

He, H., Zhang, L., Gao, Y., Ren, X., Zhang, L., Yu, G., and Wang, S.: Regional representativeness assessment and improvement of eddy flux observations in China, Sci. Total Environ., 502, 688–698, https://doi.org/10.1016/j.scitotenv.2014.09.073, 2015.

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

Hoffman, F. M., Kumar, J., Mills, R. T., and Hargrove, W. W.: Representativeness-based sampling network design for the State of Alaska, https://doi.org/10.1007/s10980-013-9902-0, 2013.

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., and Trivedi, P.: Climate change microbiology - problems and perspectives, Nat. Rev. Microbiol., 17, 391–396, https://doi.org/10.1038/s41579-019-0178-5, 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.

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.

Jongman, R. H. G., Skidmore, A. K., Mücher, C. A. S., Bunce, R. G. H., and Metzger, M. J.: Global terrestrial ecosystem observations: why, where, what and how?, in: The GEO handbook on biodiversity observation networks, Springer, Cham, 19–38, 2017.

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.

Kaufman, L. and Rousseeuw, P. J.: Finding Groups in Data: An Introduction to Cluster Analysis, John Wiley & Sons, 342 pp., 2009.

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.

Kumar, J., Hoffman, F. M., Hargrove, W. W., and Collier, N.: Understanding the representativeness of FLUXNET for upscaling carbon flux from eddy covariance measurements, Earth Syst. Sci. Data Discuss., 1–25, https://doi.org/10.5194/essd-2016-36, 2016.

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.

Le Quéré, C., Peters, G. P., Friedlingstein, P., Andrew, R. M., Canadell, J. G., Davis, S. J., Jackson, R. B., and Jones, M. W.: Fossil CO2 emissions in the post-COVID-19 era, Nat. Clim. Chang., 11, 197–199, https://doi.org/10.1038/s41558-021-01001-0, 2021.

Liaw, A. and Wiener, M.: Classification and regression by randomForest, 2002.

Lovett, G. M., Burns, D. A., Driscoll, C. T., Jenkins, J. C., Mitchell, M. J., Rustad, L., Shanley, J. B., Likens, G. E., and Haeuber, R.: Who needs environmental monitoring?, Front. Ecol. Environ., 5, 253–260, https://doi.org/10.1890/1540-9295(2007)5[253:WNEM]2.0.CO;2, 2007.

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.

Mahecha, M. D., Gans, F., Sippel, S., Donges, J. F., Kaminski, T., Metzger, S., Migliavacca, M., Papale, D., Rammig, A., and Zscheischler, J.: Detecting impacts of extreme events with ecological in situ monitoring networks, Biogeosciences, 14, 4255–4277, https://doi.org/10.5194/bg-14-4255-2017, 2017.

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.

McDermitt, D., Burba, G., Xu, L., Anderson, T., Komissarov, A., Riensche, B., Schedlbauer, J., Starr, G., Zona, D., Oechel, W., Oberbauer, S., and Hastings, S.: A new low-power, open-path instrument for measuring methane flux by eddy covariance, Appl. Phys. B, 102, 391–405, https://doi.org/10.1007/s00340-010-4307-0, 2011.

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.

Nemitz, E., Mammarella, I., Ibrom, A., Aurela, M., Burba, G. G., Dengel, S., Gielen, B., Grelle, A., Heinesch, B., Herbst, M., Hörtnagl, L., Klemedtsson, L., Lindroth, A., Lohila, A., McDermitt, D. K., Meier, P., Merbold, L., Nelson, D., Nicolini, G., Nilsson, M. B., Peltola, O., Rinne, J., and Zahniser, M.: Standardisation of eddy-

covariance flux measurements of methane and nitrous oxide, International Agrophysics, 32, 517–549, https://doi.org/10.1515/intag-2017-0042, 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, 2019.

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.

Novick, K. A., Biederman, J. A., Desai, A. R., Litvak, M. E., Moore, D. J. P., Scott, R. L., and Torn, M. S.: The AmeriFlux network: A coalition of the willing, Agric. For. Meteorol., 249, 444–456, https://doi.org/10.1016/j.agrformet.2017.10.009, 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.

Papale, D., Black, T. A., Carvalhais, N., Cescatti, A., Chen, J., Jung, M., Kiely, G., Lasslop, G., Mahecha, M. D., Margolis, H., Merbold, L., Montagnani, L., Moors, E., Olesen, J. E., Reichstein, M., Tramontana, G., Gorsel, E., Wohlfahrt, G., and Ráduly, B.: Effect of spatial sampling from European flux towers for estimating carbon and water fluxes with artificial neural networks, J. Geophys. Res. Biogeosci., 120, 1941–1957, https://doi.org/10.1002/2015jg002997, 2015.

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.

Reynolds, A. P., Richards, G., de la Iglesia, B., and Rayward-Smith, V. J.: Clustering Rules: A Comparison of Partitioning and Hierarchical Clustering Algorithms, J. Math. Model. Algorithms, 5, 475–504, https://doi.org/10.1007/s10852-005-9022-1, 2006.

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, 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.

Schimel, D. and Keller, M.: Big questions, big science: meeting the challenges of global ecology, Oecologia, 177, 925–934, https://doi.org/10.1007/s00442-015-3236-3, 2015.

Schubert, E. and Rousseeuw, P. J.: Faster k-Medoids Clustering: Improving the PAM, CLARA, and CLARANS Algorithms, https://doi.org/10.1007/978-3-030-32047-8_16, 2019.

Schubert, E. and Rousseeuw, P. J.: Fast and eager k-medoids clustering: O(k) runtime improvement of the PAM, CLARA, and CLARANS algorithms, Inf. Syst., 101, 101804, https://doi.org/10.1016/j.is.2021.101804, 2021.

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.

Sulkava, M., Luyssaert, S., Zaehle, S., and Papale, D.: Assessing and improving the representativeness of monitoring networks: The European flux tower network example, J. Geophys. Res., 116,

https://doi.org/10.1029/2010jg001562, 2011.

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.

Torgerson, W. S.: Theory and methods of scaling, 460, 1958.

Vaughan, H., Brydges, T., Fenech, A., and Lumb, A.: Monitoring long-term ecological changes through the Ecological Monitoring and Assessment Network: science-based and policy relevant, Environ. Monit. Assess., 67, 3–28, https://doi.org/10.1023/a:1006423432114, 2001.

Venables, W. N. and Ripley, B. D.: Modern applied statistics with S, 2002.

Villarreal, S., Guevara, M., Alcaraz-Segura, D., Brunsell, N. A., Hayes, D., Loescher, H. W., and Vargas, R.: Ecosystem functional diversity and the representativeness of environmental networks across the conterminous United States, Agric. For. Meteorol., 262, 423–433, https://doi.org/10.1016/j.agrformet.2018.07.016, 2018.

Wik, M., Thornton, B. F., Bastviken, D., Uhlbäck, J., and Crill, P. M.: Biased sampling of methane release from northern lakes: A problem for extrapolation, Geophys. Res. Lett., 43, 1256–1262, https://doi.org/10.1002/2015gl066501, 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.

Yang, F., Zhu, A.-X., Ichii, K., White, M. A., Hashimoto, H., and Nemani, R. R.: Assessing the representativeness

of the AmeriFlux network using MODIS and GOES data, J. Geophys. Res., 113, https://doi.org/10.1029/2007jg000627, 2008.

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.

Zhou, X., Zhang, M., Krause, S. M. B., Bu, X., Gu, X., Guo, Z., Jia, Z., Zhou, X., Wang, X., Chen, X., and Wang, Y.: Soil aeration rather than methanotrophic community drives methane uptake under drought in a subtropical forest, Sci. Total Environ., 792, 148292, https://doi.org/10.1016/j.scitotenv.2021.148292, 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.

Formatted: Caption

Page 1: [1] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Comment Subject

Page 1: [2] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Normal (Web)

Page 1: [3] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Comment Text: English (US), Left

Page 1: [4] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Authors

Page 1: [5] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Correspondence

Page 1: [6] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Footer

Page 1: [7] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Caption

Page 1: [8] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Equation

Page 1: [9] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Balloon Text

Page 1: [10] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Affiliation

Page 1: [11] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

List Paragraph

Page 1: [12] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

MS title

Page 1: [13] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Copernicus_Word_template

Page 1: [14] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Name

Page 1: [15] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Kontakt

Page 1: [16] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Header

Page 1: [18] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Betreff

Page 1: [19] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Heading 4

Page 1: [20] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Heading 3

Page 1: [21] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Heading 2

Page 1: [22] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Heading 1

Page 1: [23] Style Definition **Malone et al** **1/24/22 10:59:00 AM**

Normal: Font: 10 pt, English (UK), Justified, Line spacing: 1.5 lines

Page 1: [24] Formatted **Malone et al** **1/24/22 10:59:00 AM**

MS title

Page 1: [25] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 1: [25] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 1: [25] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 1: [25] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 1: [25] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 1: [26] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 1: [26] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 1: [27] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 1: [27] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 1: [28] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted Malone et al 1/24/22 10:59:00 AM

Page 1: [29] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 1: [30] Formatted **Malone et al** **1/24/22 10:59:00 AM**

Highlight

Page 2: [31] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 3: [32] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 15: [33] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 15: [34] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 15: [35] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 15: [36] Deleted **Malone et al** **1/24/22 10:59:00 AM**

Page 19: [37] Deleted **Malone et al** **1/24/22 10:59:00 AM**