Resolving heterogeneous fluxes from tundra halves the growing season carbon budget

Sarah M. Ludwig[1,2], Luke Schiferl[2,3], Jacqueline Hung[4], Susan M. Natali[4], and Roisin Commane[1,2]

[1]Department of Earth and Environmental Science, Columbia University, New York, NY, United States of America
[2]Lamont-Doherty Earth Observatory, Palisades, NY, United States of America
[3]Harvard John A. Paulson School of Engineering and Applied Sciences, Cambridge, MA, United States of America
[4]Woodwell Climate Research Center, Woods Hole, MA, United States of America

Correspondence: Sarah M. Ludwig (ludda.ludwig@columbia.edu)

Abstract. Landscapes are often assumed to be homogeneous when interpreting eddy covariance fluxes, which can lead to biases when gap-filling and scaling-up observations to determine regional carbon budgets. Tundra ecosystems are heterogeneous at multiple scales. Plant functional types, soil moisture, thaw depth, and microtopography, for example, vary across the landscape and influence net ecosystem exchange (NEE) of carbon dioxide (CO_2) and methane (CH_4) fluxes. With warming temperatures, Arctic ecosystems are changing from a net sink to a net source of carbon to the atmosphere in some locations, but the Arctic's carbon balance remains highly uncertain. In this study we report results from growing season NEE and CH_4 fluxes from an eddy covariance tower in the Yukon-Kuskokwim Delta in Alaska. We used footprint models and Bayesian Markov Chain Monte Carlo (MCMC) methods to unmix eddy covariance observations into constituent landcover fluxes based on high resolution landcover maps of the region. We compared three types of footprint models and used two landcover maps with varying complexity to determine the effects of these choices on derived ecosystem fluxes. We used artificially created gaps of withheld observations to compare gap-filling performance using our derived landcover-specific fluxes and traditional gap-filling methods that assume homogeneous landscapes. We also compared resulting regional carbon budgets when scaling-up observations using heterogeneous and homogeneous approaches. Traditional gap-filling methods performed worse at predicting artificially withheld gaps in NEE than those that accounted for heterogeneous landscapes, while there were only slight differences between footprint models and landcover maps. We identified and quantified hot spots of carbon fluxes in the landscape (e.g., late growing season emissions from wetlands and small ponds). We resolved distinct seasonality in tundra growing season NEE fluxes. Scaling while assuming a homogeneous landscape overestimated the growing season CO_2 sink by a factor of two and underestimated CH_4 emissions by a factor of two when compared to scaling with any method that accounts for landscape heterogeneity. We show how Bayesian MCMC, analytical footprint models, and high resolution landcover maps can be leveraged to derive detailed landcover carbon fluxes from eddy covariance timeseries. These results demonstrate the importance of landscape heterogeneity when scaling carbon emissions across the Arctic.

1 Introduction

Eddy covariance (EC) towers provide some of the longest and highest resolution timeseries of in situ observations of energy, water, and carbon fluxes. Eddy covariance flux data provide landscape-level insight into numerous ecosystem processes, such as water-use efficiency, crop yields, and carbon balances (Baldocchi, 2003; Baker and Griffis, 2005; Reichstein et al., 2007; Knauer et al., 2018). Global and regional networks of EC towers, such as FLUXNET and AmeriFlux (Novick et al., 2018; Papale, 2020), are commonly used to benchmark Earth system models, provide a priori fluxes for atmospheric inversion models, or train remote-sensing based models to scale bottom-up carbon budgets (Friend et al., 2007; Wang et al., 2007; Jung et al., 2009, 2020; Chevallier et al., 2012; Tramontana et al., 2016; Chen et al., 2018; Schiferl et al., 2022). The surface source area contributing to EC flux measurements (i.e., the footprint) is much larger than other types of direct flux measurements, such as chambers, but is spatially and temporally variable, and can change with wind direction and atmospheric stability. The dynamic spatial influence on EC fluxes is often ignored under implicit assumptions that landscapes within the EC footprint are homogeneous or spatially representative (Griebel et al., 2016; Giannico et al., 2018).

Numerous footprint models have been developed to quantify source area contribution to EC tower flux observations (Schmid, 2002) and inform interpretations and analysis of these fluxes. Aggregate footprints are commonly used to determine the general spatial extent and seasonal patterns in EC tower source areas (Amiro, 1998). When combined with landcover maps of EC tower locations, footprints have been used to filter flux observations to include only those from distinctly uniform source areas for further interpretation and analysis, though the practicality of this is highly dependent on landscape heterogeneity and the EC tower site location (Jammet et al., 2017; Juutinen et al., 2022; Beckebanze et al., 2022). Studies using concurrent chamber-based fluxes within EC tower source areas have used footprints to scale-up chamber fluxes and compare to EC tower fluxes, which can provide confidence in the flux measurements, the representativeness of the chamber fluxes, and the landcover map used (Kade et al., 2012; Stoy et al., 2013; Morin et al., 2017; Davidson et al., 2017). However, disagreement between scaled chamber and EC tower fluxes is difficult to diagnose; chamber fluxes are often limited in temporal resolution and spatial extent, and landcover maps might not capture detail or distinctions relevant for fluxes (Fox et al., 2008; Forbrich et al., 2011; Budishchev et al., 2014). Footprints have been used to identify hotspots of methane (CH_4) fluxes (Matthes et al., 2014; Rößger et al., 2019; Reuss-Schmidt et al., 2019), and in circumstances where there is a single known source location against a known or zero flux background, the footprint-weighted flux maps can derive CH_4 fluxes at these hotspots (Rey-Sanchez et al., 2022). However, footprint-weighted flux maps cannot derive actual fluxes in circumstances with multiple different CH_4 sources, or when fluxes, such as carbon dioxide (CO_2), have high temporal variability. Tuovinen et al. (2019) used footprints to weight contributions to CH_4 fluxes from land cover classifications in heterogeneous Siberian tundra, and by assuming fluxes were constant through time, was able to solve for landcover specific CH_4 fluxes using ordinary least squares.

Despite the documented effects of heterogeneous surfaces on the interpretation of fluxes, most uses of EC fluxes ignore the dynamic nature of flux source areas. For applications such as model benchmarking, bottom-up scaling, and gap-filling, the landscape around EC towers is implicitly assumed to be homogeneous. Gap-filled timeseries are often required to create seasonal or annual carbon budgets. One of the most widely used gap-filling approaches for CO_2 fluxes, marginal distribution

sampling (MDS), primarily uses the mean flux of CO_2 from similar meteorological conditions within a certain window of time, irrespective of the wind direction and source area of the gap-filled timepoint, or of the observations used to do the filling (Reichstein et al., 2005; Wutzler et al., 2018). Both model benchmarking and bottom-up carbon flux scaling rely on EC tower fluxes being spatially representative of a larger region (Williams et al., 2009). While landscape representativeness

60   or homogeneity is a reasonable assumption for some EC tower sites, such as agricultural fields, it is rarely tested explicitly with footprints. A recent study by Chu et al. (2021) tested the spatial representativeness of AmeriFlux sites using footprint climatologies, and found a minority of sites were representative of areas more than one kilometer away from the EC tower. A recent synthesis of circumpolar CH_4 fluxes excluded EC measurements that could not be attributed wholly to wetland or waterbody sources (Kuhn et al., 2021).

65   Arctic ecosystems in particular require representative carbon flux observations to accurately derive seasonal and annual budgets. Rapid arctic warming is thawing and mobilizing carbon stored in permafrost, leading to direct climate feedbacks through decomposition and indirect consequences through changing hydrology, vegetation, and disturbances (Schuur et al., 2015; Rantanen et al., 2022). There is large uncertainty in the arctic carbon budget, and it remains unclear whether the arctic is currently a carbon source or sink (McGuire et al., 2009, 2018; Natali et al., 2019, 2021; Virkkala et al., 2021; Watts

70   et al., 2021, 2023). Tundra ecosystems are extremely heterogeneous at multiple scales (Virtanen and Ek, 2014), which when combined with logistical difficulties in monitoring in the arctic, can lead to difficulties in calculating representative bottom-up carbon scaling (Goodrich et al., 2016; Lara et al., 2020; Pallandt et al., 2022). For example, bottom-up scaling models estimate twice as much CH_4 from the arctic as top-down atmospheric inversions (Thornton et al., 2016; Saunois et al., 2020).

   This study addresses how landscape heterogeneity affects gap-filling and bottom-up scaling of CO_2 and CH_4 EC fluxes. We

75   used footprint models and landcover maps to unmix EC fluxes into constituent landcover fluxes in heterogeneous tundra in the Yukon-Kuskokwim (YK) Delta, Alaska. We investigated how the choice of footprint model affects gap-filling and carbon budgets by comparing results using three of the most commonly used footprint models. We compared net ecosystem exchange (NEE) of CO_2 fluxes using both a simple and a complex landcover map to determine how the scale of heterogeneity that we consider impacts our resulting carbon budgets. Lastly, we compared gap-filled NEE fluxes and scaled-up carbon budgets to

80   an identical approach that only differs by assuming a homogenous landscape, and to a commonly used gap-filling approach (MDS), which implicitly assumes a homogeneous landscape. We discuss the implications of the resulting CO_2 and CH_4 fluxes and carbon budgets for the YK Delta and arctic carbon feedbacks.

2   Methods

2.1   Site Description

85 The study region is located in the Izaviknek and Kingaglia Uplands of the Yukon-Kuskokwim Delta in Alaska, approximately 90 km northwest of Bethel, Alaska and 110 km inland from the coast. Mean annual air temperature in Bethel was 1.2 °C for 2019 to 2020, 13.8 °C during summer (June, July, August), -11.9 °C during winter (December, January, and February), and above freezing from May to October. The study region is underlain by discontinuous permafrost, with permafrost underlying

peat plateaus and absent under wetlands and lakes (Frost et al., 2020). Thaw depths on peat plateaus averaged 30 to 40 cm in June and July 2016 to 2017 and 60 to 70 cm in September 2016 (Ludwig et al., 2022). Vegetation on the peat plateaus is heterogeneous, and is dominated by lichen (primarily *Cladonia* spp), *Sphagnum fuscum*, or low-lying shrubs, while wetland vegetation is typically *Sphagnum* and graminoid spp. (Zolkos et al., 2022). The EC tower was installed in July 2019 on a peat plateau in unburned tundra at (N 61.2548°, W 163.2589°).

We used a landcover map developed by Ludwig et al. (2022) to characterize the EC tower location and a nearby region of unburned tundra used for scaling up carbon budget (Ludwig et al., 2023a). The landcover map is 5 m x 10 m resolution and derived from Sentinel-1 synthetic aperture radar (SAR), Sentinel-2 multispectral instrument (MSI), and the ArcticDEM. Two versions of landcover were used: (1) a simple version with only four categories: surface water, tundra, wetland, and degrading permafrost, and (2) a complex version where tundra was further split into lichen tundra, shrub tundra, sedge tundra, and tundra at the edge of degrading permafrost (Fig. 1). Tundra landcover categories were primarily located on peat plateaus, and share the same dominant vegetation types of lichens, dwarf shrubs, mosses, and sedges. The differences within tundra categories were subtle; shrub tundra was often located at the edges of peat plateaus bordering and along banks with slightly larger shrubs; sedge tundra was located on peat plateau slopes that were slightly greener; lichen tundra was the least green and largest area of tundra types within the region, dominated by lichen, moss (*Sphagnum* spp. and *Dicranum* spp.), graminoids (*Carex* spp. and *Eriophorum angustifolium*) (Baillargeon et al., 2022); and edge of degraded tundra included tundra bordering degraded permafrost, often wetter, mossier, and slightly subsided. Degraded areas included isolated shallow depressions on peat plateaus, more evolved networks of flowpaths draining peat plateaus into wetlands, and recently drained waterbodies. Depending on seasonality and antecedent rain, degraded areas could have standing water, saturated soils, exposed mud, or graminoid-dominated vegetation. The wetland category included a range of wetland vegetation such as mosses, graminoids, and tall shrubs, often with complex underlying hydrology. Wetland soils were usually saturated, with small, sub-pixel channels or waterbodies undetectable at the resolution of the landcover map. Surface water includes all lakes, ponds, and streams detectable at the landcover map resolution (Ludwig et al., 2023b). There are likely smaller ponds or channels within wetlands and degraded areas, but higher resolution mapping would be needed to identify that level of heterogeneity. The full distribution of landcover areas in a 300 m radius circle around the EC tower location and in the region used for scaling is described in table 1. The scaling region was approximately 150 km^2, which is similar to the average size of a grid cell in earth system models (Williams et al., 2009).

2.2 Eddy covariance data processing

Data used in this study span from July 12th, 2019 to September 30th, 2020, though we only include May through September months. The EC tower instrumentation consisted of a Gill WindMaster Pro sonic anemometer, LI-7500DS open path analyzer for CO_2 and H2O, LI-7700 for CH_4, Vaisala HMP155 humidity and temperature probe, LI-190R quantum sensor for photosynthetically active radiation (PAR), Kipp and Zonen CNR4 four component net radiometer, and HukseFlux HFP01SC soil heat plates. All instrumentation was connected to a LI-7550 interface equipped with a LICOR SmartFlux system. The measurement height was 2.5 m above ground level. Half-hourly flux calculations were made using the eddy covariance method (Baldocchi

Table 1. Landcover category percentages in the immediate eddy covariance tower area (radius of 300 m) and in the region used to scale-up ecosystem carbon fluxes (Fig. 1).

Landcover category	Tower area	Scaling region
Lichen tundra	12 %	27 %
Shrub tundra	18 %	11 %
Sedge tundra	23 %	16 %
Edge of degraded permafrost	21 %	11 %
Degraded permafrost	3 %	5 %
Wetland	20 %	18 %
Water	3 %	12 %

et al., 1988) from 10 Hz data using the EddyPro software program (Fratini and Mauder, 2014). We used the double coordinate rotation method, spike removal, block averaging, and time lag removal by covariance maximization (Moncrieff et al., 1997).

125 We made corrections for air density fluctuations for CO_2, CH_4, and $H2O$ fluxes following Webb et al. (1980). We removed fluxes with nonstationarity (flags = 2 in the overall flag system) (Mauder and Foken, 2015). We filtered fluxes to remove times of low signal strength (< 15%) and low turbulence (friction velocity (u*) < 0.1 ms^{-1} threshold was chosen where CO_2 fluxes were independent of u*). Energy balance closure at the site was typical (70%). Lastly, we filtered fluxes to remove spikes using the double median absolute deviation method (Mauder et al., 2013). The resulting timeseries had 26% and 61% missing

130 data for CO_2 and CH_4 fluxes respectively. Due to limited access for site maintenance during the COVID-19 pandemic and the remote site location, power outages contributed to 1.5% missing data in fluxes, air temperature, and PAR. While only actual observations of air temperature and PAR were used for training gap-filling models, we used a complete timeseries of drivers for scaling and to sum fluxes to monthly carbon budgets. To interpolate missing data in air temperature and PAR for scaling we used the marginal distribution sampling and mean diurnal course method from REddyProc (Wutzler et al., 2018). Annual

135 timeseries of CO_2 fluxes, CH_4 fluxes, air temperature, and PAR observations can be found in the SI (Fig. S1a-d).

2.3 Eddy covariance footprint modeling

We compared three commonly used footprint models to determine source areas for fluxes: the Hsieh model (with the 2D extension from Detto et al. (2006)), the Kljun model, and the Kormann and Meixner model (Hsieh et al., 2000; Kormann and Meixner, 2001; Kljun et al., 2015). The Hsieh model is a hybrid approach blending a forward Lagrangian stochastic numerical

140 model with an analytical solution. The Kljun model uses multiple parameterizations of a backward Lagrangian particle model to be applicable across atmospheric stability regimes. The Kormann and Meixner model is a Eulerian analytical footprint model based on Monin-Obukhov similarity theory. All three footprint models assume Gaussian dispersion in the crosswind direction and horizontal homogeneity in turbulence effects (Schmid, 2002). Given the flat deltaic landscape and extremely short tundra canopy height relative to instrument measurement height, this site was an ideal location for footprint modeling, while still
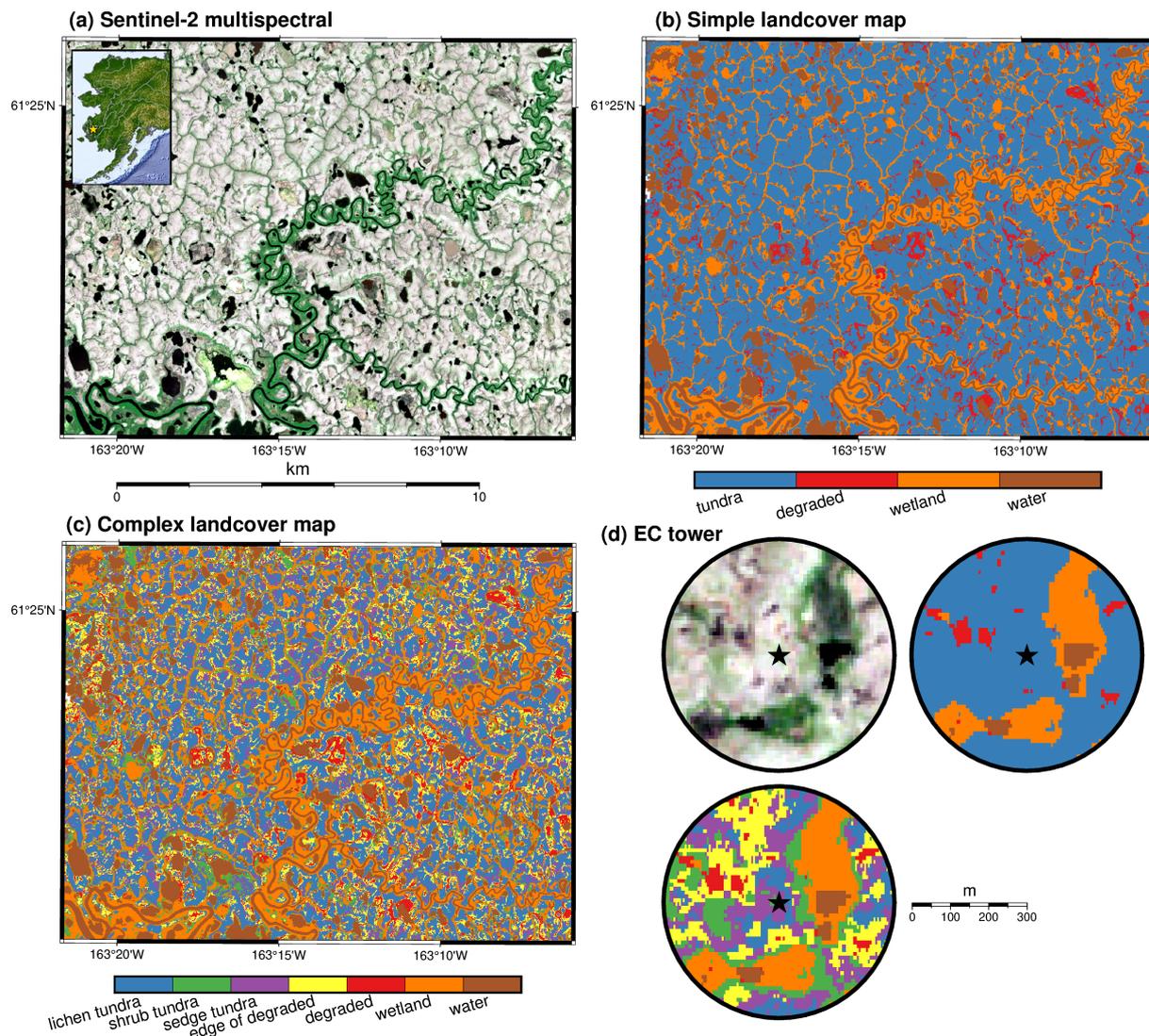
Figure 1. Scaling region within the YK Delta used in this study. Sentinel-2 RGB imagery (a) with the location of the grid within Alaska as an inset, simple landcover map (b), and complex landcover map (c). Panel (d) shows 300 m radius regions around the eddy covariance (EC) tower for the same Sentinel-2 RGB imagery as (a) in the upper left, the same simple landcover map as (b) in the upper right, and the same complex landcover map as (c) in the lower left. The EC tower is indicated by the star at N 61.2548°, W 163.2589°.

encompassing heterogeneity in CO_2 and CH_4 fluxes. We calculated a single roughness length for the site (0.02 m) from the measured wind speed and friction velocity under neutral conditions assuming a logarithmic wind profile and zero displacement height. For each half-hour flux observation, 1 x 1 m grid footprints were generated using each of the three model types, and then rotated into the wind direction. Each footprint was modeled 1000 m in the downwind direction, and 250 m to either side in the crosswind direction. These values were chosen as they were well in excess of the 90% contours of all footprints (peak

influences were < 100 m and the 90% contours averaged 200 m from the EC tower). These footprints were then reprojected to match the resolution and extent of the landcover maps at 5 x 10 m. Footprints were normalized to total 100% by dividing by the sum of the weights within each observation. The footprint weights were then summed over each landcover type (k) for each flux observation (i) as $\Omega_{i,k}$ (see equation 1 in section 2.4.1).

2.4 Unmixing and gap-filling models

We compared several modeling approaches for predicting and gap-filling the EC NEE timeseries. First, we explicitly consider landscape heterogeneity by unmixing EC tower fluxes using each of the three types of footprint models when summarized over both the simple and complex landcover map (Section 2.4.1). In order to do so, NEE fluxes were partitioned into respiration and gross primary productivity (GPP) with simple empirical models driven by PAR and air temperature. Second, we used the same method of flux partitioning, modeling, and parameter estimation to gap-fill NEE, but instead assume a homogeneous landscape. Each of the heterogeneous types of gap-filling models and the similar homogeneous variation were trained separately for each month in the growing season (May through September) to accommodate seasonality. Observations from both 2019 and 2020 were used to train the gap-filling models, though we only predicted and scaled for 2020, since the 2019 growing season was incomplete. We tested the inclusion of 2019 observations for August and September, and there was little effect on the derived landcover fluxes. Last, we compared these results to a widely-used approach by gap-filling NEE with MDS, which implicitly assumes a homogeneous landscape. CH_4 fluxes were not as temporally variable as NEE and largely unrelated to biometeorological drivers measured at the EC tower (Fig. S2). CH_4 fluxes were subsequently treated as landcover-specific constant fluxes through time and solved for separately in each month of the growing season, as has been done similarly in other studies (Rey-Sanchez et al., 2022; Tuovinen et al., 2019; Hannun et al., 2020).

2.4.1 Heterogeneous gap-filling models

Assuming that every pixel within a landcover type is characterized by a similar flux, then for a given (k^{th}) half-hour measurement, the observed EC tower NEE flux is the sum of each (i^{th}) landcover flux $(NEE_{i,k})$ times the total influence of those pixels within a footprint $(\Omega_{i,k})$ across all (P) landcovers (Eq 1).

$$NEE_{Obs,k} = \sum_{i=1}^{P} NEE_{i,k} * \Omega_{i,k} \tag{1}$$

CO_2 fluxes are often highly variable in time, especially from vegetated environments. Tundra NEE has been well characterized as the difference between respiration — modeled as an exponential function of temperature — and gross primary productivity—modeled as a light-saturating response curve often attenuated by temperature or vapor pressure deficit (Williams et al., 2006; Shaver et al., 2007; Loranty et al., 2011). For the heterogeneous gap-filling models, we structured the $(NEE_{i,k})$ fluxes from vegetated landcovers as temporally variable and dependent on air temperature $(Tair_k)$, light (PAR_k), and air temperature rescaled between 0 and 1 $(Tscale_k)$:

$$NEE_{i,k} = R_{i,k} - GPP_{i,k} \tag{2}$$

$$R_{i,k} = \alpha_i * e^{\beta_i * Tair_k} \tag{3}$$

$$GPP_{i,k} = Tscale_k * \frac{E0_i * Pmax_i * PAR_k}{Pmax_i + E0_i * PAR_k} \tag{4}$$

185

$$Tscale_k = \frac{(Tair_k - T_{min})(Tair_k - T_{max})}{(Tair_k - T_{min})(Tair_k - T_{max}) - (Tair_k - T_{opt})^2} \tag{5}$$

The parameters are the baseline respiration (α_i), the temperature sensitivity of respiration (β_i), the light-use efficiency of GPP ($E0_i$), and the maximum photosynthetic capacity ($Pmax_i$). GPP is attenuated by temperature using $Tscale_k$, where $T_{min} = -1.5\,^\circ\text{C}$, $T_{max} = 40\,^\circ\text{C}$, and $T_{opt} = 15\,^\circ\text{C}$ (Luus and Lin, 2015; Luus et al., 2017; Schiferl et al., 2022). For the simple

190 landcover map, NEE fluxes from tundra, wetland and degrading permafrost were all parameterized according to equations 2 to 5. Surface water CO_2 fluxes were parameterized as a constant flux over time. While this is likely an over-simplification, a more complex lake emissions model was not feasible because the surface waters within the footprint were too small an area and too small in footprint influence to inform a more complex model. Similarly, for the complex landcover map, all NEE fluxes from tundra landcover types as well as from wetland and degrading permafrost were structured according to equations 2 to 5 with

195 water as a constant flux.

CH$_4$ fluxes were assumed to be constant over time for each landcover type.

$$CH_{4\,Obs,k} = \sum_{i=1}^{P} CH_{4,i} * \Omega_{i,k} \tag{6}$$

The only parameters in this simpler version of unmixing are the landcover CH$_4$ fluxes themselves, $CH_{4,i}$ with the footprint influences ($\Omega_{i,k}$) as the only time-variable driver (Eq 6). All three footprint models were similarly compared for CH$_4$ fluxes.

200 Only the complex landcover map was used to unmix CH$_4$ fluxes, since the categories within tundra were known to be divergent, e.g., known very small fluxes from lichen tundra, while tundra at the edge of degraded could possibly be a large source (Ludwig et al., 2018b).

2.4.2 Parameter estimation and flux prediction

We unmixed EC tower fluxes to landcover CH$_4$ and $NEE_{i,k}$ fluxes by using a Bayesian analysis with Markov Chain Monte

205 Carlo (MCMC) simulation. We chose this method partly because unmixing approaches such as ordinary least squares (as used by Tuovinen et al. (2019)) are not applicable with the non-linear relationships used here between CO_2 and air temperature and PAR, and non-linear ordinary least squares (as used by Rößger et al. (2019)) assumes normal distributions for parameters and error variance, which is often not the case. In addition, there are several advantages to using a Bayesian approach to solve for landcover fluxes. First, we can provide prior information on flux parameters. This prior information could be specific (e.g.,

210 from chamber fluxes from landcovers within the footprints), it could be more general (e.g., dictating one landcover known to have higher GPP than another), or it could be mostly uninformative, and merely place restrictions on parameter space based on physical properties (e.g., non-negative $Pmax_i$). We used the latter approach to NEE priors for this study to demonstrate the

impacts of unmixing EC tower NEE on gap-filling accuracy and bottom-up scaling while using the fewest assumptions. Given the simpler approach used to unmix CH_4 fluxes, there were multiple solutions if all prior fluxes were strictly uninformative. We used mostly uninformative prior fluxes for landcovers known to emit CH_4 (e.g. fully saturated soils and open water) by disallowing CH_4 uptake for degraded, edge of degraded, wetland, and water landcover classes (Ludwig et al., 2018a). Peat plateau chamber flux measurements from 2017 demonstrate a very small but non-zero CH_4 flux at the driest time of the growing season (Ludwig et al., 2018b), and we assigned prior fluxes for tundra types accordingly. Prior distributions can be found in the SI (Table S1-S3). The second benefit of using Bayesian analysis with MCMC is that derived quantities and predictions of new data are inherently treated as random variables with their own probability distributions, thus enabling easy calculations of uncertainties. Therefore, we carry through uncertainty from both partitioning and gap-filling to uncertainty in predicted landcover $NEE_{i,k}$ or CH_4 fluxes, which, when summed over time and scaled up by area, leads to distributions of carbon budgets from which we can calculate explicit uncertainties.

For each month of the growing season (May to September), gap-filling models were fit separately for each footprint-type and landcover map combination. First, $NEE_{Obs,k}$ were filtered to dark data (PAR < 50 μmol m^{-2} s^{-1}) and respiration parameters (equation 3) were determined while using uninformative priors. The GPP parameters (equation 4) were then estimated using the $NEE_{Obs,k}$ from the full dataset, with uninformative GPP priors but using the posterior distributions of the respiration parameters as strict prior information for the respiration component of equation 2. CH_4 fluxes were fit separately by month as well, but all times of day were used. We used Gibbs sampler for the MCMC iterations (Just Another Gibbs Sampler; JAGS) implemented with the runjags R package (Denwood, 2016), with a burn-in of 5,000 iterations, an adaptation of 5,000 iterations, and retained 3,000 iterations in the final chains. Three parallel chains were used for each model with different initial parameter values (Table S1-3). We evaluated parameter convergence using the Gelman diagnostic (Gelman and Rubin, 1992; Brooks and Gelman, 1998). Model performance was further checked using posterior predictive checks of the mean, standard deviation, and sum of squared residuals (Gelman et al., 1996).

2.4.3 Homogeneous gap-filling models

We used several methods to gap-fill the EC tower NEE and CH_4 fluxes that both assume a homogeneous landscape/footprint for comparison. The first method is a Bayesian analysis with MCMC sampling that mirrors our landcover flux unmixing approach in every way except by assuming a homogeneous landscape. For the homogeneous Bayesian model, we assume a single landcover type everywhere that accounts for 100% of the footprint influence at every flux observation. The homogeneous landcover NEE was modeled monthly by equations 2 to 4, with the same partitioning and parameter estimation as described in section 2.4.2. The second homogeneous NEE gap-filling approach we used was MDS (Reichstein et al., 2005). We used the REddyProc package with default settings to implement the MDS gap-filling (Wutzler et al., 2018). Since the CH_4 fluxes did not have relationships to observed biometeorological drivers such as air temperature or PAR (Fig. S2), we estimated monthly budgets by calculating the monthly average CH_4 flux for each half hour of the diurnal cycle, and then applying these averages to each day within the month.

2.4.4 Artificial gaps

Artificial gaps in the NEE and CH_4 flux observation timeseries were created in order to be able to evaluate and compare gap-filling approaches. Since the MDS gap-filling method requires at least 90 days of half-hourly measurements, it could only be applied to the 2020 growing season (data in 2019 began in mid-July). Therefore, we only created artificial gaps in the 2020 growing season for comparability. Artificial gaps were generated separately for each month to ensure each portion of the growing season had a similar amount of withheld data. Between 15 and 20% of the timeseries of each month was withheld as random artificial gaps of stratified sizes, with ≈5% as larger gaps (10 observations), ≈5% as smaller gaps (4 observations), and the remainder as single gaps. The withheld drivers corresponding to the artificial gaps (PAR, air temperature, footprint weights) were then used to predict EC tower NEE or CH_4, and gap-filling methods were evaluated by calculating the root mean square error (RMSE).

2.5 Scaling up NEE and CH_4

We used the parameter posterior distributions from MCMC simulations and full timeseries of air temperature and PAR to predict complete, gap-filled, CH_4 and $NEE_{i,k}$ flux timeseries for the landcover types as described in sections 2.4.2. We then summed these distributions of half-hourly fluxes over time and multiplied by their respective areas in the scaling region (section 2.1, Fig. 1) to determine estimations of monthly carbon budgets for each landcover type. Monthly landcover carbon budgets were calculated for each footprint model and landcover map combination. Landcover carbon budgets were then summed to create monthly and growing season carbon budgets for each footprint model and map type. Monthly and growing season carbon budget distributions for the Bayesian homogeneous gap-filling models were similarly estimated. The observed EC tower NEE fluxes with MDS gap-filling were also summed over time and multiplied by the total scaling region to arrive at comparative monthly carbon budget estimates. CH_4 was presented alongside NEE in carbon budgets as CO_2-equivalents (CO_2-eq) by multiplying by a factor of 28, a conservative choice among commonly used approximations of relative global warming potentials (Bastviken et al., 2011; Stocker, 2013; Euskirchen et al., 2014; Beaulieu et al., 2020; Skytt et al., 2020). For discussion of uncertainty in carbon budgets, we calculate 89% Bayesian credible intervals (CI), which are analogous to frequentist 95% confidence intervals (Kruschke, 2014; McElreath, 2015; Hobbs and Hooten, 2015), using the highest density interval method from the 'baysetestR' package (Makowski et al., 2019).

3 Results & Discussion

3.1 Footprint influence

The most common and most influential landcover within the footprints was tundra, averaging ~70% influence over the growing season for all three footprint model types. There was a fairly even distribution between tundra category types, though sedge tundra had slightly more influence than lichen, shrub, and degraded edges. The footprint influence from wetlands was comparable to that of individual tundra types. The least represented landcovers were degraded permafrost and surface water, both of
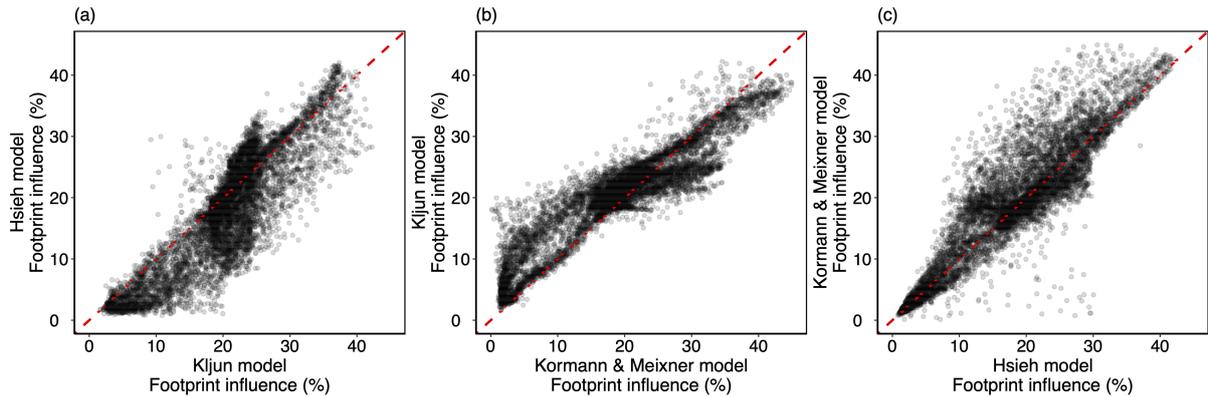
Figure 2. Scatterplot demonstrating comparison of footprint influence weights between the three models (Hsieh, Kljun, Kormann and Meixner) for lichen tundra. Other landcover footprint influence comparisons can be found in the SI (Fig. S3). The dashed red line indicates the 1:1 line.

which were between 0 to 20% influence (Fig. S3). The vast majority of footprints were a mix of landcover types, with almost no individual footprints having a landcover type with more than 75% influence (Fig. S3). There was a fair amount of agreement between the three footprint models, with the majority of footprint influences close to the 1:1 line on regressions between model types (Fig. 2, Fig. S3). Other studies that have sought to compare the ability of these footprint models to recover known flux sources have found little distinction between them, despite the differences in their methodology (Coates et al., 2021; Rey-Sanchez et al., 2022). However, the landcover influences used here were sums of all pixel influences within a landcover type, therefore small differences between models on a pixel basis were cumulative and would lead to larger discrepancies overall. There are also distinct periods of larger differences between footprint models, for example when the peak footprint influence was near the boundary between two landcover types; thus, a small shift in peak location between model types would lead to a large difference in landcover influences. A higher resolution landcover map (e.g., 3 m or smaller) would minimize some footprint model discrepancies, though this is relative to the extent of the footprints and the scale of landscape heterogeneity affecting carbon fluxes.

3.2 Model performance

All posterior predictive checks were passed (Bayesian p-values of $0.1 < p < 0.9$), and all parameters converged (Gelman diagnostics ≈ 1) for every Bayesian gap-filling model. All Bayesian gap-filling models were able to accurately reconstruct the EC tower NEE across the growing season as a function of PAR, air temperature, and source attributions (Fig. 3, Fig. S4). The only notable deviations were exclusive to outliers in EC tower NEE observations. This result is not unexpected, as EC data are often noisy. Mismatch with EC tower NEE outliers could also be a consequence of processes dominating fluxes that were not represented in our models, e.g., high CO_2 emissions from ebullition aligning with high lake influence within a footprint. When comparing performance for filling the same artificial gaps, all Bayesian models had a lower (better) RMSE than the MDS

method (Fig. 3). The Bayesian models, both heterogeneous and homogeneous, drive NEE as deterministic functions of PAR and temperature. This may be why they were more accurate than MDS, which has been shown to be biased in high-latitudes due to the effects of skewed distributions of net radiation (Vekuri et al., 2023). The heterogeneous gap-filling models often performed better than their homogeneous equivalent (Fig. 3). For most months, the heterogeneous complex map solutions had lower RMSEs than those of the simple map, though the Bayesian 89% CI overlapped between all months but May (Fig. 3). In May near the shoulder season, tundra types exhibited greater differences in seasonality. For example, the lichen tundra had very little GPP while the sedge tundra was a distinct carbon sink (Fig. 4), and this could be accommodated in the complex map while the simple map attempted to fit both to a single 'tundra' flux. While there were clear improvements in gap-filling RMSE using this unmixing method, the differences in RMSE were small relative to the magnitude of the fluxes. The drawbacks of the flux unmixing method used here are site-specific solutions and longer computation times, which increase with the landscape complexity considered. MDS remains faster to implement and could be preferred when landscape homogeneity can be safely assumed.

None of the three footprint models consistently performed better in terms of RMSE, and for most outcomes, the Bayesian 89% CI for their RMSEs overlapped (Fig. 3, Fig. S5). Given that none of the three footprint model types quantify their uncertainty, we continued to evaluate all three as an ensemble of footprint models that represents the range in footprint influence outcomes. Another way to evaluate performance of the three footprint models is by comparing their consistency in predicting landcover NEE fluxes when the underlying landcover map switches from simple to complex. The degraded permafrost, water, and wetland landcovers were identical between the two maps and ideally should have the same derived fluxes even if the tundra categories were treated differently. Similarly, the overall tundra footprint-weighted flux should match between simple and complex landcovers, even though the complex tundra was a combination of four types where there was only one tundra type for the simple map. By weighting the predicted landcover NEE by their respective footprint influences for each observation, we regressed the simple vs. complex solutions (Fig. S6-10). While the EC tower NEE and tundra total weighted NEE were very consistent between landcover maps for all footprint models, the Hsieh and Kormann and Meixner models were notably inconsistent for degraded permafrost for June and July (Fig. S7-8), while the Kljun footprint model was always distinctly consistent (Fig. S6-10). This outcome might indicate the Kljun footprint model was more representative of landcover influences at peak growing season. In the absence of extensive concurrent chamber fluxes to conclusively distinguish between derived landcovers from the footprint models, we recommend a footprint model ensemble approach.

3.3 Derived landcover fluxes

There was enough similarity between footprint model influences to yield similar patterns in derived landcover fluxes (Fig. 4, Fig. S11, Table S4-10). For example, in all three footprints both shrub tundra and tundra at the edge of degrading permafrost had higher peak carbon uptake (-0.342, -0.266, -0.308 kg-C month^{-1} m^{-2}, for Hsieh, Kljun, and Kormann and Meixner respectively) than sedge and lichen tundra (-0.175, -0.175, -0.139 kg-C month^{-1} m^{-2}, for Hsieh, Kljun, and Kormann and Meixner respectively) (Fig. 4). This aligns with previous studies that have found higher productivity in shrub tundra and areas adjacent to disturbed tundra, possibly the result of increased nutrient availability (Schuur et al., 2007; Bowden et al., 2008;
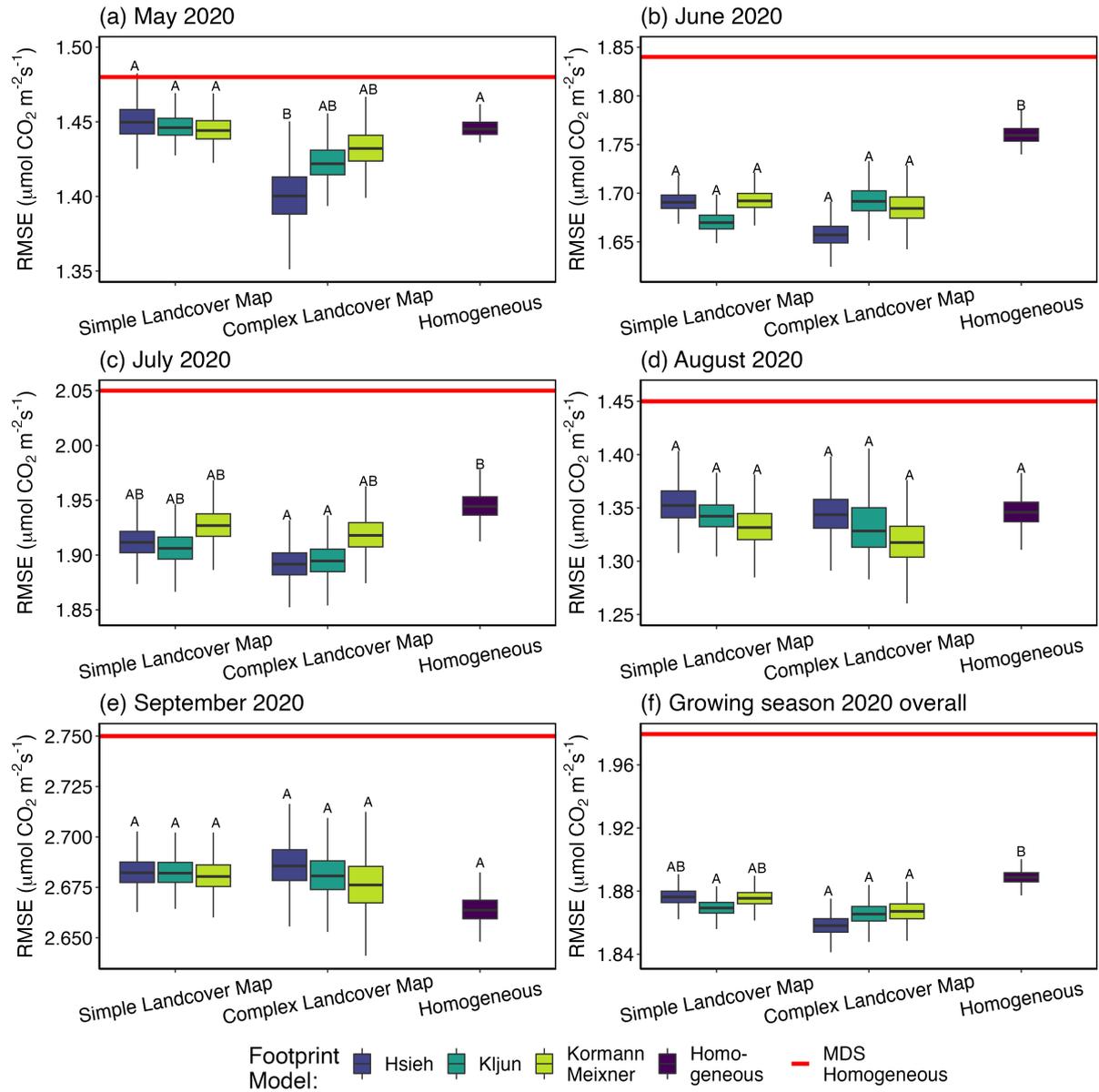
Figure 3. Monthly RMSE (a-e) and growing seasonal total (f) for 2020 from artificial gap-filling NEE fluxes. Boxes are the median and interquartile range (IQR), whiskers are 1.5*IQR, for the Bayesian model gap-filling RMSEs. The red line indicates MDS (marginal distribution sampling) gap-filling RMSEs. Distributions with overlapping Bayesian 89% credible intervals are designated with matching letters. Note the scales on y-axes are different between panels to highlight the comparability of footprint models and landcover maps within months.

Lee et al., 2011). The range of NEE fluxes derived for tundra vegetation was similar to ranges in NEE observed at other tundra sites (Euskirchen et al., 2012; Howard et al., 2020; Virkkala et al., 2022). All three footprint models also derived higher
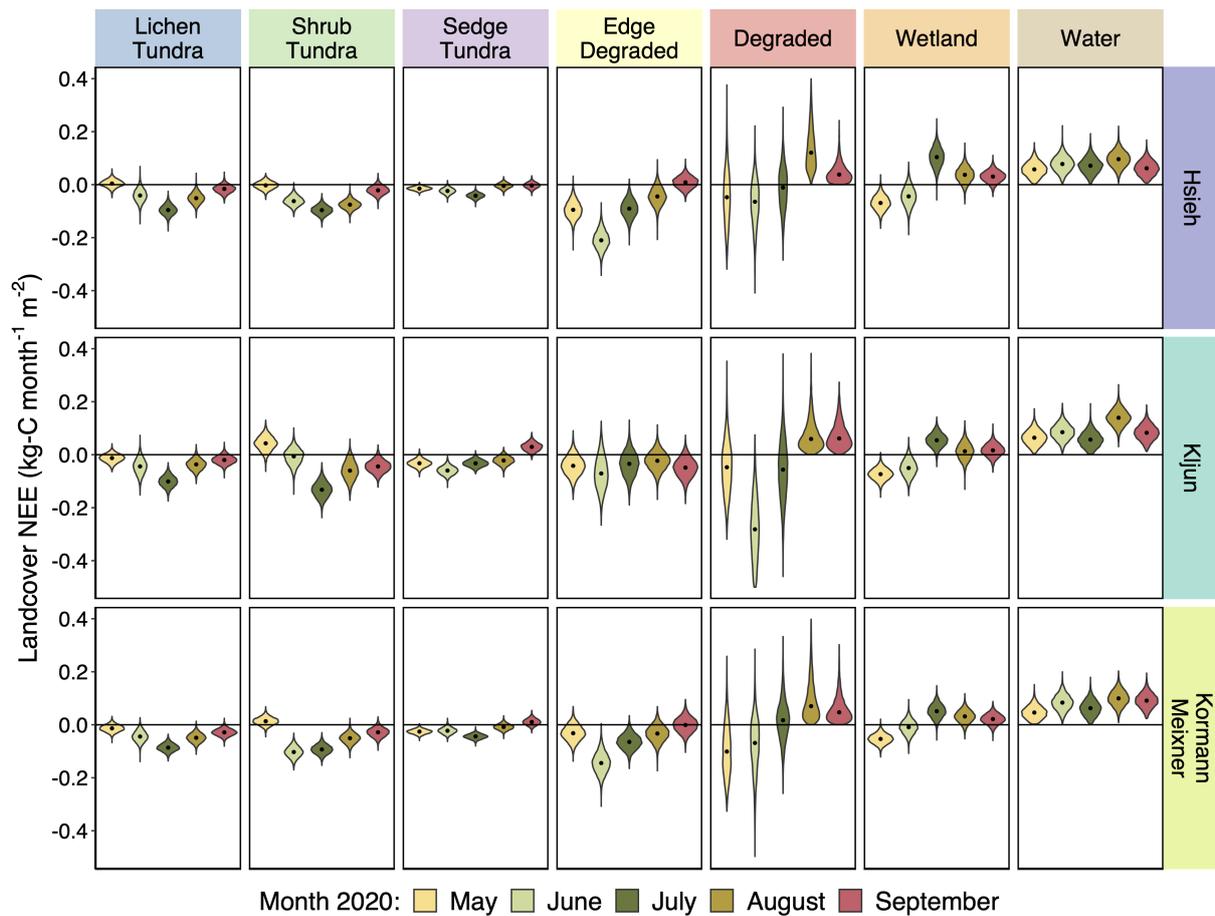
Figure 4. Monthly violin plots of predicted NEE fluxes from 2020 growing season by landcover (columns) for each of the three footprint models (rows) using the complex landcover map to unmix the EC tower fluxes. Distributions for violin plots are derived from posterior distributions of predicted NEE. Black dots indicate medians.

CO_2 and CH_4 emissions from surface water and wetlands later in the growing season (Fig. 4, Fig. 5), which could be the result of increased thaw depths contributing to greater lateral carbon transport from peat plateaus. Porewater dissolved organic
335   carbon and dissolved CO_2 and CH_4 was extremely high on peat plateaus during the growing season (Zolkos et al., 2022), and open water in both wetlands (sub-pixel) and waterbodies were likely hotspots for decomposition and outgassing (Ludwig et al., 2022). The wetlands were also characterized by deep, carbon-rich soil, which could be contributing to higher baseline respiration (Fig. 4, Table S9). The derived CH_4 fluxes from landcover classes in this study were within the ranges reported in the Boreal-Arctic Wetland and Lake Dataset (BAWLD-CH_4), including from wetlands and edge of degraded permafrost
340   (wetlands and wet tundra in BAWLD), peat plateau (dry tundra in BAWLD), and waterbody CH_4 fluxes (small peatland lakes in BAWLD) (Kuhn et al., 2021). The CO_2 fluxes reported here are similar in range to those observed in small ponds in other subarctic tundra ecosystems (Kuhn et al., 2018).

The three footprint models followed similar patterns in peat plateau seasonality as well, with NEE uptake peaking in July for most tundra types (Fig. 4). Lichen and sedge tundra were very small CH_4 sources (Fig. S11), though given the large area of lichen tundra in the landscape this resulted in a notable contribution to total CH_4 (Fig. 5) when scaling up. Shrub tundra was either zero or a very small CH_4 source, depending on the footprint model (Fig. 5). Tundra at the edge of degrading permafrost was a significant CH_4 source, and behaved more similarly to wetlands than degraded areas in terms of seasonal patterns (Fig. 5). Interestingly, degraded permafrost was a sink of CO_2 earlier in the growing season (Fig. 4), but all GPP parameters converged to zero in August and September (Fig. 4, Table S8). Degraded permafrost was a source of CH_4 early in the growing season, decreased near to zero as the depressions dried down, and then increased again later in the growing season (Fig. 5, Fig. S11). This aligns with the wettest portion of the growing season, when the small depressions of degrading permafrost become inundated as small ponds (Mullen et al., 2023), which could explain the renewed CH_4 emissions and decline in GPP.

There were differences as well between the derived landcover fluxes for the three footprint models. These differences were off-setting between adjacent landcover types. For example, degrading permafrost and tundra at the edge of degrading permafrost were always, by definition, near one another. When the Kljun model had high CO_2 uptake in degrading permafrost it had lower uptake at the edge of degrading permafrost, whereas, Hsieh and Kormann and Meixner displayed the opposite pattern (Fig. 4). This discrepancy was the result of slight differences between footprint models in peak influence positioning at the boundary of the two landcovers (Fig. 2, Fig. S3). The differences in effects of the footprint models can likely be minimized by using a relatively higher resolution map, or including spatial drivers such as LAI, soil moisture, and soil temperature, which would provide further constraints for landcover fluxes.

The complex heterogeneous models captured distinctive seasonality. Lichen and shrub tundra were net neutral in May, had peak CO_2 uptake in July, and remained small sinks in September (Fig. 4). In contrast, the sedge tundra and edge of degrading permafrost were small sinks in May, peaked earlier, and were net neutral or CO_2 sources by September (Fig. 4). Increasing the complexity of the underlying map allowed us to determine this separate but ecologically significant seasonality in peat plateau CO_2 cycling. However, there is a limit to how complex one can get. The landcover map used in this study identified two types of wetlands, one much more prevalent near the EC tower than the other. Attempting to use both wetland types failed, as the parameters for the less prevalent wetland could not to converge. We subsequently lumped both wetland categories as 'wetland'. This failure to converge serves as a check against over-fitting, which can be used in addition to comparing against withheld data via artificial gaps. In the event that a landcover with small area or influence is significant to the research questions posed and this unmixing method cannot derive a flux, then we recommend supplying stricter prior information for the landcover via chamber fluxes. By combining chamber fluxes and EC tower flux unmixing, one can leverage both the spatial coverage and temporal frequency of EC tower fluxes with the specificity of chamber fluxes.

3.4 Landcover scaling

Scaling up fluxes to the region led to distinct landcover hotspots of carbon sinks and sources, with all three footprint models having similar monthly NEE and CH_4 budgets by landcover type (Fig. 5, Fig. 6). Lichen tundra was the largest sink of carbon (the 89% CI of July carbon uptake was (-2628 to -5075 Mg-C) for the Hsieh model, (-2715 to -5452 Mg-C) for the Kljun
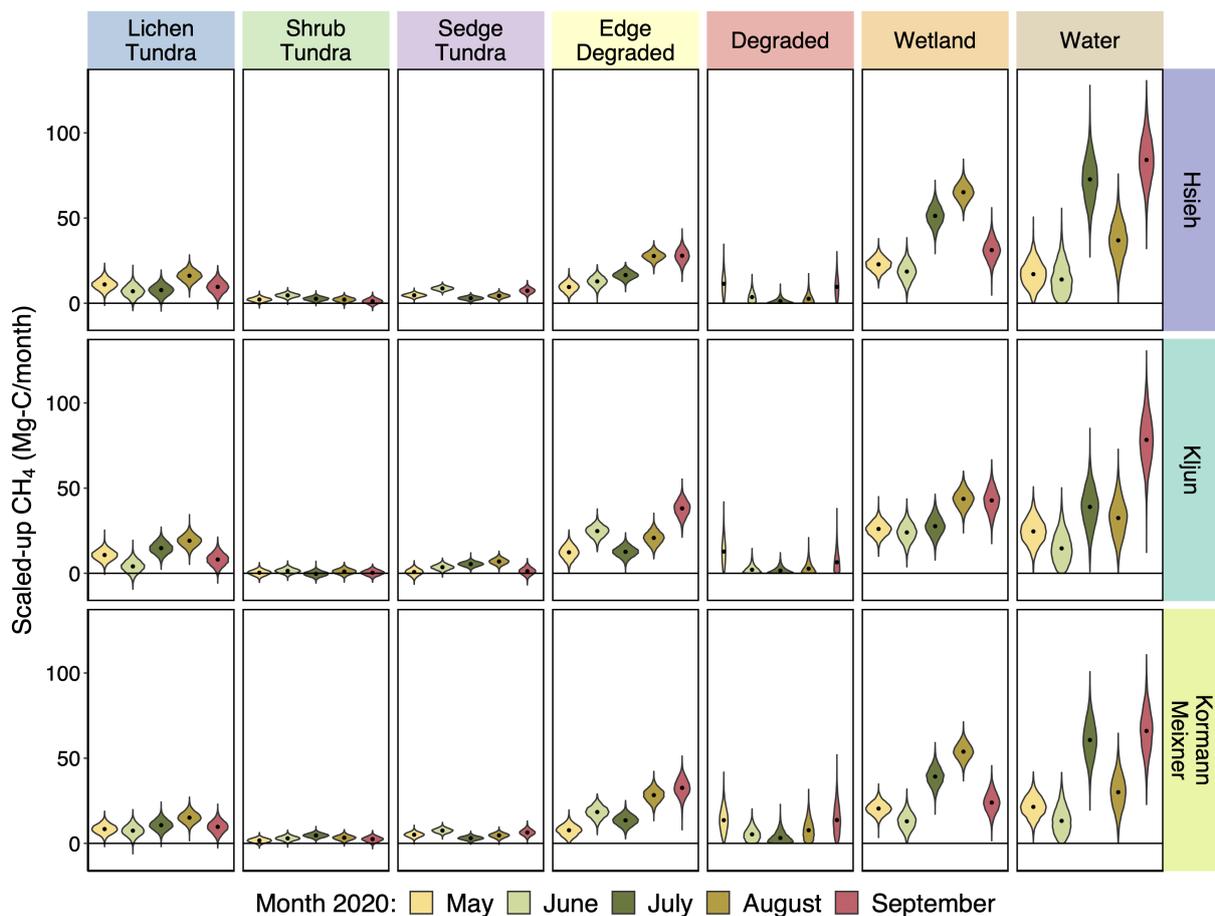
Figure 5. Monthly violin plots of CH$_4$ 2020 growing season budgets by landcovers (columns) for each footprint model (rows). Distributions for violin plots are derived from posterior distributions of predicted CH$_4$ fluxes scaled by their landcover areas in figure 1. Black dots indicate medians.

model, and (-2493 to -4428 Mg-C) for the Kormann and Meixner model), though this was driven in part by occupying the largest area in the region (Fig. 1; Table 1). Wetlands and surface waters were significant sources of both CO$_2$ and CH$_4$ in the latter half of the growing season (July-September), with large enough emissions to offset the carbon uptake in some of the peat
380   plateau landcovers (Fig. 5, Fig. 6). Wetlands account for, on average, only 7% of what the EC tower sees, but 18% of the area in the region. If we were to use a coarser landcover map, such as the recently updated circumpolar arctic vegetation map (CAVM) (Raynolds and Walker, 2022), we would have attributed 100% of the EC tower fluxes to wetland-complex vegetation, which would scale up to 26% of the region using CAVM. Given the clearly distinctive carbon dynamics between wetlands and peat plateau vegetation, using a landcover resolution appropriate to the scale of heterogeneity is important for obtaining an accurate
385   regional carbon budget and understanding the ecosystem.
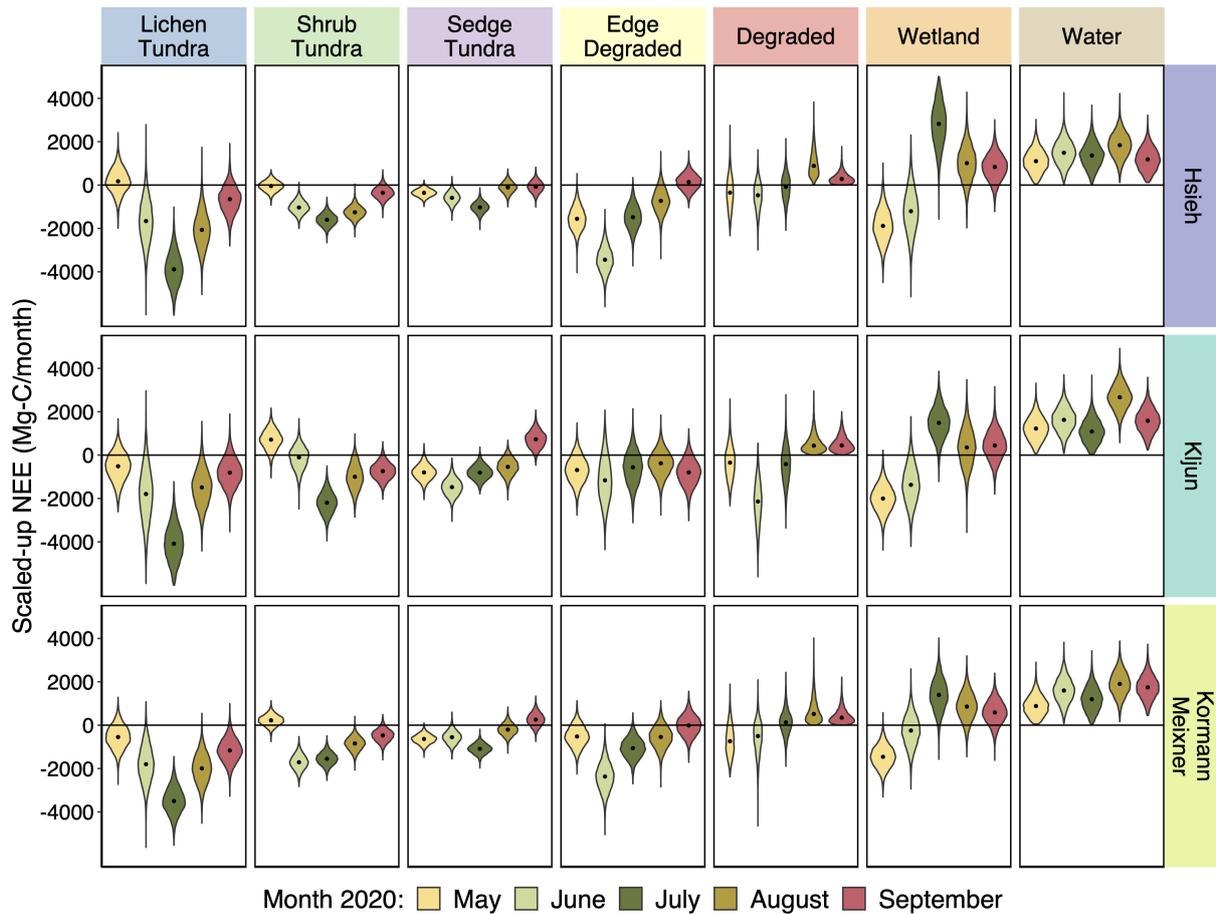
Figure 6. Monthly violin plots of 2020 growing season NEE carbon budgets by landcovers (columns) in the complex map for each footprint model (rows). Distributions for violin plots are derived from posterior distributions of predicted NEE fluxes scaled by their landcover areas in figure 1. Black dots indicate medians.

Regional surface water carbon emissions scaled-up from the EC tower fluxes are likely an overestimate, since the waterbodies within the EC tower footprints were amongst the two smallest size-classes of waterbodies in the region, which have the largest diffusive carbon fluxes (Ludwig et al., 2023b). In comparison, the coarser CAVM landcover map does not identify any surface waters in the entire scaling region, which would lead to these hotspots of emissions being completely underestimated. Scaling-up surface water carbon emissions would be better done using an approach that includes both terrestrial and aquatic landscape drivers and uses better spatial representation (e.g., Ludwig et al. (2023b)) than the area seen by a single EC tower. However, we were able to capture both plant-mediated carbon fluxes and ebullition in addition to diffusive fluxes, as well as describe the broad seasonal trends in carbon emissions using this approach.

3.5 Regional carbon scaling

Landcover carbon budgets were summed to regional carbon budgets and compared to carbon budgets estimated using the Bayesian and MDS homogeneous approaches (Fig. 7). Total NEE was similar between the three footprint models with overlapping 89% CI. Regional CH_4 budgets were also similar, with most months overlapping 89% CI between the three footprints, and a small difference in July. There was also no difference in total NEE budgets between the simple and complex map solutions. For most months the complex map solutions were slightly more uncertain, a consequence of estimating almost twice as many parameters and carrying through all of their uncertainties. The exception was May, where the simple map was more uncertain likely because grouping all tundra vegetation as one class was a poor assumption for that time. Despite the differences in methodologies, the two homogeneous approaches (Bayesian models and MDS) resulted in very similar NEE budgets to one another when scaling up to the region for June, July, and August (Fig. 7). In months closer to the shoulder season (May and September), the distributions of light and temperature were more skewed, which can be a source of bias in the MDS method and could explain the slight differences in the MDS and Bayesian homogeneous results for those months.

In contrast, all homogeneous scaled-up carbon results were quite different from the heterogeneous results. While the homogeneous approaches were worse at predicting to withheld gaps in the EC tower observations (Fig. 3), the differences in RMSE were small. However, the consequences for scaling were large. At every part of the growing season the homogeneous NEE overestimated the carbon sink relative to the heterogeneous NEE. This overestimation was smaller towards the shoulder seasons in May and September and larger in June, July, and August. Assuming homogeneity at this site meant approximating the same diurnal cycle of NEE everywhere that averages over different sink and source strengths in the landscape. Similarly, assuming a homogeneous landscape when scaling up an average CH_4 flux led to consistently underestimating the regional CH_4 emissions (Fig. 7). Because landcovers that were hotspots of emissions were farther from the footprint influence peaks, while those that exhibited larger carbon uptake were more often nearer the peaks, applying the EC tower flux to the region without accounting for footprints resulted in too much carbon uptake in NEE and too little carbon emissions in CH_4.

Throughout the growing season, incorrectly assuming a homogeneous landscape, regardless of gap-filling methodology, resulted in nearly doubling the NEE growing season carbon sink while nearly halving the CH_4 emissions (Fig. 7f). If we had assumed a homogeneous landscape we would have determined the region to be a net growing season carbon sink even after accounting for CH_4 emissions as CO_2-equivalents (89% CI: -9960 to -11919 Mg-C, Fig. 7f). We can combine the posterior distributions of scaled carbon from all three footprint model results to calculate a single CO_2-eq carbon budget estimate that accounts for across-model uncertainties. Doing this we find that growing season CH_4 emissions (mean: 16633, 89% CI: 15208 to 18212 Mg-C CO_2-eq) more than offset the CO_2 growing season sink (mean: -12512, 89% CI: -15718 to -9189 Mg-C). Similarly, Kuhn et al. (2018) found that accounting for emissions from commonly overlooked small ponds offset much of the wetland carbon sink in Northern Sweden. Other EC tower flux sites might see similar or opposite results from accounting for footprint heterogeneity (Griebel et al., 2016; Giannico et al., 2018; Reuss-Schmidt et al., 2019). A heterogeneous site with low carbon uptake or high carbon emissions located near the peak of footprint influences would overestimate carbon emissions

when scaling assuming homogeneity. Unaccounted-for heterogeneity such as this could help explain the mismatch between bottom-up and top-down carbon budgets (Thornton et al., 2016; Saunois et al., 2020).

3.6 Uncertainty

430   A benefit of using this Bayesian approach, is that uncertainties in model fit, both for respiration, GPP and constant fluxes, were carried through into uncertainties in NEE and CH_4 gap-filling and scaled-up carbon budgets. Sources of uncertainty included times and locations where the deterministic models used here were over-simplifications and failed to capture other important processes affecting carbon cycling. Instances where the landcover maps were not accurate delineations of carbon cycling are also included in this uncertainty. For example, if an underlying gradient of soil moisture were causing different

435   CH_4 fluxes within a vegetation type this would lead to greater uncertainty after the aggregation of footprint influences over the categorical map used here. Not all landcover carbon budgets were equal in terms of uncertainty; for example, among the tundra vegetation types, NEE fluxes from edge of degraded areas were the most uncertain, followed by shrub, lichen, and sedge (Fig. 4, e.g. Kljun-June standard deviations: 0.061, 0.031, 0.030, and 0.016 kg-C month^{-1} m^{-2} respectively). After scaling up to the region in (Fig. 1), lichen NEE carbon budgets were the most uncertain due to their larger area in the region, followed by

440   edge of degraded permafrost, shrub, and then sedge tundra (Fig. 6, e.g. Kljun-June standard deviations: 1199, 1009, 516, and 388 Mg-C month^{-1} respectively).

    Degraded permafrost NEE and CH_4 fluxes had the most uncertainty (Fig. 4, Fig. S11). This was likely due to a combination of their small extent and influence in the footprint supplying less signal to the EC tower fluxes, as well as the deterministic models over-simplifying carbon processes. In this case, the fluxes from degraded permafrost were distinctive enough they could

445   be determined while unmixing the EC tower fluxes despite their small area in the footprints. Within the areas of degrading permafrost there was heterogeneity in vegetation and surface water on a scale smaller than the resolution of the landcover maps, as well as more temporal dynamics related to hydrology. Similarly, NEE and CH_4 fluxes from water had relatively large uncertainty due to estimating an average flux rather than presenting diffusive, plant-mediated, and ebullitive fluxes deterministically. In lieu of representing these processes explicitly, our simpler models had greater uncertainty. The uncertainty around degraded

450   NEE and CH_4 fluxes had a smaller impact on NEE carbon budgets than other landcovers (Fig. 5, Fig. 6) due to the small area of degraded locations in the landscape (Fig. 1). Since all of the uncertainties in gap-filling fluxes and partitioning NEE were carried through into carbon budget estimates, we lose nothing from including these small areas of heterogeneity despite not representing them as well as other landcovers.

3.7 Applications and limitations of unmixing eddy covariance fluxes

455   Heterogeneity within EC tower landscapes is a common problem, and employing this flux unmixing approach at sites such as those identified by Chu et al. (2021) could improve accuracy in scaling carbon budgets and bench-marking models. Several studies have used summed spatial variables after weighting by EC footprints to relate to EC flux observations (Reuss-Schmidt et al., 2019; Xu et al., 2017; Metzger et al., 2013). While a useful way to incorporate heterogeneity, this approach reduces meaningful variation of spatial variables within footprints to single non-unique results. For example, there are multiple combi-
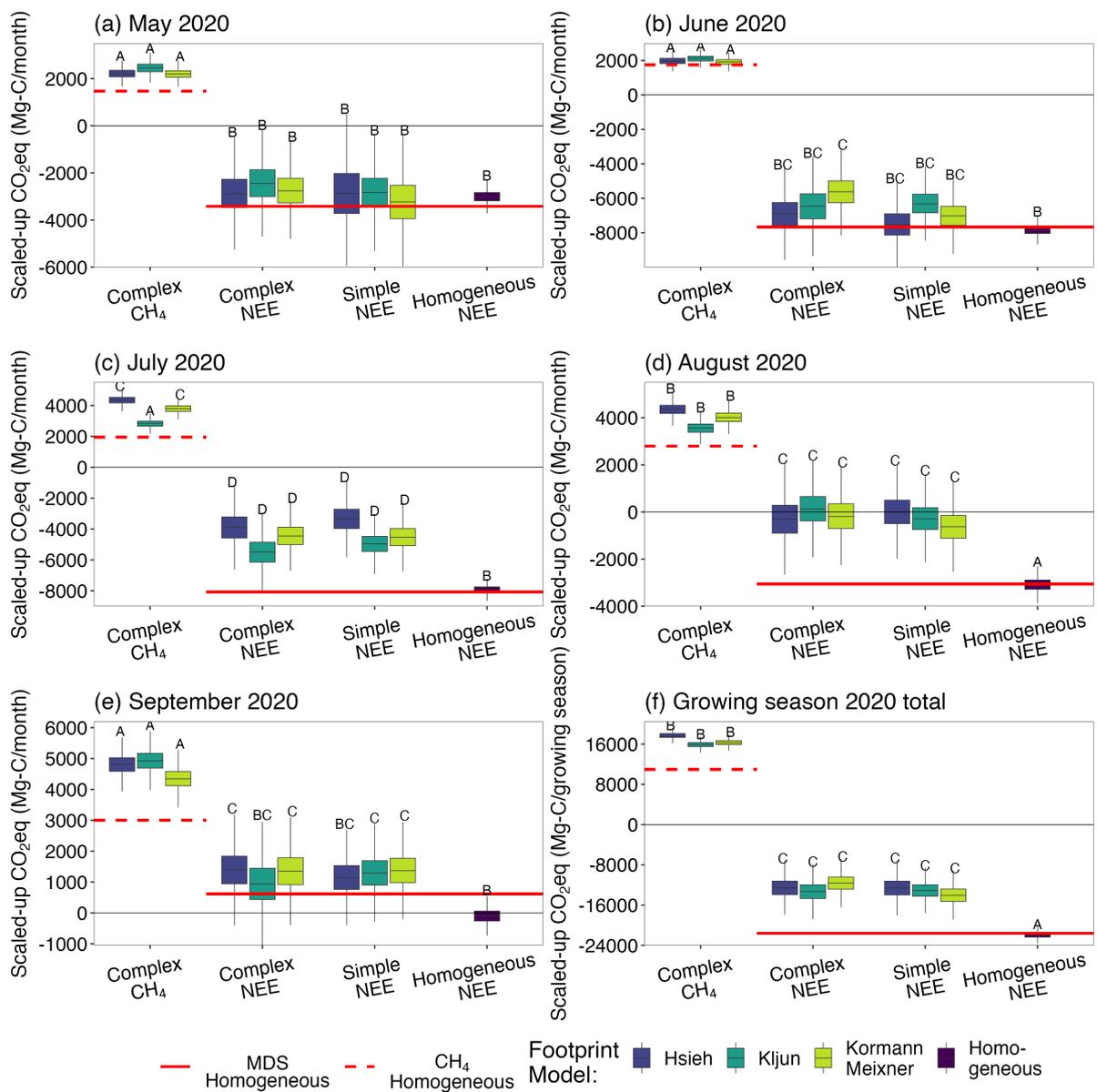
Figure 7. Monthly total (a-e) and 2020 growing season total (f) CO_2-eq carbon budgets for each gap-filling technique. Boxes are the median and interquartile range (IQR), whiskers are 1.5*IQR, for the Bayesian model gap-filling carbon budgets using the three footprint models over the complex and simple landcover maps, and without considering footprints and assuming a homogeneous landscape. The solid red line indicates MDS (marginal distribution sampling) homogeneous gap-filling NEE budgets. The dashed red line indicates the diurnal average CH4 gap-filling homogeneous budgets. Distributions with overlapping Bayesian 89% credible intervals are designated with matching letters. Note the scales on y-axes are different between panels to highlight the comparability of footprint models and landcover maps within months.

nations of footprints weights and values of the spatial drivers that could result in the same weighted sum. Statistically unmixing fluxes could yield more informative relationships to spatial drivers.

Future applications of the flux unmixing approach demonstrated in this study could incorporate spatially explicit drivers such as soil moisture and soil temperature, as well as more specific prior information from chamber fluxes. Doing so would further reduce uncertainty in landscape carbon fluxes. Seasonality could be represented through spatially explicit and temporally variable drivers such as solar induced fluorescence (SIF) (Luus et al., 2017; Schiferl et al., 2022). Interannual variability could be investigated using a hierarchical model structure by, for example, fitting an underlying distribution of a vegetation-type specific Q_{10} from which each year's specific Q_{10} is drawn. This method of interpreting EC fluxes could also be useful in sites with nested EC towers, multiple instrument heights, or where instrument heights have changed over time (e.g. (Klosterhalfen et al., 2023)). Flux data from such circumstances could be analyzed concurrently, since each observation is a function of an explicit footprint distribution. Thus, it would not matter if instrument height or position were different between observations.

An alternative model structure for GPP was investigated that uses leaf area index (LAI) as a driver (Shaver et al., 2007). In lieu of field-based LAI data, we used a timeseries of NDVI from cloud-free Sentinel-2 imagery and the empirical relationship to LAI from pan-Arctic tundra described in Shaver et al. (2013). The LAI-version GPP model failed posterior predictive checks for most months of data, and was not further pursued. This failure is likely because the approximation from NDVI was a poor representation of LAI for this site, particularly during May, August, and September where sub-pixel water presence could lead to erroneous NDVI and LAI. Furthermore, lichen and moss species dominated the vegetation biomass on peat plateaus and LAI may not be an appropriate metric in such cases. However, a spatially resolved driver such as LAI might be effective in applications for unmixing NEE at other sites.

This method of unmixing EC fluxes relies upon accurate footprint influence maps with sufficient variability over the heterogeneity in the landscape. The analysis in this study assumes the footprints were observed perfectly, i.e., footprint influence is not a random variable. For this reason, we recommend always using an ensemble of footprint models. However, for sites where the assumptions of footprint models are not met, and footprint influence maps are likely to be more error prone, this study's methodology will not work. Examples of such cases might include sites with instrument heights close to canopy heights where the effects of the surface roughness sub-layer are a concern, or anywhere the assumption of horizontal homogeneity in turbulence is invalid. In addition to valid footprint influences, this method requires variability in footprints. When footprint influences are aggregated over a landcover type for unmixing, there needs to be enough differences between observations to avoid an underdetermined dataset, where finding a solution to landcover specific fluxes won't be possible. Sites with consistent wind directions and atmospheric stability that result in very similar footprints between observations could have this issue. Small changes in the peak influence location could create enough variability between observations, even with consistent wind directions, depending on the position and scale of heterogenous landcover patches at the site.

4 Conclusions and implications

We compared multiple footprint models and landcover maps in our analysis to investigate their effects on unmixing landcover carbon fluxes. While the Kljun footprint model was the most consistent in determining fluxes when comparing outcomes using simple and complex landcover maps, there was no clear, best footprint model. We recommend including all three footprint models as an ensemble when interpreting EC fluxes. Flux estimates based on the more complex landcover map captured important differences in seasonality in tundra vegetation carbon fluxes. However, there were only minor differences in regional growing season carbon budgets between the two landcover maps, and using the more complex map had trade-offs such as greater computation time and uncertainty due to increasing the number of parameters. Investigating carbon fluxes using multiple landcover maps allows for informed lumping of landcover classes based on the resulting fluxes and the investigator's research questions.

Eddy covariance towers provide a wealth of high frequency flux data with large spatial extents. However, EC tower fluxes are under-utilized or potentially misleading if footprints are not taken into account in heterogeneous landscapes. We have demonstrated an approach to unmixing EC tower NEE and CH_4 fluxes from heterogeneous tundra, which provided detailed interpretations of landscape carbon cycling such as the detection and quantification of hot spots of carbon emissions and different timing in peak carbon uptake and senescence in tundra vegetation. We find that methods that consider footprint influences during gap-filling NEE fluxes were more accurate at predicting missing NEE fluxes than methods that assume landscape homogeneity. By using a Bayesian approach, we were able to quantify and compare uncertainties between carbon fluxes from different landcover classes. These uncertainties were carried through when gap-filling and scaling-up, providing an intrinsic estimate of uncertainty for the resulting carbon budgets. The consequence of assuming homogeneity in the landscape when gap-filling and scaling-up instead of using landcover-specific carbon fluxes was substantial: over the growing season (May to September) the homogeneous carbon budgets had half as much CH_4 emissions and twice as much net CO_2 uptake, greatly overestimating the carbon sink in the region and potential negative feedback to climate from carbon emissions. Accounting for landscape heterogeneity in carbon fluxes from EC towers could reduce uncertainty in bottom-up carbon budgets and the mismatch with top-down carbon budgets.

Code and data availability. Eddy covariance flux data, summarized footprint influences, and analysis code are located in the repository: https://github.com/arctic-carbon/heterogeneous_C_fluxes. Code for the three footprint influence models and reprojecting and summarizing over the landcover maps are located in the repository: https://github.com/arctic-carbon/eddy-footprint.

Author contributions. SMN proposed the study site and instrumentation. SML and SMN set up the EC tower and instrumentation and maintained and collected data. SML processed data, developed code, and performed the statistical analyses. SML prepared the manuscript and with contributions from all co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors acknowledge and are grateful for the opportunity to conduct research on the traditional land of the Yup'ik, who have stewarded this land through many generations. Our work would not have been possible without the support of the Yukon Delta National Wildlife Refuge, U.S. Fish and Wildlife Service. This study was supported with funding from a National Aeronautics and Space Administration FINESST grant (80NSSC19K1301) to SML, the Gordon and Betty Moore Foundation (grant #8414) and Woodwell Climate Research Center Fund for Climate Solutions to SMN, and National Science Foundation Office of Polar Programs grant (#1848620) to RC. This study was part of the NASA Arctic-Boreal Vulnerability Experiment.

References

Amiro, B.: Footprint climatologies for evapotranspiration in a boreal catchment, Agricultural and Forest Meteorology, 90, 195–201, https://doi.org/10.1016/S0168-1923(97)00096-8, 1998.

Baillargeon, N., Pold, G., Natali, S. M., and Sistla, S. A.: Lowland tundra plant stoichiometry is somewhat resilient decades following fire despite substantial and sustained shifts in community structure, Arctic, Antarctic, and Alpine Research, 54, 525–536, https://doi.org/10.1080/15230430.2022.2121246, 2022.

Baker, J. M. and Griffis, T. J.: Examining strategies to improve the carbon balance of corn/soybean agriculture using eddy covariance and mass balance techniques, Agricultural and Forest Meteorology, 128, 163–177, https://doi.org/10.1016/j.agrformet.2004.11.005, 2005.

Baldocchi, D. D.: Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future, Global Change Biology, 9, 479–492, https://doi.org/10.1046/j.1365-2486.2003.00629.x, 2003.

Baldocchi, D. D., Hincks, B. B., and Meyers, T. P.: Measuring Biosphere-Atmosphere Exchanges of Biologically Related Gases with Micrometeorological Methods, Ecology, 69, 1331–1340, https://doi.org/10.2307/1941631, 1988.

Bastviken, D., Tranvik, L. J., Downing, J. A., Crill, P. M., and Enrich-prast, A.: Freshwater Methane Emissions Offset the Continental Carbon Sink, Science, 331, 50, 2011.

Beaulieu, J. J., Waldo, S., Balz, D. A., Barnett, W., Hall, A., Platz, M. C., and White, K. M.: Methane and Carbon Dioxide Emissions From Reservoirs: Controls and Upscaling, Journal of Geophysical Research: Biogeosciences, 125, e2019JG005 474, https://doi.org/10.1029/2019JG005474, 2020.

Beckebanze, L., Rehder, Z., Holl, D., Wille, C., Mirbach, C., and Kutzbach, L.: Ignoring carbon emissions from thermokarst ponds results in overestimation of tundra net carbon uptake, Biogeosciences, 19, 1225–1244, https://doi.org/10.5194/bg-19-1225-2022, 2022.

Bowden, W. B., Gooseff, M. N., Balser, A., Green, A., Peterson, B. J., and Bradford, J.: Sediment and nutrient delivery from thermokarst features in the foothills of the North Slope, Alaska: Potential impacts on headwater stream ecosystems, Journal of Geophysical Research, 113, 2008.

Brooks, S. P. and Gelman, A.: General Methods for Monitoring Convergence of Iterative Simulations, Journal of Computational and Graphical Statistics, 7, 434–455, https://doi.org/10.1080/10618600.1998.10474787, 1998.

Budishchev, A., Mi, Y., van Huissteden, J., Belelli-Marchesini, L., Schaepman-Strub, G., Parmentier, F. J. W., Fratini, G., Gallagher, A., Maximov, T. C., and Dolman, A. J.: Evaluation of a plot-scale methane emission model using eddy covariance observations and footprint modelling, Biogeosciences, 11, 4651–4664, https://doi.org/10.5194/bg-11-4651-2014, 2014.

Chen, L., Dirmeyer, P. A., Guo, Z., and Schultz, N. M.: Pairing FLUXNET sites to validate model representations of land-use/land-cover change, Hydrology and Earth System Sciences, 22, 111–125, https://doi.org/10.5194/hess-22-111-2018, 2018.

Chevallier, F., Wang, T., Ciais, P., Maignan, F., Bocquet, M., Altaf Arain, M., Cescatti, A., Chen, J., Dolman, A. J., Law, B. E., Margolis, H. A., Montagnani, L., and Moors, E. J.: What eddy-covariance measurements tell us about prior land flux errors in CO2-flux inversion schemes, Global Biogeochemical Cycles, 26, https://doi.org/10.1029/2010GB003974, 2012.

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.,

565   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, Agricultural and Forest Meteorology, 301-302, 108 350, https://doi.org/10.1016/j.agrformet.2021.108350, 2021.

    Coates, T. W., Alam, M., Flesch, T. K., and Hernandez-Ramirez, G.: Field Testing Two Flux Footprint Models, Atmospheric Measurement
570   Techniques Discussions, pp. 1–7, https://doi.org/10.5194/amt-2021-106, 2021.

    Davidson, S., Santos, M., Sloan, V., Reuss-Schmidt, K., Phoenix, G., Oechel, W., and Zona, D.: Upscaling CH4 Fluxes Using High-Resolution Imagery in Arctic Tundra Ecosystems, Remote Sensing, 9, 1227, https://doi.org/10.3390/rs9121227, 2017.

    Denwood, M. J.: **runjags** : An *R* Package Providing Interface Utilities, Model Templates, Parallel Computing Methods and Additional Distributions for MCMC Models in **JAGS**, Journal of Statistical Software, 71, https://doi.org/10.18637/jss.v071.i09, 2016.

575   Detto, M., Montaldo, N., Albertson, J. D., Mancini, M., and Katul, G.: Soil moisture and vegetation controls on evapotranspiration in a heterogeneous Mediterranean ecosystem on Sardinia, Italy, Water Resources Research, 42, https://doi.org/10.1029/2005WR004693, 2006.

    Euskirchen, E., Bret-Harte, M., and Scott, G.: Seasonal patterns of carbon dioxide and water fluxes in three representative tundra ecosystems in northern Alaska, Ecosphere, 3, http://www.esajournals.org/doi/abs/10.1890/ES11-00202.1, 2012.

    Euskirchen, E. S., Edgar, C. W., Turetsky, M. R., Waldrop, M. P., and Harden, J. W.: Differential response of carbon fluxes to climate in
580   three peatland ecosystems that vary in the presence and stability of permafrost, Journal of Geophysical Research: Biogeosciences, 119, 1576–1595, https://doi.org/10.1002/2014JG002683, 2014.

    Forbrich, I., Kutzbach, L., Wille, C., Becker, T., Wu, J., and Wilmking, M.: Cross-evaluation of measurements of peatland methane emissions on microform and ecosystem scales using high-resolution landcover classification and source weight modelling, Agricultural and Forest Meteorology, 151, 864–874, https://doi.org/10.1016/j.agrformet.2011.02.006, 2011.

585   Fox, A. M., Huntley, B., Lloyd, C. R., Williams, M., and Baxter, R.: Net ecosystem exchange over heterogeneous Arctic tundra: Scaling between chamber and eddy covariance measurements, Global Biogeochemical Cycles, 22, https://doi.org/10.1029/2007GB003027, 2008.

    Fratini, G. and Mauder, M.: Towards a consistent eddy-covariance processing: an intercomparison of EddyPro and TK3, Atmospheric Measurement Techniques, 7, 2273–2281, https://doi.org/10.5194/amt-7-2273-2014, 2014.

    Friend, A. D., Arneth, A., Kiang, N. Y., Lomas, M., Ogée, J., Rödenbeck, C., Running, S. W., Santaren, J.-D., Sitch, S., Viovy,
590   N., Ian Woodward, F., and Zaehle, S.: FLUXNET and modelling the global carbon cycle, Global Change Biology, 13, 610–633, https://doi.org/10.1111/j.1365-2486.2006.01223.x, 2007.

    Frost, G. V., Loehman, R. A., Saperstein, L. B., Macander, M. J., Nelson, P. R., Paradis, D. P., and Natali, S. M.: Multi-decadal patterns of vegetation succession after tundra fire on the Yukon-Kuskokwim Delta, Alaska, Environmental Research Letters, 15, 025 003, https://doi.org/10.1088/1748-9326/ab5f49, 2020.

595   Gelman, A. and Rubin, D. B.: Inference from Iterative Simulation Using Multiple Sequences, Statistical Science, 7, https://doi.org/10.1214/ss/1177011136, 1992.

    Gelman, A., Meng, X.-L., and Stern, H.: Posterior Predictive Assesment of Model Fitness via Realized Discrepancies, Statistica Sinica, 6, 733–760, http://www.jstor.org/stable/24306036, 1996.

    Giannico, V., Chen, J., Shao, C., Ouyang, Z., John, R., and Lafortezza, R.: Contributions of landscape heterogeneity within
600   the footprint of eddy-covariance towers to flux measurements, Agricultural and Forest Meteorology, 260-261, 144–153, https://doi.org/10.1016/j.agrformet.2018.06.004, 2018.

Goodrich, J., Oechel, W., Gioli, B., Moreaux, V., Murphy, P., Burba, G., and Zona, D.: Impact of different eddy covariance sensors, site set-up, and maintenance on the annual balance of CO2 and CH4 in the harsh Arctic environment, Agricultural and Forest Meteorology, 228-229, 239–251, https://doi.org/10.1016/j.agrformet.2016.07.008, 2016.

605 Griebel, A., Bennett, L. T., Metzen, D., Cleverly, J., Burba, G., and Arndt, S. K.: Effects of inhomogeneities within the flux footprint on the interpretation of seasonal, annual, and interannual ecosystem carbon exchange, Agricultural and Forest Meteorology, 221, 50–60, https://doi.org/10.1016/j.agrformet.2016.02.002, 2016.

Hannun, R. A., Wolfe, G. M., Kawa, S. R., Hanisco, T. F., Newman, P. A., Alfieri, J. G., Barrick, J., Clark, K. L., DiGangi, J. P., Diskin, G. S., King, J., Kustas, W. P., Mitra, B., Noormets, A., Nowak, J. B., Thornhill, K. L., and Vargas, R.: Spatial heterogeneity in CO2, CH4,

610 and energy fluxes: insights from airborne eddy covariance measurements over the Mid-Atlantic region, Environmental Research Letters, 15, 035 008, https://doi.org/10.1088/1748-9326/ab7391, 2020.

Hobbs, N. T. and Hooten, M. B.: Bayesian Models: A Statistical Primer for Ecologists, Princeton University Press, https://www.degruyter.com/document/doi/10.1515/9781400866557/html, 2015.

Howard, D., Agnan, Y., Helmig, D., Yang, Y., and Obrist, D.: Environmental controls on ecosystem-scale cold-season methane and carbon

615 dioxide fluxes in an Arctic tundra ecosystem, Biogeosciences, 17, 4025–4042, https://doi.org/https://doi.org/10.5194/bg-17-4025-2020, 2020.

Hsieh, C.-I., Katul, G., and Chi, T.-w.: An approximate analytical model for footprint estimation of scalar fluxes in thermally stratified atmospheric flows, Advances in Water Resources, 23, 765–772, https://doi.org/10.1016/S0309-1708(99)00042-1, 2000.

Jammet, M., Dengel, S., Kettner, E., Parmentier, F.-J. W., Wik, M., Crill, P., and Friborg, T.: Year-round CH_4 and CO_2 flux dynamics in two

620 contrasting freshwater ecosystems of the subarctic, Biogeosciences, 14, 5189–5216, https://doi.org/10.5194/bg-14-5189-2017, 2017.

Jung, M., Reichstein, M., and Bondeau, A.: Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, Biogeosciences, 6, 2001–2013, https://doi.org/10.5194/bg-6-2001-2009, 2009.

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., Carvalhais, N., Chevallier, F., Gans, F., Goll, D. S., Haverd, V., Köhler, P., Ichii, K., Jain, A. K., Liu, J., Lombardozzi, D., Nabel, J. E. M. S., Nelson,

625 J. A., O'Sullivan, M., Pallandt, M., Papale, D., Peters, W., Pongratz, J., Rödenbeck, C., Sitch, S., Tramontana, G., Walker, A., Weber, U., and Reichstein, M.: Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach, Biogeosciences, 17, 1343–1365, https://doi.org/10.5194/bg-17-1343-2020, 2020.

Juutinen, S., Aurela, M., Tuovinen, J.-P., Ivakhov, V., Linkosalmi, M., Räsänen, A., Virtanen, T., Mikola, J., Nyman, J., Vähä, E., Loskutova, M., Makshtas, A., and Laurila, T.: Variation in CO_2 and CH_4 Fluxes Among Land Cover Types in Heterogeneous Arctic Tundra in

630 Northeastern Siberia, Biogeosciences Discussions, pp. 1–35, https://doi.org/10.5194/bg-2022-5, 2022.

Kade, A., Bret-Harte, M. S., Euskirchen, E. S., Edgar, C., and Fulweber, R. A.: Upscaling of CO2 fluxes from heterogeneous tundra plant communities in Arctic Alaska, Journal of Geophysical Research: Biogeosciences, 117, https://doi.org/10.1029/2012JG002065, 2012.

Kljun, N., Calanca, P., Rotach, M. W., and Schmid, H. P.: A simple two-dimensional parameterisation for Flux Footprint Prediction (FFP), Geoscientific Model Development, 8, 3695–3713, https://doi.org/10.5194/gmd-8-3695-2015, 2015.

635 Klosterhalfen, A., Chi, J., Kljun, N., Lindroth, A., Laudon, H., Nilsson, M. B., and Peichl, M.: Two-level eddy covariance measurements reduce bias in land-atmosphere exchange estimates over a heterogeneous boreal forest landscape, Agricultural and Forest Meteorology, 339, 109 523, https://doi.org/10.1016/j.agrformet.2023.109523, 2023.

Knauer, J., Zaehle, S., Medlyn, B. E., Reichstein, M., Williams, C. A., Migliavacca, M., De Kauwe, M. G., Werner, C., Keitel, C., Kolari, P., Limousin, J.-M., and Linderson, M.-L.: Towards physiologically meaningful water-use efficiency estimates from eddy covariance data, Global Change Biology, 24, 694–710, https://doi.org/10.1111/gcb.13893, 2018.

Kormann, R. and Meixner, F. X.: An Analytical Footprint Model For Non-Neutral Stratification, Boundary-Layer Meteorology, 99, 207–224, https://doi.org/10.1023/A:1018991015119, 2001.

Kruschke, J.: Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan, Academic Press, 2014.

Kuhn, M., Lundin, E. J., Giesler, R., Johansson, M., and Karlsson, J.: Emissions from thaw ponds largely offset the carbon sink of northern permafrost wetlands, Scientific Reports, 8, 9535, https://doi.org/10.1038/s41598-018-27770-x, 2018.

Kuhn, M. A., Varner, R. K., Bastviken, D., Crill, P., MacIntyre, S., Turetsky, M., Walter Anthony, K., McGuire, A. D., and Olefeldt, D.: BAWLD-CH4: a comprehensive dataset of methane fluxes from boreal and arctic ecosystems, Earth System Science Data, 13, 5151–5189, https://doi.org/10.5194/essd-13-5151-2021, 2021.

Lara, M. J., McGuire, A. D., Euskirchen, E. S., Genet, H., Yi, S., Rutter, R., Iversen, C., Sloan, V., and Wullschleger, S. D.: Local-scale Arctic tundra heterogeneity affects regional-scale carbon dynamics, Nature Communications, 11, 4925, https://doi.org/10.1038/s41467-020-18768-z, 2020.

Lee, H., Schuur, E. a. G., Vogel, J. G., Lavoie, M., Bhadra, D., and Staudhammer, C. L.: A spatially explicit analysis to extrapolate carbon fluxes in upland tundra where permafrost is thawing, Global Change Biology, 17, 1379–1393, https://doi.org/10.1111/j.1365-2486.2010.02287.x, 2011.

Loranty, M. M., Goetz, S. J., Rastetter, E. B., Rocha, A. V., Shaver, G. R., Humphreys, E. R., and Lafleur, P. M.: Scaling an Instantaneous Model of Tundra NEE to the Arctic Landscape, Ecosystems, 14, 76–93, https://doi.org/10.1007/s10021-010-9396-4, 2011.

Ludwig, S., Holmes, R. M., Natali, S. M., Mann, P. J., and Schade, J. D.: Yukon-Kuskokwim Delta fire: aquatic data, Yukon-Kuskokwim Delta Alaska, https://doi.org/10.18739/A22804Z8M, 2018a.

Ludwig, S., Holmes, R. M., Natali, S. M., Mann, P. J., Schade, J. D., Jardine, L., Melton, S., and Navarro-Perez, E.: Polaris Project 2017: Soil fluxes, carbon, and nitrogen, Yukon-Kuskokwim Delta, Alaska, https://doi.org/10.18739/A2Q23R08G, 2018b.

Ludwig, S. M., Natali, S. M., Mann, P. J., Schade, J. D., Holmes, R. M., Powell, M., Fiske, G., and Commane, R.: Using Machine Learning to Predict Inland Aquatic CO2 and CH4 Concentrations and the Effects of Wildfires in the Yukon-Kuskokwim Delta, Alaska, Global Biogeochemical Cycles, 36, e2021GB007 146, https://doi.org/10.1029/2021GB007146, 2022.

Ludwig, S. M., Natali, S. M., Schade, J. D., Powell, M., Fiske, G., Schiferl, L., and Commane, R.: CO2 and CH4 fluxes from waterbodies, and landcover map, YK Delta, Alaska, 2016-2019., https://doi.org/10.3334/ORNLDAAC/2178, 2023a.

Ludwig, S. M., Natali, S. M., Schade, J. D., Powell, M., Fiske, G., Schiferl, L. D., and Commane, R.: Scaling waterbody carbon dioxide and methane fluxes in the arctic using an integrated terrestrial-aquatic approach, Environmental Research Letters, 18, 064 019, https://doi.org/10.1088/1748-9326/acd467, 2023b.

Luus, K. A. and Lin, J. C.: The Polar Vegetation Photosynthesis and Respiration Model: a parsimonious, satellite-data-driven model of high-latitude CO2 exchange, Geoscientific Model Development, 8, 2655–2674, https://doi.org/10.5194/gmd-8-2655-2015, 2015.

Luus, K. A., Commane, R., Parazoo, N. C., Benmergui, J., Euskirchen, E. S., Frankenberg, C., Joiner, J., Lindaas, J., Miller, C. E., Oechel, W. C., Zona, D., Wofsy, S., and Lin, J. C.: Tundra photosynthesis captured by satellite-observed solar-induced chlorophyll fluorescence, Geophysical Research Letters, 44, 1564–1573, https://doi.org/10.1002/2016GL070842, 2017.

Makowski, D., Ben-Shachar, M. S., and Lüdecke, D.: bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework, Journal of Open Source Software, 4, 1541, https://doi.org/10.21105/joss.01541, 2019.

Matthes, J. H., Sturtevant, C., Verfaillie, J., Knox, S., and Baldocchi, D.: Parsing the variability in CH4 flux at a spatially heterogeneous wetland: Integrating multiple eddy covariance towers with high-resolution flux footprint analysis, Journal of Geophysical Research: Biogeosciences, 119, 1322–1339, https://doi.org/10.1002/2014JG002642, 2014.

Mauder, M. and Foken, T.: Documentation and Instruction Manual of the Eddy Covariance Software Package TK3, Arbeitsergebnisse, Universität Bayreuth, Abteilung Mikrometeorologie, ISSN 1614-8916, 62, https://doi.org/10.5194/bg-5-451-2008, 2015.

Mauder, M., Cuntz, M., Drüe, C., Graf, A., Rebmann, C., Schmid, H. P., Schmidt, M., and Steinbrecher, R.: A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements, Agricultural and Forest Meteorology, 169, 122–135, https://doi.org/10.1016/j.agrformet.2012.09.006, 2013.

McElreath, R.: Statistical Rethinking: A Bayesian Course with Examples in R and STAN, CRC Press, 2015.

McGuire, A. D., Genet, H., Lyu, Z., Pastick, N., Stackpoole, S., Birdsey, R., D'Amore, D., He, Y., Rupp, T. S., Striegl, R., Wylie, B. K., Zhou, X., Zhuang, Q., and Zhu, Z.: Assessing historical and projected carbon balance of Alaska: A synthesis of results and policy/management implications, Ecological Applications, 28, 1396–1412, https://doi.org/10.1002/eap.1768, 2018.

McGuire, D., Anderson, L., Christensen, T., Dallimore, S., Guo, L., Hayes, D. J., Heimann, M., Lorenson, T., Macdonald, R. W., and Roulet, N. T.: Sensitivity of the carbon cycle in the Arctic to climate change, Ecological Monographs, 79, 523–555, 2009.

Metzger, S., Junkermann, W., Mauder, M., Butterbach-Bahl, K., Trancón y Widemann, B., Neidl, F., Schäfer, K., Wieneke, S., Zheng, X. H., Schmid, H. P., and Foken, T.: Spatially explicit regionalization of airborne flux measurements using environmental response functions, Biogeosciences, 10, 2193–2217, https://doi.org/10.5194/bg-10-2193-2013, 2013.

Moncrieff, J. B., Massheder, J. M., de Bruin, H., Elbers, J., Friborg, T., Heusinkveld, B., Kabat, P., Scott, S., Soegaard, H., and Verhoef, A.: A system to measure surface fluxes of momentum, sensible heat, water vapour and carbon dioxide, Journal of Hydrology, 188-189, 589–611, https://doi.org/10.1016/S0022-1694(96)03194-0, 1997.

Morin, T. H., Bohrer, G., Stefanik, K. C., Rey-Sanchez, A. C., Matheny, A. M., and Mitsch, W. J.: Combining eddy-covariance and chamber measurements to determine the methane budget from a small, heterogeneous urban floodplain wetland park, Agricultural and Forest Meteorology, 237-238, 160–170, https://doi.org/10.1016/j.agrformet.2017.01.022, 2017.

Mullen, A. L., Watts, J. D., Rogers, B. M., Carroll, M. L., Elder, C. D., Noomah, J., Williams, Z., Caraballo-Vega, J. A., Bredder, A., Rickenbaugh, E., Levenson, E., Cooley, S. W., Hung, J. K. Y., Fiske, G., Potter, S., Yang, Y., Miller, C. E., Natali, S. M., Douglas, T. A., and Kyzivat, E. D.: Using High-Resolution Satellite Imagery and Deep Learning to Track Dynamic Seasonality in Small Water Bodies, Geophysical Research Letters, 50, e2022GL102 327, https://doi.org/10.1029/2022GL102327, 2023.

Natali, S. M., Watts, J. D., Rogers, B. M., Potter, S., Ludwig, S. M., Selbmann, A.-K., Sullivan, P. F., Abbott, B. W., Arndt, K. A., Birch, L., Björkman, M. P., Bloom, A. A., Celis, G., Christensen, T. R., Christiansen, C. T., Commane, R., Cooper, E. J., Crill, P., Czimczik, C., Davydov, S., Du, J., Egan, J. E., Elberling, B., Euskirchen, E. S., Friborg, T., Genet, H., Göckede, M., Goodrich, J. P., Grogan, P., Helbig, M., Jafarov, E. E., Jastrow, J. D., Kalhori, A. A. M., Kim, Y., Kimball, J. S., Kutzbach, L., Lara, M. J., Larsen, K. S., Lee, B.-Y., Liu, Z., Loranty, M. M., Lund, M., Lupascu, M., Madani, N., Malhotra, A., Matamala, R., McFarland, J., McGuire, A. D., Michelsen, A., Minions, C., Oechel, W. C., Olefeldt, D., Parmentier, F.-J. W., Pirk, N., Poulter, B., Quinton, W., Rezanezhad, F., Risk, D., Sachs, T., Schaefer, K., Schmidt, N. M., Schuur, E. A. G., Semenchuk, P. R., Shaver, G., Sonnentag, O., Starr, G., Treat, C. C., Waldrop, M. P., Wang, Y., Welker, J., Wille, C., Xu, X., Zhang, Z., Zhuang, Q., and Zona, D.: Large loss of CO2 in winter observed across the northern permafrost region, Nature Climate Change, https://doi.org/10.1038/s41558-019-0592-8, 2019.

Natali, S. M., Holdren, J. P., Rogers, B. M., Treharne, R., Duffy, P. B., Pomerance, R., and MacDonald, E.: Permafrost carbon feedbacks threaten global climate goals, Proceedings of the National Academy of Sciences, 118, https://doi.org/10.1073/pnas.2100163118, 2021.

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, Agricultural and Forest Meteorology, 249, 444–456, https://doi.org/10.1016/j.agrformet.2017.10.009, 2018.

Pallandt, M. M. T. A., Kumar, J., Mauritz, M., Schuur, E. A. G., Virkkala, A.-M., Celis, G., Hoffman, F. M., and Göckede, M.: Representativeness assessment of the pan-Arctic eddy covariance site network and optimized future enhancements, Biogeosciences, 19, 559–583, https://doi.org/10.5194/bg-19-559-2022, 2022.

Papale, D.: Ideas and perspectives: enhancing the impact of the FLUXNET network of eddy covariance sites, Biogeosciences, 17, 5587–5598, https://doi.org/10.5194/bg-17-5587-2020, 2020.

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., and Laaksonen, A.: The Arctic has warmed nearly four times faster than the globe since 1979, Communications Earth and Environment, 3, 1–10, https://doi.org/10.1038/s43247-022-00498-3, 2022.

Raynolds, M. and Walker, D.: Raster Circumpolar Arctic Vegetation Map, https://doi.org/10.17632/c4xj5rv6kv.2, 2022.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., and Valentini, R.: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm, Global Change Biology, 11, 1424–1439, https://doi.org/10.1111/j.1365-2486.2005.001002.x, 2005.

Reichstein, M., Papale, D., Valentini, R., Aubinet, M., Bernhofer, C., Knohl, A., Laurila, T., Lindroth, A., Moors, E., Pilegaard, K., and Seufert, G.: Determinants of terrestrial ecosystem carbon balance inferred from European eddy covariance flux sites, Geophysical Research Letters, 34, L01 402, https://doi.org/10.1029/2006GL027880, 2007.

Reuss-Schmidt, K., Levy, P., Oechel, W., Tweedie, C., Wilson, C., and Zona, D.: Understanding spatial variability of methane fluxes in Arctic wetlands through footprint modelling, Environmental Research Letters, 14, 125 010, https://doi.org/10.1088/1748-9326/ab4d32, 2019.

Rey-Sanchez, C., Arias-Ortiz, A., Kasak, K., Chu, H., Szutu, D., Verfaillie, J., and Baldocchi, D.: Detecting Hot Spots of Methane Flux Using Footprint-Weighted Flux Maps, Journal of Geophysical Research: Biogeosciences, 127, e2022JG006 977, https://doi.org/10.1029/2022JG006977, 2022.

Rößger, N., Wille, C., Veh, G., Boike, J., and Kutzbach, L.: Scaling and balancing methane fluxes in a heterogeneous tundra ecosystem of the Lena River Delta, Agricultural and Forest Meteorology, 266-267, 243–255, https://doi.org/10.1016/j.agrformet.2018.06.026, 2019.

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 System Science Data, 12, 1561–1623, https://doi.org/https://doi.org/10.5194/essd-12-1561-2020, 2020.

Schiferl, L. D., Watts, J. D., Larson, E. J. L., Arndt, K. A., Biraud, S. C., Euskirchen, E. S., Henderson, J. M., McKain, K., Mountain, M. E., Munger, J. W., Oechel, W. C., Sweeney, C., Yi, Y., Zona, D., and Commane, R.: Using atmospheric observations to quantify annual biogenic carbon dioxide fluxes on the Alaska North Slope, Biogeosciences, pp. 5953–5972, https://doi.org/10.5194/bg-19-5953-2022, 2022.

755 Schmid, H. P.: Footprint modeling for vegetation atmosphere exchange studies: a review and perspective, Agricultural and Forest Meteorology, 113, 159–183, https://doi.org/10.1016/S0168-1923(02)00107-7, 2002.

Schuur, E. A. G., Crummer, K. G., Vogel, J. G., and Mack, M. C.: Plant Species Composition and Productivity following Permafrost Thaw and Thermokarst in Alaskan Tundra, Ecosystems, 10, 280–292, https://doi.org/10.1007/s10021-007-9024-0, 2007.

Schuur, E. A. G., McGuire, A. D., Grosse, G., Harden, J. W., Hayes, D. J., Hugelius, G., Koven, C. D., Kuhry, P., Schädel, C., Grosse, G.,

760 Harden, J. W., Hayes, D. J., Hugelius, G., Koven, C. D., Kuhry, P., Lawrence, D. M., Natali, S. M., Olefeldt, D., Romanovsky, V. E., Schaefer, K., Turetsky, M. R., Treat, C. C., and Vonk, J. E.: Climate change and the permafrost carbon feedback, Nature, 520, 171–179, https://doi.org/10.1038/nature14338, 2015.

Shaver, G. R., Street, L. E., Rastetter, E. B., Van Wijk, M. T., and Williams, M.: Functional convergence in regulation of net CO2 flux in heterogeneous tundra landscapes in Alaska and Sweden, Journal of Ecology, 95, 802–817, https://doi.org/10.1111/j.1365-2745.2007.01259.x,

765 2007.

Shaver, G. R., Rastetter, E. B., Salmon, V., Street, L. E., Weg, M. J. V. D., Rocha, A., Wijk, M. T. V., and Williams, M.: Pan-Arctic modelling of net ecosystem exchange of CO2, Philosophical Transactions of the Royal Society B, 2013.

Skytt, T., Nielsen, S. N., and Jonsson, B.-G.: Global warming potential and absolute global temperature change potential from carbon dioxide and methane fluxes as indicators of regional sustainability – A case study of Jämtland, Sweden, Ecological Indicators, 110, 105 831,

770 https://doi.org/10.1016/j.ecolind.2019.105831, 2020.

Stocker, T.: Climate Change 2013: The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2013.

Stoy, P. C., Williams, M., Evans, J. G., Prieto-Blanco, A., Disney, M., Hill, T. C., Ward, H. C., Wade, T. J., and Street, L. E.: Upscaling Tundra CO $_2$ Exchange from Chamber to Eddy Covariance Tower, Arctic, Antarctic, and Alpine Research, 45, 275–284,

775 https://doi.org/10.1657/1938-4246-45.2.275, 2013.

Thornton, B. F., Wik, M., and Crill, P. M.: Double-counting challenges the accuracy of high-latitude methane inventories, Geophysical Research Letters, 43, 12,569–12,577, https://doi.org/10.1002/2016GL071772, 2016.

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M. A., Cescatti, A., Kiely, G., Merbold, L., Serrano-Ortiz, P., Sickert, S., Wolf, S., and Papale, D.: Predicting carbon dioxide and energy fluxes across global FLUXNET

780 sites with regression algorithms, Biogeosciences, 13, 4291–4313, https://doi.org/10.5194/bg-13-4291-2016, 2016.

Tuovinen, J.-P, Aurela, M., Hatakka, J., Räsänen, A., Virtanen, T., Mikola, J., Ivakhov, V., Kondratyev, V., and Laurila, T.: Interpreting eddy covariance data from heterogeneous Siberian tundra: land-cover-specific methane fluxes and spatial representativeness, Biogeosciences, 16, 255–274, https://doi.org/https://doi.org/10.5194/bg-16-255-2019, 2019.

Vekuri, H., Tuovinen, J.-P., Kulmala, L., Papale, D., Kolari, P., Aurela, M., Laurila, T., Liski, J., and Lohila, A.: A widely-used eddy covari-

785 ance gap-filling method creates systematic bias in carbon balance estimates, Scientific Reports, 13, 1720, https://doi.org/10.1038/s41598-023-28827-2, 2023.

Virkkala, A.-M., Aalto, J., Rogers, B. M., Tagesson, T., Treat, C. C., Natali, S. M., Watts, J. D., Potter, S., Lehtonen, A., Mauritz, M., Schuur, E. A. G., Kochendorfer, J., Zona, D., Oechel, W., Kobayashi, H., Humphreys, E., Goeckede, M., Iwata, H., Lafleur, P. M., Euskirchen,

E. S., Bokhorst, S., Marushchak, M., Martikainen, P. J., Elberling, B., Voigt, C., Biasi, C., Sonnentag, O., Parmentier, F.-J. W., Ueyama, M., Celis, G., St.Louis, V. L., Emmerton, C. A., Peichl, M., Chi, J., Järveoja, J., Nilsson, M. B., Oberbauer, S. F., Torn, M. S., Park, S.-J., Dolman, H., Mammarella, I., Chae, N., Poyatos, R., López-Blanco, E., Christensen, T. R., Kwon, M. J., Sachs, T., Holl, D., and Luoto, M.: Statistical upscaling of ecosystem CO2 fluxes across the terrestrial tundra and boreal domain: Regional patterns and uncertainties, Global Change Biology, n/a, https://doi.org/https://doi.org/10.1111/gcb.15659, 2021.

Virkkala, A.-M., Natali, S. M., Rogers, B. M., Watts, J. D., Savage, K., Connon, S. J., Mauritz, M., Schuur, E. A. G., Peter, D., Minions, C., Nojeim, J., Commane, R., Emmerton, C. A., Goeckede, M., Helbig, M., Holl, D., Iwata, H., Kobayashi, H., Kolari, P., López-Blanco, E., Marushchak, M. E., Mastepanov, M., Merbold, L., Parmentier, F.-J. W., Peichl, M., Sachs, T., Sonnentag, O., Ueyama, M., Voigt, C., Aurela, M., Boike, J., Celis, G., Chae, N., Christensen, T. R., Bret-Harte, M. S., Dengel, S., Dolman, H., Edgar, C. W., Elberling, B., Euskirchen, E., Grelle, A., Hatakka, J., Humphreys, E., Järveoja, J., Kotani, A., Kutzbach, L., Laurila, T., Lohila, A., Mammarella, I., Matsuura, Y., Meyer, G., Nilsson, M. B., Oberbauer, S. F., Park, S.-J., Petrov, R., Prokushkin, A. S., Schulze, C., St. Louis, V. L., Tuittila, E.-S., Tuovinen, J.-P., Quinton, W., Varlagin, A., Zona, D., and Zyryanov, V. I.: The ABCflux database: Arctic–boreal CO$_2$ flux observations and ancillary information aggregated to monthly time steps across terrestrial ecosystems, Earth System Science Data, 14, 179–208, https://doi.org/10.5194/essd-14-179-2022, 2022.

Virtanen, T. and Ek, M.: The fragmented nature of tundra landscape, International Journal of Applied Earth Observation and Geoinformation, 27, 4–12, https://doi.org/10.1016/j.jag.2013.05.010, 2014.

Wang, Y. P., Baldocchi, D., Leuning, R., Falge, E., and Vesala, T.: Estimating parameters in a land-surface model by applying nonlinear inversion to eddy covariance flux measurements from eight FLUXNET sites, Global Change Biology, 13, 652–670, https://doi.org/10.1111/j.1365-2486.2006.01225.x, 2007.

Watts, J., Natali, S. M., Minions, C., Risk, D., Arndt, K. A., Zona, D., Euskirchen, E. S., Rocha, A. V., Sonnentag, O., Helbig, M., Kalhori, A., Oechel, W. C., Ikawa, H., Ueyama, M., Suzuki, R., Kobayashi, H., Celis, G., Schuur, E. A., Humphreys, E. R., Kim, Y., Lee, B.-Y., Goetz, S. J., Madani, N., Schiferl, L., Commane, R., Kimball, J. S., Liu, Z., Torn, M. S., Potter, S., Wang, J. A., Jorgenson, T., Xiao, J., Li, X., and Edgar, C.: Soil respiration strongly offsets carbon uptake in Alaska and Northwest Canada, Environmental Research Letters, https://doi.org/10.1088/1748-9326/ac1222, 2021.

Watts, J. D., Farina, M., Kimball, J. S., Schiferl, L. D., Liu, Z., Arndt, K. A., Zona, D., Ballantyne, A., Euskirchen, E. S., Parmentier, F.-J. W., Helbig, M., Sonnentag, O., Tagesson, T., Rinne, J., Ikawa, H., Ueyama, M., Kobayashi, H., Sachs, T., Nadeau, D. F., Kochendorfer, J., Jackowicz-Korczynski, M., Virkkala, A., Aurela, M., Commane, R., Byrne, B., Birch, L., Johnson, M. S., Madani, N., Rogers, B., Du, J., Endsley, A., Savage, K., Poulter, B., Zhang, Z., Bruhwiler, L. M., Miller, C. E., Goetz, S., and Oechel, W. C.: Carbon uptake in Eurasian boreal forests dominates the high-latitude net ecosystem carbon budget, Global Change Biology, 29, 1870–1889, https://doi.org/10.1111/gcb.16553, 2023.

Webb, E. K., Pearman, G. I., and Leuning, R.: Correction of flux measurements for density effects due to heat and water vapour transfer, Quarterly Journal of the Royal Meteorological Society, 106, 85–100, https://doi.org/10.1002/qj.49710644707, 1980.

Williams, M., Street, L. E., Van Wijk, M. T., and Shaver, G. R.: Identifying differences in carbon exchange among arctic ecosystem types, Ecosystems, 9, 288–304, https://doi.org/10.1007/s10021-005-0146-y, 2006.

Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D. Y., Kattge, J., Leuning, R., Luo, Y., Tomelleri, E., Trudinger, C. M., and Wang, Y.-P.: Improving land surface models with FLUXNET data, Biogeosciences, 6, 1341–1359, https://doi.org/10.5194/bg-6-1341-2009, 2009.

Wutzler, T., Lucas-Moffat, A., Migliavacca, M., Knauer, J., Sickel, K., Šigut, L., Menzer, O., and Reichstein, M.: Basic and extensible post-processing of eddy covariance flux data with REddyProc, Biogeosciences, 15, 5015–5030, https://doi.org/10.5194/bg-15-5015-2018, 2018.

Xu, K., Metzger, S., and Desai, A. R.: Upscaling tower-observed turbulent exchange at fine spatio-temporal resolution using environmental response functions, Agricultural and Forest Meteorology, 232, 10–22, https://doi.org/10.1016/j.agrformet.2016.07.019, 2017.

Zolkos, S., MacDonald, E., Hung, J. K. Y., Schade, J. D., Ludwig, S., Mann, P. J., Treharne, R., and Natali, S.: Physiographic Controls and Wildfire Effects on Aquatic Biogeochemistry in Tundra of the Yukon-Kuskokwim Delta, Alaska, Journal of Geophysical Research: Biogeosciences, 127, e2022JG006 891, https://doi.org/10.1029/2022JG006891, 2022.

830