Methane fluxes from Arctic & boreal North America: Comparisons between process-based estimates and atmospheric observations

Hanyu Liu[1], Felix R. Vogel[5], Misa Ishizawa[5], Zhen Zhang[4], Benjamin Poulter[7], Doug E.J. Worthy[5], Leyang Feng[1], Anna Laure Gagné-Landmann[3], Ao Chen[1], Ziting Huang[1], Dylan C. Gaeta[1], Joe R. Melton[6], Douglas Chan[5], Vineet Yadav[2], Deborah Huntzinger[3], and Scot M. Miller[1]

[1]Department of Environmental Health and Engineering, Johns Hopkins University, Baltimore, MD, USA
[2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[3]School of Earth and Sustainability, Northern Arizona University, Flagstaff, AZ, USA
[4]Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China
[5]Environment and Climate Change Canada, Toronto, ON, Canada
[6]Environment and Climate Change Canada, Victoria, BC, Canada
[7]National Aeronautics and Space Administration, Greenbelt, MD, USA

Correspondence: Hanyu Liu (hliu154@jhu.edu)

Abstract. Methane (CH_4) flux estimates from high-latitude North American wetlands remain highly uncertain in magnitude, seasonality, and spatial distribution. In this study, we evaluate a decade (2007 – 2017) of CH_4 flux estimates by comparing 16 process-based models with atmospheric CH_4 observations collected from in situ towers. We compare the Global Carbon Project (GCP) process-based models with a model inter-comparison from a decade earlier called The Wetland and Wetland CH_4 Intercomparison of Models Project (WETCHIMP). Our analysis reveals that the GCP models have a much smaller inter-model uncertainty and have an average magnitude that is a factor of 1.5 smaller across Canada and Alaska. However, current GCP models likely overestimate wetland fluxes by a factor of two or more across Canada and Alaska based on tower-based atmospheric CH_4 observations. The differences in flux magnitudes among GCP models are more likely driven by uncertainties in the amount of soil carbon or spatial extent of inundation than in temperature relationships, such as Q_{10} factors. The GCP models do not agree on the timing and amplitude of the seasonal cycle, and we find that models with a seasonal peak in July and August show the best agreement with atmospheric observations. Models that exhibit the best fit to atmospheric observation also have a similar spatial distribution; these models concentrate fluxes near Canada's Hudson Bay Lowlands. Current, state-of-the-art process-based models are much more consistent with atmospheric observations than models from a decade ago, but our analysis shows that there are still numerous opportunities for improvement.

1 Introduction

Natural sources of CH_4 contribute $\sim$40% of total global fluxes, and wetlands are possibly the largest single source (e.g., Kirschke et al., 2013; Saunois et al., 2025). Understanding the magnitude, seasonality, and spatial distribution of wetland CH_4 fluxes is important to accurately predicting future carbon-climate feedbacks. However, the response of wetland CH_4 fluxes to temperature changes is uncertain (e.g., Zhang et al., 2023, 2017), especially in high-latitude regions where warming occurs 2-4

20   times faster than the global average (e.g., Rantanen et al., 2022).

At least some of this uncertainty is related to uncertain permafrost dynamics. Permafrost covers approximately $\sim$15% of the land in the Northern Hemisphere (Obu, 2021), and it serves as a massive reservoir for carbon. Globally, permafrost regions store about 1,000 to 1,672 peta-grams (Pg) of soil organic carbon (SOC), nearly twice the total amount of carbon in the atmosphere (Schuur et al., 2015; Hugelius et al., 2014; van Huissteden and Dolman, 2012). As permafrost thaws, it changes the soil

25   environment and triggers microbial decomposition of the stored organic matter. When the soil is wet, microbial decomposition in permafrost leads to the release of CH_4 through the process of anaerobic respiration. One study indicates that wetland CH_4 fluxes can be large enough to flip some high latitude regions from a net carbon sink to a net source (Watts et al., 2023).

To understand high-latitude wetland CH_4 fluxes and better predict future warming, process-based (bottom-up) models are important as they can be used to estimate current wetland CH_4 fluxes and project future CH_4 fluxes from regional to global

30   scales, leveraging current scientific knowledge of different biogeochemical processes (e.g., Saunois et al., 2025; Nzotungicim-paye et al., 2021; Melton et al., 2013; Zhang et al., 2017). Despite their importance, the CH_4 flux estimates from bottom-up models can have large discrepancies and uncertainties. For example, bottom-up estimates show that total global wetland fluxes range from 100 to 256 Tg CH_4 yr^{-1} (Xiao et al., 2024; Zhang et al., 2025; Saunois et al., 2025; Liu et al., 2020). In boreal North America, process-based models also estimate wetland CH_4 fluxes ranging from 13.8 to 39.6 Tg CH_4 per year (Poulter

35   et al., 2017). In addition, a recent study suggests an increase of 50 to 150% in global wetland CH_4 fluxes by 2100, a large range of numbers which points to large uncertainties in current projections (Koffi et al., 2020). Model inter-comparison projects like the Wetland and Wetland CH_4 Intercomparison of Models Project (WETCHIMP) have been used to compare the state-of-the-art wetland CH_4 flux models across different regions of the globe (Melton et al., 2013; Wania et al., 2013; Bohn et al., 2015). In more recent years, the Global Carbon Project (GCP) has been created to synthesize scientific knowledge of the global car-

40   bon cycle, and this effort includes a large ensemble of the latest process-based CH_4 flux models (Poulter et al., 2017; Zhang et al., 2025). There is limited knowledge on how these models have improved or evolved over time compared to the earlier WETCHIMP inter-comparison. Fortunately, projects like WETCHIMP and GCP make it easier to identify improvements and diagnose uncertainties in wetland flux models because all modeling groups use similar modeling protocols, meteorological inputs, and, in some cases, common inundation or wetland maps. By harmonizing inputs across models, we eliminate input-

45   driven variability due to different climate forcing data, and the remaining model spread therefore primarily reflects differences in process representations and parameterizations.

Numerous studies have also used atmospheric CH_4 to quantify CH_4 fluxes from high-latitude wetlands across North America. Early studies used a sparse network of tower observations in Canada and Alaska to quantify the magnitude, seasonality, and

spatial distribution of wetland fluxes (e.g., Worthy et al., 1998; Pickett-Heaps et al., 2011; Miller et al., 2014, 2016a; Karion
et al., 2016; Ishizawa et al., 2019). These tower-based studies provide a range of flux estimates from 14.8 to 19.5 Tg CH_4 per
year for Canada and 1.56 to 3.4 Tg CH_4 per year for the Hudson Bay Lowlands (HBL), a prominent wetland region in northern
Canada (e.g., Ishizawa et al., 2024; Miller et al., 2014; Pickett-Heaps et al., 2011; Thompson et al., 2017). Existing studies
using tower-based observations have also commented on inter-annual variability, though these studies disagree on whether
high-latitude fluxes have detectable year-to-year variability or a multi-year trend (Ishizawa et al., 2019; Sweeney et al., 2016;
Thompson et al., 2017; Ward et al., 2024). Sweeney et al. (2016) argues that there is no multi-decadal trend in CH_4 fluxes
using observations from Utqiagvik, Alaska, while inverse modeling studies by Thompson et al. (2017), Ishizawa et al. (2019),
and Ishizawa et al. (2024) identify inter-annual variability in wetland fluxes across high-latitude North America.

NASA scientists began collecting intensive aircraft-based greenhouse gas observations across Alaska in 2012, providing a
complement to the long-term tower-based network, and these aircraft campaigns have led to numerous studies on CH_4 fluxes
from that state. The authors of these studies quantify the magnitude and spatial distribution of fluxes using inverse modeling
(Chang et al., 2014; Miller et al., 2016b; Hartery et al., 2018; Sweeney et al., 2022). These estimates range from 1.48 to 2.6 Tg
CH_4 per year, a number similar to CH_4 flux totals from Canada's Hudson Bay Lowlands (Miller et al., 2016b; Hartery et al.,
2018; Chang et al., 2014; Sweeney et al., 2022). By contrast, the WETCHIMP process-based models span a much wider range
from 0.65 to 6.0 Tg CH_4 per year, a nearly nine-fold spread.

Several studies also leverage aircraft observations to conduct a detailed evaluation of flux processes and of numerous process-
based models, mostly from the WETCHIMP inter-comparison (Miller et al., 2016b; Hartery et al., 2018). For example, these
inverse modeling estimates indicate that tundra ecosystems contribute a disproportionate share of Alaskan CH_4 fluxes (often
$> 50\%$ of total Alaskan CH_4 fluxes despite their smaller areal extents). The North Slope alone accounts for $\sim 24\%$ of the total
statewide CH_4 fluxes, which is $\sim 20\%$ higher than the estimates of process-based models (Miller et al., 2016b; Hartery et al.,
2018). Collectively, these aircraft-based studies demonstrate that process-based flux estimates not only diverge substantially
from one another, but also from atmospheric constraints. The main driver of this divergence is how models represent wetland
extent and water table dynamics, and these factors have a salient influence on the magnitude and spatial distribution of high-
latitude wetland CH_4 fluxes estimated by process-based models (Miller et al., 2016b). Hartery et al. (2018) also argue that in
wetland soils, CH_4 fluxes are likely driven by near-surface soil temperature and moisture while fluxes from non-wetland soils
are more likely driven by temperature and moisture at greater depths, a difference in processes that may be key for effectively
modeling CH_4 fluxes across Alaska.

Since the early 2010s, the tower-based observing network has greatly expanded across high-latitude North America, pro-
viding a new opportunity to evaluate process-based models and to suggest future opportunities for improvement. In addition,
we can now compare two different process-based model ensembles, generated over a decade apart, to assess how process-
based estimates of high-latitude wetland fluxes have evolved over time (i.e., the WETCHIMP and GCP ensembles). In this
study, we use atmospheric CH_4 observations from tower sites to evaluate the GCP process-based models across high-latitude
North America. We specifically use four sets of analyses to compare atmospheric CH_4 observations and the GCP wetland flux
models with a goal of suggesting future improvements to these models. For each of these analyses, we run each GCP flux

estimate through an atmospheric transport model to simulate atmospheric CH_4, and we compare the results against available atmospheric CH_4 observations. First, we compare the GCP models across high latitudes against the WETCHIMP models and explore how process-based flux models have evolved over the past decade. Second, we examine how the GCP models vary in CH_4 flux magnitude and what potential factors might drive agreement or disagreement among the models. Third, we investigate differences in seasonal cycles across models that best match atmospheric observations versus models that show seasonal discrepancies with atmospheric observations. Lastly, we examine the spatial distribution of the CH_4 fluxes estimated by the GCP models and identify spatial patterns that appear to yield better agreement with the available atmospheric CH_4 data.

2 Data and methods

2.1 Atmospheric CH_4 Measurements

To better understand current wetland CH_4 fluxes, we compare GCP CH_4 flux estimates with 11 years of in situ tower data from the United States and Canada, spanning 2007 to 2017. We note that several previous studies have already used intensive aircraft campaigns to examine CH_4 fluxes across specific regions of Alaska and to evaluate process-based flux models in those regions (e.g., Miller et al., 2016b; Hartery et al., 2018; Chang et al., 2014; Sweeney et al., 2022). We build upon these existing studies by evaluating process-based models using long-term tower observation sites that are distributed cross both Canada and Alaska, and we compare and contrast our results with previous aircraft-based studies in the Results and discussion section (Sect. 3).

We also focus on the months of May through October each year. Wetland CH_4 fluxes are largest during these months, and many existing top-down studies have focused on these months for their analyses (e.g., Miller et al., 2016b; Chang et al., 2014; Pickett-Heaps et al., 2011). By contrast, the ratio of wetland fluxes to anthropogenic CH_4 emissions is much smaller in other months of the year across Alaska and Canada, making it more difficult to uniquely constrain wetland fluxes using atmospheric observations. The geographic domain of this study covers the high-latitude regions of North America, ranging from 40° N to 80° N and 170° W to 50° W.

The atmospheric data used in this study come from the NOAA Observation Package (ObsPack) CH_4 GlobalViewPlus v5.1 dataset (Di Sarra et al., 2023). There are 21 available tower sites within the study domain, and the towers provide a combination of continuous and flask measurements. We list a more detailed description of each tower site and its location in Table S1. In addition, we extract afternoon averages of the observations between 1pm and 6pm local time when the boundary layer is generally well-mixed, and we do this to reduce transport uncertainties in STILT. During this time of day, CH_4 measurements are arguably influenced by fluxes from a broader region than at night. By contrast, the atmosphere is usually stable in the morning and at night with lower boundary layer heights, making accurate atmospheric trace gas modeling challenging. As a result, we prioritize robust transport over full diurnal coverage, and this approach is similar to multiple existing top-down studies (e.g., Miller et al., 2014, 2016a; Karion et al., 2016; Ishizawa et al., 2024).

2.2 Wetland CH_4 flux model ensembles: GCP and WETCHIMP

The Global Carbon Project (GCP) includes global-scale wetland CH_4 flux models that use diverse hydrological and biogeochemical schemes (Zhang et al., 2025). The most recent GCP model ensemble includes 16 process-based models spanning the period from 2000 to 2020, though some models end earlier or later than 2020. A general description of these GCP models is provided in Table S2 and in Zhang et al. (2025). Each of the these models is run in two different ways: diagnostically and prognostically. The diagnostic runs for each model are constrained by a predefined inundation map from the product Wetland Area and Dynamics for Methane Modeling version 2 (WAD2Mv2), while the prognostic runs estimate the inundation internally using their own hydrological schemes such as soil moisture (Zhang et al., 2021). As a result, prognostic inundation is not observation-driven, and inter-model differences are driven by the hydrological scheme and climate forcing. Note that each GCP modeling group did not submit variables like soil carbon, and this fact limits our ability to diagnose disagreements in the CH_4 flux estimates from different models. More detailed descriptions of the current GCP model ensemble, including their approaches to wetland inundation and model parametrization can be found in Zhang et al. (2025).

In this study, we evaluate the 11 prognostic and 16 diagnostic models included in the GCP ensemble. Each of these models was run using two different meteorological reanalysis products to examine the effects of meteorological uncertainties on estimated CH_4 fluxes. These products include the Global Soil Wetness Project Phase 3 (GSWP3) and the Climate Research Unit (CRU) Time-Series 4.06 (Harris et al., 2022; Lange and Büchner, 2020). A recent study shows that the differences between these two climate-forcing datasets are negligible (Ito et al., 2023). Nevertheless, both datasets are included in this study to provide a comprehensive evaluation.

We also evaluate process-based wetland CH_4 fluxes using the Wetland and Wetland CH_4 Intercomparison of Models Project (WETCHIMP), which is designed to compare modeled monthly CH_4 fluxes across the globe between 1993 and 2004 (Melton et al., 2013; Wania et al., 2013). There are seven models available that provide CH_4 fluxes in the North America domain. These models are CLM4Me (Riley et al., 2011), DLEM (Tian et al., 2010), LPJ-Bern (Spahni et al., 2011), LPJ-WHyMe (Wania et al., 2010), LPJ-WSL (Hodson et al., 2011), ORCHIDEE (Ringeval et al., 2010), and SDGVM (Singarayer et al., 2011).

We regrid the GCP and WETCHIMP models into an uniform spatial resolution of $1°$ latitude by $1°$ longitude. This regridding process is performed using the "remapcon" function from the Climate Data Operators (CDO) software, which conserves the total fluxes of each model during interpolation (Schulzweida, 2023).

2.3 Anthropogenic CH_4 emissions

We include three distinct combinations of anthropogenic CH_4 flux products to highlight the variability and uncertainty in our analysis due to anthropogenic CH_4 fluxes. Bottom-up inventories such as Canada's National Inventory Report (NIR) estimate total anthropogenic fluxes of approximately 3.7 Tg of CH_4 per year (Scarpelli et al., 2021). In contrast, top-down inverse modeling and observation-constrained studies generally infer higher national totals, on the order of 5 to 7 Tg of CH_4 per year (e.g., Thompson et al., 2017; Lu et al., 2022; Scarpelli et al., 2021; Chan et al., 2020; MacKay et al., 2021; Ishizawa et al., 2024). Existing bottom-up and top-down studies show particularly large discrepancies in oil and gas producing regions of

western Canada (Ishizawa et al., 2024; Chan et al., 2020; MacKay et al., 2021; Baray et al., 2021). Collectively, the spread among bottom-up and top-down studies highlights the large uncertainty in Canadian anthropogenic CH_4 flux estimates and underscores the importance of exploring multiple flux products in our analysis.

150 We use three specific anthropogenic flux products and regrid them to a spatial resolution of $1°$ latitude by $1°$ longitude for the study domain, and we aggregate them to a monthly temporal resolution for 2007–2017:

1. CarbonTracker CH_4 2023 (Oh et al., 2023): CarbonTracker is an inverse modeling system designed to estimate CH_4 fluxes on a global scale (Oh et al., 2023).

2. A combination of the gridded U.S. Greenhouse Gas Inventory (Version 2) and a gridded inventory of Canada's an-
155 thropogenic CH_4 fluxes (Monforti Ferrario et al., 2021; Maasakkers et al., 2023; Scarpelli et al., 2021): Scarpelli et al. (2021) constructed a gridded Canadian anthropogenic flux inventory based on the Canadian National Inventory Report (NIR), the Canadian Greenhouse Gas Reporting Program (GHGRP), and other datasets to provide a detailed sectoral breakdown of fluxes. Similarly, Maasakkers et al. (2023) created a U.S. gridded inventory integrating data from the U.S. Environmental Protection Agency's (EPA) Greenhouse Gas Inventory (GHGI) to provide fluxes from different sectors.

160 3. The Copernicus Atmosphere Monitoring Service (CAMS) (Granier et al., 2019): CAMS is a global inverse modeling system that provides estimates of global atmospheric CH_4 fluxes and atmospheric mixing ratios. This product is derived from a combination of the EDGARv4.3.2 and Community Emissions Data System (CEDSv3) inventories, and the product includes estimates of fluxes from different source sectors (Granier et al., 2019).

2.4 Atmospheric modeling

165 We use the WRF-STILT (The Weather Research and Forecasting-Stochastic Time-Inverted Lagrangian Transport) to simulate the atmospheric transport of CH_4 fluxes, which has been widely used in numerous studies of regional greenhouse gas fluxes (e.g., Miller et al., 2016b; Henderson et al., 2015; McKain et al., 2015; Kort et al., 2010; Feng et al., 2023; Miller et al., 2014). STILT is a Lagrangian particle dispersion model that simulates atmospheric transport using an ensemble of tracer particles (Lin et al., 2003). For the setup here, the model releases those particles from each measurement site, and the particles travel
170 backward in time for 10 days following the wind fields in WRF meteorology. STILT uses these particle trajectories to calculate surface influence maps or footprints for each atmospheric CH_4 observation (Figure 1). These footprint maps have units of mixing ratio per unit flux (ppb per μmol m^{-2} s^{-1}) on a $1°$ by $1°$ grid. We can directly multiply these footprints by CH_4 fluxes from the process-based models to predict atmospheric CH_4 mixing ratios at each tower site. The footprints used in this study were generated as part of the NOAA CarbonTracker-Lagrange project and are available from 2007 to 2017, which defines our
175 study time frame (Hu et al., 2019).

Since CH_4 has an atmospheric lifetime of about 10 years, it can remain in the atmosphere and travel around the globe. To account for the large-scale movements of CH_4, we estimate CH_4 boundary conditions using CH_4 observations collected over the Pacific and Atlantic oceans, from high-altitude tower sites in the continental US, and from regular aircraft flights across the US and Canada. We use these observations to interpolate a curtain of CH_4 mixing ratios around the boundaries of the

180 model domain. For each STILT simulation, we sample from this boundary condition curtain based on the ending locations of the particle trajectories. This procedure thus accounts for CH_4 that enters the domain from other regions of the globe. The approach used here is identical to that used in numerous existing regional CH_4 studies (e.g., Miller et al., 2013, 2014, 2016a).

We note that the STILT particle trajectories used here from CarbonTracker-Lagrange do not include atmospheric oxidation processes. However, CH_4 oxidation by hydroxyl radicals likely has a small impact in our study given the short, 10-day time

185 frame of the regional STILT simulations used here. For example, Miller et al. (2013) argue that CH_4 mixing ratios decay less than 1 to 1.5 ppb over the first 2–3 days of STILT back trajectories given the average global-averaged lifetime of CH_4 of 7-11 years. This corresponds to less than 5% of the average modeled CH_4 mixing ratio enhancement relative to background in our study. The impact of OH in our study may be even smaller because OH mixing ratios are usually lower at high latitudes compared to the continental US. In addition, our estimated boundary conditions also account for long-range CH_4 oxidation

190 processes that occur upwind of our domain.

We combine the aforementioned modeling components using the following equation to compare atmospheric CH_4 observations with the STILT model predictions using the GCP flux models:

$$\boldsymbol{Z} \sim \mathbf{H}\big[\boldsymbol{s} + \boldsymbol{A} + \boldsymbol{B}\big] + \boldsymbol{b}. \tag{1}$$

where $\boldsymbol{Z}$ represents the atmospheric observations from the in situ towers across the US and Canada (dimensions $n \times 1$, where

195 n are the number of observations). $\mathbf{H}$ is a matrix of influence footprints assembled from the WRF-STILT model (dimensions $n \times m$, where m is the number of flux model grid boxes in space and time). Within the brackets, $\boldsymbol{s}$ refers to wetland CH_4 flux estimates from the process-based GCP or WETCHIMP models (dimensions $m \times 1$, Sect. 2.2), $\boldsymbol{A}$ refers to the anthropogenic CH_4 fluxes estimate from one anthropogenic product (dimensions $m \times 1$, Sect. 2.3), and $\boldsymbol{B}$ denotes biomass burning fluxes from the Global Fire Emissions Database (GFED v4.1) (Randerson et al., 2017) (dimensions $m \times 1$). The last variable, $\boldsymbol{b}$,

200 represents the CH_4 boundary condition (dimensions $m \times 1$).

We run STILT simulations both with and without the lake and reservoir emissions from Maasakkers et al. (2016), which contribute approximately 0.72 Tg of CH_4 per year in Canada from May to October. The WAD2M v2 inundation map (used by GCP models) represents vegetated wetlands only, but small lakes and ponds could still overlap because these features are difficult to distinguish from wetlands in satellite observations (Kyzivat and Smith, 2023). Adding a separate lake component

205 could therefore lead to partial double-counting of freshwater emissions and further increase the modeled CH_4 mixing ratios. Given this potential for double-counting, we present results without the additional lake term, while we acknowledge that adding lake and reservoir emissions would further increase modeled CH_4 mixing ratios.

Note that we only include observation sites in our analysis if those sites are predominantly influenced by CH_4 fluxes from wetlands. By contrast, we exclude urban sites and sites proximal to oil and gas operations. We specifically include sites where

210 the average ratio of modeled CH_4 from STILT using the GCP model mean to modeled CH_4 using the CAMS anthropogenic flux product is greater than 1.3 (Sect. 2.4, sect.2.2, sect.2.3). This screening means that the wetland contributions at each site are at least 30% higher than the likely influence of anthropogenic emissions. If we set a lower threshold, then we would
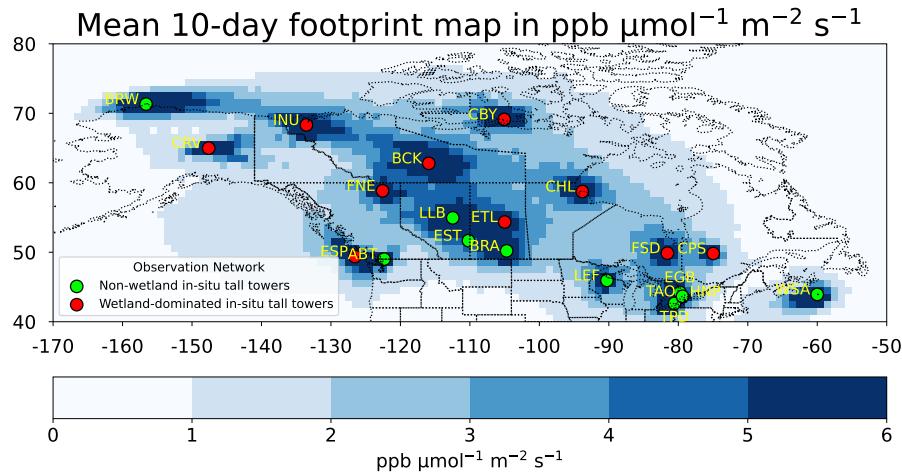
Figure 1. The US and Canadian atmospheric CH$_4$ observing network from 2007-2017. The figure also shows the WRF-STILT mean 10-day footprint map in ppb / (μmol m^{-2} s^{-1}) across the study domain of 40°N to 80°N and 170°W to 50°W, and footprints are evaluated from 2007-2017. Red circle dots show in situ tall tower sites from NOAA and Environmental Canada from the ObsPack GlobalViewPlus v5.1 dataset (Di Sarra et al., 2023). The lime-colored dots represent non-wetland sites, where the wetland-to-anthropogenic CH$_4$ concentration ratio is less than 1.3 (using anthropogenic emissions from the CAMS product). In contrast, the red-colored dots indicate wetland-dominated sites, where this ratio exceeds 1.3.

begin to include sites in urban and/or oil and gas producing areas. For example, the site with the next highest wetland-to-anthropogenic ratio is Abbotsford (ABT), which is an urban site near Vancouver, British Columbia. By contrast, if we set a
215   higher threshold, we would exclude the East Trout Lake (ETL) tower site, which is located in a sparsely populated wetland region of northern Saskatchewan. We focus on these sites because we aim to better quantify the contribution of wetlands to atmospheric CH$_4$ levels while minimizing the confounding effects of anthropogenic sources, the magnitudes of which are also uncertain. The ten sites that we include within this study are Churchill, Manitoba (CHL); Cambridge Bay, Nunavut (CBY); East Trout Lake, Saskatchewan (ETL); Estevan Point, British Columbia (ESP); Fort Nelson, British Columbia (FNE); Fraserdale,
220   Ontario (FSD); Inuvik, Northwest Territories (INU); Behchoko, Northwest Territories (BCK); Chapais, Quebec (CPS); and the Carbon in Arctic Reservoirs Vulnerability Experiment Tower, Fairbanks, Alaska (CRV) (see Table S1 for additional details). The remaining sites that are not included in this analysis are towers in urban environments (e.g., sites in the Toronto and Vancouver metropolitan areas); towers close to oil and gas production in Alberta, Canada, or Prudhoe Bay, Alaska; towers that are frequently used as clean air background sites (e.g., Sable Island, Nova Scotia), and sites proximal to intensive agriculture.
225   These sites are indicated by the solid red circles in Fig. 1.

3 Results and discussion

In this section, we compare the modeled CH_4 mixing ratios using the GCP models to atmospheric observations. We use these comparisons to evaluate the magnitude, seasonality, and spatial distribution of the GCP flux models over 2007 to 2017. In each subsection, we also speculate on the possible reasons driving the agreement or disagreements that we see in our analyses.

3.1 Comparisons between the GCP and WETCHIMP models

The GCP model ensemble is an updated version of the earlier WETCHIMP inter-comparison over a decade ago (Melton et al., 2013; Wania et al., 2013). Overall, we find that, compared to the WETCHIMP models, the GCP models have a smaller flux magnitude with reduced inter-model spread and better inter-model agreement on the spatial distribution of fluxes within our study domain.

We find that the CH_4 flux estimates from the GCP models are much smaller across most of high-latitude North America compared to the WETCHIMP models. We calculate annual CH_4 flux totals for Canada using the 11 prognostic and 16 diagnostic GCP models with both climate forcing datasets (GSWP3 and CRU), and the uncertainty bars in Fig. 2 represent the standard deviation of the CH_4 flux estimates among models within the same group. The mean annual CH_4 flux total for Canada using the 11 prognostic GCP models with CRU is 14.19 ± 7.41 Tg CH_4 per year, and the mean using the 16 diagnostic models with CRU is 12.17 ± 5.48 Tg CH_4 per year (Fig. 2). In contrast, the Canadian annual CH_4 flux total using all the WETCHIMP models with CRU meteorology is a factor of more than ~ 1.5 higher than the prognostic and diagnostic GCP models, with flux estimates of 21.50 ± 15.12 Tg CH_4 per year (based on the standard deviations of models within the same group). In Alaska, the annual CH_4 flux total estimated by the 11 prognostic GCP models with CRU is 1.31 ± 0.85 Tg CH_4 per year, whereas the seven WETCHIMP models yield a higher value of 1.66 ± 2.02 Tg CH_4 per year. We notice that the annual Canadian CH_4 flux total for the LPJ-WHyMe model from WETCHIMP is 46.25 ± 5.88 Tg CH_4 per year (Fig. S5). We subsequently exclude this model and recalculate the annual CH_4 flux total using the other six WETCHIMP models, and evaluate whether or not it brings the flux estimates similar to the GCP models. However, the annual CH_4 flux total using the other six WETCHIMP models with CRU is 17.97 ± 12.59 Tg CH_4 per year, which is still about a factor of ~ 1.4 higher than the prognostic GCP models using CRU meteorology.

In addition, the annual CH_4 flux totals estimated by the WETCHIMP models are a factor of ~ 1.3 or higher than the GCP models in the two dominant high-latitude biomes across North America (tundra and boreal forests) (Fig. 2). Across the North American boreal forests and tundra, the annual CH_4 flux totals estimated by the 11 prognostic GCP models with CRU are 10.71 ± 5.73 and 1.64 ± 1.31 Tg CH_4 per year, respectively. In comparison, the annual CH_4 flux totals estimated by the seven WETCHIMP models in these two biomes are 16.62 ± 8.55 and 2.15 ± 1.34 Tg CH_4 per year, respectively.

We also find that the CH_4 fluxes estimated by the 11 prognostic GCP models result in much lower inter-model uncertainty compared to the seven WETCHIMP models, with smaller inter-model disagreement across Canada and southern Alaska. Here, we define the uncertainty among models as the standard deviation across the models of the mean wetland flux CH_4 in May-October. To evaluate model agreement on the spatial distribution of fluxes, we compare the inter-model uncertainty or the

standard deviation of flux estimates for each individual model grid box of the GCP and WETCHIMP models. Since each

260 WETCHIMP model identifies the inundation or wetland area differently, we compare these models with the prognostic GCP models (Melton et al., 2013). Note, however, that not all of the WETCHIMP modeling groups generated their own wetland or inundation maps prognostically, and some, like LPJ-Bern and LPJ-WHyMe, use a constant, prescribed wetland map. In Figure 3, darker shades represent higher inter-model uncertainty across these process-based models. We observe that the GCP models have much lighter shades across the study domain, indicating better inter-model agreement.

265 We further find that the WETCHIMP models generally exhibit seasonal cycles that are similar to the GCP models (Figs. S1a and S1b). Most WETCHIMP models estimate peak CH_4 fluxes across Alaska and Canada in July and August, except CLM4Me (which peaks in June) and LPX-Bern (which peaks in September). This result illustrates that the seasonal cycles of the GCP models have not changed markedly from the WETCHIMP models. The WETCHIMP models already showed relatively good agreement on the seasonal cycle of fluxes, but such agreement does not guarantee accuracy, and there remains

270 scope for improvement. Furthermore, the seasonal cycle of these model estimates is likely dependent on temperature, meaning that it is arguably easier to model than other features that depend on more complex processes.

The overall reduction in inter-model uncertainties from WETCHIMP to GCP may relate to how the models estimate wetland distribution. Different WETCHIMP model yield very different estimates of maximum wetland extent – from 2.7 to 36.4×10^6 km^2 for the global extra-tropics ($> 35°$N), depending upon the model. Melton et al. (2013) explain that several WETCHIMP

275 models use a binary approach to identify wetland areas, where individual model grid boxes are either 100% wetland or 0% wetland, and these models tend to have $\sim 3-4$ times greater wetland area compared to other models (Fig. 2 and Table 2 in (Melton et al., 2013)). Other WETCHMIMP models were parameterized to match remote sensing estimates of wetland or open water extent. In contrast to WETCHIMP, the GCP model ensemble also includes diagnostic experiments in which all modeling groups used the WAD2M v2 inundation map. These efforts to create a standardized, diagnostic map of wetland extent may

280 have also influenced the prognostic GCP experiments, and modeling groups may have tuned or modified their setup to be more consistent with the diagnostic model simulations. In addition, the lower magnitude of CH_4 fluxes estimated by the GCP models (compared to the WETCHIMP models) is partly attributed to efforts by the GCP modeling group to reduce double-counting of freshwater areas (e.g., lakes and ponds) in WAD2M v2 (Zhang et al., 2021).

Note that the GCP models show lower flux magnitude and reduced inter-model spread in Canada, even when using the

285 subset of models that are common to both WETCHIMP and GCP. For diagnostic GCP runs, the overlapping model subsets with WETCHIMP are LPX-Bern (a newer version of LPJ-Bern), DLEM, ORCHIDEE, LPJ-wsl, SDGVM. For prognostic GCP runs, the common models include LPX-Bern, ORCHIDEE, LPJ-wsl, and SDGVM. Using these shared models, we find that the mean annual flux total from the WETCHIMP models is roughly 4 Tg CH_4 per year higher than the matched GCP ensemble mean, whereas in Alaska WETCHIMP is 0.11 Tg CH_4 lower (Figs. S5 and S6). In addition, we also find that the

290 GCP ensemble exhibits lower inter-model spread in Canada and broadly similar or lower spread in Alaska (Fig. S9). As a result, these analyses lead to the conclusion that the GCP ensemble is more tightly constrained than WETCHIMP over Canada when the same models are compared.
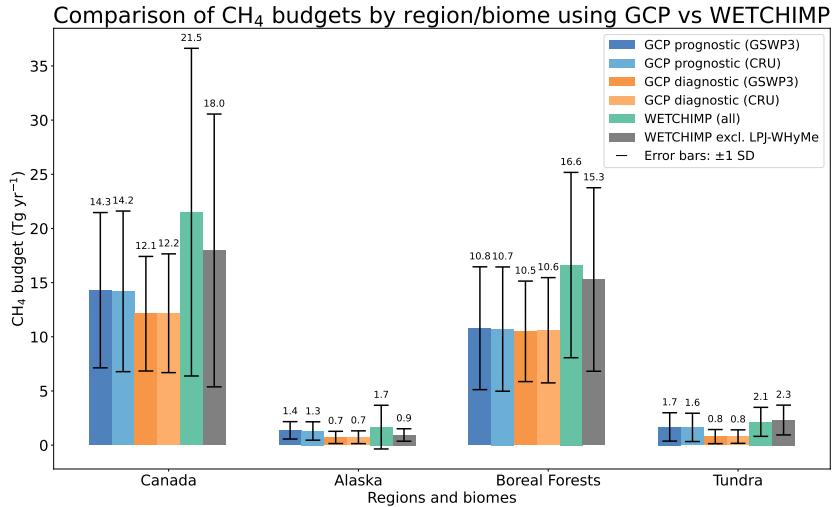
Figure 2. Annual CH$_4$ flux totals across Canada, Alaska, and several biomes. The four bars on the left of each region or biome represent the 2 different climate forcing data (GSWP3 and CRU) and prognostic versus diagnostic types for the GCP models. The green bar shows the mean annual CH$_4$ flux total using all WETCHIMP models, and the gray bar denotes the mean flux total excluding the LPJ-WHyMe model. The uncertainty bars represent the standard deviation of the CH$_4$ flux estimates among models within the same group. The unit of the annual wetland CH$_4$ flux totals is Tg CH$_4$ per year.

The reduced inter-model spread indicates greater consistency among the current GCP model outputs relative to WETCHIMP; however, reduced spread alone does not indicate improved accuracy. In the following sections, we compare the GCP and
295   WETCHIMP models with atmospheric observations to gauge whether the GCP models are indeed more skilled at capturing CH$_4$ fluxes across high-latitude North America.

3.2 Flux magnitude

We find that even though the mean wetland CH$_4$ fluxes of the GCP models are about a factor of two lower than the WETCHIMP models across northern North America, most of them are still likely an overestimate by a factor of two or more compared to
300   atmospheric CH$_4$ observations (Fig. 4). Note that we exclude lake and reservoir emissions from the following results because adding these emissions could double-count existing freshwater sources already represented in WAD2M v2 and further increase the modeled CH$_4$ mixing ratios relative to our current results.

We evaluate the magnitude of the GCP models by comparing modeled mixing ratios from STILT against observations at the tower sites. Specifically, we divide modeled CH$_4$ mixing ratios using wetland fluxes from the GCP models by the observed
305   increments, shown in Fig. 4. The modeled wetland CH$_4$ mixing ratios are calculated by passing each of the GCP models through WRF-STILT. The observed increments are calculated as the atmospheric CH$_4$ observations minus factors unrelated to wetlands – the CH$_4$ boundary condition and the contributions of anthropogenic and biomass burning fluxes at the observation sites. In Fig. 4, we compare the magnitude of the modeled wetland CH$_4$ mixing ratios and the observed increments at each

Figure 3. The inter-model standard deviation for each individual model grid box, calculated using the 11 prognostic GCP models (top) and WETCHIMP models (bottom). The inter-model uncertainty is higher for the WETCHIMP models than the GCP models. All fluxes have units μmol m^{-2} s^{-1}.

wetland-dominated in situ tower site across high-latitude North America. A factor larger than one means that the mixing ratios
310   of modeled wetland CH$_4$ using the GCP models are higher than the observed increments. By contrast, the gray dashed line at
the y-axis equal to 1 indicates a perfect alignment between the modeled wetland CH$_4$ mixing and the observed increment. The
error bars in Fig. 4 reflect the range of results when we use different anthropogenic flux estimates in the calculations (Sect.
2.3). Note that CH4MOD, DLEM, LPJ-GUESS, TEM-MDM, and TRIPLEX-GHG only have diagnostic simulations and not
prognostic simulations, and their diagnostic comparisons are represented exclusively by orange bars.
315   Interestingly, this result is not geographically uniform across high-latitude North America; the GCP models, when passed
through the WRF-STILT transport model, overshoot observations at towers in the boreal zone but not at towers in the Arctic
(Fig. S4). This result parallels earlier studies that use intensive aircraft campaign data from specific regions of Alaska. For
example, Miller et al. (2016b) estimate CH$_4$ fluxes over Alaska's North Slope that exceed most process-based model estimates
but find substantially lower fluxes than the model estimates across interior boreal and subarctic southeastern Alaska. Similarly,
320   Hartery et al. (2018) emphasize the disproportionately large contribution of Arctic Alaska to the state's total CH$_4$ fluxes, though
they do not explicitly compare their results with process-based models. Global inverse models, like those included in the most

recent Global Carbon Project CH_4 report, further reiterate these results; most yield lower wetland CH_4 fluxes across global high latitudes compared to process-based models, including across Russia, Europe, Canada, and the US (Saunois et al., 2025).

We also note that anthropogenic CH_4 fluxes pose an enormous challenge for isolating and quantifying CH_4 fluxes from wetlands, even at very remote observation sites in Canada and Alaska. The vertical bars in Fig. 4 indicate uncertainties in the results due to uncertain anthropogenic fluxes, and we observe a broad spectrum of values depending on which anthropogenic CH_4 flux estimate we use. For example, modeled mixing ratios from STILT using the GCP CH_4 model CLASSIC run prognostically are anywhere between $\sim$2.5 times higher than the observed increment to $\sim$6 times higher, depending on the choice of anthropogenic flux product. As a result, we cannot precisely constrain the optimal magnitude of wetland fluxes. These uncertainties notwithstanding, our findings still suggest that wetland fluxes estimated by the 11 prognostic and 16 diagnostic models are often higher than implied by atmospheric observations.

It is difficult to determine the specific causes that drive model disagreements over the magnitude of wetland CH_4 fluxes. However, these variations are more likely influenced by factors such as soil carbon or by the simplicity/complexity of the model structure rather than by disagreements over the effects of temperature on fluxes. We do not have a comprehensive set of modeled environmental variables (e.g., soil carbon) to conduct a systematic examination of all sources of uncertainty. However, the available model outputs allow us to reason through some key contributors to these uncertainties, such the relationships between fluxes and temperature (i.e., estimated Q_{10} values) and the effects of using a common diagnostic inundation map versus prognostically generated inundation.

To explore the temperature sensitivities of each GCP model, we fit a Q_{10} curve for each GCP model (Figs. S12 – S13). The Q_{10} parameter represents the sensitivity of wetland CH_4 fluxes to a $10°C$ increase in temperature, which provides insight into how strongly each model responds to temperature changes. A higher Q_{10} value indicates that the flux estimates are more prone to change with temperature variations. Our analysis indicates a large variation in temperature sensitivity across the prognostic and diagnostic GCP models, but there is not a strong relationship between the magnitude of wetland CH_4 fluxes estimated by these models and the estimated Q_{10} values (Figs. S12 – S13). As a result, Q_{10} does not seem to be the most important contributor driving differences in the flux magnitude of the GCP models.

We also find that uncertainties in wetland area and inundation likely contribute to but are not the primary cause of these disagreements in flux magnitude. For example, the prognostic and diagnostic models usually yield a similar magnitude of fluxes, despite of the fact that these different experiments do not use the same inundation estimates (Fig. 2). For Canada, the average total flux from the prognostic models is similar to the diagnostic models – 14.19 and 12.17 Tg per year, respectively (using GSWP3 meteorology). Similarly, the average total flux from the prognostic versus diagnostic models is nearly identical for the boreal forest biome. In some regions, the diagnostic models show greater agreement on the total annual flux than the prognostic models, but in other regions, the prognostic and diagnostic models show similar levels of inter-model agreement (Fig. 2).

Interestingly, we find models with simpler flux calculations yield flux magnitudes that agree more with atmospheric observations compared to those using more complex equations. GCP models such as LPJ-wsl, SDGVM, and JULES produce smaller flux magnitudes, and each of these models uses simple approaches to simulate CH_4 fluxes. For example, these models rely only
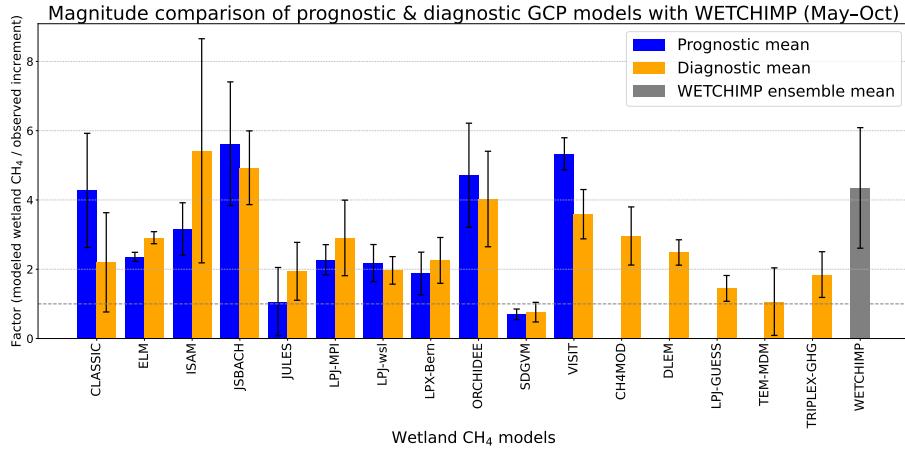
Figure 4. Comparisons between modeled mixing ratios from STILT against observations at the tower sites. The y-axis has values range from 0 to 9, representing the ratio between the modeled wetland CH_4 mixing ratios using the GCP and WETCHIMP models and the observed increment. We define the observed increment as the difference between atmospheric CH_4 observations and the sum of the boundary CH_4 levels, modeled anthropogenic CH_4 mixing ratios, and modeled biomass burning CH_4 mixing ratios. A value of 1 on the y-axis indicates perfect agreement between the modeled wetland CH_4 mixing ratios and the observed increment.

on net fluxes without accounting for specific transport pathways (e.g., ebullition, diffusion, or plant-mediated transport) (Zhang et al., 2025). In contrast, models such as VISIT, JSBACH, and ISAM have the largest flux magnitudes, and each of these models employs more complex equations that include multiple components of CH_4 fluxes, such as gross production, oxidation, and consumption. These models also simulate explicit transport pathways like ebullition, diffusion, and plant-mediated transport, alongside layered soil temperature schemes for temperature sensitivity (Zhang et al., 2025). Models with more complex representations generally require additional input data to provide more accurate flux estimates. Thus, in data-sparse regions, added process detail could potentially amplify input and parameter uncertainty and enlarge the flux spread.

3.3 Seasonality

We find that models more consistent with atmospheric observations have a distinct seasonal peak in wetland CH_4 fluxes in July and August. In contrast, models that do not agree well with atmospheric observations have a flatter seasonal cycle.

To evaluate these differences, we compare the correlation between atmospheric CH_4 observations and STILT simulations using each of the different GCP models (Fig. 5). We specifically use this analysis to explore which GCP models better capture the seasonal and spatial variability of CH_4 fluxes across our model domain. First, we calculate R^2 values for each model using a two-predictor regression model. In each regression, the first predictor variable represents modeled CH_4 mixing ratios due to wetlands using one of the GCP models, and the second predictor variable represents modeled CH_4 mixing ratios due to different anthropogenic flux products plus biomass burning from GFED (Sects. 2.3 and 2.4). The regression will scale the magnitude of the STILT model outputs to optimally match atmospheric observations. As a result, this analysis is not very sensitive to the
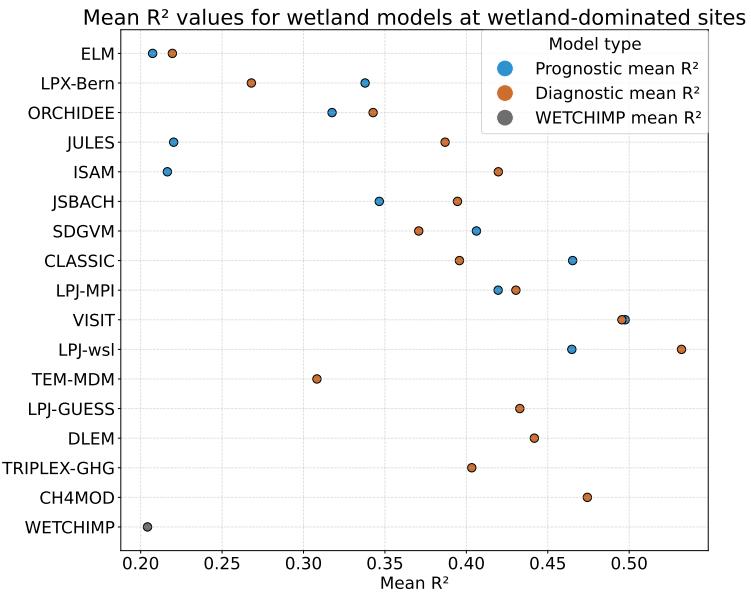
Figure 5. The correlation R^2 between modeled CH_4 mixing ratios using the process-based models and atmospheric observations. Blue dots represent the mean R^2 value for prognostic models across different climate forcing data and anthropogenic products. Orange dots represent the mean R^2 value for diagnostic models across different climate forcing data and anthropogenic products. The gray dot represents the mean R^2 value for the WETCHIMP models across different anthropogenic products. The y-axis lists all the prognostic and diagnostic GCP models and WETCHIMP models, and the x-axis shows the R^2 range for these GCP and WETCHIMP models.

absolute magnitude of the original flux estimates. Instead, the overall fit of each regression is more likely a reflection of the
375   seasonal and spatial patterns in the wetland, anthropogenic, and biomass burning flux estimates; GCP flux estimates with more
accurate seasonal and spatial variability will more likely yield higher correlation coefficients (R^2 values). Figure 5 depicts the
mean R^2 values for 16 GCP diagnostic wetland models and 11 GCP prognostic wetland models. Each model has a mean R^2
value that is averaged from the two climate forcing data (GSWP3 and CRU) and three anthropogenic flux products. These
results highlight the large variability in R^2 values across different GCP models. As shown in Fig. S7, model comparisons using
380   Root Mean Squared Error (RMSE) are identical to those using R^2, a result that further reinforces the discussion here.

       Based on this analysis, we categorize each of the diagnostic and prognostic GCP models into three groups based on how
they agree with atmospheric observations. By grouping the models, we can look for common patterns that separate models
that exhibit high R^2 values from those that exhibit lower R^2 values. Models with R^2 values greater than 0.4 are grouped into
the high R^2 group (represented by blue lines in Figs. 6a and 6b), models with R^2 values between 0.3 and 0.4 are classified as
385   the average R^2 group (represented by green lines in Figs. 6a and 6b), and models with R^2 values below 0.3 are considered as
the low R^2 group (represented by red lines in Figs. 6a and 6b). Although these cut-offs are inherently subjective, they offer a
practical framework for grouping the models and result in a similar number of models within each group.

Across the high and average R^2 groups, CH$_4$ fluxes exhibit a clear seasonal cycle, and we find that approximately 60–70% of the total fluxes from these models during the period of May to October occur during the peak summer season (June, July, and August). In these groups, the models capture the sharp rise and fall of the CH$_4$ fluxes, and they also show peak monthly percentages during July and August (Figs. 6a and 6b). This pattern aligns with the results of aircraft inversion studies that report a pronounced midsummer maximum Miller et al. (2016b); Chang et al. (2014). The low R^2 models display a much flatter seasonal pattern. These models do not capture the pronounced summer peaks observed in the high and average groups, suggesting that they may not fully capture seasonal variations in wetland fluxes.

The relationships between CH$_4$ fluxes and temperature may explain some, though not all, of the differences in seasonality among the GCP models. In our study, diagnostic SDGVM, diagnostic LPJ-MPI, diagnostic JULES, and diagnostic ISAM are the models that have high and average R^2 values (>0.35), and both have estimated Q$_{10}$ values greater than three, indicating a high sensitivity of their fluxes to temperature changes (Figs. S12 – S13). Moreover, models in the low R^2 group (<0.30) have estimated Q$_{10}$ values below 2, resulting in weaker temperature-driven mean fluxes (Figs. S12 – S13). This result shows that temperature relationships can explain at least some differences in the seasonality of the diagnostic GCP models. By comparison, existing empirical studies find a range of Q$_{10}$ values for wetlands in the Arctic region. Cao et al. (1996) suggest that a Q$_{10}$ value of 2 is calculated using a simple temperature response model, but Ito (2019); Walter and Heimann (2000) compute the Q$_{10}$ values of 3.85 and 6 using a more complicated mechanistic temperature response model. In addition, another study finds that the composition of wetlands can also yield different Q$_{10}$ values across the Arctic region. Specifically, M. Lupascu and Pancost (2012) find that wetlands that contain more Sphagnum moss can result in a Q$_{10}$ value of 8 or higher. These studies show that Q$_{10}$ values can be highly dynamic in high-latitude regions, and a Q$_{10}$ value of 6 does not necessarily mean that the temperature response model is wrong. We also examine the relationship between mean R^2 and Q$_{10}$ across models, but we find no consistent association between the two variables (Fig. S13).

Interestingly, we find that for 64% (7/11) of the models, the diagnostic version of the model yields a better fit (R^2) against atmospheric observations compared to the prognostic version of the model (Fig. 5. Prognostic versions of CLASSIC, SDGVM, LPX-Bern, and VISIT have better R^2 values compared to diagnostic versions). This result suggests that the better-performed diagnostic models may also reflect their reliance on a consistent inundation product, which potentially gives them the advantage in this evaluation framework over the more mechanistic prognostic models. In addition, process-based models with simpler and more deterministic formulations tend to produce smaller flux magnitudes and higher R^2 values compared to more complex models. This result indicates that simple formulations can effectively capture regional-to-continental flux patterns as those more complicated models (e.g., Miller et al., 2014, 2016b). However, more sophisticated process representations may become increasingly important for simulating finer-scale spatial structure or higher-frequency temporal variability of CH$_4$ fluxes.

We also find that the GCP models result in higher R^2 values and lower errors compared to the WETCHIMP models, both when comparing overlapping subsets of models and when considering their respective multi-model ensembles (Fig. S10 and S11). The ensemble of all WETCHIMP models yields a R^2 of 0.20 and an RMSE of 13.2 ppb. In contrast, the ensemble of all prognostic GCP models shows a R^2 of 0.35 with an RMSE of 11.9 ppb, while the ensemble of all diagnostic GCP models
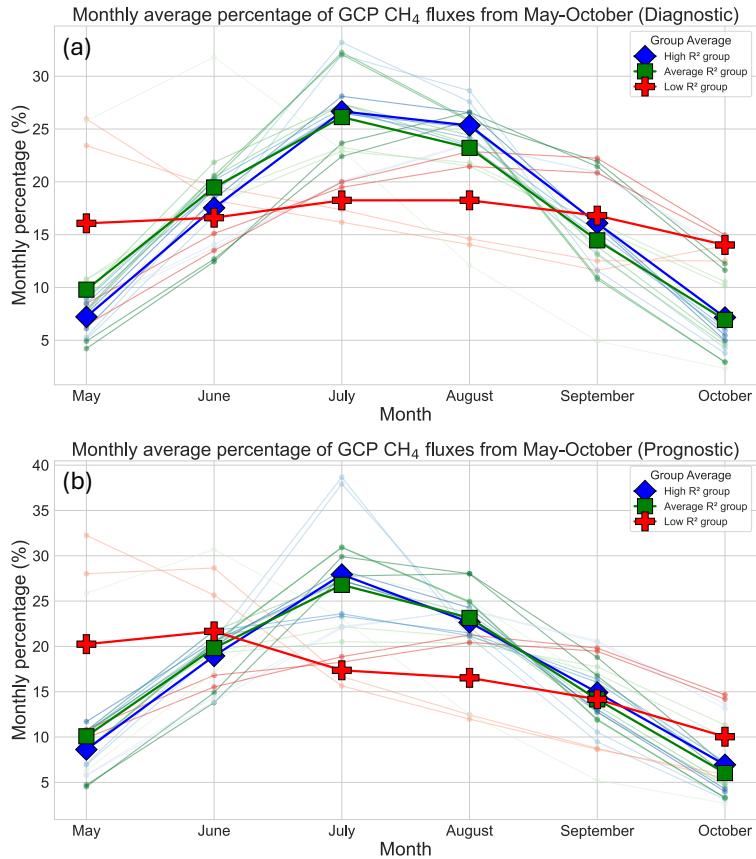
Figure 6. The seasonal cycles of the diagnostic GCP models (a) and prognostic GCP models (b) from 2007-2017. The blue, green, and red lines each represent the GCP models that have the highest, average, low R^2 values with atmospheric observations. The x-axis represents the months from May to October throughout 2007-2017, and y-axis denotes the percentages of CH_4 fluxes that occur within that month.

gives a R^2 of 0.39 with an RMSE of 11.5 ppb. These results demonstrate a clear improvement over the earlier WETCHIMP models, at least in comparisons with atmospheric observations.

3.4 Spatial distribution

We find that prognostic models that are most consistent with atmospheric observations concentrate their fluxes near the Hudson Bay Lowlands (Fig. 7a). In contrast, prognostic models with the lowest R^2 values focus their fluxes outside this key region (Fig. 7c). We focus this section on the prognostic models because the diagnostic models use the same inundation map and therefore exhibit similar spatial flux patterns. Similar to the previous analysis of seasonality, we group the prognostic models into three categories (high, average, low) depending on their R^2 values when compared against atmospheric observations. A Principal Component Analysis (PCA) highlights common spatial patterns among the models in each different group (e.g., Wold

et al., 1987; Jolliffe, 1986; Delwiche et al., 2021). The percentage of variance explained by the first principal component (PC1) shows the degree of spatial patterns shared among models in each group, and this percentage captures how consistently the models agree in their spatial flux distributions across grid boxes within the study domain. We use the PC1 explained variance as a measure of within-group spatial coherence that quantifies how much of the between-model variance in a group is captured by a single grid. We find that models in the high R^2 group have a PC1 explaining 63.5% of the variance, followed by the average R^2 group with 50.1%, and the low R^2 group with 68.9% explained variance. Although the low R^2 group shows the highest explained variance, this number does not necessarily indicate that the models in this group are more accurately capturing the true spatial patterns of the CH_4 fluxes compared to those in other groups.

We find notable common spatial features among the models in the high R^2, as seen in the PCA analysis. LPJ-wsl and CLASSIC have the highest R^2 values, and these models consistently concentrate their CH_4 fluxes in the Hudson Bay Lowlands. In contrast, JULES, ISAM, and ELM are the models with lower R^2 values. These models show large spatial discrepancies in critical wetland regions such as the Hudson Bay Lowlands, and they tend to concentrate fluxes outside of these key regions, particularly in the Great Lakes region of Canada.

An important caveat of this result is that the long-term observation network is sensitive to fluxes from some regions of high-latitude North America but not others (Fig. 1). The PCA analysis itself is unweighted, and our interpretation of spatial patterns (based on the R^2 metric) is necessarily influenced by regions with stronger observational coverage. We also note that none of the atmospheric observing towers are directly located in the Hudson Bay Lowlands, but the STILT footprints shown in Fig. 1 indicates that the network is sensitive to CH_4 fluxes from the broader region, allowing us to draw conclusions about the spatial distribution of fluxes in and around the Hudson Bay Lowlands.

4 Conclusions

This study highlights areas of convergence and disagreement among state-of-the-art process models of wetland CH_4 fluxes. We compare the estimates with atmospheric CH_4 observations between May and October in high-latitude North America. In the first section of the paper, we find that GCP models have a much smaller flux magnitude and lower inter-model uncertainty across North America compared to a previous model inter-comparison (WETCHIMP). The GCP models, when passed through an atmospheric transport model, are more consistent with atmospheric CH_4 observations compared to the WETCHIMP models, though we argue that the current GCP model ensemble is still too high across much of Canada and Alaska. In the second section of the study, we find that process-based CH_4 models that are most consistent with atmospheric observations based on our R^2 analysis exhibit the highest percentage of fluxes in July and August relative to other months and have a sharper seasonal cycle. These process-based models also concentrate their fluxes near the Hudson Bay Lowlands while less skilled models often concentrate fluxes further south near the Great Lakes.

Overall, this study highlights the opportunities to improve current process-based models of wetland CH_4 fluxes. Key areas for improvement include addressing (1) uncertainties in inundation or wetland extent and (2) improving estimated maps of soil carbon, though the latter factor was difficult to evaluate this study. We find that prognostic models show greater room for
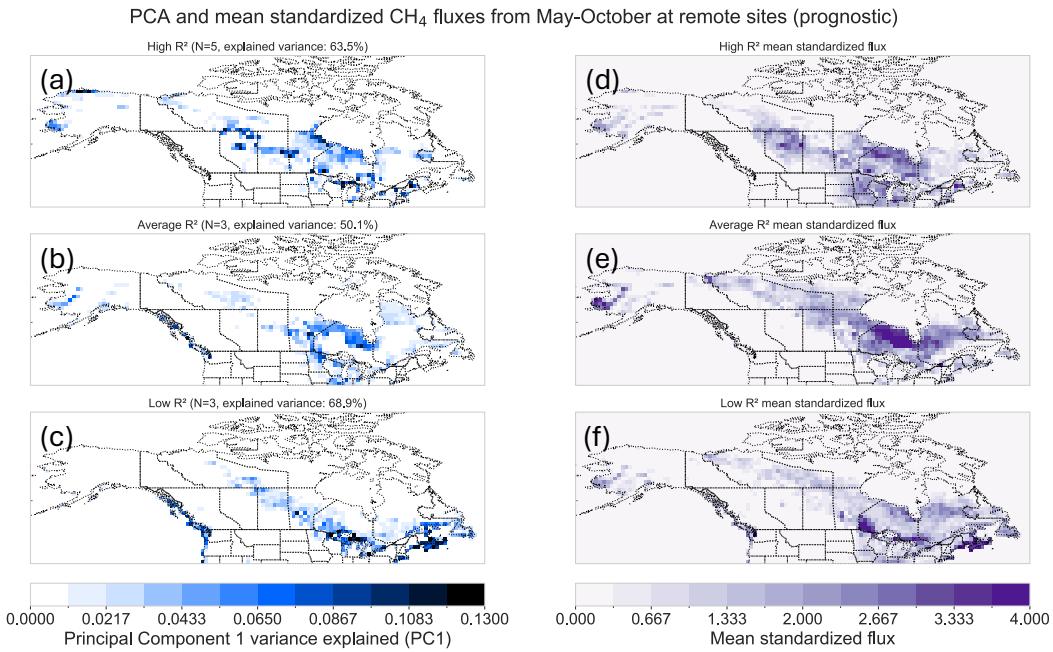
Figure 7. The PCA results and mean standardized CH$_4$ fluxes for the prognostic GCP models, run separately for each group of models – the high (a and d), average (b and e), and low (c and f) R^2 groups. The unit for PCA results is in explained variance by the first component (%), and darker (or more blue) shades represent better spatial agreements among the models within a same group.

improvement than the diagnostic models because they show less agreement with atmospheric observations based on the R^2
465 and RMSE metrics. While diagnostic models benefit from consistent inundation maps, the development of better prognostic
models is nevertheless very important because these models can be used to project future trends in wetland extent or inundation,
which is critical for future projections of CH$_4$ fluxes under ongoing climate change. Overall, we argue that the bottom-up
modeling community had made large strides in reducing inter-model uncertainties, and these improvements are consistent with
atmospheric CH$_4$ observations based on our analysis using the STILT model. We note, however, that the GCP models are
470 global and drivers vary regionally, so these conclusions apply only to our domain and time period. With that said, there is still
an enormous need for further improvements in these models to advance understanding of high-latitude wetland CH$_4$ fluxes in
a changing climate.

Data availability. We received the wetland model estimates from Zhen Zhang and the GCP modeling team, and these datasets are available upon request from the GCP modeling team. The GlobalViewPlus CH_4 ObsPack v5.1 dataset is available at https://gml.noaa.gov/ccgg/obspack/citation.php?product=obspack_ch4_1_GLOBALVIEWplus_v5.1_2023-03-08.

475

The WRF-STILT footprints for North American CH_4 monitoring sites are available at https://gml.noaa.gov/aftp/products/carbontracker/lagrange/footprints/ctl-na-v1.1/. The North American Boundary Condition product is provided by the NOAA Earth System Research Laboratory, and the dataset is available at https://gml.noaa.gov/aftp/user/arlyn/naboundary/v20190806/ROBJ/. Guidance related to these datasets can be requested from Lei Hu (lei.hu@noaa.gov) and Kathryn McKain (Kathryn.McKain@noaa.gov).

480

The CAMS global emission inventory dataset is available from the Copernicus Atmosphere Data Store. DOI:https://doi.org/10.24381/1d158bec. CarbonTracker CT-CH4-2023 data are available from NOAA's Global Monitoring Laboratory. DOI:https://doi.org/10.25925/40jt-qd67. The gridded inventory of Canada's anthropogenic CH_4 fluxes is available from the Harvard Dataverse. https://doi.org/10.7910/DVN/CC3KLO. The gridded U.S. Greenhous Gas Inventory (Version 2) can be found on Zenodo. DOI: https://doi.org/10.5281/zenodo.8367082. The Global Fire Emissions Database, Version 4 (GFEDv4) is available through the Oak Ridge National Laboratory (ORNL)

485 Distributed Active Archive Center (DAAC). DOI: https://doi.org/10.3334/ORNLDAAC/1293.

Author contributions. HL and SMM designed the study and wrote the manuscript. FRV, MI, ZZ, BP, JRM, LF, ALGL, AC, ZH, DCG, DC, VY, and DH provided feedback and comments on the manuscript. LF, AC, ZH, and DCG provided modeling support. JRM provided the WETCHIMP models. ZZ, BP, and the GCP modeling team provided the prognostic and diagnostic process-based CH_4 flux models. FRV, MI, DEJW, and DC contributed to the collection and maintenance of Canadian in situ tower CH_4 measurements included in the NOAA

490 ObsPack data product. ALGL and DH provided valuable suggestions on the Q_{10} calculations.

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

Acknowledgements. This work is funded by NASA ABoVE grant (#80NSSC22K1246) and by an NSF CAREER award (#2237404). We thank Environment And Climate Change Canada (ECCC) and NOAA Global Monitoring Laboratory for providing the GLOBALVIEWplus CH_4 ObsPack v5.1 dataset that is important for the completion of this work. We also acknowledge the use of WRF-STILT footprints data that are produced as a part of the CarbonTracker-Lagrange project with the support from NOAA's Climate Program Office and NASA's

495 Carbon Monitoring System. We acknowledge the use of the NOAA Earth System Research Laboratory's North American Boundary Condition product for CH_4, and we thank Arlyn Andrews, Kathryn McKain, and all other collaborators for providing access to the dataset. We use anthropogenic CH_4 emissions from the 2020 CAMS global emission inventory. This work contains modified Copernicus Atmosphere Monitoring Service information [2020]. Neither the European Commission nor ECMWF is responsible for any use that may be made of

500 the Copernicus information or data it contains. The CarbonTracker CT-CH4-2023 results are provided by NOAA GML, Boulder, Colorado, USA from the website at https://gml.noaa.gov/ccgg/carbontracker-ch4/. We acknowledge the use of the fire emissions from the Global Fire Emissions Database version 4 (GFED4s) described in van der Werf et al. (2017), and we regrid that dataset for this work. We also thank the Global Carbon Project (GCP) modeling team for their invaluable contributions in developing the models. In addition, we specifically

thank the following individuals for their important work in building the GCP models: Benjamin Poulter, Philippe Ciais, Joe Melton, William Riley, David Beerling, Nicola Gedney, Peter Hopcroft, Akihiko Ito, Atul Jain, Fortunat Joos, Thomas Kleinen, Tingting Li, Xiangyu Liu, Paul Miller, Changhui Peng, Shushi Peng, Zhangcai Qin, Qing Sun, Hanqin Tian, Yi Xi, Wenxin Zhang, Qing Zhu, Qiuan Zhu, and Qianlai Zhuang.

505

References

Baray, S., Jacob, D. J., Maasakkers, J. D., Sheng, J.-X., Sulprizio, M. P., Jones, D. B. A., Bloom, A. A., and McLaren, R.: Estimating 2010–2015 anthropogenic and natural methane emissions in Canada using ECCC surface and GOSAT satellite observations, Atmospheric Chemistry and Physics, 21, 18 101–18 121, https://doi.org/10.5194/acp-21-18101-2021, 2021.

Bohn, T. J., Melton, J. R., Ito, A., Kleinen, T., Spahni, R., Stocker, B. D., Zhang, B., Zhu, X., Schroeder, R., Glagolev, M. V., Maksyutov, S., Brovkin, V., Chen, G., Denisov, S. N., Eliseev, A. V., Gallego-Sala, A., McDonald, K. C., Rawlins, M. A., Riley, W. J., Subin, Z. M., Tian, H., Zhuang, Q., and Kaplan, J. O.: WETCHIMP-WSL: intercomparison of wetland methane emissions models over West Siberia, Biogeosciences, 12, 3321–3349, https://doi.org/10.5194/bg-12-3321-2015, 2015.

Cao, M., Marshall, S., and Gregson, K.: Global carbon exchange and methane emissions from natural wetlands: Application of a process-based model, Journal of Geophysical Research: Atmospheres, 101, 14 399–14 414, https://doi.org/https://doi.org/10.1029/96JD00219, 1996.

Chan, E., Worthy, D. E. J., Chan, D., Ishizawa, M., Moran, M. D., Delcloo, A., and Vogel, F.: Eight-Year Estimates of Methane Emissions from Oil and Gas Operations in Western Canada Are Nearly Twice Those Reported in Inventories, Environmental Science & Technology, 54, 14 899–14 909, https://doi.org/10.1021/acs.est.0c04117, 2020.

Chang, R. Y.-W., Miller, C. E., Dinardo, S. J., Karion, A., Sweeney, C., Daube, B. C., Henderson, J. M., Mountain, M. E., Eluszkiewicz, J., Miller, J. B., Bruhwiler, L. M. P., and Wofsy, S. C.: Methane emissions from Alaska in 2012 from CARVE airborne observations, Proceedings of the National Academy of Sciences, 111, 16 694–16 699, https://doi.org/10.1073/pnas.1412953111, 2014.

Delwiche, K. B., Knox, S. H., Malhotra, A., Fluet-Chouinard, E., McNicol, G., Feron, S., Ouyang, Z., Papale, D., Trotta, C., Canfora, E., Cheah, Y.-W., Christianson, D., Alberto, M. C. R., Alekseychik, P., Aurela, M., Baldocchi, D., Bansal, S., Billesbach, D. P., Bohrer, G., Bracho, R., Buchmann, N., Campbell, D. I., Celis, G., Chen, J., Chen, W., Chu, H., Dalmagro, H. J., Dengel, S., Desai, A. R., Detto, M., Dolman, H., Eichelmann, E., Euskirchen, E., Famulari, D., Fuchs, K., Goeckede, M., Gogo, S., Gondwe, M. J., Goodrich, J. P., Gottschalk, P., Graham, S. L., Heimann, M., Helbig, M., Helfter, C., Hemes, K. S., Hirano, T., Hollinger, D., Hörtnagl, L., Iwata, H., Jacotot, A., Jurasinski, G., Kang, M., Kasak, K., King, J., Klatt, J., Koebsch, F., Krauss, K. W., Lai, D. Y. F., Lohila, A., Mammarella, I., Belelli Marchesini, L., Manca, G., Matthes, J. H., Maximov, T., Merbold, L., Mitra, B., Morin, T. H., Nemitz, E., Nilsson, M. B., Niu, S., Oechel, W. C., Oikawa, P. Y., Ono, K., Peichl, M., Peltola, O., Reba, M. L., Richardson, A. D., Riley, W., Runkle, B. R. K., Ryu, Y., Sachs, T., Sakabe, A., Sanchez, C. R., Schuur, E. A., Schäfer, K. V. R., Sonnentag, O., Sparks, J. P., Stuart-Haëntjens, E., Sturtevant, C., Sullivan, R. C., Szutu, D. J., Thom, J. E., Torn, M. S., Tuittila, E.-S., Turner, J., Ueyama, M., Valach, A. C., Vargas, R., Varlagin, A., Vazquez-Lule, A., Verfaillie, J. G., Vesala, T., Vourlitis, G. L., Ward, E. J., Wille, C., Wohlfahrt, G., Wong, G. X., Zhang, Z., Zona, D., Windham-Myers, L., Poulter, B., and Jackson, R. B.: FLUXNET-CH$_4$: a global, multi-ecosystem dataset and analysis of methane seasonality from freshwater wetlands, Earth System Science Data, 13, 3607–3689, https://doi.org/10.5194/essd-13-3607-2021, 2021.

Di Sarra, A. G., Zahn, A., Watson, A., Ankur Desai, Karion, A., Hoheisel, A., Leskinen, A., Arlyn Andrews, Jordan, A., Colomb, A., Kers, B., Scheeren, B., Baier, B., Viner, B., Stephens, B., Daube, B., Van Der Veen, C., Labuschagne, C., Myhre, C. L., Couret, C., Miller, C. E., Choong-Hoon Lee, Lunder, C. R., Plass-Duelmer, C., Plass-Duelmer, C., Gerbig, C., Sloop, C. D., Sweeney, C., Kubistin, D., Goto, D., Jaffe, D., Heltai, D., Lowry, D., Munro, D., Worthy, D., Dlugokencky, E., Kozlova, E., Gloor, E., Cuevas, E., Hintsa, E., Kort, E., Morgan, E., Nisbet, E., Obersteiner, F., Apadula, F., Meinhardt, F., Moore, F., Vitkova, G., Giordane A. Martins, Manca, G., Zazzeri, G., Brailsford, G., Forster, G., Santoni, G., Haeyoung Lee, Boenisch, H., Moossen, H., Timas, H., Matsueda, H., Kang, H.-Y., Huilin Chen, Lehner, I., Mammarella, I., Bartyzel, J., Elkins, J. W., Jaroslaw Necki, Pittman, J., Pichon, J. M., Mÿller-Williams, J., Jgor Arduini, Turnbull, J.,

545 Miller, J. B., Lee, J., Jooil Kim, Pitt, J., DiGangi, J. P., Lavric, J., Hatakka, J., Worsey, J., Holst, J., Lehtinen, K., Kominkova, K., McKain, K., Saito, K., Davis, K., Thoning, K., TÃrseth, K., Haszpra, L., SÃrensen, L. L., Gatti, L. V., Emmenegger, L., Sha, M. K., Menoud, M., Delmotte, M., Fischer, M. L., De Vries, M., Schumacher, M., Torn, M., Popa, M. E., Leuenberger, M., Heimann, M., Heimann, M., Steinbacher, M., Schmidt, M., De Mazière, M., Lindauer, M., Mölder, M., Martin, M. Y., Ko, M.-Y., Rothe, M., Heliasz, M., Marek, M. V., Ramonet, M., Stanisavljević, M., Lopez, M., Sasakawa, M., Miles, N., Laurent, O., Hermanssen, O., Trisolino, P., Cristofanelli,

550 P., Krummel, P., Shepson, P., Smith, P., Rivas, P. P., Bergamaschi, P., Keronen, P., Keeling, R., Langenfelds, R., Weiss, R., Leppert, R., De Souza, R. A. F., Piacentino, S., Richardson, S., Biraud, S. C., Conil, S., Morimoto, S., Aoki, S., O'Doherty, S., Zaehle, S., Platt, S. M., Prinzivalli, S., Wofsy, S., Nichol, S., Schuck, T., Lauvaux, T., Röckmann, T., Seifert, T., Biermann, T., Kneuer, T., Gehrlein, T., Machida, T., Laurila, T., Aalto, T., Monteiro, V., Kazan, V., Ivakhov, V., Joubert, W., Brand, W. A., Lan, X., Niwa, Y., and Loh, Z.: Multi-laboratory compilation of atmospheric methane data for the period 1983-2021, https://doi.org/10.25925/20230301, 2023.

555 Feng, L., Tavakkoli, S., Jordaan, S. M., Andrews, A. E., Benmergui, J. S., Waugh, D. W., Zhang, M., Gaeta, D. C., and Miller, S. M.: Inter-annual variability in atmospheric transport complicates estimation of US methane emissions trends, Geophysical Research Letters, 50, e2022GL100 366, https://doi.org/https://doi.org/10.1029/2022GL100366, e2022GL100366 2022GL100366, 2023.

Granier, C., Darras, S., van der Gon, H. D., Doubalova, J., Elguindi, N., Galle, B., Gauss, M., Guevara, M., Jalkanen, J.-P., Kuenen, J., Liousse, C., Quack, B., Simpson, D., and Sindelarova, K.: The Copernicus Atmosphere Monitoring Service global and regional emissions (April

560 2019 version), https://atmosphere.copernicus.eu/sites/default/files/2019-06/cams_emissions_general_document_apr2019_v7.pdf, copernicus Atmosphere Monitoring Service (CAMS) report, 2019.

Harris, I., Jones, P., and Osborn, T.: CRU TS4.06: Climatic Research Unit (CRU) Time-Series (TS) version 4.06 of high-resolution gridded data of month-by-month variation in climate (Jan. 1901 - Dec. 2021), NERC EDS Centre for Environmental Data Analysis, https://catalogue.ceda.ac.uk/uuid/e0b4e1e56c1c4460b796073a31366980/, 2022.

565 Hartery, S., Commane, R., Lindaas, J., Sweeney, C., Henderson, J., Mountain, M., Steiner, N., McDonald, K., Dinardo, S. J., Miller, C. E., Wofsy, S. C., and Chang, R. Y.-W.: Estimating regional-scale methane flux and budgets using CARVE aircraft measurements over Alaska, Atmospheric Chemistry and Physics, 18, 185–202, https://doi.org/10.5194/acp-18-185-2018, 2018.

Henderson, J. M., Eluszkiewicz, J., Mountain, M. E., Nehrkorn, T., Chang, R. Y.-W., Karion, A., Miller, J. B., Sweeney, C., Steiner, N., Wofsy, S. C., and Miller, C. E.: Atmospheric transport simulations in support of the Carbon in Arctic Reservoirs Vulnerability Experiment

570 (CARVE), Atmospheric Chemistry and Physics, 15, 4093–4116, https://doi.org/10.5194/acp-15-4093-2015, 2015.

Hodson, E. L., Poulter, B., Zimmermann, N. E., Prigent, C., and Kaplan, J. O.: The El Niño–Southern Oscillation and wetland methane interannual variability, Geophysical Research Letters, 38, https://doi.org/https://doi.org/10.1029/2011GL046861, 2011.

Hu, L., Andrews, A. E., Thoning, K. W., Sweeney, C., Miller, J. B., Michalak, A. M., Dlugokencky, E., Tans, P. P., Shiga, Y. P., Mountain, M., Nehrkorn, T., Montzka, S. A., McKain, K., Kofler, J., Trudeau, M., Michel, S. E., Biraud, S. C., Fischer, M. L., Worthy, D. E. J.,

575 Vaughn, B. H., White, J. W. C., Yadav, V., Basu, S., and van der Velde, I. R.: Enhanced North American carbon uptake associated with El Niño, Science Advances, 5, eaaw0076, https://doi.org/10.1126/sciadv.aaw0076, 2019.

Hugelius, G., Strauss, J., Zubrzycki, S., Harden, J. W., Schuur, E. A. G., Ping, C.-L., Schirrmeister, L., Grosse, G., Michaelson, G. J., Koven, C. D., O'Donnell, J. A., Elberling, B., Mishra, U., Camill, P., Yu, Z., Palmtag, J., and Kuhry, P.: Estimated stocks of circumpolar permafrost carbon with quantified uncertainty ranges and identified data gaps, Biogeosciences, 11, 6573–6593, https://doi.org/10.5194/bg-11-6573-

580 2014, 2014.

Ishizawa, M., Chan, D., Worthy, D., Chan, E., Vogel, F., and Maksyutov, S.: Analysis of atmospheric CH_4 in Canadian Arctic and estimation of the regional CH_4 fluxes, Atmospheric Chemistry and Physics, 19, 4637–4658, https://doi.org/10.5194/acp-19-4637-2019, 2019.

Ishizawa, M., Chan, D., Worthy, D., Chan, E., Vogel, F., Melton, J. R., and Arora, V. K.: Estimation of Canada's methane emissions: inverse modelling analysis using the Environment and Climate Change Canada (ECCC) measurement network, Atmospheric Chemistry and Physics, 24, 10 013–10 038, https://doi.org/10.5194/acp-24-10013-2024, 2024.

Ito, A.: Methane emission from pan-Arctic natural wetlands estimated using a process-based model, 1901–2016, Polar Science, 21, 26–36, https://doi.org/https://doi.org/10.1016/j.polar.2018.12.001, iSAR-5/ Fifth International Symposium on Arctic Research, 2019.

Ito, A., Li, T., Qin, Z., Melton, J. R., Tian, H., Kleinen, T., Zhang, W., Zhang, Z., Joos, F., Ciais, P., Hopcroft, P. O., Beerling, D. J., Liu, X., Zhuang, Q., Zhu, Q., Peng, C., Chang, K.-Y., Fluet-Chouinard, E., McNicol, G., Patra, P., Poulter, B., Sitch, S., Riley, W., and Zhu, Q.: Cold-season methane fluxes simulated by GCP-CH$_4$ models, Geophysical Research Letters, 50, e2023GL103 037, https://doi.org/https://doi.org/10.1029/2023GL103037, e2023GL103037 2023GL103037, 2023.

Jolliffe, I. T.: Principal Component Analysis and Factor Analysis, pp. 115–128, Springer New York, New York, NY, ISBN 978-1-4757-1904-8, https://doi.org/10.1007/978-1-4757-1904-8_7, 1986.

Karion, A., Sweeney, C., Miller, J. B., Andrews, A. E., Commane, R., Dinardo, S., Henderson, J. M., Lindaas, J., Lin, J. C., Luus, K. A., Newberger, T., Tans, P., Wofsy, S. C., Wolter, S., and Miller, C. E.: Investigating Alaskan methane and carbon dioxide fluxes using measurements from the CARVE tower, Atmospheric Chemistry and Physics, 16, 5383–5398, https://doi.org/10.5194/acp-16-5383-2016, 2016.

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J. G., Dlugokencky, E. J., Bergamaschi, P., Bergmann, D., Blake, D. R., Bruhwiler, L., Cameron-Smith, P., Castaldi, S., Chevallier, F., Feng, L., Fraser, A., Heimann, M., Hodson, E. L., Houweling, S., Josse, B., Fraser, P. J., Krummel, P. B., Lamarque, J.-F., Langenfelds, R. L., Le Quéré, C., Naik, V., O'Doherty, S., Palmer, P. I., Pison, I., Plummer, D., Poulter, B., Prinn, R. G., Rigby, M., Ringeval, B., Santini, M., Schmidt, M., Shindell, D. T., Simpson, I. J., Spahni, R., Steele, L. P., Strode, S. A., Sudo, K., Szopa, S., van der Werf, G. R., Voulgarakis, A., van Weele, M., Weiss, R. F., Williams, J. E., and Zeng, G.: Three decades of global methane sources and sinks, Nature Geoscience, 6, 813–823, https://doi.org/10.1038/ngeo1955, 2013.

Koffi, E. N., Bergamaschi, P., Alkama, R., and Cescatti, A.: An observation-constrained assessment of the climate sensitivity and future trajectories of wetland methane emissions, Science Advances, 6, eaay4444, https://doi.org/10.1126/sciadv.aay4444, 2020.

Kort, E. A., Andrews, A. E., Dlugokencky, E., Sweeney, C., Hirsch, A., Eluszkiewicz, J., Nehrkorn, T., Michalak, A., Stephens, B., Gerbig, C., Miller, J. B., Kaplan, J. O., Houweling, S., Daube, B. C., Tans, P., and Wofsy, S. C.: Atmospheric constraints on 2004 emissions of methane and nitrous oxide in North America from atmospheric measurements and a receptor-oriented modeling framework, Journal of Integrative Environmental Sciences, 7, 125 – 133, https://doi.org/10.1080/19438151003767483, 2010.

Kyzivat, E. D. and Smith, L. C.: A closer look at the effects of lake area, aquatic vegetation, and double-counted wetlands on Pan-Arctic lake methane emissions estimates, Geophysical Research Letters, 50, e2023GL104 825, https://doi.org/10.1029/2023GL104825, e2023GL104825 2023GL104825, 2023.

Lange, S. and Büchner, M.: ISIMIP2a atmospheric climate input data, https://doi.org/10.48364/ISIMIP.886955, 2020.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/https://doi.org/10.1029/2002JD003161, 2003.

Liu, L., Zhuang, Q., Oh, Y., Shurpali, N. J., Kim, S., and Poulter, B.: Uncertainty quantification of global Net methane emissions from terrestrial ecosystems using a mechanistically based biogeochemistry model, Journal of Geophysical Research: Biogeosciences, 125, e2019JG005 428, https://doi.org/https://doi.org/10.1029/2019JG005428, e2019JG005428 2019JG005428, 2020.

Lu, X., Jacob, D. J., Wang, H., Maasakkers, J. D., Zhang, Y., Scarpelli, T. R., Shen, L., Qu, Z., Sulprizio, M. P., Nesser, H., Bloom, A. A., Ma, S., Worden, J. R., Fan, S., Parker, R. J., Boesch, H., Gautam, R., Gordon, D., Moran, M. D., Reuland, F., Villasana, C. A. O., and Andrews, A.: Methane emissions in the United States, Canada, and Mexico: Evaluation of national methane emission inventories and 2010–2017 sectoral trends by inverse analysis of in situ (GLOBALVIEWplus CH$_4$ ObsPack) and satellite (GOSAT) atmospheric observations, Atmospheric Chemistry and Physics, 22, 395–418, https://doi.org/10.5194/acp-22-395-2022, 2022.

M. Lupascu, J. L. Wadham, E. R. C. H. and Pancost, R. D.: Temperature sensitivity of methane production in the permafrost active layer at Stordalen, Sweden: A comparison with non-permafrost northern wetlands, Arctic, Antarctic, and Alpine Research, 44, 469–482, https://doi.org/10.1657/1938-4246-44.4.469, 2012.

Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Turner, A. J., Weitz, M., Wirth, T., Hight, C., DeFigueiredo, M., Desai, M., Schmeltz, R., Hockstad, L., Bloom, A. A., Bowman, K. W., Jeong, S., and Fischer, M. L.: Gridded national inventory of U.S. methane emissions, Environmental Science & Technology, 50, 13 123–13 133, https://doi.org/10.1021/acs.est.6b02878, 2016.

Maasakkers, J. D., McDuffie, E. E., Sulprizio, M. P., Chen, C., Schultz, M., Brunelle, L., Thrush, R., Steller, J., Sherry, C., Jacob, D. J., Jeong, S., Irving, B., and Weitz, M.: A gridded inventory of annual 2012–2018 U.S. anthropogenic methane emissions, Environmental Science & Technology, 57, 16 276–16 288, https://doi.org/10.1021/acs.est.3c05138, pMID: 37857355, 2023.

MacKay, K., Lavoie, M., Bourlon, E., Atherton, E., O'Connell, E., Baillie, J., Fougère, C., and Risk, D.: Methane emissions from upstream oil and gas production in Canada are underestimated, Scientific Reports, 11, 8041, https://doi.org/10.1038/s41598-021-87610-3, 2021.

McKain, K., Down, A., Raciti, S. M., Budney, J., Hutyra, L. R., Floerchinger, C., Herndon, S. C., Nehrkorn, T., Zahniser, M. S., Jackson, R. B., Phillips, N., and Wofsy, S. C.: Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts, Proceedings of the National Academy of Sciences, 112, 1941–1946, https://doi.org/10.1073/pnas.1416261112, 2015.

Melton, J. R., Wania, R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., Bohn, T., Avis, C. A., Beerling, D. J., Chen, G., Eliseev, A. V., Denisov, S. N., Hopcroft, P. O., Lettenmaier, D. P., Riley, W. J., Singarayer, J. S., Subin, Z. M., Tian, H., Zürcher, S., Brovkin, V., van Bodegom, P. M., Kleinen, T., Yu, Z. C., and Kaplan, J. O.: Present state of global wetland extent and wetland methane modelling: conclusions from a model inter-comparison project (WETCHIMP), Biogeosciences, 10, 753–788, https://doi.org/10.5194/bg-10-753-2013, 2013.

Miller, S. M., Wofsy, S. C., Michalak, A. M., Kort, E. A., Andrews, A. E., Biraud, S. C., Dlugokencky, E. J., Eluszkiewicz, J., Fischer, M. L., Janssens-Maenhout, G., Miller, B. R., Miller, J. B., Montzka, S. A., Nehrkorn, T., and Sweeney, C.: Anthropogenic emissions of methane in the United States, Proceedings of the National Academy of Sciences, 110, 20 018–20 022, https://doi.org/10.1073/pnas.1314392110, 2013.

Miller, S. M., Worthy, D. E. J., Michalak, A. M., Wofsy, S. C., Kort, E. A., Havice, T. C., Andrews, A. E., Dlugokencky, E. J., Kaplan, J. O., Levi, P. J., Tian, H., and Zhang, B.: Observational constraints on the distribution, seasonality, and environmental predictors of North American boreal methane emissions, Global Biogeochemical Cycles, 28, 146–160, https://doi.org/https://doi.org/10.1002/2013GB004580, 2014.

Miller, S. M., Commane, R., Melton, J. R., Andrews, A. E., Benmergui, J., Dlugokencky, E. J., Janssens-Maenhout, G., Michalak, A. M., Sweeney, C., and Worthy, D. E. J.: Evaluation of wetland methane emissions across North America using atmospheric data and inverse modeling, Biogeosciences, 13, 1329–1339, https://doi.org/10.5194/bg-13-1329-2016, 2016a.

Miller, S. M., Miller, C. E., Commane, R., Chang, R. Y.-W., Dinardo, S. J., Henderson, J. M., Karion, A., Lindaas, J., Melton, J. R., Miller, J. B., Sweeney, C., Wofsy, S. C., and Michalak, A. M.: A multiyear estimate of methane fluxes in Alaska from CARVE atmospheric observations, Global Biogeochemical Cycles, 30, 1441–1453, https://doi.org/https://doi.org/10.1002/2016GB005419, 2016b.

Monforti Ferrario, F., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Lo Vullo, E., Solazzo, E., Olivier, J., and Vignati, E.: EDGAR v6.0 Greenhouse Gas Emissions, http://data.europa.eu/89h/97a67d67-c62e-4826-b873-9d972c4f670b, eDGAR database, European Commission, 2021.

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

Obu, J.: How much of the earth's surface is underlain by permafrost?, Journal of Geophysical Research: Earth Surface, 126, e2021JF006 123, https://doi.org/10.1029/2021JF006123, 2021.

665 Oh, Y., Bruhwiler, L., Lan, X., Basu, S., Schuldt, K., Thoning, K., Michel, S. E., Clark, R., Miller, J. B., Andrews, A., Sherwood, O., Etiope, G., Crippa, M., Liu, L., Zhuang, Q., Randerson, J., van der Werf, G., Aalto, T., Amendola, S., Andra, S. C., Andrade, M., Nguyen, N. A., Aoki, S., Apadula, F., Arifin, I. B., Arnold, S., Arshinov, M., Baier, B., Bergamaschi, P., Biermann, T., Biraud, S. C., Blanc, P.-E., Brailsford, G., Chen, H., Colomb, A., Couret, C., Cristofanelli, P., Cuevas, E., Chmura, L., Delmotte, M., Emmenegger, L., Esenzhanova, G., Fujita, R., Gatti, L., Guerette, E.-A., Haszpra, L., Heliasz, M., Hermansen, O., Holst, J., Di Iorio, T., Jordan, A., Jennifer, M.-W., Karion,
670 A., Kawasaki, T., Kazan, V., Keronen, P., Kim, S.-Y., Kneuer, T., Kominkova, K., Kozlova, E., Krummel, P., Kubistin, D., Labuschagne, C., Langenfelds, R., Laurent, O., Laurila, T., Lee, H., Lehner, I., Leuenberger, M., Lindauer, M., Lopez, M., Mahdi, R., Mammarella, I., Manca, G., Marek, M. V., Mazière, M. D., McKain, K., Meinhardt, F., Miller, C. E., Mölder, M., Moncrieff, J., Moosen, H., Moreno, C., Morimoto, S., Myhre, C. L., Nahas, A. C., Necki, J., Nichol, S., O'Doherty, S., Paramonova, N., Piacentino, S., Pichon, J. M., Plass-Dülmer, C., Ramonet, M., Ries, L., di Sarra, A. G., Sasakawa, M., Say, D., Schaefer, H., Scheeren, B., Schmidt, M., Schumacher, M., Sha,
675 M. K., Shepson, P., Smale, D., Smith, P. D., Steinbacher, M., Sweeney, C., Takatsuji, S., Torres, G., Tørseth, K., Trisolino, P., Turnbull, J., Uhse, K., Umezawa, T., Vermeulen, A., Vimont, I., Vitkova, G., Wang, H.-J. R., Worthy, D., and Xueref-Remy, I.: CarbonTracker CH$_4$ 2023, https://doi.org/10.25925/40JT-QD67, 2023.

Pickett-Heaps, C. A., Jacob, D. J., Wecht, K. J., Kort, E. A., Wofsy, S. C., Diskin, G. S., Worthy, D. E. J., Kaplan, J. O., Bey, I., and Drevet, J.: Magnitude and seasonality of wetland methane emissions from the Hudson Bay Lowlands (Canada), Atmospheric Chemistry and Physics,
680 11, 3773–3779, https://doi.org/10.5194/acp-11-3773-2011, 2011.

Poulter, B., Bousquet, P., Canadell, J. G., Ciais, P., Peregon, A., Saunois, M., Arora, V. K., Beerling, D. J., Brovkin, V., Jones, C. D., Joos, F., Gedney, N., Ito, A., Kleinen, T., Koven, C. D., McDonald, K., Melton, J. R., Peng, C., Peng, S., Prigent, C., Schroeder, R., Riley, W. J., Saito, M., Spahni, R., Tian, H., Taylor, L., Viovy, N., Wilton, D., Wiltshire, A., Xu, X., Zhang, B., Zhang, Z., and Zhu, Q.: Global wetland contribution to 2000–2012 atmospheric methane growth rate dynamics, Environmental Research Letters, 12, 094 013,
685 https://doi.org/10.1088/1748-9326/aa8391, 2017.

Randerson, J., Van Der Werf, G., Giglio, L., Collatz, G., and Kasibhatla, P.: Global Fire Emissions Database, Version 4.1 (GFEDv4), https://doi.org/10.3334/ORNLDAAC/1293, 2017.

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 & Environment, 3, 168, https://doi.org/10.1038/s43247-
690 022-00498-3, 2022.

Riley, W. J., Subin, Z. M., Lawrence, D. M., Swenson, S. C., Torn, M. S., Meng, L., Mahowald, N. M., and Hess, P.: Barriers to predicting changes in global terrestrial methane fluxes: analyses using CLM4Me, a methane biogeochemistry model integrated in CESM, Biogeosciences, 8, 1925–1953, https://doi.org/10.5194/bg-8-1925-2011, 2011.

Ringeval, B., de Noblet-Ducoudré, N., Ciais, P., Bousquet, P., Prigent, C., Papa, F., and Rossow, W. B.: An attempt to quantify the impact of changes in wetland extent on methane emissions on the seasonal and interannual time scales, Global Biogeochemical Cycles, 24, https://doi.org/https://doi.org/10.1029/2008GB003354, 2010.

Saunois, M., Martinez, A., Poulter, B., Zhang, Z., Raymond, P. A., Regnier, P., Canadell, J. G., Jackson, R. B., Patra, P. K., Bousquet, P., Ciais, P., Dlugokencky, E. J., Lan, X., Allen, G. H., Bastviken, D., Beerling, D. J., Belikov, D. A., Blake, D. R., Castaldi, S., Crippa, M., Deemer, B. R., Dennison, F., Etiope, G., Gedney, N., Höglund-Isaksson, L., Holgerson, M. A., Hopcroft, P. O., Hugelius, G., Ito, A., Jain, A. K., Janardanan, R., Johnson, M. S., Kleinen, T., Krummel, P. B., Lauerwald, R., Li, T., Liu, X., McDonald, K. C., Melton, J. R., Mühle, J., Müller, J., Murguia-Flores, F., Niwa, Y., Noce, S., Pan, S., Parker, R. J., Peng, C., Ramonet, M., Riley, W. J., Rocher-Ros, G., Rosentreter, J. A., Sasakawa, M., Segers, A., Smith, S. J., Stanley, E. H., Thanwerdas, J., Tian, H., Tsuruta, A., Tubiello, F. N., Weber, T. S., van der Werf, G. R., Worthy, D. E. J., Xi, Y., Yoshida, Y., Zhang, W., Zheng, B., Zhu, Q., Zhu, Q., and Zhuang, Q.: Global Methane Budget 2000–2020, Earth System Science Data, 17, 1873–1958, https://doi.org/10.5194/essd-17-1873-2025, 2025.

Scarpelli, T., Jacob, D., Moran, M., Reuland, F., and Gordon, D.: Gridded inventory of Canada's anthropogenic methane emissions for 2018, https://doi.org/10.7910/DVN/CC3KLO, 2021.

Schulzweida, U.: Climate Data Operators (CDO) User Guide, https://code.mpimet.mpg.de/projects/cdo/embedded/cdo.pdf, 2023.

Schuur, E. A. G., McGuire, A. D., Schädel, C., Grosse, G., 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.

Singarayer, J. S., Valdes, P. J., Friedlingstein, P., Nelson, S., and Beerling, D. J.: Late Holocene methane rise caused by orbitally controlled increase in tropical sources, Nature, 470, 82–85, https://doi.org/10.1038/nature09739, 2011.

Spahni, R., Wania, R., Neef, L., van Weele, M., Pison, I., Bousquet, P., Frankenberg, C., Foster, P. N., Joos, F., Prentice, I. C., and van Velthoven, P.: Constraining global methane emissions and uptake by ecosystems, Biogeosciences, 8, 1643–1665, https://doi.org/10.5194/bg-8-1643-2011, 2011.

Sweeney, C., Dlugokencky, E., Miller, C. E., Wofsy, S., Karion, A., Dinardo, S., Chang, R. Y.-W., Miller, J. B., Bruhwiler, L., Crotwell, A. M., Newberger, T., McKain, K., Stone, R. S., Wolter, S. E., Lang, P. E., and Tans, P.: No significant increase in long-term CH_4 emissions on North Slope of Alaska despite significant increase in air temperature, Geophysical Research Letters, 43, 6604–6611, https://doi.org/https://doi.org/10.1002/2016GL069292, 2016.

Sweeney, C., Chatterjee, A., Wolter, S., McKain, K., Bogue, R., Conley, S., Newberger, T., Hu, L., Ott, L., Poulter, B., Schiferl, L., Weir, B., Zhang, Z., and Miller, C. E.: Using atmospheric trace gas vertical profiles to evaluate model fluxes: a case study of Arctic-CAP observations and GEOS simulations for the ABoVE domain, Atmospheric Chemistry and Physics, 22, 6347–6364, https://doi.org/10.5194/acp-22-6347-2022, 2022.

Thompson, R. L., Sasakawa, M., Machida, T., Aalto, T., Worthy, D., Lavric, J. V., Lund Myhre, C., and Stohl, A.: Methane fluxes in the high northern latitudes for 2005–2013 estimated using a Bayesian atmospheric inversion, Atmospheric Chemistry and Physics, 17, 3553–3572, https://doi.org/10.5194/acp-17-3553-2017, 2017.

Tian, H., Xu, X., Liu, M., Ren, W., Zhang, C., Chen, G., and Lu, C.: Spatial and temporal patterns of CH_4 and N_2O fluxes in terrestrial ecosystems of North America during 1979–2008: application of a global biogeochemistry model, Biogeosciences, 7, 2673–2694, https://doi.org/10.5194/bg-7-2673-2010, 2010.

730   van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., van Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S.: Global fire emissions estimates during 1997–2016, Earth System Science Data, 9, 697–720, https://doi.org/10.5194/essd-9-697-2017, 2017.

van Huissteden, J. and Dolman, A.: Soil carbon in the Arctic and the permafrost carbon feedback, Current Opinion in Environmental Sustainability, 4, 545–551, https://doi.org/https://doi.org/10.1016/j.cosust.2012.09.008, terrestrial systems, 2012.

735   Walter, B. P. and Heimann, M.: A process-based, climate-sensitive model to derive methane emissions from natural wetlands: Application to five wetland sites, sensitivity to model parameters, and climate, Global Biogeochemical Cycles, 14, 745–765, https://doi.org/https://doi.org/10.1029/1999GB001204, 2000.

Wania, R., Ross, I., and Prentice, I. C.: Implementation and evaluation of a new methane model within a dynamic global vegetation model: LPJ-WHyMe v1.3.1, Geoscientific Model Development, 3, 565–584, https://doi.org/10.5194/gmd-3-565-2010, 2010.

740   Wania, R., Melton, J. R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., Bohn, T., Avis, C. A., Chen, G., Eliseev, A. V., Hopcroft, P. O., Riley, W. J., Subin, Z. M., Tian, H., van Bodegom, P. M., Kleinen, T., Yu, Z. C., Singarayer, J. S., Zürcher, S., Lettenmaier, D. P., Beerling, D. J., Denisov, S. N., Prigent, C., Papa, F., and Kaplan, J. O.: Present state of global wetland extent and wetland methane modelling: methodology of a model inter-comparison project (WETCHIMP), Geoscientific Model Development, 6, 617–641, https://doi.org/10.5194/gmd-6-617-2013, 2013.

745   Ward, R. H., Sweeney, C., Miller, J. B., Goeckede, M., Laurila, T., Hatakka, J., Ivakov, V., Sasakawa, M., Machida, T., Morimoto, S., Goto, D., and Ganesan, A. L.: Increasing methane emissions and widespread cold-season release from high-Arctic regions detected through atmospheric measurements, Journal of Geophysical Research: Atmospheres, 129, e2024JD040 766, https://doi.org/https://doi.org/10.1029/2024JD040766, e2024JD040766 2024JD040766, 2024.

Watts, J. D., Farina, M., Kimball, J. S., Schiferl, L. D., Liu, Z., Arndt, K. A., Zona, D., Ballantyne, A., Euskirchen, E. S., Parmentier, 750   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/https://doi.org/10.1111/gcb.16553, 2023.

755   Wold, S., Esbensen, K., and Geladi, P.: Principal component analysis, Chemometrics and Intelligent Laboratory Systems, 2, 37–52, https://doi.org/https://doi.org/10.1016/0169-7439(87)80084-9, proceedings of the Multivariate Statistical Workshop for Geologists and Geochemists, 1987.

Worthy, D. E. J., Levin, I., Trivett, N. B. A., Kuhlmann, A. J., Hopper, J. F., and Ernst, M. K.: Seven years of continuous methane observations at a remote boreal site in Ontario, Canada, Journal of Geophysical Research: Atmospheres, 103, 15 995–16 007, 760   https://doi.org/10.1029/98JD00925, 1998.

Xiao, H., Song, C., Li, S., Lu, X., Liang, M., Xia, X., and Yuan, W.: Global wetland methane emissions from 2001 to 2020: magnitude, dynamics and controls, Earth's Future, 12, e2024EF004 794, https://doi.org/https://doi.org/10.1029/2024EF004794, 2024.

Zhang, Z., Zimmermann, N. E., Stenke, A., Li, X., Hodson, E. L., Zhu, G., Huang, C., and Poulter, B.: Emerging role of wetland methane emissions in driving 21st century climate change, Proceedings of the National Academy of Sciences, 114, 9647–9652, 765   https://doi.org/10.1073/pnas.1618765114, 2017.

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

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

Zhang, Z., Poulter, B., Melton, J. R., Riley, W. J., Allen, G. H., Beerling, D. J., Bousquet, P., Canadell, J. G., Fluet-Chouinard, E., Ciais, P., Gedney, N., Hopcroft, P. O., Ito, A., Jackson, R. B., Jain, A. K., Jensen, K., Joos, F., Kleinen, T., Knox, S. H., Li, T., Li, X., Liu, X., McDonald, K., McNicol, G., Miller, P. A., Müller, J., Patra, P. K., Peng, C., Peng, S., Qin, Z., Riggs, R. M., Saunois, M., Sun, Q., Tian, H., Xu, X., Yao, Y., Xi, Y., Zhang, W., Zhu, Q., Zhu, Q., and Zhuang, Q.: Ensemble estimates of global wetland methane emissions over

775   2000–2020, Biogeosciences, 22, 305–321, https://doi.org/10.5194/bg-22-305-2025, 2025.