Uncertainty and retrieval sensitivity in TROPOMI-based methane inversions over the North Slope of Alaska

Rebecca H. Ward[1,2], Luke M. Western[3], Rachel L. Tunnicliffe[3], Elena Fillola[4], Aki Tsuruta[2], Tuula Aalto[2], and Anita L. Ganesan[1]

[1]School of Geographical Sciences, University of Bristol, Bristol, United Kingdom
[2]Climate System Research, Finnish Meteorological Institute, 00560 Helsinki, Finland
[3]School of Chemistry, University of Bristol, Bristol, United Kingdom
[4]Department of Engineering Mathematics, University of Bristol, Bristol, United Kingdom

Correspondence: Rebecca H. Ward (rebecca.ward@fmi.fi)

Abstract.

The Arctic is experiencing unprecedented environmental changes with rapidly rising temperatures. Emissions of methane (CH_4) – a potent greenhouse gas – may be increasing from the region, making accurate monitoring essential. The TROPOspheric Monitoring Instrument (TROPOMI) instrument offers high spatial and temporal coverage of CH_4 column mole fractions. However, its data in the Arctic has historically exhibited seasonal and latitudinal biases and low-quality retrievals. A major challenge is the lack of ground-based validation data in high-latitude regions, which are used to improve satellite retrievals. This study evaluates inverse modelling to estimate CH_4 emissions using TROPOMI measurements over the North Slope of Alaska. Using two retrieval products – the operational SRON product and the scientific WFMD product from the University of Bremen – we assess the alignment of derived emissions with surface measurement-derived inversions over 2018–2020 and test their robustness through sensitivity analyses. Our results show that tundra emissions from SRON inversions align more closely with surface measurement-derived emissions than WFMD inversions. Both TROPOMI-product derived emissions have anomalously low emissions in August 2018 compared to surface measurement-derived emissions, likely due to low data density resulting from high cloud cover. TROPOMI inversions provided stronger constraints on fugitive anthropogenic emissions compared to surface inversions. However, each retrieval produced different emission estimates, highlighting retrieval-dependent differences. Sensitivity tests revealed a strong prior dependence in both retrievals, raising concerns about robustness in northern high latitudes. This study highlights the importance of using multiple retrievals and rigorous sensitivity testing in high-latitude satellite inversions.

Copyright statement. TEXT

1 Introduction

20 Temperatures in the Arctic have been increasing four times the global average (Rantanen et al., 2022), leading to unprecedented environmental changes, including the thawing of permafrost (Box et al., 2019). This thawing has the potential to trigger a positive climate feedback, releasing carbon dioxide (CO_2) and methane (CH_4) to the atmosphere, and further accelerating warming (IPCC, 2019).

Atmospheric mole fractions of CH_4, the second largest anthropogenic contributor to global warming after CO_2, are rising
25 globally (Lan and Dlugokencky, 2022), with CH_4 emissions from the Arctic particularly uncertain (Saunois et al., 2020). Regional studies have recently shown an increase in emissions from the tundra region of the North Slope of Alaska, linked to rising temperatures (Sweeney et al., 2016; Ward et al., 2024). However, global studies have not yet demonstrated this increase to be widespread from the whole Arctic region (Lan et al., 2021; Hugelius et al., 2024).

Arctic CH_4 emissions during the cold season (September–May) can comprise up to 50% of its annual emissions (Mas-
30 tepanov et al., 2008; Zona et al., 2016; Kittler et al., 2017; Treat et al., 2018; Ward et al., 2024). This season includes the zero curtain period (September–December) when the active layer of permafrost remains near $0°C$ below the frozen surface and CH_4 production still occurs (Outcalt et al., 1990). However, there are few studies of cold season CH_4 emissions despite the importance of the period (Treat et al., 2018).

Emissions of CH_4 from the Arctic have previously been estimated using a network of surface mole fraction measurements
35 in atmospheric inversions (Berchet et al., 2015; Thompson et al., 2017; Ishizawa et al., 2019; Tenkanen et al., 2021; Wittig et al., 2023). Wittig et al. (2024) demonstrated that although the current observation network can detect Arctic CH_4 emission trends, its sparse distribution poses significant challenges for accurate quantification and spatial attribution of these emissions. The coverage of this surface network is threatened further by a loss of measurement availability from Russia (Schuur et al., 2024), meaning that measurements from Siberia, with land types that do not have analogues in other regions of the Arctic, will
40 not be included in future estimates.

Satellite measurements of the column-averaged dry-air CH_4 mole fraction (XCH_4) offer a potential supplement to surface data in the Arctic. The TROPOspheric Monitoring Instrument (TROPOMI) on the Sentinel 5 Precursor satellite offers high spatial (5 km x 7 km) and temporal (daily revisit) resolutions, presenting possibilities for atmospheric monitoring. TROPOMI measurements are processed using two main retrieval methods: the operational product from the Netherlands Institute for
45 Space Research (TROPOMI-OPER) (Hu et al., 2016) and the Weighting Function Modified Differential Optical Absorption Spectroscopy (TROPOMI-WFMD) product developed by the University of Bremen (Schneising et al., 2019).

Retrievals from instruments on board satellites such as SCIAMACHY (Buchwitz et al., 2005) and GOSAT (Parker et al., 2011) have facilitated global and regional CH_4 inversions (Bergamaschi et al., 2009; Houweling et al., 2014; Alexe et al., 2015; Turner et al., 2015; Pandey et al., 2016; Wang et al., 2019; Maasakkers et al., 2019; Zhang et al., 2021; Lu et al.,
50 2022). However, few studies have utilised GOSAT or SCIAMACHY for inversions over high latitudes due to challenges from limited and biased measurements, as well as difficulties with vertical profile modelling in these regions, challenges which also apply to TROPOMI. The reliance on sunlight for passive short-wave infrared (SWIR) measurements at a high solar zenith

angle means that there are sparse measurements during the cold season (Turner et al., 2015), limiting the ability to study this potentially high-CH_4-emitting period. The steep vertical gradient of CH_4 in the stratosphere and the large contribution of the stratosphere to the total atmospheric column near the poles necessitate accurate stratospheric modelling for precise satellite retrievals (Jacob et al., 2016). Errors in stratospheric CH_4 profiles can significantly impact the inferred surface emissions from satellite data (Ostler et al., 2016). The polar vortex, a large area of low pressure and cold air surrounding the Earth's poles, can create distinct vertical CH_4 profiles and influence CH_4 transport, further complicating the modelling process in high latitudes (Ostler et al., 2014). Limited data, seasonal biases and modelling challenges have previously led to the exclusion of high-latitude satellite data from inversion studies, including those using TROPOMI measurements (Bergamaschi et al., 2009; Alexe et al., 2015; Houweling et al., 2017; Maasakkers et al., 2019).

There are a small number of inversion studies that use TROPOMI over northern high latitudes, and they reveal persistent seasonal biases (Qu et al., 2021; Tsuruta et al., 2023). Tsuruta et al. (2023) identified ongoing issues with transport model biases in CH_4 profiles, of which some have been addressed by recent updates to the TROPOMI-WFMD and TROPOMI-OPER products (Schneising et al., 2023; Lorente et al., 2021, 2023).

The ongoing validation of TROPOMI data, including comparisons with XCH_4 measurements such as the Total Carbon Column Observing Network (TCCON) (Wunch et al., 2011), the Collaborative Carbon Column Observing Network (COCCON) (Frey et al., 2019) or vertical profile measurements from AirCore devices on balloon flights (Degen et al., 2024) and aircraft campaigns (Paris et al., 2014; Narbaud et al., 2023), remains crucial for ensuring accuracy, reliability and the monitoring of biases (Turner et al., 2019). However, much of the Arctic lacks these supporting measurements, with only three Arctic TCCON stations, four Arctic COCCON stations, and sparse, discrete AirCore balloon and aircraft campaign measurements that are not continuous or long-term, leading to significant gaps in validation datasets for the region. A comprehensive validation of TROPOMI-OPER and TROPOMI-WFMD retrievals over the northern high latitudes found that the most recent versions of these products show significant improvements over previous versions, though some anomalies persist during the cold season (Lindqvist et al., 2024).

This study investigates the use of TROPOMI data for CH_4 flux estimation in the high-latitude Arctic. We compare emissions derived from satellite data, employing both TROPOMI-OPER and TROPOMI-WFMD products, with those derived from surface in situ observations. We emphasise that emissions derived using surface measurements can contain their own biases and uncertainties, however, inversions using surface measurements have been more extensively evaluated in a variety of applications. We infer emissions using a Hierarchical Bayesian Markov Chain Monte Carlo (HBMCMC) inversion over the North Slope of Alaska (NSA) region for 2018-2020. We focus on the NSA due to its high coverage of tundra and continuos permafrost, and because of increasing emissions shown in recent decades (Sweeney et al., 2016; Ward et al., 2024). Compared to other Arctic regions of interest, such as the West Siberian Lowlands, the NSA has much less overlap between anthropogenic and biogenic CH_4 sources, reducing uncertainties related to source attribution in the inversion. To evaluate the robustness of the TROPOMI inversions, we perform sensitivity tests by varying key inversion inputs and assessing their impact on the posterior emissions. Additionally, we conduct inversions combining surface and satellite data.

Sect. 2 introduces the data used in this work and a description of the inversion method and sensitivity tests. Sect. 3 describes the results of the inversions and sensitivity testing. Sect. 4 discusses the results, evaluates of the use of TROPOMI measurements over the NSA and draws comparisons to previous high-latitude TROPOMI inversion studies. Potential improvements to the inversion set-up in this and future high-latitude studies are discussed in Sect. 5.

2 Methods

Throughout this work, we define the early season as March–May, and the summertime/growing season as June–August. While the late season has been previously defined as September–December (Ward et al., 2024), we consider only September and October, as the satellite retrievals from TROPOMI over the study region do not extend beyond October.

We primarily focus on emissions from all natural sources in the NSA, referred to here as the tundra sector, due to its regional importance and its potential for increasing CH_4 emissions. We also examine anthropogenic emissions, primarily fugitive CH_4 from the Prudhoe Bay oil field. In this study the NSA region is defined as 65°N to 72°N and 172°W to 140°W.
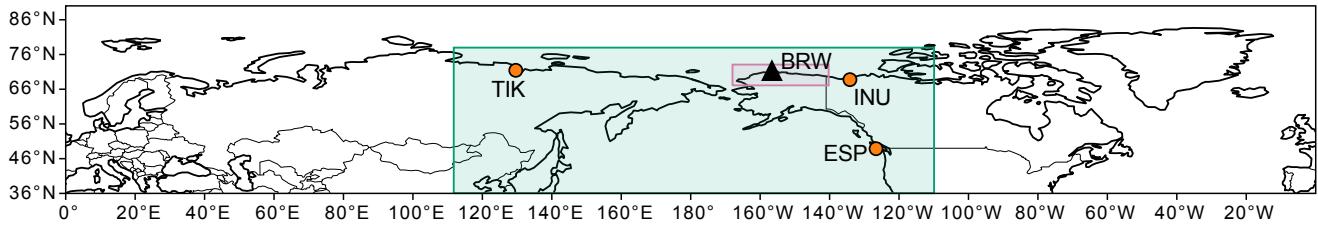
Figure 1. Regions, domains and sites used in this study. The Arctic inversion modelling domain used in Ward et al. (2024) spans the entire area shown, compared to the smaller Alaska inversion modelling domain (large green rectangle) used in this study. The location of the focus region, the NSA, as defined in this study, is shown with the smaller pink rectangle. The location of BRW is indicated by a black triangle and the three additional sites TIK, INU and ESP are shown with an orange circle.

2.1 Inversion method

The inversions in this work use the same method as Ward et al. (2024), which is described in the following. Measurements of CH_4 mole fractions (for surface inversions) or XCH_4 column-averaged mole fractions (for satellite inversions), represented by $\mathbf{y}$, are related to a vector of unknown emissions $\mathbf{x}$ and unknown boundary conditions (the background CH_4 mole fraction) $\mathbf{x_b}$ in the forward model,

$$\mathbf{y} = \mathbf{Hx} + \mathbf{H_b x_b} + \epsilon, \tag{1}$$

where $\mathbf{H}$ and $\mathbf{H_b}$ are matrices of sensitivities of the measurements to emissions and of the measurements to boundary conditions, respectively. These are produced using the NAME atmospheric transport model (described in Sect. 2.4). Model-measurement errors are represented by ϵ.

A prior modelled estimate of the measurements is obtained by combining the matrix sensitivities with prior CH_4 emissions and prior boundary conditions (described in Sect. 2.5). The inversion solves for monthly scaling factors applied to the prior emissions and boundary conditions, which produce posterior emissions and boundary conditions that best reconcile the observations with uncertainties in the measurements, transport model, and priors. We employ a hierarchical Bayesian inversion framework (Eq. 2), following the approach of Ganesan et al. (2014).

$$p(\mathbf{x}, \mathbf{x_b}, \boldsymbol{\theta} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \mathbf{x}, \mathbf{x_b}, \boldsymbol{\theta}) p(\mathbf{x} \mid \boldsymbol{\theta}) p(\mathbf{x_b} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}), \qquad (2)$$

where $p(\mathbf{y} \mid \mathbf{x}, \mathbf{x_b}, \boldsymbol{\theta})$ defines the likelihood of the measurements given the emissions and boundary conditions. The prior probability density function (PDFs) are $p(\mathbf{x} \mid \boldsymbol{\theta})$ and $p(\mathbf{x_b} \mid \boldsymbol{\theta})$. The term $p(\boldsymbol{\theta})$ specifies the prior distribution for the hyperparameters that govern the model–measurement error. This hierarchical formulation of Bayes' theorem treats the model–measurement error as an unknown variable as opposed to choosing subjective values. The hyperparameters, $\boldsymbol{\theta}$, represent both the known measurement uncertainty and the unknown model error (combined in quadrature), allowing these uncertainties to be jointly estimated with the emissions and boundary conditions during the inversion.

To avoid sampling large, non-physical negative emissions, we use a truncated normal prior PDF for both the emissions and boundary conditions. This constraint means that net uptake cannot be represented at the grid-cell level, although reductions from the prior can still capture relative decreases in emissions. The PDFs used in the inversions are described in Table 1.

Variable	Prior Probability Distribution	Parameters
Emissions scaling factor, $\mathbf{x}$	Truncated normal (truncated at 0)	$\mu = 1, \sigma = 2$
Boundary conditions scaling factor for surface inversions, $\mathbf{x_b}$	Truncated normal (truncated at 0)	$\mu = 1, \sigma = 0.02$
Boundary conditions scaling factor for satellite inversions, $\mathbf{x_b}$	Truncated normal (truncated at 0)	$\mu = 1, \sigma = 1$
Model error	Uniform	0.5 - 30 ppb

Table 1. Prior probability density functions used for each variable in the HBMCMC method.

There is no analytical solution for $\mathbf{x}$ and $\mathbf{x_b}$, we therefore use a Markov Chain Monte Carlo (MCMC) sampling approach. The emissions and boundary conditions are sampled using a No-U-Turn (NUTS) sampler (Hoffman and Gelman, 2014), while the model-measurement error is sampled using a faster Slice sampler (Neal, 2003). In each step of MCMC sampling, the PDFs of each parameter are considered by the samplers. After which, the sample is accepted or rejected based on acceptance criteria. The probability is recomputed at each step, allowing the chain to explore the parameter space, moving towards the most likely result while occasionally considering less likely outcomes through a series of acceptances and rejections of the proposed parameter space. We use a 1,500 iteration tuning phase to move away from the prior PDFs, then a 3000 iteration sampling phase, discarding the first 20% of samples as burn-in. To ensure convergence, we apply the Gelman-Rubin diagnostic (Gelman and Rubin, 1992) with a criterion of less than 1.05. The remaining samples, or the 'trace', provide the posterior PDF from which we calculate the mean scaling factor – used to scale the prior fluxes to the posterior fluxes – as well as the associated 95% uncertainty intervals. For example, for total emissions from the NSA, we calculate posterior fluxes for each

MCMC sample by multiplying the prior fluxes by each sampled scaling factor and then derive the mean and 95% uncertainty intervals (in Tg yr^{-1}) from the resulting distribution.

135   In Ward et al. (2024), an "Arctic" inversion spatial domain was used that covered the entire 360° of all latitudes greater than 36°N, selected for its ability to model the transport of particles near the poles, where particles can traverse the entire domain within the model simulation. However, this comes at a higher computational expense, making it less feasible for use with high-density TROPOMI data. This constraint is recognised as a limitation of our current study. For the current study, all of the TROPOMI inversions and sensitivity tests use a smaller "Alaska" domain (36°N to 78°N and 112°E to 110°W). All surface

140 inversions that are compared to the TROPOMI inversions also use the "Alaska" domain. Both domains are shown in Figure 1. The impact of using a different domain on the surface inversion is discussed in Sect. 3.1.

     In Ward et al. (2024), the boundary conditions in the North Slope of Alaska inversion were fixed during the inversion, as opposed to allowing variable boundary conditions. Instead a wind direction-derived boundary condition was subtracted from the observations prior to the inversion. A sensitivity analysis conducted as part of Ward et al. (2024) indicated minimal

145 difference between using fixed and variable boundary conditions. In the present study, we allow variable boundary conditions that we estimate throughout the inversions for both surface and satellite inversions to ensure consistency in the method.

     We aggregate grid cells and solve for fewer unknowns in the inversion than the output grid resolution of the transport model. The grid cells are more densely concentrated in areas of interest and more sparsely distributed further from the region of interest, resulting in a total of 297 scaling factors for the emissions for each month. In the satellite inversion, we solve the

150 boundary condition scaling factors across 4 vertical blocks – spanning 5 km intervals from 0 to 20 km – each divided into the North, East, South, and West directions, resulting in 16 scaling factors. We only use one scaling parameter in each direction (4 total) for surface inversion.

     Henceforth, the "main" inversion refers to satellite inversions that use the configuration described in this section, and any changes to this configuration are sensitivity tests (described in Sect. 2.6).

155 **2.2   Surface observations**

We use high frequency in situ atmospheric mole fraction measurements of CH$_4$ from the National Oceanic and Atmospheric Administration (NOAA) Barrow station (BRW, 71.3° N, 156.6° W) (Dlugokencky et al., 1995). We filter the BRW observations for low wind speeds and those coming from the direction of Utigiavik, a nearby town (see Ward et al., 2024). In this work, we also include a separate inversion using additional measurements from two measurement sites managed by Environment

160 and Climate Change Canada (ECCC), Inuvik (INU, 68.3° N, 133.5° W) and Estevan Point (ESP, 49.3° N, 126.5° W) (Schuldt et al., 2024), and the Finnish Meteorological Institute (FMI) site Tiksi (TIK, 71.6° N, 128.9° E) (Laurila, 2024). These sites were chosen because they are all within the study domain and have measurements within the study period (2018-2020). The locations of the sites are shown in Figure 1. The measurements from these sites are filtered using a local influence filter, which excludes observations whose NAME footprints indicate they are dominated by very local surface influence, following the

165 approach described in Ward et al. (2024). The impact of using multiple sites on the surface inversion is discussed in Sect. 3.1.

2.3 Satellite observations

We use two satellite retrieval products of column-averaged dry air mole fractions of CH_4 (XCH_4): the University of Bremen's WFM-DOAS v1.8 (Schneising et al., 2019, 2023) and the SRON operational product v2.04.00 (Lorente et al., 2021, 2023), which we term TROPOMI-WFMD and TROPOMI-OPER, respectively. For both products, the measurement error is taken from the 1-σ retrieval uncertainty.

All data have low-quality observations removed using the filters and quality flags recommended by each product, and are re-gridded to 0.234° (latitude) x 0.352° (longitude) spatial resolution (the resolution of the atmospheric transport model, see Section 2.4), using a conservative interpolation method, and averaged to an hourly temporal resolution. For both TROPOMI retrievals, we removed all retrievals over the ocean due to a lack of validation data over the ocean globally. We reduced the density of the TROPOMI observations to make it computationally feasible to run our atmospheric transport model. We prioritised the data from our main region of interest (the NSA) and retained fewer observations from areas outside the NSA but within the inversion domain, which helped to constrain boundary conditions. This resampling involved dividing the domain into regions with varying data retention rates and a random sampler. After regridding, we retained all data over the NSA, 50% from the rest of the state of Alaska, and 5% from the remaining areas. TROPOMI-WFMD observations and observational density after regridding but before and after resampling can be seen for an example month (June 2019) over the Alaska domain in Figure 2. The effects of these data retention strategies on boundary conditions were evaluated using a sensitivity test described in Sect. 2.6. After resampling, across the whole domain ∼510000 retrievals remain for TROPOMI-OPER and ∼580000 for TROPOMI-WFMD.

Due to high solar zenith angles and low sunlight, there are no observations for either TROPOMI-OPER or TROPOMI-WFMD from late October to early March. However, temporal coverage of TROPOMI-WFMD extends further into March and October than TROPOMI-OPER. As a result, TROPOMI-WFMD may have more potential to constrain emissions during the cold season. The number of retrievals over the NSA for TROPOMI-OPER and TROPOMI-WFMD (after regridding and resampling) is shown for all years in Figure S1.

The time range for our analysis is from May 2018 to October 2020. TROPOMI's fully functioning operating capacity began in late April 2018, so we do not include data from before this date.

2.4 Atmospheric transport model

To provide the sensitivity of XCH_4 observations to emissions from the surface we used the Met Office's Numerical Atmospheric-dispersion Modelling Environment (NAME) (Jones et al., 2007), a Lagrangian particle dispersion model (LPDM). We run NAME in backward mode, releasing model particles for each TROPOMI observation, which are tracked backward in time over a 30-day period. We record the sensitivity of the surface emissions (0-40 metres above ground level) and the boundary conditions to the measurements. The NAME output has a resolution of 0.234° (latitude) by 0.352° (longitude). We use the "Alaska" domain shown in Figure 1.
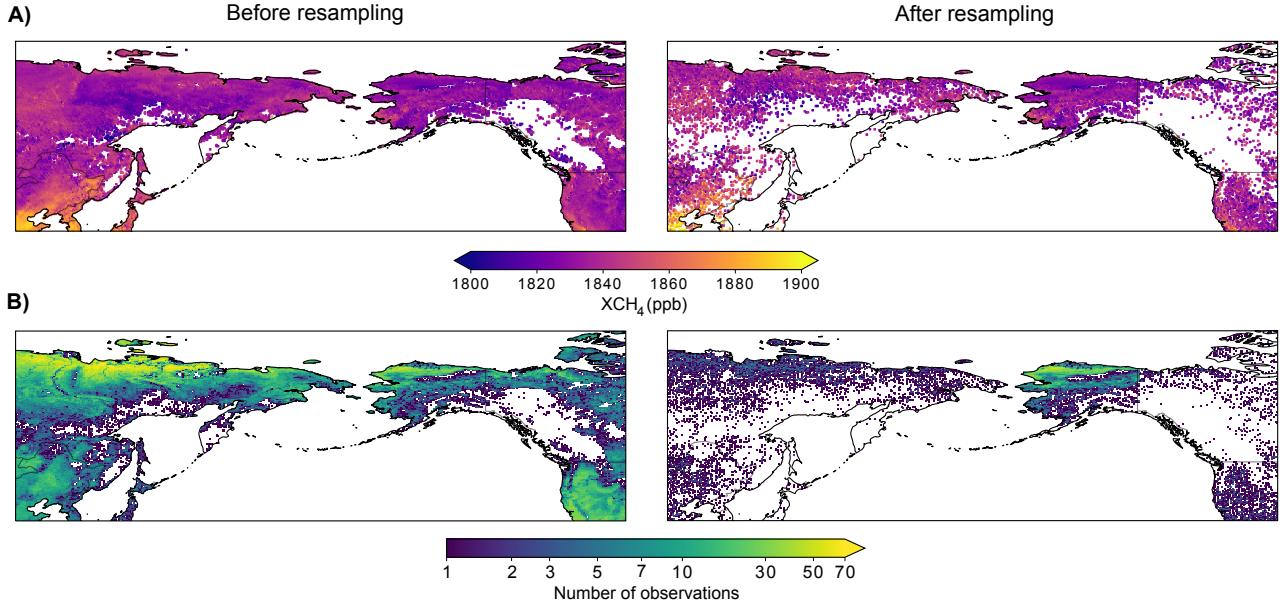
Figure 2. Example of observation coverage after regridding from the TROPOMI-WFMD retrieval for June 2019 over the Alaska study domain, before resampling (left) and after resampling (right). The top panel A) shows all observations in June 2019 and the bottom panel B) shows the number of observations (plotted on a logarithmic scale with zero observations shown in white) per $0.234°$ (latitude) x $0.352°$ (longitude) grid cell.

Meteorological inputs to NAME are provided by the Met Office's Unified Model (Walters et al., 2019), at a spatial resolution of $0.141°$ (latitude) x $0.094°$ (longitude) and 3-hourly temporal resolution. For each satellite product, we release particles uniformly across different heights defined by the levels within each retrieval. We ran NAME with a maximum altitude of 20 km, which corresponds to maximum level of 11 of 12 for TROPOMI-OPER and 19 of 20 for TROPOMI-WFMD. For the top levels that are above the maximum level of NAME, the CH_4 concentration was assumed to be equal to the corresponding value in the a priori profile from the retrieval. The average footprint sensitivity for TROPOMI (after resampling) are shown in Figure 3 and for the BRW-only inversion and the inversion using multiple surface sites in Figure S2.

To compare modelled mole fractions to satellite observations, we utilise the methodology derived in O'Dell et al. (2012) and described in Ganesan et al. (2017). We convert our modelled CH_4 for each retrieval into a column mole fraction (XCH_4^{model}) for each time index, t, as follows,

$$XCH_4^{model}|_t = \sum_1^n p_i \left[A_i \cdot CH_{4,i}^{model} + (1 - A_i) \cdot CH_{4,i}^{prior} \right],$$ (3)

where p_i is the pressure weight at each retrieval level i from the surface to the top of the atmosphere, A_i is the averaging kernel, $CH_{4,i}^{model}$ is the modelled CH_4 mole fraction and $CH_{4,i}^{prior}$ is the a priori profile. TROPOMI-WFMD retrievals provide all the
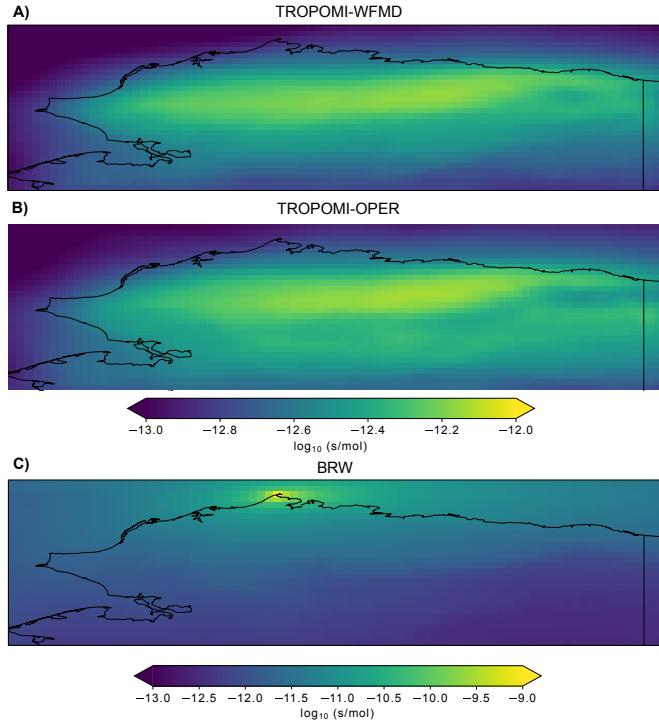
Figure 3. Average footprints in from 2018-2020 for A) TROPOMI-WFMD after resampling and B) TROPOMI-OPER after resampling and C) BRW over the North Slope of Alaska. Note that the scales for TROPOMI products and BRW are not the same. Direct comparisons of magnitude between a site footprint and a satellite footprint are not meaningful because the sensitivity of TROPOMI is averaged over both the vertical column and horizontally, whereas surface sites measure a localised point.

variables needed for Eq. 3, but for TROPOMI-OPER we calculate the pressure weights as $p_i = \frac{V_{air,i}}{\sum_i^n V_{air,i}}$, where V_{air} is the dry air sub-column.

We must adapt Eq. 3 such that we can calculate $\mathbf{XCH}_4^{model}$ in the form of the forward model, Eq. 1, defined in Sect. 3.1. Our forward modelled mole fraction at each level, i, ($CH_{4,i}^{model}$ from Eq.3) is due to emissions from the surface within our model domain, and incoming mole fractions from the boundaries,

$$CH_{4,i}^{model}|_t = \mathbf{h_i} \cdot \mathbf{q} + \mathbf{h_{b,i}} \cdot \mathbf{b}, \tag{4}$$

where $\mathbf{h_i}$ is our sensitivity to the surface derived from NAME, $\mathbf{q}$ is our emissions vector, $\mathbf{h_{b,i}}$ is our sensitivity to the boundary conditions and $\mathbf{b}$ is our boundary conditions vector.

As the inversion requires modelled mole fractions that are only dependent on NAME sensitivities, we construct a perturbed modelled column mole fraction, $XCH_{4,pert}^{model}|t$. This represents the averaging-kernel-weighted NAME contribution over the column, from the surface to the maximum level. We obtain $XCH_{4,pert}^{model}|t$ by subtracting all the retrieval-provided prior contributions, as follows, $CH_{4,i}^{prior}$, A_i and p_i are provided from the retrieval product, so the second term of Eq. 3 can be subtracted

from $\mathrm{XCH}_4^{model}|_t$. Additionally, because NAME is only run up to 20km altitude, we assume the prior profile above this level and we can also subtract this known contribution from $\mathrm{XCH}_4^{model}|_t$.

$$225 \quad \mathrm{XCH}_{4,pert}^{model}|_t = \mathrm{XCH}_4^{model}|_t - \sum_1^{maxlev} p_i(1-A_i) \cdot \mathrm{CH}_{4,i}^{prior} - \sum_{maxlev+1}^n p_i \cdot \mathrm{CH}_{4,i}^{prior}, \tag{5}$$

where maxlev is the maximum level of each retrieval for which the NAME model is run, as described above. A full derivation for Eq. 5 is found in Supplementary Note 2. Substituting Eq. 4 into Eq. 5 gives

$$\mathrm{XCH}_{4,pert}^{model}|_t = \sum_1^{maxlev} p_i \cdot A_i \cdot (\mathbf{h_i} \cdot \mathbf{q} + \mathbf{h_{b,i}} \cdot \mathbf{b}). \tag{6}$$

Over all times, t, Eq. 6 can be written in the form,

$$230 \quad \mathrm{XCH}_{4,pert}^{model} = \mathbf{Hq} + \mathbf{H_b}\mathbf{b}. \tag{7}$$

Here $\mathbf{H}$ and $\mathbf{H_b}$ are matrices of rows t, where each column is the sum of $p_i \cdot A_i \cdot \mathbf{h_i}$ and $p_i \cdot A_i \cdot \mathbf{h_{b,i}}$, respectively. Eq. 7 is now consistent with Eq. 1. In practice, because $\mathbf{q}$ and $\mathbf{b}$ are not dependent on the level, they are stacked into the vector $\mathbf{x}$ such that,

$$\mathrm{XCH}_{4,pert}^{model} = \mathbf{Hx}. \tag{8}$$

235 where each corresponding row of $\mathbf{H}$ is a weighted sum over all levels, allowing satellite observations to be used in the same inversion framework as surface measurements.

2.5 Prior boundary conditions and emissions

A priori vertical boundary conditions are derived from the CAMS global inversion-optimised greenhouse gas fluxes and concentrations v19r1 product (Segers and Houweling, 2020). This product uses surface observations in an inversion to produce
240 global atmospheric concentrations. As the product runs only until 2019, we assume that prior boundary conditions for 2020 are the same as in 2019 (discussed in Sec. 5).

Our a priori emissions field comprises emissions from three major sources over the Arctic: natural emissions from tundra, anthropogenic activities, and fires. For the main inversion, we utilise the "Late-season Zona" prior, detailed in Ward et al. (2024) and summarised in the following. For emissions from tundra, we employ a seasonal profile derived from the findings in Zona
245 et al. (2016), complemented by spatial data from the July values of the fractional wetland satellite product SWAMPs (Surface Water Microwave Product Series) (Jensen and Mcdonald, 2019). This approach is designed to allow emissions from areas of less inundated or frozen ground, as documented in Zona et al. (2016). Additionally, we incorporate anthropogenic emissions data from the Emissions Database for Global Atmospheric Research (EDGAR) v6 (Crippa et al., 2020) and fire emissions data from the Global Fire Emissions Database (GFED) v4 (Randerson et al., 2017). The a priori emissions are shown in Figure S3.
250 The majority of emissions over the NSA are from the tundra sector, with significant emissions from the anthropogenic sector over the Prudhoe Bay oil fields on the Northeast coast. Finally, there are small emissions from fires in the South of the defined

region. We do not include prior emissions from two other important sources in the NHLs: ocean and freshwater (lakes and rivers) emissions. In Ward et al. (2024), ocean emissions from Weber et al. (2019) were included as a sensitivity test, and their inclusion had no significant impact on the posterior emissions. For freshwater emissions, there is substantial spatial overlap with tundra emissions over the NSA, which can lead to double counting if both sources are included separately (Thornton et al., 2016). Including freshwater emissions in this region in the priors may therefore result in unrealistically high total prior emissions. The posterior tundra emissions can be interpreted as encompassing both tundra and freshwater methane sources.

To apportion the posterior emissions estimated in each grid cell into one of the three major source sectors, the fraction of each source in the prior emissions field was used. The distributions of each of these sectors are largely separated spatially as shown in Figure S4.

2.6 Sensitivity tests

Sensitivity tests were conducted to assess the robustness of the derived emissions to changes in the inversion setup. These tests included changing the prior used in the inversion to the "Uniform" prior from Ward et al. (2024), which uniformly distributes natural emissions from wetlands and tundra across all months, using the same spatial map as the "Late-season Zona" prior but setting the emission values to the July level for every month. We tested two different grid cell aggregation formats of 104 and 421 cells (shown in Figure S5). An additional test was carried out using a single scaling parameter for each boundary direction, such that only one block for each of the North, East, South, and West directions is solved for, in contrast to the main inversion which uses four blocks in each direction. Finally, to assess the impact of the spatial coverage of observations across the study domain, particularly on the boundary conditions, we removed TROPOMI observations from outside the state of Alaska. A summary of each sensitivity tests is shown in Table 2.

We additionally carried out an inversion using measurements from both TROPOMI and BRW. Combined TROPOMI and surface measurement inversions are referred to as TROPOMI-OPER+BRW or TROPOMI-WFMD+BRW throughout the text. We include an extra offset parameter, with a uniform PDF with a range of 0-10 ppb applied to the BRW measurements, as a method to account for systematic differences between the TROPOMI and BRW observations and in modelling each dataset by the NAME model.

Sensitivity test name	Details
Uniform prior emissions	Using prior emissions with a constant seasonal profile. Each each month has the same total emissions value.
Grid cell aggregations	Using aggregations of 104 or 421 grid cells. 297 in the main inversion.
Single BC scaling parameter	Using 1 scaling parameter in each of the N, E, S and W directions rather than 4 vertical blocks.
Alaska observations only	Using only TROPOMI observations from within the Alaska rather than from the whole domain.
Combined TROPOMI+BRW	Using observations from both TROPOMI and the BRW surface station in a combined inversion.

Table 2. Each of the sensitivity tests described in Sect. 2.6.

3 Results

3.1 Inversion using in situ measurements from surface sites
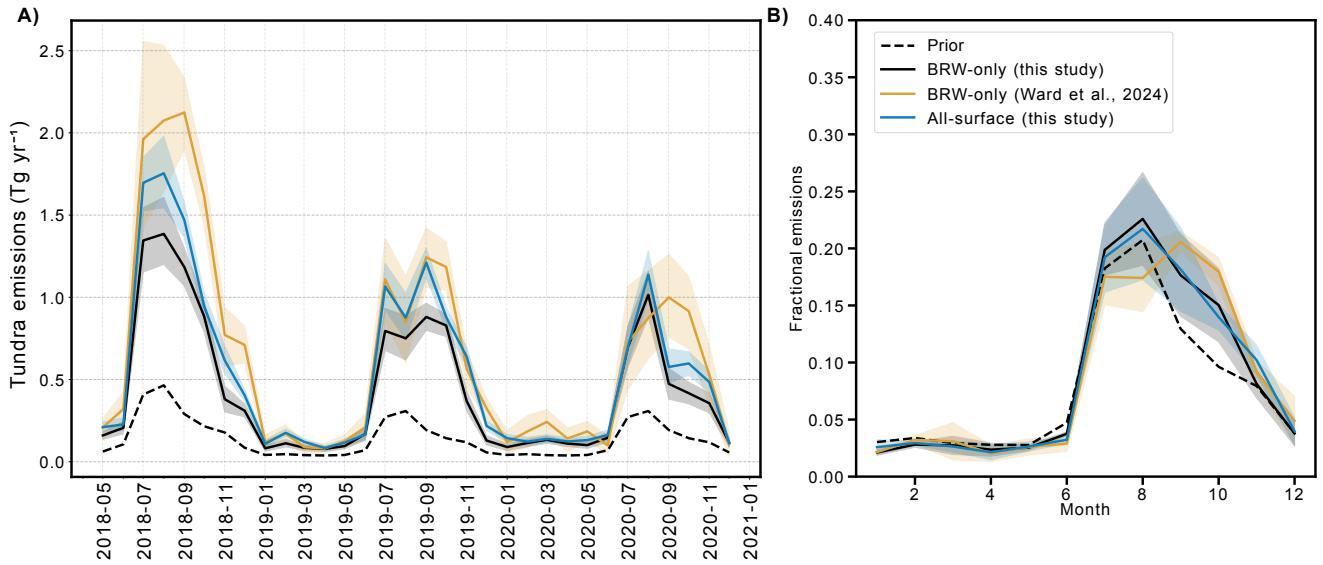
Figure 4. Tundra CH$_4$ emissions (A) and seasonal profiles (B) from the North Slope of Alaska (NSA) as derived from the BRW-only inversion for the Alaska domain (this study), in black, the BRW-only inversion for the Arctic domain ((Ward et al., 2024)), in orange, and the all-surface sites inversion for the Alaska domain (this study), in blue.

Here, satellite inversions are compared to a surface inversion using in situ measurements from BRW. We ran satellite and surface-measurement-based inversions over the same Alaska domain to compare them to the surface-measurement-based study in Ward et al. (2024), which used the more spatially extensive Arctic domain.

Figure 4A shows that, in the NSA, the impact of the size of the inversion domains on estimated tundra emissions varies year to year. While emissions for many months are consistent within uncertainties, inversions using the Arctic domain result in higher tundra emissions during the summer and late seasons of 2018, and the late seasons of 2019 and 2020. This suggests that domain size does impact the estimated tundra emissions in this region, potentially due to atmospheric circulation around the poles; however, the general findings using each domain are aligned. We observe significantly higher tundra emissions in our posterior results compared to our prior, including consistent high late-season emissions, a key finding in Ward et al. (2024). Additionally, the average seasonal emission profiles for both domains remain the same within uncertainties, as shown in Figure 4B. While we are satisfied with the similarities between the emissions derived for the NSA using either the Alaska and Arctic domains, further investigation is necessary to fully understand the implications of model domain on emission estimates from the Arctic.

A surface-based inversion was conducted using CH_4 measurements from BRW and three additional sites – INU, TIK, and ESP – within the Alaska domain, providing additional constraints outside of the NSA region on derived emissions. For 2020, the "all-surface" derived emissions agrees well with BRW-only (this study) derived emissions, as shown in Figure 4A. However, for the late season months of 2018 and most of 2019, the surface-measurement-derived emissions are larger and closer to the BRW-only (Ward et al., 2024) derived emissions, which use the larger domain. The seasonal emission profile from the surface-measurement-based inversion, shown in Figure 4B, remains consistent with the BRW-only (this study) inversion profile.

For the remainder of this study the BRW-only (this study) inversion using the Alaska domain is referred to as the BRW inversion.

3.2 Inversion using measurements from TROPOMI

3.2.1 Tundra sector emissions
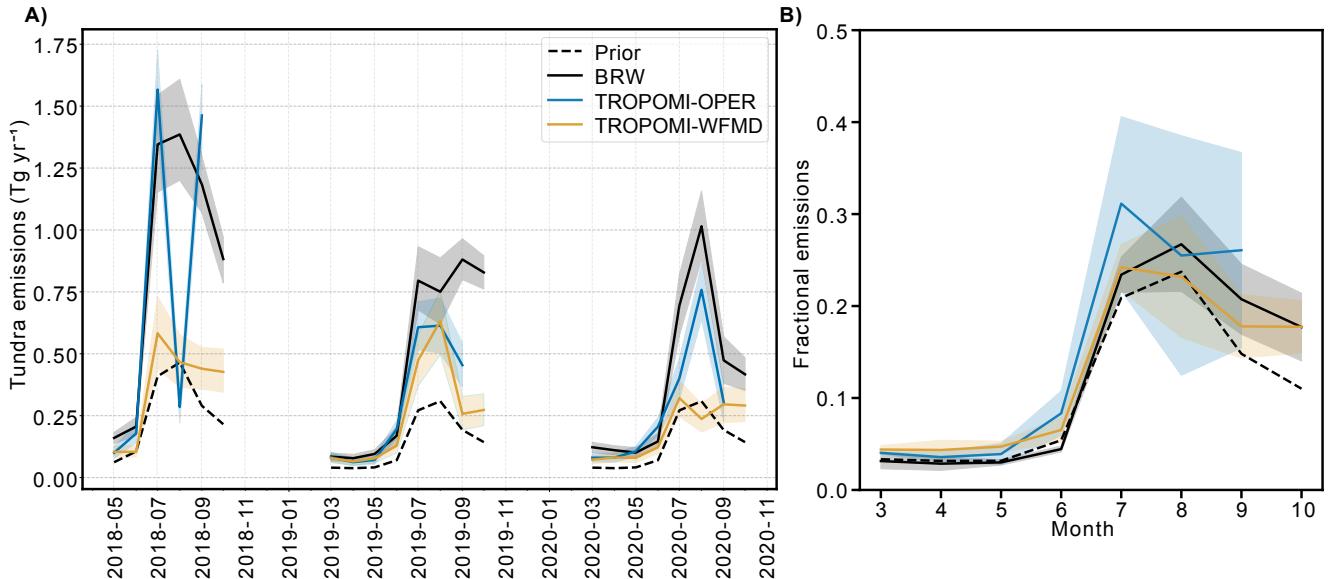
Figure 5. Posterior tundra emissions from the NSA for each of the BRW, TROPOMI-OPER and TROPOMI-WFMD inversions over the "Alaska" domain. A) shows the monthly emissions over the 2018-2020 period and B) shows average seasonal profile of emissions, expressed as a fraction of annual emissions.

The tundra emissions derived using the TROPOMI inversions over the NSA exhibit variable agreement with the BRW tundra emissions (outlined in Sect. 3.1). Figure 5A shows the total tundra emissions from the NSA for each month and Figure S6 shows the average (2018–2020) difference maps between the posterior mean and a priori emissions for the early (April–June), summer (July–August) and late (September–October) seasons for BRW, TROPOMI-OPER and TROPOMI-WFMD derived

emissions. For most months, TROPOMI-OPER tundra emissions for 2019 and 2020 are largely consistent with BRW tundra emissions within 95% uncertainty intervals. However, there is a notable exception where TROPOMI-OPER falls outside the 95% uncertainty interval, specifically in August 2018. Figure S7 shows the posterior and prior difference maps for July, August and September 2018, where TROPOMI-OPER derived emissions from the tundra region are elevated over the prior in July and September of 2018, but show a decrease in August. This anomaly is discussed in Sec. 4.3.

For most months, TROPOMI-WFMD tundra emissions are significantly lower than BRW-derived emissions within 95% uncertainty intervals. Only in March, April, May and August 2019 and in April, May, June and October 2020 do emissions agree. Otherwise, the TROPOMI-WFMD tundra emissions are less than the BRW-derived tundra emissions. Overall, the TROPOMI-OPER and TROPOMI-WFMD derived tundra emissions agree with the BRW tundra emissions for 15 of 19 months (79%) and 8 of 22 months (36%), respectively.

Comparing the derived emissions of the two TROPOMI retrievals TROPOMI-OPER has higher tundra emissions than TROPOMI-WFMD in May 2018, June 2018, September 2018, September 2019, June 2020 and August 2020. In all other months, the two TROPOMI retrievals are consistent within 95% uncertainty intervals.

3.2.2 Anthropogenic sector emissions
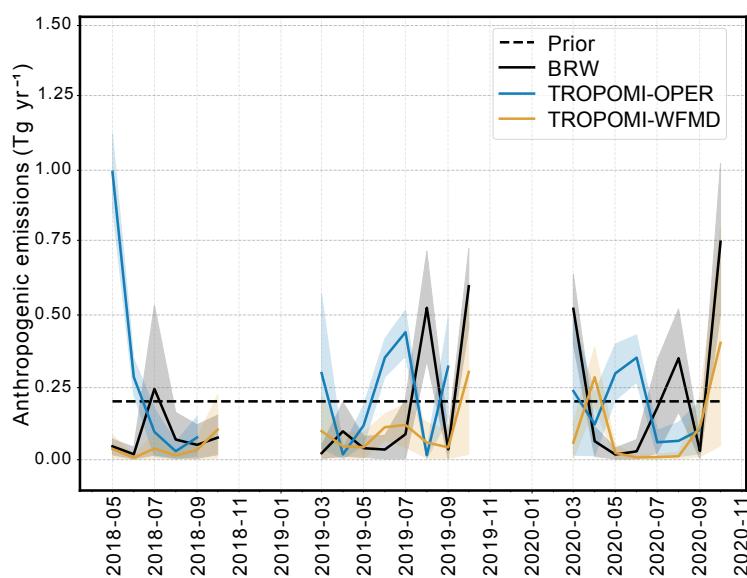
Figure 6. Posterior anthropogenic emissions from the NSA for each of the BRW, TROPOMI-OPER and TROPOMI-WFMD inversions over the "Alaska" domain for the 2018-2020 period.

There is significant variation in anthropogenic emissions across the different inversions (Figure 6). In May 2018, TROPOMI-OPER estimates of anthropogenic emissions were exceptionally high compared to the TROPOMI-WFMD and BRW esti-

mates. In June, July, and September 2019, TROPOMI-OPER derived anthropogenic emissions increase from the prior, while TROPOMI-WFMD and BRW decrease. Conversely, BRW emissions increase from the prior in August 2019, while TROPOMI-OPER and TROPOMI-WFMD decrease. In 2020, TROPOMI-OPER derived emissions increase from the prior in March, May, and June; BRW increases in March, August, and October; and TROPOMI-WFMD increases in April and October. Overall, anthropogenic emissions derived from TROPOMI-OPER agree with BRW for 8 of 19 months (42%), while TROPOMI-WFMD agree with BRW for 17 of 22 months (77%). Anthropogenic emissions in the NSA are predominantly from the Prudhoe Bay oil fields. The emissions maps (Figure S6) show that in the March-May emissions from the Prudhoe Bay oil fields are decreased for both the BRW and TROPOMI-WFMD inversions, but increased for TROPOMI-OPER.

 ## 3.3 Sensitivity tests of TROPOMI inversions
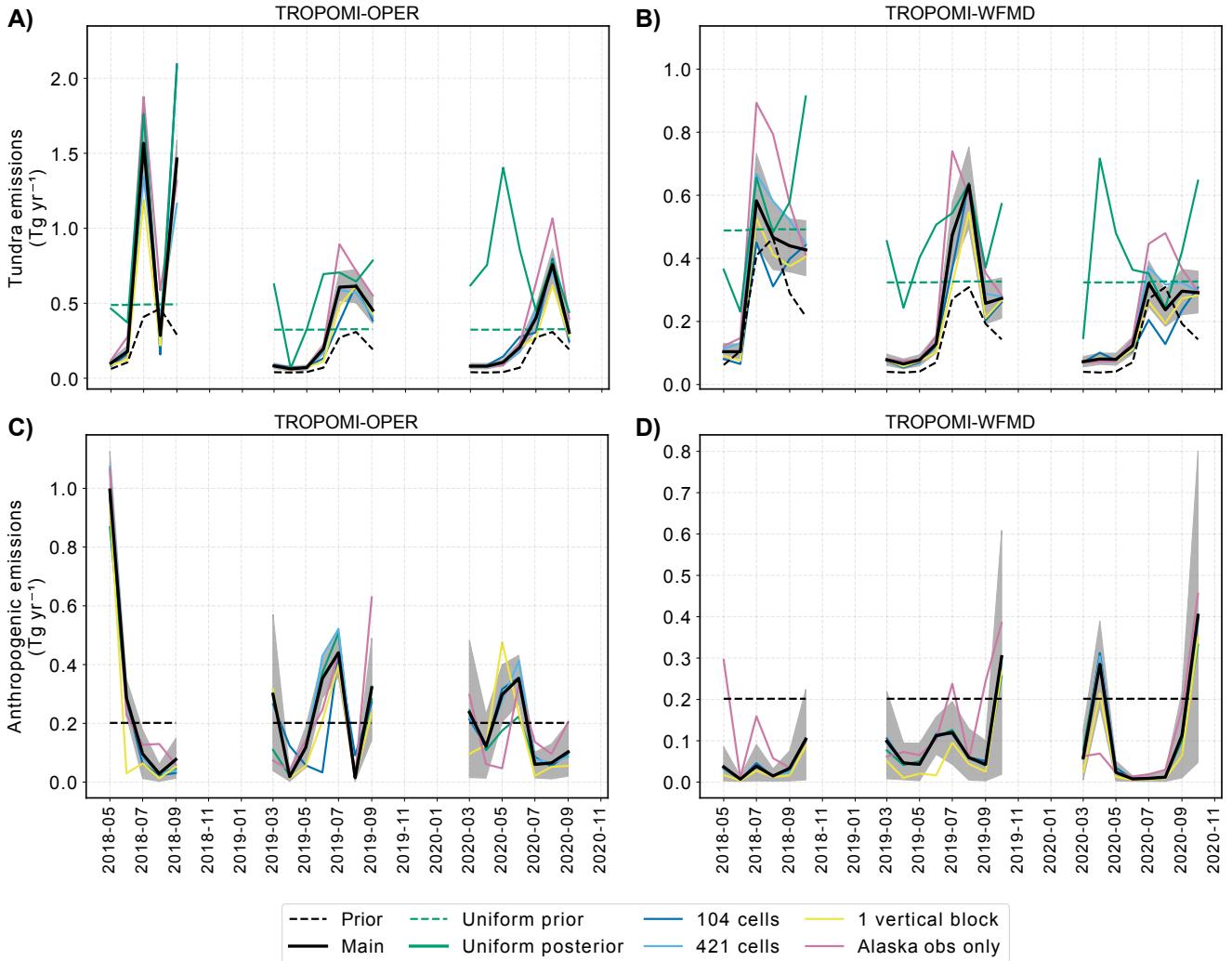
Figure 7. Sensitivity tests for the satellite-based inversions. A) and B) show derived tundra emissions from TROPOMI-OPER and TROPOMI-WFMD, respectively. C) and D) show derived anthropogenic emissions from the same inversions. Uncertainty estimates for the sensitivity tests are included in Figures S8-S12.

To assess the robustness of the satellite products to varying inversion inputs and prior choices, we carried out four sensitivity tests, as detailed in Table 2, and observed the impact on the derived tundra and anthropogenic emissions. The results of each sensitivity test are described in the following sections, with the results in Sect. 3.2 referred to as the "Main" inversion. The monthly total emissions for all tests are shown in Figure 7. We consider the monthly emissions estimate unchanged if the 95%

335 uncertainty intervals of the derived emissions from the sensitivity test and the main inversion overlap. March–September 2019 and 2020 total budgets are shown in Table S2.

3.3.1 Uniform prior emissions

A sensitivity test was conducted to assess the impact of changing seasonality of the prior emissions. For this test, the Late-Season Zona prior from the main inversion was replaced with a time-invariant (referred to as "Uniform") prior for the tundra
340 sector, as described in Sect. 2.5. This alternative prior configuration allows for emissions during both the early and late seasons. Surface inversions in Ward et al. (2024) showed that the derived emissions were not sensitive to the prior seasonal profile, as the use of the Uniform prior did not change the average posterior seasonality over the period 2000-2020.

The sensitivity test was repeated for the surface inversions detailed in Sect. 3.1. As shown in Figure S8, both the BRW-only inversion and the all-surface inversion show some sensitivity to the change in prior seasonal profile. The posterior seasonal
345 profile shifts from the Late-Season Zona inversion to the Uniform inversion, indicating a sensitivity to seasonal profile shape that did not occur in Ward et al. (2024). However, a clear seasonal profile, which peaks in the summer and is lowest in the winter, remains visible. This indicates a relatively low sensitivity to prior seasonal profile compared to the satellite inversions, described below.

Tundra emissions from TROPOMI-OPER and TROPOMI-WFMD inversions show a high sensitivity to the prior emissions
350 seasonal profile shape. Posterior tundra emissions (Figure S9) and seasonal profiles (Figure 8A and B), are significantly different between the Late-Season Zona and Uniform priors, and a clear seasonal cycle is no longer visible in the Uniform prior case.
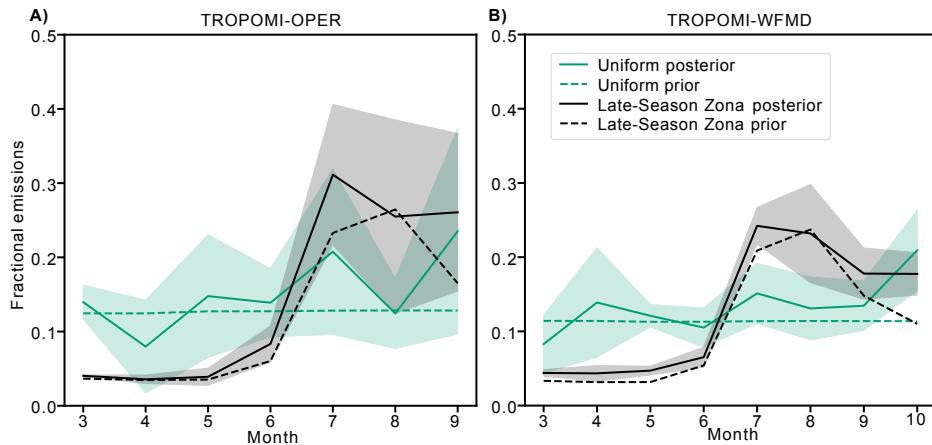
Figure 8. Uniform prior sensitivity test for the TROPOMI inversions. Average seasonal profiles of derived tundra emissions for A) TROPOMI-OPER and B) TROPOMI-WFMD for the Uniform prior (green) and Late-Season Zona prior (black). Note the Late-Season Zona prior is the same as the main inversion.

Changing the tundra sector emissions to the Uniform prior had no impact on the posterior emissions from the anthropogenic sector for any of the inversions.

3.3.2 Grid cell aggregations

As described in Sect. 2.1, we aggregate the grid cells of the transport model output in the inversion to solve for fewer parameters than its native grid resolution. In this sensitivity test, we used varying grid cell aggregations — 104 cells and 421 cells, compared to the 297 cells used in the main inversion (Figure S5). Comparisons to the main inversion are shown in Figure S10, while spatial patterns of the resulting scaling factors for selected months are shown in Figure S11.

For the derived tundra emissions, changing the number of grid cells did not result in significant alterations in the monthly emissions time series within 95% uncertainty intervals for TROPOMI-OPER, apart from in September 2018. For TROPOMI-WFMD, the summer months of 2018 and 2020 showed significant differences in the 95% uncertainty intervals between the 104 and 421 cell aggregations. The derived anthropogenic emissions were unaffected by the aggregation level for TROPOMI-WFMD. The TROPOMI-OPER inversion showed only changes for June 2019, where the 104 cell aggregation produced lower emissions than the 297- and 421-cell aggregations.

Changes in total monthly emissions are limited to a few months. For those months, there are large differences in how the prior emissions are spatially scaled across the NSA. For example, in the TROPOMI-OPER inversion for September 2018, the coarser 104-cell aggregation results in a large region of tundra being scaled up uniformly, leading to a noticeable increase in total emissions. The 297- and 421-cell aggregations allow for smaller areas to be adjusted, concentrating the scaling in fewer areas and limiting the impact on total emissions. In June 2019, scaling factors over the ocean in the TROPOMI-OPER inversion show substantial differences across the three different aggregations, suggesting that the treatment of these regions are sensitive to grid resolution.

Overall, TROPOMI-OPER inversions show greater spatial variability in scaling factors compared to TROPOMI-WFMD. For example, in August 2020 the TROPOMI-WFMD scaling factors remain relatively homogeneous, whereas TROPOMI-OPER shows more pronounced spatial heterogeneity, despite only TROPOMI-WFMD showing a change in the overall tundra emissions for this month.

3.3.3 Single boundary condition scaling parameter

We simplified the scaling factors for the boundary conditions compared to the main inversion by solving for a single scaling parameter for each of the North, East, South, and West boundaries, instead of the four vertical blocks in the main inversion. This tested whether four vertical blocks accounted for potential inaccuracies in modelling of the boundary conditions, by allowing for adjustments to the vertical profiles at the boundaries. Comparisons to the main inversion are shown in Figure S12.

For the derived tundra emissions, using a single vertical block resulted in a decrease in derived emissions but the decrease is not outside of the 95% uncertainty intervals for either inversion. For the derived anthropogenic emissions, when going from four to one vertical block, the TROPOMI-OPER emissions decreased in June 2018 and June 2019 and the TROPOMI-WFMD emissions decreased in June 2019. This suggests that allowing for vertical variability in the boundary conditions during the

inversion has a small but statistically insignificant impact on the tundra emissions, and a small impact on the anthropogenic emissions.

3.3.4 Observations only over Alaska

To assess the impact on derived emissions and boundary conditions of keeping observations across the whole domain, we
390 removed TROPOMI data from anywhere outside of Alaska. Comparisons to the main inversion are shown in Figure S13.

For the derived tundra emissions, using observations from Alaska resulted in increased emissions estimates outside of 95% uncertainty intervals in the summer months for both TROPOMI-OPER and TROPOMI-WFMD. For the derived anthropogenic emissions, using observations from Alaska only decreased emissions in May 2020 for TROPOMI-OPER and decreased emissions in May 2018 and April 2020 for TROPOMI-WFMD.

395 With Alaska measurements only, we observed lower boundary conditions. The exclusion of observations near the domain boundaries likely reduced the constraint on the boundary conditions and may have caused the model to underestimate influences from outside the inversion domain, contributing to the observed lowering of the boundary values.

3.4 Forward modelling of Barrow mole fractions with TROPOMI derived emissions
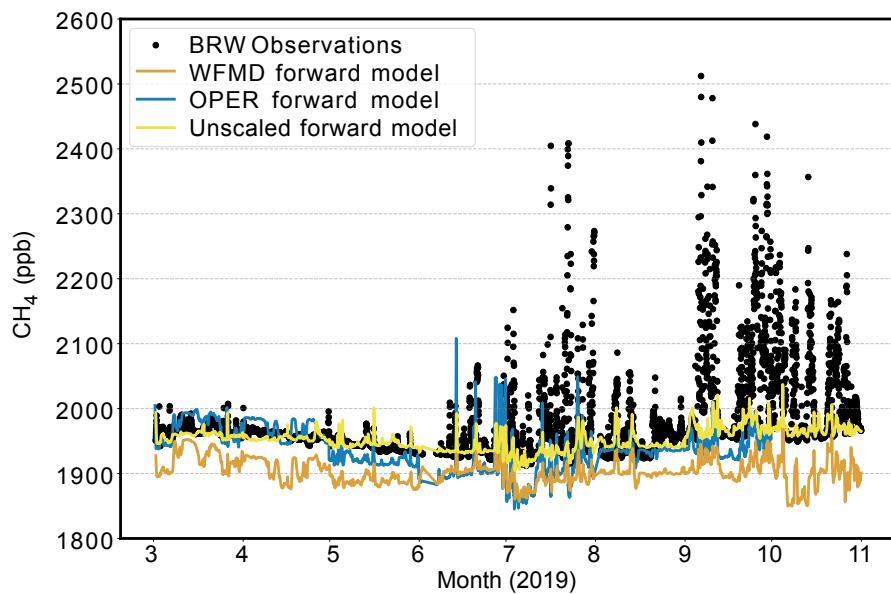
Figure 9. Forward modelled mole fractions at BRW for 2019. The forward model using TROPOMI derived emissions and boundary condition scaling factors are shown in blue for TROPOMI-OPER (blue) and TROPOMI-WFMD (orange). The forward model using the prior is shown in yellow and the BRW observed mole fractions are shown by black dots.

To validate the TROPOMI-derived emissions, we forward modelled the TROPOMI posterior emissions to generate modelled
mole fractions at BRW, as BRW observations were not assimilated in these inversions. The a priori emissions and boundary
conditions, $\mathbf{x}$ and $\mathbf{x_b}$, from the forward model in Equation 1 were scaled by the mean posterior emissions scaling factors and
posterior boundary condition scaling factors, derived in the TROPOMI inversions. We compared these scaled forward model
estimates to the prior forward model outputs.

Our results indicate that the TROPOMI scaled forward model provides a poorer fit to the BRW observations than the unscaled
prior forward model, as shown for 2019 in Figure 9. For the TROPOMI scaled forward model, the baseline is consistently
negatively offset for TROPOMI-WFMD and negatively offset in the summer months and positively offset in the early season
months for TROPOMI-OPER.

Additionally, forward models incorporating two of the sensitivity tests from Sect. 3.3 are shown in Figure S14 for August
2019: one vertical block boundary conditions, and Alaska observations only. The one vertical block test increases the baseline
closer to the observations for TROPOMI-WFMD. However, for TROPOMI-OPER the baseline decreases further from the
observations. Limiting the measurement dataset to observations over Alaska increased the baseline for TROPOMI-OPER. For
TROPOMI-WFMD, limiting measurements to Alaska only produced an anomalous trough in the baseline around 2019-08-15,
which requires further investigation.

3.5 Combined TROPOMI and Barrow inversion

We conducted combined TROPOMI and surface measurement inversions, incorporating an additional offset parameter to our
hierarchical model to account for any potential monthly biases between modelled satellite and surface mole fractions due to
calibration differences, as detailed in Sect. 2.6.

For August 2019, the mean inferred offset was 37 ppb between BRW and TROPOMI-WFMD, and 6 ppb between BRW
and TROPOMI-OPER. These values closely match the differences between the TROPOMI-scaled forward model and BRW
observations shown in Figure S14A for the same month, indicating that the offset parameter effectively corrects for this monthly
bias.

For the TROPOMI-OPER+BRW inversion, we observed the same derived tundra emissions to the TROPOMI-OPER inver-
sion within 95% uncertainty intervals for most months. However, in September 2019, the TROPOMI-OPER+BRW derived
tundra emissions increased compared to the BRW inversion and TROPOMI-OPER inversion outside of the 95% uncertainty
intervals (Figure 10A). The TROPOMI-WFMD+BRW derived tundra emissions were greater than TROPOMI-WFMD derived
tundra emissions, and closer to BRW derived tundra emissions for the summer and late season months of 2018, the late sea-
son months of 2019 and August 2020 (Figure 10B). Figure S6 shows the average seasonal difference between the posterior
and prior emissions from the tundra region for the TROPOMI-OPER+BRW. which are spatially similar to those from the
TROPOMI-OPER inversion. For the TROPOMI-WFMD+BRW inversion, emissions from the tundra region are larger than the
prior, whereas emissions from the TROPOMI-WFMD inversion are smaller across the tundra region.

The anthropogenic tundra emissions are shown in Figure S15. In May 2018, TROPOMI-OPER had exceptionally high
emissions. However, in the TROPOMI-OPER+BRW the emissions are substantially lower and align with the prior emissions.
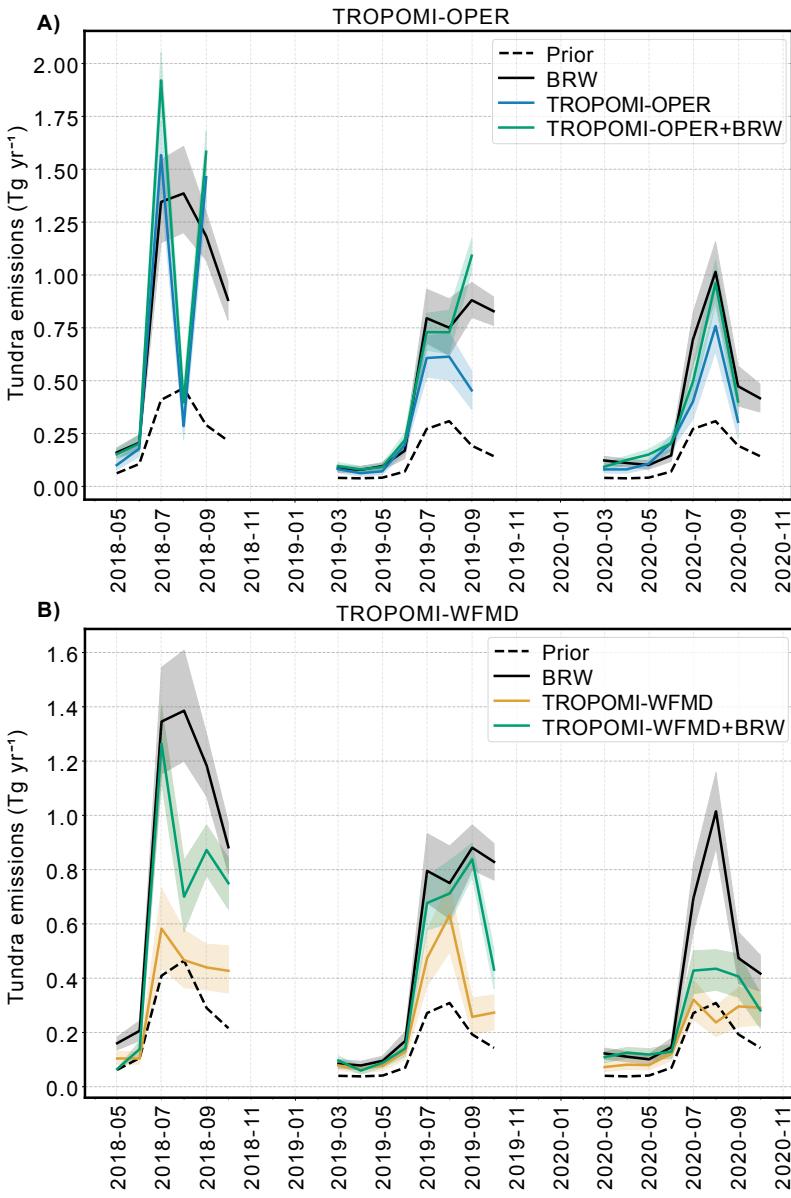
Figure 10. Combined TROPOMI+BRW inversions. A) and B) show the tundra sector emissions for the TROPOMI-OPER+BRW and TROPOMI-WFMD+BRW inversions, respectively.

Otherwise, all other months remain consistent with the TROPOMI-OPER only anthropogenic emissions, apart from June 2020 where the emissions decrease from the prior. For the TROPOMI-WFMD+BRW inversion, the derived anthropogenic emissions do not change from the TROPOMI-WFMD only inversion apart from May 2018, March 2020 and August 2020.

3.6 Fit of posterior modelled XCH$_4$ to measured XCH$_4$

The root mean square error (RMSE) between the modelled and measured CH$_4$ mole fractions is used to assess the fit of the model to the observations. Table 3 shows the RMSEs for TROPOMI observations, and Table 4 for BRW observations, for each inversion in this study. Both prior (before inference) and posterior (after inference) RMSEs are reported. A lowering of the
440   RMSE from the prior to the posterior for the same observations indicates a better fit, though it does not account for potential overfitting. Across all inversions, the RMSE decreases from prior to posterior for both TROPOMI and BRW, indicating improved agreement with the observations. TROPOMI-WFMD has consistently lower posterior RMSEs than TROPOMI-OPER, but also consistently higher prior RMSEs.

The RMSE ranges quantify the variability among the inversion experiments. For TROPOMI-OPER, the prior RMSE spans
445   23.7 to 33.7 ppb (a 10 ppb spread), which is larger than the corresponding TROPOMI-WFMD prior spread of 7.1 ppb (26.1 to 33.2 ppb). This indicates that TROPOMI-OPER is more sensitive to prior assumptions and inversion set-up. After assimilation, however, both products converge to similarly narrow posterior ranges. The OPER spread contracts to 2 ppb (16.1 to 18.1 ppb), an 80% reduction, and TROPOMI-WFMD to 2.6 ppb (14.5 to 17.1 ppb), a 63% reduction.

Changes in RMSE across the sensitivity tests reflect differences in the number of adjustable parameters and observations.
450   For instance, the Alaska-only inversion has fewer observations, resulting in lower RMSEs. The 421 grid cells aggregation results in lower RMSEs than the main inversion due to increased spatial flexibility, allowing more grid cells to be optimised. In contrast, the single vertical block test reduces the degrees of freedom for adjusting boundary conditions, leading to a higher RMSE.

For BRW, the BRW-only Alaska domain inversion has the lowest posterior RMSE. The larger Arctic domain inversion, and
455   the multiple sites and combined TROPOMI inversion, increases the RMSE.

Inversion	OPER prior	OPER posterior	WFMD prior	WFMD posterior
Main	24.8	17.3	26.2	16.0
Uniform prior	33.7	17.3	33.2	16.1
104 cells	23.7	17.2	26.2	16.0
421 cells	24.8	17.1	26.2	15.9
Alaska only	25.8	16.1	27.0	14.5
1 vertical block	24.8	18.1	26.1	17.1
TROPOMI+BRW	24.7	17.3	26.2	16.1

Table 3. TROPOMI RMSE (modelled mole fraction - measured mole fraction), in pbb, for each inversion in this study for prior modelled mole fraction and posterior modelled mole fraction. The RMSEs for the main inversion in Sect. 3.2, is shown in the first row. The following rows are for the sensitivity tests in Sect. 3.3.

Inversion	BRW prior	BRW posterior
BRW only (this study)	42.1	21.0
BRW only (Ward et al., 2024)	42.3	23.4
All surface (this study)	43.3	26.4
TROPOMI-OPER+BRW	43.2	24.6
TROPOMI-WFMD+BRW	43.5	24.3

Table 4. BRW RMSE (modelled mole fraction - measured mole fraction) for each inversion in this study for prior modelled mole fraction and posterior modelled mole fraction.

3.7 Comparison of inversion uncertainties

In our inversion method, the TROPOMI model-measurement error is a quadrature combination of the measurement error taken from the retrieval and an estimated model error (see Sect. 2.1). Figures S16–S18 show the monthly average measurement error, inferred model error (mean of all MCMC samples), and combined model-measurement error for BRW, TROPOMI-OPER, and TROPOMI-WFMD, respectively.

Over 2018–2020, the retrieval-based measurement error is much lower for TROPOMI-OPER (3 ppb; 1–6 ppb) than for TROPOMI-WFMD (15 ppb; 10–19 ppb; the reported values are the mean error over the full study period, with the numbers in parentheses representing the mean lower and upper bounds of the monthly 95% uncertainty intervals). The inferred model error is higher for TROPOMI-OPER (16 ppb; 13–21 ppb) than for TROPOMI-WFMD (5 ppb; 2–10 ppb). These differences lead to similar combined model-measurement errors for both products: 17 ppb (14–20 ppb) for TROPOMI-OPER and 16 ppb (11–21 ppb) for TROPOMI-WFMD. The combined model-measurement error for both TROPOMI products varies seasonally, with higher uncertainties during the early and late seasons compared to summer.

For the BRW-only inversions, the average model-measurement error is 13 ppb (5–25 ppb). When BRW is combined with TROPOMI, the model-measurement error increases to 18 ppb (9–32 ppb) in the TROPOMI-WFMD+BRW inversion and to 15 ppb (7–28 ppb) in the TROPOMI-OPER+BRW inversion. Across all BRW inversions, the model error component dominates the model-measurement error, and seasonal variation is less pronounced than using TROPOMI.

4 Discussion

This section evaluates the use of TROPOMI-derived CH_4 emissions for high-latitude inversions, with a particular focus on their applicability over the NSA. Sect. 4.1 compares the TROPOMI inversions with surface-based inversions. Sect. 4.2 evaluates the sensitivity of derived emissions to prior emissions and inversion set-up. Sect. 4.3 discusses the challenges of using TROPOMI

products in this region. Sect 4.4 discusses cold season TROPOMI coverage. Finally, sect 4.5 compares inversions in this study to previous TROPOMI inversions over the Arctic.

4.1 Emissions derived from TROPOMI and surface-based inversions

TROPOMI-derived emissions can be compared with emissions derived from surface inversions to assess whether different inversion approaches yield consistent results. Here, TROPOMI-derived emissions are considered consistent with surface inversion if they align with the BRW-derived emissions within 95% uncertainty intervals. However, as discussed in Sect. 2, the surface inversions are themselves sensitive to some aspects of the inversion set-up and inputs. Despite the limitations, in situ measurement-based inversions are a well-established method for emissions quantification in the Arctic and are valuable for comparison with TROPOMI inversions.

The main inversion in Sect. 3.2 shows that the TROPOMI-OPER tundra emissions are consistent with BRW-derived emissions for multiple months and exhibit similar spatial emission patterns. In contrast, TROPOMI-WFMD exhibits different spatial patterns and shows less consistency with BRW-derived emissions. While TROPOMI-OPER demonstrates greater consistency during the early season and summer, its coverage ends in September, missing the potentially high-emitting late-season period (Zona et al., 2016; Ward et al., 2024). In the combined inversions the TROPOMI-WFMD product shows improved agreement with BRW derived tundra emissions when combined with the BRW data, as shown in Sect. 3.5. For TROPOMI-OPER, combining the BRW data results in minimal changes to derived tundra emissions, though spatial emissions patterns more closely resemble those from the BRW only inversion. For both TROPOMI products, forward modelling (Sect. 3.4) indicates that combined inversions require an offset to account for bias between satellite and surface data. This is effectively handled using an offset parameter in 3.5.

Average NAME footprint sensitivities (Figure 3A) could potentially be used as an indicator of how strongly posterior emissions are constrained, with weaker footprints expected to produce fluxes closer to the prior. However, our results show that this relationship does not hold consistently for individual years. For example, in 2019 the TROPOMI-WFMD footprints are lower than those of TROPOMI-OPER, yet the TROPOMI-WFMD tundra emissions are least constrained to the prior in the year (Figure 5). In contrast, in 2018 and 2020 the TROPOMI-WFMD tundra emissions remain closer to the prior, even though the footprint magnitudes between the two products are more similar. However, the most pronounced spatial difference in sensitivity is that TROPOMI-OPER exhibits higher footprint values over southern NSA regions, which may allow emissions from these areas to be more strongly constrained. One explanation for the difference in the footprints between the two TROPOMI products is the number of observations remaining after the retrieval and QA filtering. The footprint sensitivities alone cannot explain the inversion result, which also depends on many factors including the retrieved column values, uncertainties, potential biases in the data, and the treatment of different vertical levels in NAME (as discussed in Section 5).

For the anthropogenic sector, the main inversion (Sect. 3.2) indicates that TROPOMI-WFMD derived emissions are more consistent with BRW derived emissions than TROPOMI-OPER, reversing the pattern observed for tundra emissions. Anthropogenic emissions in the NSA are dominated by fugitive releases at the Prudhoe Bay oil fields, where methane is emitted as discrete point sources rather than the widespread, diffuse emissions from tundra. The BRW station is situated approximately

300 km north-west from Prudhoe Bay, and measurements are less sensitive to emissions from this source compared to the tundra south of the station. In contrast, TROPOMI directly samples over Prudhoe Bay, offering stronger constraints when clear-sky observations are available. This is confirmed by the combined inversions, where adding BRW data to the TROPOMI-only inversions caused minimal changes to anthropogenic emissions, indicating TROPOMI provides the dominant constraint. However, despite more coverage over Prudhoe Bay oil fields, TROPOMI-WFMD and TROPOMI-OPER derived different anthropogenic emissions many months, indicating retrieval-driven differences.

The intra-annual variability of anthropogenic emissions remains uncertain. The prior (EDGAR v6) assumes no monthly variation, yet all inversions in this study show some degree of monthly variability, suggesting dynamics not captured by the inventory. Hu et al. (2025) identified a winter peak in oil and gas methane emissions in the contiguous U.S., likely linked to higher extraction and transmission rates during the heating season. State-level activity data from the U.S. Energy Information Administration show similar trends in Alaska, including a notable spike in May 2018 – consistent with high emissions inferred by TROPOMI-OPER (U.S. Energy Information Administration, 2025). TROPOMI-WFMD aligns most closely with overall seasonal activity patterns, while TROPOMI-OPER and BRW show elevated emissions during June–August, although not consistently in the same months. To further study these differences between the TROPOMI-OPER and BRW, we examined the BRW mole fractions and corresponding NAME footprints where the posterior emissions diverge. In August 2019 and 2020, where TROPOMI-OPER infers low anthropogenic emissions and BRW infers higher emissions, the posterior BRW mole fractions reproduce the observations well. However, in June 2019 and 2020, when TROPOMI-OPER infers higher anthropogenic emissions, the posterior modelled observations from BRW show a poorer fit. In some cases, this can be explained by NAME footprints exhibiting limited sensitivity over land, however, there are periods where the footprint patterns do not explain the mismatch, indicating additional factors may contribute. These diagnostics suggest that TROPOMI-OPER maybe be more sensitive to anthropogenic emissions than BRW in during some summer months. Lindqvist et al. (2024) reported a small positive bias in summer and a negative bias in autumn for TROPOMI-OPER relative to TCCON at other northern high-latitude stations, which could contribute to the lower anthropogenic emissions inferred by TROPOMI-OPER in August relative to June. However, without independent validation data over the North Slope of Alaska, it is not possible to determine whether the observed differences reflect genuine emission variability or retrieval artefacts.

Finally, updates in EDGAR 2024 (Crippa et al., 2024) suggest a ∼60% reduction in emissions totals over Alaska compared to EDGAR v6. If more accurate, it implies that the prior total used here may overstate actual emissions total, while the inferred seasonal variability remains unclear.

4.2 Sensitivity of derived emissions to priors and inversion set-up

To assess the robustness of TROPOMI-based inversions, sensitivity tests varying the inversion setup were conducted (Sect. 3.3). Derived emissions are considered robust if they do not change when the inversion set up changes, meaning that the results are stable and reliable across different modelling configurations. The stability would indicate that the inversion method and TROPOMI data can consistently produce accurate emissions estimates despite variations in the setup, thereby enhancing the confidence in the derived emissions. TROPOMI derived tundra emissions showed significant changes with alterations to the

prior seasonal cycle. Both the early season (March–May) and late season (September–October) periods were especially affected when the prior emissions were given a uniform seasonal profile with unrealistically high emissions during the cold season (see Table 2). This highlights the importance of using a well-informed prior, especially in the cold season, consistent with findings by Tenkanen et al. (2021). However, the BRW inversion was much less sensitive to the uniform prior and still produced plausible posterior emissions despite the unrealistic emissions in the cold season. Overall, the strong dependence of TROPOMI-derived tundra emissions on the prior suggests a lack of robustness, a concern that could extend to other high-latitude Arctic regions. Tests involving grid cell aggregations, boundary condition vertical blocks, and Alaska observations showed smaller or insignificant changes in derived tundra emissions. However, the grid cell aggregations sensitivity test showed spatial differences, with TROPOMI-OPER showing more spatial variation than TROPOMI-WFMD. The RMSE spread across the sensitivity tests shows that TROPOMI-OPER has higher prior sensitivity than TROPOMI-WFMD, but after the inversion, both products converge to similarly narrow posterior RMSE ranges, indicating that the inversion reduces experiment-to-experiment variability. For anthropogenic emissions, all sensitivity tests resulted in changes in TROPOMI-derived emissions for very few months. Unlike for tundra emissions, no sensitivity test directly altered the prior anthropogenic emissions, meaning their sensitivity to prior inputs was not assessed, and this should be included in future work.

An important consideration is how observations from outside the NSA influence derived emissions within it. For the surface inversions, excluding sites outside Alaska (i.e. using only BRW on the same domain) resulted in lower tundra emissions compared to the inversion using more sites. In contrast, for the TROPOMI inversions, restricting observations to Alaska alone, rather than the whole inversion domain, led to increased tundra emissions in some summer months, accompanied by a change in the optimised boundary conditions relative to the prior boundary conditions. Although these setups are not directly comparable, these contrasting results suggest that information from outside the region of interest plays different roles in the surface and satellite inversions.

4.3 TROPOMI retrieval and coverage limitations over the NSA

Data providers have corrected for bias in both TROPOMI products used in this study to improve accuracy at high latitudes, and the retrievals have been validated using XCH_4 measurements from TCCON sites (Lorente et al., 2021; Schneising et al., 2019; Lorente et al., 2023; Schneising et al., 2023). However, the geographic locations of TCCON, COCCON and AirCore sites relative to the NSA pose significant limitations for bias correction. The nearest TCCON site, Eureka in Nunavut, Canada, is about 2000 km away and the nearest AirCore measurements are in Europe. The closest ground-based column XCH_4 measurements come from a COCCON station in Fairbanks, Alaska, located approximately 300 km south of the NSA, and COCCON stations are not currently used for the correction or validation of either of the TROPOMI products used in this study. As Lorente et al. (2023) notes, anomalies have been detected in the TROPOMI-OPER product that are not near to TCCON stations, which is likely to affect the TROPOMI-WFMD product. Consequently, despite the presence of a few ground-based measurement stations in the Arctic, potential improvements or anomalies over the NSA, as well as other high-latitude regions such as the East Siberian Lowlands and the Taymyr Peninsula (Ward et al., 2024), could be overlooked.

Compared to previous estimates of emissions in the Arctic in August 2018 which typically has high emissions (Tenkanen et al., 2021; Wittig et al., 2023; Tsuruta et al., 2023), a pronounced trough is observed in TROPOMI-OPER tundra emissions over the NSA during August 2018. August is a month of low data coverage over the NSA (Figure S1), and this is likely due to persistent high cloud cover. Both TROPOMI-OPER and TROPOMI-WFMD implement cloud filtering using data from the Visible Infrared Imaging Radiometer Suite (VIIRS) onboard the Suomi-NPP satellite (Hutchison and Cracknell, 2005), to ensure high-quality retrievals under clear-sky conditions. During the 2018–2020 study period, August 2018 experienced the highest cloud coverage, with clouds present during 88% of daylight hours (Figure S19). The sparse data in August 2018 likely results in the inversion adhering closely to the prior in this month, producing lower emissions estimates than in surrounding months due to insufficient observational constraint. This underlines the need to continue and expand in situ measurements in important but persistently cloudy methane emitting regions. Future satellite missions, such as the Franco-German Methane Remote Sensing LIDAR Mission (MERLIN) (Ehret et al., 2017), planned to launch in 2028, will potentially have higher data density in cloudy scenes, as the high resolution LIDAR beam will reach the surface through gaps in optically thick clouds.

Despite lower reported measurement errors and prior RMSE, TROPOMI-OPER consistently results in higher posterior RMSE and larger inferred model error compared to TROPOMI-WFMD. This discrepancy might indicate that TROPOMI-OPER measurement errors are underestimated, leading the inversion to assign more of the model-measurement error to model errors over measurement errors. The greater spatial and temporal variability observed in TROPOMI-OPER scaling factors and monthly emissions estimates suggests it contains finer-scale variations that are more challenging for the inversion to resolve. In contrast, TROPOMI-WFMDs smoother spatial patterns and higher retrieval uncertainties may provide the inversion with more flexibility to fit the observations, resulting in lower posterior RMSE and smaller model error estimates.

4.4 Cold season coverage improvements and remaining challenges

The cold season in the Arctic tundra has been shown to be a larger contributor to annual CH_4 emissions than previously thought (Mastepanov et al., 2008; Zona et al., 2016; Kittler et al., 2017; Treat et al., 2018). Compared to previous satellite products, the increased data coverage of TROPOMI during this period, capturing high-resolution measurements at the start of the late season and end of the early season, potentially enables improved monitoring of this understudied season. This is especially valuable given the practical challenges of deploying additional in situ measurement sites due to harsh cold season conditions and the remoteness of Arctic tundra. The main inversions in this study found that after imposing prior tundra emissions with higher late season emissions derived from Zona et al. (2016), the posterior tundra emissions in September–October increased further from the prior, aligning with these previous studies that find unexpectedly large emissions from the cold season.

However, despite additional data coverage, the tundra emissions derived from the TROPOMI inversions demonstrated a lack of robustness in the early and late seasons. This lack of robustness may stem from several factors, including issues with XCH_4 retrievals during these periods due to higher solar zenith angles, albedo effects from snow cover, and seasonal stratospheric influences, such as the polar vortex. These challenges are reflected in the model errors derived in the inversions. In the TROPOMI inversions, the model errors are consistently higher during the early and late seasons compared to the growing season (Figures S17 and S18). In contrast, BRW model errors show year-to-year variability when model errors are highest

(Figure S16) – not consistently in the cold season. This indicates that there may be a systematic difference in how well the transport model captures boundary-layer versus upper-atmosphere dynamics.

4.5 Comparison to other northern high-latitude TROPOMI inversions

Tsuruta et al. (2023) conducted inversions using previous versions of the TROPOMI-WFMD (v1.2) and TROPOMI-OPER (v01.02.02 and v01.03.01) products for 2018 to assess the utility of TROPOMI products at high latitudes. Unlike our study, which concentrates on the NSA, their research divides the analysis into larger geographical areas circumpolar north of 45∘N. Their inversion uses the CTE-CH$_4$ inverse model, which employs an ensemble Kalman filter and an Eulerian atmospheric transport model, which significantly differs from that used here.

Thompson et al. (2025) recently applied a different LPDM, FLEXPART, to estimate methane emissions from the West Siberian Lowlands – another key northern high-latitude source region. To reduce computational costs, they developed a retrieval averaging routine that aggregates satellite observations onto a coarser spatial grid, based on the standard deviation of the original observations. This contrasts with our observation filtering approach, which prioritised measurements over our specific region of interest. However, their inversion domain is considerably smaller than ours, meaning that a larger proportion of available observations are relevant to their domain by default. The inversion system used by Thompson et al. (2025), FLEXINVERT, is a variational inversion method that uses a conjugate gradient minimisation with the FLEXPART LPDM.

A direct comparison of the derived emissions derived here and those from Tsuruta et al. (2023) and Thompson et al. (2025) is not possible due to differing geographical focusses. However, similar to Tsuruta et al. (2023), we find that TROPOMI-derived tundra (or natural in Tsuruta et al. (2023)) emissions are lower than those from in situ site-based inversions. Notably, our results show larger deviations from the prior than both Tsuruta et al. (2023) and Thompson et al. (2025), who report posterior emissions more closely aligned with their prior estimates. This discrepancy may reflect differences in the accuracy of the a priori emissions, or the relative strength of prior constraints in the inversion frameworks, which can limit how much the posterior can deviate from the prior.

The treatment of model–measurement errors differs between our inversion, Tsuruta et al. (2023) and Thompson et al. (2025), yet they all converge on similar total model–measurement uncertainties, with Tsuruta et al. (2023) imposing the highest minimum error. Tsuruta et al. (2023) fix their transport model error at 15 ppb and impose a 5 ppb minimum retrieval error (total $\geq$ 20 ppb). Thompson et al. (2025) prescribe combined model–measurement errors of 14–20 ppb based on retrieval errors and an estimate of background uncertainty. In contrast, our hierarchical inversion treats the model–measurement error as a hyperparameter that is inferred to reduce subjective assumptions and values fall in the 11–21 ppb range (Sec. 3.7).

In all three studies, it is assumed that the model-measurement errors are uncorrelated. For future TROPOMI inversions, Tsuruta et al. (2023) recommended increasing model-measurement uncertainty or accounting for model-measurement error correlation in both space and time. The HBMCMC method in this work assumes that uncertainties are uncorrelated both spatially and temporally. In reality, model-measurement errors in satellite data could be highly spatially correlated. From the transport model side, systematic biases in meteorological input data, such as wind speed and direction, can affect multiple measurements in the same area. Additionally, simplified parameterisation of fine-scale processes like turbulence and convection

can lead to spatially correlated errors. For satellite measurements, retrievals in close proximity could be impacted by surface albedo biases, such as from snow cover. Calibration issues, such as instrument calibration drifts and inconsistent calibration corrections, can introduce systematic biases over large areas. Data processing, such as spatial interpolation of XCH$_4$ values, can also introduce correlations by spreading errors from one location to adjacent areas.

As discussed previously, transport models exhibit seasonal biases and inaccuracies in CH$_4$ profiles at high latitudes (Tsuruta et al., 2023). Most TROPOMI-based inversions use Eulerian transport models, but this work, alongside Thompson et al. (2025), employs an LPDM. LPDMs become linearly computationally expensive with increased satellite data due to the rapid growth in the number of particles modelled as resolution increases. In contrast, most Eulerian models use a fixed spatial emissions grid, making them more efficient for large datasets (although some Eulerian models have used nested grids (Lunt et al., 2019; Nesser et al., 2024)). However, they also must account for the chemical loss of CH$_4$ through reactions with OH, introducing uncertainties due to the spatial and temporal variability of different OH fields (Zhao et al., 2020). Improving the computational cost of producing footprint sensitivities is the subject of ongoing work (e.g., Fillola et al., 2023), which could result in the use of TROPOMI data on larger spatial domains – such as full-Arctic model simulations – being feasible using LPDM backward-running simulations.

5   Potential improvements to inversions using TROPOMI

Given the challenges of installing additional in situ measurement stations at high latitudes in the near term, improving high-latitude inversions using TROPOMI data is crucial. This section discusses various strategies to refine TROPOMI inversions in this study, and more generally in terms of model development, especially related to vertical profiles and uncertainty estimates, and improvement to the retrieval products.

Improvements could be made to the modelling of atmospheric transport by the NAME LPDM (described in Sect. 2.4). Currently, our configuration of NAME does not simulate the atmosphere above 20 km. In our setup, particles are released up to the 20 km limit, and for levels above this, we assume the modelled XCH$_4$ concentrations match the corresponding levels in the CH$_4$ a priori profile (taken from the satellite retrieval). The TROPOMI-WFMD product is defined on 20 levels, using equal pressure intervals. The TROPOMI-OPER product, however, is defined on 12 equal pressure levels. Because of this configuration, the NAME model for TROPOMI-OPER was cut off after the 11th level at an equivalent altitude of 16 km. In contrast, the configuration for TROPOMI-WFMD allowed full modelling up to 20 km. Consequently, a larger portion of the TROPOMI-OPER column is approximated by the CH$_4$ a priori profile. Although stratospheric contributions to the estimation of XCH$_4$ are minor compared to the tropospheric contributions (Lindqvist et al., 2024; Tsuruta et al., 2025), it can play a major role in specific conditions such as the polar vortex, and may lead to systematic biases of estimated XCH$_4$ values. An improvement for estimating TROPOMI-OPER XCH$_4$ values would be to interpolate the retrieval onto levels extending up to 20 km, similar to those used by TROPOMI-WFMD. Or, a further improvement, which would mitigate the need to replace the top levels of the modelled XCH$_4$ with the CH$_4$ a priori profile, would be to run NAME to a much higher altitude. Running

NAME to a higher altitude would require a significant computational burden that was not feasible for this study, but should be considered for future work.

In this work, a priori boundary conditions are derived from CAMS v19r1 (Segers and Houweling, 2020), which only runs until 2019. Therefore, we apply the 2019 boundary conditions also for 2020. We acknowledge this as limitation, given the significant increase in global background concentrations from 2019 to 2020. However, the boundary conditions are optimised each month throughout the inversion, and the primary aim of this study is a comparison of satellite and surface observations in the same inversion framework, rather than to quantify absolute emissions. Future inversions will use more up-to-date versions of CAMS.

Not accounting for correlated errors could lead to an underestimation of TROPOMI model-measurement uncertainties, as discussed in Sect. 4.5, giving the inversion overestimated confidence in the transport model and TROPOMI observations. However, implementing spatial correlations requires introducing non-diagonal matrices. As our inversion method requires the inversion of matrices with every iteration, including spatial correlations would require further computational approximations. In the near term, implementing a minimum model-measurement error, such as in Tsuruta et al. (2023), could be a practical modification to ensure that uncertainties are not underestimated. Additionally, the measurement errors provided from the TROPOMI-OPER product are likely underestimated (mean of 3 ppb). This issue could be addressed to improve the reliability of the data product.

Enhanced filtering of TROPOMI measurements could be used to reduce noise in the retrievals. A blended albedo filter, as suggested by Wunch et al. (2011), which combines surface albedo in the NIR and SWIR to filter out snow-covered scenes, could be implemented to refine the TROPOMI data quality. Lorente et al. (2021) recommends that the albedo is limited to 0.85 to eliminate scenes that contribute to seasonal biases. This would require careful optimisation for the NSA, where there could be a large loss of data for the early and late seasons, which is a crucial for understanding NSA budget. Efforts should be made to improve retrieval quality outside of growing seasons in northern high latitudes. Since the inversions in this work were carried out, new versions of both TROPOMI-OPER and TROPOMI-WFMD have been released, which should be used in future inversions.

6 Conclusions

This study assessed the robustness of TROPOMI-derived CH_4 emissions using inverse modelling over the high Arctic. The focus was on tundra emissions from the North Slope of Alaska, comparing emissions derived using the TROPOMI-OPER and TROPOMI-WFMD products with those derived from in situ surface measurements, as well as multiple sensitivity tests. A comparison of TROPOMI-based inversions with surface-based inversions provides a valuable consistency check in regions lacking ground-based column XCH_4 validation. However, as demonstrated in this study, both satellite and surface-based inversions are sensitive to changes in inversion inputs, which has implications for other Arctic regions without ground-based validation data, such as in Siberia.

The TROPOMI-OPER-derived tundra emissions aligned with surface measurement-derived emissions for 15 of the 19 months in the inversion period (79%). TROPOMI-WFMD-derived emissions aligned with surface-derived emissions for only 8 of 22 months (36%). Both TROPOMI products derived anomalously low emissions in August 2018, compared to those derived using surface measurements, likely due to low data density due to high cloud coverage. Future satellite missions with active LIDAR sensors, such as MERLIN, may offer higher data coverage in broken-cloud conditions due to a small Lidar ground spot (Ehret et al., 2017). Sensitivity tests revealed that TROPOMI-derived tundra emissions are much more sensitive to the choice of prior emissions compared to the surface-derived emissions, indicating a lack of robustness in our TROPOMI inversions. Combining TROPOMI data with in situ observations increased alignment of tundra emission estimates for TROPOMI-WFMD but had minimal impact on TROPOMI-OPER.

Emissions from the anthropogenic sector were also assessed, with the largest emissions coming from fugitive emissions from the Prudhoe Bay oil fields. In this case, TROPOMI-WFMD-derived emissions agreed more closely with BRW-derived emissions than the TROPOMI-OPER-derived emissions. However, emissions derived from the combined inversions indicate that TROPOMI observations provide more constraint to anthropogenic emissions than the BRW observations. The intra-annual variability of anthropogenic emissions in the NSA remains uncertain, as the posterior emissions from each inversion diverge from one another and from the seasonally constant prior emissions inventory.

Compared to the strong influence of prior emissions, other sensitivity tests affected emissions in fewer months and to a lesser extent, but still highlighted important sensitivities in the model setup. Changing the grid cell aggregation affected both the total and spatial distribution of emissions in a some months, showing that grid configuration can influence both emission magnitudes and patterns. Resampling TROPOMI observations to include those only over Alaska led to higher summer emissions and indicated less constraint on the boundary conditions compared to using observations across the whole inversion domain. Finally, allowing for vertical variability in boundary condition optimisation had minimal impact on the resulting emissions.

Our work supports the recommendations of Tsuruta et al. (2023), including increasing the model-measurement uncertainty or accounting for model-measurement error correlation in both space and time. Furthermore, we suggest the use of a minimum measurement error for the TROPOMI-OPER product, given the low measurement errors provided. Additionally, we recommend that TROPOMI inversions in high northern latitudes using various inversion frameworks and transport models should be intercompared. These intercomparisons should employ the same prior emissions maps to isolate biases in the models and inversion frameworks. We encourage rigorous sensitivity testing with TROPOMI inversions, in order to evaluate the robustness of derived emissions, particularly regarding alterations to prior emissions, which, as demonstrated in this study, can have a significant impact. We encourage the continued use of LPDMs in TROPOMI inversions as this becomes more computationally feasible with the implementation of emulator models. Alongside this, it is crucial to continue and expand ground-based column measurements from instruments such as AirCore or TCCON, which are instrumental in validating the vertical profiles of the satellite retrievals and transport models.

The North Slope of Alaska is a key northern high-latitude region undergoing rapid environmental change, with methane emissions from natural sources already potentially increasing. While TROPOMI XCH$_4$ data is valuable for monitoring CH$_4$

emissions with inversions, the interpretation of posterior emissions requires careful consideration of the limitations to ensure accurate assessments in northern high latitudes.

745 *Code and data availability.* In-situ CH_4 mole fractions at BRW, TIK, INU and ESP are available for download from the ObsPack Data Portal at https://gml.noaa.gov/ccgg/obspack/ (last access: October 2024). The University of Bremen's WFM-DOAS v1.8 TROPOMI product is available for download from https://www.iup.uni-bremen.de/carbon_ghg/products/tropomi_wfmd/index_v18.php (last access: July 2023). The SRON Operational v2.04.00 TROPOMI product is available for download from https://ftp.sron.nl/open-access-data-2/TROPOMI/ tropomi/ch4/19_446/ (last access: August 2023). Emissions estimates from all inversions and sensitivity tests are available at https://doi.
750 org/10.5281/zenodo.16630729 (Ward, R. H., 2025). All the model inputs and code will be provided on request from the corresponding author.

Author contributions. RHW and ALG designed the research and RHW carried out the development of TROPOMI model code, experiments, visualisation, and original draft preparation with supervision from ALG. LMW contributed supervisory input throughout and provided critical review and feedback. RLT developed and supported the TROPOMI model code. EMF contributed the concept and implementation of the TROPOMI resampling routine. AT and TA provided critical review and feedback on the analysis and manuscript. All authors contributed to the manuscript through review and editing.

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

Acknowledgements. RHW was supported by a NERC GW4+ Doctoral Training Partnership studentship from the Natural Environment Research Council [NE/S007504/1]. We thank the teams responsible for the TROPOMI instrument, which was developed through a collaboration between Airbus Netherlands, KNMI, SRON, and TNO, and was commissioned by the Netherlands Space Office (NSO) and the European Space Agency (ESA). We thank those responsible for the development of the retrieval products - the operational product by SRON and the WFM-DOAS algorithm by the University of Bremen. We also thank Oliver Schneising and Michael Buchwitz of the WFM-DOAS team for discussions on TROPOMI data availability over the study region. We thank those at the National Oceanic and Atmospheric Administration (NOAA), Environment and Climate Change Canada (ECCC) and Finnish Meteorological Institute (FMI) that perform high-quality, continuous CH_4 measurements at the in-situ sites used in this study and make them publicly available to use via the ObsPack. We thank all those that have contributed to the Bristol Atmospheric Chemistry Research Group's code repository, past and present. The inversions and analysis in this work was carried out using the computational facilities of the Advanced Computing Research Centre at the University of Bristol (http://www.bris.ac.uk/acrc/). NAME model usage was made possible by the University of Bristol's NAME license from the UK Met Office and was run on JASMIN, the UK's collaborative data analysis environment (https://www.jasmin.ac.uk). We thank the European Space Agency ESRIN Contract No: 4000137895/22/I-AG MethaneCAMP, AO/1-10901/21/I-DT AMPAC-Net, ESA AO/1-11844/23/I-NS ESA SMART-CH4, Research Council Finland projects 337552 (ACCC), 359196 (FAME) and 364975 (CHARM), and EU-Horizon IM4CA (101183460), for financial support.

References

Alexe, M., Bergamaschi, P., Segers, A., Detmers, R., Butz, A., Hasekamp, O., Guerlet, S., Parker, R., Boesch, H., Frankenberg, C., Scheepmaker, R. A., Dlugokencky, E., Sweeney, C., Wofsy, S. C., and Kort, E. A.: Inverse modelling of CH_4 emissions for 2010–2011 using different satellite retrieval products from GOSAT and SCIAMACHY, Atmospheric Chemistry and Physics, 15, 113–133, https://doi.org/10.5194/acp-15-113-2015, 2015.

Berchet, A., Pison, I., Chevallier, F., Paris, J.-D., Bousquet, P., Bonne, J.-L., Arshinov, M. Y., Belan, B. D., Cressot, C., Davydov, D. K., Dlugokencky, E. J., Fofonov, A. V., Galanin, A., Lavrič, J., Machida, T., Parker, R., Sasakawa, M., Spahni, R., Stocker, B. D., and Winderlich, J.: Natural and anthropogenic methane fluxes in Eurasia: a mesoscale quantification by generalized atmospheric inversion, Biogeosciences, 12, 5393–5414, https://doi.org/10.5194/bg-12-5393-2015, 2015.

Bergamaschi, P., Frankenberg, C., Meirink, J. F., Krol, M., Villani, M. G., Houweling, S., Dentener, F., Dlugokencky, E. J., Miller, J. B., Gatti, L. V., Engel, A., and Levin, I.: Inverse modeling of global and regional CH4 emissions using SCIAMACHY satellite retrievals, Journal of Geophysical Research: Atmospheres, 114, https://doi.org/10.1029/2009JD012287, 2009.

Box, J. E., Colgan, W. T., Christensen, T. R., Schmidt, N. M., Lund, M., Parmentier, F.-J. W., Brown, R., Bhatt, U. S., Euskirchen, E. S., Romanovsky, V. E., Walsh, J. E., Overland, J. E., Wang, M., Corell, R. W., Meier, W. N., Wouters, B., Mernild, S., Mård, J., Pawlak, J., and Olsen, M. S.: Key indicators of Arctic climate change: 1971–2017, Environmental Research Letters, 14, https://doi.org/10.1088/1748-9326/aafc1b, 2019.

Buchwitz, M., de Beek, R., Noël, S., Burrows, J. P., Bovensmann, H., Bremer, H., Bergamaschi, P., Körner, S., and Heimann, M.: Carbon monoxide, methane and carbon dioxide columns retrieved from SCIAMACHY by WFM-DOAS: year 2003 initial data set, Atmospheric Chemistry and Physics, 5, 3313–3329, https://doi.org/10.5194/acp-5-3313-2005, 2005.

Crippa, M., Solazzo, E., Huang, G., Guizzardi, D., Koffi, E., Muntean, M., Schieberle, C., Friedrich, R., and Janssens-Maenhout, G.: High resolution temporal profiles in the Emissions Database for Global Atmospheric Research, Scientific Data, 7, 121, https://doi.org/10.1038/s41597-020-0462-2, 2020.

Crippa, M., Guizzardi, D., Pagani, F., Banja, M., Muntean, M., Schaaf, E., Monforti-Ferrario, F., Becker, W., Quadrelli, R., Risquez Martin, A., Taghavi-Moharamli, P., Köykkä, J., Grassi, G., Rossi, S., Melo, J., Oom, D., Branco, A., San-Miguel, J., Manca, G., Pisoni, E., Vignati, E., and Pekar, F.: GHG emissions of all world countries, Tech. rep., Luxembourg (Luxembourg), ISBN 978-92-68-20572-3 (online),978-92-68-20573-0 (print), 2024.

Degen, J., Baier, B., Blazaki, K., Chen, H., Engel, A., Heikkinen, P., Karhu, J., Kivi, R., Leuenberger, M., Meixner, K., Nyfeler, P., Sweeney, C., Heuven, S. v., Zanchetta, A., and Laube, J.: Short-term variability and uncertainties of trace gases in the boreal summer UTLS from AirCore measurements during the OSTRICH campaign, https://doi.org/10.5194/egusphere-egu24-15443, 2024.

Dlugokencky, E. J., Steele, L. P., Lang, P. M., and Masarie, K. A.: Atmospheric methane at Mauna Loa and Barrow observatories: Presentation and analysis of in situ measurements, Journal of Geophysical Research: Atmospheres, 100, 23 103–23 113, https://doi.org/10.1029/95JD02460, 1995.

Ehret, G., Bousquet, P., Pierangelo, C., Alpers, M., Millet, B., Abshire, J., Bovensmann, H., Burrows, J., Chevallier, F., Ciais, P., Crevoisier, C., Fix, A., Flamant, P., Frankenberg, C., Gibert, F., Heim, B., Heimann, M., Houweling, S., Hubberten, H., Jöckel, P., Law, K., Löw, A., Marshall, J., Agusti-Panareda, A., Payan, S., Prigent, C., Rairoux, P., Sachs, T., Scholze, M., and Wirth, M.: MERLIN: A French-German Space Lidar Mission Dedicated to Atmospheric Methane, Remote Sensing, 9, 1052, https://doi.org/10.3390/rs9101052, 2017.

Fillola, E., Santos-Rodriguez, R., Manning, A., O'Doherty, S., and Rigby, M.: A machine learning emulator for Lagrangian particle dispersion model footprints: a case study using NAME, Geoscientific Model Development, 16, 1997–2009, https://doi.org/10.5194/gmd-16-1997-2023, 2023.

Frey, M., Sha, M. K., Hase, F., Kiel, M., Blumenstock, T., Harig, R., Surawicz, G., Deutscher, N. M., Shiomi, K., Franklin, J. E., Bösch, H., Chen, J., Grutter, M., Ohyama, H., Sun, Y., Butz, A., Mengistu Tsidu, G., Ene, D., Wunch, D., Cao, Z., Garcia, O., Ramonet, M., Vogel, F., and Orphal, J.: Building the COllaborative Carbon Column Observing Network (COCCON): long-term stability and ensemble performance of the EM27/SUN Fourier transform spectrometer, Atmospheric Measurement Techniques, 12, 1513–1530, https://doi.org/10.5194/amt-12-1513-2019, 2019.

Ganesan, A. L., Rigby, M., Zammit-Mangion, A., Manning, A. J., Prinn, R. G., Fraser, P. J., Harth, C. M., Kim, K.-R., Krummel, P. B., Li, S., Mühle, J., O'Doherty, S. J., Park, S., Salameh, P. K., Steele, L. P., and Weiss, R. F.: Characterization of uncertainties in atmospheric trace gas inversions using hierarchical Bayesian methods, Atmospheric Chemistry and Physics, 14, 3855–3864, https://doi.org/10.5194/acp-14-3855-2014, 2014.

Ganesan, A. L., Rigby, M., Lunt, M. F., Parker, R. J., Boesch, H., Goulding, N., Umezawa, T., Zahn, A., Chatterjee, A., Prinn, R. G., Tiwari, Y. K., van der Schoot, M., and Krummel, P. B.: Atmospheric observations show accurate reporting and little growth in India's methane emissions, Nature Communications, 8, 836, https://doi.org/10.1038/s41467-017-00994-7, 2017.

Gelman, A. and Rubin, D. B.: Inference from Iterative Simulation Using Multiple Sequences, Statistical Science, 7, 457–472, https://doi.org/10.1214/ss/1177011136, 1992.

Houweling, S., Krol, M., Bergamaschi, P., Frankenberg, C., Dlugokencky, E. J., Morino, I., Notholt, J., Sherlock, V., Wunch, D., Beck, V., Gerbig, C., Chen, H., Kort, E. A., Röckmann, T., and Aben, I.: A multi-year methane inversion using SCIAMACHY, accounting for systematic errors using TCCON measurements, Atmospheric Chemistry and Physics, 14, 3991–4012, https://doi.org/10.5194/acp-14-3991-2014, 2014.

Houweling, S., Bergamaschi, P., Chevallier, F., Heimann, M., Kaminski, T., Krol, M., Michalak, A. M., and Patra, P.: Global inverse modeling of CH_4 sources and sinks: an overview of methods, Atmospheric Chemistry and Physics, 17, 235–256, https://doi.org/10.5194/acp-17-235-2017, 2017.

Hoffman, M. D. and Gelman, A.: The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo, The Journal of Machine Learning Research, 15, 1593–1623, https://doi.org/10.5555/2627435.2638586, 2014.

Hu, H., Hasekamp, O., Butz, A., Galli, A., Landgraf, J., Aan De Brugh, J., Borsdorff, T., Scheepmaker, R., and Aben, I.: The operational methane retrieval algorithm for TROPOMI, Atmospheric Measurement Techniques, 9, 5423–5440, https://doi.org/10.5194/amt-9-5423-2016, 2016.

Hu, L., Andrews, A. E., Montzka, S. A., Miller, S. M., Bruhwiler, L., Oh, Y., Sweeney, C., Miller, J. B., McKain, K., Ibarra Espinosa, S., Davis, K., Miles, N., Mountain, M., Lan, X., Crotwell, A., Madronich, M., Mefford, T., Michel, S., and Houwelling, S.: An Unexpected Seasonal Cycle in U.S. Oil and Gas Methane Emissions, Environmental Science & Technology, https://doi.org/10.1021/acs.est.4c14090, 2025.

Hugelius, G., Ramage, J., Burke, E., Chatterjee, A., Smallman, T. L., Aalto, T., Bastos, A., Biasi, C., Canadell, J. G., Chandra, N., Chevallier, F., Ciais, P., Chang, J., Feng, L., Jones, M. W., Kleinen, T., Kuhn, M., Lauerwald, R., Liu, J., López-Blanco, E., Luijkx, I. T., Marushchak, M. E., Natali, S. M., Niwa, Y., Olefeldt, D., Palmer, P. I., Patra, P. K., Peters, W., Potter, S., Poulter, B., Rogers, B. M., Riley, W. J., Saunois, M., Schuur, E. a. G., Thompson, R. L., Treat, C., Tsuruta, A., Turetsky, M. R., Virkkala, A.-M., Voigt, C., Watts, J., Zhu, Q., and

Zheng, B.: Permafrost Region Greenhouse Gas Budgets Suggest a Weak CO2 Sink and CH4 and N2O Sources, But Magnitudes Differ Between Top-Down and Bottom-Up Methods, Global Biogeochemical Cycles, 38, https://doi.org/10.1029/2023GB007969, 2024.

Hutchison, K. D. and Cracknell, A. P.: Visible Infrared Imager Radiometer Suite: A New Operational Cloud Imager, CRC Press, Boca Raton, ISBN 978-0-429-11856-2, https://doi.org/10.1201/9781420023398, 2005.

850   IPCC: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)], Cambridge University Press, Cambridge, UK and New York, NY, USA, p. 755, https://doi.org/10.1017/9781009157964, 2019.

Ishizawa, M., Chan, D., Worthy, D., Chan, E., Vogel, F., and Maksyutov, S.: Analysis of atmospheric CH4 in Canadian Arctic and estimation of the regional CH4 fluxes, Atmospheric Chemistry and Physics, 19, 4637–4658, https://doi.org/10.5194/acp-19-4637-2019, 2019.

855   Jacob, D. J., Turner, A. J., Maasakkers, J. D., Sheng, J., Sun, K., Liu, X., Chance, K., Aben, I., McKeever, J., and Frankenberg, C.: Satellite observations of atmospheric methane and their value for quantifying methane emissions, Atmospheric Chemistry and Physics, 16, 14 371–14 396, https://doi.org/10.5194/acp-16-14371-2016, 2016.

Jensen, K. and Mcdonald, K.: Surface Water Microwave Product Series Version 3: A Near-Real Time and 25-Year Historical Global Inundated Area Fraction Time Series From Active and Passive Microwave Remote Sensing, IEEE Geoscience and Remote Sensing Letters,
860   16, 1402–1406, https://doi.org/10.1109/LGRS.2019.2898779, 2019.

Jones, A., Thomson, D., Hort, M., and Devenish, B.: The U.K. Met Office's Next-Generation Atmospheric Dispersion Model, NAME III, in: Air Pollution Modeling and Its Application XVII, edited by Borrego, C. and Norman, A.-L., pp. 580–589, Springer US, Boston, MA, 2007.

Kittler, F., Heimann, M., Kolle, O., Zimov, N., Zimov, S., and Göckede, M.: Long-Term Drainage Reduces CO 2 Uptake and CH 4 Emis-
865   sions in a Siberian Permafrost Ecosystem: Drainage impact on Arctic carbon cycle, Global Biogeochemical Cycles, 31, 1704–1717, https://doi.org/10.1002/2017GB005774, 2017.

Lan, X. and Dlugokencky, E. J.: Trends in globally-averaged CH4, N2O, and SF6 determined from NOAA Global Monitoring Laboratory measurements. Version 2024-07, https://doi.org/https://doi.org/10.15138/P8XG-AA10, 2022.

Lan, X., Nisbet, E. G., Dlugokencky, E. J., and Michel, S. E.: What do we know about the global methane budget? Results from four decades
870   of atmospheric CH4 observations and the way forward, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 379, https://doi.org/10.1098/rsta.2020.0440, 2021.

Laurila, T.: Atmospheric CH4 product, Tiksi (10.0 m), 2010-09-01–2022-04-09, European ObsPack, https://hdl.handle.net/11676/pQuLCZ0dgFVBHp0ZiSHB4vzm, 2024.

Lindqvist, H., Kivimäki, E., Häkkilä, T., Tsuruta, A., Schneising, O., Buchwitz, M., Lorente, A., Martinez Velarte, M., Borsdorff, T., Alberti,
875   C., Backman, L., Buschmann, M., Chen, H., Dubravica, D., Hase, F., Heikkinen, P., Karppinen, T., Kivi, R., McGee, E., Notholt, J., Rautiainen, K., Roche, S., Simpson, W., Strong, K., Tu, Q., Wunch, D., Aalto, T., and Tamminen, J.: Evaluation of Sentinel-5P TROPOMI Methane Observations at Northern High Latitudes, Remote Sensing, 16, 2979, https://doi.org/10.3390/rs16162979, 2024.

Lorente, A., Borsdorff, T., Butz, A., Hasekamp, O., aan de Brugh, J., Schneider, A., Wu, L., Hase, F., Kivi, R., Wunch, D., Pollard, D. F., Shiomi, K., Deutscher, N. M., Velazco, V. A., Roehl, C. M., Wennberg, P. O., Warneke, T., and Landgraf, J.: Methane retrieved from
880   TROPOMI: improvement of the data product and validation of the first 2 years of measurements, Atmospheric Measurement Techniques, 14, 665–684, https://doi.org/10.5194/amt-14-665-2021, 2021.

Lorente, A., Borsdorff, T., Martinez-Velarte, M. C., and Landgraf, J.: Accounting for surface reflectance spectral features in TROPOMI methane retrievals, Atmospheric Measurement Techniques, 16, 1597–1608, https://doi.org/10.5194/amt-16-1597-2023, 2023.

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.

Lunt, M. F., Palmer, P. I., Feng, L., Taylor, C. M., Boesch, H., and Parker, R. J.: An increase in methane emissions from tropical Africa between 2010 and 2016 inferred from satellite data, Atmospheric Chemistry and Physics, 19, 14 721–14 740, https://doi.org/10.5194/acp-19-14721-2019, 2019.

Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Scarpelli, T. R., Nesser, H., Sheng, J.-X., Zhang, Y., Hersher, M., Bloom, A. A., Bowman, K. W., Worden, J. R., Janssens-Maenhout, G., and Parker, R. J.: Global distribution of methane emissions, emission trends, and OH concentrations and trends inferred from an inversion of GOSAT satellite data for 2010–2015, Atmospheric Chemistry and Physics, 19, 7859–7881, https://doi.org/10.5194/acp-19-7859-2019, 2019.

Mastepanov, M., Sigsgaard, C., Dlugokencky, E. J., Houweling, S., Ström, L., Tamstorf, M. P., and Christensen, T. R.: Large tundra methane burst during onset of freezing, Nature, 456, 628–630, https://doi.org/10.1038/nature07464, 2008.

Narbaud, C., Paris, J.-D., Wittig, S., Berchet, A., Saunois, M., Nédélec, P., Belan, B. D., Arshinov, M. Y., Belan, S. B., Davydov, D., Fofonov, A., and Kozlov, A.: Disentangling methane and carbon dioxide sources and transport across the Russian Arctic from aircraft measurements, Atmospheric Chemistry and Physics, 23, 2293–2314, https://doi.org/10.5194/acp-23-2293-2023, 2023.

Neal, R. M.: Slice sampling, The Annals of Statistics, 31, 705–767, https://doi.org/10.1214/aos/1056562461, 2003.

Nesser, H., Jacob, D. J., Maasakkers, J. D., Lorente, A., Chen, Z., Lu, X., Shen, L., Qu, Z., Sulprizio, M. P., Winter, M., Ma, S., Bloom, A. A., Worden, J. R., Stavins, R. N., and Randles, C. A.: High-resolution U.S. methane emissions inferred from an inversion of 2019 TROPOMI satellite data: contributions from individual states, urban areas, and landfills, Atmospheric Chemistry and Physics, pp. 1–36, https://doi.org/10.5194/acp-24-5069-2024, 2024.

O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS CO_2 retrieval algorithm – Part 1: Description and validation against synthetic observations, Atmospheric Measurement Techniques, 5, 99–121, https://doi.org/10.5194/amt-5-99-2012, 2012.

Ostler, A., Sussmann, R., Rettinger, M., Deutscher, N. M., Dohe, S., Hase, F., Jones, N., Palm, M., and Sinnhuber, B.-M.: Multistation intercomparison of column-averaged methane from NDACC and TCCON: impact of dynamical variability, Atmospheric Measurement Techniques, 7, 4081–4101, https://doi.org/10.5194/amt-7-4081-2014, 2014.

Ostler, A., Sussmann, R., Patra, P. K., Houweling, S., De Bruine, M., Stiller, G. P., Haenel, F. J., Plieninger, J., Bousquet, P., Yin, Y., Saunois, M., Walker, K. A., Deutscher, N. M., Griffith, D. W. T., Blumenstock, T., Hase, F., Warneke, T., Wang, Z., Kivi, R., and Robinson, J.: Evaluation of column-averaged methane in models and TCCON with a focus on the stratosphere, Atmospheric Measurement Techniques, 9, 4843–4859, https://doi.org/10.5194/amt-9-4843-2016, 2016.

Outcalt, S. I., Nelson, F. E., and Hinkel, K. M.: The zero-curtain effect: Heat and mass transfer across an isothermal region in freezing soil, Water Resources Research, 26, 1509–1516, https://doi.org/10.1029/WR026i007p01509, 1990.

Pandey, S., Houweling, S., Krol, M., Aben, I., Chevallier, F., Dlugokencky, E. J., Gatti, L. V., Gloor, E., Miller, J. B., Detmers, R., Machida, T., and Röckmann, T.: Inverse modeling of GOSAT-retrieved ratios of total column CH_4 and CO_2 for 2009 and 2010, Atmospheric Chemistry and Physics, 16, 5043–5062, https://doi.org/10.5194/acp-16-5043-2016, 2016.

Paris, J.-D., Belan, B., Ancellet, G., Nédélec, P., Arshinov, M., Pruvost, A., Berchet, A., Arzoumanian, E., Pison, I., Ciais, P., and Law, K.: YAK-AEROSIB: continental scale aircraft measurements of trace gases gases above Siberia, 2014.

Parker, R., Boesch, H., Cogan, A., Fraser, A., Feng, L., Palmer, P. I., Messerschmidt, J., Deutscher, N., Griffith, D. W. T., Notholt, J., Wennberg, P. O., and Wunch, D.: Methane observations from the Greenhouse Gases Observing SATellite: Comparison to ground-based TCCON data and model calculations, Geophysical Research Letters, 38, https://doi.org/10.1029/2011GL047871, 2011.

Qu, Z., Jacob, D. J., Shen, L., Lu, X., Zhang, Y., Scarpelli, T. R., Nesser, H., Sulprizio, M. P., Maasakkers, J. D., Bloom, A. A., Worden, J. R., Parker, R. J., and Delgado, A. L.: Global distribution of methane emissions: a comparative inverse analysis of observations from the TROPOMI and GOSAT satellite instruments, Atmospheric Chemistry and Physics, 21, 14 159–14 175, https://doi.org/10.5194/acp-21-14159-2021, 2021.

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, 1–10, https://doi.org/10.1038/s43247-022-00498-3, 2022.

Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond, P. A., Dlugokencky, E. J., Houweling, S., Patra, P. K., Ciais, P., Arora, V. K., Bastviken, D., Bergamaschi, P., Blake, D. R., Brailsford, G., Bruhwiler, L., Carlson, K. M., Carrol, M., Castaldi, S., Chandra, N., Crevoisier, C., Crill, P. M., Covey, K., Curry, C. L., Etiope, G., Frankenberg, C., Gedney, N., Hegglin, M. I., Höglund-Isaksson, L., Hugelius, G., Ishizawa, M., Ito, A., Janssens-Maenhout, G., Jensen, K. M., Joos, F., Kleinen, T., Krummel, P. B., Langenfelds, R. L., Laruelle, G. G., Liu, L., Machida, T., Maksyutov, S., McDonald, K. C., McNorton, J., Miller, P. A., Melton, J. R., Morino, I., Müller, J., Murguia-Flores, F., Naik, V., Niwa, Y., Noce, S., O'Doherty, S., Parker, R. J., Peng, C., Peng, S., Peters, G. P., Prigent, C., Prinn, R., Ramonet, M., Regnier, P., Riley, W. J., Rosentreter, J. A., Segers, A., Simpson, I. J., Shi, H., Smith, S. J., Steele, L. P., Thornton, B. F., Tian, H., Tohjima, Y., Tubiello, F. N., Tsuruta, A., Viovy, N., Voulgarakis, A., Weber, T. S., van Weele, M., van der Werf, G. R., Weiss, R. F., Worthy, D., Wunch, D., Yin, Y., Yoshida, Y., Zhang, W., Zhang, Z., Zhao, Y., Zheng, B., Zhu, Q., Zhu, Q., and Zhuang, Q.: The Global Methane Budget 2000–2017, Earth System Science Data, 12, 1561–1623, https://doi.org/10.5194/essd-12-1561-2020, 2020.

Schneising, O., Buchwitz, M., Reuter, M., Bovensmann, H., Burrows, J. P., Borsdorff, T., Deutscher, N. M., Feist, D. G., Griffith, D. W. T., Hase, F., Hermans, C., Iraci, L. T., Kivi, R., Landgraf, J., Morino, I., Notholt, J., Petri, C., Pollard, D. F., Roche, S., Shiomi, K., Strong, K., Sussmann, R., Velazco, V. A., Warneke, T., and Wunch, D.: A scientific algorithm to simultaneously retrieve carbon monoxide and methane from TROPOMI onboard Sentinel-5 Precursor, Atmospheric Measurement Techniques, 12, 6771–6802, https://doi.org/10.5194/amt-12-6771-2019, 2019.

Schneising, O., Buchwitz, M., Hachmeister, J., Vanselow, S., Reuter, M., Buschmann, M., Bovensmann, H., and Burrows, J. P.: Advances in retrieving XCH$_4$ and XCO from Sentinel-5 Precursor: improvements in the scientific TROPOMI/WFMD algorithm, Atmospheric Measurement Techniques, 16, 669–694, https://doi.org/10.5194/amt-16-669-2023, 2023.

Schuldt, K. N., Aalto, T., Andrade, M., Arlyn Andrews, Apadula, F., Jgor Arduini, Arnold, S., Baier, B., Bäni, L., Bartyzel, J., Bergamaschi, P., Biermann, T., Biraud, S. C., Pierre-Eric Blanc, Boenisch, H., Brailsford, G., Brand, W. A., Brunner, D., Bui, T. P. V., Van Den Bulk, P., Benoit Burban, Francescopiero Calzolari, Chang, C. S., Huilin Chen, Lukasz Chmura, St. Clair, J. M., Sites Climadat, Coletta, J. D., Colomb, A., Condori, L., Conen, F., Conil, S., Couret, C., Cristofanelli, P., Cuevas, E., Curcoll, R., Daube, B., Davis, K. J., Dean-Day, J. M., Delmotte, M., Ankur Desai, DiGangi, E., DiGangi, J. P., Elsasser, M., Emmenegger, L., Forster, G., Frumau, A., Fuente-Lastra,

960   M., Galkowski, M., Gatti, L. V., Gehrlein, T., Gerbig, C., Francois Gheusi, Gloor, E., Goto, D., Hammer, S., Hanisco, T. F., Haszpra, L., Hatakka, J., Heimann, M., Heliasz, M., Heltai, D., Henne, S., Hensen, A., Hermans, C., Hermansen, O., Hoheisel, A., Holst, J., Di Iorio, T., Iraci, L. T., Ivakhov, V., Jaffe, D. A., Jordan, A., Joubert, W., Kang, H.-Y., Karion, A., Kazan, V., Keeling, R. F., Keronen, P., Kers, B., Jooil Kim, Klausen, J., Kneuer, T., Ko, M.-Y., Kolari, P., Kominkova, K., Kort, E., Kozlova, E., Krummel, P. B., Kubistin, D., Kulawik, S. S., Kumps, N., Labuschagne, C., Lan, X., Langenfelds, R. L., Lanza, A., Laurent, O., Laurila, T., Lauvaux, T., Lavric,

965   J., Choong-Hoon Lee, Lee, J., Lehner, I., Lehtinen, K., Leppert, R., Leskinen, A., Leuenberger, M., Levin, I., Levula, J., Lindauer, M., Lindroth, A., Mikaell Ottosson Löfvenius, Loh, Z. M., Lopez, M., Lowry, D., Lunder, C. R., Machida, T., Mammarella, I., Manca, G., Manning, A., Marek, M. V., Marklund, P., Marrero, J. E., Martin, D., Martin, M. Y., Giordane A. Martins, Matsueda, H., De Mazière, M., McKain, K., Meinhardt, F., Menoud, M., Jean-Marc Metzger, Miles, N. L., Miller, C. E., Miller, J. B., Mölder, M., Monteiro, V., Montzka, S., Moore, F., Moossen, H., Moreno, C., Morgan, E., Josep-Anton Morgui, Morimoto, S., Müller-Williams, J., Munro, D.,

970   Mutuku, M., Myhre, C. L., Jaroslaw Necki, Nichol, S., Nisbet, E., Niwa, Y., Njiru, D. M., Noe, S. M., O'Doherty, S., Obersteiner, F., Parworth, C. L., Peltola, O., Peters, W., Philippon, C., Piacentino, S., Pichon, J. M., Pickers, P., Pitt, J., Pittman, J., Plass-Dülmer, C., Platt, S. M., Popa, M. E., Prinzivalli, S., Ramonet, M., Richardson, S. J., Louis-Jeremy Rigouleau, Rivas, P. P., Röckmann, T., Rothe, M., Yves-Alain Roulet, Ju-Mee Ryoo, Santoni, G., Di Sarra, A. G., Sasakawa, M., Schaefer, H., Scheeren, B., Schmidt, M., Schuck, T., Schumacher, M., Seifert, T., Sha, M. K., Shepson, P., Daegeun Shin, Sloop, C. D., Smale, D., Smith, P. D., Sørensen, L. L., De Souza,

975   R. A. F., Spain, G., Steger, D., Steinbacher, M., Stephens, B., Sweeney, C., Taipale, R., Takatsuji, S., Thoning, K., Timas, H., Torn, M., Trisolino, P., Turnbull, J., Van Der Veen, C., Vermeulen, A., Viner, B., Vitkova, G., De Vries, M., Watson, A., Weiss, R., Weyrauch, D., Wofsy, S. C., Worsey, J., Worthy, D., Xueref-Remy, I., Yates, E. L., Dickon Young, Yver-Kwok, C., Zaehle, S., Zahn, A., Zazzeri, G., Zellweger, C., and Miroslaw Zimnoch: Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1983-2023; obspack_ch4_1_GLOBALVIEWplus_v7.0_2024-10-29, https://doi.org/10.25925/20241001, 2024.

980 Schuur, E. A. G., Pallandt, M., and Göckede, M.: Russian collaboration loss risks permafrost carbon emissions network, Nature Climate Change, pp. 1–2, https://doi.org/10.1038/s41558-024-02001-6, 2024.

Segers, A. and Houweling, S.: CAMS global inversion-optimised greenhouse gas fluxes and concentrations, version v19r1, Atmosphere Data Store (ADS), https://doi.org/10.24381/ed2851d2, accessed: 2020-12-28, 2020.

Sweeney, C., Dlugokencky, E., Miller, C. E., Wofsy, S., Karion, A., Dinardo, S., Chang, R. Y.-W., Miller, J. B., Bruhwiler, L., Crotwell,

985   A. M., Newberger, T., McKain, K., Stone, R. S., Wolter, S. E., Lang, P. E., and Tans, P.: No significant increase in long-term CH4 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.

Tenkanen, M., Tsuruta, A., Rautiainen, K., Kangasaho, V., Ellul, R., and Aalto, T.: Utilizing Earth Observations of Soil Freeze/Thaw Data and Atmospheric Concentrations to Estimate Cold Season Methane Emissions in the Northern High Latitudes, Remote Sensing, 13,

990   https://doi.org/10.3390/rs13245059, 2021.

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.

Thompson, R. L., Krishnankutty, N., Pisso, I., Schneider, P., Stebel, K., Sasakawa, M., Stohl, A., and Platt, S.: Efficient use of a La-

995   grangian Particle Dispersion Model for atmospheric inversions using satellite observations of column mixing ratios, EGUsphere, pp. 1–28, https://doi.org/10.5194/egusphere-2025-147, 2025.

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

Treat, C. C., Bloom, A. A., and Marushchak, M. E.: Nongrowing season methane emissions–a significant component of annual emissions across northern ecosystems, Global Change Biology, 24, 3331–3343, https://doi.org/10.1111/gcb.14137, 2018.

Tsuruta, A., Kivimäki, E., Lindqvist, H., Karppinen, T., Backman, L., Hakkarainen, J., Schneising, O., Buchwitz, M., Lan, X., Kivi, R., Chen, H., Buschmann, M., Herkommer, B., Notholt, J., Roehl, C., Té, Y., Wunch, D., Tamminen, J., and Aalto, T.: CH4 Fluxes Derived from Assimilation of TROPOMI XCH4 in CarbonTracker Europe-CH4: Evaluation of Seasonality and Spatial Distribution in the Northern High Latitudes, Remote Sensing, 15, 1620, https://doi.org/10.3390/rs15061620, 2023.

Tsuruta, A., Kuze, A., Shiomi, K., Kataoka, F., Kikuchi, N., Aalto, T., Backman, L., Kivimäki, E., Tenkanen, M. K., McKain, K., García, O. E., Hase, F., Kivi, R., Morino, I., Ohyama, H., Pollard, D. F., Sha, M. K., Strong, K., Sussmann, R., Te, Y., Velazco, V. A., Vrekoussis, M., Warneke, T., Zhou, M., and Suto, H.: Global CH Fluxes Derived from JAXA/GOSAT Lower Tropospheric Partial Column Data and the CTE-CH Atmospheric Inverse Model, EGUsphere, 2025, 1–46, https://doi.org/10.5194/egusphere-2025-159, 2025.

Turner, A. J., Jacob, D. J., Wecht, K. J., Maasakkers, J. D., Lundgren, E., Andrews, A. E., Biraud, S. C., Boesch, H., Bowman, K. W., Deutscher, N. M., Dubey, M. K., Griffith, D. W. T., Hase, F., Kuze, A., Notholt, J., Ohyama, H., Parker, R., Payne, V. H., Sussmann, R., Sweeney, C., Velazco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D.: Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data, Atmospheric Chemistry and Physics, 15, 7049–7069, https://doi.org/10.5194/acp-15-7049-2015, 2015.

Turner, A. J., Frankenberg, C., and Kort, E. A.: Interpreting contemporary trends in atmospheric methane, Proceedings of the National Academy of Sciences, 116, 2805–2813, https://doi.org/10.1073/pnas.1814297116, 2019.

U.S. Energy Information Administration: Alaska Natural Gas Gross Withdrawals (Million Cubic Feet), https://www.eia.gov/dnav/ng/hist/n9010ak2M.htm, 2025.

Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Carslaw, K., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Jones, C., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whitall, M., Williams, K., and Zerroukat, M.: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, Geoscientific Model Development, 12, 1909–1963, https://doi.org/10.5194/gmd-12-1909-2019, 2019.

Wang, F., Maksyutov, S., Tsuruta, A., Janardanan, R., Ito, A., Sasakawa, M., Machida, T., Morino, I., Yoshida, Y., Kaiser, J. W., Janssens-Maenhout, G., Dlugokencky, E. J., Mammarella, I., Lavric, J. V., and Matsunaga, T.: Methane Emission Estimates by the Global High-Resolution Inverse Model Using National Inventories, Remote Sensing, 11, 2489, https://doi.org/10.3390/rs11212489, 2019.

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, https://doi.org/10.1029/2024JD040766, 2024.

Ward, R. H.: CH4 emissions estimates from "Uncertainty and retrieval sensitivity in TROPOMI-based methane inversions over the North Slope of Alaska" [data set], https://doi.org/10.5281/zenodo.16630729, 2025.

Weber, T., Wiseman, N. A., and Kock, A.: Global ocean methane emissions dominated by shallow coastal waters, Nature Communications, 10, 2019.

Wittig, S., Berchet, A., Pison, I., Saunois, M., Thanwerdas, J., Martinez, A., Paris, J.-D., Machida, T., Sasakawa, M., Worthy, D. E. J., Lan, X., Thompson, R. L., Sollum, E., and Arshinov, M.: Estimating Methane Emissions in the Arctic nations using surface observations from 2008 to 2019, Atmospheric Chemistry and Physics, 23, 6457–6485, https://doi.org/10.5194/acp-23-6457-2023, 2023.

Wittig, S., Berchet, A., Pison, I., Saunois, M., and Paris, J.-D.: Surface networks in the Arctic may miss a future *methane bomb*, Atmospheric Chemistry and Physics, pp. 6359–6373, https://doi.org/10.5194/acp-24-6359-2024, 2024.

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The Total Carbon Column Observing Network, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 369, 2087–2112, https://doi.org/10.1098/rsta.2010.0240, 2011.

Zhang, Y., Jacob, D. J., Lu, X., Maasakkers, J. D., Scarpelli, T. R., Sheng, J.-X., Shen, L., Qu, Z., Sulprizio, M. P., Chang, J., Bloom, A. A., Ma, S., Worden, J., Parker, R. J., and Boesch, H.: Attribution of the accelerating increase in atmospheric methane during 2010–2018 by inverse analysis of GOSAT observations, Atmospheric Chemistry and Physics, 21, 3643–3666, https://doi.org/10.5194/acp-21-3643-2021, 2021.

Zhao, Y., Saunois, M., Bousquet, P., Lin, X., Berchet, A., Hegglin, M. I., Canadell, J. G., Jackson, R. B., Dlugokencky, E. J., Langenfelds, R. L., Ramonet, M., Worthy, D., and Zheng, B.: Influences of hydroxyl radicals (OH) on top-down estimates of the global and regional methane budgets, Atmospheric Chemistry and Physics, 20, https://doi.org/10.5194/acp-20-9525-2020, 2020.

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