https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


# High-resolution remote sensing and machine-learning-based upscaling of methane fluxes: a case study in the Western Canadian tundra

Kseniia Ivanova[1], Anna-Maria Virkkala[2], Victor Brovkin[3], Tobias Stacke[3], Barbara Widhalm[4], Annett Bartsch[4], Carolina Voigt[5,6,7], Oliver Sonnentag[7], and Mathias Göckede[1]

[1] Department for Biogeochemical Signals, Max Planck Institute for Biogeochemistry, Jena, Germany
[2] Woodwell Climate Research Centre, Falmouth, MA, USA
[3] Department Climate Dynamics, Max Planck Institute for Meteorology, Hamburg, Germany
[4] b.geos, Industriestrasse 1, 2100 Korneuburg, Austria
[5] Permafrost Research Section, Alfred Wegener Institute Helmholtz Center for Polar and Marine Research, Potsdam, Germany
[6] Institute of Soil Science, University of Hamburg, Hamburg, Germany
[7]Université de Montréal, Montréal, Québec, Canada

*Correspondence to*: Kseniia Ivanova (kivanova@bgc-jena.mpg.de)

**Abstract.** Arctic methane ($CH_4$) budgets are uncertain because field measurements often capture only fragments of the wet-to-dry gradient that control tundra $CH_4$ fluxes. Wet hotspots are over-represented, while dry, net-sink sites are under-sampled. We paired over 13,000 chamber flux measurements during peak growing season in July (2019-2024) from Trail Valley Creek in the western Canadian Arctic with co-registered remotely sensed predictor variables to test how spatial resolution (1 m vs. 10 m) and choice of machine-learning algorithm shape upscaled $CH_4$ flux maps over our 3.1 km$^2$ study domain. Four algorithms for $CH_4$ flux scaling (Random Forest (RF), Gradient Boosting Machine (GBM), Generalised Additive Model (GAM), and Support Vector Regression (SVR)) were tuned using the same stack of multispectral indices, terrain derivatives and a six-class landscape classification. Tree-based models such as RF and GBM offered the best balance of 10-fold cross-validated $R^2$ ($\leq 0.75$) and errors, so RF and GBM were used in a subsequent step for upscaling to the study area. With 1 m resolution, GBM captured the full range of microtopographic extremes and predicted a mean July flux of 99 mg $CH_4$ m$^{-2}$ month$^{-1}$. In contrast, RF, which smoothed local extremes, yielded an average flux of 519 mg $CH_4$ m$^{-2}$ month$^{-1}$. The disagreement between flux estimates using GBM and RF correlated mainly with the Normalized Difference Water Index (NDWI), a moisture proxy, and was most pronounced in waterlogged, low-lying areas. Aggregating predictors to 10 m averaged the sharp metre-scale flux highs in hollows and lows on ridges, narrowing the GBM-RF difference to ~75 mg $CH_4$ m$^{-2}$ month$^{-1}$ while broadening the overall flux distribution with more intermediate values. At 1 m, microtopography is the main driver. At 10 m, moisture proxies explained about half of the variance. Our results demonstrate that: (i) sub-metre predictors are indispensable for capturing the wet-dry microtopography and its $CH_4$ signals, (ii) upscaling algorithm selection strongly controls prediction spread and uncertainty once that microrelief is resolved, and (iii) coarser grids smooth local microtopographic details, resulting in flattened $CH_4$ flux peaks and wider distribution. All factors combined lead to potentially large differences in scaled $CH_4$ flux budgets, calling for a careful selection of scaling approaches, spatial predictor layers (e.g., vegetation, moisture, topography), and grid resolution. Future work should couple ultra-high-resolution imagery with temporally dynamic indices to reduce upscaling bias along Arctic wetness gradients.

## 1 Introduction

The Arctic is warming nearly four times faster than the global average due to Arctic amplification feedbacks (Previdi et al., 2021; AMAP, 2021; Rantanen et al., 2022; Ballinger et al., 2020). This rapid warming is of particular concern due to the substantial quantities of organic carbon stored in wetland ecosystems of the circumpolar permafrost region (Hugelius et al., 2014; Schuur et al., 2015; Turetsky et al., 2020; Olefeldt et al., 2016). Thaw exposure may mobilize part of the previously frozen carbon as methane ($CH_4$), a greenhouse gas 28-34 times more potent than $CO_2$ over 100 years (Koven et al., 2011; Etminan et al., 2016; Nisbet et al., 2019; Saunois et al., 2020). Rising temperatures, therefore, risk to trigger a positive feedback in which permafrost degradation elevates $CH_4$ emissions, further intensifying warming (Schuur et al., 2015; Walter Anthony et al., 2018; Turetsky et al., 2020; Natali et al., 2021).

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


High-resolution $CH_4$ flux measurements in tundra ecosystems remain sparse even during the growing season due to the Arctic's remoteness, harsh climate, and logistical challenges (e.g., lengthy travel times, high fieldwork costs, sparse infrastructure, and challenging equipment maintenance), which limits the number of long-term monitoring sites. The primary tools for plot- to ecosystem scale $CH_4$ flux observations are flux chambers (Subke et al., 2021) and eddy covariance techniques, respectively (Matthes et al., 2014; Baldocchi, 2003); however, the time window to conduct growing season chamber campaigns is usually limited to a few months between June and September, and locations in the Arctic featuring eddy covariance towers are few (Vogt et al., 2025). As a consequence, most synthesis studies aiming at constraining $CH_4$ budgets in the high northern latitudes must rely on a limited database biased toward high-emitting sites near research stations and often overlooking areas with net $CH_4$ uptake (Mastepanov et al., 2013; Varner et al., 2021; Kuhn et al., 2021; Voigt et al., 2023c). Most tundra chamber campaigns collect data only for short intervals, typically from a single day up to a few weeks during the growing season, and many are conducted in just one growing season without repeated multi-year sampling or covering winter fluxes, which limits their value for model benchmarking and interannual analysis (Varner et al., 2021; Kuhn et al., 2021; Räsänen et al., 2021; Mastepanov et al., 2013; Treat et al., 2018).

Even where flux data exist, $CH_4$ fluxes can shift within metres because the relative position and seasonal movement of the water table and the frost table create mosaics of anoxic ($CH_4$ - producing) and oxic ($CH_4$ - oxidising) soil (Frolking et al., 2011). These redox contrasts are further modulated by microtopography, plant functional type, and surface moisture (Mastepanov et al., 2013; Pirk et al., 2015; Olefeldt et al., 2021). Because the water table and frost table rarely coincide at the same depth across tundra microtopography, neighbouring microsites can experience very different oxic–anoxic conditions. Across the Arctic tundra, surface types range from water-saturated zones, such as sedge fens, polygon centres, troughs and thaw slumps, to better-drained features like hummocky ridges, palsas and gravelly uplands. These elements cover the entire $CH_4$ flux range, with microtopographically lower, wetter zones acting as strong sources and microtopographically elevated, better-aerated zones often functioning as net sinks (Räsänen et al., 2021; Bao et al., 2021; Yuan et al., 2024). Such small-scale heterogeneity frequently occurs within a single 10 m pixel, so coarse maps or remote-sensing data products can combine zones of strong $CH_4$ emission and neighbouring areas that act as net $CH_4$ sinks (Knox et al., 2019; Treat et al., 2018). Without spatially explicit methods that resolve this fine-scale heterogeneity, upscaling can introduce systematic biases. It may overestimate $CH_4$ emissions when dry areas that act as sinks are overlooked or underestimate them when narrow wet trenches surrounding dry patches are missed (Räsänen et al., 2021; Treat et al., 2018).

Ultra-high-resolution (~1 m) imagery from drones or commercial satellites can directly resolve fine-scale vegetation patterns and microtopographic features (e.g., hummocks and hollows) in heterogeneous tundra landscapes, for example mapping plant communities on dry polygon rims versus wet sedge hollows and other microrelief features that correspond to $CH_4$ "hotspots" and "cold spots", respectively. Studies show that 1 m data products capture these plot-scale variations and can reveal flux heterogeneity at meter scales (Wangari et al., 2023; Ludwig et al., 2024; Lehmann et al., 2016; Ström et al., 2005; Davidson et al., 2017; Becker et al., 2008). However, working with spatially ultra-high-resolution data presents significant challenges. The acquisition and processing of sub-meter imagery through drones or advanced satellites and LiDAR are both costly and labour-intensive; such datasets are rarely available as dense, multi-date image stacks and cannot be easily collected over large areas (Scheller et al., 2022; Karim et al., 2024; Anderson & Gaston, 2013). Moreover, ultra-high resolution can even introduce noise and not necessarily lead to a better representation of environmental conditions (Riihimäki et al., 2021).

By contrast, high resolution (~10 m) predictors such as Sentinel-2 multispectral imagery and ArcticDEM terrain products are freely available and cover the entire Arctic with regular revisits with a standardized approach (Drusch et al., 2012; Porter et al., 2023). However, the coarse 10 m resolution has a clear disadvantage because individual microtopographic features (e.g., hummocks, hollows) and landforms (e.g., dry palsas, wet trenches, etc.) that control small-scale variability in $CH_4$ fluxes are aggregated into single pixels, blurring the fine-scale patterns of emission and uptake (Räsänen & Virtanen, 2019).

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


Machine-learning (ML) models such as random forests (RF), gradient-boosting machines (GBM), generalized additive models (GAM), and support-vector regression (SVR) can integrate predictors derived from remote sensing products with flux measurements to upscale $CH_4$ from plot- and ecosystem- to landscape scales (Yuan et al., 2024; Chen et al., 2024; Zhang et al., 2020, Ying et al. 2025). Tree ensembles (RF, GBM) are particularly well suited for capturing complex interactions and handle multicollinearity, while GAMs have the advantage of yielding interpretable smooth functions, and SVR excels with limited nonlinear data (Wood, 2017; Smola & Schölkopf, 2004; Zhang et al., 2019). Model choice, predictor resolution and limited training data still generate large spreads in upscaled Arctic tundra $CH_4$ fluxes, with ensemble estimates differing by roughly 25–50 % of the mean depending on the study (Peltola et al., 2019; McNicol et al., 2023; Chen et al., 2024; Räsänen et al., 2021). Quantifying and reducing these uncertainties are essential for robust $CH_4$ budgets.

Here, we address these methodological challenges in a study aiming at upscaling CH4 emissions in a heterogeneous tundra landscape in the western Canadian Arctic by pairing >13,000 peak growing season (July) chamber measurements collected over five years with matched 1 m and 10 m remote sensing predictors and training four ML algorithms (RF, GBM, GAM, SVR). Our overarching aim is to reduce uncertainties in peak-season (July) $CH_4$ budgets for the 3.1 km² heterogeneous tundra around the Trail Valley Creek Research Station. We address this aim through four specific questions:

- Which remotely-sensed vegetation, moisture, and topographic characteristics best explain July $CH_4$ fluxes across a wet-to-dry micro-site gradient?
- Does replacing freely available 10 m data (Sentinel-2, ArcticDEM) with sub-metre imagery from drones and airborne lidar lead to a detectable improvement in prediction accuracy and spatial detail?
- How do the four ML algorithms differ in predicted net flux magnitudes and spatial patterns?
- How do model choice, grid resolution, and their interaction shape the spatial patterns and uncertainty of our upscaled $CH_4$ flux maps?

Optimizing a data-driven upscaling approach based on these questions allows us to produce July $CH_4$ flux maps with pixel-level uncertainty, improving peak-season emission estimates and guiding where additional measurements or higher-resolution imagery would most reduce prediction error.

## 2 Materials and Methods

### 2.1 Study site

The study site is the undulating tundra landscape of the Trail Valley Creek (TVC) Research Station, about 55 km north of the town of Inuvik, NT, in the western Canadian Arctic east of the Mackenzie River Delta (Fig. 1). TVC lies in the Southern Arctic ecozone and contains continuous permafrost, with thickness ranging from 100 to 150 m (Marsh et al., 2008). Our analyses focus on a ~3.1 km² section of this 57 km$^2$ basin with elevations ranging from 41 to 102 m a.s.l. The 1991 - 2020 climate normals for Inuvik are a mean annual air temperature of –7 °C, mean annual precipitation of ~250 mm, and a frost-free period (the interval with minimum air temperatures above 0 °C) of roughly 78 days (Environment and Climate Change Canada, 2024). The vegetation at TVC is highly diverse, reflecting the microtopography and moisture gradients. Isolated patches of white and black spruce (*Picea glauca*, P. *mariana*) occur in valley bottoms and on slopes. Tall shrub tundra, dominated by green alder (*Alnus alnobetula*) and featuring scattered willows and dwarf birch, can be found on hill slopes and alongside streams. Riparian zones feature dense willow thickets reaching up to 2 metres in height. Upland areas support dwarf shrub tundra with dense stands of dwarf birch (*Betula glandulosa*), Labrador tea (*Ledum palustre*) and mountain cranberry (*Vaccinium vitis-idaea*), interspersed with mosses and lichens. Flat, poorly drained areas are dominated by tussock-forming sedges (*Eriophorum* and *Carex*), alongside moss and scattered shrubs. Exposed uplands and polygon rims are covered by lichen mats and low dwarf shrubs. Mosses, especially *Sphagnum* and *Polytrichum* species, are prevalent in wetter microhabitats. Snow depth and winter soil temperatures

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


are highest in the tall shrub and tussock zones and lowest in the lichen tundra (Grünberg et al., 2020; Marsh et al., 2010).

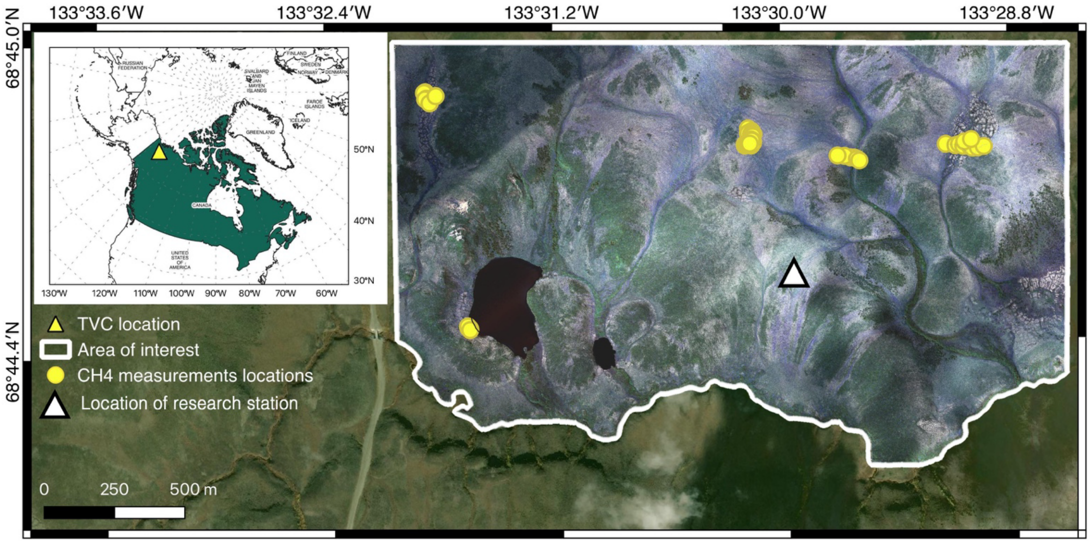


**Figure 1. Map of the study area showing the location of the area of interest (outlined with white polygon), with CH4 flux measurement locations marked with yellow circles. The inset map in the upper left corner highlights the region in which the Trail Valley Creek (TVC) research station is located (marked with a yellow triangle on the overview map and a white triangle on the detailed map). Background satellite imagery sources: © Maxar 2025, provided by Esri, acquired on 12 July 2024. Area of interest aerial imagery: Rettelbach et al., 2024.**

### 2.2 Data sources

#### 2.2.1 $CH_4$ flux data

We used a combination of continuous and campaign-based $CH_4$ flux measurements to capture spatial and temporal variability in $CH_4$. The dataset includes previously published automated chamber observations made in 2019 and 2021 (Voigt et al., 2023a), and campaign-based manual chamber observations made in 2019 (Voigt et al., 2023b) and in 2022 to 2024 (Ivanova et al., 2025). Manual chamber measurements from 2022 to 2024 were collected as part of this study. The main measurement protocols, chamber specifications, instrumentation, and flux calculation methods for each campaign are summarized in Table 1.

**Table 1. Summary of $CH_4$ flux measurement protocols and instrumentation used at TVC, 2019–2024.**

| **Year** | 2019, 2021 | 2019 | 2022–2024 |
|---|---|---|---|
| **Method** | Automated | Manual | Manual |
| **Number of microsites** | 18 | 13 | 37 |
| **Chamber size and shape** | 30-45 L, hemispherical | 17 L, cylindrical | 17 L, cylindrical |
| **Gas analyzer** | Los Gatos Research Enhanced Performance Greenhouse Gas Analyser (Rackmount GGA-24EP 911-0010, Los Gatos) | Picarro G4301 GasScouter (Picarro, Inc., Santa Clara, CA, USA) | LI-COR LI-7810 Trace Gas Analyzer (LI-COR Biosciences, Lincoln, NE, USA) |
| **Measurement frequency** | 1 Hz | 1 Hz | 1 Hz |
| **Enclosure time** | 3 min | 5 min | 2–4 min |
| **Flux calculation method** | Linear regression (default); exponential fit for large fluxes | Linear or nonlinear regression with the Math Works Inc., Natick, MA, USA) | Linear regression with bootstrapping (R) |
| **Reference** | Voigt et al., 2023a | Voigt et al., 2023b | Ivanova & Göckede, 2025 |

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


The complete dataset included 13,384 $CH_4$ flux measurements collected between 1 July and 31 July (2019-2024) under both light and dark conditions. Of these, 1,093 fluxes were measured manually using closed chambers, while 12,291 were collected using an automated chamber system (Fig. 1). For each flux measurement, ancillary data recorded include coordinates, PAR (measured as photosynthetic photon flux density (PPFD; µmol $m^{-2}$ $s^{-1}$), air temperature, land cover type, and time of day (when available).

### 2.2.2 Climatic data

Air temperature (AT) data were obtained from the Trail Valley Creek meteorological station operated by Environment and Climate Change Canada – Meteorological Service of Canada (ECCC, 2024). The station is located within the study area at 68°44′46.8″ N, 133°30′06.4″ W, at an elevation of 85 m a.s.l. (Climate ID: 220N005; WMO ID: 71683; TC ID: XTV). The original data were recorded at hourly resolution and were downsampled to 3-hour intervals to match the temporal resolution of the model predictions.

### 2.2.3 Remotely sensed data

We assembled two separate but equivalent predictor stacks, one with a cell size of 1 m and one with 10 m. Both cover the same area of interest (AOI, Fig. 1), use the same map projection, and pass through the same preprocessing code (Ivanova et al., 2025). We defined the AOI along natural drainage lines on three sides and the Inuvik-Tuktoyaktuk Highway that forms its western border. An image stack is a set of co-registered raster layers (multispectral indices and terrain derivatives) that share the same grid and extent. The 1 m stack is based on the RGB + NIR orthomosaic captured on 22 August 2018 by Rettelbach et al. (2024) and on the 1 m LiDAR-derived digital terrain model (DTM) from Lange et al. (2021). From these layers we derived the Normalized Difference Vegetation Index (NDVI, Rouse 1974) and the Normalized Difference Water Index (NDWI, Gao, 1996; McFeeters, 1996) as proxies for biomass and surface moisture, respectively. Sentinel-2 index extraction of NDVI and NDWI was implemented in Google Earth Engine (Gorelick et al., 2017). Slope, aspect, the topographic position index (TPI, 30 m window), and the topographic wetness index (TWI) were calculated with Whitebox Tools (Lindsay, 2016). The 10 m stack contains the same set of variables but at coarser resolution. It contains data from Sentinel-2 Level-2A scenes collected between 2015 and 2024 (Copernicus, 2024). ArcticDEM (2 m) was resampled to 10 m to match this grid. All cloud-free Sentinel-2 scenes from July – August 2018 were averaged to create a composite that aligns with the 2018 drone campaign. A complete overview of all predictor variables, including data descriptions, resolution, temporal variability, and references, is provided in Appendix Table A1.

Based on these predictors (multispectral indices and terrain derivatives), we produced a site-specific landscape classification map at both 1 m and 10 m resolution using a Random Forest approach (Breiman, 2001). The same training dataset and classification parameters were applied to both resolutions to ensure comparability. The classification is based on six land cover types in Grünberg et al. (2020) and includes: Water, Lichens, Tussock, Dwarf shrubs, Tall shrubs + trees, and Sedges. Reference points for training were selected using a combination of our own field observations. At 1 m resolution, the Tall shrubs + trees class was merged with Dwarf shrubs due to the absence of chamber measurements within that class, resulting in five effective classes at 1 m and six at 10 m. Water pixels were masked prior to analysis, and all statistical comparisons were restricted to terrestrial classes. A full description of the classification workflow is provided in Text A1 and Table A2. Same as for NDVI and NDWI, Sentinel-2 index extraction for landscape classification was implemented in Google Earth Engine.

In addition, we also explored the potential of two datasets that are particularly relevant for Arctic-scale applications. The Circumarctic Land cover Units (CALU) (Bartsch et al., 2024) provides a 10 m classification of vegetation physiognomy and soil moisture regimes across the circumpolar Arctic tundra. The data product is based on the fusion of Sentinel-1 and Sentinel-2 imagery and was calibrated using over 3,500 field samples of soil and vegetation properties. One of the key strengths of CALU is that it captures spatial gradients in surface wetness while using a consistent classification scheme across all Arctic regions.

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


This makes it possible to directly compare classes between distant sites across the Arctic, which is rarely achievable with site-specific classifications. 20 of 23 land cover units are found across the AOI, but only 5 of those were covered by $CH_4$ measurements. The complete legend of CALU classes used in this study, including definitions, their occurrence within the AOI, and whether $CH_4$ flux measurements are available for each class, is presented in Table A3. Additionally, we considered a radar interferometric (InSAR) dataset derived from Sentinel-1 data for 2018 - 2023 (Widhalm et al. 2025), which captures seasonal ground subsidence rates in thawing degree days domain associated with thaw table (the uppermost soil that freezes and thaws each year). The magnitude of the subsidence rates reflects soil moisture gradients (Widhalm et al. 2025).

Finally, we assessed the benefit of incorporating time-specific spectral indices (NDVI and NDWI) extracted from Sentinel-2 scenes close to each chamber measurement. We compared the effect of using these time-matched indices versus a seasonal composite (July – August 2018) to test whether short-term variability in vegetation and moisture status improves model skill. Although this approach relies on satellite scenes taken within a limited time window and may not align perfectly with the exact in situ measurement date, it still offers a more detailed representation of changes in surface conditions than seasonal averages.

All four of the additional predictors (CALU, InSAR, and temporal NDVI and NDWI) were included in sensitivity tests to evaluate their predictive power. However, they were not included in the main upscaling workflow or resolution comparison, as they are only available at a single resolution and would thus bias inter-resolution comparisons.

To isolate the effects of scale, we selected 1 m and 10 m resolutions, as chamber fluxes represent sub-metre patches, whereas most Arctic land cover products have coarser resolutions (>10 m) (Bartsch et al., 2016). By keeping the processing steps identical and changing only the grid size, we were able to compare the results directly. As shown in Fig. 2, narrow, wet features such as polygonal trenches are captured at 1 m but blended at 10 m, which alters both the NDVI and the landscape classification. Since many high $CH_4$ fluxes originate from these small wet zones, aggregation at a coarser resolution obscures their contribution. Some of the remaining model disagreement may also be due to the limited representation of extremely wet or complex terrain in the training data, which reduces the model's generalisability.

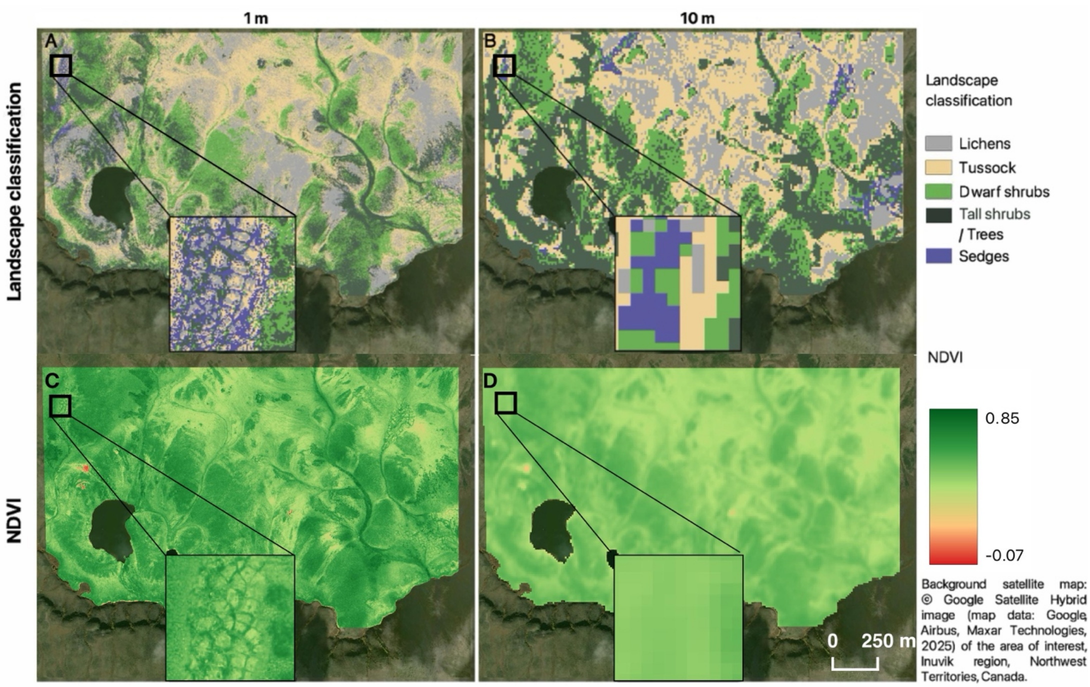


**Figure 2. Site-specific landscape classification (LC) and Normalized Difference Vegetation Index (NDVI) at two spatial resolutions: 1 m (panels A and C) and 10 m (panels B and D). Panels A and B show LC maps, while panels C and D show NDVI. Each panel includes a black-framed inset highlighting a representative polygonal mire. Narrow, waterlogged microtopographic features such as wet**

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


**trenches remain distinct at 1 m resolution but blend into mixed pixels at 10 m. Background imagery: © Google Satellite Hybrid (Maxar, 2025).**

### 2.3 Statistical analyses

The statistical analysis was structured into five sequential stages: (1) data preparation, (2) model training and evaluation, (3) spatial prediction, (4) temporal aggregation and interpretation, and (5) variable importance analysis (Figure 3). All steps were applied identically to the 1 m and 10 m datasets to enable direct comparison of model behaviour and prediction outcomes across spatial resolutions. The overall workflow is summarized in Fig. 3. The analysis was implemented in R 4.3.2 (R Core Team, 2024).
The first step consisted in the preparation of the predictor datasets to explain spatio-temporal variability in $CH_4$ fluxes. In total ten predictors were used: air temperature (AT), photosynthetically go radiation (PAR), cumulative thawing-degree-days (TDD), NDVI, NDWI, slope, aspect, the 30 m topographic position index, the topographic wetness index (TWI) and a six-class landscape classification (see Appendix Table A1 for details).
Second, we evaluated four modelling families for their ability to predict $CH_4$ fluxes: random forests (RF), gradient-boosting machines (GBM), generalized additive models (GAM), and support-vector regression (SVR). RF is a ML algorithm that builds multiple decision trees on bootstrapped data. The mean of their outputs is then calculated. The averaging reduces noise and the method reports easy-to-read variable-importance scores (Breiman 2001; Prasad et al. 2006). GBM also uses trees but adds them one after another. Each new tree learns from the errors of the current ensemble, which often reduces bias but requires careful tuning to avoid over-fitting (Friedman 2001; Elith et al. 2008). Similar RF, GBM handles mixed predictor types, outliers, missing values, and nonlinear relationships without preprocessing (Elith et al., 2008). GAM is a statistical technique that fits a smooth curve to each predictor and then combines these curves to create a composite curve. The curves demonstrate how $CH_4$ changes with each driver and provide reliable predictions beyond the training range (Hastie & Tibshirani 1990; Wood 2017). SVR is a ML algorithm that fits a flexible line or surface that best follows the data while allowing small errors within a defined range. It uses a mathematical function called a kernel to handle weak non-linear patterns, and is particularly effective when the dataset is small or the relationships are not strongly linear (Cortes & Vapnik 1995; Smola & Schölkopf 2004). Each model was implemented using the **caret** package in R (Kuhn 2008) for tuning and evaluation via stratified 10-fold cross-validation. We used the R-packages ***ranger*** for RF (Wright & Ziegler 2017), ***gbm*** for GBM (Greenwell et al. 2022), ***kernlab*** for SVR (Karatzoglou et al. 2004), and ***mgcv*** for GAM (Wood 2017). Root-mean-square error (RMSE) between measured and predicted $CH_4$ fluxes was the primary comparison metric because it penalizes large deviations more strongly than a mean-absolute error (Chai & Draxler 2014).
Models were also trained with additional predictors, including CALU land cover, InSAR-derived subsidence, and temporally dynamic NDVI and NDWI extracted close to the $CH_4$ flux measurement dates.
Third, two best-performing models (RF and GBM) were applied to a complete spatial predictor stack, a multi-layer raster covering the entire study area without gaps. The stack included two types of layers. Static layers, such as NDVI, NDWI, slope, aspect, TPI 30 m, TWI, and land cover, remained unchanged throughout July. In contrast, the meteorological layers (AT, TDD, PAR) were spatially uniform but temporally dynamic. A temporal loop progressed from 1 July at 00:00 to 31 July at 23:59 in three-hour steps. At each time step, the corresponding values of AT, PAR, and TDD were inserted into their respective layers in the stack. The model then generated an instantaneous $CH_4$ flux raster in mg $CH_4$ $m^{-2}$ $h^{-1}$ using the ***terra*** package (Hijmans 2023). This routine resulted in 248 flux rasters for the whole month of July, produced per year and spatial resolution. To ensure consistency, areas with missing input values (e.g., water bodies) were excluded from predictions. In total, 5,952 $CH_4$ flux rasters were generated (248 time steps × 2 models × 2 resolutions × 6 years).
Fourth, the predicted raster time series was aggregated using arithmetic operations in the ***terra*** package. Averaging over all time steps resulted in July mean flux maps, while summing and multiplying by three converted instantaneous rates into cumulative monthly fluxes. Six-year means and interannual variation (2019 to 2024) were calculated. Spatial mismatches due to scale effects were examined by differencing the 1 m and 10 m mean-flux maps cell by cell. Similarly, differences between model families

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


were mapped to capture structural uncertainty. To interpret these mismatches, we calculated Pearson correlations between flux difference maps (resolution-based or model-based) and individual predictor layers.

Fifth, we conducted a separate variable importance analysis to identify the most influential predictors in each model. Variable importance scores were extracted from the tuned RF and GBM models using permutation importance (***ranger*** package; Wright & Ziegler, 2017) and relative influence (***gbm*** package; Greenwell et al., 2022), respectively. These scores were used to assess the consistency of predictor relevance across models and spatial resolutions.

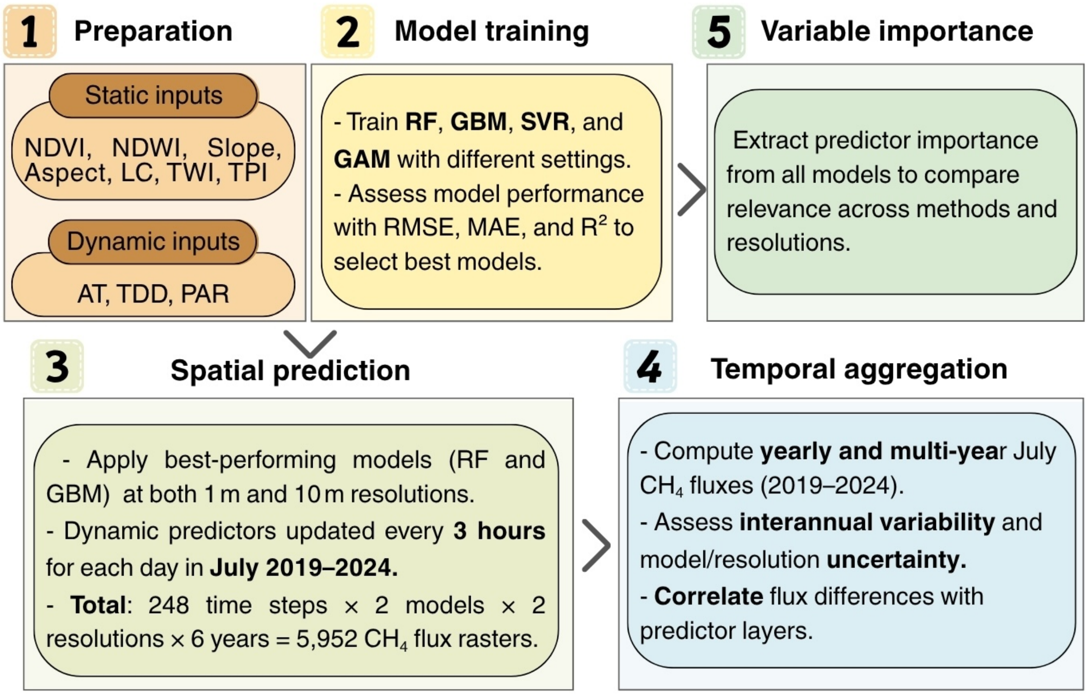


**Figure 3. Workflow for modelling and upscaling $CH_4$ fluxes in the designated study area. The analysis was performed separately for 1 m and 10 m spatial resolutions and comprised five primary stages: (1) Predictor preparation. (2) The training and tuning of models. (3) Spatial prediction. (4) Temporal aggregation and evaluation. (5) Variable importance.**

## 3 Results and Discussion

### 3.1 Correlation between observed $CH_4$ fluxes and single remote sensing parameters

This exploratory analysis examined how observed July $CH_4$ fluxes correlate with individual environmental variables at two spatial scales (1 m and 10 m) to identify significant controls of $CH_4$ flux and how spatial resolution affects their predictive power. Although the observed associations were generally weak, several clear patterns emerged across landscape classes and environmental gradients.

Seasonal subsidence showed the strongest positive correlation, underscoring the explanatory power of this parameter for moisture availability and related enhancements in $CH_4$ fluxes (Table 2). This is in line with observations linking InSAR-derived subsidence to elevated $CH_4$ fluxes in Arctic ecosystems (Sjögersten et al., 2023). Several moisture-related indices (NDWI, TWI, TPI) show higher correlations at 10 m than at 1 m, because 10 m aggregation smooths microtopographic noise while 1 m retains over-detailed, heterogeneous signals. This indicates that coarser resolution better captures landscape-scale hydrological gradients. This finding is supported by Ruhoff et al. (2011), who demonstrated that TWI values stabilize and become more spatially coherent at coarser resolutions, and by Riihimäki et al. (2021), who showed that TWI's ability to predict soil moisture

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


improves when derived from coarser DEMs (e.g., 10-30 m). Conversely, the correlation with aspect weakened at 10 m, compared to 1 m resolution, likely due to the loss of microtopographic detail when pixels are aggregated, as shown previously (Schoorl et al., 2000; Vaze et al., 2010).

Temporally matched NDVI and NDWI show weaker correlation with $CH_4$ fluxes compared to static indices. The reason may be the limited effective temporal resolution of Sentinel-2: although the constellation has a nominal 5-day revisit, persistent Arctic cloud cover often stretches the cloud-free gap well beyond 10 days (Runge & Grosse, 2019), producing a temporal mismatch with chamber measurements.

**Table 2. Spearman rank-correlation coefficients (ρ) between July $CH_4$ flux and environmental predictors at 1 m and 10 m spatial resolution. Positive values indicate that higher predictor values coincide with higher $CH_4$ emissions, negative values indicate the opposite. For vegetation and surface wetness predictors (NDVI, NDWI), both static (July 2018) and temporally matched values (Sentinel-2 scenes within ±10 days of each chamber measurement) are shown. The column "10 m, temporal" reflects those temporally matched predictors. For predictors derived from static landscape characteristics (e.g., TWI, Slope, TPI, Subsidence), 10 m and 10 m-temporal columns are merged as they do not vary in time. Significance levels: $^{**}p < 0.01$, $^{***}p < 0.001$.**

<table>
<tr><th>Group</th><th>Predictor</th><th>1 m</th><th>10 m</th><th>10 m, temporal</th></tr>
<tr><td>Vegetation</td><td>NDVI</td><td>-0.289***</td><td>-0.295***</td><td>-0.082**</td></tr>
<tr><td rowspan="2">Surface wetness and soil moisture</td><td>NDWI</td><td>0.141***</td><td>0.24***</td><td>-0.013</td></tr>
<tr><td>TWI</td><td>0.027**</td><td colspan="2">0.235***</td></tr>
<tr><td rowspan="3">Topography</td><td>Slope</td><td>-0.187***</td><td colspan="2">-0.238***</td></tr>
<tr><td>Aspect</td><td>0.14***</td><td colspan="2">0.035***</td></tr>
<tr><td>TPI</td><td>-0.162***</td><td colspan="2">-0.327***</td></tr>
<tr><td>Ground subsidence</td><td>Cumulative seasonal subsidence</td><td></td><td colspan="2">0.534***</td></tr>
</table>

We examined $CH_4$ flux variation across the landscape classes and CALU units (Fig. 4A and Table B1). For example, sedge-dominated landscape classes had the highest mean $CH_4$ flux (0.87 – 0.94 mg $CH_4$ $m^{-2}$ $h^{-1}$). Elevated fluxes in these systems are likely driven by plant transport through aerenchymatous tissue during which $CH_4$ produced at depth bypasses the oxic zones, and enhanced $CH_4$ production resulting from high plant productivity and increased substrate availability via root exudates (Olefeldt et al., 2013; Kwon et al., 2017). Tussock areas displayed the lowest flux values, with on average minor uptake of $CH_4$ (-0.02 mg $CH_4$ $m^{-2}$ $h^{-1}$). These patterns were consistent with observations by Voigt et al. (2023c).

All pairwise differences between $CH_4$ flux distributions for the 1 m and 10 m products were statistically significant (Wilcoxon rank-sum test, $p < 0.0001$). However, this result should be interpreted with caution due to the large sample sizes (even subtle differences can appear significant). This also applies to the correlations reported in Table 2. In some cases, the differences in median fluxes were small (e.g., sedges), while in others, the resolution shift results in more substantial changes (e.g., dwarf shrubs: median increased from 0.05 to 0.19 mg $CH_4$ $m^{-2}$ $h^{-1}$). In some cases, the flux sign even changed, for instance, lichen-dominated areas shift from weak uptake to weak emission. These shifts likely reflect the effects of aggregation, where coarser

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


resolution mixes surface types or blends microsites with different flux patterns.

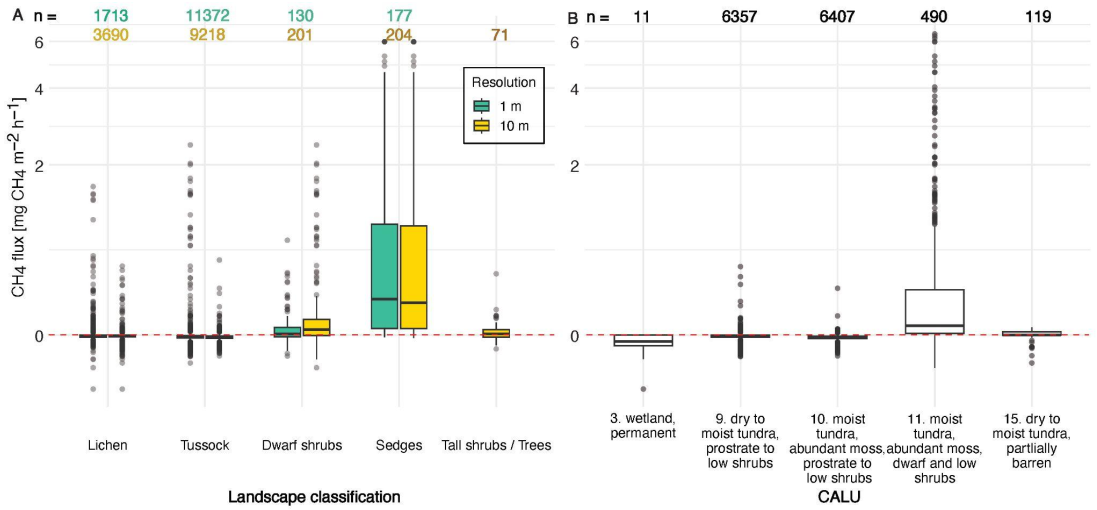


**Figure 4. Comparison of observed $CH_4$ fluxes across site-specific landscape classification at two spatial resolutions and CALU vegetation classes. Panel A: $CH_4$ fluxes across five site-specific landscape classes with existing $CH_4$ flux measurements. Measurements were aggregated separately for 1 m and 10 m spatial resolution. Panel B: $CH_4$ fluxes grouped by CALU (Circumarctic Land cover Units) classes. $CH_4$ fluxes differed significantly between most CALU classes ($p < 0.001$, pairwise Wilcoxon test), except classes 10 and 3 (wetlands), where no significant difference was observed ($p = 0.054$). Boxplots show the distribution of fluxes for each group. horizontal lines represent medians, boxes indicate the interquartile range, and whiskers extend to 1.5× the IQR. The red dashed lines indicate zero fluxes.**

To test whether a published pan-Arctic vegetation scheme would yield similar flux stratification, we overlaid our measurements on the CALU map (Fig. 4B). CALU vegetation classes differed significantly in $CH_4$ flux, except between moist moss tundra, abundant moss, prostrate to low shrubs (class 10) and permanent wetlands (class 3) (Fig. 4B, Table B2). Within CALU classes, average $CH_4$ fluxes ranged from slight uptake in wetland class (-0.09 mg $CH_4$ $m^{-2}$ $h^{-1}$) to moderate emissions in moist tundra, abundant moss, dwarf and low shrubs (CALU 11) (0.46 mg $CH_4$ $m^{-2}$ $h^{-1}$). Unexpectedly, the permanent wetland class showed the $CH_4$ uptake. This category only included one area, where dry lichen areas dominate most of the area. Moreover, the 10 m resolution of CALU likely leads to mixed pixels, where wetter spots were averaged with drier surroundings, reducing the apparent $CH_4$ emissions. In contrast, many wet areas at our site were too small to be resolved as wetlands in CALU and were instead classified into other categories.

Overlay analysis (Fig. 5) showed that each of our landscape classes included 6-11 CALU classes (with coverage > 1 %), typically dominated by moist tundra, abundant moss, dwarf and low shrubs (CALU 11). This reflects differences in classification approaches: CALU aimed at representing vegetation diversity and wetness gradients across the entire Arctic (Bartsch et al., 2024), whereas the site-specific landscape classification was explicitly built for $CH_4$ flux modelling and therefore integrates fine-scale microtopography, surface-moisture patterns, and local vegetation.

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


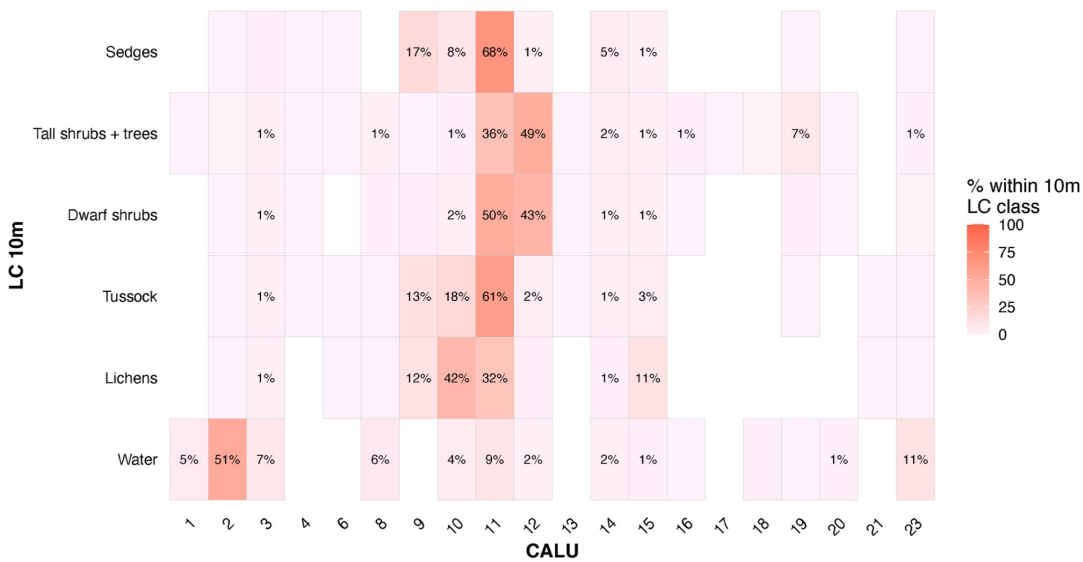


**Figure 5. Pixel-wise cross-comparison between two 10 m land-cover products for the TVC study area. LC 10 m (this study): a site-specific Sentinel-2 + ArcticDEM classification built (see SI Text 1). CALU (Circumarctic Land cover Units): published pan-Arctic landcover units (full legend in Table A3). Each tile shows the fraction of pixels of a given CALU class that fall into that LC 10 m class; row totals, therefore, equal 100 %. Values ≥ 0.5 % are printed inside the tiles. Tiles that are coloured but unlabelled occur (< 0.5 %), while blank tiles indicate class pairs that do not intersect within the AOI.**

However, even within each CALU or LC class, flux variance remained high, underlining that vegetation type alone cannot capture the full pattern of $CH_4$ fluxes without considering microtopography and moisture indices. Similar to the pan-Arctic synthesis by Olefeldt et al. (2013), our findings support the view that the effects of key environmental parameters on $CH_4$ flux should be considered jointly rather than independently.

### 3.2 Evaluation of Model Accuracy

Due to the relatively low mean $CH_4$ flux across all sites (0.102 mg $CH_4$ $m^{-2}$ $h^{-1}$), the emphasis of our model evaluation was placed on absolute errors (MAE) rather than the fraction of explained variance (Table 2). Our cross-validated modelling framework achieved predictive performance ($R^2$ from 0.53 to 0.87, Table 2) comparable to recent $CH_4$ upscaling studies in the Arctic-boreal region, including both chamber- (e.g., Virkkala et al., 2023; Räsänen et al., 2021) and eddy covariance-based studies (e.g., McNicol et al., 2023; Chen et al., 2024; Peltola et al., 2019; Tramontana et al., 2016).

Model evaluation at 1 m resolution revealed that SVR achieved the highest $R^2$ of 0.87, indicating strong predictive power. However, this was accompanied by substantial errors (RMSE = 115 %, MAE = 21.5 % of mean $CH_4$ flux), suggesting high sensitivity to skewed distributions and outliers, a known limitation of SVR when modelling non-Gaussian ecological data (Smola & Schölkopf, 2004). In contrast, RF showed both high accuracy and robustness, combining high $R^2$ with the lowest errors among tested algorithms. This confirms the algorithm's strength in capturing nonlinear interactions while being less sensitive to noise and overfitting, as highlighted in ecological applications (Belgiu & Drăguţ, 2016; Räsänen et al., 2021; Cutler et al., 2007). GBM also showed strong performance, with low errors and consistent $R^2$ values, reflecting its capability to efficiently leverage key predictors (Kämäräinen et al., 2023; Natekin & Knoll, 2013). GAM, in contrast, had the weakest performance among all models at 1 m resolution, with the lowest $R^2$ (0.62), highest RMSE (0.077), and highest MAE (0.025). This likely reflects the model's limited ability to capture sharp spatial variability in $CH_4$ fluxes when localized structure is strong. GAMs rely on detecting smooth nonlinear effects, but when predictors become noisy or spatially complex, the fitted splines lack the detail needed for accurate prediction (Wood, 2017).

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


At 10 m resolution, RF not only achieved the lowest mean absolute and root-mean-square errors, but its $R^2$ and error metrics also changed the least when we varied resolution or added temporally dynamic predictors, indicating the most consistent performance in our experiments (Table 3).

GBM showed similarly low errors but a slightly lower $R^2$ (0.57). SVR achieved the highest $R^2$ (0.68), but this was offset by much higher prediction errors, indicating poor generalization despite high apparent fit. GAM performed worst, with the lowest $R^2$ (0.53) and the highest RMSE (0.13).

**Table 3. Performance of four models at 1 m and 10 m spatial resolutions. Metrics include $R^2$ (coefficient of determination), MAE (mean absolute error), and RMSE (root mean square error). Bold values represent the best score for each metric within each resolution. The "10 m" scenario includes models with temporally stable normalized difference vegetation index (NDVI) and normalized difference water index (NDWI), while "10 m_temporal" refers to models using temporally dynamic indices, matched to the closest available date of in-field $CH_4$ flux measurements.**

| Model Type | Resolution | $R^2$ | MAE | RMSE |
|---|---|---|---|---|
| GAM | 1 m | 0.616 | 0.025 | 0.077 |
| | 10 m | 0.527 | 0.027 | 0.126 |
| | 10 m_temporal | 0.645 | 0.022 | 0.084 |
| GBM | 1 m | 0.625 | 0.008 | 0.012 |
| | 10 m | 0.570 | 0.008 | 0.013 |
| | 10 m_temporal | 0.689 | 0.117 | 0.024 |
| RF | 1 m | 0.744 | 0.006 | 0.010 |
| | 10 m | 0.650 | 0.007 | 0.012 |
| | 10 m_temporal | 0.751 | 0.016 | 0.105 |
| SVR | 1 m | 0.868 | 0.019 | 0.078 |
| | 10 m | 0.682 | 0.022 | 0.117 |
| | 10 m_temporal | 0.668 | 0.022 | 0.124 |

The observed drop in SVR and GAM performance at 10 m resolution reflects the Modifiable Areal Unit Problem (MAUP; Dark & Bram, 2007). This MAUP effect increased within-pixel heterogeneity as resolution coarsens and smooths fine-scale ecological signals. SVR, which is designed to fit narrow margins around nonlinear data patterns, becomes unstable when the training data lacks localized structure (Smola & Schölkopf, 2004). Similarly, GAM relies on the ability to detect smooth nonlinear effects in the predictors. When predictor distributions become more centralized and less variable, as often observed under coarser spatial resolutions (Riihimäki et al., 2019), GAM loses explanatory power because the fitted splines have insufficient detail to represent the underlying relationships (Wood, 2017). These scale-induced effects lead to higher errors and reduced $R^2$, particularly for models that depend on localized relationships. RF and GBM remained more stable across resolutions. The RF algorithm is less affected by noisy or imbalanced training data and overfitting, due to its ensemble structure combining multiple decision trees trained on bootstrapped subsets of data and predictor variables (Belgiu and Drăguţ, 2016). Based on these results, we selected RF and GBM for further analysis as the most reliable combination of accuracy and cross-resolution stability.

Including the temporal variability of the NDVI and NDWI values led to an average increase in $R^2$ of approximately 0.11 for the GAM, GBM and RF models at a resolution of 10 m (Table 3). SVR was the exception, showing a slight decrease in $R^2$ with no reduction in errors. For the GAM model, this increase in explanatory power was accompanied by lower RMSE and MAE values, indicating more accurate and robust performance. In contrast, both RF and GBM showed a higher $R^2$, but also exhibited increased absolute errors, which may indicate overfitting to temporally dynamic predictors. This likely reflects the tendency of ensemble models to capture noise in dynamic inputs when training data is limited (Barry and Elith, 2006; Chollet Ramampiandra et al., 2023; Reichstein et al., 2019). Similar behaviour has been observed in other ecosystem carbon flux modelling studies, for

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


example, in neural network models that overfit to lagged meteorological inputs (Papale & Valentini, 2003). The GAM model likely benefited from its ability to represent gradual ecological shifts through penalised smoothers, which reduces sensitivity to noise (Berbesi & Pritchard, 2023). The limited improvement in performance for SVR may be due to its sensitivity to data structure and lower flexibility when modelling smooth temporal trends in ecological datasets (Smola & Schölkopf, 2004).

### 3.3 Impact of model and resolution selection on $CH_4$ flux predictions

Different ML models can produce distinct spatial predictions even when trained on the same input data. Although well-documented, most ML models are not easily interpretable. We therefore compare their spatial predictions and simple diagnostics to assess reliability and guide model choice for $CH_4$ upscaling. Our comparison of upscaled $CH_4$ flux fields produced by the RF and GBM models showed that algorithm choice remained an important influence on spatial variability in predicted $CH_4$ fluxes (Fig. 6). The GBM model generated higher local contrast and more pronounced extremes, especially at 1 m resolution, with pronounced peaks in wet, topographically complex areas, reflecting its greater sensitivity to extreme values and local predictor variation. RF produced smoother, noise-resistant distributions, aligning with its known strength in generalizing across heterogeneous landscapes (Räsänen et al., 2021; Cutler et al., 2007). Although RF performs robustly and is widely used (Cutler et al., 2007), our results show that different algorithms can yield substantially different spatial patterns. This highlights the value of including multiple model types, not only for optimizing performance, but also for quantifying model-driven uncertainty in $CH_4$ flux upscaling.

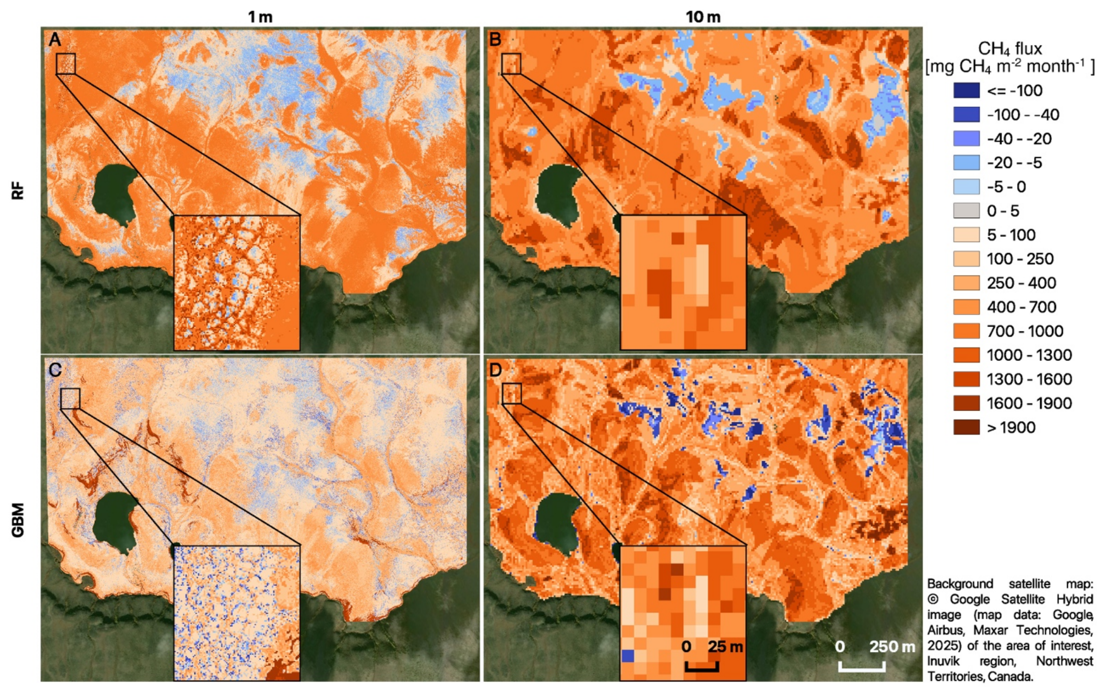


**Figure 6. Predicted mean monthly $CH_4$ fluxes (mg $CH_4$ $m^{-2}$ $month^{-1}$) for July (averaged over 2019–2024), generated by two machine-learning models: Random Forest (RF, panels A and B) and Gradient Boosting Machine (GBM, panels C and D). The panels A and C shows predictions at 1 m resolution, and panels B and D right column at 10 m resolution. Each panel contains a black-framed zoom window, which enlarges a representative section of the polygonal mire. Visual comparison of the two insets illustrates how the fine wet-to-dry microtopography resolved at 1 m is smoothed when aggregated to 10 m. Background imagery: © Google Satellite Hybrid (Maxar, 2025).**

Interestingly, although GBM exhibited more spatial flux variability, the mean fluxes predicted by GBM were consistently lower than those of RF. At 1 m, GBM averages 98.7 mg $CH_4$ $m^{-2}$ $month^{-1}$, whereas RF averages 518.6 mg; at 10 m the values rise to

608.8 mg and 683.4 mg, respectively (Fig. 7A). Net-sink pixels accounted for 10.0 % (RF) and 9.5 % (GBM) of the 1 m domain, but only 4.9 % (RF) and 4.4 % (GBM) at 10 m. $CH_4$ sink areas were spatially limited and highly sensitive to scale. Pixels acting as net $CH_4$ sinks (i.e. with negative monthly fluxes) were located on well-drained polygon rims and other lichen-dominated uplands where oxygen remained available throughout the summer. This allowed highly efficient methanotrophs to oxidise $CH_4$ faster than it was produced (Biasi et al., 2008). Resolving these units at a scale of 1 m showed that they covered around 10% of the scene and significantly reduced the landscape-mean flux. However, coarsening to 10 m mixed the aerobic patches with adjacent wet hollows, reducing their mapped extent to approximately 4.5% and erasing many uptake pixels. A comparable effect has been observed when chamber data were averaged across broader physiographic units, shifting site-level balances from weak sinks to slight sources (Zona et al., 2016). This pronounced scale effect is consistent with pan-Arctic syntheses, indicating that, although they cover only a small fraction of the surface, aerated uplands can offset a significant proportion of wetland emissions, yet they are often obscured in coarse products and regional budgets (Olefeldt et al., 2013; Kuhn et al., 2021). Our findings support recent assessments that retaining metre-scale information on microtopography, vegetation, and soil moisture is essential for capturing sink behaviour and ultimately for refining carbon budgets in permafrost regions, which currently indicate a small terrestrial $CO_2$ sink and a wetland $CH_4$ source (Treat et al., 2024).

Minimum and maximum fluxes remain similar between models at each resolution, indicating disagreement mainly in intermediate values. The residual model disagreement is driven less by the number of sink pixels than by their intensity. Minimum fluxes predicted by GBM were consistently more negative than those from RF, with extremes of -147 mg $CH_4$ $m^{-2}$ $month^{-1}$ (1 m) and -330 mg $CH_4$ $m^{-2}$ $month^{-1}$ (10 m), compared to -45 and -33 mg $CH_4$ $m^{-2}$ $month^{-1}$ in RF, respectively. This suggests that GBM may emphasize $CH_4$ sink strength more than RF, even though the spatial extent of sinks is similar across models.

At 1 m, GBM often responds more strongly to localized environmental extremes. These include areas with much higher soil moisture, surface temperature spikes, or abrupt changed in microtopography that may only occur at the meter scale. This is due to its sequential learning process, which can emphasize subtle but high-impact predictors. RF, in contrast, smooths local extremes and yields more conservative area means. Because GBM-1 m produced a markedly lower AOI mean than RF, we treat this behaviour as a potential systematic bias toward stronger sinks and hotspots. We therefore use RF-1 m as the reference budget estimate and retain GBM-1 m as a sensitivity case to bracket structural uncertainty. At 10 m, aggregation reduces fine-scale contrasts and the RF-GBM predictions converge. Pixel-wise standard deviations (Fig. 7B) reveal that RF is temporally more stable, while GBM is more sensitive to inter-annual variation, particularly in wet or geomorphically complex areas.

Additional analysis of spatial differences between models (Fig. B1) showed that several predictors were moderately to strongly correlated with the differences between RF and GBM predictions. At 1 m resolution, the strongest correlation was observed for NDWI (-0.53), indicating that model disagreement was most pronounced in wetter areas. NDVI (0.49) and land-cover type (0.41) also showed strong positive correlations with model differences, suggesting greater divergence in vegetated zones and across cover transitions. For the 10 m products, aspect (0.43) became the only predictor for model differences above 0.4, implying that model choice matters most on directionally exposed terrain once fine micro-relief is lost. Across both resolutions, NDWI exhibited consistent negative correlations, implying that divergences are magnified in wetter and concave landforms that tend to accumulate water or thaw differently. These findings are in line with Tagesson et al. (2013), who showed that adding satellite-derived NDWI improves $CH_4$ flux modelling by capturing moisture-driven variability.

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


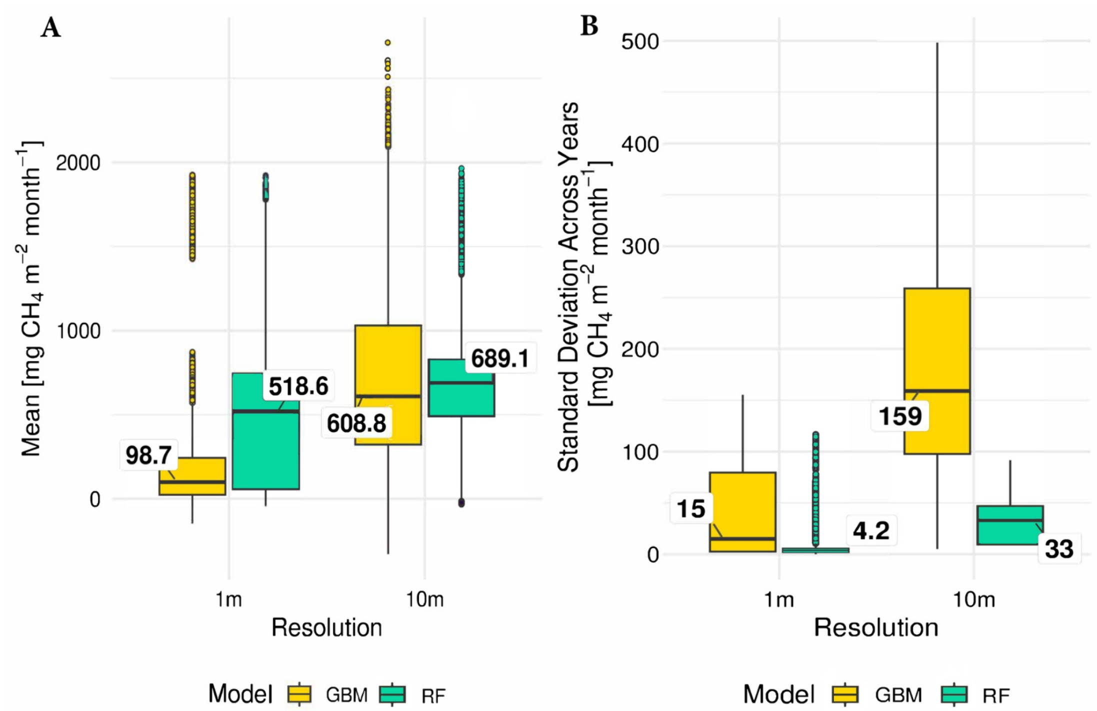


**Figure 7. Mean $CH_4$ flux and interannual variability across the study area. (A) Mean monthly $CH_4$ flux predicted by RF and GBM models, averaged over the entire area of interest for the period July 2019 to 2024. (B) Pixel-wise interannual standard deviation of predicted $CH_4$ fluxes for July months from 2019 to 2024, calculated separately for each model and resolution. Each boxplot shows the distribution of values across all pixels: the box spans the interquartile range (IQR, 25th to 75th percentile), the horizontal line within the box indicates the median, and whiskers extend to 1.5×IQR. Points beyond the whiskers represent potential outliers.**

In addition, the observed difference in mean fluxes may be partly due to the limited representation of extremely wet or topographically complex areas in the training dataset, reducing the generalization of the models under these conditions.

### 3.4 Parameters importance in $CH_4$ flux prediction

Analysis of the relative importance of the predictors revealed fundamental differences between the RF and GBM models, and how these differences change when moving from 1 m to 10 m resolution (Table B3). Significance was assessed using the permutation method for each model and scale combination.

At the 1 m resolution, RF distributed importance fairly evenly across the topographic parameters. TPI (~22 %), Aspect (~21 %) and Slope (~18 %) show comparable high influence, followed by Landscape class (~16 %). All other predictors contributed less than 10 %, and meteorological drivers collectively stay under that level. This broad topography-centred profile is consistent with the tendency of random forests to spread importance across correlated terrain drivers because of their random feature-selection mechanism (Räsänen et al., 2021; Cutler et al., 2007).

GBM showed a different pattern: again, no single parameter dominates, but five drivers spread across different input categories (Slope, Landscape class, AT, TDD and NDWI) each explained about 14-16 % of the total, and none exceeds 20 %. This flatter profile is based on the boosting process. Each new tree fixes the errors left by the previous one, so different predictors take turns improving the model (Friedman, 2001). When several drivers reduce error by a similar amount, the model splits importance among them (Kämäräinen et al., 2023).

When the resolution was coarsened to 10 m, pixel aggregation smoothed micro-relief, and both algorithms shifted toward moisture-integrating drivers as primary explanatory influences. In RF, NDWI (~25 %) and Landscape class (~25 %) emerge as joint leaders, NDVI rises to ~12 %, and all topographic parameters drop below 8 %. In GBM, the re-organisation is even stronger:

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


the moisture indicators NDWI (~25 %) and TWI (~19 %) together explained almost half of the total importance, while landscape class follows at ~11 % and Slope and Aspect fall below 7 %. This pattern agrees with field evidence that moisture proxies dominate $CH_4$-flux prediction at coarser resolution, where fine-scale topographic details are lost (Tagesson et al., 2013; Wangari et al., 2023). NDWI and TWI both integrate water content over several pixels, making them potential surrogates for local water-table height and the extent of anoxic microsites that drive methanogenesis. NDWI is also sensitive to vegetation water and phenology, allowing it to track water-table depth in peatlands (Kalacska et al., 2018). TWI, which maps landscape-scale water accumulation and thus redox and gas-diffusion controls, aligns with syntheses showing that water-table fluctuations set the size of anoxic zones and largely govern $CH_4$ production and emission (Kaiser et al., 2018; Cui et al., 2024). Landscape class and NDVI contributed complementary information on vegetation type and biomass, which modulate both substrate supply and methane oxidation. In practical terms, upscaling to 10 m can still capture landscape-scale $CH_4$ patterns, but only if robust moisture indices such as NDWI and TWI were included; purely geometric terrain drivers lose most of their explanatory power once microtopography is averaged out.

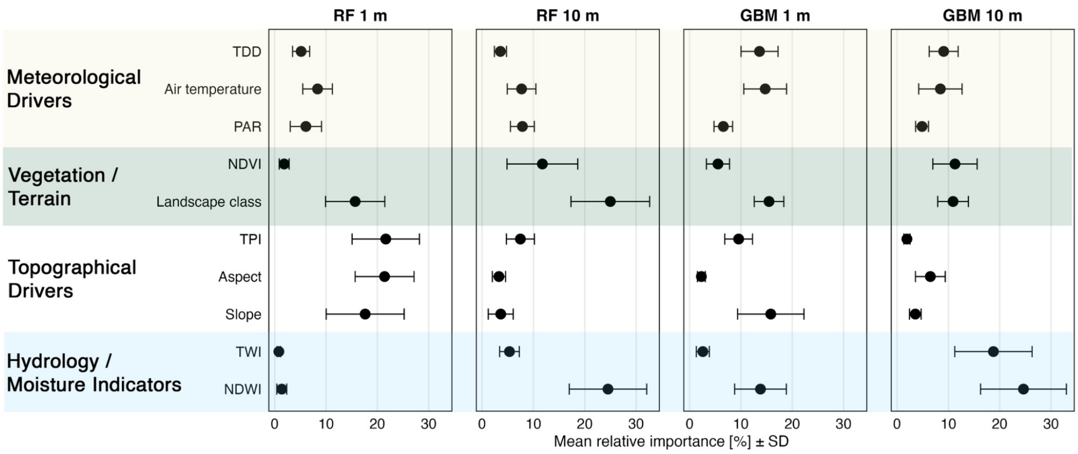


**Figure 8 Mean relative importance (± SD) of environmental predictors for $CH_4$ fluxes across two machine-learning models –Random Forest (RF) and Gradient Boosting Machine (GBM) – evaluated at 1 m and 10 m spatial resolutions. Importance was estimated by bootstrap resampling (n = 100) and is expressed as a percentage of total importance within each model. Predictors are grouped into four categories: Meteorological drivers (thawing degree days, air temperature, photosynthetically active radiation), Vegetation/Terrain (Normalized Difference Vegetation Index, landscape class), Topography (Topographic Position Index, aspect, slope), and Hydrology/Moisture (Topographic Wetness Index, Normalized Difference Water Index). Abbreviations: TDD – thawing degree days; PAR – photosynthetically active radiation; NDVI – Normalized Difference Vegetation Index; TPI – Topographic Position Index; TWI – Topographic Wetness Index; NDWI – Normalized Difference Water Index.**

### 3.5 Study limitations

The study's limitations include unbalanced sampling across landscape types and under-representation of extremely wet or microtopographically complex areas. In particular, several landscape and CALU classes remain undersampled, which limits the model's ability to predict fluxes across all types. Additionally, soil temperature and moisture, which are known to control $CH_4$ fluxes (Wille et al., 2008; Mastepanov et al., 2013), were not included, but are planned for integration in future model development.

## 4 Conclusion

Spatial resolution emerged as the important factor determining the predictive power data-driven upscaled $CH_4$ flux patterns, exerting a stronger influence than model choice. At a resolution of 1 m, fine-scale heterogeneity was captured at a high degree of detail, making it possible for models to distinguish between local sources and sinks of $CH_4$. At 10 m, micro features merge into mixed pixels, boosting mean fluxes and variability. This resulted in fine-scale sinks and hotspots disappearing, and in some

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


cases, fluxes being misclassified as a source of $CH_4$ in dry areas. Consequently, 10 m models produced higher mean fluxes and broader flux distributions. However, some of these high values may be due to mixed-pixel artefacts rather than true local emissions.

Although different models varied significantly in their estimates, RF and GBM provided the most consistent and reliable upscaling results. However, their robustness should be verified by targeted sensitivity analyses. Significance of model predictors was found to be strongly scale-dependent. At a resolution of 1 m, the models derived most of their explanatory power from microtopographic metrics, which capture the detailed elevation contrasts that distinguish between hummocks and hollows, as well as localising $CH_4$ hotspots. However, after aggregation to 10 m, these relief cues were diluted, causing a change in ranking: moisture proxies NDWI and TWI became the principal drivers, together accounting for almost half of the explained variance. This transition from terrain- to moisture-controlled importance highlights the fact that fine-scale mapping requires detailed topographic data, whereas regional upscaling must prioritize robust hydrological indices. For AOI budgets we report RF-1 m as the reference and use GBM-1 m as a sensitivity bound due to its amplification of metre-scale extremes.

Our study findings imply that resolution is not simply a case of 'the higher, the better', and similarly, more complex ML methods may not necessarily yield better predictions. Although 1 m models captured fine-scale heterogeneity, 10 m models with temporally dynamic predictors improve explanatory power but increase prediction errors, likely due to overfitting to short-term fluctuations. This suggests that, in some cases, high-resolution models can outperform ultra-high resolution ones, particularly when enhanced with well-timed spectral information – though caution is needed to balance fit with generalisability.

Subsidence, derived from InSAR, showed the highest correlation with observed $CH_4$ fluxes of all the tested predictors, emphasising its value as a spatial proxy for soil moisture. It should therefore be included directly in $CH_4$ upscaling workflows, particularly in permafrost landscapes where moisture conditions were key drivers of emissions.

Future work should expand sampling into underrepresented landscape and vegetation classes, high-emission zones, methane uptake regions, and winter fluxes, and incorporate temporally dynamic predictors. Integrating theory-guided time-series modelling approaches informed by ecological theory could enhance both the interpretability and accuracy of $CH_4$ forecasts under complex seasonal dynamics, particularly when data availability is limited.

## Appendix A. Predictors from remote sensing and meteorological data

**Table A1. Overview of predictor variables used in the $CH_4$ flux models. This table lists all environmental predictor variables considered in the modelling framework. For each parameter, the spatial resolution (for remote sensing layers), source, short description, and formulas for calculations are presented (where applicable). Parameters are grouped into six thematic categories: Meteorological Drivers (e.g., PAR, AT, TDD), Vegetation / Land Cover (e.g., NDVI, landscape classification, CALU), Hydrology / Moisture Indicators (e.g., NDWI, TWI), Topography (e.g., slope, aspect, TPI), and Surface Deformation (subsidence). Each variable is marked as either static (unchanging during the study period) or dynamic (time-specific).**

| Parameter | Resolution | Derived from | Description | Temporal variability | Parameter type |
|---|---|---|---|---|---|
| Photosynthetically Active Radiation (PAR) | 1 km | NASA Langley Research Center (2024) | Extracted as a predictor variable for $CH_4$ flux models. | Dynamic | Meteorological Drivers |
| NDVI | 1 m | Rettelbach et al. (2024) | Ultra-high resolution NDVI derived from drone imagery. $NDVI = \frac{NIR - Red}{NIR + Red}$ (A1) | Static | Vegetation /Terrain |
| NDWI | 1 m | Rettelbach et al. (2024) | Ultra-high resolution NDWI derived from drone imagery. $NDWI = \frac{Green - NIR}{Green + NIR}$ (A2) | Static | Hydrology / Moisture Indicators |

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


| **Parameter** | **Resolution** | **Derived from** | **Description** | **Temporal variability** | **Parameter type** |
|---|---|---|---|---|---|
| Landscape classification | 1 m | Rettelbach et al. (2024), Lange et al., 2021 | Landscape classification performed using 1 m drone imagery & ALS-derived DTM (Appendix B). | Static | Vegetation /Terrain |
| NDVI | 10 m | Sentinel-2 [2019 - 2024] (mean for July - August 2018). | Extracted from the composite Sentinel-2 image for July - August 2018. $NDVI = \frac{NIR - Red}{NIR + Red}$ (A1) | Static | Vegetation /Terrain |
| NDWI | 10 m | Sentinel-2 [2019 - 2024] (mean for July - August 2018). | Extracted from the composite Sentinel-2 image for July - August 2018. $NDWI = \frac{Green - NIR}{Green + NIR}$ (A2) | Static | Hydrology / Moisture Indicators |
| NDVI | 10 m | Sentinel-2 [2019 - 2024] (Single-date, closest to flux measurement). | Extracted from single-date, closest to flux measurement. $NDVI = \frac{NIR - Red}{NIR + Red}$ (A1) | Dynamic | Vegetation /Terrain |
| NDWI | 10 m | Sentinel-2 [2019 - 2024] (Single-date, closest to flux measurement). | Extracted from single-date, closest to flux measurement. $NDWI = \frac{Green - NIR}{Green + NIR}$ (A2) | Dynamic | Hydrology / Moisture Indicators |
| Landscape classification | 10 m | Copernicus Sentinel-2 data [2018], ArcticDEM v4 (Porter et al., 2023) | Landscape classification performed using Sentinel-2 indices (2018) and terrain derivatives of ArcticDEM (Appendix B). | Static | Vegetation /Terrain |
| Slope | 1 m | Lange et al., 2021 | Measures the rate of change of elevation in the direction of steepest descent. It controls how gravity drives the movement of water and materials across the surface and through the soil. Slope affects surface and subsurface water flow, soil water content, erosion potential, and soil formation. It is a key factor in hydrological and geomorphological processes.<br><br>Derived from DTM (Wilson & Gallant, 2000). | Static | Topographical parameters |
| Aspect | 1 m | Lange et al., 2021 | Aspect is the compass direction that a slope faces. It is defined as the orientation of the line of steepest descent and is typically measured in degrees clockwise from north. Aspect affects local microclimate, radiation exposure, snowmelt timing, and vegetation structure, making it a common site descriptor in ecological and geomorphological studies.<br>Derived from DTM (Wilson & Gallant, 2000). | Static | Topographical parameters |

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


| Parameter | Resolution | Derived from | Description | Temporal variability | Parameter type |
|---|---|---|---|---|---|
| TWI | 1 m | Lange et al., 2021 | TWI combines upslope catchment area and slope to model potential soil moisture accumulation. It is commonly used to identify areas potentially prone to saturation and water accumulation.<br><br>$TWI = ln\frac{a}{tan\,b}$,<br>- a = *upslope contributing area per unit contour length*<br>- b = *local slope angle*<br><br>Derived from DTM (Wilson & Gallant, 2000). | Static | Topographical parameters |
| TPI_30m | 1 m | Lange et al., 2021 | The Topographic Position Index (TPI) quantifies the elevation of a cell relative to the mean elevation of surrounding cells, allowing differentiation between ridges, valleys, and flat areas. We computed TPI using a 30 m circular moving window, meaning that for each location, its elevation was compared to the average of all surrounding elevations within a 30 m radius. This window size smooths out small-scale variation and captures broader landform patterns.Derived from DTM (Weiss, 2001). | Static | Topographical parameters |
| Slope | 10 m (from 2 m) | ArcticDEM v4 (Porter et al., 2023) | Measures the rate of change of elevation in the direction of steepest descent. It controls how gravity drives the movement of water and materials across the surface and through the soil. Slope affects surface and subsurface water flow, soil water content, erosion potential, and soil formation. It is a key factor in hydrological and geomorphological processes.<br><br>Derived from DTM (Wilson & Gallant, 2000). | Static | Topographical parameters |
| Aspect | 10 m (from 2 m) | ArcticDEM v4 (Porter et al., 2023) | Aspect is the compass direction that a slope faces. It is defined as the orientation of the line of steepest descent and is typically measured in degrees clockwise from north. Aspect affects local microclimate, radiation exposure, snowmelt timing, and vegetation structure, making it a common site descriptor in ecological and geomorphological studies.<br>Derived from DTM (Wilson & Gallant, 2000). | Static | Topographical parameters |
| TWI | 10 m (from 2 m) | ArcticDEM v4 (Porter et al., 2023) | TWI combines upslope catchment area and slope to model potential soil moisture accumulation. It is commonly | Static | Topographical parameters |

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


| Parameter | Resolution | Derived from | Description | Temporal variability | Parameter type |
|---|---|---|---|---|---|
| | | | used to identify areas potentially prone to saturation and water accumulation.<br><br>$TWI = ln\frac{a}{tan\,b}$, (A3)<br>- a = *upslope contributing area per unit contour length*<br>- b = *local slope angle*<br><br>Derived from DTM (Wilson & Gallant, 2000). | | |
| TPI_30m | 10 m (from 2 m) | ArcticDEM v4 (Porter et al., 2023) | The Topographic Position Index (TPI) quantifies the elevation of a cell relative to the mean elevation of surrounding cells, allowing differentiation between ridges, valleys, and flat areas. We computed TPI using a 30 m circular moving window, meaning that for each location, its elevation was compared to the average of all surrounding elevations within a 30 m radius. This window size smooths out small-scale variation and captures broader landform patterns. Derived from DTM (Weiss, 2001). | Static | Topographical parameters |
| Subsidence | 10 m | Copernicus Sentinel-1/2 data | Seasonal deformation has been derived from Sentinel-1 time series (2018 - 2023) using SAR Interferometry. Six years have been averaged to reduce noise. The seasonal deformation rates in thawing degree days domain represent near surface soil moisture spatial patterns. (Widhalm et al., 2025) | Static | Surface Deformation |
| CALU | 10 m | CALU<br><br>(Bartsch et al., 2024) | The Circumarctic Landcover Units provide a consistent high-resolution land cover classification across the entire Arctic tundra. CALU defines **23 units of similar reflectance derived from multispectral (Sentinel-2) and C-band SAR (Sentinel-1) data**. The classification reflects wetness gradients, shrub density, moss abundance, and surface moisture (Bartsch et al., 2024). | Static | Vegetation /Terrain |
| AT | Point | Trail Valley Creek meteorological station (Climate ID: 220N005; WMO ID: 71683; TC ID: XTV). | Hourly air temperature measured at 2 m above ground level. Used as a dynamic meteorological driver for $CH_4$ flux models. | Dynamic | Meteorological Drivers |

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


| **Parameter** | **Resolution** | **Derived from** | **Description** | **Temporal variability** | **Parameter type** |
|---|---|---|---|---|---|
| Thawing Degree Days (TDD) | Point | Trail Valley Creek meteorological station (Climate ID: 220N005; WMO ID: 71683; TC ID: XTV). | Cumulative positive air temperature sum (above 0 °C) used as a proxy for thaw energy and season length. Calculated per flux measurement period based on air temperature from meteorological station.<br><br>$TDD = \sum_{i=1}^{n} max(T_{mean,i}, 0)$ , (A4)<br><br>• $T_{mean,i}$ = mean daily air temperature on day *i*<br>• n = number of days in the accumulation period<br>• The **max** function ensures only temperatures above 0 °C are counted | Dynamic | Meteorological Drivers |

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


**Text A1. Landscape classification**

To classify land cover in the TVC area, we employed a supervised classification approach using multi-source remote sensing data at 1 m and 10 m resolutions. The classification process was implemented in Google Earth Engine (GEE), enabling large-scale data processing. A Random Forest (RF) classifier was chosen due to its ability to handle high-dimensional data, its resistance to overfitting, and its suitability for land cover mapping. By applying a consistent classification framework at both 1 m and 10 m resolutions, this study enables direct comparisons of classification performance across spatial scales,

Training and Validation Data

The classification was trained using manually collected validation points that were assigned to six distinct land cover classes: Dwarf Shrub, Tall Shrub, Sedges, Tussock, Lichen, and Water. To ensure statistical robustness, 80 % of the validation points were used for model training, while the remaining 20 % were reserved for accuracy assessment.

Remote Sensing Data and Feature Extraction

To optimize classification accuracy, we integrated spectral, texture, and topographic features derived from multiple remote sensing sources. Sentinel-2 optical imagery at 10 m resolution was used for broad-scale classification, with images acquired during the 2018 growing season (25 June–4 September 2018) to ensure that differences in land cover classification were due to spatial resolution rather than changing environmental conditions, matching the same summer period as the 1 m drone survey. Topographic features were extracted from ArcticDEM (2-m resolution) (Porter et al., 2023). At finer spatial scales, we incorporated ultra-high resolution drone imagery (1 m and 10 cm) from Rettelbach et al. (2024) and a digital terrain model (DTM) (Lange et al., 2021).

To further enhance classification accuracy, we performed a Gray-Level Co-occurrence Matrix (GLCM) texture analysis of NDVI, allowing us to incorporate information on vegetation heterogeneity. A 2 × 2 kernel was used for 10 m classification, while a 20 × 20 kernel was applied at 1 m resolution to capture 20 m spatial patterns.

**Table A2. Parameters used for the landscape classification**

| Parameter | Description | Formula (if applicable) | Resolution |
|---|---|---|---|
| NDVI | Measures vegetation greenness | $\frac{NIR - RED}{NIR + RED}$ | 10 m, 1 m |
| NDWI | Identifies water and moisture content | $\frac{Green - NIR}{Green + NIR}$ | 10 m, 1 m |
| EVI | Improves sensitivity to high biomass | $2.5 \times \frac{NIR - RED}{NIR + 6 \times RED - 7.5 \times Blue + 1}$ | 10 m, 1 m |
| SAVI | Reduces soil brightness effects | $\frac{(NIR - RED) \times (1 + L)}{NIR + RED + L}$, where L = 0.5 | 10 m, 1 m |
| GLCM Entropy | Measures randomness in pixel intensity | Derived from NDVI | 10 m, 1 m |
| GLCM Contrast | Captures local texture variation | Derived from NDVI | 10 m, 1 m |
| GLCM Homogeneity | Measures uniformity in image texture | Derived from NDVI | 10 m, 1 m |
| Slope | Measures terrain steepness | Derived from DEM | 2 m, 1 m |
| Aspect | Identifies terrain orientation | Derived from DEM | 2 m, 1 m |
| TPI 6m | Detects local terrain position | Elevation - Mean(Elevation within 6m radius) | 2 m, 1 m |
| TPI 30m | Identifies broader-scale landforms | Elevation - Mean(Elevation within 30m radius) | 2 m, 1 m |
| TWI | Estimates soil moisture potential | $ln(\frac{A}{tan\ (\beta)})$, where | 2 m, 1 m |

| Parameter | Description | Formula (if applicable) | Resolution |
|---|---|---|---|
| | | A = specific contributing area<br>$\beta$ = slope in radians | |
| Band parameters | Captures spectral variation in different wavelengths | mean and sd for each pixel of RGB and NIR bands | 10 m |
| Band parameters | Captures spectral variation in different wavelengths | pixel value of RGB and NIR bands | 10 m |

Classification Model and Accuracy Assessment

The Random Forest classifier was trained separately for 10 m Sentinel-2 data and 1 m drone-based data, with 200 decision trees used in both cases. The trained models were then applied to classify the entire dataset. The overall accuracy was 0.76 for 1 m resolution and 0.71 for 10 m resolution.

Export

Final classified maps at 10 m and 1 m resolutions were exported as GeoTIFF files for further analysis and comparison.

**Table A3. Description of Circumarctic Land Cover Units (CALU) present in the study area. Class names and definitions are taken from Bartsch et al. (2024). Additional columns indicate (i) whether the class is present within the area of interest (AOI), and (ii) whether $CH_4$ flux measurements are available for this class.**

| CALU class | Description | Present in AOI | $CH_4$ measurements available |
|---|---|---|---|
| 1 | Water | yes | |
| 2 | shallow water/abundant macrophytes | yes | |
| 3 | wetland, permanent | yes | yes |
| 4 | wet to aquatic tundra (seasonal), abundant moss | yes | |
| 5 | moist to wet tundra, abundant moss, prostrate shrubs | | |
| 6 | dry to moist tundra, partially barren, prostrate shrubs | yes | |
| 7 | dry tundra, abundant lichen, prostrate shrubs | | |
| 8 | dry to aquatic tundra, dwarf shrubs (& sparse tree cover along treeline) | yes | |
| 9 | dry to moist tundra, prostrate to low shrubs | yes | yes |
| 10 | moist tundra, abundant moss, prostrate to low shrubs | yes | yes |
| 11 | moist tundra, abundant moss, dwarf and low shrubs | yes | yes |
| 12 | moist tundra, dense dwarf and low shrubs (& sparse tree cover along treeline) | yes | |
| 13 | moist to wet tundra, dense dwarf and low shrubs (& sparse tree cover along treeline) | yes | |
| 14 | moist tundra, low shrubs | yes | |
| 15 | dry to moist tundra, partially barren | yes | yes |
| 16 | moist tundra, abundant forbs, dwarf to tall shrubs | yes | |
| 17 | recently burned or flooded, partially barren | yes | |
| 18 | forest (deciduous) with dwarf to tall shrubs | yes | |
| 19 | forest (mixed) with dwarf to tall shrubs | yes | |
| 20 | forest (needle leave) with dwarf and low shrubs | yes | |
| 21 | partially barren | yes | |
| 22 | snow/ice | | |
| 23 | other (incl. shadow) | yes | |

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


## Appendix B. Results

**Table B1. Summary statistics of observed $CH_4$ fluxes (mg $CH_4$ $m^{-2}$ $h^{-1}$) across site-specific landscape classes at 1 m and 10 m spatial resolutions. The table reports the number of observations (n), mean, first quartile (Q1), third quartile (Q3), minimum, and maximum $CH_4$ flux values for each landscape class at both resolutions.**

| Landscape class | Resolution | n | Mean | Q1 | Q3 | Min | Max |
|---|---|---|---|---|---|---|---|
| Lichen | 1 m | 1713 | 0.002 | -0.02 | 0 | -0.48 | 1.62 |
| | 10 m | 3690 | -0.011 | -0.02 | 0 | -0.48 | 0.62 |
| Tussock | 1 m | 11372 | -0.016 | -0.03 | -0.01 | -0.24 | 2.41 |
| | 10 m | 9218 | -0.020 | -0.03 | -0.01 | -0.18 | 0.68 |
| Dwarf shrub | 1 m | 130 | 0.053 | -0.02 | 0.06 | -0.18 | 0.89 |
| | 10 m | 201 | 0.19 | -0.01 | 0.13 | -0.28 | 2.41 |
| Tall shrub | 10 m | 71 | 0.024 | -0.02 | 0.05 | -0.12 | 0.54 |
| Sedges | 1 m | 177 | 0.94 | 0.06 | 1.09 | -0.02 | 6.39 |
| | 10 m | 204 | 0.87 | 0.05 | 1.07 | -0.03 | 6.39 |

**Table B2. Summary statistics of observed $CH_4$ fluxes (mg $CH_4$ $m^{-2}$ $h^{-1}$) across CALU classes. The table reports the number of observations (n), mean, first quartile (Q1), third quartile (Q3), minimum, and maximum $CH_4$ flux values for each landscape class at both resolutions. Class descriptions are available in Bartsch et al. (2024).**

| CALU class | n | Mean | Q1 | Q3 | Min | Max |
|---|---|---|---|---|---|---|
| 3. Permanent wetland | 11 | -0.09 | -0.09 | 0 | -0.48 | 0 |
| 9. Dry to moist tundra, prostrate to low shrubs, tussocks | 6357 | -0.01 | -0.02 | 0 | -0.19 | 0.61 |
| 10. Moist tundra, abundant moss, prostrate to low shrubs, tussocks | 6407 | -0.02 | -0.03 | -0.01 | -0.18 | 0.41 |
| 11. Moist tundra, abundant moss, dwarf and low shrubs, tussocks | 490 | 0.46 | 0.01 | 0.39 | -0.28 | 6,39 |
| 15. Moist to wet tundra, abundant lichen, in some cases partially barren (disturbed). | 119 | 0 | 0 | 0.03 | -0.24 | 0.06 |

**Figure B1. Pearson correlation between spatial differences in $CH_4$ flux predictions and selected environmental predictors. The heatmap shows correlations between $CH_4$ prediction differences and predictor variables, separately for: Model influence (left block): differences between RF and GBM predictions at the same resolution (1 m or 10 m), Resolution influence based on 1 m predictors (middle block): differences between 1 m and 10 m model predictions using predictors extracted from 1 m products, Resolution influence based on 10 m predictors (right block): differences between 1 m and 10 m model predictions using 10 m products. Each cell represents Pearson's r between predictor values and flux differences across 30402 pixels for 10 m resolution and 3050788 pixels for 1 m resolution. Positive values indicate that higher predictor values are associated with greater $CH_4$ prediction mismatches.**

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


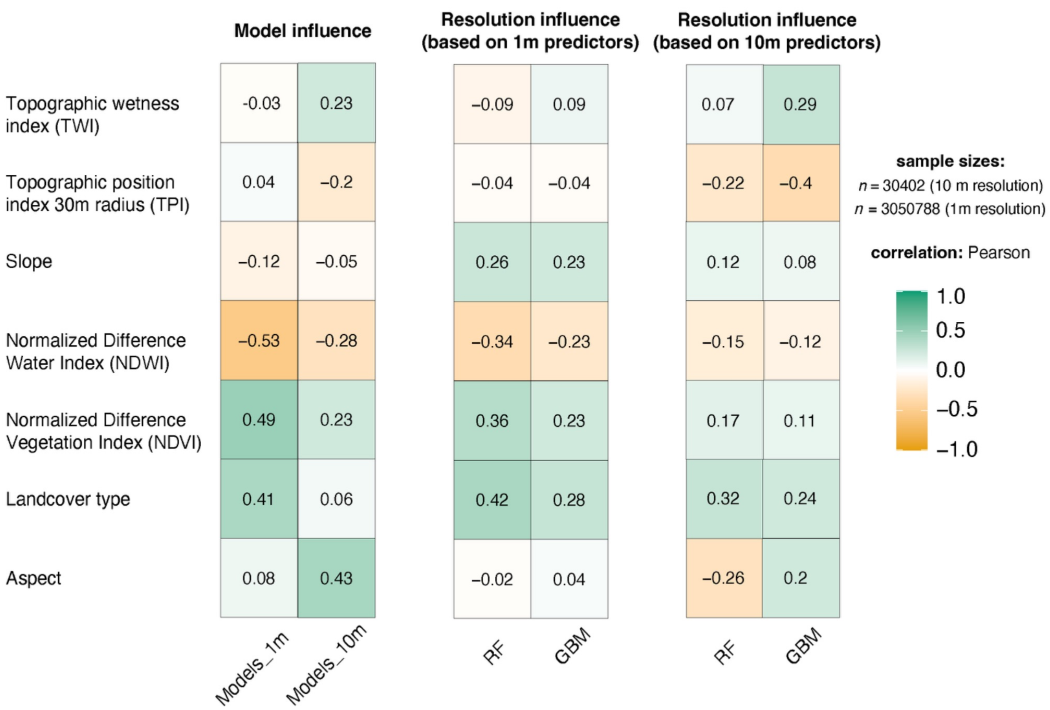


**Table B3. Relative importance [%] of environmental predictors for $CH_4$ flux models across spatial resolutions and algorithms. The table shows the variable importance (in %) for each predictor derived from Random Forest (RF) and Gradient Boosted Machine (GBM) models at 1 m and 10 m spatial resolution. Predictors are grouped by thematic category (e.g., Meteorological, Topographic). Importance values reflect the mean contribution of each predictor to the model performance and standard deviations (± SD).**

| Group | Parameter | RF 1 m | RF 10 m | GBM 1 m | GBM 10 m |
|---|---|---|---|---|---|
| **Meteorological Drivers** | Air temperature | 8.4 ± 2.9 | 7.7 ± 2.8 | 14.7 ± 4.2 | 8.4 ± 4.2 |
| | PAR | 6.1 ± 3 | 7.8 ± 2.3 | 6.6 ± 1.8 | 4.9 ± 1.3 |
| | TDD | 5.2 ± 1.7 | 3.6 ± 1.2 | 13.6 ± 3.6 | 9.1 ± 2.8 |
| **Hydrology / Moisture Indicators** | NDWI | 1.4 ± 1 | 24.5 ± 7.5 | 13.8 ± 5 | 24.6 ± 8.4 |
| | TWI | 0.8 ± 0.3 | 5.3 ± 1.9 | 2.6 ± 1.3 | 18.8 ± 7.5 |
| **Topographical parameters** | Aspect | 21.4 ± 5.7 | 3.3 ± 1.3 | 2.3 ± 0.8 | 6.5 ± 2.9 |
| | Slope | 17.6 ± 7.6 | 3.6 ± 2.4 | 15.8 ± 6.5 | 3.6 ± 1.1 |
| | TPI | 21.6 ± 6.6 | 7.5 ± 2.7 | 9.6 ± 2.7 | 1.9 ± 0.6 |
| **Vegetation / Terrain** | NDVI | 1.9 ± 1 | 11.7 ± 6.9 | 5.5 ± 2.3 | 11.3 ± 4.3 |
| | Landscape class | 15.7 ± 5.8 | 25 ± 7.7 | 15.5 ± 2.9 | 10.9 ± 3 |

**Author contributions**

KI, MG, AB, BW, VB, and TS received funding from the European Research Council Synergy Grant Q-Arctic (grant no. 951288). MG further acknowledges financial support from the EU HORIZON EUROPE programme (project GreenFeedBack,

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


grant agreement number 101056921), the European Space Agency (AMPAC-net project, contract number 4000137912/22/I-DT) and the German Ministry for Research, Technology and Space (MOMENT project, Grant No. 03F0931G). AV received support from the Gordon and Betty Moore Foundation (grant no. 8414), and funding catalyzed by the TED Audacious Project (Permafrost Pathways). CV received financial support from the Research Council of Finland project MUFFIN (no. 332196) and the European Research Council Starting Grant COLDSPOT (no. 101163177). OS acknowledges financial support through the Canada Research Chair (CRC-2018-00259) and NSERC Discovery Grants program (DGPIN-2018-05743 awarded, ArcticNet, a Network of Centres of Excellence Canada (grant no. P216), Canada First Research Excellence Fund's Global Water Futures program (Northern Water Futures), and the Polar Continental Shelf Program (608-20 and 602-21).

**Competing interests**

The authors declare that they have no conflict of interest.

**Acknowledgments**

We are grateful for the support of all research teams and the assistance provided at Trail Valley Creek Research Station in organizing the fieldwork. We are grateful to Dr. Theresia Yazbeck for insightful feedback that helped improve the manuscript. Parts of the text were language-edited with DeepL, and some analysis and plotting scripts were revised for clarity with assistance from ChatGPT (OpenAI; accessed January 2025). All outputs were reviewed and verified by the authors, who take full responsibility for the content.

We acknowledge the Trail Valley Creek orthophotos, digital surface models, and 3D point clouds provided by the MOSES airborne campaign and published by Rettelbach et al. (2024) on PANGAEA (DOI: 10.1594/PANGAEA.961942). ArcticDEM provided by the Polar Geospatial Center under NSF-OPP awards 1043681, 1559691, 1542736, 1810976, and 2129685. We acknowledge the use of Copernicus Sentinel-2 data (Contains modified Copernicus Sentinel data [2018 - 2024]). We acknowledge the use of NASA POWER data, provided by the NASA Langley Research Center (LaRC) through the POWER Project, funded by the NASA Earth Science/Applied Science Program.

**Financial support**

KI, MG, AB, BW, VB, and TS received funding from the European Research Council Synergy Grant Q-Arctic (grant no. 951288). AV received support from the Gordon and Betty Moore Foundation (grant no. 8414), and funding catalyzed by the TED Audacious Project (Permafrost Pathways). CV received financial support from the Research Council of Finland project MUFFIN (no. 332196) and the European Research Council Starting Grant COLDSPOT (no. 101163177). OS acknowledges financial support through the Canada Research Chair (CRC-2018-00259) and NSERC Discovery Grants program (DGPIN-2018-05743 awarded, ArcticNet, a Network of Centres of Excellence Canada (grant no. P216), Canada First Research Excellence Fund's Global Water Futures program (Northern Water Futures), and the Polar Continental Shelf Program (608-20 and 602-21).

**Code availability**

The code used for $CH_4$ flux modelling, resolution comparison, and upscaling across Arctic wetland landscapes is publicly available on Zenodo (Ivanova et al., 2025b): https://doi.org/10.5281/zenodo.15399084

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


**Data availability**

The $CH_4$ flux predictions, spatial difference maps, and input dataset used in this study are publicly available on Zenodo (Ivanova et al., 2025a): https://doi.org/10.5281/zenodo.15753253

**References**

Abdelmajeed, A. Y. A. and Juszczak, R.: Challenges and Limitations of Remote Sensing Applications in Northern Peatlands: Present and Future Prospects, Remote Sensing, 16, 591, https://doi.org/10.3390/rs16030591, 2024.

AMAP: Arctic climate change update 2021: Key trends and impacts. Summary for policy-makers, Arctic Monitoring and Assessment Programme (AMAP), Tromsø, Norway, 2021.

Anderson, K. and Gaston, K. J.: Lightweight unmanned aerial vehicles will revolutionize spatial ecology, Frontiers in Ecol & Environ, 11, 138–146, https://doi.org/10.1890/120150, 2013.

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

Ballinger, T. J., Overland, J. E., Wang, M., Bhatt, U. S., Hanna, E., Hanssen-Bauer, I., Kim, S.-J., Thoman, R. L., and Walsh, J. E.: Arctic Report Card 2020: Surface Air Temperature, https://doi.org/10.25923/GCW8-2Z06, 2020.

Barry, S. and Elith, J.: Error and uncertainty in habitat models, Journal of Applied Ecology, 43, 413–423, https://doi.org/10.1111/j.1365-2664.2006.01136.x, 2006.

Bartsch, A., Efimova, A., Widhalm, B., Muri, X., Von Baeckmann, C., Bergstedt, H., Ermokhina, K., Hugelius, G., Heim, B., and Leibman, M.: Circumarctic land cover diversity considering wetness gradients, Hydrol. Earth Syst. Sci., 28, 2421–2481, https://doi.org/10.5194/hess-28-2421-2024, 2024.

Becker, T., Kutzbach, L., Forbrich, I., Schneider, J., Jager, D., Thees, B., and Wilmking, M.: Do we miss the hot spots? – The use of very high resolution aerial photographs to quantify carbon fluxes in peatlands, Biogeosciences, 5, 1387–1393, https://doi.org/10.5194/bg-5-1387-2008, 2008.

Belgiu, M. and Drăguţ, L.: Random forest in remote sensing: A review of applications and future directions, ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24–31, https://doi.org/10.1016/j.isprsjprs.2016.01.011, 2016.

Biasi, C., Meyer, H., Rusalimova, O., Hämmerle, R., Kaiser, C., Baranyi, C., Daims, H., Lashchinsky, N., Barsukov, P., and Richter, A.: Initial effects of experimental warming on carbon exchange rates, plant growth and microbial dynamics of a lichen-rich dwarf shrub tundra in Siberia, Plant Soil, 307, 191–205, https://doi.org/10.1007/s11104-008-9596-2, 2008.

Berbesi, L. and Pritchard, G.: Modelling Energy Data in a Generalized Additive Model—A Case Study of Colombia, Energies, 16, 1929, https://doi.org/10.3390/en16041929, 2023.

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

Chai, T. and Draxler, R. R.: Root mean square error (RMSE) or mean absolute error (MAE)?, https://doi.org/10.5194/gmdd-7-1525-2014, 28 February 2014.

Chen, S., Liu, L., Ma, Y., Zhuang, Q., and Shurpali, N. J.: Quantifying Global Wetland Methane Emissions With In Situ Methane Flux Data and Machine Learning Approaches, Earth's Future, 12, e2023EF004330, https://doi.org/10.1029/2023EF004330, 2024.

Chollet Ramampiandra, E., Scheidegger, A., Wydler, J., and Schuwirth, N.: A comparison of machine learning and statistical species distribution models: Quantifying overfitting supports model interpretation, Ecological Modelling, 481, 110353, https://doi.org/10.1016/j.ecolmodel.2023.110353, 2023.

Cortes, C. and Vapnik, V.: Support-vector networks, Mach Learn, 20, 273–297, https://doi.org/10.1007/BF00994018, 1995.

Cui, S., Liu, P., Guo, H., Nielsen, C. K., Pullens, J. W. M., Chen, Q., Pugliese, L., and Wu, S.: Wetland hydrological dynamics and methane emissions, Commun Earth Environ, 5, 470, https://doi.org/10.1038/s43247-024-01635-w, 2024.

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., and Lawler, J. J.: RANDOM FORESTS FOR CLASSIFICATION IN ECOLOGY, Ecology, 88, 2783–2792, https://doi.org/10.1890/07-0539.1, 2007.

Dark, S. J. and Bram, D.: The modifiable areal unit problem (MAUP) in physical geography, Progress in Physical Geography: Earth and Environment, 31, 471–479, https://doi.org/10.1177/0309133307083294, 2007.

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

Dias, A. T. C., Hoorens, B., Van Logtestijn, R. S. P., Vermaat, J. E., and Aerts, R.: Plant Species Composition Can Be Used as a Proxy to Predict Methane Emissions in Peatland Ecosystems After Land-Use Changes, Ecosystems, 13, 526–538, https://doi.org/10.1007/s10021-010-9338-1, 2010.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., and Bargellini, P.: Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services, Remote Sensing of Environment, 120, 25–36, https://doi.org/10.1016/j.rse.2011.11.026, 2012.

Elith, J., Leathwick, J. R., and Hastie, T.: A working guide to boosted regression trees, Journal of Animal Ecology, 77, 802–813, https://doi.org/10.1111/j.1365-2656.2008.01390.x, 2008.

Etminan, M., Myhre, G., Highwood, E. J., and Shine, K. P.: Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing, Geophysical Research Letters, 43, https://doi.org/10.1002/2016GL071930, 2016.

Friedman, J. H.: Greedy function approximation: A gradient boosting machine., Ann. Statist., 29, https://doi.org/10.1214/aos/1013203451, 2001.

Frolking, S., Talbot, J., Jones, M. C., Treat, C. C., Kauffman, J. B., Tuittila, E.-S., and Roulet, N.: Peatlands in the Earth's 21st century climate system, Environ. Rev., 19, 371–396, https://doi.org/10.1139/a11-014, 2011.

Gao, B.: NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space, Remote Sensing of Environment, 58, 257–266, https://doi.org/10.1016/S0034-4257(96)00067-3, 1996.

Gachibu Wangari, E., Mwangada Mwanake, R., Houska, T., Kraus, D., Gettel, G. M., Kiese, R., Breuer, L., and Butterbach-Bahl, K.: Identifying landscape hot and cold spots of soil greenhouse gas fluxes by combining field measurements and remote sensing data, Biogeosciences, 20, 5029–5067, https://doi.org/10.5194/bg-20-5029-2023, 2023.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, Remote Sensing of Environment, 202, 18–27, https://doi.org/10.1016/j.rse.2017.06.031, 2017.

Greenwell, B., Boehmke, B., Cunningham, J., and Developers, G.: gbm: Generalized boosted regression models, 2022.

Grünberg, I., Wilcox, E. J., Zwieback, S., Marsh, P., and Boike, J.: Linking tundra vegetation, snow, soil temperature, and permafrost, Biogeosciences, 17, 4261–4279, https://doi.org/10.5194/bg-17-4261-2020, 2020.

Hastie, T. and Tibshirani, R.: Generalized Additive Models, Statist. Sci., 1, https://doi.org/10.1214/ss/1177013604, 1986.

Hijmans, R. J.: terra: Spatial data analysis, 2023.

Hugelius, G., Strauss, J., Zubrzycki, S., Harden, J. W., Schuur, E. A. G., Ping, C.-L., Schirrmeister, L., Grosse, G., Michaelson, G. J., Koven, C. D., O'Donnell, J. A., Elberling, B., Mishra, U., Camill, P., Yu, Z., Palmtag, J., and Kuhry, P.: Estimated stocks of circumpolar permafrost carbon with quantified uncertainty ranges and identified data gaps, Biogeosciences, 11, 6573–6593, https://doi.org/10.5194/bg-11-6573-2014, 2014.

Ivanova, K. and Göckede, M.: Processing of the carbon gas chamber flux, with automatic window detection and manual improvement, , https://doi.org/10.5281/zenodo.16732353, 2025.

Ivanova, K., Virkkala, A.-M., Voigt, C., Sonnentag, O., Brovkin, V., Stacke, T., Bartsch, A., Widhalm, B., & Goeckede, M. (2025a). CH4 Flux Dataset and Upscaling Maps for TVC, Canada, 2019–2024 [Data set]. Zenodo. https://doi.org/10.5281/zenodo.15753253

Ivanova, K., Virkkala, A.-M., Brovkin, V., Stacke, T., Widhalm, B., Bartsch, A., Voigt, C., Sonnentag, O., & Goeckede, M. (2025b). Modelling and Comparing Methane Flux Upscaling at 1m and 10m Resolution in Trail Valley Creek. Zenodo. https://doi.org/10.5281/zenodo.15399084

Janssen, P. H. M. and Heuberger, P. S. C.: Calibration of process-oriented models, Ecological Modelling, 83, 55–66, https://doi.org/10.1016/0304-3800(95)00084-9, 1995.

Kaiser, K. E., McGlynn, B. L., and Dore, J. E.: Landscape analysis of soil methane flux across complex terrain, Biogeosciences, 15, 3143–3167, https://doi.org/10.5194/bg-15-3143-2018, 2018.

Kalacska, M., Arroyo-Mora, J. P., Soffer, R. J., Roulet, N. T., Moore, T. R., Humphreys, E., Leblanc, G., Lucanus, O., and Inamdar, D.: Estimating Peatland Water Table Depth and Net Ecosystem Exchange: A Comparison between Satellite and Airborne Imagery, Remote Sensing, 10, 687, https://doi.org/10.3390/rs10050687, 2018.

Kämäräinen, M., Tuovinen, J.-P., Kulmala, M., Mammarella, I., Aalto, J., Vekuri, H., Lohila, A., and Lintunen, A.: Spatiotemporal lagging of predictors improves machine learning estimates of atmosphere–forest $CO_2$ exchange, Biogeosciences, 20, 897–909, https://doi.org/10.5194/bg-20-897-2023, 2023.

Karatzoglou, A., Smola, A., Hornik, K., and Zeileis, A.: kernlab – An S4 package for kernel methods in R, Journal of Machine Learning Research, 5, 1457–1460, 2004.

Karim, M. R., Reza, M. N., Jin, H., Haque, M. A., Lee, K.-H., Sung, J., and Chung, S.-O.: Application of LiDAR Sensors for Crop and Working Environment Recognition in Agriculture: A Review, Remote Sensing, 16, 4623, https://doi.org/10.3390/rs16244623, 2024.

Koven, C. D., Ringeval, B., Friedlingstein, P., Ciais, P., Cadule, P., Khvorostyanov, D., Krinner, G., and Tarnocai, C.: Permafrost carbon-climate feedbacks accelerate global warming, Proc. Natl. Acad. Sci. U.S.A., 108, 14769–14774, https://doi.org/10.1073/pnas.1103910108, 2011.

Kuhn, M.: Building Predictive Models in R Using the caret Package, J. Stat. Soft., 28, https://doi.org/10.18637/jss.v028.i05, 2008.

Kuhn, M. A., Varner, R. K., Bastviken, D., Crill, P., MacIntyre, S., Turetsky, M., Walter Anthony, K., McGuire, A. D., and Olefeldt, D.: BAWLD-$CH_4$ : a comprehensive dataset of methane fluxes from boreal and arctic ecosystems, Earth Syst. Sci. Data, 13, 5151–5189, https://doi.org/10.5194/essd-13-5151-2021, 2021.

Lange, S., Grünberg, I., Anders, K., Hartmann, J., Helm, V., and Boike, J.: Airborne Laser Scanning (ALS) Point Clouds of Trail Valley Creek, NWT, Canada (2018), https://doi.org/10.1594/PANGAEA.934387, 2021.

Lehmann, J., Münchberger, W., Knoth, C., Blodau, C., Nieberding, F., Prinz, T., Pancotto, V., and Kleinebecker, T.: High-Resolution Classification of South Patagonian Peat Bog Microforms Reveals Potential Gaps in Up-Scaled CH4 Fluxes by use of Unmanned Aerial System (UAS) and CIR Imagery, Remote Sensing, 8, 173, https://doi.org/10.3390/rs8030173, 2016.

Lindsay, J. B.: Whitebox GAT: A case study in geomorphometric analysis, Computers & Geosciences, 95, 75–84, https://doi.org/10.1016/j.cageo.2016.07.003, 2016.

Ludwig, S. M., Schiferl, L., Hung, J., Natali, S. M., and Commane, R.: Resolving heterogeneous fluxes from tundra halves the growing season carbon budget, Biogeosciences, 21, 1301–1321, https://doi.org/10.5194/bg-21-1301-2024, 2024.

Marsh, P., Bartlett, P., MacKay, M., Pohl, S., and Lantz, T.: Snowmelt energetics at a shrub tundra site in the western Canadian Arctic, Hydrological Processes, 24, 3603–3620, https://doi.org/10.1002/hyp.7786, 2010.

Mastepanov, M., Sigsgaard, C., Tagesson, T., Ström, L., Tamstorf, M. P., Lund, M., and Christensen, T. R.: Revisiting factors controlling methane emissions from high-Arctic tundra, Biogeosciences, 10, 5139–5158, https://doi.org/10.5194/bg-10-5139-2013, 2013.

Matthes, J. H., Sturtevant, C., Verfaillie, J., Knox, S., and Baldocchi, D.: Parsing the variability in $CH_4$ flux at a spatially heterogeneous wetland: Integrating multiple eddy covariance towers with high-resolution flux footprint analysis, JGR Biogeosciences, 119, 1322–1339, https://doi.org/10.1002/2014JG002642, 2014.

McFEETERS, S. K.: The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features, International Journal of Remote Sensing, 17, 1425–1432, https://doi.org/10.1080/01431169608948714, 1996.

McNicol, G., Fluet-Chouinard, E., Ouyang, Z., Knox, S., Zhang, Z., Aalto, T., Bansal, S., Chang, K., Chen, M., Delwiche, K., Feron, S., Goeckede, M., Liu, J., Malhotra, A., Melton, J. R., Riley, W., Vargas, R., Yuan, K., Ying, Q., Zhu, Q., Alekseychik, P., Aurela, M., Billesbach, D. P., Campbell, D. I., Chen, J., Chu, H., Desai, A. R., Euskirchen, E., Goodrich, J., Griffis, T., Helbig, M., Hirano, T., Iwata, H., Jurasinski, G., King, J., Koebsch, F., Kolka, R., Krauss, K., Lohila, A., Mammarella, I., Nilson, M., Noormets, A., Oechel, W., Peichl, M., Sachs, T., Sakabe, A., Schulze, C., Sonnentag, O., Sullivan, R. C., Tuittila, E., Ueyama, M., Vesala, T., Ward, E., Wille, C., Wong, G. X., Zona, D., Windham-Myers, L., Poulter, B., and Jackson, R. B.: Upscaling Wetland Methane Emissions From the FLUXNET-CH4 Eddy Covariance Network (UpCH4 v1.0): Model Development, Network Assessment, and Budget Comparison, AGU Advances, 4, e2023AV000956, https://doi.org/10.1029/2023AV000956, 2023.

Natekin, A. and Knoll, A.: Gradient boosting machines, a tutorial, Front. Neurorobot., 7, https://doi.org/10.3389/fnbot.2013.00021, 2013.

Nisbet, E. G., Manning, M. R., Dlugokencky, E. J., Fisher, R. E., Lowry, D., Michel, S. E., Myhre, C. L., Platt, S. M., Allen, G., Bousquet, P., Brownlow, R., Cain, M., France, J. L., Hermansen, O., Hossaini, R., Jones, A. E., Levin, I., Manning, A. C., Myhre, G., Pyle, J. A., Vaughn, B. H., Warwick, N. J., and White, J. W. C.: Very Strong Atmospheric Methane Growth in the 4 Years 2014–2017: Implications for the Paris Agreement, Global Biogeochemical Cycles, 33, 318–342, https://doi.org/10.1029/2018GB006009, 2019.

Olefeldt, D., Turetsky, M. R., Crill, P. M., and McGuire, A. D.: Environmental and physical controls on northern terrestrial methane emissions across permafrost zones, Global Change Biology, 19, 589–603, https://doi.org/10.1111/gcb.12071, 2013.

Olefeldt, D., Goswami, S., Grosse, G., Hayes, D., Hugelius, G., Kuhry, P., McGuire, A. D., Romanovsky, V. E., Sannel, A. B. K., Schuur, E. A. G., and Turetsky, M. R.: Circumpolar distribution and carbon storage of thermokarst landscapes, Nat Commun, 7, 13043, https://doi.org/10.1038/ncomms13043, 2016.

Papale, D. and Valentini, R.: A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization, Global Change Biology, 9, 525–535, https://doi.org/10.1046/j.1365-2486.2003.00609.x, 2003.

Peltola, O., Vesala, T., Gao, Y., Räty, O., Alekseychik, P., Aurela, M., Chojnicki, B., Desai, A. R., Dolman, A. J., Euskirchen, E. S., Friborg, T., Göckede, M., Helbig, M., Humphreys, E., Jackson, R. B., Jocher, G., Joos, F., Klatt, J., Knox, S. H., Kowalska, N., Kutzbach, L., Lienert, S., Lohila, A., Mammarella, I., Nadeau, D. F., Nilsson, M. B., Oechel, W. C., Peichl, M., Pypker, T., Quinton, W., Rinne, J., Sachs, T., Samson, M., Schmid, H. P., Sonnentag, O., Wille, C., Zona, D., and Aalto, T.: Monthly gridded data product of northern wetland methane emissions based on upscaling eddy covariance observations, Earth Syst. Sci. Data, 11, 1263–1289, https://doi.org/10.5194/essd-11-1263-2019, 2019.

Pohl, S. and Marsh, P.: Modelling the spatial–temporal variability of spring snowmelt in an arctic catchment, Hydrol. Process., 20, 1773–1792, https://doi.org/10.1002/hyp.5955, 2006.

Porter, C., Howat, I., Noh, M.-J., Husby, E., Khuvis, S., Danish, E., Tomko, K., Gardiner, J., Negrete, A., Yadav, B., Klassen, J., Kelleher, C., Cloutier, M., Bakker, J., Enos, J., Arnold, G., Bauer, G., Morin, P., and Polar Geospatial Center: ArcticDEM - Mosaics, Version 4.1 (1.0), https://doi.org/10.7910/DVN/3VDC4W, 2023.

Prasad, A. M., Iverson, L. R., and Liaw, A.: Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction, Ecosystems, 9, 181–199, https://doi.org/10.1007/s10021-005-0054-1, 2006.

Previdi, M., Smith, K. L., and Polvani, L. M.: Arctic amplification of climate change: a review of underlying mechanisms, Environ. Res. Lett., 16, 093003, https://doi.org/10.1088/1748-9326/ac1c29, 2021.

Rantanen, M., Karpechko, A. Yu., 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, Commun Earth Environ, 3, 168, https://doi.org/10.1038/s43247-022-00498-3, 2022.

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


Räsänen, A. and Virtanen, T.: Data and resolution requirements in mapping vegetation in spatially heterogeneous landscapes, Remote Sensing of Environment, 230, 111207, https://doi.org/10.1016/j.rse.2019.05.026, 2019.

Räsänen, A., Manninen, T., Korkiakoski, M., Lohila, A., and Virtanen, T.: Predicting catchment-scale methane fluxes with multi-source remote sensing, Landscape Ecol, 36, 1177–1195, https://doi.org/10.1007/s10980-021-01194-x, 2021.

Rettelbach, T., Nitze, I., Grünberg, I., Hammar, J., Schäffler, S., Hein, D., Gessner, M., Bucher, T., Brauchle, J., Hartmann, J., Sachs, T., Boike, J., and Grosse, G.: Super-high-resolution aerial imagery, digital surface models and 3D point clouds of Trail Valley Creek, Canada, https://doi.org/10.1594/PANGAEA.961942, 2024.

Riihimäki, H., Kemppinen, J., Kopecký, M., and Luoto, M.: Topographic Wetness Index as a Proxy for Soil Moisture: The Importance of Flow-Routing Algorithm and Grid Resolution, Water Resources Research, 57, e2021WR029871, https://doi.org/10.1029/2021WR029871, 2021.

Rouse, J., Haas, R., Schell, J., and Deering, D.: Monitoring vegetation systems in the great plains with ERTS. 1974, in: Proceeding of erts-1 symposium, Citation Key: rouse1974monitoring, 1974.

Ruhoff, A. L., Castro, N. M. R., and Risso, A.: Numerical modelling of the topographic wetness index: An analysis at different scales, International Journal of Geosciences, 02, 476–483, 2011.

Runge, A. and Grosse, G.: Comparing Spectral Characteristics of Landsat-8 and Sentinel-2 Same-Day Data for Arctic-Boreal Regions, Remote Sensing, 11, 1730, https://doi.org/10.3390/rs11141730, 2019.

Scheller, J. H., Mastepanov, M., and Christensen, T. R.: Toward UAV-based methane emission mapping of Arctic terrestrial ecosystems, Science of The Total Environment, 819, 153161, https://doi.org/10.1016/j.scitotenv.2022.153161, 2022.

Schoorl, J. M., Sonneveld, M. P. W., and Veldkamp, A.: Three-dimensional landscape process modelling: the effect of DEM resolution, Earth Surf. Process. Landforms, 25, 1025–1034, https://doi.org/10.1002/1096-9837(200008)25:9<1025::AID-ESP116>3.0.CO;2-Z, 2000.

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

Sjögersten, S., Ledger, M., Siewert, M., De La Barreda-Bautista, B., Sowter, A., Gee, D., Foody, G., and Boyd, D. S.: Optical and radar Earth observation data for upscaling methane emissions linked to permafrost degradation in sub-Arctic peatlands in northern Sweden, Biogeosciences, 20, 4221–4239, https://doi.org/10.5194/bg-20-4221-2023, 2023.

Smola, A. J. and Schölkopf, B.: A tutorial on support vector regression, Statistics and Computing, 14, 199–222, https://doi.org/10.1023/B:STCO.0000035301.49549.88, 2004.

Ström, L., Mastepanov, M., and Christensen, T. R.: Species-specific Effects of Vascular Plants on Carbon Turnover and Methane Emissions from Wetlands, Biogeochemistry, 75, 65–82, https://doi.org/10.1007/s10533-004-6124-1, 2005.

Subke, J.-A., Kutzbach, L., and Risk, D.: Soil chamber measurements, in: Springer handbook of atmospheric measurements, Springer, 1603–1624, 2021.

Tagesson, T., Mastepanov, M., Mölder, M., Tamstorf, M. P., Eklundh, L., Smith, B., Sigsgaard, C., Lund, M., Ekberg, A., Falk, J. M., Friborg, T., Christensen, T. R., and Ström, L.: Modelling of growing season methane fluxes in a high-Arctic wet tundra ecosystem 1997–2010 using in situ and high-resolution satellite data, Tellus B: Chemical and Physical Meteorology, 65, 19722, https://doi.org/10.3402/tellusb.v65i0.19722, 2013.

The MathWorks Inc. (2020). MATLAB version: 9.9 (R2020b), Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M. A., Cescatti, A., Kiely, G., Merbold, L., Serrano-Ortiz, P., Sickert, S., Wolf, S., and Papale, D.: Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms, Biogeosciences, 13, 4291–4313, https://doi.org/10.5194/bg-13-4291-2016, 2016.

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


Treat, C. C., Virkkala, A., Burke, E., Bruhwiler, L., Chatterjee, A., Fisher, J. B., Hashemi, J., Parmentier, F. W., Rogers, B. M., Westermann, S., Watts, J. D., Blanc-Betes, E., Fuchs, M., Kruse, S., Malhotra, A., Miner, K., Strauss, J., Armstrong, A., Epstein, H. E., Gay, B., Goeckede, M., Kalhori, A., Kou, D., Miller, C. E., Natali, S. M., Oh, Y., Shakil, S., Sonnentag, O., Varner, R. K., Zolkos, S., Schuur, E. A. G., and Hugelius, G.: Permafrost Carbon: Progress on Understanding Stocks and Fluxes Across Northern Terrestrial Ecosystems, JGR Biogeosciences, 129, e2023JG007638, https://doi.org/10.1029/2023JG007638, 2024.

Turetsky, M. R., Abbott, B. W., Jones, M. C., Anthony, K. W., Olefeldt, D., Schuur, E. A. G., Grosse, G., Kuhry, P., Hugelius, G., Koven, C., Lawrence, D. M., Gibson, C., Sannel, A. B. K., and McGuire, A. D.: Carbon release through abrupt permafrost thaw, Nat. Geosci., 13, 138–143, https://doi.org/10.1038/s41561-019-0526-0, 2020.

Vaze, J., Teng, J., and Spencer, G.: Impact of DEM accuracy and resolution on topographic indices, Environmental Modelling & Software, 25, 1086–1098, https://doi.org/10.1016/j.envsoft.2010.03.014, 2010.

Vogt, J., Pallandt, M. M. T. A., Basso, L. S., Bolek, A., Ivanova, K., Schlutow, M., Celis, G., Kuhn, M., Mauritz, M., Schuur, E. A. G., Arndt, K., Virkkala, A.-M., Wargowsky, I., and Göckede, M.: ARGO: ARctic greenhouse Gas Observation metadata version 1, Earth Syst. Sci. Data, 17, 2553–2573, https://doi.org/10.5194/essd-17-2553-2025, 2025.

Voigt, C., Sonnentag, O., Nesic, Z., and Hould Gosselin, G.: Atmospheric methane consumption by upland soils - dataset 1: high temporal resolution methane fluxes, https://doi.org/10.1594/PANGAEA.953117, 2023a.

Voigt, C., Sonnentag, O., Virkkala, A.-M., Bennett, K. A., Marushchak, M. E., Saarela, T., Wilcox, E. J., Chevrier-Dion, C., and Marquis, C.: Atmospheric methane consumption by upland soils - dataset 2: campaign-based manual chamber measurements, https://doi.org/10.1594/PANGAEA.953119, 2023b.

Voigt, C., Virkkala, A.-M., Hould Gosselin, G., Bennett, K. A., Black, T. A., Detto, M., Chevrier-Dion, C., Guggenberger, G., Hashmi, W., Kohl, L., Kou, D., Marquis, C., Marsh, P., Marushchak, M. E., Nesic, Z., Nykänen, H., Saarela, T., Sauheitl, L., Walker, B., Weiss, N., Wilcox, E. J., and Sonnentag, O.: Arctic soil methane sink increases with drier conditions and higher ecosystem respiration, Nat. Clim. Chang., 13, 1095–1104, https://doi.org/10.1038/s41558-023-01785-3, 2023c.

Weiss, A. D.: Topographic position and landforms analysis, 2001.

Widhalm, B., Bartsch, A., Strozzi, T., Jones, N., Khomutov, A., Babkina, E., Leibman, M., Khairullin, R., Göckede, M., Bergstedt, H., Von Baeckmann, C., and Muri, X.: InSAR-derived seasonal subsidence reflects spatial soil moisture patterns in Arctic lowland permafrost regions, The Cryosphere, 19, 1103–1133, https://doi.org/10.5194/tc-19-1103-2025, 2025.

Wille, C., Kutzbach, L., Sachs, T., Wagner, D., and Pfeiffer, E.: Methane emission from Siberian arctic polygonal tundra: eddy covariance measurements and modelling, Global Change Biology, 14, 1395–1408, https://doi.org/10.1111/j.1365-2486.2008.01586.x, 2008.

Willmott, C. and Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Clim. Res., 30, 79–82, https://doi.org/10.3354/cr030079, 2005.

Wilson, J. P. and Gallant, J. C.: Terrain analysis: principles and applications, John Wiley & Sons, 2000.

Wood, S. N.: Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models, Journal of the Royal Statistical Society Series B: Statistical Methodology, 73, 3–36, https://doi.org/10.1111/j.1467-9868.2010.00749.x, 2011.

Wood, S. N.: Generalized Additive Models: An Introduction with R, 2nd ed., Chapman and Hall/CRC, https://doi.org/10.1201/9781315370279, 2017.

Wright, M. N. and Ziegler, A.: ranger : A Fast Implementation of Random Forests for High Dimensional Data in C++and R, J. Stat. Soft., 77, https://doi.org/10.18637/jss.v077.i01, 2017.

Ying, Q., Poulter, B., Watts, J. D., Arndt, K. A., Virkkala, A.-M., Bruhwiler, L., Oh, Y., Rogers, B. M., Natali, S. M., Sullivan, H., Armstrong, A., Ward, E. J., Schiferl, L. D., Elder, C. D., Peltola, O., Bartsch, A., Desai, A. R., Euskirchen, E., Göckede, M., Lehner, B., Nilsson, M. B., Peichl, M., Sonnentag, O., Tuittila, E.-S., Sachs, T., Kalhori, A., Ueyama, M., and Zhang, Z.:

https://doi.org/10.5194/egusphere-2025-3968
Preprint. Discussion started: 28 August 2025


WetCH$_4$ : a machine-learning-based upscaling of methane fluxes of northern wetlands during 2016–2022, Earth Syst. Sci. Data, 17, 2507–2534, https://doi.org/10.5194/essd-17-2507-2025, 2025.

Zhang, C., Comas, X., and Brodylo, D.: A Remote Sensing Technique to Upscale Methane Emission Flux in a Subtropical Peatland, JGR Biogeosciences, 125, e2020JG006002, https://doi.org/10.1029/2020JG006002, 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, Proc. Natl. Acad. Sci. U.S.A., 113, 40–45, https://doi.org/10.1073/pnas.1516017113, 2016.

NASA Langley Research Center. (2024). Prediction Of Worldwide Energy Resources (POWER) Project. The data was obtained from the POWER Project's Hourly 2.5.1 version on January 24, 2025, from https://power.larc.nasa.gov.

Environment and Climate Change Canada (ECCC): Weather station data for trail valley creek (climate ID: 220N005, WMO ID: 71683), northwest territories, canada, 2003–2024, 2024.

R Core Team (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/