Disaggregating the carbon exchange of degrading permafrost peatlands using Bayesian deep learning

Norbert Pirk[1], Kristoffer Aalstad[1], Erik Schytt Mannerfelt[1], François Clayer[2], Heleen de Wit[2], Casper T. Christiansen[3], Inge Althuizen[4], Hanna Lee[5], Sebastian Westermann[1]

[1]Department of Geosciences, University of Oslo, Oslo, Norway
[2]Norwegian Institute for Water Research (NIVA), Oslo, Norway
[3]University of Copenhagen, Copenhagen, Denmark
[4]NORCE Norwegian Research Centre, Bergen, Norway
[5]Norwegian University of Science and Technology, Trondheim, Norway

Key Points:

- Eddy covariance fluxes are disaggregated for different surfaces using Bayesian neural networks to derive uncertainty-aware carbon balances
- While palsa areas have a near-zero annual methane balance, the fens and ponds that form upon palsa degradation emit large amounts of methane
- Fens compensate for methane emissions with strong annual CO_2 sinks, while ponds appear as strong, yet uncertain, CO_2 emission hotspots

Corresponding author: Norbert Pirk, norbert.pirk@geo.uio.no

Abstract

Extensive regions in the permafrost zone are projected to become climatically unsuitable to sustain permafrost peatlands over the next century, suggesting transformations in these landscapes that can leave large amounts of permafrost carbon vulnerable to post-thaw decomposition. We present three years of eddy covariance measurements of CH_4 and CO_2 fluxes from the degrading permafrost peatland Iškoras in Northern Norway, which we disaggregate into separate fluxes of palsa, pond, and fen areas using information provided by the dynamic flux footprint in a novel ensemble-based Bayesian deep neural network framework. The three-year mean CO_2-equivalent flux is estimated to be 106 $gCO_2\ m^{-2}\ yr^{-1}$ for palsas, 1780 $gCO_2\ m^{-2}\ yr^{-1}$ for ponds, and -31 $gCO_2\ m^{-2}\ yr^{-1}$ for fens, indicating that possible palsa degradation to thermokarst ponds would strengthen the local greenhouse gas forcing by a factor of about 17, while transformation into fens would slightly reduce the current local greenhouse gas forcing.

Plain Language Summary

Arctic and sub-arctic regions on the southern border of the permafrost zone often feature peatlands with a patchy surface of peat mounds, thaw ponds, and surrounding fens. As the permafrost underneath peat mounds thaws, these areas transform and can change their emission or uptake of greenhouse gases like CO_2 and methane. Assessing this gas exchange on the patchy surface is difficult because our measurement techniques cannot directly observe the variability in space and time. We collected three years of gas exchange measurements at a Norwegian permafrost peatland and developed a new method using a collection of uncertainty-aware neural networks to predict the greenhouse gas exchange of different surface types. Our work suggests that large amounts of methane are emitted by ponds and fens, while the elevated peat mounds have almost no methane emissions. For CO_2, we see that ponds are strong emitters, while fens take up substantial amounts as their vegetation absorbs this gas. We are still unsure when the peat mounds will collapse and if they turn into ponds or fens, but we can say that pond formation would give a 17 fold increase in greenhouse gas emissions, while fen formation would slightly reduce today's emissions of permafrost peatlands.

1 Introduction

Permafrost peatlands are considered to be some of the most dynamic and rapidly changing ecosystems in the permafrost zone (Olefeldt et al., 2016). These ecosystems cover large areas in often harsh and inaccessible arctic regions, playing an important role in the global carbon cycle as they have historically accumulated large amounts of soil organic carbon which is vulnerable to microbial re-mobilization upon climate warming (Oechel et al., 2000; Schuur et al., 2015). The characteristic palsa peat mounds—elevated by excess ground ice in the permafrost—are increasingly subject to thawing and degradation, resulting in thermokarst pond or wetland formation (Luoto & Seppälä, 2003; Sannel & Kuhry, 2011; Grosse et al., 2013; Borge et al., 2017; Martin et al., 2021). While irreversible palsa collapse is widely anticipated in large regions of the pan-Arctic over the coming century (Aas et al., 2019), the dynamics of the degradation as well as the state and fate of their carbon stocks remain elusive.

The emission and uptake of the greenhouse gases carbon dioxide (CO_2) and methane (CH_4) is associated with a stark heterogeneity as a result of the patchy surface cover consisting of different plant and microbial communities, which can inform possible future trends of land-atmosphere interactions through space-for-time substitutions (e.g., Jiao et al., 2023). While moist or inundated areas typically feature microbes with anaerobic metabolisms producing CH_4, drier areas are dominated by aerobic soil respiration that produces CO_2 and may even consume considerable amounts of atmospheric CH_4 (Voigt et al., 2019). Thermokarst ponds are typically supersaturated in dissolved CO_2 and CH_4,

and emissions can be associated with diffusive fluxes (Matveev et al., 2016) as well as gas ebullition (Walter et al., 2006; Serikova et al., 2019). The pathways of gas production, consumption, dissolution, transport, and emission in permafrost peatlands depend on a complex interplay of biogeochemical processes that are influenced by a plethora of interacting environmental factors, including soil, surface, and atmospheric conditions. As a result, CO_2 and CH_4 exchange can vary tremendously on small spatio-temporal scales, which complicates the representative (unbiased) quantification of the greenhouse gas budgets of permafrost peatlands based on sparse flux measurements. Consequently, the quantification of the greenhouse gas balances not only reflects the ecosystem in its environmental setting, but can also depend strongly on the method used to estimate them.

Using the manual chamber technique Nykänen et al. (2003) estimated that palsa surfaces with shrub vegetation in northern Finland were sinks of (atmospheric) carbon, whereas palsa surfaces with sparse vegetation were carbon sources. In their study, the annual emissions of CH_4 ranged from 1.0 gC m^{-2} yr^{-1} on top of the palsas to 24.7 gC m^{-2} yr^{-1} at the palsa margins. However, the manual chamber technique only allows for sporadic spatio-temporal sampling, and unwanted disturbances by the manual deployment of the chamber are inevitable (Kutzbach et al., 2007), adding uncertainty to the estimated annual budgets. Year-round automatic chamber measurements in the Stordalen permafrost peatland in Sweden indicate CO_2 sinks and CH_4 sources in the mire, with a net carbon balance of -13 gC m^{-2} yr^{-1} for palsa areas and -91 gC m^{-2} yr^{-1} for fen areas (Holmes et al., 2022). While such estimates are derived from near-continuous flux measurements, the long-term presence of the chamber base can disturb the ecosystem as it locally increases air temperature similar to an open-top chamber (Frei et al., 2020) and a fair number of automatic chambers are required to obtain spatially representative estimates. Moreover, potentially important flux hotspots in wetlands can be unsuitable for the operation of an automatic flux chamber, because water levels can be too high or too variable. Measurements with the micro-meteorological eddy covariance (EC) technique (Baldocchi, 2020) in a nearby palsa-dominated area in Stordalen indicate a sink of CO_2 with an annual balance amounting to between -20 and -95 gC m^{-2} yr^{-1} across the years, as well as relatively stable CH_4 emissions of between 18 and 22 gC m^{-2} yr^{-1} (Christensen et al., 2012). The anemometer and gas analyzer needed for EC measurements can in principle be operated year-round, but unfavorable micro-meteorological conditions due to a lack of stationarity or weak turbulent mixing will inevitably cause gaps in the flux time series. As these gaps tend to occur systematically, e.g., in very stable conditions during nights or wintertime, gap-filling is needed to avoid biased seasonal or annual flux budgets. While EC measurements are widely regarded as the most accurate flux measurements on the landscape scale, it must be noted that the flux footprint of the measurements changes continuously depending on the wind conditions. So unless the ecosystem around the flux tower can be considered spatially homogeneous, the flux time series will contain confounding effects of both spatial and temporal variability. A possible indication of this effect can be seen in the EC measurements from a Siberian palsa mire reported by Olchev et al. (2022), where CH_4 fluxes show alternating uptake and release fluxes, possibly as a result of the surface heterogeneity. A simple flux footprint disaggregation by wind sector can be a viable option in special cases (Griebel et al., 2016; Pirk et al., 2023), but a generally accepted disaggregation method remains lacking.

Levy et al. (2020) present a Bayesian method to infer spatial heterogeneity in surface fluxes from individual control variables at an EC tower using the information provided by the temporally changing footprint. As the surface fluxes of CO_2 and CH_4 result from a complex interplay of biogeochemical processes, the more advanced parameterizations encoded in process-based land-surface models (Qiu et al., 2018; Lawrence et al., 2019) or non-linear data-driven models such as deep neural networks (Krizhevsky et al., 2012; LeCun et al., 2015; Murphy, 2023) can be appropriate options. Neural networks can in principle approximate any functional relationship between inputs (predictors) and outputs (fluxes) (Hornik et al., 1989), but their parameters (weights of the net-

work edges and bias terms of the nodes) are less interpretable than those used in process-based models (Rudin, 2019). In practice, finding a suitable network architecture for a given problem can be challenging and training these networks may require vast amounts of data for complex relationships. To alleviate these challenges, one may incorporate Bayesian inference into the training process of the network by treating the model parameters as random variables with probability distributions representing their uncertainty. Such Bayesian neural networks (BNNs) produce uncertainty-aware outputs and—while being an old concept (Neal, 1996; MacKay, 2003)—are becoming increasingly popular in machine learning where uncertainty awareness is becoming a vital consideration (Ghahramani, 2015). In Earth system science, BNNs remain a relatively unexplored topic (Clare et al., 2022; Lopez-Gomez et al., 2022), despite the popularity of deep learning (Reichstein et al., 2019).

Here, we present three years of EC fluxes of CO_2 and CH_4 collected at a permafrost peatland in northern Norway. We develop a new flux disaggregation method using ensemble-based Bayesian deep learning with predictors from in-situ measurements and satellite remote sensing to estimate uncertainty-aware fluxes separately for palsa, pond, and fen areas. We use historic aerial photography as well as a modern drone-based survey of terrain changes to characterize the permafrost degradation and inform future scenarios for the carbon balance of permafrost peatlands through a space-for-time substitution.

2 Materials and Methods

2.1 Site description and surface characterization

Our study is conducted at the Iškoras permafrost peatland (69.34°N, 25.30°E, 380 m a.s.l., shown in Figure S1 in the Supplement), which is located on the Finnmarksvidda mountain plateau in northern Norway. The climate here is classified as subarctic or polar, with a mean annual air temperature of -1.2°C and a mean annual precipitation of 417 mm for the period 1991-2020 (measured at weather station SN97251 approximately 15 km north of Iškoras). The site lies just above the current tree line with mountain birch trees, and features typical upland tundra vegetation. Shrubs and lichens dominate dry areas, while sedges and mosses dominate in wetter areas. The site features sporadic permafrost with organic-rich peat soils and active layer depths of up to 90 cm on the palsas.

We geo-referenced a historic aerial photograph taken in 1955 by the Norwegian Mapping Authority (Figure 1b, Kartverket survey WF-688 H-13) and conducted multiple drone surveys during our three-year study period. We produced digital elevation models using the structure-from-motion technique (Ullman, 1979) from our drone imagery from 2019 and 2022, which we subtracted to estimate the surface subsidence. We classified the landscape at the site into three discrete surface types (palsas, ponds, fens) based on a visual inspection of our ortho-rectified drone imagery from 2019 (Figure 1a).

2.2 Eddy covariance flux measurements

The EC flux system shown in Figure 2a was established at the Iškoras site in March 2019. The data period in the present study covers three years, i.e., until March 2022. The EC system consists of a CSAT3 three-dimensional sonic anemometer (Campbell Scientific, USA), an Li-7200 closed-path infrared gas analyzer for CO_2, as well as an Li-7700 open path gas analyzer for CH_4 (both Li-Cor, USA). The system is supplied with an off-grid power supply based on a wind generator and solar panels. An electronic relay is used to turn off the EC system when the battery voltage of the power supply becomes too low, while the sensors and logger of the ancillary measurements (described in Text S2 in the Supplement), which consume only a fraction of the power of the EC system, continue to operate. We processed the EC raw data to 30 minute flux estimates following the conventional EC methodology (Gu et al., 2012). After filtering the flux time series for unfavorable measurement conditions, e.g., due to a lack of stationarity or turbulent mix-

ing (see Text S1 in the Supplement for details), we are left with 11 334 and 4 743 valid half-hourly flux estimates for CO_2 and CH_4, respectively.

The total flux estimate from EC is the mathematical convolution of the surface flux distribution with the flux footprint function, which we estimate using the flux footprint model by Kljun et al. (2015). The resulting footprint weight-maps are combined with the surface type classification (see Section 2.1 and Figure 1a) to estimate the weight of the contribution of each of the three surface types (w_{palsa}, w_{ponds}, w_{fen}) to each 30-minute EC flux estimate. The resulting average distribution (climatology) of these dynamic footprint weights is shown in Figure 1d.

2.3 Bayesian neural networks

For our BNN flux disaggregation model we use a fully-connected feedforward neural network, also known as a multilayer perceptron, with ten predictors as inputs (ancillary variables) and one total flux as output (either CO_2 or CH_4, so we train two BNNs separately), as depicted in Figure 2b. We use a total of 300 nodes placed in five hidden layers with respectively 96, 48, 12, 48, and 96 nodes per layer, resembling the architecture of an auto-encoder (Goodfellow et al., 2016). This architecture results in a total of $N_p = 11\,919$ parameters (network weights and biases) collectively denoted through the random vector $\boldsymbol{\theta} \in \mathbb{R}^{N_p}$ that we infer. As is usually the case in deep learning there are thus more parameters than data points (Murphy, 2022), in which case the Bayesian approach adopted herein helps to regularize the problem and avoid overfitting (MacKay, 2003; Murphy, 2023). At each node, the inputs are multiplied by weights, summed, combined with an additive bias term, and passed through an activation function to produce the node's output. For all the hidden layers we employ the widely used Rectified Linear Unit (ReLU) non-linear activation function defined as $\text{ReLU}(x) = \max(0, x)$. Our BNN can be thought of as having two output layers with linear activation functions: the first output layer consists of three nodes, predicting the CO_2 or CH_4 fluxes for each of the three surface types based on the dynamic inputs $\mathbf{x}(t)$ and the (static) uncertain network parameters $\boldsymbol{\theta}$. In the second output layer, these fluxes are averaged using the deterministic dynamic footprint weights $\mathbf{w}_s(t)$ for each surface type at the corresponding 30-minute interval (see Section 2.2), predicting the total flux F_{total} that can be compared to observations from the EC system, i.e.,

$$F_{\text{total}}(\mathbf{w}_s, \mathbf{x}, \boldsymbol{\theta}) = w_{\text{palsa}}\, F_{\text{palsa}}(\mathbf{x}, \boldsymbol{\theta}) + w_{\text{ponds}}\, F_{\text{ponds}}(\mathbf{x}, \boldsymbol{\theta}) + w_{\text{fen}}\, F_{\text{fen}}(\mathbf{x}, \boldsymbol{\theta}) \tag{1}$$

This innovative network architecture with two consecutive output layers serves as the basis for the envisioned flux disaggregation between the three surface types. Note that this disaggregation assumes that the within-class flux is spatially homogeneous and can thus only estimate the spatial average of within-class flux dynamics.

The BNN parameters are initialized by drawing from a standard normal distribution as the (weakly informative) prior distribution. Unlike the more conventional approach of training a neural network by optimizing the parameters via backpropagation, we train our network parameters using (approximate) Bayesian inference techniques developed for geophysical data assimilation (Evensen et al., 2022), namely an iterative ensemble Kalman method (Emerick & Reynolds, 2013) (see details in Text S3 in the Supplement). To better represent the typically multi-modal posterior parameter distributions (Izmailov et al., 2021), we repeat the BNN training 100 times with different random seeds to capture local modes and combine these 100 local ensembles to form one global ensemble, as a so-called deep ensembles approximation (Lakshminarayanan et al., 2017; Wilson & Izmailov, 2020) of the posterior predictive distribution (see Text S3 in the Supplement). For the ensemble data assimilation-based BNN training, we assume a typical zero mean additive Gaussian observation error model with observation error standard deviations

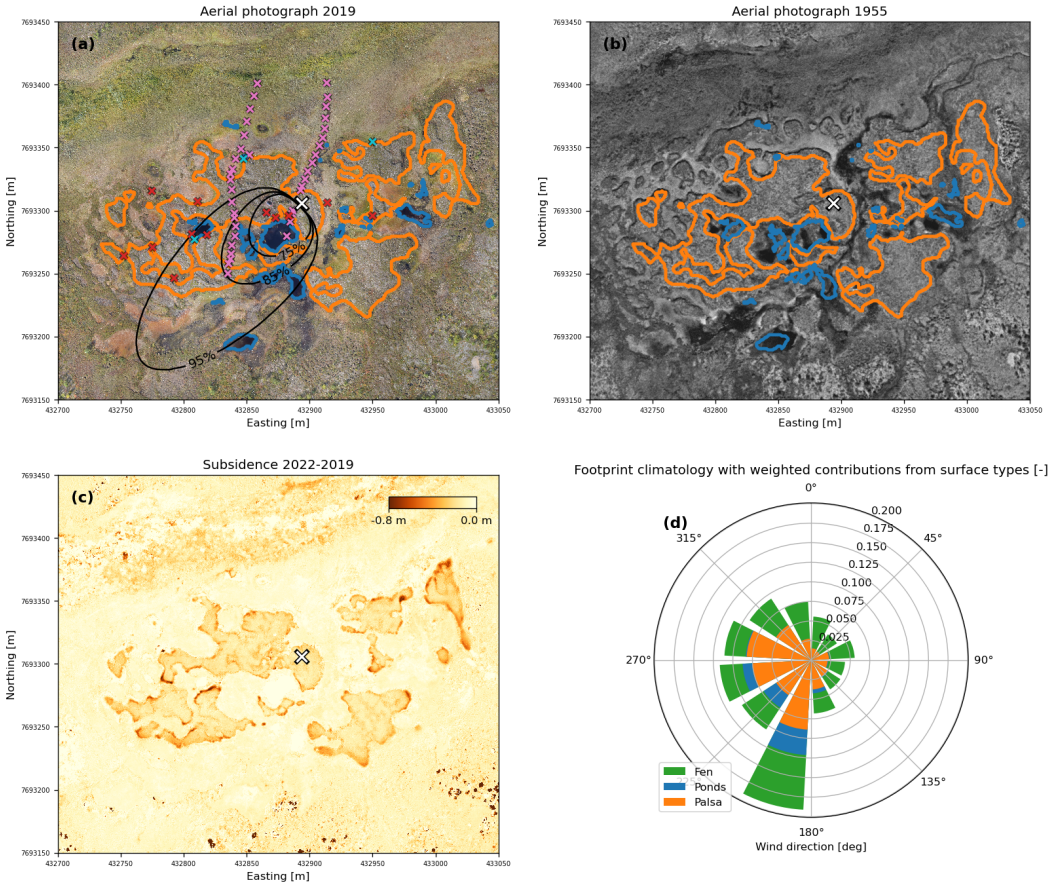

Figure 1. Surface characterization of the Iškoras permafrost peatland. (a): Ortho-rectified aerial photographs from 2019 with contour lines for palsa (orange) and pond (blue) areas (all other areas are classified as fen). The white cross marks the location of the flux tower, from which the black lines show an example of the 75%, 85%, and 95% contours of the cumulative flux footprint function for 25 July 2019, 16:00 UTC. Colored crosses indicate independent validation flux measurements by chambers (red), dissolved gas concentrations in ponds (cyan), and snowpack profiles (pink). (b): Ortho-rectified aerial photographs taken in 28 July 1955, overlain with the same contour lines for palsa and pond extent in 2019. (c): Vertical difference between elevation models from September 2022 and 2019 indicating surface subsidence. (d): Averaged footprint weights of all valid flux measurements plotted by the corresponding wind sectors. Colors indicate the footprint-weighted contribution of each surface type.

of 0.1 μmol m^{-2} s^{-1} for CO_2 and 2.5 nmol m^{-2} s^{-1} for CH_4. As predictors, we use air, surface, and soil temperature (T_{air}, T_{surf}, and T_{soil}, respectively), vapor pressure deficit (VPD), shortwave and longwave incoming radiation (SW_{in} and LW_{in}, respectively), albedo, fractional snow-covered area (FSCA), the Normalized Difference Vegetation index (NDVI), and soil volumetric water content (VWC), estimated from in-situ measurements and remote sensing data as described in Text S2 in the Supplement. Figure S2 in the Supplement shows the scatter plot matrix of the predictors, fluxes, and footprint weights, which are all archived and avalable (Pirk, 2023). As the predictors are available continuously for the entire three year campaign, the BNN can simultaneously perform both flux disaggregation and gap-filling.

A performance evaluation for the prediction of the total flux using a common 80%-20% train-test split (Murphy, 2022) indicates good prediction accuracy and generalization, with normalized root mean square error values of between 5 and 11% for both train and test datasets (Figure S4 in the Supplement). The coefficients of determination (R^2) for CO_2 range between 0.7 and 0.8. For CH_4, we notice that while the dynamics of daily average fluxes are well captured in the BNN model ($0.71 < R^2 < 0.72$), the model performs notably worse for the dynamics of the instantaneous 30-minute fluxes ($0.28 < R^2 < 0.29$), see discussion in Section 3.3.

2.4 Independent flux validation

As an independent validation of the disaggregated flux results, we conducted manual measurement campaigns to estimate fluxes on the plot scale using flux chambers, dissolved gas concentrations, and snowpack gradients. The sampling locations were distributed around the EC tower as shown in Figure 1a. Chamber flux measurements of CO_2 and CH_4 were performed in palsa and fen areas on 2-3 July 2021, and 11 September 2021. We used a Li-7810 gas analyzer (Li-Cor, USA), with a plexiglass chamber covering 25$\times$ 25 cm^2, and followed Pedersen et al. (2010) to estimate fluxes from the measured concentration sequences. Dissolved concentrations of CO_2 and CH_4 were measured in the surface waters in three ponds with the acidified headspace technique (Valiente et al., 2022) at five occasions during the snow-free season, ranging between 40 and 520 μmol L^{-1} for CO_2, and between 1.1 and 26 μmol L^{-1} for CH_4. Pond fluxes were estimated from these dissolved gas concentrations following the methodology in Clayer et al. (2021), using the surface renewal gas exchange model by MacIntyre et al. (2010) for the gas transfer velocity, accounting for the small pond sizes (Vachon & Prairie, 2013) and the typically low wind speeds at Iškoras (Crusius & Wanninkhof, 2003). We conducted a survey of snowpack CH_4 concentrations on 14 March 2023, to estimate the magnitude and direction of wintertime CH_4 fluxes for all three surface types. Here, we used a portable CH_4 laser spectrometer (MIRA Strato, Aeris Technologies, USA) and estimated diffusive CH_4 emission from snowpack concentration gradients using the methodology described in Pirk et al. (2016) (see Figures S5 in the Supplement for examples of concentration time series as well as snowpack density and temperature profiles).

3 Results and discussion

3.1 Disaggregated fluxes of palsas, ponds, and fens

The estimated mean flux dynamics shown in Figure 3(a-f) indicate a clear separation of fluxes from the three surface types achieved by our BNN model. All surface types show a seasonal cycle of CO_2 and CH_4 flux dynamics. In summertime, after snow melt-out, CO_2 fluxes exhibit diurnal cycles as expected for northern latitude ecosystems. For CH_4, there is a relatively weak indication of such diurnal cycles for pond and fen surfaces, possibly due to the diurnal cycle of the ground temperature regulating CH_4 production rates.

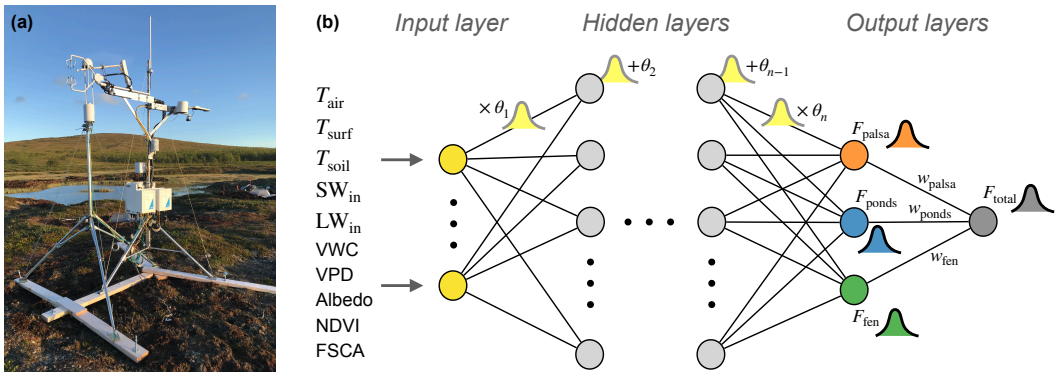

Figure 2. Flux estimation system. (a): The eddy covariance tower at the Iškoras permafrost peatland. (b): Conceptual architecture of our Bayesian neural network with uncertainty-aware parameters $\boldsymbol{\theta}$ to estimate fluxes F of three different surface types and their weighted average with weights $\mathbf{w}_s$ based on the footprint model.

The dry palsa areas feature CO_2 fluxes in a range between -2.4 and 1.4 μmol m^{-2} s^{-1}. Palsa CH_4 fluxes are relatively insignificant, but show persistent negative CH_4 fluxes (average of -1.8 nmol m^{-2} s^{-1}). This CH_4 uptake is consistent with our validation measurements with chambers in summertime (average of -0.95 nmol m^{-2} s^{-1}) and snowpack gas gradients in wintertime (average of -0.04 nmol m^{-2} s^{-1}) as shown in Figure S6 in the Supplement.

Thermokarst ponds are estimated to emit CO_2 at rates of up to 5.0 μmol m^{-2} s^{-1}, which is relatively high compared to fluxes from thermokarst ponds in a Canadian permafrost peatland (around 2.8 μmol m^{-2} s^{-1}) documented by Matveev et al. (2016). The magnitude and seasonal pattern of our BNN emission estimates are in very good agreement with our independent flux estimates from the dissolved gas measurements (Figure S6 in the Supplement). The observed diurnal cycle of the summertime CO_2 emission from the ponds could be due to periodic overturning of the water column or photo dissociation of dissolved organic carbon in the surface water. Parts of this carbon can originate from collapsing palsa edges releasing labile organic carbon (Patzner et al., 2022), as documented for our site in Figure 1c. Pond CH_4 fluxes reach up to 73 nmol m^{-2} s^{-1}, which is on the lower end of the range supported by our dissolved gas measurements, which may be attributed to differences between ponds as no water samples could be taken from the largest pond that dominates the pond flux signal in our EC measurements. The relative seasonal CH_4 flux patterns still agree well, also with the wintertime snowpack flux estimates (average of 1.4 nmol m^{-2} s^{-1}) as shown in Figure S6 in the Supplement. For reference, Matveev et al. (2016) report maximum diffusive CH_4 emissions from the aforementioned Canadian thermokarst ponds of around 120 nmol m^{-2} s^{-1}, which is notably higher than our maximum pond CH_4 flux.

Among the three surface classes, fens feature the largest maximum CO_2 uptake (7.0 μmol m^{-2} s^{-1}, i.e., three times higher uptake flux than the palsas) and the largest maximum CH_4 release (134 nmol m^{-2} s^{-1}, i.e., almost two times higher release flux than the ponds). The estimated CO_2 fluxes are consistent with our chamber fluxes in summertime (Figure S6 in the Supplement), and the large CH_4 fluxes are also corroborated by summertime chamber fluxes (average of 145 nmol m^{-2} s^{-1}) and snowpack fluxes in wintertime (average of 8.0 nmol m^{-2} s^{-1}).

The annual budgets of all these fluxes are relatively similar across the three years of our measurement campaign, which is in line with findings from other multi-year flux

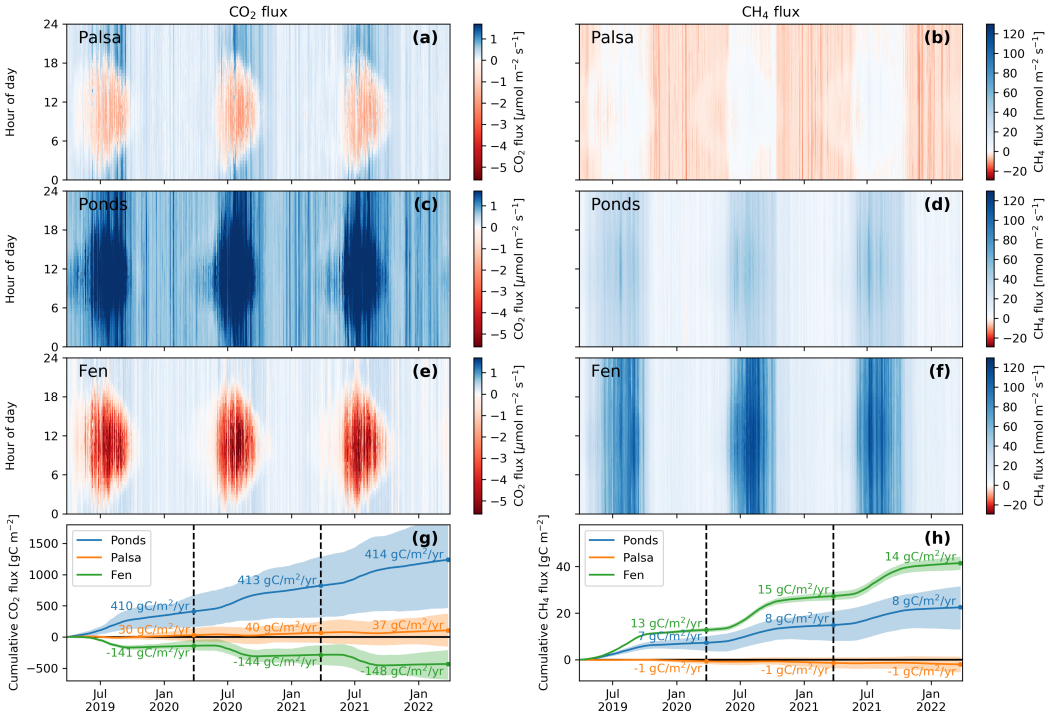

Figure 3. Flux dynamics and budgets. (a-f): Posterior mean CO_2 (left) and CH_4 (right) flux predictions for the three surface types as fingerprint plots using the same color scale. (g-h): Corresponding cumulative CO_2 and CH_4 fluxes with uncertainty bands representing the posterior ensemble's interquartile range. Stipulated vertical lines indicate one-year intervals for which the numbers indicate annual budgets.

studies in the sub-Arctic (e.g., Christensen et al., 2012). The cumulative carbon balance of each surface type is dominated by CO_2 fluxes (Figure 3(g-h)). Palsa surfaces are moderate carbon sources (35 gC m^{-2} yr^{-1} on average), while fen areas feature a strong carbon sink (−131 gC m^{-2} yr^{-1} on average, after accounting for CH_4 release). Ponds, while only a small area in the EC footprint and therefore most uncertain, are strong carbon emission hotspots, releasing on average 420 gC m^{-2} yr^{-1} to the atmosphere as CO_2 and CH_4 combined. For CH_4, the relative difference between surface types is larger than for CO_2, with fens emitting most (on average 14 gC m^{-2} yr^{-1}), followed by ponds (7.5 gC m^{-2} yr^{-1}), and a small CH_4 sink in palsa areas (−0.7 gC m^{-2} yr^{-1}). These CH_4 annual budgets are in general similar to those reported in a arctic-boreal synthesis compiled by Kuhn et al. (2021), which reports CH_4 emissions with an interquartile range between 4.5 and 29 gC m^{-2} yr^{-1} for fens, 3.3 and 52 gC m^{-2} yr^{-1} for diffusion and ebullition from small peatland ponds, and −0.3 to 0.5 gC m^{-2} yr^{-1} for dry tundra.

3.2 Climate feedbacks and geomorphological trajectory

Thermokarst ponds and lakes are currently estimated to cover about 7% of the permafrost region, and wetland thermokarst landscapes (including fens) cover a similar proportion of about 8% (Olefeldt et al., 2016). However, most areas of Fennoscandia and Western Siberia are projected to become climatically unsuitable to sustain permafrost peatlands over the next century, suggesting transformations in the landscapes that can leave an estimated 39 Gt of permafrost carbon (equivalent to twice the amount of carbon stored in European forests) vulnerable to post-thaw decomposition (Hugelius et al.,

2020; Fewster et al., 2022). Our CO_2 and CH_4 flux budgets allow for a direct assessment of the fate of these carbon stocks in permafrost peatlands, and their effect on the atmospheric energy budget through greenhouse gas forcing. To this end, we combine the mean annual CO_2 and CH_4 budgets of each surface type to CO_2-equivalent fluxes using a 100-year global warming potential for CH_4 of 27 (Forster et al., 2021). Thus, the three-year mean (interquartile range) CO_2-equivalent flux is estimated to be 106 $gCO_2\ m^{-2}\ yr^{-1}$ (-252 to 469 $gCO_2\ m^{-2}\ yr^{-1}$) for palsas, 1780 $gCO_2\ m^{-2}\ yr^{-1}$ (725 to 2834 $gCO_2\ m^{-2}\ yr^{-1}$) for ponds, and -31 $gCO_2\ m^{-2}\ yr^{-1}$ (-375 to 274 $gCO_2\ m^{-2}\ yr^{-1}$) for fens. The present day areal fraction of surface types contributing to the EC signal (52% palsa, 7% ponds, and 41% fen according to our EC footprint climatology shown in Figure 1d) suggests that the Iškoras site is currently a source of atmospheric carbon with a CO_2-equivalent flux of 167 $gCO_2\ m^{-2}\ yr^{-1}$. Using a space-for-time substitution, the ratios of the CO_2-equivalent fluxes indicate that palsa degradation to thermokarst ponds would lead to a 17 fold increase in the local greenhouse gas forcing, while palsa transformation into fens would reduce the local greenhouse gas forcing to slightly negative values.

Simulations by Aas et al. (2019) indicate that the degradation of permafrost peatlands in northern Norway is likely to accelerate in the next three to four decades. If and when palsa degradation creates ponds or fens will depend on the degradation rate and the amount of excess ice at the site (fast collapse at ice-rich sites likely favors pond formation). The geophysical interplay of processes causing lake formation and subsequent terrestrialization through drainage and infilling with fen vegetation is, however, still hard to predict. This complexity is exemplified by Nitze et al. (2018) who report both increasing and decreasing trends in limnicity (i.e., the areal fraction of ponds and lakes) for different parts of the permafrost region. Moreover, even an overall constant limnicity in a region can mask extensive lake drainage combined with new thermokarst lake formation (Sannel & Kuhry, 2011). Future studies combining even more Earth observations could help to refine our understanding of permafrost peatlands and upscale our benchmark of the greenhouse gas exchange beyond our EC footprint.

3.3 Bayesian deep learning for flux disaggregation

Using the BNN model for flux gap-filling and disaggregation yields considerable flexibility, generalization, and predictive accuracy, but comes at the cost of limited interpretability (Rudin, 2019), as the parameters of this black box model are not directly associated with any real-world process. At the same time, to the best of our knowledge, no interpretable mechanistic modeling approaches exist for EC flux disaggregation. In the interim, the BNN proposed herein serves as an uncertainty- and sparsity-aware data-driven approach that can help guide future method developments. For example, this flux disaggregation approach can be used to validate emerging drone data assimilation-based flux estimation methods (Pirk et al., 2022), guide land surface model developments (Aas et al., 2019), and incorporate uncertainty in flux gap filling approaches (Pirk et al., 2023).

Neural networks are in principle universal function approximators (Hornik et al., 1989), but we must assume that unobserved fluxes follow the same predictor relationship as the observed fluxes. The good generalization seen in our train-test split evaluation (Figure S4 in the Supplement) suggests that our CO_2 and CH_4 flux datasets are sufficiently representative. Interestingly, our BNN model for CH_4 flux has a low value of R^2 (around 0.3) for instantaneous CH_4 fluxes compared to daily average fluxes (R^2 around 0.7), which could be related to "unpredictable" ebullition events, or, more generally, to important but unobserved predictor variables. Compared to CO_2, CH_4 dynamics likely have a stronger dependency on processes occurring in the soil (Treat et al., 2015), where conditions change on small spatial scales and only few sensors were available, so that many of the control mechanisms are only indirectly captured through our surface and atmospheric sensors. A spatially distributed network of soil sensors could be employed to reduce the associated uncertainty.

Despite the partly opposing flux directions (release vs uptake) between the three surface types, the BNN flux disaggregation yields realistic flux magnitudes without clear indications of equifinality problems (e.g., large fluxes in opposing directions). Future work could explore other network architectures for flux data analysis, such as recurrent or convolutional networks, combined with marginal likelihood methods for hyperparameter and architecture optimization (Murphy, 2023). Another aspect with potential for improvement is the assumption of deterministic footprint weights $\mathbf{w}_s$ between the ultimate layers of the BNN. Here, future studies could use an ensemble of footprint models representing the uncertainty in the footprint input parameters using uncertain hyperparameters that are then inferred from the measurements together with the other network parameters. Such approaches could result in even better calibration of flux uncertainties.

4 Conclusions

Representative, high resolution, and uncertainty-aware flux estimates are invaluable to confidently assess land-atmosphere interactions in heterogeneous and dynamic ecosystems like permafrost peatlands. To achieve this goal, we developed an ensemble-based BNN model for EC flux disaggregation, which we compared against three other flux estimation methods. These independent flux estimates are compatible with our BNN results, but their large spatial variability also demonstrate the challenges to obtain landscape-scale flux measurements with manual sampling techniques.

Our BNN results indicate that while palsa areas have a near-zero annual CH_4 balance, the fens and ponds that form upon palsa degradation emit large amounts of CH_4. Fens compensate this greenhouse gas forcing with a strong annual CO_2 sink, while ponds are also strong—yet uncertain—CO_2 emission hotspots. Our flux results indicate that palsa degradation to thermokarst ponds would lead to a 17 fold increase in the local greenhouse gas forcing, while transformation into fens would reduce the local greenhouse gas forcing.

5 Open Research

Processed flux and ancillary data are archived and available at doi.org/10.5281/zenodo.7913027

Acknowledgments

We would like to thank our engineers John Hulth, Luc Girod, and Trond Eiken for their help with field measurements and drone imagery processing, as well as Jacqueline Knutson for water sample collection. The work was supported by the Research Council of Norway (projects #301552, #294948, #160016, and #323945). This work is a contribution to the strategic research initiative LATICE (#UiO/GEO103920), the Center for Biogeochemistry in the Anthropocene, as well as the Center for Computational and Data Science at the University of Oslo. The study contains modified Copernicus Sentinel data [Year 2019, 2020, 2021, 2022] obtained from the Google Earth Engine.

References

Aas, K. S., Martin, L., Nitzbon, J., Langer, M., Boike, J., Lee, H., ... Westermann, S. (2019). Thaw processes in ice-rich permafrost landscapes represented with laterally coupled tiles in a land surface model. *The Cryosphere*, *13*(2), 591–609. doi: 10.5194/tc-13-591-2019

Baldocchi, D. D. (2020). How eddy covariance flux measurements have contributed to our understanding of *Global Change Biology*. *Global Change Biology*, *26*(1), 242–260. doi: 10.1111/gcb.14807

Borge, A. F., Westermann, S., Solheim, I., & Etzelmüller, B. (2017). Strong degra-

dation of palsas and peat plateaus in northern Norway during the last 60 years. *The Cryosphere*, *11*(1), 1–16. doi: 10.5194/tc-11-1-2017

Christensen, T. R., Jackowicz-Korczyński, M., Aurela, M., Crill, P., Heliasz, M., Mastepanov, M., & Friborg, T. (2012). Monitoring the Multi-Year Carbon Balance of a Subarctic Palsa Mire with Micrometeorological Techniques. *AMBIO*, *41*(S3), 207–217. doi: 10.1007/s13280-012-0302-5

Clare, M. C. A., Sonnewald, M., Lguensat, R., Deshayes, J., & Balaji, V. (2022). Explainable Artificial Intelligence for Bayesian Neural Networks: Toward Trustworthy Predictions of Ocean Dynamics. *Journal of Advances in Modeling Earth Systems*, *14*(11). doi: 10.1029/2022MS003162

Clayer, F., Thrane, J.-E., Brandt, U., Dörsch, P., & Wit, H. A. (2021). Boreal Headwater Catchment as Hot Spot of Carbon Processing From Headwater to Fjord. *Journal of Geophysical Research: Biogeosciences*, *126*(12). doi: 10.1029/2021JG006359

Crusius, J., & Wanninkhof, R. (2003). Gas transfer velocities measured at low wind speed over a lake. *Limnology and Oceanography*, *48*(3), 1010–1017. doi: 10.4319/lo.2003.48.3.1010

Emerick, A. A., & Reynolds, A. C. (2013). Ensemble smoother with multiple data assimilation. *Computers & Geosciences*, *55*, 3–15. doi: 10.1016/j.cageo.2012.03.011

Evensen, G., Vossepoel, F. C., & van Leeuwen, P. J. (2022). *Data Assimilation Fundamentals: A Unified Formulation of the State and Parameter Estimation Problem*. Springer. doi: 10.1007/978-3-030-96709-3

Fewster, R. E., Morris, P. J., Ivanovic, R. F., Swindles, G. T., Peregon, A. M., & Smith, C. J. (2022). Imminent loss of climate space for permafrost peatlands in Europe and Western Siberia. *Nature Climate Change*, *12*(4), 373–379. doi: 10.1038/s41558-022-01296-7

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., ... Zhang, H. (2021, 10). Chapter 7: The Earth's energy budget, climate feedbacks, and climate sensitivity. doi: 10.25455/wgtn.16869671.v1

Frei, E. R., Schnell, L., Vitasse, Y., Wohlgemuth, T., & Moser, B. (2020). Assessing the effectiveness of in-situ active warming combined with open top chambers to study plant responses to climate change. *Frontiers in Plant Science*, *11*. doi: 10.3389/fpls.2020.539584

Ghahramani, Z. (2015). Probabilistic machine learning and artificial intelligence. *Nature*, *521*(7553), 452–459. doi: 10.1038/nature14541

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

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

Grosse, G., Jones, B., & Arp, C. (2013). 8.21 Thermokarst Lakes, Drainage, and Drained Basins. In *Treatise on Geomorphology* (pp. 325–353). Elsevier. doi: 10.1016/B978-0-12-374739-6.00216-5

Gu, L., Massman, W. J., Leuning, R., Pallardy, S. G., Meyers, T., Hanson, P. J., ... Yang, B. (2012). The fundamental equation of eddy covariance and its application in flux measurements. *Agricultural and Forest Meteorology*, *152*, 135–148. doi: 10.1016/j.agrformet.2011.09.014

Holmes, M. E., Crill, P. M., Burnett, W. C., McCalley, C. K., Wilson, R. M., Frolking, S., ... Chanton, J. P. (2022). Carbon Accumulation, Flux, and Fate in Stordalen Mire, a Permafrost Peatland in Transition. *Global Biogeochemical Cycles*, *36*(1). doi: 10.1029/2021GB007113

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359–366. doi: 10.1016/0893-6080(89)90020-8

Hugelius, G., Loisel, J., Chadburn, S., Jackson, R. B., Jones, M., MacDonald, G., ... Yu, Z. (2020). Large stocks of peatland carbon and nitrogen are vulnerable to permafrost thaw. *Proceedings of the National Academy of Sciences*, *117*(34), 20438–20446. doi: 10.1073/pnas.1916387117

Izmailov, P., Vikram, S., Hoffman, M. D., & Wilson, A. G. G. (2021). What are bayesian neural network posteriors really like? In M. Meila & T. Zhang (Eds.), *Proceedings of the 38th international conference on machine learning* (Vol. 139, pp. 4629–4640). PMLR.

Jiao, Y., Davie-Martin, C. L., Kramshøj, M., Christiansen, C. T., Lee, H., Althuizen, I. H., & Rinnan, R. (2023). Volatile organic compound release across a permafrost-affected peatland. *Geoderma*, *430*, 116355. doi: 10.1016/j.geoderma.2023.116355

Kljun, N., Calanca, P., Rotach, M. W., & Schmid, H. P. (2015). A simple two-dimensional parameterisation for Flux Footprint Prediction (FFP). *Geoscientific Model Development*, *8*(11), 3695–3713. doi: 10.5194/gmd-8-3695-2015

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In F. Pereira, C. Burges, L. Bottou, & K. Weinberger (Eds.), *Advances in neural information processing systems* (Vol. 25).

Kuhn, M. A., Varner, R. K., Bastviken, D., Crill, P., MacIntyre, S., Turetsky, M., ... Olefeldt, D. (2021). Bawld-ch_4: a comprehensive dataset of methane fluxes from boreal and arctic ecosystems. *Earth System Science Data*, *13*(11), 5151–5189. doi: 10.5194/essd-13-5151-2021

Kutzbach, L., Schneider, J., Sachs, T., Giebels, M., Nykänen, H., Shurpali, N. J., ... Wilmking, M. (2007). Co_2 flux determination by closed-chamber methods can be seriously biased by inappropriate application of linear regression. *Biogeosciences*, *4*(6), 1005–1025. doi: 10.5194/bg-4-1005-2007

Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. In I. Guyon et al. (Eds.), *Advances in neural information processing systems* (Vol. 30). Curran Associates, Inc.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., ... Zeng, X. (2019). The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245–4287. doi: 10.1029/2018MS001583

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. doi: 10.1038/nature14539

Levy, P., Drewer, J., Jammet, M., Leeson, S., Friborg, T., Skiba, U., & van Oijen, M. (2020). Inference of spatial heterogeneity in surface fluxes from eddy covariance data: A case study from a subarctic mire ecosystem. *Agricultural and Forest Meteorology*, *280*, 107783. doi: 10.1016/j.agrformet.2019.107783

Lopez-Gomez, I., Christopoulos, C., Langeland Ervik, H. L., Dunbar, O. R. A., Cohen, Y., & Schneider, T. (2022). Training Physics-Based Machine-Learning Parameterizations With Gradient-Free Ensemble Kalman Methods. *Journal of Advances in Modeling Earth Systems*, *14*(8). doi: 10.1029/2022MS003105

Luoto, M., & Seppälä, M. (2003). Thermokarst ponds as indicators of the former distribution of palsas in Finnish Lapland: Thermokarst Development from Palsas. *Permafrost and Periglacial Processes*, *14*(1), 19–27. doi: 10.1002/ppp.441

MacIntyre, S., Jonsson, A., Jansson, M., Aberg, J., Turney, D. E., & Miller, S. D. (2010). Buoyancy flux, turbulence, and the gas transfer coefficient in a strati-

fied lake. *Geophysical Research Letters*, *37*(24). doi: 10.1029/2010GL044164
MacKay, D. J. C. (2003). *Information theory, inference, and learning algorithms*. Cambridge University Press.
Martin, L. C. P., Nitzbon, J., Scheer, J., Aas, K. S., Eiken, T., Langer, M., ... Westermann, S. (2021). Lateral thermokarst patterns in permafrost peat plateaus in northern norway. *The Cryosphere*, *15*(7), 3423–3442. doi: 10.5194/tc-15-3423-2021
Matveev, A., Laurion, I., Deshpande, B. N., Bhiry, N., & Vincent, W. F. (2016). High methane emissions from thermokarst lakes in subarctic peatlands: Methane Emissions from Peatland Thermokarst Lakes. *Limnology and Oceanography*, *61*(S1), S150-S164. doi: 10.1002/lno.10311
Murphy, K. P. (2022). *Probabilistic machine learning: An introduction*. MIT Press.
Murphy, K. P. (2023). *Probabilistic machine learning: Advanced topics*. MIT Press.
Neal, R. M. (1996). *Bayesian Learning for Neural Networks*. Springer. doi: 10.1007/978-1-4612-0745-0
Nitze, I., Grosse, G., Jones, B. M., Romanovsky, V. E., & Boike, J. (2018). Remote sensing quantifies widespread abundance of permafrost region disturbances across the Arctic and Subarctic. *Nature Communications*, *9*(1), 5423. doi: 10.1038/s41467-018-07663-3
Nykänen, H., Heikkinen, J. E. P., Pirinen, L., Tiilikainen, K., & Martikainen, P. J. (2003). Annual CO $_2$ exchange and CH $_4$ fluxes on a subarctic palsa mire during climatically different years. *Global Biogeochemical Cycles*, *17*(1). doi: 10.1029/2002GB001861
Oechel, W. C., Vourlitis, G. L., Hastings, S. J., Zulueta, R. C., Hinzman, L., & Kane, D. (2000). Acclimation of ecosystem CO2 exchange in the Alaskan Arctic in response to decadal climate warming. *Nature*, *406*(6799), 978–981. doi: 10.1038/35023137
Olchev, A., Zyrianov, V., Panov, A., Satosina, E., Mukhartova, I., Novenko, E., & Prokushkin, A. (2022). Seasonal Variability of Carbon Dioxide and Methane Fluxes in a Subarctic Palsa Mire in North-Central Siberia. In *ECAS 2022* (p. 52). doi: 10.3390/ecas2022-12837
Olefeldt, D., Goswami, S., Grosse, G., Hayes, D., Hugelius, G., Kuhry, P., ... Turetsky, M. R. (2016). Circumpolar distribution and carbon storage of thermokarst landscapes. *Nature Communications*, *7*(1), 13043. doi: 10.1038/ncomms13043
Patzner, M. S., Logan, M., McKenna, A. M., Young, R. B., Zhou, Z., Joss, H., ... Bryce, C. (2022). Microbial iron cycling during palsa hillslope collapse promotes greenhouse gas emissions before complete permafrost thaw. *Communications Earth & Environment*, *3*(1), 76. doi: 10.1038/s43247-022-00407-8
Pedersen, A. R., Petersen, S. O., & Schelde, K. (2010). A comprehensive approach to soil-atmosphere trace-gas flux estimation with static chambers. *European Journal of Soil Science*, *61*(6), 888–902. doi: 10.1111/j.1365-2389.2010.01291.x
Pirk, N. (2023). *Resources for "Disaggregating the carbon exchange of degrading permafrost peatlands using Bayesian deep learning" [Dataset]*. Zenodo. doi: 10.5281/zenodo.7913027
Pirk, N., Aalstad, K., Westermann, S., Vatne, A., van Hove, A., Tallaksen, L. M., ... Katul, G. (2022). Inferring surface energy fluxes using drone data assimilation in large eddy simulations. *Atmospheric Measurement Techniques*, *15*(24), 7293–7314. doi: 10.5194/amt-15-7293-2022
Pirk, N., Aalstad, K., Yilmaz, Y. A., Vatne, A., Popp, A. L., Horvath, P., ... Tallaksen, L. M. (2023). *Snow-vegetation-atmosphere interactions in alpine tundra* (Preprint). Biogeochemistry: Air - Land Exchange. doi: 10.5194/bg-2023-21
Pirk, N., Tamstorf, M. P., Lund, M., Mastepanov, M., Pedersen, S. H., Mylius, M. R., ... Christensen, T. R. (2016). Snowpack fluxes of methane and carbon

dioxide from high Arctic tundra. *Journal of Geophysical Research: Biogeosciences*, *121*(11), 2886–2900. doi: 10.1002/2016JG003486

Qiu, C., Zhu, D., Ciais, P., Guenet, B., Krinner, G., Peng, S., ... Ziemblinska, K. (2018). ORCHIDEE-PEAT (revision 4596), a model for northern peatland CO_2, water, and energy fluxes on daily to annual scales. *Geoscientific Model Development*, *11*(2), 497–519. doi: 10.5194/gmd-11-497-2018

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195–204. doi: 10.1038/s41586-019-0912-1

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206–215. doi: 10.1038/s42256-019-0048-x

Sannel, A. B. K., & Kuhry, P. (2011). Warming-induced destabilization of peat plateau/thermokarst lake complexes. *Journal of Geophysical Research*, *116*(G3), G03035. doi: 10.1029/2010JG001635

Schuur, E. A. G., McGuire, A. D., Schädel, C., Grosse, G., Harden, J. W., Hayes, D. J., ... Vonk, J. E. (2015). Climate change and the permafrost carbon feedback. *Nature*, *520*(7546), 171–179. doi: 10.1038/nature14338

Serikova, S., Pokrovsky, O. S., Laudon, H., Krickov, I. V., Lim, A. G., Manasypov, R. M., & Karlsson, J. (2019). High carbon emissions from thermokarst lakes of Western Siberia. *Nature Communications*, *10*(1), 1552. doi: 10.1038/s41467-019-09592-1

Treat, C. C., Natali, S. M., Ernakovich, J., Iversen, C. M., Lupascu, M., McGuire, A. D., ... Waldrop, M. P. (2015). A pan-arctic synthesis of ch4 and co2 production from anoxic soil incubations. *Global Change Biology*, *21*(7), 2787-2803. doi: https://doi.org/10.1111/gcb.12875

Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *203*(1153), 405–426. doi: 10.1098/rspb.1979.0006

Vachon, D., & Prairie, Y. T. (2013). The ecosystem size and shape dependence of gas transfer velocity versus wind speed relationships in lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, *70*(12), 1757–1764. doi: 10.1139/cjfas-2013-0241

Valiente, N., Eiler, A., Allesson, L., Andersen, T., Clayer, F., Crapart, C., ... Hessen, D. O. (2022). Catchment properties as predictors of greenhouse gas concentrations across a gradient of boreal lakes. *Frontiers in Environmental Science*, *10*, 880619. doi: 10.3389/fenvs.2022.880619

Voigt, C., Marushchak, M. E., Mastepanov, M., Lamprecht, R. E., Christensen, T. R., Dorodnikov, M., ... Biasi, C. (2019). Ecosystem carbon response of an Arctic peatland to simulated permafrost thaw. *Global Change Biology*, *25*(5), 1746–1764. doi: 10.1111/gcb.14574

Walter, K. M., Zimov, S. A., Chanton, J. P., Verbyla, D., & Chapin, F. S. (2006). Methane bubbling from Siberian thaw lakes as a positive feedback to climate warming. *Nature*, *443*(7107), 71–75. doi: 10.1038/nature05040

Wilson, A. G., & Izmailov, P. (2020). Bayesian deep learning and a probabilistic perspective of generalization. In H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, & H. Lin (Eds.), *Advances in neural information processing systems* (Vol. 33, pp. 4697–4708). Curran Associates, Inc.