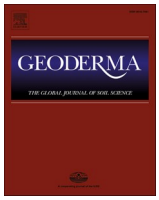
Geoderma 438 (2023) 116652

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma



High resolution mapping shows differences in soil carbon and nitrogen stocks in areas of varying landscape history in Canadian lowland tundra

Julia Wagner [a,b,*], Victoria Martin [c], Niek J. Speetjens [d], Willeke A'Campo [a], Luca Durstewitz [a], Rachele Lodi [e], Michael Fritz [f], George Tanski [d,f], Jorien E. Vonk [d], Andreas Richter [c], Annett Bartsch [g], Hugues Lantuit [f], Gustaf Hugelius [a,b]

[a] Department of Physical Geography, Stockholm University, Stockholm, Sweden
[b] Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden
[c] University of Vienna Centre for Microbiology and Environmental System Science, Vienna, Austria
[d] Vrije Universiteit Amsterdam, Department of Earth Sciences, Amsterdam, Netherlands
[e] Ca' Foscari University, Department of Environmental Sciences, Informatics and Statistics, and CNR Polar Sciences Institute, Venice, Italy
[f] Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Potsdam, Germany
[g] b.geos GmbH, Korneuburg, Austria

ARTICLE INFO

Handling Editor: A. Agnelli

Keywords:
Random forest
Machine learning
Soil organic carbon
Tundra
Permafrost

ABSTRACT

Soil organic carbon (SOC) in Arctic coastal polygonal tundra is vulnerable to climate change, especially in soils with occurrence of large amounts of ground ice. Pan-arctic studies of mapping SOC exist, yet they fail to describe the high spatial variability of SOC storage in permafrost landscapes. An important factor is the landscape history which determines landform development and consequently the spatial variability of SOC. Our aim was to map SOC stocks, and which environmental variables that determine SOC, in two adjacent coastal areas along Canadian Beaufort Sea coast with different glacial history. We used the machine learning technique random forest and environmental variables to map the spatial distribution of SOC stocks down to 1 m depth at a spatial resolution of 2 m for depth increments of 0–5, 5–15, 15–30, 30–60 and 60–100 cm.

The results show that the two study areas had large differences in SOC stocks in the depth 60–100 cm due to high amounts of ground ice in one of the study areas. There are also differences in variable importance of the explanatory variables between the two areas. The area low in ground ice content had with 66.6 kg C/m^{-2} more stored SOC than the area rich in ground ice content with 40.0 kg C/m^{-2}. However, this SOC stock could be potentially more vulnerable to climate change if ground ice melts and the ground subsides. The average N stock of the area low in ground ice is 3.77 kg m^{-2} and of the area rich in ground ice is 3.83 kg m^{-2}.

These findings support that there is a strong correlation between ground ice and SOC, with less SOC in ice-rich layers on a small scale. In addition to small scale studies of SOC mapping, detailed maps of ground ice content and distribution are needed for a validation of large-scale quantifications of SOC stocks and transferability of models.

1. Introduction

Approximately 22% of the Earth's surface is defined as permafrost regions (Obu, 2021). These areas store large amounts of soil organic carbon (SOC), approximately twice as much carbon as in the atmosphere (Strauss et al., 2021). The upper three meters of permafrost region soils store 1035 ± 150 Pg of soil organic carbon (Hugelius et al., 2014). Air temperatures in high latitudes have increased more than double as the global average (Meredith et al., 2022). This warming of global air temperatures causes permafrost thaw at a global scale (Biskaborn et al., 2019) which leads to SOC stocks stored in the soil becoming potentially available for decomposition and emission of greenhouse gases, which may shift permafrost ecosystems from carbon sinks to sources (Schuur et al., 2015). An Increase in ground temperatures has been observed in the past 20 years, in particular at arctic coasts (Miner et al., 2022).

The degree of thaw and the subsequent greenhouse gas emissions

* Corresponding author at: Stockholm University Department of Physical Geography, Svante Arrhenius väg 8, SE-106 91 Stockholm, Sweden.
E-mail address: julia.wagner@natgeo.su.se (J. Wagner).

https://doi.org/10.1016/j.geoderma.2023.116652
Received 1 August 2022; Received in revised form 21 August 2023; Accepted 23 August 2023
Available online 4 September 2023

depend on different complex factors which are interlinked. Warming leads to the more frequent occurrence of abrupt thaw events, especially in areas with large amounts of ground ice. These include ice wedge polygon degradation and development of thermokarst lakes or wetlands which may accelerate the release of CO2 or methane (Turetsky et al., 2020).Together with soil properties and geomorphological shifts, vegetation dynamics are key controls of carbon cycling and nitrogen cycling. Therefore, high resolution maps of N stocks are needed as a complement to SOC stocks. Vegetation changes in arctic tundra are complex (Heijmans et al., 2022; Van Der Kolk et al., 2016). In contrast to an anticipated broadscale trend of arctic greening (Arndt et al., 2019; Berner et al., 2020), Arctic browning (Myers-Smith et al., 2020) can be observed on smaller spatial and temporal scales, especially in ice-rich lowland tundra (Heijmans et al., 2022) caused by the occurrence of fires and abrupt thaw events (Phoenix and Bjerke, 2016). The soil–plant feedback under climate change remains complex, in particular with regard to interactions between roots (Blume-Werry et al., 2018), soil and microbially mediated processes (Ernakovich et al., 2022). Increased nutrient availability is thought to be linked to warming, because enhanced microbial activity accelerates the rate of soil organic matter decomposition (Wild et al., 2016; Wilson et al., 2022).

The long-term landscape development determines the distribution and characteristics of periglacial landforms (French, 2007). In our study area, the Canadian Beaufort coast, glacial history played a major role in shaping the landscape and soil formation. The area was influenced by the Laurentide ice sheet during the Last Glacial Maximum (LGM) at 16.2 BP (Fritz et al., 2012). Its limits were located west of Herschel Island (Fig. 1). Ice wedge polygon development, ground ice formation and degradation differed during the late Pleistocene and early Holocene between areas inside and outside of the glacial limit (Fritz et al., 2012; 2016). Ice wedge polygons can be divided into two main categories with low-centered polygons being the non-degraded form consisting of a waterlogged or -filled center and elevated rims and high-centered being the degraded form of ice wedge polygons consisting of an elevated center and an often water-filled trough around it (French, 2007; Liljedahl et al., 2016).

Maps of SOC storage exist for the entire Arctic, yet often fail to capture the high spatial variability of permafrost landscapes and soils (Hugelius et al., 2014; Mishra et al., 2021). Different processes dominate the distribution of soil properties from the pedon scale, via landfrom scale to landscape scale. These range from the influence of cryoturbation on scales from centimeters to meters, ice wedge dynamics on l scales of meters to tens of meters and landscape scale changes on scales of several hundred meters (Siewert et al. 2021). Because multiple processes at different scales are involved, and because of very high spatial heterogeneity, it is challenging to quantify the potential release of carbon from permafrost thaw. This is especially the case for abrupt thaw features related to thermokarst such as retrogressive thaw slumps and thermokarst lake expansion (Turetsky et al., 2020). Digital soil mapping generates model based maps for soil properties by finding relationships and spatial patterns between the soil forming factors and the soil properties (Lagacherie, 2008). Machine learning techniques, such as random forest classifiers are often used to perform these analyses (Heung et al., 2016; McBratney et al., 2003; Wadoux et al., 2020a).

Tobleŕs first law of Geography, “everything is related to everything else, but near things are more related than distant things” (Tobler, 1970) is the theoretical basis for various digital soil mapping methods, especially variogram-based mapping such as Kriging. Siewert et al. (2021) provide evidence that this is not applicable in complex tundra environments. There are rapid, sometimes repetitive, changes of soil properties over very short distances due to patterned ground and associated variations in e.g. ground ice content, organic layer thickness and the occurrence of cryoturbation. Machine learning techniques that find relationships between environmental factors and the target variable without consideration of spatial structures or autocorrelation have the potential to overcome many of the issues associated with SOC mapping in tundra environments. It has been shown that they can resolve small-scale soil variations, which cannot be mapped with variogram-based methods.

Many soil mapping approaches on local (Siewert, 2018) as well as

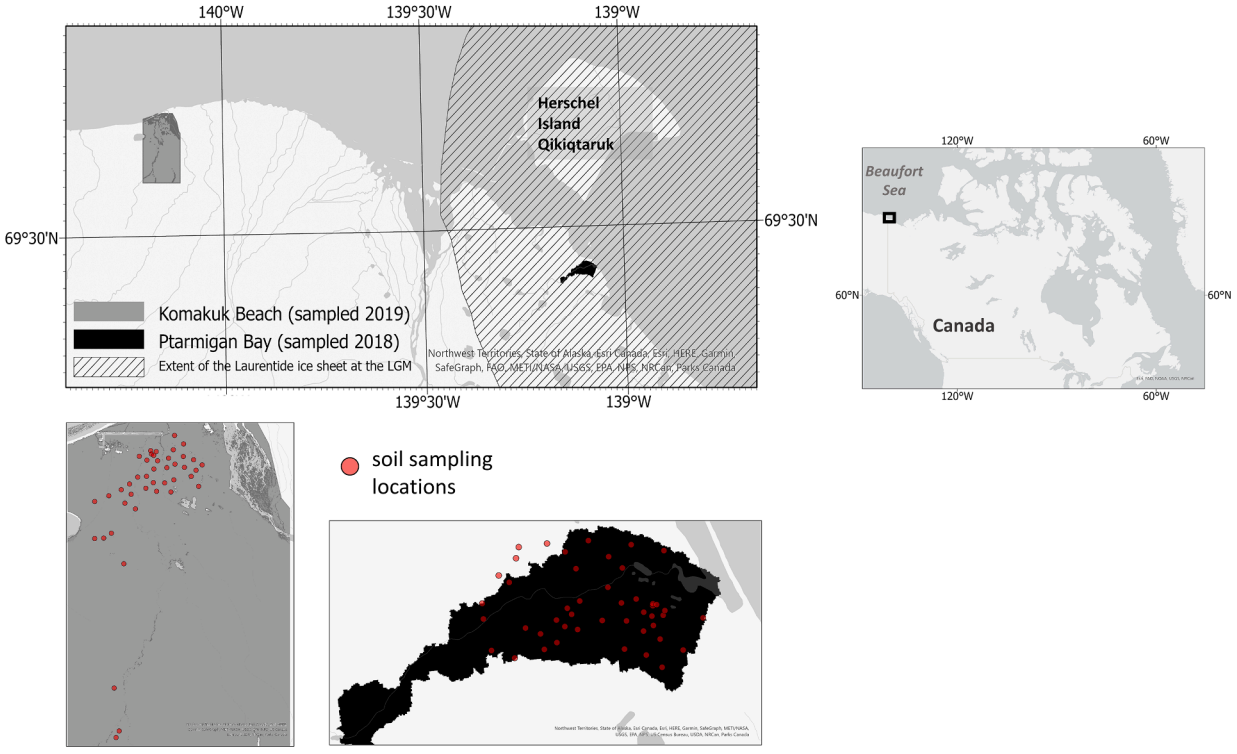


Fig. 1. Locations and overview of the two study areas sampled during the two field campaigns in 2018 and 2019 (coordinate system of the map on the left: WGS 84 UTM Zone 7 N, Basemap: Esri, 2023).

regional (Hanes et al., 2022; Ramcharan et al., 2018) and global scale (Poggio et al., 2021)- are based on machine learning. Studies have demonstrated that in arctic tundra environments random forest models have the best performance (Siewert, 2018; Siewert et al., 2021).

In our study we sampled two small study areas, representative of lowland coastal tundra ecosystems along the Canadian Beaufort coast, to investigate

(1) the distribution of SOC and N stocks within the first meter of soil,
(2) which environmental parameters control SOC and N stocks distribution and
(3) to assess whether contrasting glacial histories of the two study areas left an imprint on the prevailing SOC stocks.

To answer these questions, we used environmental variables ranging from a spatial resolution of 1.24 m up until 20 m resampled to a resolution of 2 m and applied a random forest (Breiman, 2001) model and further assessed the uncertainty of the model by using the quantile regression forest method (Vaysse and Lagacherie, 2017) following the workflow of Yigini et al (2018).

2. Material and methods

2.1. Study area and background

The study areas Ptarmigan Bay and Komakuk Beach are located on the Canadian Beaufort Sea coast which is the terrestrial extension of the Canadian Beaufort shelf. The plain stretches ca. 300 km long from the Alaskan border in the west until the Mackenzie River delta in the east. The study area in the east, Ptarmigan Bay, lies within the limits of the Laurentide ice sheet during the LGM, while the study area in the west, Komakuk Beach, lies outside those ice sheet limits. The study areas were affected by periglacial processes over long timescales in the late Quaternary and are composed of ice wedge polygonal lowland tundra (Rampton, 1982). The Ptarmigan Bay area is defined by lacustrine, fluvial and morainal deposits (Rampton, 1982). The Komakuk Beach area lies between two alluvial fans. The surficial geology is defined by lacustrine and fluvial sediments in the lowland area and colluvial sediments in the upland area (Rampton, 1982).

Both study areas have a tundra climate. The mean annual temperatures (1972–2000) at the nearby climate stations Shingle Point and Komakuk Beach are −9.9 °C and −11.0 °C and the mean annual precipitation is 254 and 161 mm, respectively. The average summer temperatures during June, July and August are 8.6 °C (±1.7 °C) at Shingle point and 6.0 °C (±1.6 °C) Komakuk Beach (Government of Canada, 2022).

Reported average active layer depths in the literature are 30—40 cm on Herschel Island – Qikiqtaruk (Siewert et al 2021), and generally < 50 cm along the Yukon mainland coast (Wolter et al., 2018). This is in agreement with our measurements of on average 33 cm at Ptarmigan Bay in 2018 and 35 cm at Komakuk Beach in 2019. Average amounts of volumetric ground ice in the area are 46%, but also up to 74% in some areas (Couture et al., 2018; Couture and Pollard, 2017). The regional vegetation is defined as low shrub tundra, subzone E in the Circum Arctic Vegetation Map (Walker et al., 2005). The species and plant functional type composition of the vegetation is strongly linked to micro topography and influenced by ice wedge degradation. Elevated areas such as high-center polygon centers and low-center polygons rims accommodate dwarf and low shrubs such as *Betula nana* or *Salix* sp. Whereas lower areas such as low-center polygon centers are dominated by gramindoids, especially sedges (*Carex* sp.) (Wolter et al., 2016).

The landscape typically consists of low-centered and high-centered polygons. The former being the non-degraded form composed of an elevated circular rim around a lowered often water-filled center. The latter represents the degraded form characterized by a high drained center surrounded by often water-filled troughs around them (French, 2007; Liljedahl et al., 2016). The transformation of low-centered polygons into high-centered polygons is characterized by the following sequential processes: (i) the melt of ground ice, (ii) the formation of wet troughs (iii) the increased drainage of polygon centers and the thermal erosion which occurs along ice wedges, and (iv) the redistribution from rim material towards the center, as ice wedges grow wider in size (Wolter et al., 2018). The degradation of ice wedge polygons is mostly initiated or accelerated by geomorphic disturbances (Wolter et al., 2018). When waterlogged conditions are sustained, ice wedge polygons remain stable, however, the current climate change may induce geomorphic changes accelerating the degradation process (Liljedahl et al., 2016; Wolter et al., 2018).

2.2. Field campaign and soil data

Soil sampling was carried out in the two coastal areas using a stratified random sampling design based on catchment area and Quaternary geology (Rampton, 1982) at the Canadian Beaufort coast. The sampling sites were registered in the field with a handheld GPS. During a field campaign in August 2018, 40 sites were sampled in Ptarmigan Bay, followed by a field campaign during August 2019 where 46 sites were sampled in Komakuk Beach. The sampling in both years was carried out at the end of the summer near the time of maximum annual thaw of the active layer depth. We sampled the sites using a combination of sampling by soil horizon from open soil pits in the active layer and at fixed depth interval with a permafrost corer (Palmtag et al., 2022). At each site a one-meter-wide soil pit was dug until the permafrost table was reached. The samples of the active layer were taken using either fixed-volume cylinders or, where soil density was low (usually in the organic horizons) by cutting blocks and measuring their dimensions. At sites that were too wet to be sampled from a soil pit, the upper soil was sampled from blocks that were cut out from the top. The frozen part of the profile was cored by hammering a steel pipe (outer diameter: 4.2 cm) in 10 cm depth increments into the ground. Whenever possible, all sites were sampled to a 100 cm depth. Sampling was stopped before that depth if a massive ice wedge was reached. Additional information about each sampling site, such as description about visible ice content, vegetation composition, amount of roots and landforms were noted in the field. Further, photos of the soil pit, the site itself from the top, and the surroundings of the locations were taken with a digital camera.

2.3. Soil chemical analysis

Samples were oven-dried at 65 °C for at least 72 h. Bulk density (g cm^{-3}) was calculated by dividing the dry weight (DW, g)) by the volume of the sample (VOL, cm^3).

(1) DBD [g cm^{-3}] = DW/VOL.

All samples were analyzed for C % and N % content at the Centre for Microbiology and Environmental Systems Science, University of Vienna via EA1110 (CE instrument, Italy) coupled to a continuous-flow isotopic ratio mass spectrometer. Isotopic signatures ($\delta^{13}C$, ‰) were assessed by relating the abundance of ^{13}C to ^{12}C in reference to Vienna Pee Dee Belemnite (VPDB) standard (IRMS, DeltaPlus, Finnigan MAT). The loss on ignition method (Heiri et al., 2001) was performed on the samples after drying them at 105 °C. In a first step they were burnt at 550 °C for four hours to obtain soil organic matter (SOM) content. In a next step they were burnt at 900 °C for two hours to obtain inorganic carbon content. The latter was negligibly low (below 0.5 %), so we assume the total carbon content measured equals the SOC content.

2.4. Pre-processing of soil data

Due to frost heave processes, soil horizons in permafrost soils often vary horizontally (Siewert et al., 2016). Therefore, the depth of the

horizons were calculated using the notes from the field and by digitizing the soil horizon extent onto perspective corrected pictures of the soil pits (Supplementary material Fig. S1). The SOC stocks (kg C/m^{-2}) (2) and total nitrogen stocks (3) (kg N/m^{-2}) were calculated for each sampling depth using the fraction of organic carbon (% OC)/ total nitrogen (% TN), dry bulk density (DBD, g cm^{-3}), coarse fragments > 2 mm (CF, weight %) and thickness of the depth (T, cm) multiplied by 10 for unit conversion (Palmtag et al., 2018; Siewert et al., 2015):

(2) SOC [kg m^{-2}] = (DBD * %OC *((1-CF) *T)*10.
(3) TN [kg m^{-2}] = (DBD * %TN *((1-CF)*T)*10.

In some cases, sufficient sample material was not available to perform laboratory analyses for some depth increments. For Komakuk Beach approx. 12 % and for Ptarmigan Bay 25 % of the data is based on values that were gap-filled from qualitative observations (field notes) and interpolation from adjacent depth increments. Depending on the position of the missing sample within the soil profile, rules were set up to estimate missing values. If samples above and below were available, the mean of those samples was used. In some cases, the same values as the sample below or above was used or the mean of similar soil horizons from other soil profiles. The soil data was then recalculated to the depth increments that are also present in the soil grids dataset (Poggio et al., 2021) (0–5, 5–15, 15–30, 30–60 and 60–100 cm) using the same approach as in Siewert et al. (2016). The sampling depths were split into artificial depths of 1 mm (Wickham, 2011) and aggregated to the needed depth increments. All data calculations were performed in R statistical software (R Development Core Team, 2021).

A total of 43 sites were used for the mapping of the Komakuk Beach area, three sites which were outside the mapping extent were excluded. For Ptarmigan Bay 40 sites plus six additional sites were used. The six additional sites were not part of the stratified random sampling design and were subjectively chosen based on being representative landforms during the same field campaign.

2.5. Covariates

Environmental covariates (Table 1) for the digital soil mapping were selected in accordance with the SCORPAN modelling framework (McBratney et al., 2003). The concept of covariates representing the soil forming factors is based on the traditional CLORPT model (Jenny, 1994) defined by the factors climate (CL), organisms (O), relief (R), parent material (P) and time (T). This model was updated for digital soil mapping and includes the factors soil (S), climate (C), organisms (O), relief (R), parent material (P), age (A) and the spatial location (N), also known as SCORPAN-model (McBratney et al., 2003). High resolution 8-band and panchromatic worldview-3 images (for Ptarmigan Bay: DigitalGlobe 18 July 2018 and for Komakuk Beach: DigitalGlobe 12 July 2019) were acquired for each study area at the timing of the respective field season to show the conditions exactly at the time of the sampling. From these images 17 spectral indices were derived using the package RStoolbox (Leutner et al., 2022) in R from which the NDVI and GEMI indices were used for the final mapping due to a high level of intercorrelation between the indices. Here we used the function findCorrelation from the package caret (Kuhn, 2022) and a cutoff value of > 0.8, indicating a high correlation of the variables. Landcover information was used from a Landcover dataset based on Sentinel 1 and 2 data from Bartsch et al. (2019a) which is an indicator for different soil properties in tundra environments (Bartsch et al., 2019b).

Studies have shown that C-band synthetic aperture radar can be used as an indicator for carbon stocks estimation and prediction across the arctic tundra landscape (Bartsch et al., 2016). Therefore, a dataset that contains the C-HH (Widhalm et al., 2019) and C-VV winter backscatter from Sentinel-1 preprocessed following the schemes detailed in Widhalm et al. (2018) and Bartsch et al. (2020) were used as covariates. In further explanations the variables S1 C-HH and S1 C-VV (Table. 1) are referred to as radar backscatter.

Another covariate was the elevation derived from the Arctic DEM at 2-meter resolution (Porter et al., 2018). The ArcticDEM is a digital surface model obtained from high resolution satellite imagery (Worldview-1,2,3 and GeoEye-1). Although covariates derived from a DEM are commonly used in digital soil mapping (e.g. Lamichhane et al., 2019; McBratney et al., 2003) we did not further use terrain parameters such as slope, aspect or curvature due to the very flat terrain of the study areas and a low accuracy of the DEM over the area. All covariates were resampled to a resolution of 2 m to account for the uncertainty in GPS measurements, but to still create high resolution maps.

2.6. Predictive model

We trained a Random Forest algorithm for each depth increment using the package randomForest implemented in the package caret which offers functions for data splitting, pre-processing, feature selection, model tuning and variable importance estimation (Kuhn, 2022). Random Forest is a technique that builds an ensemble of several decision trees (Breiman, 2001). To reduce the prediction error, random forests are based on the concepts of bagging (bootstrap and aggregating). Based on bootstrap samples of the training data a large amount of trees is built which is later averaged by aggregating (Breiman, 2001, 1996; Wadoux et al., 2019). Three parameters can be modified in a random forest model. The first parameter is ntree which is the amount of computed trees and which we set to 500. The second is mtry which is the amount of covariates randomly selected at each split. Here we tested several numbers of mtry by using grid search and the model with the lowest RMSE was chosen for prediction. The last parameter is the nodesize, which defines the minimum amount of training data to continue splitting of the tree. This was kept at its default value of 5. The variable importance was calculated using the varImp function of the caret

Table 1
Description of the raster layers used as covariates within the SCORPAN digital soil mapping framework (McBratney et al., 2003). All layers were resampled to a resolution of 2 m. The table shows only the covariates that were used for the models. Further spectral indices derived from the Worldview-3 images and individual bands were removed due to intercorrelation (r > 0.8).

SCORPAN Category	Variable	Source	Data source	Resolution	Acronym
Soil	Blue	this study	Worldview-3 image	1.24 m	B
Organisms	Near Infraread 2	this study	Worldview-3 image	1.24 m	NIR2
Organisms	Global Environmental Monitoring Index	(Pinty and Verstraete, 1992)	Worldview-3 image	1.24 m	GEMI
Organisms, Soil	Normalised Difference Vegetation Index	(Rouse et al., 1974)	Worldview-3 image	1.24 m	NDVI
Organisms, Soil, Parent material, Age	Landcover data	(Bartsch et al., 2019a)	Landcover derived from Sentinel 1 and 2 data	20 m	LC
Relief (terrain), Soil, Organisms, Climate	Normalized C-VV-winter backscatter	this study	Sentinel 1	10 m	S1 C-VV
Relief (terrain), Soil, Organisms, Climate	Normalized C-HH winter backscatter	(Widhalm et al., 2019)	Sentinel 1	20 m	S1 C-HH
Relief (terrain)	Elevation	(Porter et al., 2018)	ArcticDEM	2 m	DEM

package. The function findCorrelation of the caret package was used to remove continuous variables that are highly intercorrelated ($r > 0.8$) to reduce redundancy in the model. The chosen covariates to be the least intercorrelated were: B, NIR2, GEMI, NDVI, LC, S1 C-VV, S1 C-HH, elevation (Table 1). The landcover data was not included in this process and all landcover data was added to the predictor variables. Each landcover class was converted into a binary layer [0,1 – presence/absence] to handle uneven distribution of classes across the study areas. The classes of interest that occur in the study areas are as follows: 1: Sparse vegetation (without shrubs), mostly sandy soil; flood plains, recent landslides, also within fire scars, 2: Dry cryptogamic-crust or sparse vegetation, 3: Graminoid, prostrate dwarf shrub, patterned ground, partially bare, 4: Dry to moist prostrate to erect dwarf shrub tundra, 5: Moist to wet graminoid prostrate to erect dwarf shrub tundra, 13: Disturbed, including forest fire scars, seasonally inundated areas and landslide scars. They are referred to as LC_PB1-13 and LC_KB1-13 in the following.

The random forest model was internally validated using leave-one-out crossvalidation (LOOCV) due to the limited amount of sampling sites. The final prediction was computed for the catchment area for Ptarmigan Bay. For Komakuk Beach the actual predicted area is a selected area in the lowland.

In addition to the validation of the model itself, the uncertainty of each predicted value was assessed using the quantile regression forest (QRF) method following the workflow of Yigini et al. (2018). The resulting datasets provide an uncertainty measure at the pixel – level that depicts the range of uncertainty across the whole study areas. This method combines an estimation of the model uncertainty and the uncertainty caused by the variations of the data. Use of QRF is recommended when only a limited amount of soil data is available and no independent validation dataset is available (Vaysse and Lagacherie, 2017). Both of these conditions apply to this study. In the first step of QRF the sensitivity of the model to the data was tested by running several repetitions of the algorithm. We did 20 repetitions based on 70% of the data and used 30% for testing. The sensitivity of the model to available data was created based on the standard deviation of all 20 predictions. In the next step quantile regression forest was used to assess the uncertainty of the model caused by variations of the data by generating a probability distribution of the values that each pixel can take. The total uncertainty is the combined uncertainty. Yigini et al. (2018) also mention that further errors are introduced by uncertainties in the soil data and uncertainties in the covariates which were not further estimated in this work.

In a final step we performed an additional analysis for the SOC stock data by adding the accumulated SOC stock predictions from the upper depth as predictor variable to the model for the respective lower depth interval. For simplification in following explanations this variable is called “accumulated SOC stocks”. The SOC stocks were mapped in kg C/m^{-2}, but the final results were converted to kg C/m^{-3} for a better comparison due to the different thicknesses of the depth intervals.

3. Results

3.1. Model performance

The results of the crossvalidation show stable model results, with all RMSE values being lower than the standard-deviation. The model for the study area Ptarmigan bay depth interval 0 – 5 cm for example had an RMSE value of 0.67, while the standard-deviation of the sampling values is 0.95. For models where the variable “accumulated SOC stocks” were included, performance was further improved, showing a lower RMSE value than the respective model where it was not included. This was the case for all the models of Komakuk Beach but not for the soil layer below 15 cm at Ptarmigan Bay (15–30, 30–60 and 60–100 cm).

3.2. Variable importance

At Ptarmigan Bay, radar backscatter (a proxy for micro topography) was an important variable for SOC stocks in the upper depth increments 0–5 and 5–15 cm and for the lowest depth 60–100 cm. Vegetation indices were the most important for the depth intervals from 15 to 60 cm (Table 2,Fig. 4). The models for the N stocks have a similar behavior. Radar backscatter was among the most important variables for depth 15–30 cm and it was also important for depth 0–5 and 60–100 cm. For the SOC-stocks as well as for the N stocks landcover (LC, Table 1) was the second most important variable after radar backscatter when looking at the depth 60–100 cm. At Komakuk Beach radar backscatter (C-HH) together with the LC were important variables across all depths. However, for the SOC stocks, NDVI was the most important for the upper soil and the worldview-3B for the lowest depth interval. For the N stocks worldview-3B was important for the lowest depth interval was well. For the upper depth intervals LC was more important than radar backscatter.

When adding the accumulated SOC stocks as predictor variable, this variable is selected as the second most important variable for depth 5–15 cm at Ptarmigan Bay and the second most for depth 15–30 cm. For depth 30–60 cm the soil was not selected as important and GEMI, NIR2 and NDVI were the most important variables, which is the same result as without the soil variable. This applies as well to the depth 60–100 cm where radar backscatter and LC were most important. At Komakuk Beach the accumulated SOC stocks was the most important variable for depth 5–15, 30–60 and 60–100 cm, but not for depth 15–30 cm.

The models for the SOC content (Supplementary material Figs. S14–15) show similar patterns as the models for the SOC stocks with in general higher importance of the radar backscatter at Ptarmigan Bay and lower relevance at Komakuk Beach. A difference can be seen that for 30–60 and 60–100 cm the elevation is the most important variable at Komakuk Beach. Relevance of LC and radar backscatter decreases with depth whereas at Ptarmigan Bay particularly radar backscatter is important across all depths, except 0–5 cm.

3.3. Predictions

The highest amounts of SOC stocks and N stocks can be found in depth 15–30 cm in both study areas (Fig. 2). At Ptarmigan Bay SOC and N stocks are on average higher in all depth intervals except 60–100 cm where the average SOC and N stocks are higher at Komakuk Beach than at Ptarmigan Bay (Supplementary material. Table S 1). At Ptarmigan Bay more sites were cored, but often the coring hit an ice wedge which led to a higher presence of zero-values in the depth interval 60–100 cm. In this depth interval SOC and N stock values show a substantially higher variability at Ptarmigan Bay than at Komakuk Beach (Fig. 2). A higher variability can be also observed in the SOC stock data at depth 15–30 cm (Fig. 2 B).

The prediction maps reflect the values of the input data and show high SOC stocks and N stocks in depth 15–30 cm. The area in the northwest of the stream is dominated by dry high-center polygons where higher values are predicted in contrast to the wetter fen area east of the stream. Here lower SOC stock and N stock values were predicted. At Ptarmigan Bay in depth 15–30 cm there is a trend visible with higher SOC and N-stock values towards the west of the study area. This trend is not visible across the other depth intervals. In general the centers of high-centered polygons and the rims of low-centered polygons show higher predicted SOC and N stock values, whereas areas of low-centered polygon centers show lower predicted values.

The SOC content is the highest in 0–5 and 5–15 cm depth and decreases substantially with the 15–30 cm depth interval (Fig. 2 A). At Ptarmigan Bay there is a lower spread in the data for 0–5 and 5–15 cm than for Komakuk Beach. For 30–60 and 60–100 cm the spread is higher at Ptarmigan Bay and for 15–30 cm both areas have a relatively high spread. At Komakuk Beach the total study area is 20.7 km^2 and the average SOC stock is 66.6 kg m^{-2} and the average N stock is 3.77 kg m^{-2}

Table 2
Model results of random forest models. The table shows the mtry at which the best performing model was chosen according to the RMSE and the three most important variables of each model. The standard deviation was added to put the RMSE value into context.

model	mtry	RMSE	MAE	StDev	Var1	Var2	Var3
Ptarmigan Bay							
SOC stocks							
PB_0-5 cm	5	0.67	0.67	0.95	DEM	S1 C-HH	NDVI
PB_5-15 cm	11	2.10	2.10	2.51	S1 C-HH	NIR2	GEMI
PB_15-30 cm	4	4.66	4.66	6.48	GEMI	NDVI	B
PB_30-60 cm	2	6.51	6.51	8.24	GEMI	NIR2	NDVI
PB_60-100 cm	7	11.76	11.79	13.03	S1 C-HH	S1 C-VV	LC_PB1
SOC stocks (prediction of soil above included)							
PB_5-15 cm	14	1.84	1.84	2.51	SOC stock above	S1 C-HH	GEMI
PB_15-30 cm	7	4.66	4.66	6.48	GEMI	SOC stock above	B
PB_30-60 cm	2	6.53	6.53	8.24	GEMI	NIR2	NDVI
PB_60-100 cm	3	12.03	12.03	13.03	S1 C-HH	S1 C-VV	LC_PB1
N stocks							
PB_0-5 cm	9	0.06	0.07	0.09	DEM	B	S1 C-HH
PB_5-15 cm	2	0.16	0.16	0.21	NIR2	LC_PB2	NDVI
PB_15-30 cm	4	0.28	0.28	0.42	B	GEMI	S1 C-VV
PB_30-60 cm	2	0.39	0.39	0.5	GEMI	NDVI	NIR2
PB_60-100 cm	9	0.72	0.72	0.81	S1 C-HH	LC_PB1	NDVI
SOC content							
PB_0-5 cm	2	5.54	5.54	6.9	B	NDVI	LC_PB3
PB_5-15 cm	2	5.02	5.02	5.72	S1 C-VV	B	NDVI
PB_15-30 cm	2	9.09	9.09	10.21	LC_PB4	S1 C-VV	B
PB_30-60 cm	8	7.87	7.87	9.19	B	S1 C-HH	DEM
PB_60-100 cm	4	10.12	10.12	11.43	S1 C-HH	DEM	NIR2
Komakuk Beach							
SOC stocks							
KB_0-5 cm	7	0.78	0.78	1.04	NDVI	NIR2	S1 C-HH
KB_5-15 cm	2	1.53	1.53	1.82	S1 C-HH	NIR2	LC_KB4
KB_15-30 cm	2	4.04	4.04	5.94	LC_KB5	NIR2	DEM
KB_30-60 cm	2	6.06	6.06	7.35	S1 C-HH	LC_KB4	DEM
KB_60-100 cm	2	9.80	9.80	11.97	B	GEMI	LC_KB5
SOC stocks (prediction of soil above included)							
KB_5-15 cm	14	1.31	1.31	1.82	SOC stock above	LC_KB5	NIR2
KB_15-30 cm	2	3.90	3.90	5.94	S1 C-HH	NIR2	LC_KB5
KB_30-60 cm	12	5.04	5.04	7.35	SOC stock above	NIR2	B
KB_60-100 cm	2	9.65	9.65	11.97	SOC stock above	B	GEMI
N stocks							
KB_0-5 cm	2	0.05	0.05	0.07	LC_KB4	NIR2	NDVI
KB_5-15 cm	2	0.12	0.12	0.13	LC_KB4	DEM	NIR2
KB_15-30 cm	2	0.23	0.23	0.3	NIR2	S1 C-HH	LC_KB5
KB_30-60 cm	2	0.36	0.36	0.41	LC_KB4	S1 C-HH	B
KB_60-100 cm	2	0.61	0.61	0.72	B	NDVI	GEMI
SOC content							
KB_0-5 cm	2	9.81	9.81	12.12	S1 C-HH	GEMI	NIR2
KB_5-15 cm	2	10.85	10.85	12.47	LC_KB4	NDVI	LC_KB3
KB_15-30 cm	2	9.82	9.82	13.28	NDVI	GEMI	LC_KB3
KB_30-60 cm	2	7.01	7.01	10.03	DEM	NDVI	S1 C-HH
KB_60-100 cm	2	6.66	6.66	8.96	DEM	NDVI	NIR2

in the upper meter. The Ptarmigan Bay study area is 2.8 km^2 and the average SOC stock is 40.0 kg m^{-2} and the average N stock is 3.83 kg m^{-2} in the upper meter.

3.4. Uncertainty

The results of the uncertainty modelling based on quantile regression forest shows that higher accuracy is achieved close to the sampling locations and in areas were the spectral properties of the ground surface are similar to sampling locations (Fig. 6). Especially for Komakuk Beach, fen areas with partially inundated surface have the highest uncertainty. In Ptarmigan Bay, areas with higher uncertainty are more widespread across the area. The uncertainty differences in the depth 60–100 cm for Ptarmigan Bay are the highest. Observed vs. predicted value plots (Fig. 3) show a fairly good prediction of values around the mean value of the target variable, whereas an over-prediction of low values and an under-prediction of high values can be observed. This applies in particular to the zero-values of the depth interval 60–100 cm which were substantially over predicted.

4. Discussion

We have developed high resolution maps of SOC (Fig. 5) and N stocks and SOC content (Supplementary material Figs. S2–5) for two coastal areas along Canadian Beaufort Sea coast using a random forest algorithm. We show the distribution of C and N across the landscape and also discuss vertical differences in the upper 1 m of soil. We assessed the sensitivity of the model to available data and uncertainty of the model with quantile regression forest (Vaysse and Lagacherie, 2017; Yigini et al., 2018). We also developed models for SOC stocks where the soil

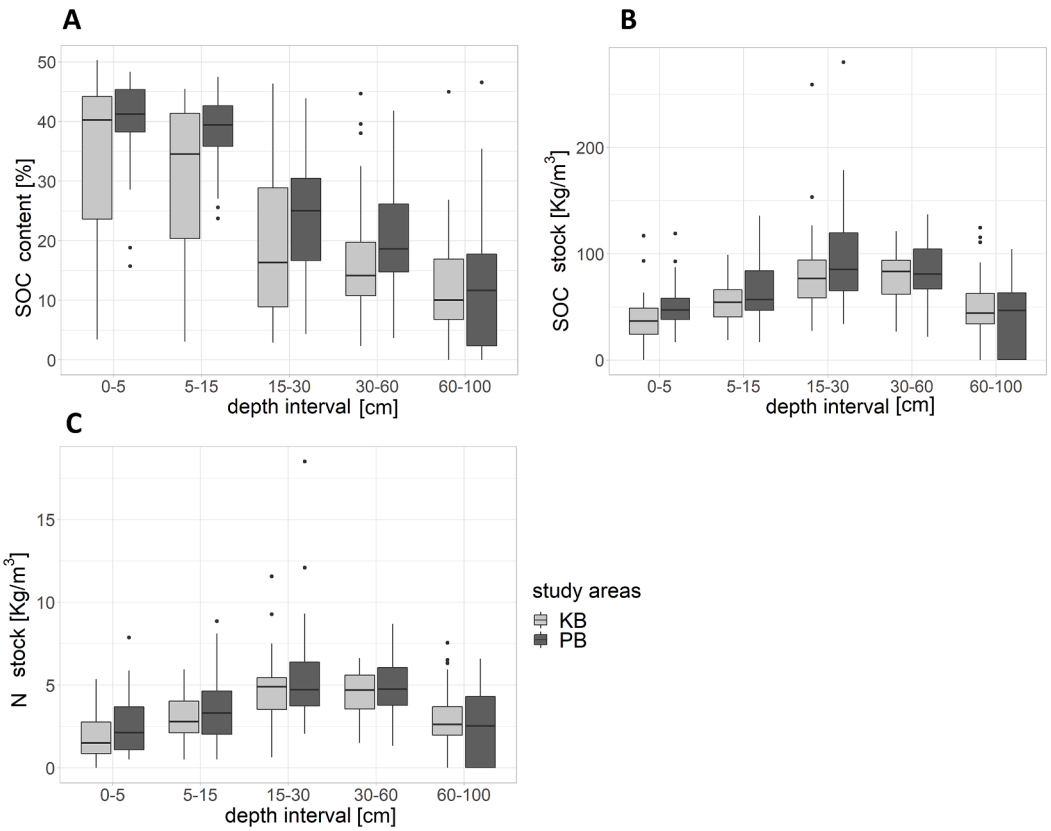


Fig. 2. Value distribution of SOC content (A), SOC stocks (B) and N stocks (C) for Komakuk Beach (KB) and Ptarmigan Bay (PB) across all sampled sites.

itself is added a predictor variable in the form of the accumulated SOC stocks. This is in line with the SCORPAN model (McBratney et al., 2003), as the soil itself is suggested to be an explanatory variable as well. With the help of the Random Forest models, it is further possible to identify important parameters that influence the target variables spatially across the landscape and also vertically within the soil profile.

4.1. *Variable importance*

A previous study introduced the usage of Sentinel 1C - HH winter backscatter to map SOC stocks in tundra environments (Bartsch et al., 2016). That study found a negative linear relationship between SOC stocks for 0–30 and 30–100 cm depth and early winter radar backscatter. The hypothesized mechanism is that varied microtopography (roughness) governs in early winter, when the other factors such as soil moisture and snow structure are minimal. A low roughness (low backscatter) indicates high SOC (Bartsch et al., 2016).

Our findings support that early winter radar backscatter can be used to map SOC stocks in tundra environments. It is a proxy for surface roughness, aboveground remains of vascular plants and near-surface soil texture and represents only volume scattering and surface roughness (Bartsch et al., 2016). In contrast, the signal during the summer is influenced by liquid water. High backscatter can be caused by open water and wetlands, but also by areas with high surface roughness. High SOC stocks are associated with areas of low surface roughness, therefore winter backscatter is used as it represents similar conditions as dry soil (Bartsch et al., 2016).

Our results underline the usage of radar backscatter as one predictor among several other predictors in a machine learning framework. In our study, radar backscatter, specifically C-HH, was important for the topsoil in both study areas, and even for the depth interval 60–100 cm at Ptarmigan Bay. This link between surface roughness and C stocks reflects an interaction between micro- topography (pronounced IWP rims at the surface) and the presence of massive ice wedges at depth (which reduce C and N stocks). Plotting linear fits between SAR backscatter and SOC stocks in surface (5–15 cm) and deep soils (30–60 cm) shows concurring increases of surface soil C with surface roughness, but opposing patterns for deep soil (Fig. 7). We interpret the opposing patterns for deep soil as a higher presence of ice wedges in the subsoil at Ptarmigan bay (commonly associated to LCPs), while the more mature HCPs at Komakuk Beach typically do not express massive ice wedge polygons in the upper meter of soil (but massive ice is likely present at greater depths), For the surface soils, weak positive trends indicate higher SOC stock values with increased surface expression of microtopography. This is in line with discussions in Bartsch et al. (2016) who suggest higher SOC stocks in cryoturbated soils where surface roughness and therefore radar backscatter is often higher. However, the correlation between SOC stocks and radar backscatter in our study is not strongly linear (Fig. 7) and further assessment of links between landform properties and C stock distributions are needed.

For Ptarmigan Bay, vegetation indexes (NDVI, NIR2 and GEMI) were among the most important variables for the SOC stocks across depths from 5 to 60 cm. Previous studies show that NDVI is linked to SOC stock distribution in permafrost regions and use NDVI as single predictor (Horwath Burnham and Sletten, 2010) or among other predictors (Siewert, 2018; Wu et al., 2022). Further studies in other climatic areas use NDVI as predictor for SOC stocks and could identify NDVI as one of the most important predictors (Gomes et al., 2019; Yang et al., 2016; Zhou et al., 2019). GEMI is a non-linear vegetation index and is similar to NDVI, but less sensitive to atmospheric effects. The importance of these vegetation-related variables shows that SOC stocks accumulation in these depths are related to plant activity and indirectly to rooting depth (Grosse et al., 2011). The latter is limited by the active layer depth which was around 30 cm during the time of the sampling. The influence of vegetation indices on the mainly frozen depth 30–60 cm might be caused by the strong link to near-surface organic matter by providing

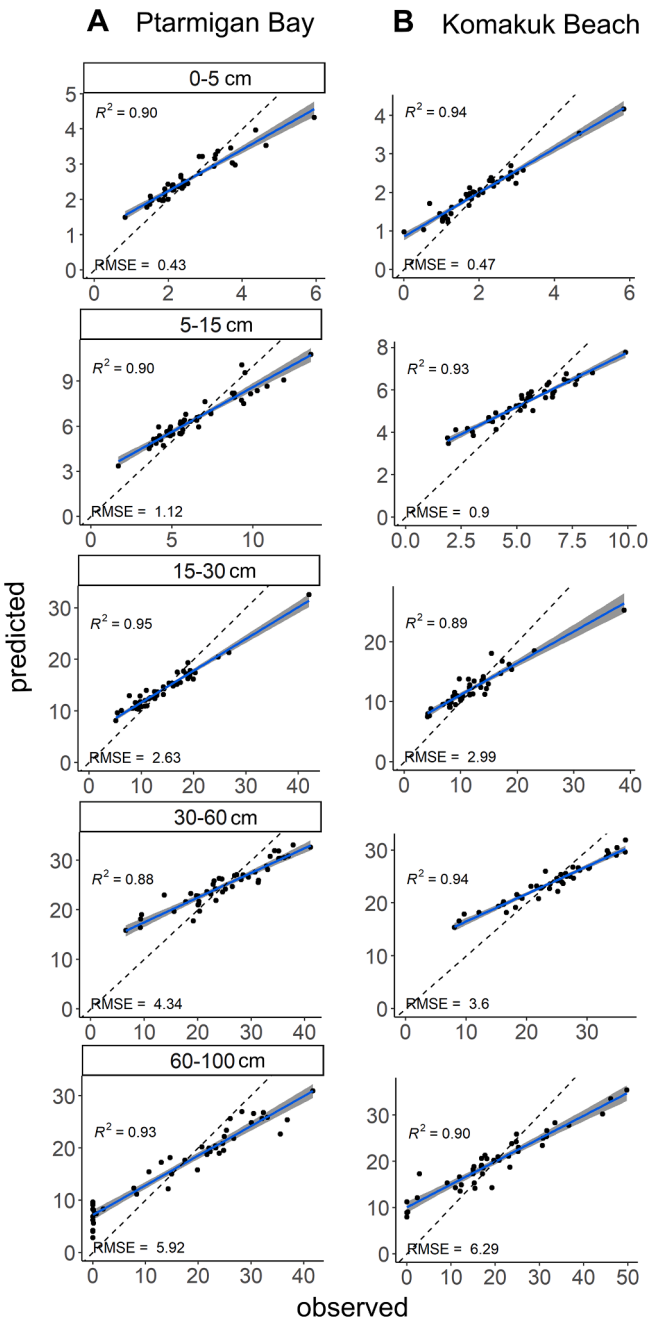


Fig. 3. Observed SOC stocks [kg m^{-2}] vs predicted SOC stocks [kg m^{-2}] by the random forest model for each depth interval at A Ptarmigan Bay and B Komakuk Beach. The 1:1 dashed line is added for reference and the blue line is the regression line. RMSE from observed and predicted values and adjusted R^2 from a linear model between observed and predicted values were added. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

inputs of SOC into the upper permafrost through cryoturbation (Ping et al., 2015). For Komakuk Beach however, landcover and microtopography (radar backscatter) related parameters were more important for the SOC stocks across all depths. NIR2 and GEMI were the second most important variables for the upper soil (0–5, 5–15 and 15–30 cm) and the lower soil (60–100 cm). This shows that vegetation, but also factors related to cryoturbation and permafrost processes are equally important at Komakuk Beach. At Ptarmigan Bay there is a difference between the different depths with more importance of vegetation related parameters in the middle and microtopography related parameters in the upper soil and lower soil linked to the ice wedges.

For some models the blue band of the worldview-3 image was chosen to be the among the most important variables. Water shows a higher reflectance in the blue band and therefore flooded areas and saturated soils are reflected in the blue band. The moisture conditions have a high influence on SOC stock accumulation and decomposition (van Huissteden, 2020a).

The high importance of the overlying soil layers for all depths at Komakuk Beach in contrast to Ptarmigan Bay indicates a tighter coupling between surface soil and deep frozen soils at Komakuk Beach (Supplementary material Figs. 10 and 11). At Ptarmigan Bay there was only high importance of the accumulated SOC stocks until depth 15–30 cm which indicates only an importance of the topsoil within the Active Layer. Here the influence of ground ice and its related processes and properties (van Huissteden, 2020b) on the deeper depth increments could play a more important role on SOC stocks than the upper soil.

At Komakuk Beach the reason for the influence of the overlying soil C on deeper layers could be caused by less ground ice in the upper meter or because the landforms are older and the SOC has accumulated deeper into the soil over time. According to Fritz et al. (2012) there has been extensive syngenetic peat growth at Komakuk Beach until 6700 cal years BP, and a gradual slowing (afterwards) with lower carbon accumulation rate during the last 4000 years (Fritz et al., 2016). However, there has been still regular peat growth initiations during the last 300 years at a studied ice wedge polygon at Komakuk Beach (Wolter et al., 2018), which shows continuous peat growth. A studied ice wedge polygon at Ptarmigan Bay shows that peat formation started within the last 1600–600 years especially with a buildup of peat in ice wedge polygons during at least 1100 years (Wolter et al., 2018). This indicates that the peat at Ptarmigan Bay is younger and soils at Komakuk Beach formed over longer time which is why the accumulated SOC stocks are important for deeper soil. The longer period of peat accumulation at Komakuk Beach also indicates that ice wedges, and therefore massive ground ice, might be found deeper than our 1 m sampling depth. However, this cannot be confirmed by our study, since no data from deeper soils was collected in this study. The results of our study with 66.6 kg $C/m^{-2}(-|-)$ at Komakuk Beach and 40.03 kg $C/m^{-2}(-|-)$ at Ptarmigan Bay is in contrast to the predicted values by the Northern Circumpolar Soil Carbon Database (NCSCD) (Hugelius et al., 2013). Here the average SOC stocks at Komakuk Beach are 18.1 kg C/m^{-2} and 92.8 kg C/m^{-2} at Ptarmigan Bay. Our results indicate less SOC storage where ground ice content is high.

4.2. *Uncertainties and limitations*

Our study shows relatively high uncertainties reflected by the RMSE values in relation to the standard deviation. This supports previous findings of a high variability of soil properties on a small scale in northern latitudes, especially in permafrost areas (e.g. Ping et al., 2015; Siewert et al., 2021, 2016).

One limitation of our study might be the amount of sampling points which could be collected in the field due to very challenging logistical constraints. The small amount of field sampling points lead to the selection of "leave-one-out" cross validation as internal validation method, which was used as well in a study by Wu et al. (2022).

By analyzing the variables chosen as most important in a random forest model, inferences can be made regarding which soil forming processes or mechanisms most affect SOC or N distribution. Studies use variable importance to interpret their relationships with soils and soil processes (e.g. McNicol et al., 2019). In contrast some studies suggest to not use variable importance to draw causal relationships between the target variable and the variable importance (Fourcade et al., 2018; Wadoux et al., 2020b). However, these studies test model performance by using pseudovariables in machine learning models. To be able to interpret and relate variable importance to soil processes and properties expert knowledge is necessary. This includes knowledge on which

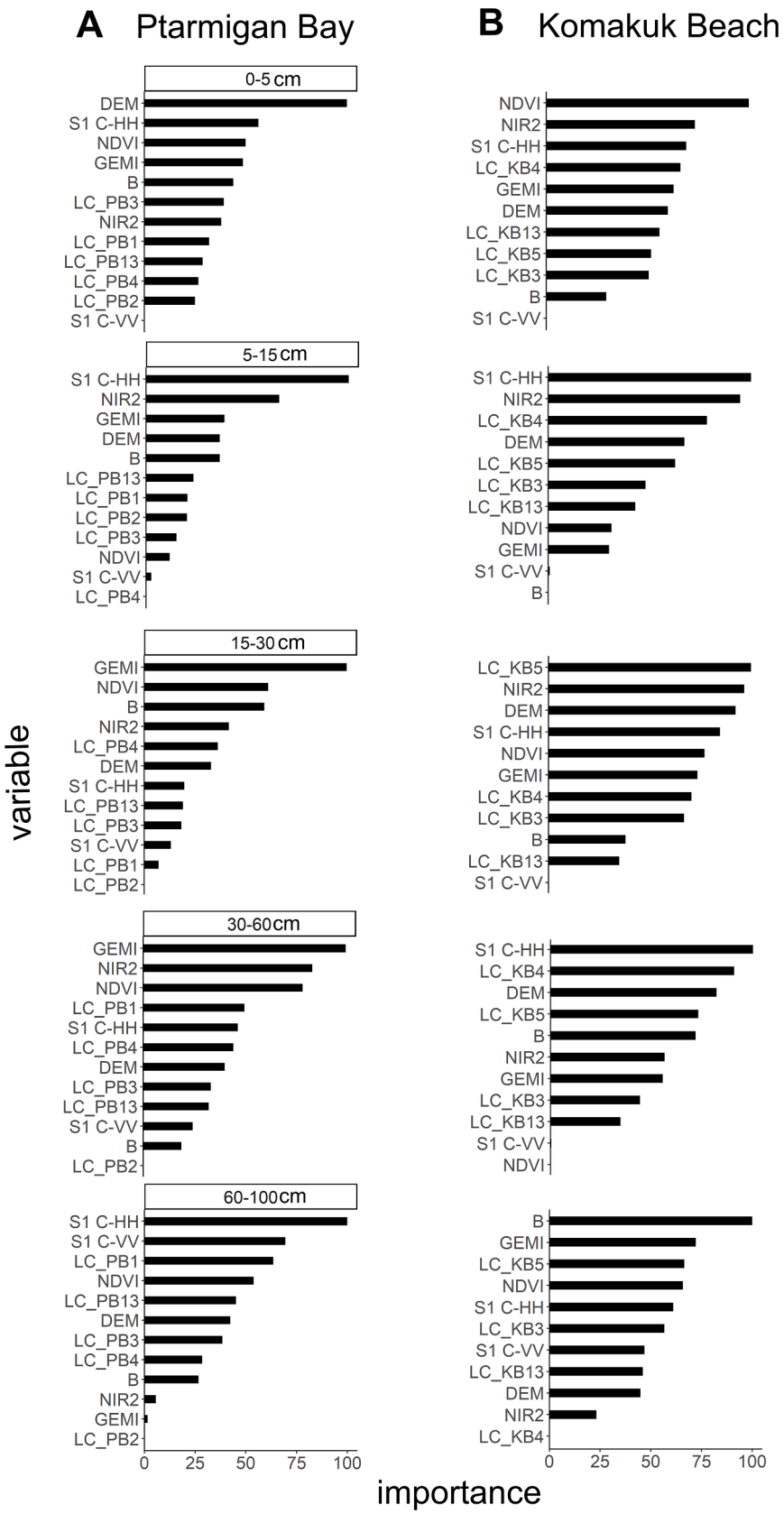


Fig. 4. Variable Importance for the random forest models for each depth interval (0–5, 5–15, 15–30, 30–60, 60–100 cm) at A Ptarmigan Bay and B Komakuk Beach (for explanations of abbreviations see Table 1 and section 2.6 for the different landcover classes. LC stands for landcover, PB for Ptarmigan Bay and the number for the landcover class described in section 2.6).

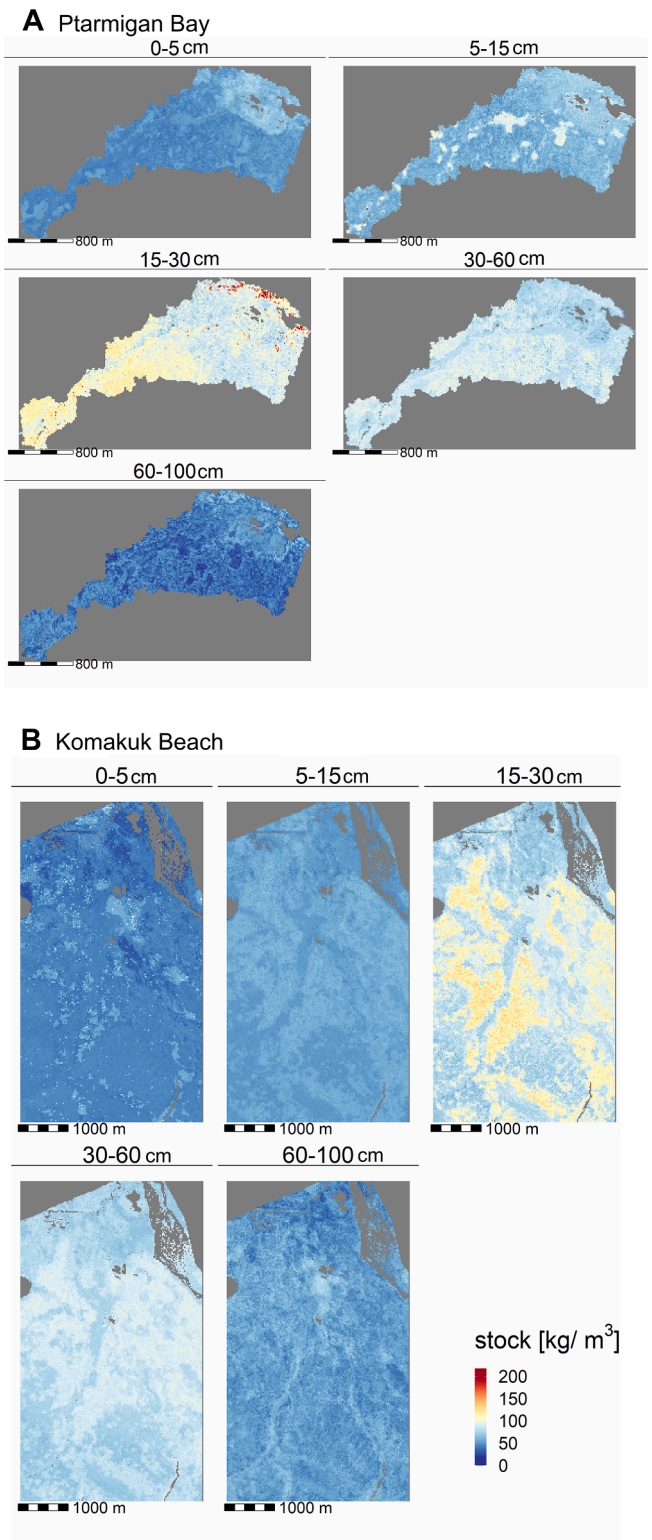


Fig. 5. SOC stock predictions for the five different depth intervals 0–5, 5–15, 15–30, 30–60 and 60–100 cm at A Ptarmigan Bay and B Komakuk Beach. The predictions were made in kg C/m^{-2}, but for a better comparability the values were converted into kg C/m^{-3}.

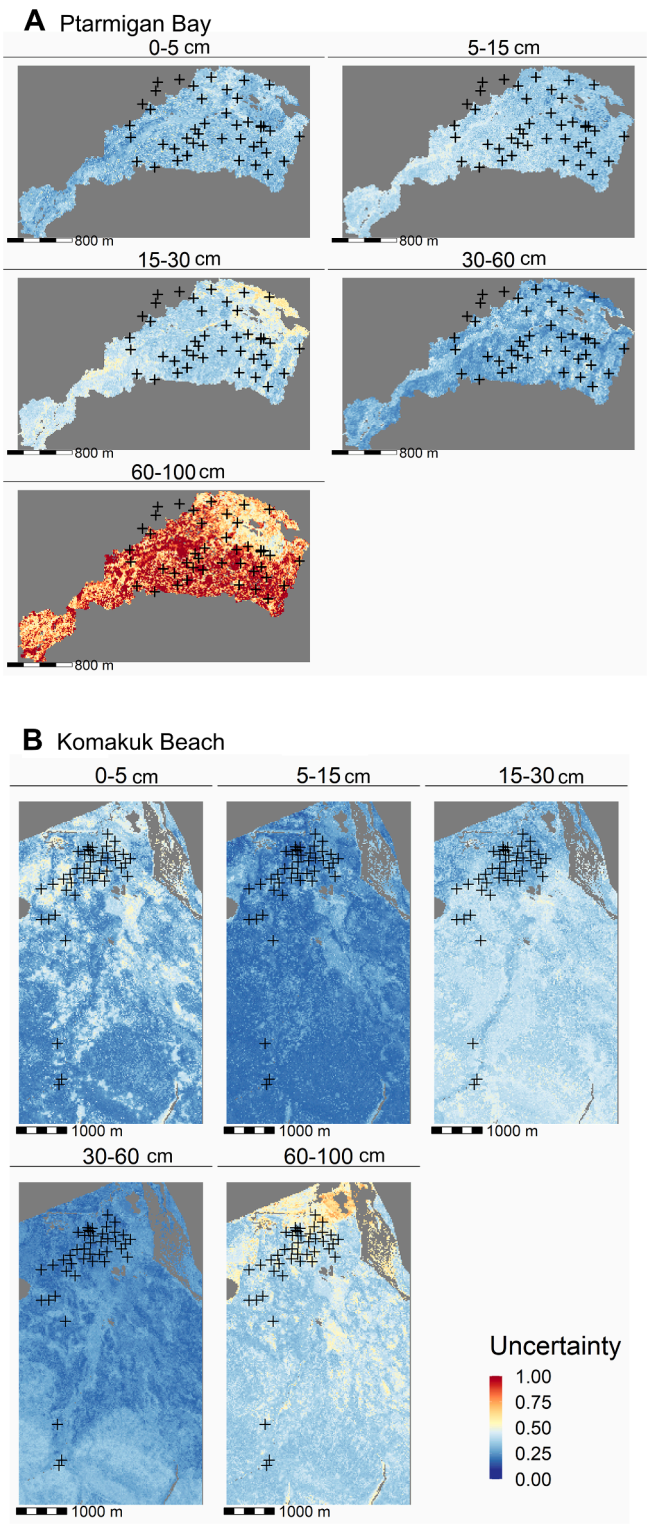


Fig. 6. Uncertainty of SOC stock predictions (fraction of mean from 0 to 1) at A Ptarmigan Bay and B Komakuk Beach using the Quantile Regression Forest method, where crosses indicate the soil sampling points. Note that each uncertainty map is based on an individual uncertainty assessment for that specific depth model and should be interpreted separately. The maps were combined in this figure with the same legend to be able to show all figures in the manuscript.

variables are related to the target variable. In the case of permafrost SOC stocks Sentinel 1 backscatter (Bartsch et al., 2016) or NDVI (Horwath Burnham and Sletten, 2010) have been proven to be related to SOC stock distribution. Another advantage of random forest models is their robustness and good performance in general (e.g. Li et al., 2011;

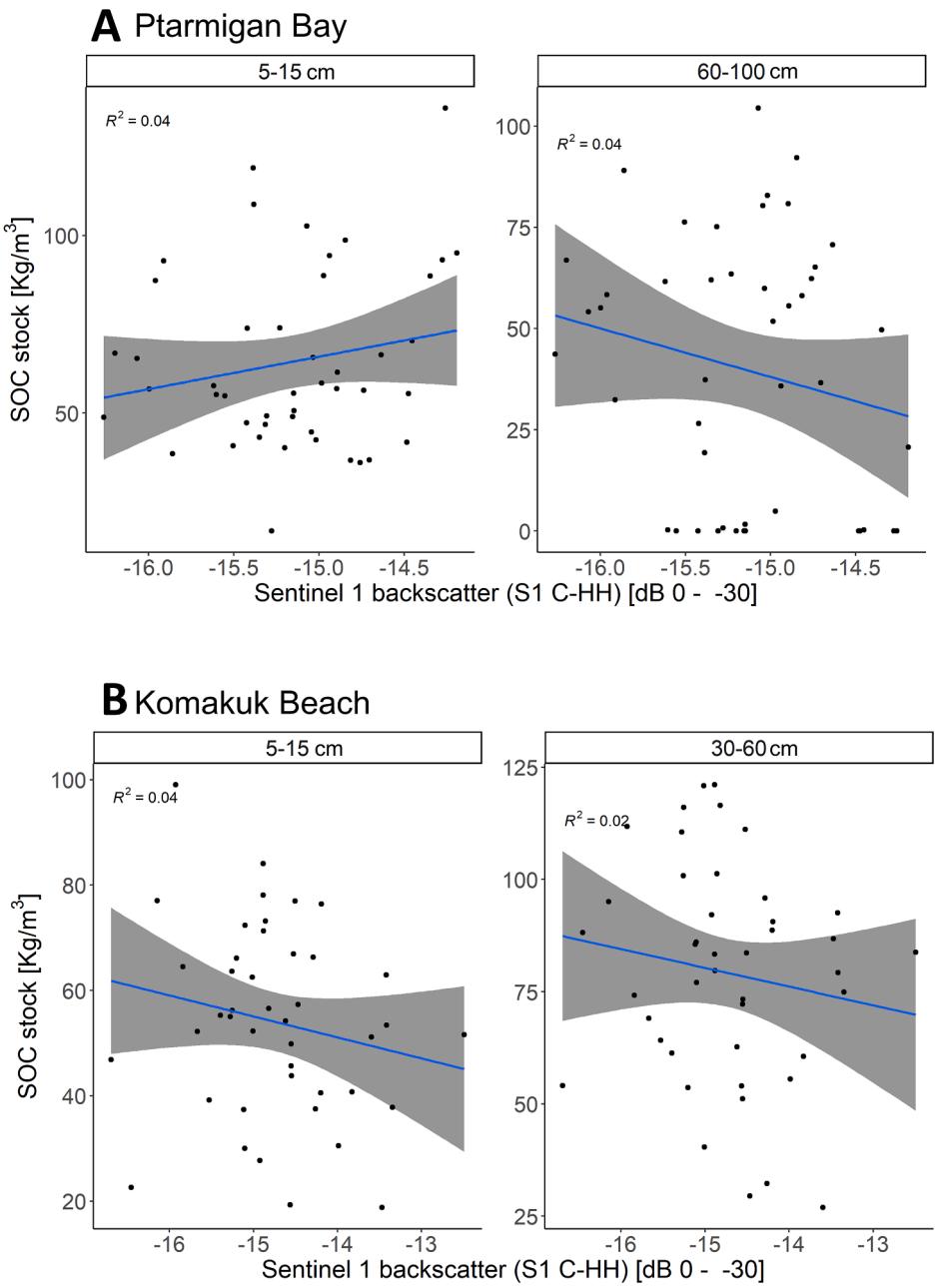


Fig. 7. Sentinel 1C-HH winter backscatter vs SOC stocks for A Ptarmigan Bayand B Komakuk Beachfor selected depth intervals where Sentinel 1C-HH was chosen to be the most important variable by the random forest model. Sentinel 1C-HH backscatter values are in sigma 0 at 30° and range in the full dataset from 0 to −30 dB (Widhalm et al. 2019).

Tziachris et al., 2019), but especially in permafrost environments (Siewert, 2018) which is why it is recently among the most popular used algorithms (Wadoux et al., 2020a). It performs well around the mean of the dataset (Fig. 3), however, a limitation in random forest models is that they cannot be used for extrapolation outside the range of the dataset. High values are under predicted and low values over predicted (Fig. 3). Especially the latter lead in our case to an over prediction of zero values in depth 60–100 cm at Ptarmigan Bay (3).

4.3. Implications and further research

Our study underlines the importance of ground ice as one major factor influencing the SOC stocks distribution in permafrost areas in lowland tundra ecosystems. Melting of ground ice can cause abrupt thaw events that potentially change SOC stocks and fluxes irreversibly (Turetsky et al., 2020). Both our study areas show similar surface properties such as typical lowland tundra landforms and associated plant communities. However, the subsurface structure with regard to soil properties and ground ice varied greatly between both areas in the upper meter with higher amounts of ground ice at Ptarmigan Bay than at Komakuk Beach. To be able to transfer our models and data to areas with no soil data availability, information on ice wedge structure and ground ice content is needed, in addition to organic layer thickness and peat depth. This information could be helpful to place pseudo-points based on expert knowledge that can generate possible new data points for digital soil mapping studies as shown by Koch et al. (2019). Data augmentation which is commonly used in deep learning applications (Padarian et al., 2019) could be an additional step in the pre-processing of predictor variables to make models more robust and potentially transferable to other study sites. Further, 3D models could be developed

to represent changes with depth (e.g. Adhikari et al., 2013; Zhang et al., 2020), though the algorithm needs to be selected carefully and the results assessed properly whether they can reflect variation with depth (Ma et al., 2021).

To gain detailed knowledge on ground-ice content and distribution, geophysical surveys could help to improve the knowledge on ice distribution (Schennen et al., 2022; Schwamborn et al., 2002).

In total, more SOC is stored in the upper meter at Komakuk Beach than at Ptarmigan Bay. However, considering the higher amounts of ground ice at Ptarmigan Bay, degradation of ice wedge polygons could make the soil carbon at this site more vulnerable to warming, particularly through abrupt thaw events. Siewert et al. (2021) strongly recommended landform-scale maps, which our study now provides. This will provide necessary information for modeling studies that assess subgrid-variability (Beer, 2016) as well as potential carbon fluxes in ice wedge polygon landscapes.

5. Conclusion

Our study provides one of the first high spatial and vertical resolution datasets of SOC stock and N stock in a coastal lowland tundra area in northwestern Canada. We also present maps of SOC stock, N stock and SOC content distribution at a spatial scale of 2 m which represents landform- and sub-landform-scale. This study partially captures the high variability in permafrost areas at a sub-regional scale and demonstrates that permafrost soils are highly complex.

Though landforms in both study areas are characterized by HCP and LCP being typical for lowland tundra, their developmental stages vary between the two study sites which has an impact on the C storage in the upper meter of the soil. The during the LGM previously glaciated area contains 40.0 kg C/m^{-2}, 3.83 kg N m^{-2} and more ground ice whereas the unglaciated area contains 66.6 kg C/m^{-2} and 3.77 kg N m^{-2} in the upper meter. Our study may contribute to improved understanding of C and nutrient fluxes and ice wedge polygon degradation on fine spatial scales.

Our results show that the random forest machine learning algorithm performs well for landscape scale mapping in areas with limited data availability. It is further possible to identify important parameters selected by these models. A limitation of random forest models is that they overpredict zero-values in ice rich depth intervals. High resolution maps of ground ice distribution and content, that we currently lack, are therefore vital to properly map SOC stocks for those depth intervals that contain a lot of ice. We therefore recommend to improve data on ground ice content as well as peat and organic layer thickness to refine the transferability of the models to other areas where no data is available. Geophysical techniques in combination with satellite data may here be valuable tools.

Funding

This study is part of the Nunataryuk project, which received funding under the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 773421. It was further funded by the Swedish Research Council (grant 2018–04516) and the ESA DUE GlobPermafrost project (contract number 4000116196/15/I–NB).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data, models and outputs will be archived openly in the Bolin Centre Database (https://bolin.su.se/data/).

Acknowledgements

We are grateful towards everyone who contributed to the project that made this publication possible. We thank the Yukon Territorial Government, Yukon Parks (Herschel Island Qikiqtaryuk Territorial Park) and the Aurora Research Institute for their support during the project. We also would like to thank J. Gimsa, S. Stettner and A. Beamish for assistance in the field during the field campaign in 2018 and S. McLeod, P. Archie and F. Dillon for their helpful insights and support in the field during the field campaigns in 2018 and 2019.

Author contributions

JW is the lead author responsible for the main scientific conceptualization, laboratory work setup, writing of the manuscript and co-responsible for the fieldwork setup. GH is responsible as the main supervising authority for this research project, co-responsible for the fieldwork setup and provided main conceptual content and guidance and input during manuscript writing. GH also provided main financial and managerial support. All co-authors contributed to data interpretation and writing of the manuscript. GT and AR contributed to fieldwork planning, logistics, and execution mainly of the field campaign in 2018 and to a minor extent to the field campaign in 2019 and contributed with conceptualization of the fieldwork setup and scientific input. NJS, RL and VM contributed to fieldwork planning, logistics, and execution of both field campaigns in 2018 and 2019 and provided scientific input. RL and VM provided minor manuscript writing input. WAC and LD contributed to fieldwork planning, logistics, and execution of the field campaign in 2019 and contributed with minor manuscript writing input. LD processed the soil samples from the field campaign in 2019. AB contributed with Sentinel-1 and 2 products. MF and HL provided scientific input throughout the project, planning of fieldwork and logistics and expertise about the research area.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.geoderma.2023.116652.

References

Adhikari, K., Kheir, R.B., Greve, M.B., Bøcher, P.K., Malone, B.P., Minasny, B., McBratney, A.B., Greve, M.H., 2013. High-Resolution 3-D Mapping of Soil Texture in Denmark. Soil Sci. Soc. Am. J. 77, 860–876. https://doi.org/10.2136/sssaj2012.0275.

Arndt, K.A., Santos, M.J., Ustin, S., Davidson, S.J., Stow, D., Oechel, W.C., Tran, T.T.P., Graybill, B., Zona, D., 2019. Arctic greening associated with lengthening growing seasons in Northern Alaska. Environ. Res. Lett. 14 (12), 125018.

Bartsch, A., Widhalm, B., Pointner, G., Ermokhina, K., Leibman, M., Heim, B., 2019a. Landcover derived from Sentinel-1 and Sentinel-2 satellite data (2015-2018) for subarctic and arctic environments. https://doi.org/10.1594/PANGAEA.897916.

Bartsch, A., Widhalm, B., Kuhry, P., Hugelius, G., Palmtag, J., Benjamin Siewert, M., 2016. Can C-band synthetic aperture radar be used to estimate soil organic carbon storage in tundra? Biogeosciences 13, 5453–5470. https://doi.org/10.5194/bg-13-5453-2016.

Bartsch, A., Leibman, M., Strozzi, T., Khomutov, A., Widhalm, B., Babkina, E., Mullanurov, D., Ermokhina, K., Kroisleitner, C., Bergstedt, H., 2019. Seasonal progression of ground displacement identified with satellite radar interferometry and the impact of unusually warm conditions on permafrost at the Yamal Peninsula in 2016. Remote Sens. 11, 1–25. https://doi.org/10.3390/rs11161865.

Bartsch, A., Widhalm, B., Leibman, M., Ermokhina, K., Kumpula, T., Skarin, A., Wilcox, E.J., Jones, B.M., Frost, G.V., Höfler, A., Pointner, G., 2020. Feasibility of tundra vegetation height retrieval from Sentinel-1 and Sentinel-2 data. Remote Sens. Environ. 237, 111515 https://doi.org/10.1016/j.rse.2019.111515.

Berner, L.T., Massey, R., Jantz, P., Forbes, B.C., Macias-Fauria, M., Myers-Smith, I., Kumpula, T., Gauthier, G., Andreu-Hayles, L., Gaglioti, B.V., Burns, P., Zetterberg, P., D'Arrigo, R., Goetz, S.J., 2020. Summer warming explains widespread but not uniform greening in the Arctic tundra biome. Nat. Commun. 11, 1–12. https://doi.org/10.1038/s41467-020-18479-5.

Biskaborn, B.K., Smith, S.L., Noetzli, J., Matthes, H., Vieira, G., Streletskiy, D.A., Schoeneich, P., Romanovsky, V.E., Lewkowicz, A.G., Abramov, A., Allard, M., Boike, J., Cable, W.L., Christiansen, H.H., Delaloye, R., Diekmann, B., Drozdov, D., Etzelmüller, B., Grosse, G., Guglielmin, M., Ingeman-Nielsen, T., Isaksen, K.,

Ishikawa, M., Johansson, M., Johannsson, H., Joo, A., Kaverin, D., Kholodov, A., Konstantinov, P., Kröger, T., Lambiel, C., Lanckman, J.P., Luo, D., Malkova, G., Meiklejohn, I., Moskalenko, N., Oliva, M., Phillips, M., Ramos, M., Sannel, A.B.K., Sergeev, D., Seybold, C., Skryabin, P., Vasiliev, A., Wu, Q., Yoshikawa, K., Zheleznyak, M., Lantuit, H., 2019. Permafrost is warming at a global scale. Nat. Commun. 10, 1–11. https://doi.org/10.1038/s41467-018-08240-4.

Blume-Werry, G., Lindén, E., Andresen, L., Classen, A.T., Sanders, N.J., von Oppen, J., Sundqvist, M.K., Collins, B., 2018. Proportion of fine roots, but not plant biomass allocation below ground, increases with elevation in arctic tundra. J. Veg. Sci. 29 (2), 226–235.

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24, 123–140. https://doi.org/10.1007/bf00058655.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32.

Couture, N.J., Irrgang, A., Pollard, W., Lantuit, H., Fritz, M., 2018. Coastal Erosion of Permafrost Soils Along the Yukon Coastal Plain and Fluxes of Organic Carbon to the Canadian Beaufort Sea. J. Geophys. Res. Biogeosci. 123 (2), 406–422.

Couture, N.J., Pollard, W.H., 2017. A Model for Quantifying Ground-Ice Volume, Yukon Coast, Western Arctic Canada. Permafr. Periglac. Process. 28, 534–542. https://doi.org/10.1002/ppp.1952.

Ernakovich, J.G., Barbato, R.A., Rich, V.I., Schädel, C., Hewitt, R.E., Doherty, S.J., Whalen, E., Abbott, B.W., Barta, J., Biasi, C., Chabot, C.L., Hultman, J., Knoblauch, C., Vetter, M.Y.L., Leewis, M.-C., Liebner, S., Mackelprang, R., Onstott, T. C., Richter, A., Schütte, U.E., Siljanen, H.M.P., Taş, N., Timling, I., Vishnivetskaya, T. A., Waldrop, M.P., Winkel, M., 2022. Microbiome assembly in thawing permafrost and its feedbacks to climate. Glob. Chang. Biol. 28 (17), 5007–5026.

Esri. "Ligh gray base" [basemap]. Scale Not Given. " ESRI World Light Gray Base Map" February 09, 2023. https://www.arcgis.com/home/item.html?id=94eb28b8734a4844bc30cf66332424db. (Feb 09, 2023).

Fourcade, Y., Besnard, A.G., Secondi, J., 2018. Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. Glob. Ecol. Biogeogr. 27, 245–256. https://doi.org/10.1111/geb.12684.

French, H.M., 2007. The Periglacial Environment, The Periglacial Environment: Third Edition. John Wiley & Sons Ltd, West Sussex, England. https://doi.org/10.1002/9781118684931.

Fritz, M., Wetterich, S., Schirrmeister, L., Meyer, H., Lantuit, H., Preusser, F., Pollard, W. H., 2012. Eastern Beringia and beyond: Late Wisconsinan and Holocene landscape dynamics along the Yukon Coastal Plain. Canada. Palaeogeogr. Palaeoclimatol. Palaeoecol. 319–320, 28–45. https://doi.org/10.1016/j.palaeo.2011.12.015.

Fritz, M., Wolter, J., Rudaya, N., Palagushkina, O., Nazarova, L., Obu, J., Rethemeyer, J., Lantuit, H., Wetterich, S., 2016. Holocene ice-wedge polygon development in northern Yukon permafrost peatlands (Canada). Quat. Sci. Rev. 147, 279–297. https://doi.org/10.1016/j.quascirev.2016.02.008.

Gomes, L.C., Faria, R.M., de Souza, E., Veloso, G.V., Schaefer, C.E.G.R., Filho, E.I.F., 2019. Modelling and mapping soil organic carbon stocks in Brazil. Geoderma 340, 337–350. https://doi.org/10.1016/j.geoderma.2019.01.007.

Government of Canada, 2022. Canadian Climate Normals & Averages [WWW Document]. Climate. URL http://climate.weather.gc.ca/climate_normals/index_e.html (accessed 4.30.22).

Grosse, G., Harden, J., Turetsky, M., McGuire, A.D., Camill, P., Tarnocai, C., Frolking, S., Schuur, E.A.G., Jorgenson, T., Marchenko, S., Romanovsky, V., Wickland, K.P., French, N., Waldrop, M., Bourgeau-Chavez, L., Striegl, R.G., 2011. Vulnerability of high-latitude soil organic carbon in North America to disturbance. J. Geophys. Res. Biogeosci. https://doi.org/10.1029/2010JG001507.

Hanes, C.C., Wotton, M., Woolford, D.G., Martell, D.L., Flannigan, M., 2022. Mapping organic layer thickness and fuel load of the boreal forest in Alberta, Canada. Geoderma 417, 115827. https://doi.org/10.1016/j.geoderma.2022.115827.

Heijmans, M.M.P.D., Magnússon, R., Lara, M.J., Frost, G. V., Myers-Smith, I.H., van Huissteden, J., Jorgenson, M.T., Fedorov, A.N., Epstein, H.E., Lawrence, D.M., Limpens, J., 2022. Tundra vegetation change and impacts on permafrost. Nat. Rev. Earth Environ. https://doi.org/10.1038/s43017-021-00233-0.

Heiri, O., Lotter, A.F., Lemcke, G., 2001. Loss on ignition as a method for estimating organic and carbonate content in sediments: Reproducibility and comparability of results. J. Paleolimnol. 25, 101–110. https://doi.org/10.1023/A:1008119611481.

Heung, B., Ho, H.C., Zhang, J., Knudby, A., Bulmer, C.E., Schmidt, M.G., 2016. An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. Geoderma 265, 62–77. https://doi.org/10.1016/j.geoderma.2015.11.014.

Horwath Burnham, J., Sletten, R.S., 2010. Spatial distribution of soil organic carbon in northwest Greenland and underestimates of high Arctic carbon stores. Global Biogeochem. Cycles 24. https://doi.org/10.1029/2009GB003660.

Hugelius, G., Bockheim, J.G., Camill, P., Elberling, B., Grosse, G., Harden, J.W., Johnson, K., Jorgenson, T., Koven, C.D., Kuhry, P., Michaelson, G., Mishra, U., Palmtag, J., Ping, C.-L., O'Donnell, J., Schirrmeister, L., Schuur, E.A.G., Sheng, Y., Smith, L.C., Strauss, J., Yu, Z., 2013. A new data set for estimating organic carbon storage to 3 m depth in soils of the northern circumpolar permafrost region. Earth Syst. Sci. Data 5, 393–402. https://doi.org/10.5194/essd-5-393-2013.

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., Kuhry, P., 2014. 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.

Jenny, H., 1994. Factors of soil formation: A system of quantitative pedology, [Republica. ed, Dover books on earth sciences. Dover Publications, New York.

Koch, J., Stisen, S., Refsgaard, J.C., Ernstsen, V., Jakobsen, P.R., Højberg, A.L., 2019. Modeling depth of the redox interface at high resolution at national scale using random forest and residual gaussian simulation. Water Resour. Res. 55, 1451–1469. https://doi.org/10.1029/2018WR023939.

Kuhn, M., 2022. caret: Classification and Regression Training. R package version 6.0-92.

Lagacherie, P., 2008. Digital soil mapping: A state of the art. In: Hartemink, A.E., McBratney, A., Mendonça-Santos, M.d.L. (Eds.), Digital Soil Mapping with Limited Data. Springer Netherlands, Dordrecht, pp. 3–14.

Lamichhane, S., Kumar, L., Wilson, B., 2019. Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review. Geoderma 352, 395–413.

Leutner, B., Horning, N., Schwalb-Willmann, J., Hijmans, R.J., 2022. RStoolbox: Tools for Remote Sensing Data Analysis. Version 0.3.0.

Li, J., Heap, A.D., Potter, A., Daniell, J.J., 2011. Application of machine learning methods to spatial interpolation of environmental variables. Environ Model Softw. 26, 1647–1659. https://doi.org/10.1016/j.envsoft.2011.07.004.

Liljedahl, A.K., Boike, J., Daanen, R.P., Fedorov, A.N., Frost, G.V., Grosse, G., Hinzman, L.D., Iijma, Y., Jorgenson, J.C., Matveyeva, N., Necsoiu, M., Raynolds, M. K., Romanovsky, V.E., Schulla, J., Tape, K.D., Walker, D.A., Wilson, C.J., Yabuki, H., Zona, D., 2016. Pan-Arctic ice-wedge degradation in warming permafrost and its influence on tundra hydrology. Nat. Geosci. 9, 312–318. https://doi.org/10.1038/ngeo2674.

Ma, Y., Minasny, B., McBratney, A., Poggio, L., Fajardo, M., 2021. Predicting soil properties in 3D: Should depth be a covariate? Geoderma 383, 114794. https://doi.org/10.1016/j.geoderma.2020.114794.

McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3–52. https://doi.org/10.1016/S0016-7061(03)00223-4.

McNicol, G., Bulmer, C., D'Amore, D., Sanborn, P., Saunders, S., Giesbrecht, I., Arriola, S. G., Bidlack, A., Butman, D., Buma, B., 2019. Large, climate-sensitive soil carbon stocks mapped with pedology-informed machine learning in the North Pacific coastal temperate rainforest. Environ. Res. Lett. 14 (1), 014004.

Meredith, M., Sommerkorn, M., Cassotta, S., Derksen, C., Ekaykin, A., Hollo, A., Kofinas, G., Mackintosh, A., Melbourne-Thomas, J., Muelbert, M.M.C., Ottersen, G., Pritchard, H., Schuur, E.A.G., 2022. Polar Regions. In: The Ocean and Cryosphere in a Changing Climate. Cambridge University Press, pp. 203–320. https://doi.org/10.1017/9781009157964.005.

Miner, K.R., Turetsky, M.R., Malina, E., Bartsch, A., Tamminen, J., McGuire, A.D., Fix, A., Sweeney, C., Elder, C.D., Miller, C.E., 2022. Permafrost carbon emissions in a changing Arctic. Nat. Rev. Earth Environ. 3, 55–67. https://doi.org/10.1038/s43017-021-00230-3.

Mishra, U., Hugelius, G., Shelef, E., Yang, Y., Strauss, J., Lupachev, A., Harden, J.W., Jastrow, J.D., Ping, C.L., Riley, W.J., Schuur, E.A.G., Matamala, R., Siewert, M., Nave, L.E., Koven, C.D., Fuchs, M., Palmtag, J., Kuhry, P., Treat, C.C., Zubrzycki, S., Hoffman, F.M., Elberling, B., Camill, P., Veremeeva, A., Orr, A., 2021. Spatial heterogeneity and environmental predictors of permafrost region soil organic carbon stocks. Sci. Adv. 7 https://doi.org/10.1126/sciadv.aaz5236.

Myers-Smith, I.H., Kerby, J.T., Phoenix, G.K., Bjerke, J.W., Epstein, H.E., Assmann, J.J., John, C., Andreu-Hayles, L., Angers-Blondin, S., Beck, P.S.A., Berner, L.T., Bhatt, U. S., Bjorkman, A.D., Blok, D., Bryn, A., Christiansen, C.T., Cornelissen, J.H.C., Cunliffe, A.M., Elmendorf, S.C., Forbes, B.C., Goetz, S.J., Hollister, R.D., de Jong, R., Loranty, M.M., Macias-Fauria, M., Maseyk, K., Normand, S., Olofsson, J., Parker, T. C., Parmentier, F.-J., Post, E., Schaepman-Strub, G., Stordal, F., Sullivan, P.F., Thomas, H.J.D., Tømmervik, H., Treharne, R., Tweedie, C.E., Walker, D.A., Wilmking, M., Wipf, S., 2020. Complexity revealed in the greening of the Arctic. Nat. Clim. Chang. 10 (2), 106–117.

Obu, J., 2021. How Much of the Earth's Surface is underlain by Permafrost? J. Geophys. Res. Earth Surf. 126, 1–5. https://doi.org/10.1029/2021JF006123.

Padarian, J., Minasny, B., McBratney, A.B., 2019. Using deep learning for digital soil mapping. Soil 5, 79–89. https://doi.org/10.5194/soil-5-79-2019.

Palmtag, J., Cable, S., Christiansen, H.H., Hugelius, G., Kuhry, P., 2018. Landform partitioning and estimates of deep storage of soil organic matter in Zackenberg. Greenland Cryosphere 12, 1735–1744. https://doi.org/10.5194/tc-12-1735-2018.

Palmtag, J., Obu, J., Kuhry, P., Richter, A., Siewert, M.B., Weiss, N., Westermann, S., Hugelius, G., 2022. A high spatial resolution soil carbon and nitrogen dataset for the northern permafrost region based on circumpolar land cover upscaling. Earth Syst. Sci. Data 14, 4095–4110. https://doi.org/10.5194/essd-14-4095-2022.

Phoenix, G.K., Bjerke, J.W., 2016. Arctic browning: extreme events and trends reversing arctic greening. Glob. Chang. Biol. 22, 2960–2962. https://doi.org/10.1111/gcb.13261.

Ping, C.L., Jastrow, J.D., Jorgenson, M.T., Michaelson, G.J., Shur, Y.L., 2015. Permafrost soils and carbon cycling. Soil 1 (1), 147–171.

Pinty, B., Verstraete, M.M., 1992. GEMI: a non-linear index to monitor global vegetation from satellites. Vegetatio 101, 15–20. https://doi.org/10.1007/BF00031911.

Poggio, L., De Sousa, L.M., Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Ribeiro, E., Rossiter, D., 2021. SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. Soil 7, 217–240. https://doi.org/10.5194/soil-7-217-2021.

Porter, C., Morin, P., Howat, I., Noh, M.J., Bates, B., Peterman, K., Keesey, S., Schlenk, M., Gardiner, J., Tomko, K., Willis, M., Kelleher, C., Cloutier, M.., Husby, E., Foga, S., Nakamura, H., Platson, M., Wethington, M.J., Williamson, C., Bauer, G., Enos, J., Arnold, G., Kramer, W., Becker, P., Doshi, A., D'Souza, C., Cummens, P., Laurier, F., Bojesen, M., 2018. "ArcticDEM", https://doi.org/10.7910/DVN/OHHUKH, Harvard Dataverse, V1, [accessed on 12.2.2018].

R Development Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramcharan, A., Hengl, T., Nauman, T., Brungard, C., Waltman, S., Wills, S., Thompson, J., 2018. Soil Property and Class Maps of the Conterminous United States

at 100-Meter Spatial Resolution. Soil Sci. Soc. Am. J. 82, 186–201. https://doi.org/10.2136/sssaj2017.04.0122.
Rampton, V.R., 1982. Surficial geology. Yukon Coastal Plain. https://doi.org/10.4095/126954.
Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS: Greenbelt, NASA SP-351. Proc. Third Earth Resour. Technol. Satell. 1 Symp. 309–317.
Schennen, S., Wetterich, S., Schirrmeister, L., Schwamborn, G., Tronicke, J., 2022. Seasonal Impact on 3D GPR Performance for Surveying Yedoma Ice Complex Deposits. Front. Earth Sci. 10, 1–14. https://doi.org/10.3389/feart.2022.741524.
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., Vonk, J.E., 2015. Climate change and the permafrost carbon feedback. Nature 520, 171–179. https://doi.org/10.1038/nature14338.
Schwamborn, G.J., Dix, J.K., Bull, J.M., Rachold, V., 2002. High-resolution seismic and ground penetrating radar-geophysical profiling of a thermokarst lake in the western Lena Delta, Northern Siberia. Permafr. Periglac. Process. 13, 259–269. https://doi.org/10.1002/ppp.430.
Siewert, M.B., 2018. High-resolution digital mapping of soil organic carbon in permafrost terrain using machine learning: A case study in a sub-Arctic peatland environment. Biogeosciences 15, 1663–1682. https://doi.org/10.5194/bg-15-1663-2018.
Siewert, M.B., Lantuit, H., Richter, A., Hugelius, G., 2021. Permafrost Causes Unique Fine-Scale Spatial Variability Across Tundra Soils. Global Biogeochem. Cycles 35, e2020GB006659. https://doi.org/10.1029/2020GB006659.
Siewert, M.B., Hanisch, J., Weiss, N., Kuhry, P., Maximov, T.C., Hugelius, G., 2015. Comparing carbon storage of Siberian tundra and taiga permafrost ecosystems at very high spatial resolution. J. Geophys. Res. Biogeosciences 120, 1973–1994. https://doi.org/10.1002/2015JG002999.
Siewert, M.B., Hugelius, G., Heim, B., Faucherre, S., 2016. Landscape controls and vertical variability of soil organic carbon storage in permafrost-affected soils of the Lena River Delta. Catena 147, 725–741. https://doi.org/10.1016/j.catena.2016.07.048.
Strauss, J., Abbott, B.W., Hugelius, G., Schuur, E., Treat, C., Fuchs, M., Schädel, C., Ulrich, M., Turetsky, M., Keuschnig, M., Biasi, C., Yang, Y., Grosse, G., 2021. 9. Permafrost, in: Recarbonizing Global Soils – A Technical Manual of Recommended Management Practices. FAO, pp. 127–147.
Tobler, W.R., 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. Econ. Geogr. 46, 234. https://doi.org/10.2307/143141.
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., McGuire, A.D., 2020. Carbon release through abrupt permafrost thaw. Nat. Geosci. 13, 138–143. https://doi.org/10.1038/s41561-019-0526-0.
Tziachris, P., Aschonitis, V., Chatzistathis, T., Papadopoulou, M., 2019. Assessment of spatial hybrid methods for predicting soil organic matter using DEM derivatives and soil parameters. Catena 174, 206–216. https://doi.org/10.1016/j.catena.2018.11.010.
Van Der Kolk, H.J., Heijmans, M.M.P.D., Van Huissteden, J., Pullens, J.W.M., Berendse, F., 2016. Potential Arctic tundra vegetation shifts in response to changing temperature, precipitation and permafrost thaw. Biogeosciences 13, 6229–6245. https://doi.org/10.5194/bg-13-6229-2016.
van Huissteden, J., 2020a. In: Thawing Permafrost: Permafrost Carbon in a Warming Arctic. Springer International Publishing, Cham, pp. 179–274.
van Huissteden, J., 2020b. In: Thawing Permafrost: Permafrost Carbon in a Warming Arctic. Springer International Publishing, Cham, pp. 107–177.
Vaysse, K., Lagacherie, P., 2017. Using quantile regression forest to estimate uncertainty of digital soil mapping products. Geoderma 291, 55–64. https://doi.org/10.1016/j.geoderma.2016.12.017.
Wadoux, A.M.J.C., Brus, D.J., Heuvelink, G.B.M., 2019. Sampling design optimization for soil mapping with random forest. Geoderma 355, 113913. https://doi.org/10.1016/j.geoderma.2019.113913.
Wadoux, A.M.J.C., Minasny, B., McBratney, A.B., 2020a. Machine learning for digital soil mapping: Applications, challenges and suggested solutions. Earth-Science Rev. 210, 103359 https://doi.org/10.1016/j.earscirev.2020.103359.
Wadoux, A.M.J.C., Samuel-Rosa, A., Poggio, L., Mulder, V.L., 2020b. A note on knowledge discovery and machine learning in digital soil mapping. Eur. J. Soil Sci. 71, 133–136. https://doi.org/10.1111/ejss.12909.
Walker, D.A., Raynolds, M.K., Daniëls, F.J.A., Einarsson, E., Elvebakk, A., Gould, W.A., Katenin, A.E., Kholod, S.S., Markon, C.J., Melnikov, E.S., Moskalenko, N.G., Talbot, S.S., Yurtsev, B.A., The other members of the CAVM Team, 2005. The Circumpolar Arctic vegetation map. J. Veg. Sci. 16 (3), 267–282.
Wickham, H., 2011. The split-apply-combine strategy for data analysis. J. Stat. Softw. 40, 1–29. https://doi.org/10.18637/jss.v040.i01.
Widhalm, B., Bartsch, A., Goler, R., 2018. Simplified normalization of C-band synthetic aperture radar data for terrestrial applications in high latitude environments. Remote Sens. 10, 1–18. https://doi.org/10.3390/rs10040551.
Widhalm, B., Bartsch, A., Goler, R., 2019. Normalized C-HH backscatter from Sentinel-1 (December, 2014-2017) for selected tundra regions. https://doi.org/10.1594/PANGAEA.897046.
Wild, B., Gentsch, N., Capek, P., Diáková, K., Alves, R.J.E., Bárta, J., Gittel, A., Hugelius, G., Knoltsch, A., Kuhry, P., Lashchinskiy, N., Mikutta, R., Palmtag, J., Schleper, C., Schnecker, J., Shibistova, O., Takriti, M., Torsvik, V.L., Urich, T., Watzka, M., Šantrůcková, H., Guggenberger, G., Richter, A., 2016. Plant-derived compounds stimulate the decomposition of organic matter in arctic permafrost soils. Sci. Rep. 6, 1–11. https://doi.org/10.1038/srep25607.
Wilson, R.M., Hough, M.A., Verbeke, B.A., Hodgkins, S.B., Tyson, G., Sullivan, M.B., Brodie, E., Riley, W.J., Woodcroft, B., McCalley, C., Dominguez, S.C., Crill, P.M., Varner, R.K., Frolking, S., Cooper, W.T., Chanton, J.P., Saleska, S.D., Rich, V.I., Tfaily, M.M., 2022. Plant organic matter inputs exert a strong control on soil organic matter decomposition in a thawing permafrost peatland. Sci. Total Environ. 820, 152757 https://doi.org/10.1016/j.scitotenv.2021.152757.
Wolter, J., Lantuit, H., Fritz, M., Macias-Fauria, M., Myers-Smith, I., Herzschuh, U., 2016. Vegetation composition and shrub extent on the Yukon coast, Canada, are strongly linked to ice-wedge polygon degradation. Polar Res. 35 (1), 27489.
Wolter, J., Lantuit, H., Wetterich, S., Rethemeyer, J., Fritz, M., 2018. Climatic, geomorphologic and hydrologic perturbations as drivers for mid- to late Holocene development of ice-wedge polygons in the western Canadian Arctic. Permafr. Periglac. Process. 29, 164–181. https://doi.org/10.1002/ppp.1977.
Wu, T., Wang, D., Mu, C., Zhang, W., Zhu, X., Zhao, L., Li, R., Hu, G., Zou, D., Chen, J., Wei, X., Wen, A., Shang, C., La, Y., Lou, P., Ma, X., Wu, X., 2022. Storage, patterns, and environmental controls of soil organic carbon stocks in the permafrost regions of the Northern Hemisphere. Sci. Total Environ. 828, 154464 https://doi.org/10.1016/j.scitotenv.2022.154464.
Yang, R.-M., Zhang, G.-L., Liu, F., Lu, Y.-Y., Yang, F., Yang, F., Yang, M., Zhao, Y.-G., Li, D.-C., 2016. Comparison of boosted regression tree and random forest models for mapping topsoil organic carbon concentration in an alpine ecosystem. Ecol. Ind. 60, 870–878. https://doi.org/10.1016/j.ecolind.2015.08.036.
Yigini, Y., Olmedo, G.F., Reiter, S., Baritz, R., Viatkin, K., Vargas, R., 2018. Soil Organic Carbon Mapping Cookbook 2nd edition. FAO.
Zhang, Y., Ji, W., Saurette, D.D., Easher, T.H., Li, H., Shi, Z., Adamchuk, V.I., Biswas, A., 2020. Three-dimensional digital soil mapping of multiple soil properties at a field-scale using regression kriging. Geoderma 366, 114253. https://doi.org/10.1016/j.geoderma.2020.114253.
Zhou, Y., Hartemink, A.E., Shi, Z., Liang, Z., Lu, Y., 2019. Land use and climate change effects on soil organic carbon in North and Northeast China. Sci. Total Environ. 647, 1230–1238. https://doi.org/10.1016/j.scitotenv.2018.08.016.