Martin Karlson[1*], and David Bastviken[1]

[1] Department of Thematic Studies/Environmental Change, Linköping University, Linköping, Sweden

Corresponding author: Martin Karlson (martin.karlson@liu.se)

Key Points:

- Peatland types in five contrasting high-latitude areas was mapped with high accuracy using freely available Earth Observation data
- Wide-scale mapping efforts should consider stratified classification implementation to account for variability in peatland characteristics
- Availability of high-quality reference data distributed throughout high-latitude regions is key for achieving wide-scale peatland mapping

Abstract

Mapping wetland types in northern-latitude regions with Earth Observation (EO) data is important for several practical and scientific applications, but at the same time challenging due to the variability and dynamic nature in wetland features introduced by differences in geophysical conditions. The objective of this study was to better understand the ability of Sentinel-1, Sentinel-2 and terrain derivatives derived from Copernicus DEM to distinguish three main peatland types, two upland classes, and surface water, in five contrasting landscapes located in the northern parts of Alaska, Canada and Scandinavia. The study also investigated the potential benefits for classification accuracy of using regionalized classification models constructed from region-specific training data compared to a global classification model based on pooled reference data from all five sites. Overall, the results show high promise for classifying peatland types and the three other land cover classes using the fusion approach that combined all three EO data sources (Sentinel-1, Sentinel-2 and terrain derivatives). Overall accuracy for the individual sites ranged between 84% to 92%. Class specific accuracies for the peatland types were also high overall, but differed between the five sites as well as between the three classes bog, fen and swamp. A key finding is that the regionalized classification models consistently outperformed the global classification by producing significantly higher classification accuracies for all five sites. This opens for promising progress in terms of identifying effective approaches for stratifying northern-latitude areas for continental scale peatland classification.

Plain Language Summary

This is optional but will help expand the reach of your paper. Information on writing a good plain language summary is available here.

1. Introduction

Accurate and up-to-date land cover maps with large area coverage have a broad range of important practical and scientific applications, for example in envi-

ronmental monitoring and management, biogeochemical modelling, and climate change impact assessments (Saah et al., 2019). Their value as sources of spatially explicit information of the extent, distribution and temporal dynamics of land cover classes is particular high in areas experiencing rapid environmental change, such as high-latitude regions (>60°N; Hinzman et al., 2005; Intergovernmental Panel on Climate Change, 2014). Frequent production of land cover maps enables systematic assessments of the impacts on landscapes resulting from, for example, changes in temperature, precipitation, fire, and human management (Pastick et al., 2019; Wulder et al., 2018). In recent years, the production of broad scale land cover maps has benefited greatly from several key technical developments, including improved capabilities of Earth Observation (EO) data, such as those provided by optical and synthetic aperture radar (SAR) systems, as well digital elevation models (DEM) and other ancillary datasets. Furthermore, accessibility of EO data repositories and cloud computing resources have increased greatly, as exemplified by the wide use of Google Earth Engine that has opened unprecedented opportunities for large area land cover mapping (Gorelick et al., 2017).

Several land cover products that cover the whole or parts of the high-latitude regions have been released in recent years, and their accuracy is often relatively high for general land cover classes (e.g., forest, water and barren land) that often produce distinct signatures in the EO data. Examples of recent global products include the Copernicus Global Land Cover (Buchhorn et al., 2020) and the WorldCover map produced by the European Space Agency (Zanaga et al., 2021). Many countries in these regions also have national land cover products managed by land surveying organizations or related stakeholders. However, such broad scale land cover class maps often conceal important thematic details rendering them less useful for specific applications that require a higher level of detail when characterizing the land cover. A prime example of such a broad land cover class is wetlands that in reality can include a large number of subclasses characterized by differences in vegetation composition, hydrology and soils, as well as ecological and biogeochemical functioning (Kreplin et al., 2021). National wetland inventories based on field surveys and manual interpretation of high-resolution imagery are available in some countries, including the U.S. (Wilen & Bates, 1995), Canada (Ducks Unlimited Canada, 2022) and Sweden (Gunnarsson & Löfroth, 2009), but these generally have incomplete spatial coverage and are sometimes outdated. In addition, these products often apply different wetland classification systems that can make regional comparisons problematic.

A consequence of limitations in existing land cover maps and wetland inventories, combined with the inherent inaccessibility of such areas making field surveys challenging, is that the extent of high-latitude wetlands is still highly uncertain and even more so when considering the abundance and geographical distribution of different wetland types (Bruhwiler et al., 2021; Melton et al., 2013). Such information is of vital importance for monitoring and managing these areas, which are often biodiversity hotspots and play a key role in local livelihoods. In addition, wetlands in these regions are large sources of green-

house gases (GHGs) to the atmosphere (Gorham, 1991; Hugelius et al., 2020). These wetland GHG fluxes are at the same time sensitive to climate change related impacts, including permafrost thaw, temperature increase, precipitation change and shifts in surface water regimes (Bruhwiler et al., 2021; Kreplin et al., 2021; Pastick et al., 2019). For example, as much as 99% of the arctic wetlands are underlain by permafrost, suggesting that large areas could be particularly vulnerable to regime shifts depending on the magnitude and effects of climate forcing (Kåresdotter et al., 2021), including the effects of permafrost melting, with uncertain yet potentially very large effects on wetland landscapes and their biogeochemical functioning and GHG emissions (Schuur et al., 2015).

Consequently, improved land cover products that differentiate between wetland types is of increasing importance for understanding and predicting future changes in the high-latitude regions, including GHG fluxes and feedback mechanisms (Bruhwiler et al., 2021; Olefeldt et al., 2021). Of particular concern in the context of carbon storage and GHG flux regulation is the balance between C storage and GHG emissions in different wetland types. In peatlands, dominating the high-latitude wetlands, the production of organic material exceeds the decomposition because of consistently waterlogged conditions, thereby promoting the formation and accumulation peat. While peatlands are estimated to cover only about 2.8% of the Earth´s land surface (Xu et al., 2018), it is the terrestrial ecosystem that stores and sequesters the highest amount of carbon per unit area. High-latitude peatlands are also one of the largest natural sources of methane (CH_4) to the global CH_4 budget (Kirschke et al., 2013; Saunois et al., 2016), but these emissions are at present poorly constrained (Peltola et al., 2019; Saunois et al., 2020; Thornton et al., 2016). Furthermore, CH_4 emissions differ largely between wetland types and locations in high-latitude areas (Bao et al., 2021). For example, mean emissions from fens can be more than twice as high as from bogs. There are also relatively large differences in CH_4 emissions among peatland types in different geographical regions with the highest reported emissions generally in lower latitude environments (Bao et al., 2021). Accordingly, the balance between C sequestration (net C exchange) and CH_4 emission from peatlands located throughout high-latitude regions are mainly the result of environmental heterogeneity, which influence the main controlling factors for peat accumulation, CH_4 production and CH_4 emission pathways, including water table depth, soil temperature and composition of the vegetation cover (see Bao, Jia and Xu 2021 and references therein). This balance is important, and it has been suggested that peatlands having a net C uptake can still contribute a net global warming potential (GWP) due to the CH_4 emissions (Bäckstrand et al., 2010). To assess peatland C sequestration and GHG emissions more accurately, higher resolution land cover information is vital.

Wetland mapping using remote sensing has engaged the scientific community for several decades (Guo et al., 2017; Rundquist et al., 2001), but many challenges still remain despite great advances in sensor capabilities and methods for EO data processing interpretation (A. L. Gallant, 2015). One of the main inherent challenges is that wetlands often are dynamic with potentially large

fluctuations in size and water levels through time, which results in constantly changing spectral signatures between seasons and years. Furthermore, distinguishing specific wetland types using remote sensing observations is far from a trivial task. For example, separating swamps from upland forests can be difficult because these classes are both characterized by a relatively dense canopy and therefore produce a similar appearance in both optical and SAR observations (Merchant et al., 2020). In addition, characteristic wetland vegetation species, such as mosses and sedges, are often spectrally similar at the spatial scales of most EO data available free-of-charge, making the separation between wetland types difficult. Broad-scale mapping wetland types in high-latitude regions presents additional challenges related to geographical differences in climate and local terrain. This can result in large regional variations in vegetation conditions and species composition (Hermosilla et al., 2022), as well as in hydrological regimes, peat accumulation and permafrost presence. For example, vegetation composition in different wetland types can vary along environmental gradients resulting in inconsistent spectral signatures and phenology, further complicating the production of large area products (Räsänen & Virtanen, 2019; Raynolds et al., 2019). Such variations can be compensated for to some extent by including ancillary datasets in the classification model used, such as gridded temperature and precipitation products and elevation derivatives (Mahdianpari et al., 2017). Another approach that has been proven successful is to divide the area to be mapped by the extent of satellite image tiling system (e.g., Landsat 8), administrative units (e.g., provinces) or eco-climatic regions, and perform the classification within those (Hermosilla et al., 2022; Mahdianpari, Salehi, et al., 2020; Mahdianpari et al., 2021; Mao et al., 2020). In this way, the classification algorithm can be trained using regionalized reference data that ensures the locally relevant representation of wetland types and thereby limit potential confusions between classes resulting from heterogeneity in environmental conditions.

Consequently, high-latitude peatlands can differ in structure and composition between geographical regions depending on for example local climatic conditions and topography, which translate into heterogeneous signatures in EO data. This challenge needs to be addressed when designing strategies for regional or continental scale mapping initiatives. This study therefore aims to clarify to what extent regionalized classification model implementation is preferable over a generalized classification model implementation when mapping main peatland types in high-latitude landscapes. Differences between these two approaches include both how reference data is used and how EO data sources should be combined in different geographic regions. To better understand these issues, we analysed peatland classification model implementation in five high-latitude landscapes in Alaska, Canada and Sweden with contrasting environmental conditions and terrain. The study also analyses the consistency with which peatland types in these different regions interact with a range of predictor variables derived from EO data. Specifically, this study focuses on the following questions:

1. How accurately can peatland types in different high-latitude regions be

classified using multi-source remote sensing data?

2. How does classification accuracy differ when comparing results from classification models trained by regionalized compared to generalized reference data?
3. How do the five sites differ in terms of the contribution of EO data sources and remotely sensed predictor variables to peatland type classification accuracy?

2. Materials and Methods

2.1 Peatland and Upland Classes

The classification scheme used in this study follows the Canadian Wetland Classification System to define main peatland types, with some modifications to the local conditions of the test sites and the available reference data. This system includes the two main peatland types bog and fen. It also includes a class for treed wetland, referred to as swamps, that can be underlain by peat, mineral soils or a combination of the two. In addition, the Canadian system includes the wetland class marsh, which are temporally flooded areas on mineral soils generally dominated by grasses and sedges. The marsh class was not considered in this study because of the limited occurrence and extent in the test areas that impeded the collection of a sufficiently large reference data sample. Two broad upland classes were considered in the study, including forest and other open land. In addition, open water was also included in our classification scheme. The following sections describe the peatland and upland classes in more detail. Descriptions of the peatland classes are based primarily on Warner and Rubec (1997).

2.1.1 Bog

Bogs are ombrogeneous peatlands that are often raised relative to the surrounding terrain, causing the hydrology to be largely disconnected from groundrunoff- and surface water. Precipitation therefore constitutes the primary source of water and nutrients. Bogs are relatively flat and open areas, and with a water table close to the surface. The upper peat layer of bogs is generally poor in nutrients and the surface water is acidic, forming a vegetation cover dominated by sphagnum mosses, often with a sparse cover of ericaeous shrubs and stunted trees.

2.1.2 Fen

Fens are minerotrophic peatlands that receives water from more nutrient-rich sources, including surface water and groundwater. This makes them less acidic, as well as more productive and biodiverse compared to bogs. Surface water often flow through channels and open water bodies that can produce specific surface patterns that are characteristic to fens. The water table can fluctuate a few centimetres around the surface of a fen and the vegetation cover is often a mix between brown mosses and graminoid species, including sedges and herbs.

A sparse cover of shrubs and trees can also be present in fens. The location of vegetation species in fens often reflect the depth of the water table, with graminoids in the wetter parts and shrubs and trees in the drier parts.

2.1.3 Swamp

Swamps can occur in areas with both peat and mineral soils, or a mix of the two soil types. However, swamps on mineral soils also tend to accumulate woody peat and was therefore included as a peatland type in this study. Characteristic to swamps is that they have a relatively dense vegetation cover ($> 30\%$) consisting of coniferous or deciduous trees, or tall shrubs. In addition, swamps experience periodic inundation depending on season, while the subsurface is continuously waterlogged.

2.1.4 Upland Classes

The forest class includes areas dominated by both coniferous and deciduous trees, or a mix of the two. Tree canopy cover can vary significantly between regions depending on local environmental conditions and definitions used by the reference data sources. Other open land is an aggregated class that primarily includes areas of rock outcrop and low growing vegetation.

2.1.5 Water

The water class include permanent water bodies such as ponds, lakes, streams and rivers. It also includes shallow water, which is defined as a separate wetland class in the Canadian wetland classification system.

2.2 Study Areas and Reference Data Sources

Five sites were used as areas to develop and test the peatland classification models, including sites in Alaska, Canada and Sweden (Figure 1). These sites represent different climatic, biogeographical and terrain conditions in high-latitude regions that make them interesting cases for comparison. Availability of reference data of sufficiently high quality was also an important aspect of the site selection process. Each test site covers the extent of two neighbouring Sentinel-1 scenes (253×340 km) acquired in the same orbit on the same date.

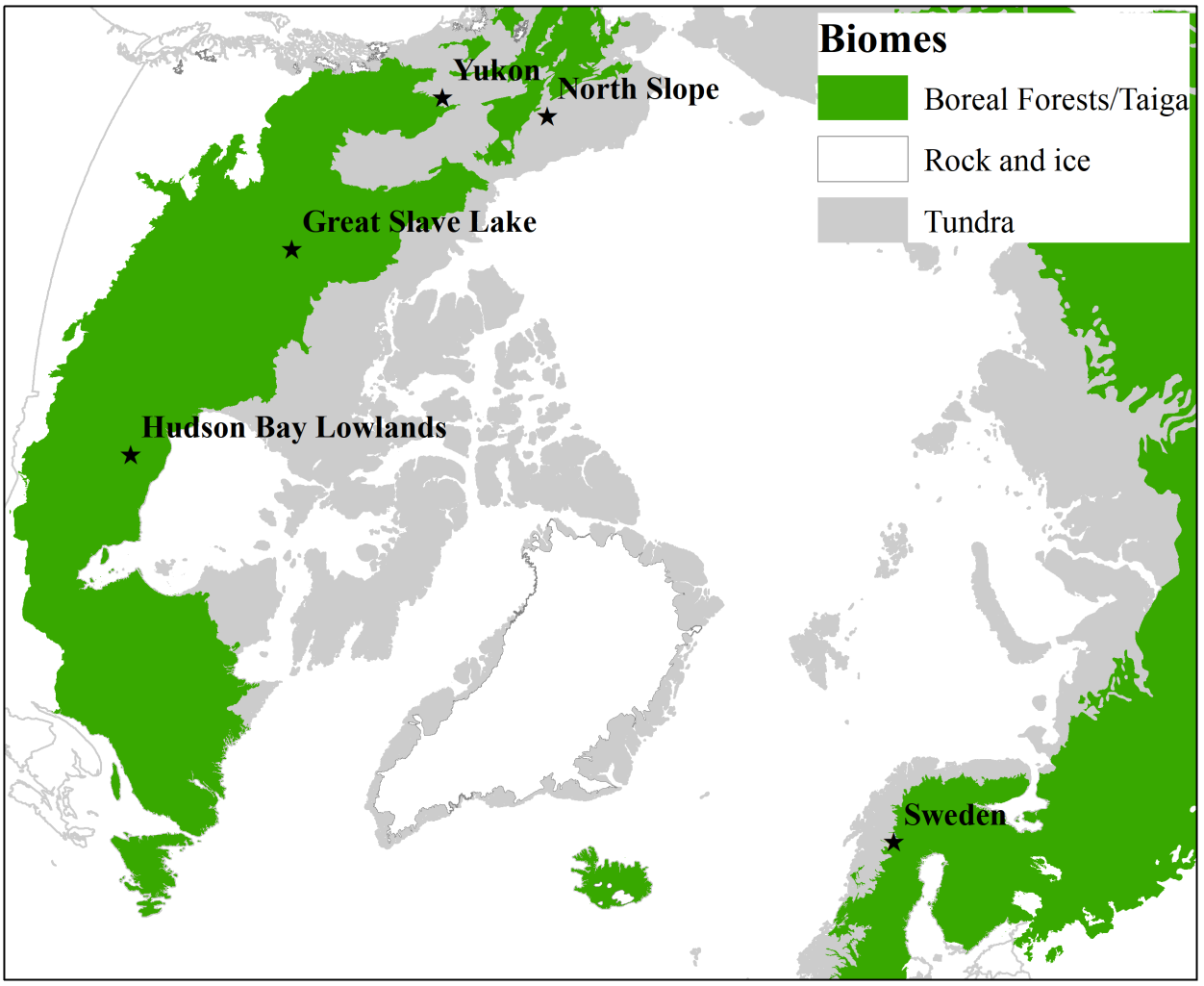

Figure 1. Map showing locations of the five test areas. The biome base-map is based on the Terrestrial Ecoregions of the World by the World Wildlife Foundation (Olson et al., 2001).

The following sections provide descriptions of the environmental conditions of the five sites and the respective reference data sets used for collecting training and validation data. For each site description, information about elevation was derived from the Copernicus DEM (Fahrland et al., 2020) and general climate conditions (temperature and precipitation) represent mean values (1970-2000) extracted from the WorldClim (V2) database (Fick & Hijmans, 2017).

Several sources and types of reference data covering different parts and extents of the five sites were consulted when collecting training and validation data for developing and testing the classification models to ensure a sufficient data

sample (Foody et al., 2016). These reference data do not implement consistent definitions of the peatland and upland classes, and class harmonization was needed to enable consistent use of class definitions in the mapping procedure. The procedure for class harmonization and the following development of the reference dataset is described in section 2.3.

2.2.1. Great Slave Lake, Canada

This site centres on the Great Slave Lake located in the Northwest Territories in central Canada. The areas surrounding the lake includes both boreal forest, primarily in the western and southern parts, and a mix of taiga and tundra in the northern and eastern parts, and contain large extents of peatland complexes (Olson & Dinerstein, 2002; Piper, 2016). Jack pine and spruce are the dominant tree species. Geologically, this area is located along the boundary of the Canadian Shield and the Interior Platform. Mean annual temperature (1970-2000) in the mapping area ranges between -7.3°C and -1.9°C (mean -3.6°C) and mean annual precipitation ranges between 217 mm and 357 mm (mean 259 mm). This is a relatively flat area with mean elevation of 206.5 m above sea level (min: 125.3, max: 445.3, std: 50.0). The northern parts of the area are within the extensive discontinuous permafrost zone (50-90%), whereas the southern portion is underlain by sporadic permafrost that covers between 10-50% of the land area.

The reference data consisted of the Great Slave Lake ecosystem map with a spatial resolution of 12.5 m and a minimum mapping unit of 0.2 ha (Bourgeau-Chavez et al., 2016, 2019). This product covers most of the test area and includes nine classes for wetland and upland areas, with a reported overall classification accuracy of 89.5%, and is based on multi-temporal PALSAR and Landsat-5 imagery (Bourgeau-Chavez et al., 2016). This dataset also includes a large number of polygons used for training and validation of the classification distributed throughout the test area, which were manually delineated based on high-resolution WorldView-2 imagery and field observations in 2014-2018. Peatland types in these data follow the Canadian classification system with the additional separation between open fens and treed fens.

2.2.2. Hudson Bay Lowlands, Canada

The Hudson Bay lowlands is the largest peatland area in Canada and located in northern Ontario between the Canadian Shield and the southern shore of the Hudson Bay. This site is a relatively flat landscape with mean elevation of 122.2 m a.s.l. (min: 0; max: 2593.7; std: 301.8). The northern part of the lowlands border Manitoba to the northwest. The area belongs to the boreal forest and taiga eco-zones and the climate is largely controlled by the water surface of the Hudson Bay, which often generates large amounts of rain during the summer (Abraham & Keddy, 2005). Mean annual temperature (1970-2000) in the test site ranges between -5.8°C and -2.0°C (mean -3.8°C) and mean annual precipitation ranges between 388 mm and 513 mm (mean 447 mm). Most of the area (90-100%) is underlain by continuous permafrost.

The main reference data used for this site is the wetland product distributed by

Land Information Ontario (Ontario Ministry of Natural Resources and Forestry, 2019). This product is primarily based on the Far North Land Cover map (Ontario Ministry of Natural Resources and Forestry, 2015). The Far North Land Cover map has a minimal mapping unit of about 1 ha and was created from Landsat imagery (2005-2011), DEM data and geological maps by using a semi-automated method with a relatively large component of visual interpretation to improve accuracy. Peatland classes in the wetland product include bog, fen and swamp. The Far North Land Cover map was also used to identify areas of the two upland classes.

For Hudson Bay Lowlands, we also used the wetland map by Franklin et al. (2018), which was created using RADARSAT-2 and Landsat 8 data and has a classification accuracy above 90%. This map covers a relatively small area south of the test site and was consequently not directly used as reference data. However, the landscape compositions of these two areas are very similar and the Franklin et al. (2018) map provided important insights to the visual appearance in high-resolution imagery of the different peatland types in the test area.

2.2.3. North Slope, Alaska

This test site is located in the eastern parts of North Slope Borough, Alaska, with the Beaufort Sea to the north and the Canadian border to the east. Much of the area is gently rolling tundra going from the Arctic coastal plains to the foothills of Brooks Range in the south. Mean elevation in the test site is 641.5 m a.s.l. (min: 0, max: 2433, std: 478.4). Some areas of boreal forest, mainly white spruce, and taiga can be found south of the Brooks Range. The central part is a designated national park, including the Gates of the Arctic National Park and Reserve and the Arctic National Wildlife Refuge. Mean annual temperature (1970-2000) ranges between -13.2°C and -4.7°C (mean -9.5°C) and mean annual precipitation ranges between 114 mm and 308 mm (mean 191 mm). The whole test site area is underlain by continuous permafrost (Jorgenson et al., 2008).

Four sources were used as reference data for the North Slope test site. The vegetation cover map of Walker and Raynolds (2017) covers the Upper Kuparuk River basin located in the centre of the test site and was created from satellite imagery (Landsat and SPOT) and existing geo-botanical maps. The classification scheme of this map is different from the Canadian system and the peatland classes include tussock/non-tussock sedge- dwarf shrub- moss tundra (bog), dwarf to low-shrub sedge- moss tundra (bog) and sedge- moss tundra (fen). We also used high-resolution vegetation community maps by Greaves et al. (2019) that cover three sites (3-6 km^2) near the Toolik research station in the central parts of the test area. These maps are based on LiDAR data and aerial photography collected in 2013 and were verified to an overall accuracy of 57% using field observations. Peatland classes include moist non-tussock tundra (bog), shrubby- moist non-tussock tundra (bog), raised areas in wet tundra (bog) and wet tundra (fen). In addition, we used the continuous foliar cover maps at 10 m spatial resolution covering North American Beringia (Nawrocki, 2021; Nawrocki et al., 2020). These maps represent the fractional coverage of

a range of vegetation species produced by a machine learning approach based on field data and multi-seasonal spectral and environmental covariates mainly derived from satellite data, including Sentinel-2. We used the sphagnum and wetland sedge maps to guide the identification of bogs and fens, respectively. Lastly, we used the U.S. Fish and Wildlife Service National Wetland Inventory dataset (version 2.0), which is primarily based on visual interpretation of aerial photography (Wilen & Bates, 1995), to specifically identify freshwater forested/shrub wetland areas (swamp) because this class was absent from the other three reference dataset.

2.2.4. Norrbotten County, Sweden

The Norrbotten county test site is located in the most northern parts of Sweden. Small portions in the north-western and north-eastern parts of the test area also extends into Norway and Finland, respectively. The western parts of the test area are mountains and alpine tundra, whereas the eastern parts is boreal forest consisting primarily of a mix of spruce, pines and birch. Mean annual temperature (1970-2000) in area ranges between -5.4°C and 3.5°C (mean -1.1°C) and mean annual precipitation ranges between 277 mm and 809 mm (mean 477 mm). The terrain is hilly in the western parts towards Norway and relatively flat in the eastern parts with a mean elevation of 500.7 m.a.s.l (min: 0, max: 2083.4, std: 284.6). The main part to the east is lowland with sporadic permafrost (10-50%), whereas the western areas characterized by higher elevation is often within the discontinuous permafrost zone (50-90%) (Gisnås et al., 2017).

The main reference data for this site is the Swedish national wetland inventory (VMI) which was produced through visual interpretation of aerial photography collected between 1981 and 2005, and field visits for verification (Gunnarsson & Löfroth, 2009). This dataset uses a classification schema that corresponds well with the Canadian system, but also includes a high number of subclasses (Naturvårdsverket, 2019). The main wetland types in the dataset include bogs (Swedish classification “mosse”), fens (“kärr”) and swamps (“sumpskog”). We also used the Swedish National Land Cover data from 2020 (Naturvårdsverket, 2020; Nilsson et al., 2020) for collecting reference data for upland area classes.

2.2.6. Yukon, Canada

This site is located in the south-western parts of Yukon, Canada, centred on Dawson City with the Alaska border to the west and the Klondike Plateau to the south. The northern boundary is roughly made up of the Ogilvie River. Most of the area is tundra, but the southern part includes large extents of boreal forests with mainly white spruce and pine. Relatively dense forest areas can also be found in the river valleys throughout the test site. The climate is subarctic with mean annual temperature (1970-2000) between -16.4°C and -2.2°C (mean -5.9°C) and mean annual precipitation between 196 mm and 1066 mm (mean 304 mm). The landscape includes plateaus and mountains with a mean elevation of 881.7 (min: 251.2, max: 2593.7, std: 301.8). This area is within the extensive discontinuous permafrost zone where the coverage is between 50-90%.

The primary reference data for Yukon test sites was a wetland map covering the Indian River watershed located south of Dawson City in the central parts of the test (Government of Yukon, 2018). This map follows the Canadian classification system and is based on visual interpretation of Geoeye-1 and SPOT-6 imagery using a mapping scale of 1:10000. It was validated using field observations in 2016 and 2017. The map also includes forest type classes, as well as other upland classes. Similar to the North Slope test site, we also used the continuous foliar cover maps at 10 m spatial resolution (Nawrocki, 2021; Nawrocki et al., 2020), described in section 2.1.3. to identify bog and fen areas. Lastly, we used the vegetation inventory map (40k) produced through visual interpretation of aerial photography and digital elevation models at a scale of 1:400000 (Government of Yukon, 2017) to guide the identification of upland class areas.

2.3 Preparation of Training and Testing Dataset

As described in the previous sections, we consulted several sources of reference data when compiling the datasets used for training and testing the classification models. These reference data sources originated from three countries and had been created for different purposes using various data sources, methods and classification systems, and a standardization was therefore needed. This primarily included the translation of peatland classes in the Alaskan and Swedish reference data to the Canadian classification system used in this study, as described in sections 2.1.3. and 2.1.4.

For each test site, 200 randomly distributed polygons per class were manually delineated based on overlaying the reference datasets on very high-resolution satellite imagery available in ArcGIS Pro, which resulted in 1200 polygons per site used for training and testing the classification models. The procedure of combining reference data with visual interpretation of VHR imagery enabled selection of polygons with a high likelihood of representing the actual class on the ground (Mahdianpari, Brisco, et al., 2020). This procedure also ensured exclusion of areas affected by land cover change or fires during the time between reference data development and acquisition of EO data used in this study from the training and validation dataset.

2.4 Processing of Sentinel-1 and Sentinel-2 Data

2.4.1 Sentinel-1

The Sentinel-1 system provides C-band synthetic aperture radar (SAR) data at dual polarization and a spatial resolution of 10 m. A monthly time-series of Sentinel-1 data from 2020 was downloaded from the Copernicus Open Access Hub for each site. We used the Ground Range Detected (GRD) interferometric wide (IW) product acquired in the VV-VH polarization mode. SAR preprocessing was done using the S1 Toolbox and included thermal noise removal, radiometric calibration and terrain correction using the Copernicus DEM. A Refined Lee filter (Yommy et al., 2015) was then applied to remove speckle noise. Two neighbouring and pre-processed scenes were mosaicked to cover each site. As Sentinel-1 predictor variables, we created an annual time series from 2020

consisting of one SAR observation per month in both VV and VH polarizations (Table 1). From the monthly observation time series, we also calculated temporal metric mosaics, including minimum, maximum, mean, median and standard deviation values at pixel level for both VV and VH polarizations.

Table 1. Predictor variables derived from Sentinel-1 VV and VH polarizations. Temporal metrics (min, max, mean, median and standard deviation) was calculated on the pixel level.

Predictor variables	Description
Monthly observations (January – December 2020)	One backscatter observation close to the middle of the m
Minimum	The minimum backscatter value extracted from the 2020
Maximum	The maximum backscatter value extracted from the 2020
Mean	The mean backscatter value extracted from the 2020 tim
Median	The median backscatter value extracted from the 2020 ti
Standard deviation	The standard deviation backscatter value extracted from

2.4.2 Sentinel-2

Sentinel-2 mosaics at 20 m spatial resolution based on time series covering the period July 1 to October 31 2020 was created for each test area using the Sentinel-2 Global Mosaic (S2GM; V1.3) service (Kirches, 2020). The input to S2GM consists of Sentinel-2 Level 2A surface reflectance products provided by the European Space Agency (ESA) Copernicus data hub processed with the Sen2Cor (v2.9) method. Sen2Cor also produces the associated product scene classification, aerosol optical depth and water vapour used as input by S2GM. The S2GM service is based on a compositing approach that selects the best pixel observation during the predefined time-period using a set of criteria considering spectra, viewing geometry and Sen2Cor scene classification (Frantz et al., 2017). Two different composite algorithms are used in S2MG depending on the number of valid observations at pixel-level, including Medoid composite (Flood, 2013) when the number of valid observation is four or more, and Short Term Composite (Kirches, 2020; Roy et al., 2011) when the four observation threshold is not met. All resulting spectral bands and were resampled to 30 m pixel size using bilinear interpolation and used as predictor variables (Table 2). From the 30 m Sentinel-2 mosaics, we also produced a number of vegetation indices, including Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI) and Red Edge NDVI (RENDVI).

Table 2. Predictor variables based on the Sentinel-2 mosaic from July to October, 2020.

Predictor variable	**Central wavelengths (bands) and equations (vegetation indices)**	**Referenc**
Band 2	Blue (490 nm)	-
Band 3	Green (560 nm)	-

Predictor variable	Central wavelengths (bands) and equations (vegetation indices)	Referenc
Band 4	Red (665 nm)	-
Band 5	Vegetation red edge (705 nm)	-
Band 6	Vegetation red edge (740 nm)	-
Band 7	Vegetation red edge (783 nm)	-
Band 8	Near infrared (842 nm)	-
Band 8A	Near infrared (865 nm)	-
Band 11	Short wave infrared (1610 nm)	-
Band 12	Short wave infrared (2190 nm)	-
NDVI	(B8 - B4) / (B8 + B4)	(Rouse et
NDWI	(B8 – B12) / (B8 + B12)	(Gao, 199
RENDVI	(B8 - B5) / (B8 + B5)	(Forkuor e

2.5 Processing of Elevation Data

The 30 m Copernicus DEM (v 2.1) with global coverage released in 2020 was used for terrain correction of the Sentinel-1 data and for extracting a range of terrain derivatives used as predictors in the peatland classification models (Table 3). This is an edited version of the WorldDEM^{TM} that was produced from SAR data acquired by the TANDEM-X satellite system between 2010-2015 (Fahrland et al., 2020). In the following sections, we describe the terrain derivatives used as predictor variables in the classification models.

Table 3. Terrain derivatives used in the classification model.

Predictor variable	Description
Depth to water Index (DTW)	Difference in elevation along the lowest slope path between a specific are
Water table elevation	Elevation difference between DEM and DTW at a specific area
Topographic wetness index (TWI)	Tendency of an area to accumulate surface water runoff based on local sl
Topographic Position Index (TPI)	Topographic position of an area relative to the surrounding landscape of
Curvature	Measure of the local variations in curvature of an area.
Landform classes	Specific topographic features of the Earth surface that together make up

2.5.1 Depth to Water Index

Depth to water index (DTW) has been proposed as a useful indicator for identifying areas with a high likelihood of being saturated (Murphy et al., 2008, 2009). It has previously been used to delineate wetlands in different types of boreal landscapes (Ågren et al., 2014; Lidberg et al., 2020; Oltean et al., 2016; White et al., 2012), but rarely in higher-latitude regions. The main assumptions behind DTW is that areas closer to surface water have a higher probability of being wet when considering proximity as a function of elevation difference and distance (Murphy et al., 2008). DTW is conceptualized as the elevation difference along the least-slope path from an area to the nearest surface water. Its calcula-

tion requires two primary inputs, including i) a terrain slope layer derived from the original DEM and ii) a layer representing surface water, including streams and waterbodies. DTW is then derived from the following least-cost function (Equation 1):

$$\text{DTW } (m) = c \sum \frac{\Delta_{\text{zi}}}{\Delta_{\text{xi}}} a_i \quad (1)$$

where c represents the cell size of the DEM, $\Delta_{\text{zi}}/\Delta_{\text{xi}}$ is the slope of cell_{i} along the least cost path, a is 1 when the path is parallel along the cell boundary or $2^{0.5}$ when the path crosses diagonally. Regarding input data, streams delineation was done using a filled version of the Copernicus DEM and the D8 algorithm (O´Callaghan & Mark, 1984) to determine flow direction, flow accumulation and slope gradient. We then extracted streams from the flow accumulation layer using stream initiation thresholds between 0.5 ha – 100 ha (Lidberg et al., 2020). These stream layers were each combined with the water body layer developed by Pekel et al. (2016) to create surface water layers for use in the DTW calculation. By subtracting the DTW layers from the original DEM, it is also possible to derive an estimate of the water-table elevation (WTE; White et al., 2012)). Consequently, WTE layers were calculated using each of the DTW layers as input.

2.5.2 Topographic Wetness Index

The hydrologically based topographic wetness index (TWI) quantifies the tendency of an area to accumulate water depending on size of the upslope catchment area and local slope (Beven & Kirby, 1979). It has been widely used in different types of techniques for predict location of wet areas (Grabs et al., 2009; Mattivi et al., 2019). TWI is calculated as follows:

$$TWI = \ln\left(\frac{\alpha}{\tan(\beta)}\right) \quad (2)$$

where is the upslope catchment area draining through a specific point and is the local slope angle. We used the *r.watershed* package in Geographic Resource Analysis Support System (GRASS 7.6.1) originally developed by the U.S. Army Corps of Engineers (Neteler et al., 2012) and the multiple flow direction algorithm (Freeman, 1991; Quinn et al., 1991) for calculating TWI.

2.5.3 Topographic Position Index

The topographic position index (TPI) is computed from a comparison between the elevation of each cell (e_0) in a DEM to the mean elevation ($\bar{\text{e}}$) in a neighbourhood with a specified radius (R), according to Equation 3 and Equation 4 below (J. P. Gallant & Wilson, 2000; Guisan et al., 1999).

$$TPI = \ e_0 - \quad (3)$$

$$z = \ \frac{1}{n_R} \ \sum_i \ _{\text{R}} z_i \ (4)$$

It thus characterizes the topographic position of an area relative to the elevation context, with positive and negative values meaning that a specific area is

located higher or lower compared to it average surrounds, respectively. We used rectangular neighbourhood model of five different sizes (3×3, 5×5, 9×9, 15×15, 30×30) in the calculation.

2.5.4 Curvature

Curvature indicates the degree to which the surface of an area is upwardly convex, upwardly concave or flat relative to the surroundings. It can be a useful indicator for delineating areas that are saturated with water or channelized (Ågren et al., 2014; O'Neil et al., 2018, 2019). For each cell in a DEM, a fourth-order polynomial is fit to a surface consisting of a 3×3 cell window centred on the cell for calculation. The curvature value is then acquired by taking the sum of the second derivatives of this surface (Moore et al., 1991). The curvature calculation was done in ArcGIS Pro 2.8.

2.5.5 Landform Classification

Landforms are conceptual descriptions of geomorphic or terrain features of Earth's surface. Most methods for deriving landform classifications from DEMs are based on characterizing the topographic, spatial and contextual attributes of an area, including elevation, shape, relief, relative position and size, using various algorithms (Mokarram & Sathyamoorthy, 2018). Examples of landform classes include depressions, plains, valleys and plateaus. Knowledge about the spatial distribution of landforms can be useful for inferring hydrological and biogeochemical processes that operate in an area and that shapes its ecological conditions and characteristics (Moore et al., 1991).

Landform maps were produced for each site using the GRASS tool *r.geomorphon* (Jasiewicz & Stepinski, 2013). This tool characterizes the relative elevation of each cell in a DEM based on a comparison with the elevation of at least eight neighbouring cells using a three-value scale (higher, lower and same). The primary output of *r.geomorphon* is a visibility neighbourhood pattern where each cell is assigned one of 498 potential pattens, or geomorphons. The landform of each cell is then determined based on a lookup-table. This tool requires input on the two scalar parameters lookup distance (L) and flatness threshold (t). We set L to 9 and t to the default value 1 based on a previous assessment showing the suitability of these values in the context of high-latitude land cover mapping (Karlson et al., 2021).

2.6 Classification and Accuracy Assessment

For classifying peatland types (bog, fen and swamp), upland classes (forest and other open land) and water, the random forest (RF) algorithm (Breiman, 2001) implemented in ArcGIS Pro 2.8 was used. RF is a supervised machine learning method that has proven to be among the most accurate when used for land cover classification (Belgiu & Drăguţ, 2016; Pelletier et al., 2016), including distinguishing between wetland types (Mahdianpari et al., 2021; Merchant et al., 2020). RF also has several practical benefits, such as allowing both quantitative and categorical predictor variables, having few parameters needing tuning, as

well as minimizing the risk of overfitting the model by using an ensemble of decision trees that are constructed from random subsets of the training data and predictor variables. It also provides metrics for assessing predictor variable importance, which can contribute to improved understanding of the classification model.

All classifications were done on the object level using the reference polygons described in section 2.3. The object average predictor variable values provided the input to the classification models. Classification models were developed using reference data for each of the five sites separately (regionalized models) and by combining the reference data from all sites (global model). We evaluated four classification scenarios that combined predictor variables from different EO data sources: i) Sentinel-1 and terrain derivatives, ii) Sentinel-2 and terrain derivatives, iii) Seninel-1, Sentinel-2 and terrain derivatives, and iv) Sentinel-1 and Sentinel-2, without terrain derivatives. This procedure provided insight to the relative contribution of the EO data sources to the classifications

We used a repeated random sub-sampling cross-validation approach to assess the accuracy of the different classification models and input data scenarios (Richter et al., 2012; Zou et al., 2021). Specifically, 80% randomly selected reference polygons were used for training the classification models and the remaining 20% for validation. This randomized sub-sampling procedure was repeated 100 times to ensure a high likelihood that each reference polygon was included in the validation sub-sample. The summed output statistics from this procedure were used for calculating metrics for the accuracy assessment and for quantifying the relative importance of individual predictor variables. Classification accuracy metrics included overall accuracy (OA), kappa statistic (K; Cohen, 1960) and the class specific producer´s (PA) and user´s accuracy (UA; Story & Congalton, 1986). Predictor variable importance was quantified from the mean decrease in accuracy measure provided by the Random Forest algorithm (Breiman, 2001). This measure quantifies the reduction in classification accuracy that results when excluding each individual predictor variable from the classification model.

3. Results

3.1 Accuracies of Regionalized and Generalized Classification Models

The outcomes from the four classification scenarios representing different combination of EO data sources were highly consistent in terms of OA for both the individual sites (regionalized models) and the generalized classification model (Figure 2). Specifically, the combination of all three data sources (S-1, S-2 and DEM) produced the highest OA in all cases, whereas the second most accurate classification scenario was the S-1 and DEM combination overall. Including terrain derivatives from DEM increased OA between 2.8% (North Slope) and 10.6% (Yukon) for the regionalized classification models, and by 7.6% for the generalized model. Furthermore, the combination of S-1 and DEM consistently outperformed the S-2 and DEM combination, with increases in OA between 0.3% (Hudson Bay Lowlands) and 7.0% (Sweden) for the individual sites, and

a 2.4% increase for the generalized model.

Overall, the results show clear benefits in terms of higher OA when training classification models with regionalized reference data compared to pooling it to produce a generalized classification model. When comparing the most accurate classification scenario (S-1 + S-2 + DEM), OA was between 4.6-12.3% higher for the individual sites compared to the generalized classification model. On the other hand, using the most accurate classification scenario produced clear differences in OA between the individual sites with the highest OA for Sweden (91.9%), followed by Great Slave Lake (88.2%), North Slope (87.2%), Yukon (85.4%) and Hudson Bay Lowlands (84.2%). The generalized classification model resulted in an OA of 79.6%.

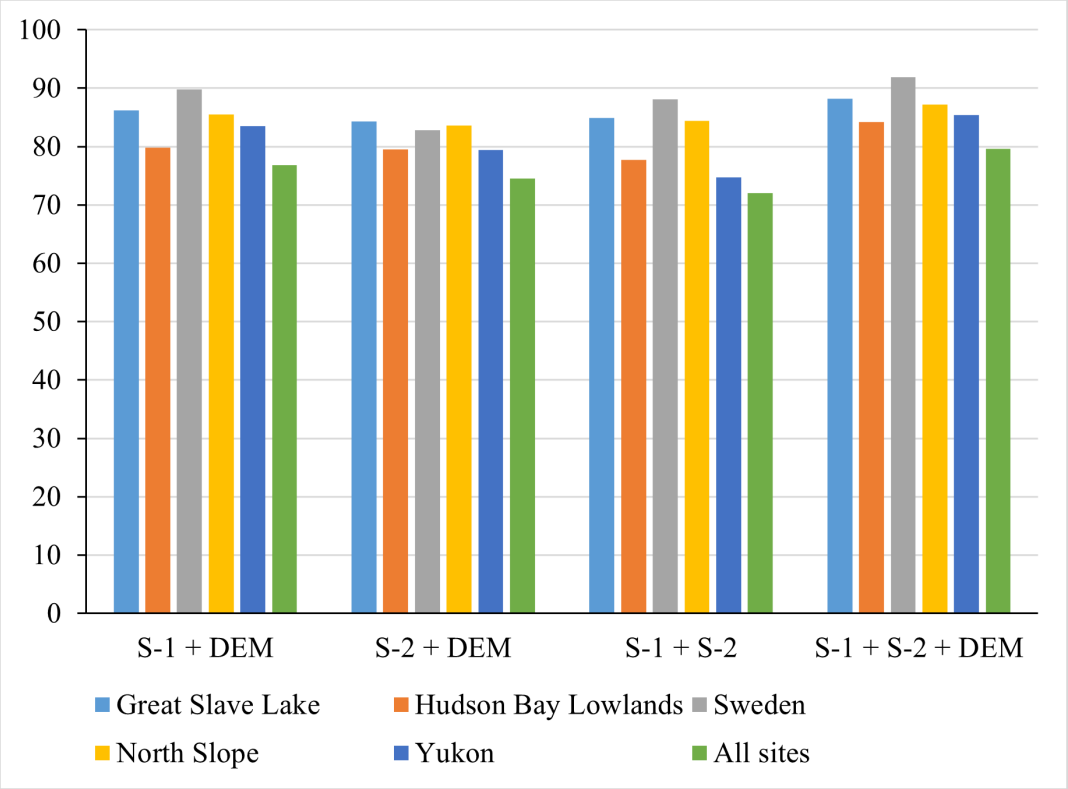

Figure 2. Overall classification accuracy resulting from regionalized and generalized classification models using four classification scenarios based on different EO data sources.

3.2 Class Specific Comparison

The classification accuracies for the peatland and upland classes produced by the different classification scenarios are presented in Table 3. In general, the classification accuracy of peatland types was relatively high, but differed markedly,

both in relation to each other and between the five sites. Bog was the most accurately classified peatland type in Hudson Bay Lowlands, Sweden and Yukon with UA ranging between 85.6% and 91.1%. In Great Slave Lake and North Slope, fens had highest UA of the peatland types at 90.7% and 87.5%, respectively. Swamp had the lowest UA of peatland types in all sites except North Slope and Yukon. Similar to the OA assessment, inclusion of terrain derivatives from DEM had large positive influence on the accuracies for all peatland types, but more for fens and swamps compared to bogs. This pattern was even clearer when assessing peatland type accuracies resulting from the generalized classification model. The latter also resulted in significantly lower peatland type accuracies in relation to those obtained from the regionalized classification models. Unsurprisingly, open water was the most accurate land cover class in all sites, as well as when the generalized classification model was used. The class other open land followed a similar pattern with relatively high UA. For the forest class, UA differed between sites. Lowest UA for the forest class was in Great Slave Lake and Hudson Bay Lowlands where the swamp class had relatively low UA, indicating a confusion between these two classes.

Table 4. Class specific accuracy for the five sites and the generalized classification model. Included are also results from the four classification scenarios, with the most accurate (S-1 + S-2 + DEM) indicated in bold.

@ >p(- 26) * >p(- 26) * >p(- 26) * >p(- 26) * >p(- 26) * >p(- 26) * >p(- 26) * >p(- 26) * >p(- 26) * >p(- 26) * >p(- 26) * >p(- 26) * @ **Site** & **Datasets** & **Bog** & **Fen** & **Swamp** & **Water** & **Other open land** & **Forest** & & & & & &
& & **PA** & **UA** & **PA** & **UA** & **PA** & **UA** & **PA** & **UA** & **PA** & **UA** & **PA** & **UA**

Great Slave Lake

& S-1 + DEM & 85.4 & 86.5 & 88.3 & 88.0 & 74.8 & 70.5 & 99.0 & 100 & 92.5 & 96.5 & 77.3 & 77.1
& S-2 + DEM & 82.5 & 77.5 & 79.9 & 89.3 & 73.3 & 70.8 & 97.8 & 96.5 & 96.8 & 98.7 & 76.8 & 75.4
& S-1 + S-2 + DEM & **87.7** & **89.5** & **90.9** & **90.7** & **74.8** & **72.2** & **99.1** & **100** & **98.9** & **98.4** & **77.5** & **78.4**
& S-1 + S-2 & 83.1 & 86.6 & 80.9 & 85.8 & 72.8 & 65.2 & 98.5 & 100 & 97.1 & 97.8 & 75.7 & 74.4

Hudson Bay lowlands

& S-1 + DEM & 80.2 & 77.6 & 69.0 & 72.9 & 71.8 & 66.2 & 100 & 100 & 73.9 & 78.5 & 83.0 & 83.0
& S-2 + DEM & 83.0 & 74.9 & 67.7 & 73.2 & 66.2 & 63.0 & 99.8 & 100 & 80.3 & 84.8 & 79.6 & 79.4
& S-1 + S-2 + DEM & **86.3** & **85.6** & **71.6** & **79.1** & **78.8** & **69.7** & **100** &

100 & **84.0** & **83.4** & **82.7** & **85.3**
& S-1 + S-2 & 72.4 & 81.8 & 71.4 & 61.9 & 70.4 & 60.9 & 100 & 100 & 78.8 & 80.0 & 72.8 & 81.2

Sweden

& S-1 + DEM & 90.0 & 86.1 & 86.4 & 90.3 & 83.8 & 83.2 & 99.0 & 99.7 & 95.1 & 95.8 & 85.1 & 84.2
& S-2 + DEM & 79.7 & 78.3 & 78.2 & 76.9 & 65.9 & 65.2 & 98.7 & 99.0 & 97.3 & 96.9 & 74.8 & 78.1
& S-1 + S-2 + DEM & **92.3** & **91.1** & **88.9** & **89.6** & **89.5** & **82.2** & **97.0** & **100** & **98.3** & **97.4** & **84.2** & **90.3**
& S-1 + S-2 & 90.7 & 89.0 & 84.3 & 89.3 & 80.7 & 72.0 & 99.5 & 99.2 & 99.0 & 99.2 & 75.4 & 80.5

North Slope

& S-1 + DEM & 75.5 & 71.7 & 81.6 & 84.3 & 74.4 & 85.6 & 99.8 & 99.3 & 92.1 & 85.7 & 89.2 & 84.9
& S-2 + DEM & 75.7 & 73.1 & 76.6 & 75.8 & 75.8 & 78.5 & 98.1 & 97.6 & 90.1 & 90.3 & 87.2 & 88.1
& S-1 + S-2 + DEM & **79.1** & **74.6** & **84.0** & **87.5** & **76.9** & **86.9** & **99.5** & **97.2** & **93.3** & **87.6** & **92.3** & **89.8**
& S-1 + S-2 & 71.4 & 74.0 & 80.8 & 79.6 & 75.8 & 78.5 & 99.0 & 97.6 & 92.4 & 85.0 & 86.8 & 90.3

Yukon

& S-1 + DEM & 84.8 & 82.9 & 71.0 & 73.2 & 74.9 & 79.4 & 97.4 & 96.7 & 91.6 & 83.9 & 83.7 & 85.4
& S-2 + DEM & 80.5 & 68.0 & 59.7 & 63.5 & 63.1 & 70.9 & 95.8 & 95.5 & 92.1 & 90.3 & 87.5 & 87.5
& S-1 + S-2 + DEM & **86.2** & **81.7** & **73.2** & **74.1** & **70.1** & **76.3** & **97.7** & **95.2** & **90.8** & **90.8** & **84.4** & **82.2**
& S-1 + S-2 & 72.3 & 76.5 & 57.3 & 64.3 & 59.9 & 54.2 & 98.2 & 93.3 & 87.2 & 87.6 & 73.3 & 70.2

All

& S-1 + DEM & 71.3 & 69.6 & 67.1 & 69.7 & 67.4 & 65.7 & 97.6 & 97.6 & 79.6 & 80.8 & 77.8 & 77.5
& S-2 + DEM & 70.3 & 64.5 & 63.8 & 68.7 & 57.8 & 54.6 & 98.8 & 97.8 & 85.6 & 86.6 & 69.2 & 73.4
& S-1 + S-2 + DEM & **75.0** & **72.3** & **69.8** & **75.3** & **69.4** & **66.5** & **98.8** & **97.6** & **87.7** & **87.2** & **76.8** & **78.4**
& S-1 + S-2 & 61.4 & 67.7 & 61.3 & 56.9 & 57.7 & 53.8 & 99.0 & 97.5 & 85.6

& 82.1 & 68.2 & 74.1

In Figure 3, S-1 backscatter in VH and VV polarization from the three peatland types are plotted over time for the year 2020 to show potential differences in the seasonal dynamics in vegetation structure and soil moisture condition. The temporal pattern of the VH backscatter response from peatland types show strong similarities for all sites except Great Slave Lake, with clear peaks in April around when snow melting starts, which results in more heterogeneous ground conditions and higher backscatter. The backscatter peak is followed by the growing period with relatively high backscatter levels for all peatland types resulting from presence of vegetation and higher soil moisture. In Great Slave Lake, the backscatter peak of all peatland types occurs later in the growing season compared to the other sites. For the VV polarization, the temporal pattern is also similar between sites, but peak backscatter generally occurs later in the growing season compared to VH. Overall, peak backscatter levels for the peatland types differ markedly between sites, especially in VH polarization, with differences of up to 5 dB, suggesting significant difference in vegetation composition and surface condition, including soil moisture regimes.

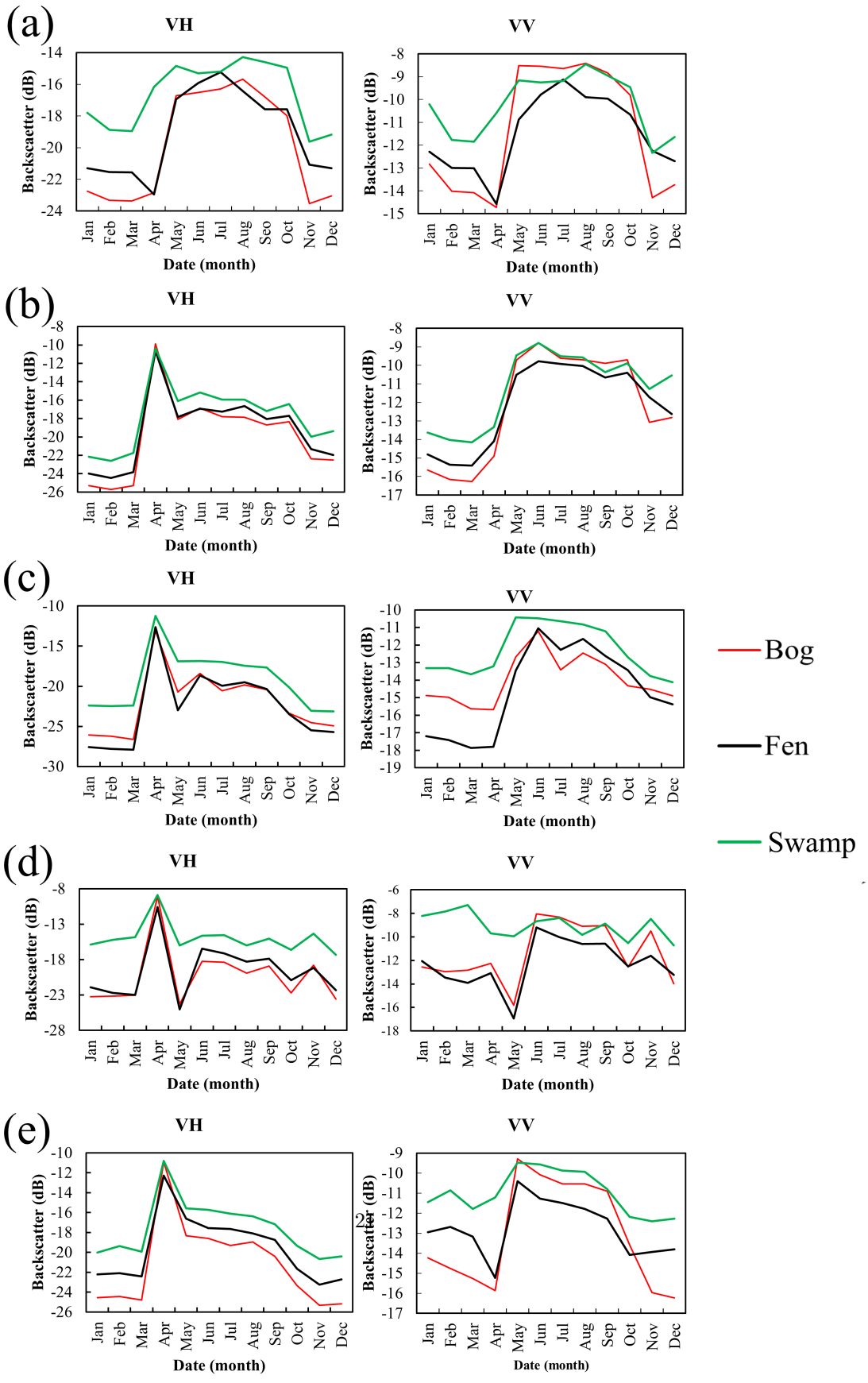
(a)
VH
VV
(b)
VH
VV
(c)
VH
VV
(d)
VH
VV
(e)
VH
VV
Backscatter (dB)
Date (month)
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Seo
Oct
Nov
Dec
Bog
Fen
Swamp

Figure 3. Seasonal patterns in S-1 backscatter from three peatland types in the five sites. (a) Great Slave Lake, (b) Hudson Bay lowlands, (c) North Slope, (d) Sweden, and (e) Yukon.

3.3 Predictor Variable Importance

The relative contribution of the 20 most important predictor variables to the most accurate classification models (S1 + S2 + DEM) for the individual sites and the generalized model are shown in Figure 4. A noticeable result is that the landform classification was the single most important predictor for all five sites in the regionalized classification models, but it was not even among the top 20 predictors in the generalized classification model. NDWI also ranked high for all sites except Yukon. In addition, predictors based on S-2 had especially high importance in Great Slave Lake, Hudson Bay Lowlands and North Slope, whereas predictors based on S-1 were generally ranked higher in Sweden and Yukon. Except for landform classification, terrain derivatives were not ranked high in predictor importance in any of the sites, except for Hudson Bay Lowlands where several versions of DTW were among those ranked among the top 20 predictors. The ranking of predictor importance shows a clearly different pattern for the generalized classification model. Specifically, this model showed smaller differences between predictors in terms of their contribution to the classification results, and no terrain derivative was ranked among the 20 most important.

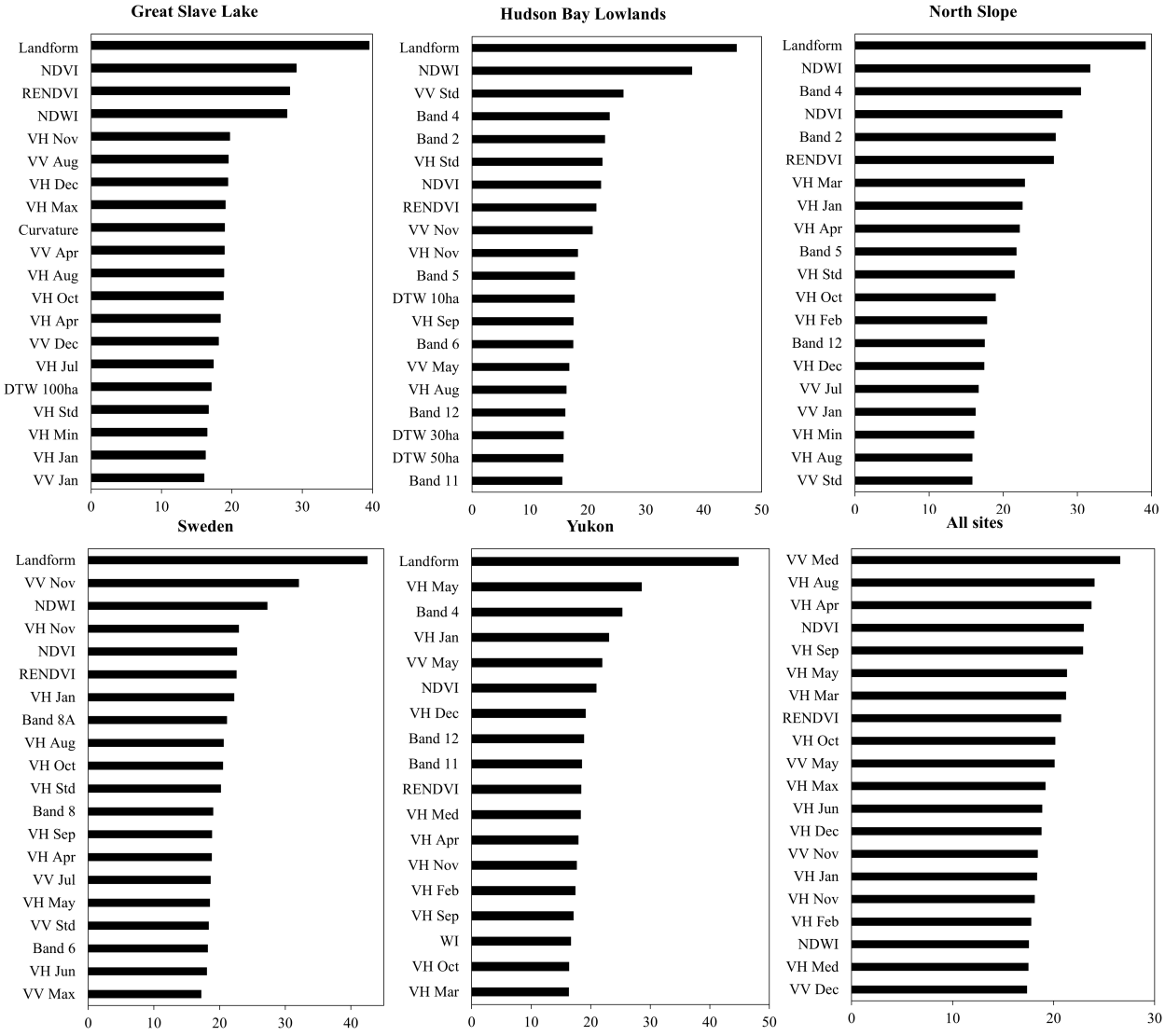

Figure 4. Contribution of the top 20 predictor variables to the most accurate classification scenario (S-1 + S-2 + DEM) for the five sites and the generalized classification model.

4. Discussion

Distinguishing detailed wetland types using EO data is known to be a challenge, especially when the objective is large area mapping (Ozesmi & Bauer, 2002). This is manifested in the available land cover products with global or continental scale coverage that generally only include an aggregated wetland class at

best (Phiri et al., 2020). In such products, the wetland class is often also among those with the lowest classification accuracy (Hermosilla et al., 2022). A key challenge for mapping wetland types is that features used to distinguish them in the field, including vegetation composition, hydrological characteristics and soil properties, usually are difficult to resolve at the scale at which satellite systems suitable for large-area applications make observations. Another important challenge is that wetland types can differ between regions, in particular vegetation composition and soil moisture regimes, depending on local geographical conditions, including topography and climate (Mahdianpari et al., 2021).

4.1 Influence of EO Data Sources on Classification Accuracy

Selecting between EO data sources or data fusion strategies is a key consideration for land cover mapping activities that affects both the capabilities to detect relevant land cover features and data processing requirements. Such a choice also controls the spatiotemporal scales covered by the mapping activity and the quality of the resulting map product, including classification accuracy. This study focused on EO data from Sentinel-1, Sentinel-2 and Copernicus DEM mainly because of the consistent global coverage and relatively high spatiotemporal resolution that potentially could enable continental scale mapping of high-latitude peatland types, given sufficient availability of reference data to train classification models, as well as EO data processing capabilities.

The peatland type and upland classifications produced in this study demonstrated varying accuracies that were highly dependent on the EO data sources used for input to the classification models. Our results show that significant improvements in both OA and class-specific accuracies (UA and PA) were achieved for all five sites, as well as the generalized model, when using the classification scenario that combined predictor variables from S-1 and S-2, and terrain derivatives, albeit at varying levels. These results add to the growing body of research showing that combining or fusing different types of EO data sources, including optical, SAR and elevation data, generally improves land cover classification accuracy in high-latitude regions (Hermosilla et al., 2022). This is especially the case for applications targeting detailed classes such as wetland or peatland types (Karlson et al., 2019; Mahdianpari et al., 2021; Merchant et al., 2020). The improved accuracies suggest that these three EO data sources contribute with complementary information to the classification that each characterizes different features of both peatland types and upland classes that facilitates their separation. The links between predictor variables based on optical, SAR and elevation data, and vegetation properties, soil moisture conditions, and local hydrology, respectively, are well established (A. L. Gallant, 2015; Ozesmi & Bauer, 2002). However, the accuracy improvements (4.6-12.3% increase) resulting from their combination in this study are notable when compared to previous research. For example in Yukon, Canada, Merchant et al. (2020) reported smaller improvements in accuracy when adding terrain derivatives to a classification model based on optical and SAR data, but also that it improved the visual appearance of the resulting map by reducing edge effects (Merchant

et al., 2020). Similar benefits of improved map appearance, but also increasing classification accuracy (+6% in OA), resulting from including terrain derivatives were also reported by Karlson et al. (2019) focusing on northern Sweden. Furthermore, Mahdianpari et al. (2021) found a 2% increase in OA when adding terrain derivatives in the production of a Canada wide wetland map. Potential explanations to the relatively high accuracy improvements achieved in this study could be that we used another set of terrain derivatives that may better capture local topographic and hydrologic features that are more closely related to peatland occurrence (Mahdianpari et al., 2021), and the generally high quality of the Copernicus DEM (Karlson et al., 2021).

Class specific accuracies of both peatland types and upland classes reported here are relatively high in general, but they also differ clearly between the five sites. Highest accuracies for the three peatland types were achieved in the Great Slave Lake and Sweden sites. Such between-site differences in class specific accuracies can be the result of regional peatland characteristics producing more separable signatures in EO data in some sites facilitating their accurate mapping (Mahdianpari et al., 2021). Another possible explanation is differences in the quality of the reference data used for training the regionalized classification models (Hermosilla et al., 2022; Mahdianpari, Brisco, et al., 2020; Mellor et al., 2015). Specifically, the reference data in Great Slave Lake and Sweden were drawn from consistent datasets with full spatial coverage of the sites, produced from field surveys and manual interpretation of high-resolution imagery conducted by local experts. We judge these reference data to be of higher quality than the other sites where several diverse sources of reference data with potential limitations in both quality and spatial extent needed to be used to produce a sufficiently large sample. We believe that the visual interpretation of high-resolution imagery as a quality check employed in this study improved the reference data quality in these sites, but also acknowledge that the lower reference data quality may be a source of error in the classifications. Availability of high-quality reference data needed for detailed land cover mapping, including peatland type classification, is limited in many parts of high-latitude regions. This presents a key challenge that needs to be addressed in the pursuit of continental scale map products.

The use of different EO data as input to the classification models also affected peatland accuracy differently between sites. A notable example is that peatland type accuracies benefited more from including terrain derivatives at the Hudson Bay Lowlands and North Slope sites, compared to the other ones. Both these sites are characterized by relatively flat landscapes located on coastal plains, high occurrence of peatlands and low tree cover. While the reasons for this effect are difficult to pinpoint, we speculate that an important factor is the higher capability of Copernicus DEM to characterize flatter areas accurately compared to rugged terrain (Karlson et al., 2021), thereby also better capturing topographic features that separate peatland types in this geographical setting. Swamp was the peatland type that benefited most in accuracy when including terrain derivatives in the classification model, largely because of the improved separation from the forest class. This effect is expected since these two classes

usually occupy topographically different areas in the landscape that can be captured by terrain derivatives, such as landform classes. A drawback with Copernicus DEM is the relatively coarse spatial resolution of 30 m. This substantially increases the minimal mapping unit that could be achieved using only the higher spatial resolution S-1 and S-2 data and can result in small and narrow peatland areas being missed in the map product. The Arctic DEM (Porter et al., 2018) is an interesting higher resolution option to Copernicus DEM, but this product contains frequent data voids that needs to be managed using a gap-filling technique before using it for generating terrain derivatives.

4.2 Regional and Global Classification Models

Our study further shows the clear benefits of training classification models with regionalized reference data compared to pooling it into a global model for classifying large areas with potentially high variability in geographical conditions. The regionalized classification models applied at the five sites consistently produced superior classification accuracies, both for peatland types and upland classes. These results are in line with those of Mahdianpari et al. (2021) and Hermosilla et al. (2021) who used contrasting stratification approaches in their Canada-wide mapping projects, including stratifications based on eco-regions and a tiling system (150 $\times$ 150 km grid), respectively. Irrespective of the approach applied, a stratified implementation ensures, albeit at varying degrees, that classification models are trained using reference data that are representative of local land cover characteristics and conditions (Townshend et al., 2012). Hence, a main benefit with a stratified model implementation is that it compensates for potential regional variability in spectral and terrain signatures of land cover classes that can negatively affect classification accuracies. Another benefit with regionalized classification models is that they ensures the use of predictor variables with the highest potential to characterize local land cover characteristics (Hermosilla et al., 2022). Future research should focus on finding the most effective strategy for stratifying large areas in high-latitude regions to facilitate the mapping of wetland and peatland types, but also other detailed land cover classes. While both the eco-region and tiling approaches have shown promise in the context of wetland mapping, another potential way is to base the stratification on is the 'wetscapes' concept introduced by Olefeldt et al. (2021). This concept is based on grouping and mapping spatially co-occurring wetland classes that have similarities in biogeochemical functioning, such as methane fluxes, into regions with characteristic wetland class compositions.

4.3 Regional Variations in Predictor Variable Importance

Quantifying predictor variable importance can facilitate the construction of parsimonious classification models and provides the basis for limiting data processing requirements by identifying redundant predictor variables. In our results, predictor variable importance differed largely both between the regionalized and global classification model, as well as between the five sites, but some general observations can be made. Specifically, landform classification was ranked as the most important predictor variable for all five sites but did not contribute

nearly as much to the global classification model (ranked as number 21). The relatively low importance in the global classification model suggests that the peatland types, as well as the upland classes, tend to occupy different landforms at least in some of the sites. The five sites are characterized by largely different landform compositions because of local topography that can control peatland formation. For example, the slope class dominates in North Slope, Sweden and Yukon, whereas its extent is more limited in Great Slave Lake and Hudson Bay Lowlands. Peatland types can also be characterized by difference in their respective surface topography, slope type and landscape position (Tiner 2014), but this finer level classification was not included in the reference data used in this study. Examples of such detailed peatland types include raised bog (dome-shaped topography), blanked bog (formed on slope and extends over rolling terrain), hillslope fen (located on the side of a hill) and stream valley slope fens (located in a stream valley, typically in mountainous terrain). These local topographic surface characteristics and landscape positions of peatland types are to a large extent site specific and controlled by a combination of climatic conditions, geological factors and local topography (Gorham, 1957; Vitt, 1994). This potential between-site variability may not be apparent in reference datasets but provides further motivation for using regionalized classification models. In addition to landform classification, no terrain derivative provided significant contribution to either the regionalized or global classification models, which suggests that many are redundant and can be left out from the DEM processing.

Spectral vegetation indices derived from the growing season mosaics, including NDVI, NDWI and RENDVI, generally contributed more to the classifications compared to the individual S-2 bands for both the regionalized models and the global model. However, both the vegetation indices and the individual bands ranked highest are based on wavelengths that extends from the red to the SWIR portion of the electromagnetic spectrum (0.665 to 2.190 m). The ability of these wavelengths for identifying peatland types, as well as other upland classes, in similar high-latitude environment settings have been demonstrated in previous research (Mahdianpari, Brisco, et al., 2020; Merchant et al., 2017, 2020). Concerning the S-1 based predictor variables, monthly observations contributed more to the regionalized models compared to the temporal metrics, but the specific months and polarizations differed largely between the five sites. This provides further evidence of high between site variability in spectral and temporal characteristics of both peatland types and upland classes, possibly due to variations in vegetation phenology and soil moisture. Consequently, these results also point to the benefits that can be achieved when using a regionalized mapping approach.

5. Conclusions

The main objective of this study was to compare the performance of peatland type classification models based on different EO data inputs and strategies for model implementation in five contrasting high-latitude landscapes. Our results

provide clear indications that regionalized classification models represents a path for more accurate EO-based land cover assessments, at the high class-detail and spatial resolution required to link land cover with various biogeochemical processes, including C sequestration and GHG emissions. Such regionalized classification models enable taking into account regionally specific characteristics of peatland types and can optimize the use of limited reference data. Our results further demonstrate the benefits of combining different types of EO data, including optical and SAR data and terrain derivatives, in classification models to better separate detailed land cover classes. A key task for future research is to identify an effective strategy for stratifying high-latitude areas for effective classification model implementation. Another remaining challenges includes ensuring availability of detailed reference data for classification model training and validation that are representative for all main regions in the high-latitude circumpolar regions and increasing the spatial resolution enough to cover small but biogeochemically important areas in landscapes. However, the rapid development towards higher resolution EO systems and low-cost airborne areal imaging (e.g., with drones) is promising and increases the likelihood that dynamic high-resolution continental scale mapping of detailed land cover classes is soon possible.

Acknowledgements

The authors would like to acknowledge the support from the Swedish Research Council Formas under the grants 2017-01944, 2018-01794, 2018-00570. We also would like to acknowledge the European Space Agency and the Copernicus Programme for provision of Earth Observation datasets. No potential conflict of interest was reported by the author(s).

Data Availability Statement

The Earth Observation (EO) and reference data used for land cover mapping in the five sites will be available in a digital repository by acceptance and DOI will be provided. This dataset contains pre-processed EO data (Sentinel-1, Sentinel-2 and terrain derivatives from Copernicus DEM) and reference data used in the study. Further details about and citations to the different reference data sources are provided in section 2.2. (Study Areas and Reference Data Source).

References

Abraham, K. F., & Keddy, C. J. (2005). The Hudson Bay Lowlands. In L. . Fraser & P. . Keddy (Eds.), *The World's Largest Wetlands: Ecology and Conservation* (pp. 118–148). Cambridge, U.K.: Cambridge University Press.Ågren, A. M., Lidberg, W., Strömgren, M., Ogilvie, J., & Arp, P. A. (2014). Evaluating digital terrain indices for soil wetness mapping – A Swedish case study. *Hydrol. Earth Syst. Sci.*, *18*(9), 3623–3634. https://doi.org/10.5194/hess-18-3623-2014Bäckstrand, K., Crill, P. M., Jackowicz-Korczyñski, M., Mastepanov, M., Christensen, T. R., & Bastviken, D. (2010). Annual carbon gas budget for a subarctic peatland, Northern Sweden. *Biogeosciences*, *7*(1), 95–108. https://doi.org/10.5194/bg-7-95-2010Bao,

T., Jia, G., & Xu, X. (2021). Wetland heterogeneity determines methane emissions: A pan-Arctic synthesis. *Environmental Science & Technology*, *55*(14), 10152–10163. https://doi.org/10.1021/acs.est.1c01616Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 24–31. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2016.01.011Beven, K. J., & Kirby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, *24*(1), 43–69.Bourgeau-Chavez, L. L., Endres, S., Powell, R., Battaglia, M. J., Benscoter, B., Turetsky, M., et al. (2016). Mapping boreal peatland ecosystem types from multitemporal radar and optical satellite imagery. *Canadian Journal of Forest Research*, *47*(4), 545–559. https://doi.org/10.1139/cjfr-2016-0192Bourgeau-Chavez, L. L., Graham, J. A., Endres, S., French, N. H. F., Battaglia, M., Hansen, D., & Tanzer, D. (2019). ABoVE: Ecosystem Map, Great Slave Lake Area, Northwest Territories, Canada, 1997-2011 [Dataset]. ORNL Distributed Active Archive Center. https://doi.org/10.3334/ORNLDAAC/1695Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324Bruhwiler, L., Parmentier, F.-J. W., Crill, P., Leonard, M., & Palmer, P. I. (2021). The Arctic carbon cycle and its response to changing climate. *Current Climate Change Reports*, *7*(1), 14–34. https://doi.org/10.1007/s40641-020-00169-5Buchhorn, M., Lesiv, M., Tsendbazar, N.-E., Herold, M., Bertels, L., & Smets, B. (2020). Copernicus Global Land Cover Layers—Collection 2. *Remote Sensing*. https://doi.org/10.3390/rs12061044Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.Ducks Unlimited Canada. (2022). Canadian Wetland Inventory [Dataset]. Retrieved from https://www.ducks.ca/initiatives/canadian-wetland-inventory/Fahrland, E., Jacob, P., Schrader, H., & Kahabka, H. (2020). *Copernicus digital elevation model - Product handbook*. Potsdam, Germany. Retrieved from https://spacedata.copernicus.eu/documents/20126/0/GEO1988-CopernicusDEM-SPE-002_ProductHandbook_I1.00.pdfFick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086Flood, N. (2013). Seasonal composite Landsat TM/ETM+ images using the medoid (a multi-dimensional median). *Remote Sensing*. https://doi.org/10.3390/rs5126481Foody, G. M., Pal, M., Rocchini, D., Garzon-Lopez, C. X., & Bastin, L. (2016). The sensitivity of mapping methods to reference data quality: Training supervised image classifications with imperfect reference data. *ISPRS International Journal of Geo-Information*. https://doi.org/10.3390/ijgi5110199Forkuor, G., Dimobe, K., Serme, I., & Tondoh, J. E. (2018). Landsat-8 vs. Sentinel-2: examining the added value of sentinel-2's red-edge bands to land-use and land-cover mapping in Burkina Faso. *GIScience & Remote Sensing*, *55*(3), 331–354. https://doi.org/10.1080/15481603.2017.1370169Franklin, S. E., Skeries, E. M., Stefanuk, M. A., & Ahmed, O. S. (2018). Wetland classification using Radarsat-2 SAR quad-polarization and Landsat-8 OLI spectral response data: A case study in the Hudson Bay Lowlands

Ecoregion. *International Journal of Remote Sensing*, *39*(6), 1615–1627. https://doi.org/10.1080/01431161.2017.1410295Frantz, D., Röder, A., Stellmes, M., & Hill, J. (2017). Phenology-adaptive pixel-based compositing using optical Earth observation imagery. *Remote Sensing of Environment*, *190*, 331–347. https://doi.org/https://doi.org/10.1016/j.rse.2017.01.002Freeman, T. G. (1991). Calculating catchment area with divergent flow based on a regular grid. *Computers & Geosciences*, *17*(3), 413–422. https://doi.org/https://doi.org/10.1016/0098-3004(91)90048-IGallant, A. L. (2015). The challenges of remote monitoring of wetlands. *Remote Sensing*, *7*(8). https://doi.org/10.3390/rs70810938Gallant, J. P., & Wilson, J. C. (2000). Primary topographic attributes. In J. P. Wilson & J. C. Gallant (Eds.), *Terrain Analysis: Principles and Applications* (pp. 51–85). Hoboken, New Jersey: John Wiley.Gao, B. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, *58*(3), 257–266. https://doi.org/https://doi.org/10.1016/S0034-4257(96)00067-3Gisnås, K., Etzelmüller, B., Lussana, C., Hjort, J., Sannel, A. B. K., Isaksen, K., et al. (2017). Permafrost map for Norway, Sweden and Finland. *Permafrost and Periglacial Processes*, *28*(2), 359–378. https://doi.org/https://doi.org/10.1002/ppp.1922Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/https://doi.org/10.1016/j.rse.2017.06.031Gorham, E. (1957). The development of peat lands. *The Quarterly Review of Biology*, *32*(2), 145–166. Retrieved from http://www.jstor.org/stable/2816118Gorham, E. (1991). Northern peatlands: Role in the carbon cycle and probable responses to climatic warming. *Ecological Applications*, *1*(2), 182–195. https://doi.org/https://doi.org/10.2307/1941811Government of Yukon. (2017). Vegetation Inventory - 40K [Dataset]. Retrieved from https://mapservices.gov.yk.ca/arcgis/rest/services/GeoYukon/GY_Forestry/MapServer/23Government of Yukon. (2018). *Mapping and classifying wetlands in the Indian River Valley, Yukon [Dataset]*. Retrieved from https://yukon.ca/sites/yukon.ca/files/emr/emr-forms/emr-mapping-classifying-wetlands-indian-river-valley-yukon-final-report.pdfGrabs, T., Seibert, J., Bishop, K., & Laudon, H. (2009). Modeling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model. *Journal of Hydrology*, *373*(1), 15–23. https://doi.org/https://doi.org/10.1016/j.jhydrol.2009.03.031Greaves, H. E., Eitel, J., Vierling, L., Boelman, N., Griffin, K., Magney, T., & Prager, C. (2019). High-resolution vegetation community maps, Toolik Lake area, Alaska, 2013-2015 [Dataset]. ORNL Distributed Active Archive Center. https://doi.org/10.3334/ORNLDAAC/1690Guisan, A., Weiss, S. B., & Weiss, A. D. (1999). GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology*, *143*(1), 107–122. https://doi.org/10.1023/A:1009841519580Gunnarsson, U., & Löfroth, M. (2009). *Våtmarkinventeringen - resultat från 25 års inventeringar*. Stockholm, Sweden. Retrieved from https://www.naturvardsverket.se/globalassets/media/publikationer-pdf/5900/978-91-620-5925-5.pdfGuo, M., Li, J., Sheng, C., Xu, J., & Wu, L.

(2017). A review of wetland remote sensing. *Sensors*. https://doi.org/10.3390/s17040777Hermosilla, T., Wulder, M. A., White, J. C., & Coops, N. C. (2022). Land cover classification in an era of big and open data: Optimizing localized implementation and training data selection to improve mapping outcomes. *Remote Sensing of Environment*, *268*, 112780. https://doi.org/https://doi.org/10.1016/j.rse.2021.112780Hinzman, L. D., Bettez, N. D., Bolton, W. R., Chapin, F. S., Dyurgerov, M. B., Fastie, C. L., et al. (2005). Evidence and implications of recent climate change in northern Alaska and other Arctic regions. *Climatic Change*, *72*(3), 251–298. https://doi.org/10.1007/s10584-005-5352-2Hugelius, G., Loisel, J., Chadburn, S., Jackson, R. B., Jones, M., MacDonald, G., et al. (2020). Large stocks of peatland carbon and nitrogen are vulnerable to permafrost thaw. *Proceedings of the National Academy of Sciences*, *117*(34), 20438 LP – 20446. https://doi.org/10.1073/pnas.1916387117Intergovernmental Panel on Climate Change. (2014). *Climate Change 2014 – Impacts, Adaptation and Vulnerability: Part B: Regional Aspects: Working Group II Contribution to the IPCC Fifth Assessment Report: Volume 2: Regional Aspects* (Vol. 2). Cambridge, UK: Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9781107415386Jasiewicz, J., & Stepinski, T. F. (2013). Geomorphons — a pattern recognition approach to classification and mapping of landforms. *Geomorphology*, *182*, 147–156. https://doi.org/https://doi.org/10.1016/j.geomorph.2012.11.005Jorgenson, T., Yoshikawa, K., Kanevskiy, M., Shur, Y., Romanovsky, V., Marchenko, S., et al. (2008). Permafrost characteristics of Alaska. In D. . Kane & K. . Hinkel (Eds.), *Proceedings of the Ninth International Conference on Permafrost* (pp. 121–122). Fairbanks, Alaska: Institute of Northern Engineering, University of Alaska Fairbanks.Kåresdotter, E., Destouni, G., Ghajarnia, N., Hugelius, G., & Kalantari, Z. (2021). Mapping the vulnerability of Arctic wetlands to global warming. *Earth's Future*, *9*(5), e2020EF001858. https://doi.org/https://doi.org/10.1029/2020EF001858Karlson, M., Gålfalk, M., Crill, P., Bousquet, P., Saunois, M., & Bastviken, D. (2019). Delineating northern peatlands using Sentinel-1 time series and terrain indices from local and regional digital elevation models. *Remote Sensing of Environment*, *231*, 111252. https://doi.org/10.1016/j.rse.2019.111252Karlson, M., Bastviken, D., & Reese, H. (2021). Error characteristics of pan-Arctic digital elevation models and elevation derivatives in northern Sweden. *Remote Sensing*. https://doi.org/10.3390/rs13224653Kirches, G. (2020). *Algorithm Theoretical Basis Document - Copernicus Sentinel-2 Global Mosaic (S2GM) within the Global Land Component of the Copernicus Land Service*. Retrieved from https://land.copernicus.eu/imagery-in-situ/global-image-mosaics/sites/default/files/S2GM-ATBD-BC-v1.3.3.pdfKirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J. G., Dlugokencky, E. J., et al. (2013). Three decades of global methane sources and sinks. *Nature Geoscience*, *6*, 813. Retrieved from http://dx.doi.org/10.1038/ngeo1955Kreplin, H. N., Santos Ferreira, C. S., Destouni, G., Keesstra, S. D., Salvati, L., & Kalantari, Z. (2021). Arctic wetland system dynamics under climate warming. *WIREs Water*, *8*(4), e1526. https://doi.org/https://doi.org/10.1002/wat2.1526Lidberg,

W., Nilsson, M., & Ågren, A. (2020). Using machine learning to generate high-resolution wet area maps for planning forest management: A study in a boreal forest landscape. *Ambio*, *49*(2), 475–486. https://doi.org/10.1007/s13280-019-01196-9Mahdianpari, M., Salehi, B., Mohammadimanesh, F., & Motagh, M. (2017). Random forest wetland classification using ALOS-2 L-band, RADARSAT-2 C-band, and TerraSAR-X imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *130*, 13–31. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2017.05.010Mahdianpari, M., Salehi, B., Mohammadimanesh, F., Brisco, B., Homayouni, S., Gill, E., et al. (2020). Big data for a big country: The first generation of Canadian wetland inventory map at a spatial resolution of 10-m using Sentinel-1 and Sentinel-2 data on the Google Earth Engine cloud computing platform. *Canadian Journal of Remote Sensing*, *46*(1), 15–33. https://doi.org/10.1080/07038992.2019.1711366Mahdianpari, M., Brisco, B., Granger, J. E., Mohammadimanesh, F., Salehi, B., Banks, S., et al. (2020). The second generation Canadian wetland inventory map at 10 meters resolution using Google Earth Engine. *Canadian Journal of Remote Sensing*, *46*(3), 360–375. https://doi.org/10.1080/07038992.2020.1802584Mahdianpari, M., Brisco, B., Granger, J., Mohammadimanesh, F., Salehi, B., Homayouni, S., & Bourgeau-Chavez, L. (2021). The third generation of pan-Canadian wetland map at 10 m resolution using multisource Earth observation data on cloud computing platform. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 8789–8803. https://doi.org/10.1109/JSTARS.2021.3105645Mao, D., Wang, Z., Du, B., Li, L., Tian, Y., Jia, M., et al. (2020). National wetland mapping in China: A new product resulting from object-based and hierarchical classification of Landsat 8 OLI images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *164*, 11–25. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2020.03.020Mattivi, P., Franci, F., Lambertini, A., & Bitelli, G. (2019). TWI computation: a comparison of different open source GISs. *Open Geospatial Data, Software and Standards*, *4*(1), 6. https://doi.org/10.1186/s40965-019-0066-yMellor, A., Boukir, S., Haywood, A., & Jones, S. (2015). Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin. *ISPRS Journal of Photogrammetry and Remote Sensing*, *105*, 155–168. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2015.03.014Melton, J. R., Wania, R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., et al. (2013). Present state of global wetland extent and wetland methane modelling: conclusions from a model inter-comparison project (WETCHIMP). *Biogeosciences*, *10*(2), 753–788. https://doi.org/10.5194/bg-10-753-2013Merchant, M., Adams, J. R., Berg, A. A., Baltzer, J. L., Quinton, W. L., & Chasmer, L. E. (2017). Contributions of C-band SAR data and polarimetric decompositions to subarctic boreal peatland mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(4), 1467–1482. https://doi.org/10.1109/JSTARS.2016.2621043Merchant, M., Haas, C., Schroder, J., Warren, R., & Edwards, R. (2020). High-latitude wetland

mapping using multidate and multisensor Earth observation data: A case study in the Northwest Territories. *Journal of Applied Remote Sensing*, *14*(3), 1–18. https://doi.org/10.1117/1.JRS.14.034511Mokarram, M., & Sathyamoorthy, D. (2018). A review of landform classification methods. *Spatial Information Research*, *26*(6), 647–660. https://doi.org/10.1007/s41324-018-0209-8Moore, I. D., Grayson, R. B., & Ladson, A. R. (1991). Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, *5*(1), 3–30. https://doi.org/https://doi.org/10.1002/hyp.3360050103Murphy, P. N. C., Ogilvie, J., Castonguay, M., Zhang, C., Meng, F.-R., & Arp, P. A. (2008). Improving forest operations planning through high-resolution flow-channel and wet-areas mapping. *The Forestry Chronicle*, *84*(4), 568–574. https://doi.org/10.5558/tfc84568-4Murphy, P. N. C., Ogilvie, J., & Arp, P. (2009). Topographic modelling of soil moisture conditions: A comparison and verification of two models. *European Journal of Soil Science*, *60*(1), 94–109. https://doi.org/https://doi.org/10.1111/j.1365-2389.2008.01094.xNaturvårdsverket. (2019). Våtmarksinventeringen VMI: Nydigitaliserad information [Dataset]. Retrieved from https://gpt.vic-metria.nu/data/land/VMI/Naturvårdsverket. (2020). *Nationella marktäckesdata 2018 baskikt [Dataset]*. Stockholm, Sweden. Retrieved from https://www.naturvardsverket.se/upload/sa-mar-miljon/kartor/NMD-Produktbeskr-NMD2018Basskikt-v2-2.pdfNawrocki, T. W. (2021). Continuous foliar cover of vegetation for North American Beringia (1.0) [Dataset]. Zendo. https://doi.org/https://doi.org/10.5281/zenodo.4770218Nawrocki, T. W., Carlson, M. L., Osnas, J. L. D., Trammell, E. J., & Witmer, F. D. W. (2020). Regional mapping of species-level continuous foliar cover: Beyond categorical vegetation mapping. *Ecological Applications*, *30*(4), e02081. https://doi.org/https://doi.org/10.1002/eap.2081Neteler, M., Bowman, M. H., Landa, M., & Metz, M. (2012). GRASS GIS: A multi-purpose open source GIS. *Environmental Modelling & Software*, *31*, 124–130. https://doi.org/https://doi.org/10.1016/j.envsoft.2011.11.014Nilsson, M., Ahlkrona, E., Jönsson, C., & Allard, A. (2020). *Regionala jämförelser mellan Nationella Marktäckedata och fältdata från Riksskogstaxeringen och NILS*. Umeå, Sweden.O'Neil, G. L., Goodall, J. L., & Watson, L. T. (2018). Evaluating the potential for site-specific modification of LiDAR DEM derivatives to improve environmental planning-scale wetland identification using Random Forest classification. *Journal of Hydrology*, *559*, 192–208. https://doi.org/https://doi.org/10.1016/j.jhydrol.2018.02.009O'Neil, G. L., Saby, L., Band, L. E., & Goodall, J. L. (2019). Effects of LiDAR DEM smoothing and conditioning techniques on a topography-based wetland identification model. *Water Resources Research*, *55*(5), 4343–4363. https://doi.org/https://doi.org/10.1029/2019WR024784O´Callaghan, J. F., & Mark, D. M. (1984). The extraction of drainage networks from digital elevation data. *Computer Vision Graphics and Image Processing*, *27*(3), 323–344.Olefeldt, D., Hovemyr, M., Kuhn, M. A., Bastviken, D., Bohn, T. J., Connolly, J., et al. (2021). The boreal–Arctic wetland and lake dataset (BAWLD). *Earth Syst. Sci. Data*, *13*(11), 5127–5149. https://doi.org/10.5194/essd-13-5127-2021Olson,

D. M., & Dinerstein, E. (2002). The Global 200: Priority ecoregions for global conservation. *Annals of the Missouri Botanical Garden*, *89*(2), 199–224. https://doi.org/10.2307/3298564Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., et al. (2001). Terrestrial ecoregions of the world: A new map of life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, *51*(11), 933–938. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2Oltean, G. S., Comeau, P. G., & White, B. (2016). Linking the depth-to-water topographic index to soil moisture on boreal forest sites in Alberta. *Forest Science*, *62*(2), 154–165. https://doi.org/10.5849/forsci.15-054Ontario Ministry of Natural Resources and Forestry. (2015). Far North Landcover [Dataset]. Retrieved from https://data.ontario.ca/dataset/far-north-land-coverOntario Ministry of Natural Resources and Forestry. (2019). Wetlands [Dataset]. Retrieved from https://geohub.lio.gov.on.ca/datasets/mnrf::wetlands/aboutOzesmi, S. L., & Bauer, M. E. (2002). Satellite remote sensing of wetlands. *Wetlands Ecology and Management*, *10*(5), 381–402. https://doi.org/10.1023/A:1020908432489Pastick, N. J., Jorgenson, M. T., Goetz, S. J., Jones, B. M., Wylie, B. K., Minsley, B. J., et al. (2019). Spatiotemporal remote sensing of ecosystem change and causation across Alaska. *Global Change Biology*, *25*(3), 1171–1189. https://doi.org/https://doi.org/10.1111/gcb.14279Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422. https://doi.org/10.1038/nature20584Pelletier, C., Valero, S., Inglada, J., Champion, N., & Dedieu, G. (2016). Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. *Remote Sensing of Environment*, *187*, 156–168. https://doi.org/https://doi.org/10.1016/j.rse.2016.10.010Peltola, O., Vesala, T., Gao, Y., Räty, O., Alekseychik, P., Aurela, M., et al. (2019). Monthly gridded data product of northern wetland methane emissions based on upscaling eddy covariance observations. *Earth Syst. Sci. Data*, *11*(3), 1263–1289. https://doi.org/10.5194/essd-11-1263-2019Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V. R., Murayama, Y., & Ranagalage, M. (2020). Sentinel-2 data for land cover/use mapping: A review. *Remote Sensing*. https://doi.org/10.3390/rs12142291Piper, L. (2016). Great Slave Lake. In *The Canadian Encyclopedia*. Retrieved from https://www.thecanadianencyclopedia.ca/en/article/great-slave-lakePorter, C., Morin, P., Howat, I., Noh, M.-J., Bates, B., Peterman, K., et al. (2018). ArcticDEM. Harvard Dataverse. https://doi.org/doi:10.7910/DVN/OHHUKHQuinn, P., Beven, K., Chevallier, P., & Planchon, O. (1991). The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological Processes*, *5*(1), 59–79. https://doi.org/https://doi.org/10.1002/hyp.3360050106Räsänen, A., & Virtanen, T. (2019). Data and resolution requirements in mapping vegetation in spatially heterogeneous landscapes. *Remote Sensing of Environment*, *230*, 111207. https://doi.org/https://doi.org/10.1016/j.rse.2019.05.026Raynolds, M. K., Walker, D. A., Balser, A., Bay, C., Campbell, M., Cherosov,

M. M., et al. (2019). A raster version of the Circumpolar Arctic Vegetation Map (CAVM). *Remote Sensing of Environment*, *232*, 111297. https://doi.org/https://doi.org/10.1016/j.rse.2019.111297Richter, K., Hank, T. B., Mauser, W., & Atzberger, C. (2012). Derivation of biophysical variables from Earth observation data: Validation and statistical measures. *Journal of Applied Remote Sensing*, *6*(1), 1–23. https://doi.org/10.1117/1.JRS.6.063557Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. In *NASA Special Publication* (pp. 309–317). Washington DC.Roy, D. P., Ju, J., Kommareddy, I., Hansen, M., Vermote, E., Zhang, C., & Kommareddy, A. (2011). *Web Enabled Landsat Dta (WELD) Products - Algorithm Theoretical Basis Document*. Retrieved from http://globalmonitoring.sdstate.edu/projects/weld/WELD_ATBD.pdfRundquist, D. C., Narumalani, S., & Narayanan, R. M. (2001). A review of wetlands remote sensing and defining new considerations. *Remote Sensing Reviews*, *20*(3), 207–226. https://doi.org/10.1080/02757250109532435Saah, D., Tenneson, K., Matin, M., Uddin, K., Cutter, P., Poortinga, A., et al. (2019). Land cover mapping in data scarce environments: Challenges and opportunities. *Frontiers in Environmental Science*, *7*. https://doi.org/10.3389/fenvs.2019.00150Saunois, M., Bousquet, P., Poulter, B., Peregon, A., Ciais, P., Canadell, J. G., et al. (2016). The global methane budget 2000–2012. *Earth Syst. Sci. Data*, *8*(2), 697–751. https://doi.org/10.5194/essd-8-697-2016Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., et al. (2020). The global methane budget 2000–2017. *Earth Syst. Sci. Data*, *12*(3), 1561–1623. https://doi.org/10.5194/essd-12-1561-2020Schuur, E. A. G., McGuire, A. D., Schädel, C., Grosse, G., Harden, J. W., Hayes, D. J., et al. (2015). Climate change and the permafrost carbon feedback. *Nature*, *520*, 171. Retrieved from http://dx.doi.org/10.1038/nature14338Story, M., & Congalton, R. G. (1986). Accuracy assessment: A user's perspective. *Photogrammetric Engineering and Remote Sensing*, *52*(3), 397–399.Thornton, B. F., Wik, M., & Crill, P. M. (2016). Double-counting challenges the accuracy of high-latitude methane inventories. *Geophysical Research Letters*, *43*(12), 12569–12577. https://doi.org/10.1002/2016GL071772Townshend, J. R., Masek, J. G., Huang, C., Vermote, E. F., Gao, F., Channan, S., et al. (2012). Global characterization and monitoring of forest cover using Landsat data: opportunities and challenges. *International Journal of Digital Earth*, *5*(5), 373–397. https://doi.org/10.1080/17538947.2012.713190Vitt, D. H. (1994). An overview of factors that influence the development of Canadian peatlands. *Memoirs of the Entomological Society of Canada*, *126*(S169), 7–20. https://doi.org/DOI: 10.4039/entm126169007-1Walker, D. A., & Raynolds, M. K. (2017). Maps of vegetation types and physiographic features, Kuparuk River basin, Alaska [Dataset]. ORNL Distributed Active Archive Center. https://doi.org/10.3334/ORNLDAAC/1378Warner, B. G., & Rubec, C. D. A. (1997). *The Canadian Wetland Classification System, Second Edition*. Waterloo, Ontario.White, B., Ogilvie, J., Campbell, D. M. H. M. H., Hiltz, D., Gauthier, B., Chisholm, H. K. H., et al. (2012). Using the cartographic depth-to-water index to locate small streams

and associated wet areas across landscapes. *Canadian Water Resources Journal / Revue Canadienne Des Ressources Hydriques*, *37*(4), 333–347. https://doi.org/10.4296/cwrj2011-909Wilen, B. O., & Bates, M. K. (1995). The US Fish and Wildlife Service's National Wetlands Inventory Project. *Vegetatio*, *118*(1), 153–169. https://doi.org/10.1007/BF00045197Wulder, M. A., Li, Z., Campbell, E. M., White, J. C., Hobart, G., Hermosilla, T., & Coops, N. C. (2018). A national assessment of wetland status and trends for Canada's forested ecosystems using 33 Years of Earth observation satellite data. *Remote Sensing*. https://doi.org/10.3390/rs10101623Xu, J., Morris, P. J., Liu, J., & Holden, J. (2018). PEATMAP: Refining estimates of global peatland distribution based on a meta-analysis. *CATENA*, *160*, 134–140. https://doi.org/https://doi.org/10.1016/j.catena.2017.09.010Yommy, A. S., Liu, R., & Wu, A. S. (2015). SAR image despeckling using Refined Lee filter. In *2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics* (Vol. 2, pp. 260–265). https://doi.org/10.1109/IHMSC.2015.236Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., et al. (2021). ESA WorldCover 10 m 2020 v100. https://doi.org/10.5281/zenodo.5571936Zou, Z., DeVries, B., Huang, C., Lang, M. W., Thielke, S., McCarty, G. W., et al. (2021). Characterizing wetland inundation and vegetation dynamics in the Arctic coastal plain using recent satellite data and field photos. *Remote Sensing*. https://doi.org/10.3390/rs13081492