https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

Open Access Earth System Science Data Discussions

Fusing Regional and Global Datasets to Develop a Composite Land Cover Product Across High Latitudes

Valeria Briones[1], Hélène Genet[2], Elchin E. Jafarov[1], Brendan M. Rogers[1], Jennifer D.Watts[1], Anna-Maria Virkkala [1] Annett Bartsch[3], Benjamin C. Maglio[2], Joshua M. Rady[1], Susan M. Natali[1]

[1] Woodwell Climate Research Center, Falmouth, MA, 02540, USA
[2] Institute of Arctic Biology, University of Alaska Fairbanks, AK 99775 Fairbanks, USA
[3] b.geos, Industriestrasse 1, 2100 Korneuburg, Austria

Correspondence to: Valeria Briones (vbriones@woodwellclimate.org)

Abstract

Rapid warming across the Arctic is the primary driver of widespread permafrost thaw, with far-reaching consequences for local ecosystem resilience, the regional carbon budget, and the global climate system. Because permafrost characteristics and vulnerability are tightly linked to land cover, particularly vegetation type and surface properties, understanding these dynamics requires accurate and detailed land cover information. Spatial variation in vegetation cover influences energy balance, snow insulation, and soil moisture, factors that directly affect permafrost stability. Consequently, high-resolution land cover products are essential for assessing the ecological impacts of permafrost thaw and for improving the representation of permafrost-related processes in predictive models. However, many global land cover datasets fail to capture the spatial heterogeneity and fine-scale ecological features that influence permafrost dynamics, while more detailed regional products often lack coverage across broader, continental extents. This gap presents a challenge for large-scale assessments of permafrost vulnerability under accelerating climate change.

To create a spatially cohesive land cover map that accurately represents the distribution of ecosystems across the Arctic-Boreal region, we integrated existing global and regional land cover datasets using a workflow including machine learning techniques. This approach seamlessly combines diverse data sources, enhancing representation and accuracy. The resulting map represents high-latitude land cover types at a 1km spatial resolution, better capturing the spatial heterogeneity of the landscape compared to coarser resolution land surface products, with a total of 35 land cover classes, including 20 forest types (e.g., Larch, Birch, Mixed forests), 6 shrubland classes, and wetlands subdivided into bog, fen, and marsh. To achieve this, we used a global land cover map, the European Space Agency Climate Change Initiative Land Cover data (ESA CCI-LC), as the base map and integrated regional maps across the circumpolar region with finer-resolution land cover information to capture the diversity of land cover types. This approach ensured consistent classification across geopolitical boundaries, while incorporating representative vegetation communities at a region-specific level. We show that regional land cover products can be successfully fused to yield a higher-resolution thematic content at the circumpolar scale in comparison to existing global products. The hybrid land cover product can be freely access via https://doi.org/10.5281/zenodo.15231293 (Briones et al 2025).

1. Introduction

As an expression of the interactions between climate, disturbances, landform, soil characteristics, land cover is a critical driver of the spatial heterogeneity of numerous ecological processes. Yet, existing land cover products often face trade-offs between spatial coverage and detailed classification. Global products provide cohesive spatial coverage but often represent land cover using broad generalized classes that may not capture the

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

heterogeneity and complexity of high-latitude systems. In contrast, regional products are often offered at a finer spatial resolution, with detailed classifications, but with limited spatial coverage. These limitations often present a challenge for large scale ecological investigations of the interactions between land cover and climate, disturbances, permafrost, energy balance, or biogeochemical cycles, that require both a cohesive product with detailed classification. The exponential increase in field observations are allowing significant advances in understanding and quantifying the role of vegetation composition on ecological processes for an ever increasing number of land cover classes across the Arctic (e.g.Euskirchen et al., 2016; Melvin et al., 2015; Oehri et al., 2022). However, the coarse classifications of circumpolar land cover products represent a significant boundary in upscaling new understanding about the influence of local vegetation composition at regional and global scales.

These applications require detailed representations of vegetation communities and composition to improve the accuracy of climate projections, permafrost modeling, and disturbance impact analysis. However, many models continue to rely on coarse land cover classifications, which limits their ability to capture the heterogeneity of vegetation and, in turn, reduces confidence in projections of ecosystem vulnerability and climate feedbacks. This is particularly problematic when scaling from detailed site-level observations to regional or circumpolar models, as much of the ecological heterogeneity, critical differences in vegetation composition and structure between land cover types, is lost in coarse-resolution datasets.

An approach to address this gap is the fusion of global and regional land cover products, an approach that has not yet been widely implemented. By integrating the broad spatial coverage of global datasets with the finer-scale detail of regional classifications, this method can produce harmonized datasets that retain both spatial continuity of the global products and ecological relevance of the local and regional classifications (Pérez-Hoyos et al., 2012; Luo et al., 2024; Wang et al., 2023). In this study, we present a novel approach to address this challenge, focusing on the circumpolar high latitudes, where ecosystem models require relatively fine-scale land cover information to improve projections of carbon and energy fluxes. In high-latitude regions, climate change is reshaping the intricate interplay of biophysical and biogeochemical processes that regulate terrestrial carbon and energy balances (Heimann & Reichstein, 2008), contributing a positive feedback to warming of the global climate system (Intergovernmental Panel On Climate Change, 2022). Climate warming and changes in precipitation are affecting disturbance regimes across the Arctic and boreal biomes, including permafrost thaw and wildfires, increasing the vulnerability of ecosystems to state change (Johnstone et al., 2016; Schuur et al., 2022). Interactions between climate and disturbances are impacting vegetation composition, which in turn modulate ecosystem structure and functions, defining ecosystem vulnerability to change. These changes in vegetation affect land cover distribution and can have significant impacts on climate feedback through above and below-ground carbon dynamics (Poulter et al., 2015) and surface energy exchange (Oehri et al., 2022; Thompson et al., 2004). For example, changes in Arctic-Boreal Zone (ABZ) vegetation composition are closely tied to carbon fluxes and stocks (e.g., Balshi et al., 2009; Mack et al., 2021; Virkkala et al., 2018), and increases in tree and shrub density and decrease in land surface albedo, accelerating regional warming (Betts, 2000; Bonfils et al., 2012; Miller & Smith, 2012), altering carbon and nitrogen feedbacks as well as snow and permafrost dynamics (Elmendorf et al., 2012; Tape et al., 2006; J. A. Wang et al., 2020; Zhang et al., 2013). As Arctic warming accelerates, vegetation shifts and permafrost thaw, mapping land cover in high-latitude systems is becoming increasingly important to predict the implications of these changes at the regional and global scale (Horvath et al., 2021; Kåresdotter et al., 2021).

Remotely sensed observations serve as a major data source for land cover mapping, allowing monitoring across global ecosystems (Joshi et al., 2016; Rogan & Chen, 2004). There have been extensive efforts to map vegetation communities across the Arctic biome to capture the heterogeneity and diversity of tundra vegetation communities (Bartsch et al., 2024; M. K. Raynolds et al., 2019; Walker et al., 2005). However, there is often a gap between the level of detail in land cover classification between these region-specific and global products, which mostly do not capture the diversity of these high-latitude vegetation communities, especially across wetlands and tundra (Bartsch et al., 2016). Advances in wetland maps have been made recently, including the Boreal-Arctic Wetland Lake Database (BAWLD, Olefeldt et al. 2021), which provides wetland classes characterized by different hydrology, species composition, and methane emissions. This classification includes 5 distinct wetland types (bog, fen, marsh, permafrost bog, tundra wetland), but offered at a coarse 1° spatial resolution (Olefeldt et al., 2021).

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

There have been great advances in the quantity and quality of land cover maps representative of the Arctic biome at finer-scale resolutions (Raynolds et al. 2019, Liu et al. 2023, Bartsch et al. 2024). one of which includes the 10-m resolution CircumArctic Land Cover Units map (CALU, Bartsch et al. 2024) representing the area north of the latitudinal treeline with 23 land cover classes and (Bartsch et al., 2024). However, none of these maps cover the boreal biome. Despite the increasing availability of land cover products at regional and global scales with varying resolutions, there remains a need for a spatially continuous and harmonized classification that reflects the latest advances in land cover characterization and detection across the permafrost zone. Such a dataset must provide the necessary detail to represent the influence of land cover distribution on ecological processes while accurately capturing the spatial heterogeneity consistently across regional and geopolitical boundaries across the ABZ. In this study, we address the need for a wall-to-wall land cover map, by presenting our workflow for developing a hybrid land cover product. We fuse existing global and regional products through machine learning techniques, producing a detailed circumpolar land cover map at a 1 km^2 spatial resolution, specifically optimized for modeling applications. While our primary focus is on the ABZ, the methods and approaches are broadly applicable to other regions. Additionally, we assess the limitations of existing land cover products and identify opportunities for enhancing the representation of high-latitude terrestrial systems and beyond.

2. Material & Methods

2.1 Study area

Our study region encompasses the ABZ, a region characterized by diverse vegetation composition spanning the northern tundra and boreal biomes (Fig. 1, Dinerstein et al., 2017). The ABZ includes approximately 4.98×10^6 km^2 of tundra ecosystems , 11.93×10^6 km^2 of boreal ecosystems, and 2.45×10^6 km^2 of other ecosystems (e.g. ice, barren …) (Neigh et al., 2013). The boreal zone transitions from tundra at its northern boundary to temperate forest, steppe, or prairie in the south. Boreal forests dominate this domain, with coniferous species including spruce, pine, larch, and fir constituting the majority (~61 %), followed by mixed wood forests (~24 %) (Neigh et al., 2013; USGCRP, 2018). Additionally, hardwood tree species (~3 %), including birch, alder, and aspen, contribute to the region's vegetation diversity (USGCRP, 2018). Wetlands are also extensive across the ABZ made up of 3.2×10^6 $km^2$of the domain, and if comprised of fens (29 %), non-permafrost bogs (28 %), permafrost bogs (27 %), marshes (5 %) and tundra wetlands (12 %) (Olefeldt *et al* 2021). Tundra ecosystems are largely treeless dominated by herbaceous plants, mosses, lichens and dwarf shrubs (USGCRP, 2018; Epstein et al., 2004). Vegetation structure ranges from low-lying moss and lichen-dominated areas in colder, drier regions to more productive graminoid and shrub tundra in milder zones. Tundra vegetation plays a critical role in modulating surface energy balance, snow insulation, and permafrost stability, making it a key element in predicting feedbacks to the global climate system (Walker et al., 2005; Schuur et al., 2022).

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

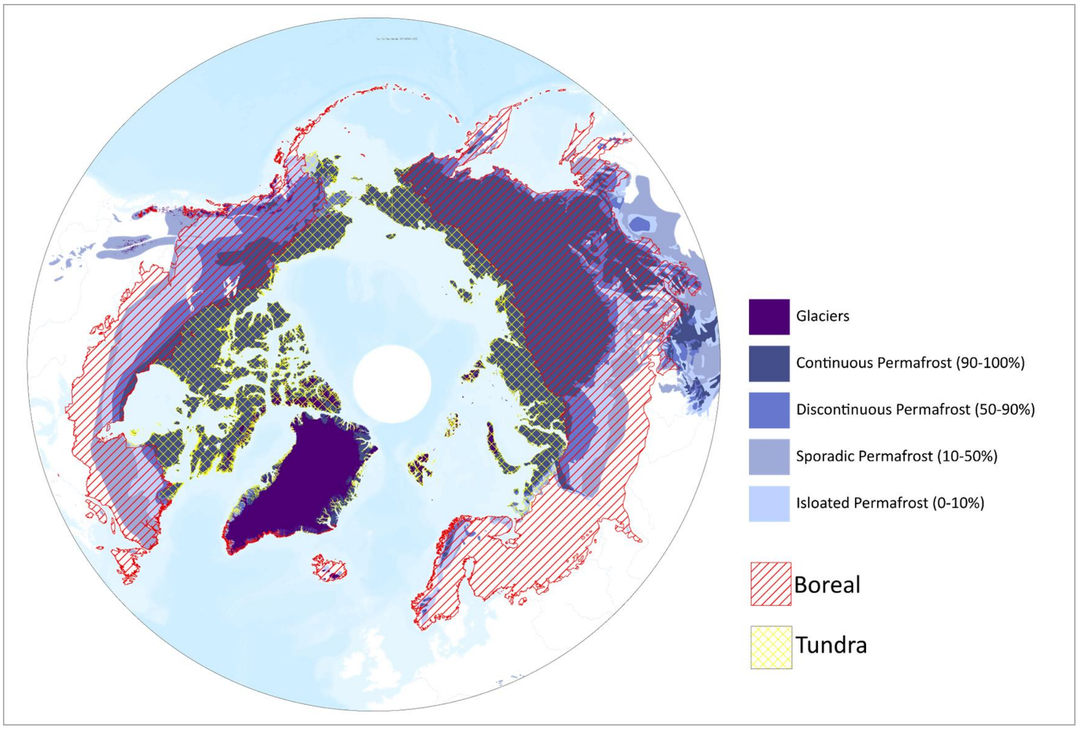

Figure 1: The permafrost map (Brown et al., 1997) overlaid over the study region highlighting the Tundra and Boreal extent of the study region, with the Arctic and Boreal regions highlighted in red. Map adapted from (USGCRP, 2018).

2.2 Methods Overview

Our goal was to develop a hybrid land cover product that combines the fine-scale land cover information from regional products with a global land cover layer that provides consistency across regional boundaries, in our case for the circumpolar ABZ. The workflow (Fig. 2) is comprised of five major steps: (1) pre-processing and re-projecting select land cover products to a 1 km grid using a majority rule approach, (2) reclassifying and combining global and regional products to a common legend based on agreement between global and regional products, (3) reclassifying pixels that were inconsistent between global and regional products using random forest machine learning and ancillary data variables, (4) post-classification comparison, and (5) final product compilation at 1 km resolution.

2.3 Land Cover Products and Classes

To develop the hybrid land cover product, we used a global map as the base layer to ensure a spatially continuous land cover classification. Region-specific land cover products were then selected and integrated to provide a finer-scale land cover classification scheme and to inform the reclassification process for the final product (Table 1). We chose the The European Space Agency (ESA) Climate Change Initiative Land Cover (CCI-LC) (ESA CCI-LC) as the global base map as it represents the relevant key land cover classes representative of ecosystem dynamics used in other studies related to carbon flux modeling (Virkkala et al., 2021). The global ESA CCI-LC is an annual 300 m spatial resolution land cover dataset with a time series of annual maps from 1992-2020. The map includes 22 general land cover classes defined using the Land Cover Classification System (ESA, 2016), of which include 7 forest type (i.e. Tree cover, needle leaved, Tree cover broadleaved), as well as shrublands, mosaic vegetation and sparse vegetation, with an overall accuracy of 71.5 % (ESA, 2017) (Table S2). To represent Arctic tundra vegetation community types, we incorporated the CALU map, which is an extension to the ESA Permafrost CCI at a 10 m resolution (ESA CCI Permafrost, Bartsch et al., 2024). The CALU map includes 23 land cover units derived from a fusion of Sentinel-1 and Sentinel-2 and separates between water bodies, wetlands, and soil moisture conditions useful for the parameterization of permafrost processes. Of the 23 vegetation classes within CALU there are 20 vegetated units including 3 wetlands classes, 2 lichen/moss, 10 classes of various shrub tundra types, 3 forest types and classes for graminoids and barren tundra. CALU is able to capture the fine-scale heterogeneity of the Arctic landscapes, which allows for a better representation of high-latitude systems and processes. We grouped the 20 vegetated land cover types in CALU into five major land cover types based on the CAVM map units: barrens, graminoid tundra, prostrate-shrub tundra, erect-shrub tundra, tall-shrub tundra, and wet-sedge tundra (M. K. Raynolds et al., 2019). CALU also has three forested land cover classifications (evergreen, deciduous, mixed), which we further classified to a specific forest type (i.e. black spruce, birch forest) depending on the classification of overlapping regional land cover products. Since CALU primarily represents the tundra biome, its applicability is limited and does not adequately capture the land cover types found within the expansive boreal region, which constitutes the majority of our study area (Table S3).

Regional products used for the boreal biome in the study were selected based on their spatial resolution, coverage, and the representativeness of land cover classes. Across North America, the main land cover products selected include the Landscape Fire and Resource Management Planning Tools (LANDFIRE) Existing Vegetation Type (EVT) map covering all of Alaska as well as a 90 km buffer zone across the Alaska-Canada border. The LANDFIRE EVT map is derived from field observations and Landsat images at a 30 m resolution and includes over 130 land cover classes for Alaska (e.g., North American Arctic lichen tundra, Western North American boreal mesic birch-aspen forest) (LANDFIRE, 2023, Rollins, 2009) (Please see Table S4 for a full list of land cover classes). For Canada, we selected the Virtual Land Cover Engine (VLCE) product mapped annually from 1984 to 2022 across Canada's forested systems, derived from Landsat surface-reflectance best-available-pixel image composites (Hermosilla et al., 2022). The VLCE map consists of 12 land cover classes at a 30-m spatial resolution, including 3 forested classes (broadleaf, coniferous, mixedwood), and 2 wetland classes (wetland, wetland-treed) (Hermosilla et al., 2018) (Table S5). To supplement the Canada VLCE, we included the lead tree species map of Canada, consisting of 37 tree species mapped across Canada at a 30-m spatial resolution, with a reported overall accuracy of 70.3 % (brus) (Table S6). The leading tree species map provides the distribution of dominant tree species per pixel derived from Landsat image composites and ancillary predictor layers such as data derived from Canada's Forest Inventory plots, and provides an overall accuracy of approximately 93 % (Hermosilla et al., 2022; Stinson et al., 2016; White et al., 2014).

We selected several regional products to cover Arctic-boreal Eurasia For Europe, we selected the CORINE Land Cover dataset, which is a pan-European product at a 100-m spatial resolution produced from Landsat and Spot images (CORINE, 2018) (Table S7). The CORINE product contains 44 unique land cover classes (level 3) including 3 main forest types (mixed, coniferous, and broad-leaved forest) and 5 wetland classes (inland marshes, peat bogs, salt marshes, intertidal flats, and salines) with a reported 85 % overall accuracy, with the accuracy of individual classes ranging from 95 % to 70 % (Aune-Lundberg and Strand 2021). To enhance the classification of specific forest types across Fennoscandia, we incorporated the Tree Species Map of Europe (Table S8). This map was developed by integrating plot data from the International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP-Level I) and National Forest Inventory (NFI) data, resulting in a 1 km resolution product representing 20 unique species groups across Europe, with an estimated accuracy of 43–57 % (Brus et al., 2012). To supplement the classification of wetlands across Fennoscandia, we integrated the European Environmental Agency (EEA) Extended wetland ecosystem layer map, which includes 20 wetland classes across Europe at a 100 m spatial resolution for the year 2018 (EEA, 2018). For Russia, we used the land cover map of Northern Eurasia, which was created in 2000 from SPOT-4 Vegetation products and contains 26 classes at a 1km spatial resolution (Bartalev et al., 2003) (Table S9). To supplement the product with specific forest species, we incorporated the dominant tree species of Russian forests map from the Space Research Institute of the Russian Academy of Sciences which consists of 20 forest classes at a spatial resolution of 230 m (Balashov et al., 2020; Bartalev et al., 2016).

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

Table 1: Characteristics of the land cover products used in this study.

Dataset	Spatial Resolution	Extent	Zone	Reference Year(s)	Number of Classes	Reference
ESA CCI Land Cover 2016	300 m	Global	Global	2016	22	Lamarche et al., 2017
ESA CCI Permafrost Circumpolar Arctic Land cover Units (CALU)	10 m	Circumpolar	Arctic	2016-2023*	23	Bartsch et al. 2024
LANDFIRE	30 m	Alaska	Arctic-Boreal	2023	130	Rollins, 2009
Leading Tree Species	30 m	Canada	Boreal	2019	37	Hermosilla et al., 2022
Canada Virtual Land Cover Engine (VLCE)	30 m	Canada	Boreal	2016	23	Hermosilla et al., 2018
Canadian Wetland Inventory Map	10 m	Canada	Arctic-Boreal	2000-2016*	4	Mahdianpari et al., 2020
Tree Species Map for European Forests	1 km	Europe	Boreal	-	20	Brus et al., 2012
Corine Land Cover (CLC)	100 m	Europe	Boreal	2018	44	CORINE, 2018
European Environment Agency (EEA) Extend wetland ecossytem layer	100 m	Europe	Boreal	2018	20	EEA, 2018
Land Cover of Northern Eurasia	250	Russia	Arctic-Boreal	2000	26	Bartalev et al., 2003
Russia's Forests Map	300 m	Russia	Arctic-Boreal	2016	20	Bartalev et al. 2016

***** Denotes land cover products that are available at an annual time scale, but a single year was selected for the project.

2.4 Pre-Processing and Initial Agreement Classification

The datasets used in this study were inconsistent in their temporal and spatial resolutions as well as the land cover classifications. We addressed these differences by pre-processing the datasets to a uniform projection and spatial resolution (1 km), as well as developing a crosswalk, as discussed in the following section, of land cover classes to translate from the original classes of each regional product to that of the final product. To account for the geographical extents of each product, we approached this reclassification region by region: Alaska, Canada, Iceland, Fennoscandia, Russia, and Greenland. The final list of classes and their descriptions can be found in Table 2.

Each land cover product was reprojected to WGS 1984 NSIDC EASE-Grid Global to match the projection of our input data and resampled to 1 km using a majority-rule approach. We developed a crosswalk between the regional products and the global base map to align with the general land cover classification (i.e., herbaceous). This involved grouping similar land cover classes (i.e., grasslands, croplands) within the more broad group (i.e., herbaceous). The reclassification created a harmonized classification scheme while maintaining the main representative land cover classes (forest, shrubland, herbaceous/grassland, barren, and wetland).

The first step of reclassification is based on the agreement of major land cover classifications between global and regional maps (Fig. 2). A given pixel is grouped into a new classification if both the global and regional map pixels agree at the specific land cover classification, such as evergreen needleleaf forest, broadleaf forest, shrubland, herbaceous, barren or wetland. The classification is then further refined by reassigning each pixel to a final land cover class based on its regional classification. For example, if both the global and regional maps classify the pixel as coniferous forest and a supplemental regional product identifies it as a Black Spruce Forest, then the final classification is Black Spruce Forest (the general schematic can be seen in Fig. 2). However, if the global

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

dataset classifies a pixel as evergreen needleleaf forest while the regional product identifies it as a broadleaf forest, they are in disagreement and require further processing. When such disagreements occur, we implement a random forest classification to determine the final classification of a given pixel. This further refines the classification to the final target legend of the hybrid product for a region.

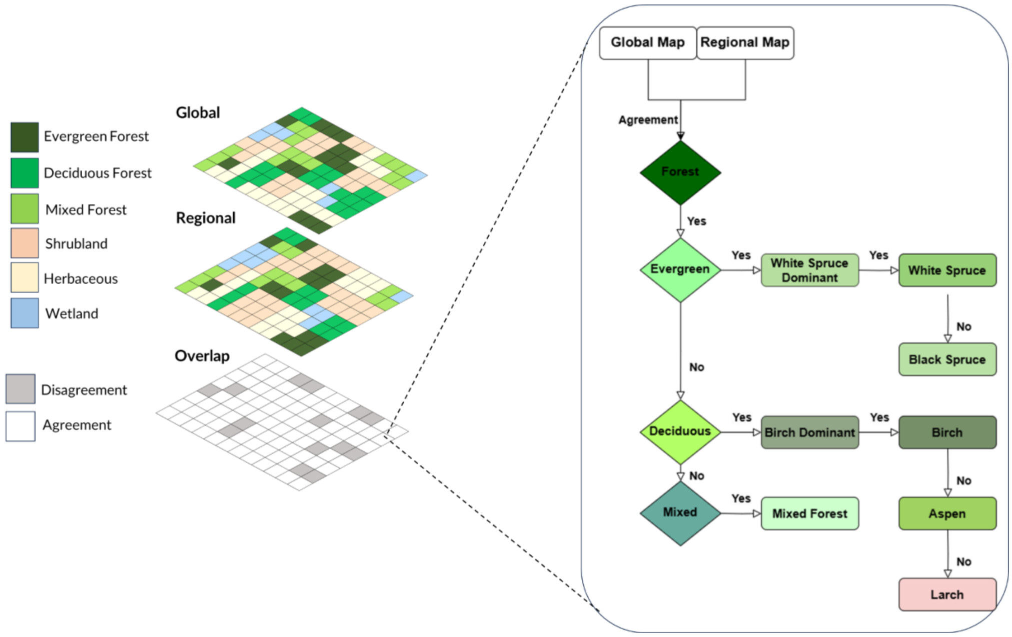

Figure 2: An overview of the crosswalk between the global base map (ESA CCI-LC) and regional map(s) land cover classifications used to develop the classification scheme for the final hybrid land cover product classifications, specifically in the case when there is overlap agreement between products.

2.5 Random Forest Classification

When global and regional land cover products showed inconsistencies in classification for a given pixel, we applied a machine-learning approach to ensure consistency in vegetation classes. This was achieved by training the model with successfully classified pixels from the initial agreement step and incorporating predictor data. Random Forest is one of the most widely used supervised machine learning algorithms for land cover mapping and classification because of its ability to handle noisy and multi-source datasets (Jin et al., 2018; Maxwell et al., 2019). The random forest method is an ensemble-based classifier that uses decision trees for training and prediction. In Google Earth Engine, we used the random forest classifier. We optimized the number of trees (ntrees), selecting values between 200 and 300 based on the lowest out-of-bag (OOB) error across regions. Other hyperparameters, such as mtry (the number of variables considered at each split), were left at their default values, approximately the square root of the total number of predictor variables used in the model, which ranged from 2 to 3. We also calculated the importance score of the predictor data within the random forest classifier. The value of the importance scores are not uniform, but instead change depending on the number of sampling data and variables included within the classifier (C. Liu et al., 2020). The importance scores for each major region can be found in Fig. S1.

Predictor included within the random forest model are those strongly linked to vegetation dynamics and characteristics. Climate variables, specifically the 19 Bioclimatic variables from WorldClim, were included to represent aspects of temperature, precipitation and seasonality (Hijmans et al., 2005). Other variables included topography (i.e. elevation, slope, aspect), as well as descriptors of vegetation and moisture including Normalized

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

Difference Vegetation Index (NDVI) and Normalized Difference Infrared Index (NDII) (Table S1). Training and testing data came from pixels that agreed during the initial classification agreement step between the global land cover and regional layers. We created a collection of samples that were then randomly split into groups of training (80 %) and testing (20 %) data. In total, between 500-1,200 pixels per land cover class were collected within each region. We adapted each random forest machine learning model to be region-specific to map the distribution of land cover class.

2.6 Assessment

The accuracy of the final hybrid land cover product (including the classifications from random forest) was dependent on the quality, level of detail, and spatial extent of the regional land cover products. We compared the proportion of agreement between the global base dataset (ESA CCI-LC) and each respective regional dataset. This comparison was conducted at a coarser land cover classification level than our final hybrid product (e.g., agreement between deciduous forest type was compared), as the ESA CCI-LC does not provide detailed information on, e.g., forest species or wetland types comparable to regional products. For each region we used the confusion matrix reports to estimate accuracy metrics such as the overall accuracy (OA), producer accuracy (PA), and user accuracy (UA) (Olofsson et al., 2014) which can be found in Fig. S2. In addition, we estimated the proportion of each land cover class of the hybrid product to compare to the global and regional products for each region, and show the accuracy metrics for the random forest classifications (Sect. 2.5).

3. Results

We present our final hybrid land cover product at a 1-km spatial resolution, developed by fusing both global and regional land cover datasets across the circumpolar region (Fig. 3). You can find an in-depth classification for each individual major region within Fig. S3.

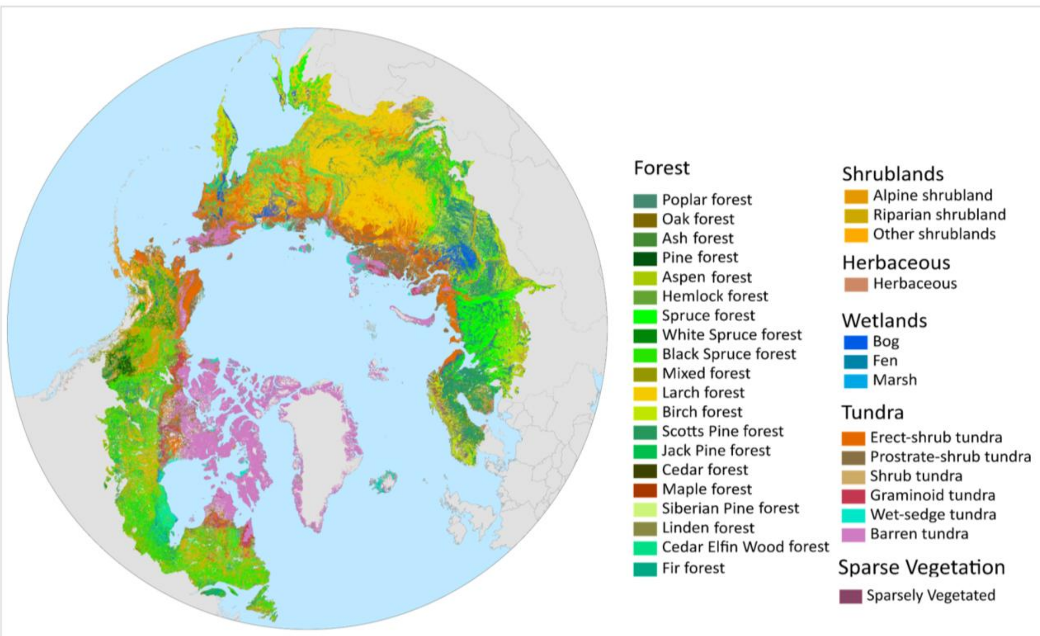

Figure 3: Final Hybrid product at 1 km resolution across the Arctic-Boreal Zone, with the integrated Circumpolar Arctic Land cover Units Map (CALU) map covering the pan-Arctic.

3.1 Overlap

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

When assessing the overlap between the global and regional products across the ABZ, our results show a majority agreement for each region (Fig. 4). In Alaska, the global and regional products show a 82.68 % agreement (17.32 % disagreement) across the land cover products (Fig. 4a), while in Canada there is a 79.86 % agreement (20.14 % disagreement) (Fig. 4b). There is a 79.51 % agreement (20.49 % disagreement) between global and regional products for Fennoscandia (Fig. 4c), while Russia has a 90.82 % agreement (9.18 % disagreement) amongst the land cover products, showing a large majority agreement across the region (Fig. 4d).

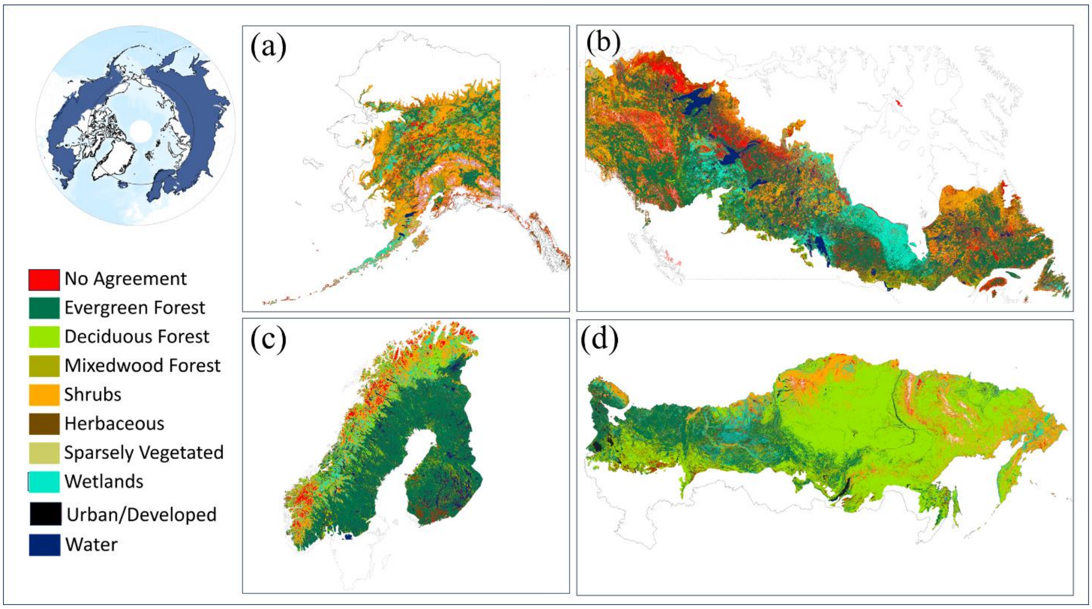

Figure 4: Agreement between the global ESA CCI-LC map and region-specific land cover products including: LANDFIRE EVT, VLCE, CORINE, Bartalev et al. 2003) across the circumpolar region at a 1 km spatial resolution. Major regions include Alaska (a), Canada (b), Fennoscandia (c), and Russia (d).

3.2 Accuracy

The pixels classified using the random forest machine learning model were further assessed for each of the major regions. The producer's and user's accuracy metrics were derived from the confusion matrix (Fig. 5). Due to the fact that a majority of the pixels within each regional map were classified within the first agreement-based classification step, the performance of the random forest classifier was relatively accurate in classifying the remaining unclassified pixels.

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

Open Access Earth System Science Data Discussions

50-70% 71-80% 81-94% ≥95%

		Alaska		Canada		Fennoscandia		Russia		Iceland	
		UA	PA	UA	PA	UA	PA	UA	PA	UA	PA
	Land Cover Class	-									
Deciduous Forest	Poplar forest	98.00%	100.00%	98.00%	100.00%	-	-	-	-	-	-
	Ash Forest	-	-	-	-	82.00%	88.00%	-	-	-	-
	Aspen Forest	99.00%	100.00%	94.00%	96.00%	-	-	92.00%	98.00%	-	-
	Larch forest			99.00%	100.00%			89.00%	97.00%	-	-
	Birch forest	94.00%	91.00%	95.00%	99.00%	87.00%	85.00%	84.00%	92.00%	-	-
	Maple	-	-	91.00%	98.00%	-	-	100.00 %	99.00%	-	-
	Linden	-	-	-	-	-	-	100.00 %	100.00%	-	-
Evergreen Forest	Fir forest	-	-	98.00%	-	-	-	95.00%	94.00%	-	-
	Oak forest	-	-	-	-	-	-	100.00 %	100.00%	-	-
	Pine Forest	-	-	97.00%	100.00%	87.00%	93.00%	-	-	-	-
	Hemlock forest	99.00%	100.00%	100.00%	100.00%	-	-	-	-	-	-
	Spruce forest	-	-	-	-	-	-	89.00%	65.00%	-	-
	White Spruce forest	97.00%	97.00%	98.00%	98.00%	-	-	-	-	-	-
	Black Spruce forest	95.00%	97.00%	100.00%	99.00%	-	-	-	-	-	-
	Scotts Pine forest	-	-	-	-	88.00%	83.00%	71.00%	89.00%	-	-
	Jack Pine forest	-	-	99.00%	100.00%	-	-	-	-	-	-
	Cedar forest	-	-	-	-	-	-	-	-	-	-
	Siberian Pine	-	-	-	-	-	-	91.00%	77.00%	-	-
	Cedar Elfin Wood	-	-	-	-	-	-	95.00%	100.00%	-	-
	Mixed forest	100.00%	98.00%	94.00%	84.00%	90.00%	90.00%	82.00%	77.00%	-	-
Shrublands	Alpine shrubland	99.00%	100.00%	94.00%	98.00%	-	-	-	-	-	-
	Riparian shrubland	100.00%	99.00%	91.00%	95.00%	-	-	97.00%	89.00%	-	-
	Other shrublands	93.00%	81.00%	97.00%	88.00%	86.00%	81.00%	86.00%	85.00%	-	-
	Erect-shrub tundra	77.00%	89.00%	91.00%	88.00%	88.00%	85.00%	87.00%	76.00%	94.00%	99.00%
	Prostrate-shrub tundra	96.00%	100.00%	-	-	90.00%	84.00%	89.00%	92.00%	100.00%	100.00%
	Shrub tundra	-	-	91.00%	92.00%	-	-	-	-	100.00%	100.00%
Herbaceous	Herbaceous	94.00%	99.00%	100.00%	100.00%	82.00%	93.00%	94.00%	92.00%	92.00%	85.00%
	Graminoid tundra	94.00%	96.00%	87.00%	91.00%	89.00%	86.00%	79.00%	88.00%	-	-
Sparse Vegetation	Sparsely Vegetated	100.00%	100.00%	97.00%	83.00%	89.00%	94.00%	-	-	-	-
	Barren tundra	89.00%	89.00%	96.00%	81.00%	100.00%	66.00%	88.00%	76.00%	100.00%	100.00%
Wetlands	Bog	88.00%	86.00%	99.00%	100.00%	85.00%	83.00%	85.00%	89.00%	77.00%	86.00%
	Fen	91.00%	89.00%	100.00%	100.00%	98.00%	87.00%	89.00%	95.00%	95.00%	50.00%
	Marsh	100.00%	99.00%	100.00%	98.00%	100.00%	69.00%	94.00%	100.00%	83.00%	85.00%
	Wet-sedge tundra	92.00%	90.00%	95.00%	95.00%	91.00%	92.00%	91.00%	89.00%	100.00%	100.00%
	Developed	100.00%	88.00%	-	-	-	-	99.00%	100.00%		

Figure 5: Confusion Matrix for the Random Forest Model used to classify unclassified pixels from the initial agreement classification step for all major regions, displaying the User's Accuracy (UA), and Producer's Accuracy (PA).

For class-specific accuracies of the unclassified pixels, predominantly forested land cover classes are mapped with very high accuracies (>82 %). The class accuracies of herbaceous vegetation, wetlands, and tundra classes within the ABZ transition zone were classified with moderate accuracy (>50 %). Specifically, herbaceous vegetation and shrubs have high confusion errors with forest and herbaceous vegetation classes. This is expected considering the spectral similarity of these classes. Some land cover classes generated more confusion within the classifier, such as Barren tundra across most regions except Iceland, with the lowest PA at 66 % in Fennoscandia and 76 % in Russia. Within these regions the classification of Barren tundra often had higher confusion errors with other tundra classes (i.e. Graminoid tundra). Detailed confusion matrices for each region are shown in Fig. S2.

3.3 Assessment

When comparing the proportion of major land cover classes across the boreal domain (Fig. 6c, d) of the final hybrid product at a 1 km resolution, Alaska shows a forest composition of 44% in comparison to 44 % and 48 % for the regional LANDFIRE and global ESA CCI-LC maps, respectively. Total forest composition in the boreal region of Canada was 50 % in our final product compared to 50 % and 71 % for the regional VLCE and ESA CCI-LC maps, respectively. Boreal Alaska shrublands represented 44 % of the total vegetation cover compared to 41 % and 24 % represented in LANDFIRE and ESA CCI-LC, respectively. Total shrublands were less prominent across boreal Canada with values of 23 % in the hybrid product compared to 21 % and 8 % from the VLCE and ESA CCI-LC maps, respectively.

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

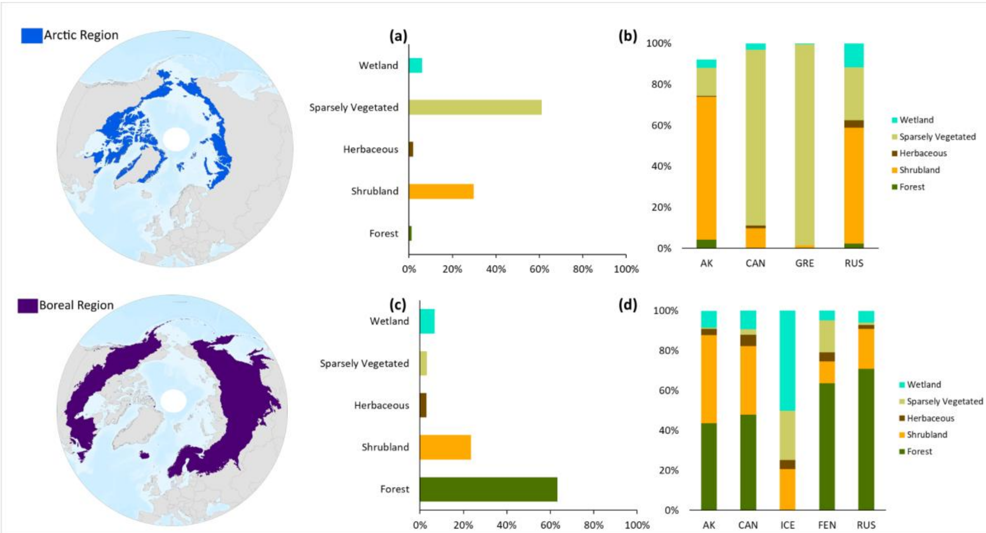

Figure 6: Assessment of major land cover class types for our hybrid land cover product grouped by biome, Arctic (a,b) and Boreal (c,d). Panels show the total proportion of each class over the entire extent (a,c), and further subdivided by major region (b,d) including Alaska (AK), Canada (CAN), Iceland (ICE), Greenland (GRE), Fennoscandia (FEN), and Russia (RUS).

We compared the representation of five dominant land cover classes (forest, shrubland, herbaceous, sparsely vegetated, and wetlands) of the hybrid land cover product land cover types to the global and regional land cover products in their original spatial resolution. Figure 7 shows the comparison of land cover products and their classification at four established research sites including a (a) shrub tussock tundra at Eight Mile Lake in Alaska (63.8784° N, -149.2536°W), (b), black spruce peat plateau at Scotty Creek in Canada (61.3079°, 121.299° W), (c) Larch forest at Yakutsk Spasskaya in Russia (62.55° N, 129.241°E) and (d) shrub tundra at Seida in Russia (67.05° N, 62.94°E).

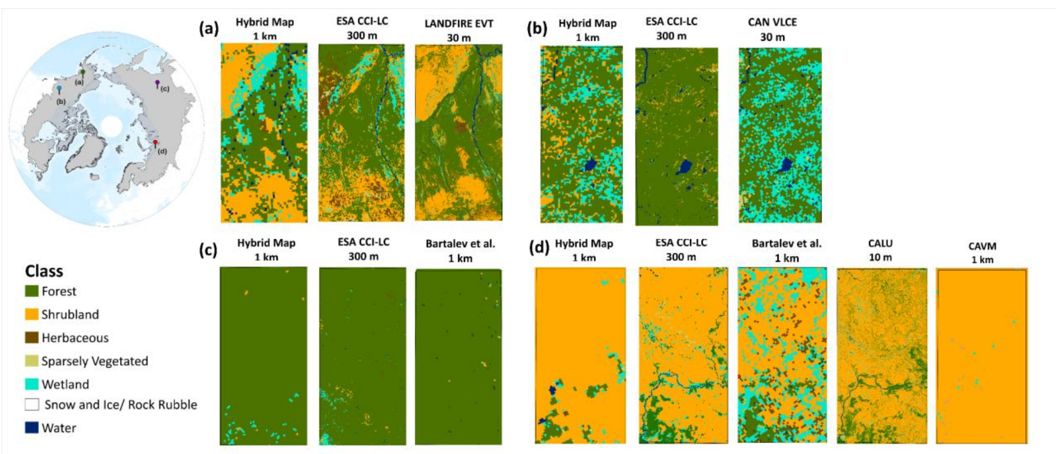

Figure 7: Comparison of the final hybrid land cover product to the regional and global products used within the study. The sites selected includes (a) shrub tundra at Eight Mile Lake in Alaska, (b) Black Spruce peat plateau at Scotty Creek in Canada, (c) Larch forest at Yakutsk Spasskaya in Russia, and (d) a shrub tundra at Seida in Russia.

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

4. Discussion

The development of the hybrid land cover product at a 1 km spatial resolution provides a valuable synthesis of both global and regional land cover datasets across the ABZ. The results indicate a high degree of agreement between global and regional products, yet notable differences emerge both regionally and across specific land cover classes.

4.1 Integration Approach for a Hybrid Land Cover Product

We present an integration approach to generate a hybrid land cover product at a 1 km spatial resolution, specifically tailored to vegetation classifications for regional model applications and other research investigations. Our main objective was to introduce an adaptable approach that enables the combination of multiple land-cover datasets with varying spatial resolutions, thematic content, and sources, resulting in a unified hybrid classification system. Our final product exhibited reasonable accuracy across regions when compared to both the global and corresponding regional maps.

The general harmonization approach of our workflow is reproducible and can be applied to other study regions at both regional and global scales. To accurately represent the diverse vegetation communities of the ABZ, integrating multiple local and regional land cover maps is essential. Compared to the Arctic region, where substantial efforts have been made to represent tundra vegetation in detail and at high spatial resolutions (Bartsch et al., 2024; A. Liu et al., 2023; Raynolds & Walker, 2008), boreal vegetation has remained more challenging to classify consistently. As new or updated regional and global products become available, our approach remains flexible enough to incorporate these data, leading to continuous refinement and enhancement of the final product and providing more detailed land cover classifications.

4.2 Regional Disparities in Agreement and Model Performance

Agreement between global and regional products varies by region, with Alaska and Canada exhibiting higher disagreement rates (17.32 % and 20.14 %, respectively) compared to Russia (9.49 %). This suggests that regional land cover datasets in North America capture fine-scale variations in vegetation classes that global products fail to represent. Russia performs relatively better in terms of agreement rate, likely due to limited availability in region-specific land cover products, therefore reducing the chance of disagreement across various products, compared to regions such as Alaska and Canada, which have more available land cover products at fine spatial resolution. Across the Alaskan Boreal extent, in areas where there was higher disagreement between global (ESA CCI-LC) and regional products (LANDFIRE), the greatest source of disagreement came from misclassification between coniferous forest and mixed forest, at 13 % of pixel disagreement. Similarly, across the Canadian Boreal extent, the areas of highest disagreement were those between shrublands (VLCE) and coniferous forest (ESA CCI-LC) at 8 %, and disagreement between coniferous forest and mixed forest classifications at 6 %. This disagreement among similar land cover classes has often been attributed to confusion from similar spectral signatures in the satellite imagery. This is especially true for differences between mixed and coniferous forest, as well as deciduous forest, shrublands, shrub-covered wetlands, and herbaceous classes (Latifovic et al., 2017; Wang et al., 2019).

Comparisons of land cover class representation reveal discrepancies in the total coverage of different classes across datasets. For instance, boreal Alaska's forest composition aligns closely between the hybrid and LANDFIRE datasets (both at 44 %), yet the global ESA dataset estimates forest extent at 48 %. Similarly, shrubland representation is higher in the hybrid product (44 %) compared to ESA (24 %), suggesting the importance of integrating regional datasets to more accurately reflect dominant vegetation classifications. These differences in classification can be observed across the ABZ as shown in Fig. 6, specifically showcasing the general underestimation of wetlands within the ESA CCI-LC product. These differences have significant implications for ecological modeling, as over- or underestimation of land cover classes directly affects simulations of ecosystem processes, such as carbon fluxes and vegetation dynamics (Jung et al., 2006; Pérez-Hoyos et al., 2012).Wetlands, in particular, remain one of the most challenging land cover types to classify accurately due to their fine-scale heterogeneity, seasonal variability, and often ambiguous spectral signatures. Our hybrid product takes advantage of regional classification schemes and ancillary data layers to improve the delineation of wetland classes. For example, in North America across the boreal extent, our hybrid map estimates wetlands at 21 %, which is closely aligned with the 22 % wetland extent reported in BAWLD, while our hybrid map shows 21 % and BAWLD 28 % across Eurasia boreal extent (Olefeldt et al., 2021). This consistency suggests that our approach can effectively capture the spatial distribution of wetlands, enhancing the reliability of wetland representation for process-based models and climate assessments. A more refined understanding of wetland extent and distribution is particularly critical for permafrost modeling and methane emission projections, where wetland dynamics play a disproportionate role in influencing carbon-climate feedbacks.

Certain regions exhibit higher levels of disagreement, particularly in transitional zones where land cover classes are more challenging to delineate (Heiskanen, 2008; Herold et al., 2008; Prestele et al., 2016). For example, shrub tundra and herbaceous wetlands demonstrate higher classification errors due to spectral similarities with adjacent vegetation classes (Latifovic et al., 2017). The confusion matrix highlight how barren tundra, particularly in Fennoscandia and Russia, is frequently misclassified as graminoid tundra by the random forest model. This misclassification underscores the limitations of optical remote sensing in distinguishing subtle land cover transitions.

The application of the random forest classifier was crucial for resolving unclassified pixels, especially in regions with sparse or inconsistent data coverage. Notably, locations such as black spruce peat plateaus in Canada and larch forests in Russia required machine learning-based estimations due to their unique spectral and structural characteristics. While the classifier performed well for forested classes (>82 % accuracy), moderate accuracies (>50 %) were observed for herbaceous wetlands and tundra, reflecting inherent classification challenges in these environments.

Relevance for Carbon Modeling and Ecosystem Dynamics

From a carbon modeling perspective, land cover maps that contain a detailed and accurate classification of various vegetation community types are important due to their unique contributions to carbon and methane dynamics. Forested and shrubland regions play a crucial role in CO_2 sequestration due to their high biomass density and substantial carbon storage in both aboveground and belowground pools. Forests act as significant carbon sinks, with global estimates indicating that they sequester approximately 7.6 ± 49 Pg C annually (Pan et al., 2011). Boreal and temperate forests, in particular, store vast amounts of carbon in biomass and soils, with permafrost-associated forests having long-term carbon storage potential (Hugelius et al., 2020), while shrublands also contribute to carbon dynamics by storing carbon in woody biomass and organic soils. Shrub expansion in high-latitude tundra regions due to climate change has been linked to increased CO_2 uptake during the growing season (Myers-Smith et al., 2011), however, these ecosystems can also act as sources of methane under conditions of water saturation, particularly in permafrost regions (Treat et al., 2018). In contrast, wetlands and peatlands, which store about one-third of the world's soil carbon (Yu, 2012), are also significant sources of methane emissions, especially under warmer and wetter conditions (Turetsky et al., 2014).

Improving the spatial resolution of wetland land cover mapping can significantly enhance model accuracy by reducing the amount of heterogeneity not captured within pixels, leading to more accurate ecosystem representation (e.g., Kuhn et al., 2021). Land cover products, such as BAWLD and the Circumpolar Arctic Land Cover Unit (CALU) dataset, provide refined classifications of wetland, lake, and river ecosystems (Bartsch et al., 2024; Olefeldt et al., 2021). These datasets incorporate expert knowledge and spatial data to differentiate landscape classes based on distinct hydrological and biogeochemical characteristics, enabling improved modeling of current and future methane emissions. However, global models disagree as to the magnitude and spatial distribution of emissions, due to uncertainties in wetland area and emissions per unit area (Bohn et al 2015).

Accurate representation of these dynamics in carbon cycle models is crucial for improving projections of future climate, and understanding these distributions accurately is fundamental for process-based models that

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

simulate greenhouse gas fluxes in Arctic and boreal ecosystems. The hybrid land cover product, with its improved representation of fine-scale vegetation classes, enhances the capacity of these models to capture spatial variability in these complex systems. Despite its advantages, uncertainties remain with this approach. Differences in time periods and methodologies across the global and regional datasets introduce inconsistencies in land cover classification. Additionally, classification errors due to spectral similarities between vegetation classes highlight the need for continued refinement, potentially through the integration of ancillary datasets such as LiDAR or SAR imagery.

Uncertainties

While integrating multiple land cover datasets may enhance spatial representation, it introduces challenges related to classification consistency/harmonization, scale and accurate representation of changes in vegetation dynamics. The fusion of global and regional land cover products can lead to boundary mismatches, particularly in transition zones. Studies have shown that the classification accuracy of global land cover products varies across different regions. For example the ESA CCI-LC product shows an overally accuracy of 63.5 % in the Arctic region, thus highlights the need for careful consideration when integrating more detailed regional products to supplement the classification of these regions (Liang et al 2019).

Our hybrid land cover product's 1 km resolution balances spatial detail with computational efficiency, but may not capture fine scale heterogeneity in certain landscapes. Less represented land cover types, such as fens or patchy wetlands may be underrepresented when aggregated to coarser resolutions. Resampling from finer resolutions (i.e. 10 m to 10 km) can significantly alter the proportion of these heterogeneous classes. For instance, the proportion of wetland cover in the CALU dataset decreases from approximately 9 % at 10 m resolution to 6.3 % at 1km and further to 5.5 % at 10 km. These changes should be considered when integrating datasets at varying resolutions or when applying the datasets to broad regions.

Given these uncertainties, users applying this dataset for carbon flux modeling, biodiversity assessments, or land cover change analyses should account for potential biases introduced by classification errors, resolution limitations, and temporal inconsistencies. While this hybrid dataset improves Arctic and boreal land cover representation compared to global datasets alone, ongoing refinements are needed to enhance the accuracy of underrepresented classes and transitional zones. Understanding these dataset uncertainties is essential for informed application in ecological and climate research. By acknowledging and addressing these limitations, users can better interpret the data and contribute to ongoing efforts to refine land cover mapping in Arctic and boreal regions.

Conclusion

Accurate land cover maps are essential for understanding ecosystem structure, dynamics, and change, yet comprehensive, high-resolution maps remain scarce across many of Earth's biomes, including the ABZ. To address this gap, we present a new hybrid land cover dataset spanning the entire ABZ at a moderate spatial resolution of 1 km. This circumpolar product integrates and harmonizes multiple existing global and regional land cover datasets, improving representation of key vegetation types, including shrub tundra and boreal forest communities, which are often underrepresented or misclassified in coarser-resolution products. Our dataset distinguishes 35 land cover classes tailored to ecological and modeling applications, offering improved accuracy and spatial consistency across geopolitical boundaries.

The underlying methodology combines a multi-step integration process with machine learning-based refinement, leveraging agreement between validated products to enhance reliability. While this workflow supports the dataset's development, our primary contribution is the hybrid product itself—designed to support ecosystem modeling, permafrost and carbon assessments, and land-atmosphere interaction studies. This dataset serves as a valuable resource for the scientific community working in the ABZ, and future updates may incorporate additional observational inputs to further improve resolution and thematic detail.

Author contributions

VB conceptualized the paper with the support of EJ, HG, BR. Co-authors EJ, HG, BR, BM, JW, SN, AB, AMV, JR reviewed and provided feedback and content to the manuscript.

Competing Interests

The authors declare that they have no conflict of interest.

Acknowledgements

This work was funded by the Quadrature Climate Foundation (grant no. 01-21-000094) and funding catalyzed by the TED Audacious Project (Permafrost Pathways), with additional support from Google.org's Impact Challenge for Climate Innovation Program, and the Bonanza Creek Long Term Ecological Research Program (grant no. DEB-2224776).

Data Availability

The dataset: A hybrid circumpolar 1 km land cover product cover product has been made available to the Zenodo platform and can be accessed via https://doi.org/10.5281/zenodo.15231293 (Briones et al., 2025). The data is available in GeoTIFF format at a 1 km spatial resolution.

Supplementary Materials

References

Aune-Lundberg, L., and Strand, G.-H.:The content and accuracy of the CORINE Land Cover dataset for Norway. International Journal of Applied Earth Observation and Geoinformation, 96, 102266. https://doi.org/10.1016/j.jag.2020.102266, 2021.

Balashov, I., Bartalev, S., Bartalev, S., Burtsev, M., Vorushilov, I., Egorov, V., Kashnitskii, A., Khovratovich, T., Khvostikov, S., Kobets, D., Loupian, E., Saigin, I., Senko, K., Stytsenko, F., Sychugov, I., and Zharko, V.: Vega-Les Information System. Actual Features and Future Evolution. IOP Conference Series: Earth and Environmental Science, 507(1), 012002. https://doi.org/10.1088/1755-1315/507/1/012002, 2020.

Balshi, M. S., Mcguire, A. D., Duffy, P., Flannigan, M., Kicklighter, D. W., & Melillo, J.:Vulnerability of carbon storage in North American boreal forests to wildfires during the 21st century. Global Change Biology, 15(6), 1491–1510. https://doi.org/10.1111/j.1365-2486.2009.01877.x, 2009.

Bartalev, S. A., Belward, A. S., Erchov, D. V., & Isaev, A. S.: A new SPOT4-VEGETATION derived land cover map of Northern Eurasia. International Journal of Remote Sensing, 24(9), 1977–1982. https://doi.org/10.1080/0143116031000066297, 2003.

Bartalev, S. A., Plotnikov, D. E., & Loupian, E. A.: Mapping of arable land in Russia using multi-year time series of MODIS data and the LAGMA classification technique. Remote Sensing Letters, 7(3), 269–278.

https://doi.org/10.1080/2150704X.2015.1130874, 2016.

Bartsch, A., Efimova, A., Widhalm, B., Muri, X., von Baeckmann, C., Bergstedt, H., Ermokhina, K., Hugelius, G., Heim, B., & Leibman, M.: Circumarctic land cover diversity considering wetness gradients. Hydrology and Earth System Sciences, 28(11), 2421–2481. https://doi.org/10.5194/hess-28-2421-2024, 2024.

Bartsch, A., Höfler, A., Kroisleitner, C., and Trofaier, A.: Land cover mapping in northern high latitude permafrost regions with satellite data: Achievements and remaining challenges. Remote Sensing, 8(12), 979. https://doi.org/10.3390/rs8120979, 2016.

Betts, R. A.: Offset of the potential carbon sink from boreal forestation by decreases in surface albedo. Nature, 408(6809), Article 6809. https://doi.org/10.1038/35041545. 2000.

Bohn, T. J., Melton, J. R., Ito, A., Kleinen, T., Spahni, R., Stocker, B. D., Zhang, B., Zhu, X., Schroeder, R., Glagolev, M. V., et al.:WETCHIMP-WSL: Intercomparison of wetland methane emissions models over West Siberia. Biogeosciences, 12(11), 3321–3349. https://doi.org/10.5194/bg-12-3321-2015, 2015

Bonfils, C. J. W., Phillips, T. J., Lawrence, D. M., Cameron-Smith, P., Riley, W. J., and Subin, Z. M.: On the influence of shrub height and expansion on northern high latitude climate. Environmental Research Letters, 7(1), 015503. https://doi.org/10.1088/1748-9326/7/1/015503, 2012.

Briones V, Genet H, Jafarov E E, Rogers B, Watts J D, Virkkala A-M, Bartsch A, Maglio B C, Natali S M and Rady J.: Hybrid Land Cover Product: A hybrid circumpolar 1 km land cover product Online: https://zenodo.org/records/15231293. 2025.

Brown, J., Sidlauskas, F. J., and Delinski, G.: Circum-Arctic map of permafrost and ground ice conditions, map, National Snow and Ice Data Center, Boulder, CO, USA, 1997.

Brus, D. J., Hengeveld, G. M., Walvoort, D. J. J., Goedhart, P. W., Heidema, A. H., Nabuurs, G. J., and Gunia, K.: mapping of tree species over Europe. European Journal of Forest Research, 131(1), 145–157. https://doi.org/10.1007/s10342-011-0513-5, 2012.

Carman, T. B., Genet, H., Rutter, R. A., Jafarov, E., Euskirchen, E., Maglio, B., Clein, J., Briones, V., and others: DVMDOSTEM: a terrestrial ecosystem model designed to represent Arctic, boreal and permafrost ecosystem dynamics, in review, J. Open Source Software., available at: https://joss.theoj.org/papers/de6adf41412b7028d7be0a2064ea78cf, 2025.

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., Hahn, N., et al.: An ecoregion-based approach to protecting half the terrestrial realm. BioScience, 67(6), 534–545. https://doi.org/10.1093/biosci/bix014, 2017.

Elmendorf, S. C., Henry, G. H. R., Hollister, R. D., Björk, R. G., Boulanger-Lapointe, N., Cooper, E. J., Cornelissen, J. H. C., Day, T. A., Dorrepaal, E., Elumeeva, T. G., et al.:Plot-scale evidence of tundra vegetation change and links to recent summer warming. Nature Climate Change, 2(6), 453–457. https://doi.org/10.1038/nclimate1465, 2012.

Epstein, H. E., Calef, M. P., Walker, M. D., Stuart Chapin III, F., and Starfield, A. M.: Detecting changes in Arctic tundra plant communities in response to warming over decadal time scales. Global Change Biology, 10(8), 1325–1334. https://doi.org/10.1111/j.1529-8817.2003.00810.x, 2004.

ESA: Land Cover CCI Product User Guide Version 2.0, available at: http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf, last access: 16 April 2024.

European Environment Agency: CORINE Land Cover 2018 (raster 100 m), Europe, 6-yearly – version 2020_20u1, May 2020, https://doi.org/10.2909/960998c1-1870-4e82-8051-6485205ebbac, 2020.

European Environment Agency (EEA): Extended wetland ecosystem layer, Prod-ID: DAT-276-en, published 30 June 2022, available at: https://www.eea.europa.eu/en/datahub/datahubitem-view/b9399908-557a-47a8-954a-958dabeaf1b6, last access: 05 June 2024.

Euskirchen, E. S., Bennett, A. P., Breen, A. L., Genet, H., Lindgren, M. A., Kurkowski, T. A., McGuire, A. D., and Rupp, T. S.: Consequences of changes in vegetation and snow cover for climate feedbacks in Alaska and northwest Canada. Environmental Research Letters, 11(10), 105003. https://doi.org/10.1088/1748-9326/11/10/105003, 2016.

Heimann, M., and Reichstein, M.: Terrestrial ecosystem carbon dynamics and climate feedbacks. Nature, 451(7176), Article 7176. https://doi.org/10.1038/nature06591, 2008.

Heiskanen, J.: Evaluation of global land cover data sets over the tundra–taiga transition zone in northernmost Finland. International Journal of Remote Sensing, 29(13), 3727–3751. https://doi.org/10.1080/01431160701871104, 2008.

Hermosilla, T., Wulder, M. A., White, J. C., and Coops, N. C.: Land cover classification in an era of big and open

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

data: Optimizing localized implementation and training data selection to improve mapping outcomes. Remote Sensing of Environment, 268, 112780. https://doi.org/10.1016/j.rse.2021.112780, 2022.

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., and Hobart, G. W.: Disturbance-informed annual land cover classification maps of Canada's forested ecosystems for a 29-Year Landsat time series. Canadian Journal of Remote Sensing, 44(1), 67–87. https://doi.org/10.1080/07038992.2018.1437719, 2018.

Herold, M., Mayaux, P., Woodcock, C. E., Baccini, A., and Schmullius, C.: Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. Remote Sensing of Environment, 112(5), 2538–2556. https://doi.org/10.1016/j.rse.2007.11.013, 2008.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology, 25(15), 1965–1978. https://doi.org/10.1002/joc.1276, 2005.

Horvath, P., Tang, H., Halvorsen, R., Stordal, F., Tallaksen, L. M., Berntsen, T. K., and Bryn, A.: Improving the representation of high-latitude vegetation distribution in dynamic global vegetation models. Biogeosciences, 18(1), 95–112. https://doi.org/10.5194/bg-18-95-2021, 2021.

Hugelius, G., Loisel, J., Chadburn, S., Jackson, R. B., Jones, M., MacDonald, G., Marushchak, M., Olefeldt, D., Packalen, M., Siewert, M. B., Treat, C., Turetsky, M., Voigt, C., and Yu, Z.: Large stocks of peatland carbon and nitrogen are vulnerable to permafrost thaw. Proceedings of the National Academy of Sciences, 117(34), 20438–20446. https://doi.org/10.1073/pnas.1916387117, 2020.

Intergovernmental Panel on Climate Change.: Climate change and land: IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781009157988, 2022.

Jin, Y., Liu, X., Chen, Y., and Liang, X.: Land-cover mapping using Random Forest classification and incorporating NDVI time-series and texture: A case study of central Shandong. International Journal of Remote Sensing, 39, 1–21. https://doi.org/10.1080/01431161.2018.1490976, 2018.

Johnstone, J. F., Allen, C. D., Franklin, J. F., Frelich, L. E., Harvey, B. J., Higuera, P. E., Mack, M. C., Meentemeyer, R. K., Metz, M. R., Perry, G. L., Schoennagel, T., and Turner, M. G.: Changing disturbance regimes, ecological memory, and forest resilience. Frontiers in Ecology and the Environment, 14(7), 369–378. https://doi.org/10.1002/fee.1311, 2016.

Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M., Kuemmerle, T., Meyfroidt, P., Mitchard, E., Reiche, J., Ryan, C., and Waske, B.: A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring, Remote Sens., 8, 70, https://doi.org/10.3390/rs8010070, 2016.

Jung, M., Henkel, K., Herold, M., and Churkina, G.: Exploiting synergies of global land cover products for carbon cycle modeling. Remote Sensing of Environment, 101(4), 534–553. https://doi.org/10.1016/j.rse.2006.01.020, 2006.

Kåresdotter, E., Destouni, G., Ghajarnia, N., Hugelius, G., and Kalantari, Z.: Mapping the vulnerability of Arctic wetlands to global warming. Earth's Future, 9(5), e2020EF001858. https://doi.org/10.1029/2020EF001858, 2021.

Kuhn, M. A., Varner, R. K., Bastviken, D., Crill, P., MacIntyre, S., Turetsky, M., Walter Anthony, K., McGuire, A. D., and Olefeldt, D.: BAWLD-CH_4: A comprehensive dataset of methane fluxes from boreal and Arctic ecosystems. Earth System Science Data, 13(11), 5151–5189. https://doi.org/10.5194/essd-13-5151-2021, 2021.

Lamarche, C., Santoro, M., Bontemps, S., D'Andrimont, R., Radoux, J., Giustarini, L., Brockmann, C., Wevers, J., Defourny, P., and Arino, O.: Compilation and validation of SAR and optical data products for a complete and global map of inland/ocean water tailored to the climate modeling community. Remote Sensing, 9(1), Article 1. https://doi.org/10.3390/rs9010036, 2017.

Latifovic, R., Pouliot, D., and Olthof, I.: Circa 2010 land cover of Canada: Local optimization methodology and product development. Remote Sensing, 9(11), Article 11. https://doi.org/10.3390/rs9111098, 2017.

Liang, L., Liu, Q., Liu, G., Li, H., and Huang, C.: Accuracy evaluation and consistency analysis of four global land cover products in the Arctic region. Remote Sensing, 11(12), Article 12. https://doi.org/10.3390/rs11121396, 2019.

Liu, A., Moore, J. C., and Chen, Y.: PInc-PanTher estimates of Arctic permafrost soil carbon under the GeoMIP G6solar and G6sulfur experiments. Earth System Dynamics, 14(1), 39–53. https://doi.org/10.5194/esd-14-39-2023, 2023.

Liu, C., Li, W., Zhu, G., Zhou, H., Yan, H., and Xue, P.: Land use/land cover changes and their driving Factors in the Northeastern Tibetan Plateau based on geographical detectors and Google Earth Engine: A case study

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

in Gannan Prefecture. Remote Sensing, 12(19), Article 19. https://doi.org/10.3390/rs12193139, 2020.

Luo, Y., Zhu, Z., Zhao, W., Li, M., Chen, J., Zhao, P., Sun, L., Zhang, Y., Duanmu, Z., and Chen, J.: Hybrid global annual 1-km IGBP land cover maps for the period 2000–2020. Journal of Remote Sensing, 4, 0122. https://doi.org/10.34133/remotesensing.0122, 2024.

Mack, M. C., Walker, X. J., Johnstone, J. F., Alexander, H. D., Melvin, A. M., Jean, M., and Miller, S. N.: Carbon loss from boreal forest wildfires offset by increased dominance of deciduous trees. Science, 372(6539), 280–283. https://doi.org/10.1126/science.abf3903, 2021.

Mahdianpari, M., Salehi, B., Mohammadimanesh, F., Brisco, B., Homayouni, S., Gill, E., DeLancey, E. R., and Bourgeau-Chavez, L.: 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/10.1080/07038992.2019.1711366, 2020.

Maxwell, A. E., Strager, M. P., Warner, T. A., Ramezan, C. A., Morgan, A. N., and Pauley, C. E.: Large-area, high spatial resolution land cover mapping using Random Forests, GEOBIA, and NAIP orthophotography: findings and recommendations. Remote Sensing, 11(12), Article 12. https://doi.org/10.3390/rs11121409, 2019.

Melvin, A. M., Mack, M. C., Johnstone, J. F., David McGuire, A., Genet, H., and Schuur, E. A. G.: Differences in ecosystem carbon distribution and nutrient cycling linked to forest tree species composition in a mid-successional Boreal forest. Ecosystems, 18(8), 1472–1488. https://doi.org/10.1007/s10021-015-9912-7, 2015.

Miller, P. A., and Smith, B.: Modelling tundra vegetation response to recent Arctic warming. AMBIO, 41(3), 281–291. https://doi.org/10.1007/s13280-012-0306-1, 2012.

Myers-Smith, I. H., Forbes, B. C., Wilmking, M., Hallinger, M., Lantz, T., Blok, D., Tape, K. D., Macias-Fauria, M., Sass-Klaassen, U., Lévesque, E., Boudreau, S., Ropars, P., Hermanutz, L., Trant, A., Collier, L. S., Weijers, S., Rozema, J., Rayback, S. A., Schmidt, N. M., and Hik, D. S.: Shrub expansion in tundra ecosystems: Dynamics, impacts and research priorities. Environmental Research Letters, 6(4), 045509. https://doi.org/10.1088/1748-9326/6/4/045509, 2011.

Neigh, C. S. R., Nelson, R. F., Ranson, K. J., Margolis, H. A., Montesano, P. M., Sun, G., Kharuk, V., Næsset, E.,

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

Wulder, M. A., and Andersen, H.-E.: Taking stock of circumboreal forest carbon with ground measurements, airborne and spaceborne LiDAR. Remote Sensing of Environment, 137, 274–287. https://doi.org/10.1016/j.rse.2013.06.019, 2013.

Oehri, J., Schaepman-Strub, G., Kim, J.-S., Grysko, R., Kropp, H., Grünberg, I., Zemlianskii, V., Sonnentag, O., Euskirchen, E. S., Reji Chacko, M., Muscari, G., Blanken, P. D., Dean, J. F., Di Sarra, A., Harding, R. J., Sobota, I., Kutzbach, L., Plekhanova, E., Riihelä, A., and Boike, J., et al.: Vegetation type is an important predictor of the Arctic summer land surface energy budget, Nature Communications, 13, 6379, https://doi.org/10.1038/s41467-022-34049-3, 2022.

Olefeldt, D., Hovemyr, M., Kuhn, M. A., Bastviken, D., Bohn, T. J., Connolly, J., et al.: The Boreal–Arctic Wetland and Lake Dataset (BAWLD). Earth System Science Data, 13(11), 5127–5149. https://doi.org/10.5194/essd-13-5127-2021, 2021.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., and Wulder, M. A.: Good practices for estimating area and assessing accuracy of land change. Remote Sensing of Environment, 148, 42–57. https://doi.org/10.1016/j.rse.2014.02.015, 2014.

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., et al.: A large and persistent carbon sink in the world's forests. Science, 333(6045), 988–993. https://doi.org/10.1126/science.1201609, 2011.

Pérez-Hoyos, A., García-Haro, F. J., and San-Miguel-Ayanz, J.: A methodology to generate a synergetic land-cover map by fusion of different land-cover products. International Journal of Applied Earth Observation and Geoinformation, 19, 72–87. https://doi.org/10.1016/j.jag.2012.04.011, 2012.

Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., Bontemps, S., Boettcher, M., Brockmann, C., Defourny, P., Hagemann, S., Herold, M., Kirches, G., Lamarche, C., Lederer, D., Ottlé, C., Peters, M., and Peylin, P.: Plant functional type classification for earth system models: Results from the European Space Agency's Land Cover Climate Change Initiative. Geoscientific Model Development, 8(7), 2315–2328. https://doi.org/10.5194/gmd-8-2315-2015, 2015.

Prestele, R., Alexander, P., Rounsevell, M. D. A., Arneth, A., Calvin, K., Doelman, J., et al.: Hotspots of uncertainty in land-use and land-cover change projections: A global-scale model comparison. Global Change Biology, 22(12), 3967–3983. https://doi.org/10.1111/gcb.13337, 2016.

Raynolds, M. K. and Walker, D. A.: Circumpolar relationships between permafrost characteristics, NDVI, and

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

Open Access Earth System Science Data Discussions

Arctic vegetation types, in: Proceedings of the Ninth International Conference on Permafrost, Fairbanks, Alaska, USA, 29 June–3 July 2008, Institute of Northern Engineering, University of Alaska Fairbanks, 1469–1474, 2008.

Rogan, J., and Chen, D.: Remote sensing technology for mapping and monitoring land-cover and land-use change. Progress in Planning, 61(4), 301–325. https://doi.org/10.1016/S0305-9006(03)00066-7, 2004.

Rollins, M. G.: LANDFIRE: A nationally consistent vegetation, wildland fire, and fuel assessment. International Journal of Wildland Fire, 18(3), 235–249. https://doi.org/10.1071/WF08088, 2009.

Schuur, E. A. G., Abbott, B. W., Commane, R., Ernakovich, J., Euskirchen, E., Hugelius, G., et al.: Permafrost and climate change: Carbon cycle feedbacks from the warming Arctic. Annual Review of Environment and Resources, 47(Volume 47, 2022), 343–371. https://doi.org/10.1146/annurev-environ-012220-011847, 2022.

Stinson, G., Magnussen, S., Boudewyn, P., Eichel, F., Russo, G., Cranny, M., and Song, A.: Canada. In National Forest Inventories (pp. 233–247). Springer, Cham. https://doi.org/10.1007/978-3-319-44015-6_12, 2016.

Tape, K., Sturm, M., and Racine, C.: The evidence for shrub expansion in northern Alaska and the Pan-Arctic. Global Change Biology, 12(4), 686–702. https://doi.org/10.1111/j.1365-2486.2006.01128.x, 2006.

Thompson, C., Beringer, J., Iii, F. S. C., and McGuire, A. D.: Structural complexity and land-surface energy exchange along a gradient from Arctic tundra to boreal forest. Journal of Vegetation Science, 15(3), 397–406. https://doi.org/10.1111/j.1654-1103.2004.tb02277.x, 2004.

Treat, C. C., Marushchak, M. E., Voigt, C., Zhang, Y., Tan, Z., Zhuang, Q., Virtanen, T. A., Räsänen, A., Biasi, C., Hugelius, G., Kaverin, D., Miller, P. A., Stendel, M., Romanovsky, V., Rivkin, F., Martikainen, P. J., and Shurpali, N. J.: Tundra landscape heterogeneity, not interannual variability, controls the decadal regional carbon balance in the Western Russian Arctic. Global Change Biology, 24(11), 5188–5204. https://doi.org/10.1111/gcb.14421, 2018.

Turetsky, M. R., Kotowska, A., Bubier, J., Dise, N. B., Crill, P., Hornibrook, E. R. C., Minkkinen, K., Moore, T. R., Myers-Smith, I. H., Nykänen, H., Olefeldt, D., Rinne, J., Saarnio, S., Shurpali, N., Tuittila, E.-S., Waddington, J. M., White, J. R., Wickland, K. P., and Wilmking, M.: A synthesis of methane emissions from 71 northern, temperate, and subtropical wetlands. Global Change Biology, 20(7), 2183–2197. https://doi.org/10.1111/gcb.12580, 2014.

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

USGCRP.: Second state of the carbon cycle report (pp. 1–470). U.S. Global Change Research Program, Washington, DC. https://carbon2018.globalchange.govhttps://carbon2018.globalchange.gov/chapter/11, 2018.

Virkkala, A.-M., Aalto, J., Rogers, B. M., Tagesson, T., Treat, C. C., Natali, S. M., et al.: Statistical upscaling of ecosystem CO2 fluxes across the terrestrial tundra and boreal domain: Regional patterns and uncertainties. Global Change Biology, 27(17), 4040–4059. https://doi.org/10.1111/gcb.15659, 2021.

Virkkala, A.-M., Virtanen, T., Lehtonen, A., Rinne, J., and Luoto, M.: The current state of CO2 flux chamber studies in the Arctic tundra: A review. Progress in Physical Geography: Earth and Environment, 42(2), 162–184. https://doi.org/10.1177/0309133317745784, 2018.

Walker, D. A., Raynolds, M. K., Daniëls, F. J. A., Einarsson, E., Elvebakk, A., Gould, W. A., Katenin, A. E., Kholod, S. S., Markon, C. J., Melnikov, E. S., Moskalenko, N. G., Talbot, S. S., and Yurtsev, B. A.: The Circumpolar Arctic vegetation map, J. Veg. Sci., 16, 267–282, https://doi.org/10.1111/j.1654-1103.2005.tb02365.x, 2005.

Wang, F., Kroeger, K. D., Gonneea, M. E., Pohlman, J. W., and Tang, J.: Water salinity and inundation control soil carbon decomposition during salt marsh restoration: An incubation experiment. Ecology and Evolution, 9(4), 1911–1921. https://doi.org/10.1002/ece3.4884, 2019.

Wang, J. A., Sulla-Menashe, D., Woodcock, C. E., Sonnentag, O., Keeling, R. F., and Friedl, M. A.: Extensive land cover change across Arctic–Boreal Northwestern North America from disturbance and climate forcing. Global Change Biology, 26(2), 807–822. https://doi.org/10.1111/gcb.14804, 2020.

Wang, L., Arora, V. K., Bartlett, P., Chan, E., and Curasi, S. R.: Mapping of ESA's Climate Change Initiative land cover data to plant functional types for use in the CLASSIC land model. Biogeosciences, 20(12), 2265–2282. https://doi.org/10.5194/bg-20-2265-2023, 2023.

White, J. C., Wulder, M. A., Hobart, G. W., Luther, J. E., Hermosilla, T., Griffiths, P., Coops, N. C., Hall, R. J., Hostert, P., Dyk, A., and Guindon, L.: Pixel-based image compositing for large-area dense time series applications and science. Canadian Journal of Remote Sensing, 40(3), 192–212. https://doi.org/10.1080/07038992.2014.945827, 2014.

Yu, Z. C.: Northern peatland carbon stocks and dynamics: A review. Biogeosciences, 9(10), 4071–4085. https://doi.org/10.5194/bg-9-4071-2012, 2012.

https://doi.org/10.5194/essd-2025-226
Preprint. Discussion started: 15 May 2025

741 Zhang, W., Miller, P. A., Smith, B., Wania, R., Koenigk, T., and Döscher, R.: Tundra shrubification and tree-line advance amplify Arctic climate warming: Results from an individual-based dynamic vegetation model. Environmental Research Letters, 8(3), 034023. https://doi.org/10.1088/1748-9326/8/3/034023, 2013.